Lecture Notes in Networks and Systems 65

H. S. Saini
R. K. Singh
Girish Kumar
G. M. Rather
K. Santhi *Editors*

Innovations in Electronics and Communication Engineering

Proceedings of the 7th ICIECE 2018

Springer

Lecture Notes in Networks and Systems

Volume 65

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

More information about this series at http://www.springer.com/series/15179

H. S. Saini · R. K. Singh ·
Girish Kumar · G. M. Rather ·
K. Santhi
Editors

Innovations in Electronics and Communication Engineering

Proceedings of the 7th ICIECE 2018

Springer

Editors
H. S. Saini
Guru Nanak Institutions
Ibrahimpatnam, Telangana, India

Girish Kumar
Indian Institute of Technology Bombay
Mumbai, Maharashtra, India

K. Santhi
Department of ECE
Guru Nanak Institutions Technical Campus
Ibrahimpatnam, Telangana, India

R. K. Singh
Guru Nanak Institutions Technical Campus
Ibrahimpatnam, Telangana, India

G. M. Rather
Department of ECE
NIT Srinagar
Srinagar, Jammu and Kashmir, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-13-3764-2 ISBN 978-981-13-3765-9 (eBook)
https://doi.org/10.1007/978-981-13-3765-9

Library of Congress Control Number: 2018964232

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Organizing Committee

Editorial Board Members

Patrons

Sardar Tavinder Singh Kohli, Chairman, Guru Nanak Institutions
Sardar Gagandeep Singh Kohli, Vice-Chairman, Guru Nanak Institutions

Conference Chair

Dr. H. S. Saini, Managing Director, Guru Nanak Institutions
Dr. M. Ramalinga Reddy, Director, GNITC

Conference Co-Chairs

Prof. R. K. Singh, Associate Director, GNITC
Dr. S. V. Ranganayakulu, Dean—R&D, GNITC

Steering Committee

Dr. S. Sreenatha Reddy, Principal, GNIT
Prof. Parthsaradhy, Associate Director, GNITC
Dr. Rishi Sayal, Associate Director, GNITC
Dr. K. Chanthirasekharan, Dean (Academics)

Convenors

Dr. S. J. Sugumar, HOD—ECE
Dr. K. Santhi, Professor, GNITC

Committee Members

Prof. S. Maheshwara Reddy, HOD—ECE, GNITC
Prof. B. Kedharnath, HOD—ECE, GNIT
Mr. Binod Kumar, Associate Professor, GNITC

Dr. Rashid Mohmood, Professor, ECE
Dr. R. Prabhakar, Professor, ECE
Prof. A. Mohan, GNITC
Dr. Kushboo Pachori, Associate Professor, ECE
Dr. Harpreeth Kaur, Associate Professor, ECE
Prof. Anitha Swamidas, GNITC
Mr. Sandeep Patil, Associate Professor
Mr. B. Sridhar, Associate Professor, GNITC
Mr. Md. Shoukath Ali, Associate Professor, GNITC
Mr. D. Surendra Rao, Associate Professor, GNITC
Mr. K. Shashidhar, Associate Professor, GNITC
Mr. R. P. Naraiah, Associate Professor, GNITC

Advisory Board—International

Dr. Xiao-zhi Gao, Finland University
Prof. Libor Rufer, CIMA, Micro and Nano Systems Groups, France
Dr. Vishal M. Patel, State University of New Jersey, USA
Dr. Savatri Bevinakopp, MIT, Australia
Dr. C. H. Suresh Babu, K.K. University, Abha, KSA
Shri. Rajashekar N., INPHI, UK
Dr. Sergey Mosin, Vladimir State University, Russia
Dr. Katina Michel, University of Wollongong, Australia
Dr. Ahmed Faheem Zobaa, BU, UK
Dr. Akhtar Kalam, VU, Australia
Dr. Afredo Vaccaro, US, Italy
Dr. David YU, UMW, USA
Dr. Dimitri Vinnikov, TUT, Estonia
Dr. Gorazad Stumberger, UM, Slovenia
Dr. Jagadish Nayak, BITS, Pilani, Dubai Campus
Dr. Lausiong Hoe, Multimedia University, Malaysia
Dr. Hussian Shreef, UK, Malaysia
Dr. Murad Al-Shibli, Head—EMET, Abu Dhabi
Dr. Nesimi Ertugrul, UA, Australia
Richarad Blanchard, LBU, UK
Dr. Shashi Paul, DM, UK
Dr. Zhao Xu, HKPU, Hong Kong
Dr. Ahmed Zobba, Brunel University, UK
Dr. Adel Nasiri, UMW, USA
Dr. P. N. Sugunathan, NTU, Singapore
Dr. Hiroya Fuji Saki, University of Tokyo
Dr. Fawnizu Azmadi Husaain, Universiti Teknologi Petronas, Malaysia
Dimov Stojce Lice, DUT, Durban, South Africa

Advisory Board—National

Dr. Sheela Rani, Vice Chancellor, Sathyabama University, Chennai
Dr. M. Mukunda Rao, Professor Emeritus, IIT Chennai
Dr. N. V. L. Narasimha Murthy, IIT Bhubaneswar
Dr. M. Ravi Babu, IIT Ropar
Dr. Srinivasulu Thadishetti, Principal, K.U
Dr. Mahaveer Singh, Sr. Principal Scientist, National Physical Laboratory, CSIR, New Delhi
Dr. B. N. Bhandari, JNTUH, Hyderabad
Dr. L. Pratap Reddy, JNTUH, Hyderabad
Dr. D. Srinivasa Rao, JNTUH, Hyderabad
Dr. Ajish K. Abraham, HOD—Electronics, AIISH, Mysore
Dr. P. Lakshmi Narayana, Osmania University, Hyderabad
Dr. Manoj Kumar, Principal—DAVIET, Jalandhar
Dr. Pamela Chawla, Director, Institute of Engineering, MAU, HP
Dr. Novalina Jacob, Scientist, ADRIN, Hyderabad
Dr. Satish J. Sharma, RTM Nagpur University Nagpur
Dr. A. Bhujanga Rao, Department of Instrumentation, A.U. Visakhapatnam
Dr. N. K. Narayanan, Kannur University, Kerala
Dr. N. M. Renukappa, Professor and Scientist, JSS Technical Institutions, Mysore

Conference Committee Members

Budget

Prof. R. K. Singh
Mr. Sandeep Patil

Brochure, Web site, Easy Chair/E-mail Creation

Prof. V. Bhagya Raju
Mr. B. Sridhar
Mr. Srinivas Nanda

E-mail Campaigning Team

Prof. A. Mohan
Prof. Binod Kumar
Mr. L. Siva
Mr. K. Srinivas

Physical Campaigning Team

Prof. V. Bhagya Raju
Mr. D. Surendra Rao

Applications for Funding

Mr. K. Shashidhar

Preconference Tutorials

Mr. B. Sridhar
Ms. K. Ramyasri
Mr. K. Srinivas
Mr. K. V. Abhi
Ms. K. Madhavi
Mr. Md. Mohiddin (GNIT)
Mr. Saibabu

Receiving Papers and Acknowledgment, Attending Queries of Authors

Dr. K. Santhi
Mrs. C. Sailaja
Mr. Srinivas Nanda

Conversion of Papers

Dr. S. J. Sugumar
Dr. Prabhakar
Dr. Rashid Mahmood
Dr. Khushboo
Prof. Binod Kumar
Prof. Anita Swamidas

VIP Committee

Dr. K. Santhi
Mrs. C. Sailaja
Mr. Srinivas Nanda

Invitation Preparation and Distribution

Dr. K. Santhi
Mrs. C. Sailaja

Registration (Online/Spot)

Mr. Ch. Raja Rao
Mrs. K. Yashodha
Mrs. S. Swetha

Ms. S. Aparna
Mr. A. V. Rameswar Rao

Reception

Mrs. Anitha Swamidas
Prof. Binod Kumar
Ms. Simar Preeth Kaur
Mrs. B. Mythily Devi (GNIT)

Conference Office

Dr. Binod Kumar
Mrs. P. Shravya

VC, MD, JNTUH Messages

Prof. A. Mohan

Banners

Mr. K. Krishna Kumar

Proceedings

Dr. Binod Kumar
Mr. Naraiah
Mr. L. Siva

Transport Committee

Mr. D. Surendra Rao (VIP)
Mr. N. V. S. Moorthy (Participants)
Mr. V. Shivanna
Mr. T. A. P. Noor Mohamed

Program Committee

Prof. A. Mohan
Mr. K. Sashidhar
Dr. Harpreet Kaur
Mr. K. Krishna Kumar
Mr. H. Sharath Chandra (GNIT)
Mr. A. Aravind (GNIT)
Mr. Sriram (IT services)

Purchase Committee

Prof. S. Maheswara Reddy
Mr. D. Naresh
Mr. Rameshwara Rao
Mr. Saibabu

Certificates

Dr. Prabhakar
Mr. L. Shiva
Mrs. Nagamani
Mrs. G. Suma Latha

Food Committee

Dr. Prabhakar
Mrs. B. Anitha (VIP Refreshment and Lunch)
Mr. R. Gopinath (VIP Refreshment and Lunch)
Mr. K. Raju
Mr. Chinna Narasimhulu
Mr. Y. Jeevan

Accommodations

Dr. Prabhakar
Dr. Kushboo Pachori
Mr. Chinna Narasimhulu
Mr. K. Raju

Inauguration and Valedictory Function

Prof. R. K. Singh
Mr. Sandeep Patil
Dr. K. Santhi
Dr. S. J. Sugumar

Photographs and Video

Mr. O. Ravinder

Press Release and Media

Prof. A. Mohan
Mr. B. Sridhar

Conference Report

Dr. K. Santhi
Dr. S. J. Sugumar
Mrs. C. Sailaja

After Conference Attending Queries of Authors

Dr. K. Santhi (Springer)
Mrs. C. Sailaja (Springer)
Mr. Shoukath Ali (UGC-IJRAR)
Prof. A. Mohan (IJET)
Mr. L. Shiva (IJET)
Dr. Kushboo (JIECE)

Keynote Speech Arrangements

Mrs. G. Kiranmayee
Simar Preeth Kaur

VIP Accompany

Mr. Md. Shoukath Ali (Dr. Razali Ngah)
Mrs. G. Kiranmayee (Dr. Kim's Wife)
Dr. S. J. Sugumar (Dr. Nandy)
Mrs. Anitha Swamidas (Dr. K. B. Jayanthi)
Student (Dr. Parisa Beham)
(Dr. Danial Elizebeth Rani)
(Dr. Girish Kumar)
Ms. B. Aruna (Dr. Girish Kumar's Wife)

Reviewers

Internal Reviewers

Prof. R. K. Singh, Associate Director
Dr. S. V. Ranganayakulu, Dean/R&D, GNI
Prof. M. A. Kadar Baba, HOD/ECE, GNITC
Dr. K. Satish Reddy, Professor/ECE
Dr. K. Santhi, Professor/ECE
Dr. S. Sukumar, Professor/ECE
Dr. Syed Jalal Ahmad, Professor/ECE
Prof. Anita Swamidas, Professor/ECE
Prof. CH. Hemasundara Rao, Professor/ECE
Prof. N. V. Purnachandar Rao, Professor/ECE

Prof. A. Mohan, Professor/ECE
Mr. V. Bhagya Raju, Professor/ECE
Mr. K. Shashidhar, Associate Professor/ECE
Mr. B. Sridhar, Associate Professor/ECE
Mrs. G. Kiran Maye, Associate Professor/ECE
Mr. Bachu Srinivas, Associate Professor/ECE
Dr. N. Satish Kumar, Professor/ECE, GNIT
Dr. R. Vinoth Kanna, HOD/EEE, GNIT

External Reviewers

Dr. Nazib Uddin, HOD, Professor, NIT Srinagar
Dr. Gulal Mohemmed Rather, Professor, NIT Srinagar
Challa S. Sastry, No. 318 (E Block), Department of Mathematics, IIIT Hyderabad
Dr. Suryakanth V. Gangashetty, Assistant Professor, IIIT Hyderabad
Dr. Somasundaram, Professor, IISC, Bangalore
Dr. B. Thayaneswaran, Assistant Professor, Sona College of Technology, Salem, Tamil Nadu
Dr. C. Rajasekaran, Professor/ECE, KSRCT, Namakkal, Tamil Nadu
Arul Jyothi, Associate Professor, SREC, Coimbatore
M. Nirmala, Associate Professor, KSRCT
Dr. Shiva Subramanyam, KSRIT, Tamil Nadu
Dr. Supha Laxmi, Professor and Head, Paavai Institutions, Tamil Nadu
Dr. M. Maheswaran, Kongu Engineering College
Dr. Gunavathi, Associate Professor, VIT, Vellore
Dr. Maheshwar, Professor, Sri Krishna College of Engineering, Coimbatore
Dr. Malmurugan, Professor, PAL Institutions
Dr. M. V. Jithin Kumar, Research Coordinator, Heera College of Engineering & Technology, Kerala
Gnanou Florence Sudha, Pondicherry Engineering College, Pondicherry
Dr. J. Yogapriya, Dean (R&D), Kongunadu College of Engineering and Technology, Trichy, Tamil Nadu
Mrs. Rohini R. Mergu, Assistant Professor, Walchand Institute of Technology, Solapur
Dr. Jakulin Moni, Professor, Karunya University, Tamil Nadu
Dr. V. R. Vijay Kumar, Professor, Anna University, Coimbatore
Dr. Kanchana Bhaskar, Professor Dean Academics, VIT, Chennai
Dr. S. P. Singh, MGIT, Hyderabad
Dr. K. Jayashangkar, Principal, Vasavi College of Engineering, Hyderabad
Dr. B. L. Raju, Principal, ACE College of Engineering, Hyderabad
Dr. P. G. Geetha, Professor, Sri Vidyaniketan, Tirupati
Dr. R. P. Ramkumar, Professor/CSE/MREC
Dr. Pasha, HOD, Professor, BVRIT
Dr. B. Lokeswar Rao, Professor, CMRCET, Hyderabad
Dr. Nageswara Rao, Professor, TKRCE, Hyderabad

Dr. Karunakar, Associate Professor, GITAM, Visakhapatnam
Dr. M. Suresh, Professor, CMRCET, Hyderabad
Dr. C. V. Narasimhulu, Professor, GCET, Hyderabad
Dr. A. S. N. Murthy, Professor, BVRIT, Hyderabad
Dr. P. Narahari Shastry, Professor, CBIT
Dr. Mahesh, Associate Professor
Dr. Geetha, Professor
Dr. B. K. Madhavi, Professor
Dr. Keerti Kumar, Professor
Dr. D. Venkat Reddy, Professor, MGIT
Sanket Goel, Associate Professor
Dr. D. Jayanthi, Professor, GRIET
Dr. I. V. Subba Reddy, Professor, Gitam University
Dr. Gulam Rasol Bagh, NIT Srinagar
Dr. Ramprasad, KHIT, Guntur
Dr. Usha sree, Professor, VJIET
Prof. P. V. Rao, Professor, VBIT
Dr. Prabha Selvaraj, Professor, CSE, MRIET
Dr. Brahma Reddy, Professor, VBIT
Dr. Rangaiah, Professor, RRCE
Dr. Pushamala, Professor, DSCE University
Dr. Humaira Nishat, Professor, CVR College of Engineering
Dr. Kanaka Durga, Professor and HOD, MVSR
Dr. Abbas Ahmed
Dr. Prakash Rao, Professor, St. Peters Engineering College
Dr. Keerthikumar, Professor, ECE Department
Dr. Shankar Babu, Professor, TKR Engineering College
Dr. I. A. Pasha, BVRIT
A. Raja Sekhar, Professor and HOD, Methodist College of Engineering
Dr. Sharma, Professor, Sri Indu College of Engineering
Dr. Kezia Joseph, Professor, Stanley College of Engineering
Dr. N. Sathish Kumar, Professor/ECE, Sri Ramakrishna Engineering College
Ms. S. Thenmozhi

Advisory Board

International Advisory Board for ICIECE 2017

Dr. Xiao-Zhi Gao, Finland University
Prof. Libor Rufer, CIMA, Micro and Nano Systems Groups, France
Dr. Vishal M. Patel, State University of New Jersey, USA
Dr. Savatri Bevinakopp, MIT, Australia
Dr. C. H. Suresh Babu, K.K. University, Abha, KSA

Shri. Rajashekhar N., INPHI, UK
Dr. Surgey Mosin, Vladimir State University, Russia
Dr. Katina Michael, University of Wollongong, Australia
Dr. Ahmed Faheem Zobaa, BU, UK
Dr. Akhtar Kalam, VU, Australia
Dr. Afredo Vaccaro, US, Italy
Dr. David Yu, UMW, USA
Dr. Dimitri Vinnikov, TUT, Estonia
Dr. Gorazad Stumberger, UM, Slovenia
Dr. Jagadish Nayak, BITS Pilani, Dubai Campus
Dr. Lausiong Hoe, Multimedia University, Malaysia
Dr. Hussian Shareef, UK, Malaysia
Dr. Murad Al-Shibli, Head, EMET, Abu Dhabi
Dr. Nesimi Ertugrul, UA, Australia
Dr. Richarad Lanchard, LBU, UK
Dr. Shashi Paul, DM, UK
Dr. Zhao Xu, HKPU, Hong Kong
Dr. Ahmed Zobaa, Brunel University, UK
Dr. Adel Nasiri, UMW, USA
Dr. P. N. Sugunathan, NTU, Singapore
Dr. Hiroya Fuji Saki, University of Tokyo
Dr. Fawnizu Azmadi Hussin, Universiti Teknologi Petronas, Malaysia
Dimov Stojce Ilcev, DUT, Durban, South Africa

National Advisory Board for ICIECE 2017

Dr. Mohammed Zafar Ali Khan, Associate Professor and HOD, EE Department, IIT Hyderabad
Dr. Gulam Mohammed Rather, Professor—ECE, NIT Srinagar
Dr. N. Bheem Rao, Associate Professor, NIT Warangal
Dr. P. Radha Krishna, Principal, Research Scientist, Infosys Labs, Hyderabad
Dr. M. Madhavi Latha, Professor—ECE, JNTU, Hyderabad
Dr. V. Malleswara Rao, Professor and HOD—ECE, Gitam University, Visakhapatnam
Dr. T. Srinivasulu, Professor, Professor and Dean, Faculty of Engineering, K.U., Warangal
Dr. S. Ramanarayana Reddy, HOD—CSE, IGDTU, Delhi
Dr. N. V. L. Narasimha Murthy, IIT Bhubaneswar
Dr. P. V. Rao, Raja Rajeswari College of Engineering, Bangalore
Dr. B. N. Bhandari, JNTUH, Hyderabad
Dr. P. Lakshmi Narayana, Osmania University, Hyderabad
Dr. Pamela Chawla, Director, Institute of Engineering, MAU, HP

Dr. Novalina Jacob, Scientist, ADRIN, Hyderabad
Dr. M. Lakshminarayana, Scientist-H (Retd) and Former Associate Director, DLRL
Ms. Madhumita Chakravarti, Scientist-G
Dr. Girish Kumar, Professor, IIT, Mumbai
Dr. Ajaz Hussain Mir, Professor, NIT, Srinagar

Preface

The Seventh International Conference on Innovations in Electronics and Communication Engineering (ICIECE 2018) organized by the Department of Electronics and Communication Engineering, Guru Nanak Institutions Technical Campus, was conducted in association with ISRO on July 27–28, 2018. The conference has grown exponentially over the years and has become a platform for scientists, researchers, academicians, and students to present their ideas and share their cutting-edge research in various fields of electronics and communication engineering. The focus for this year conference was "State of Technical Education all over the Globe and Smart Technologies," and we had distinguished speakers from India and abroad who shared innovative solutions and technologies.

Over 300 papers were received, of which 69 papers have been selected for Springer Conference Proceedings publication. The papers selected were presented by the authors during the conference. Parallel sessions were conducted to accommodate all the authors and to give them ample time to discuss their ideas. We are happy to note that all the authors were satisfied with the arrangements and encouraged us to conduct such conferences in the future as well.

We would like to thank all the keynote speakers, participants, speakers for a preconference tutorial session, session chairs, committee members, reviewers, international and national board members, Guru Nanak Institutions management, and all the people who have directly or indirectly contributed to the success of this conference. The editors would also like to thank Springer editorial team for their support and for publishing the papers as part of the "Lecture Notes in Networks and Systems" series continuously for last 3 years.

Ibrahimpatnam, India	H. S. Saini
Ibrahimpatnam, India	R. K. Singh
Mumbai, India	Girish Kumar
Srinagar, India	G. M. Rather
Ibrahimpatnam, India	K. Santhi

Contents

Part V Miscelleneous

Editors and Contributors

About the Editors

Dr. H. S. Saini is Managing Director for the Guru Nanak Institutions with Ph.D. in the field of computer science. He has over 22 years of experience at university/college level in teaching UG/PG students and has guided several B.Tech., M.Tech., and Ph.D. projects. He has published/presented more than 30 high-quality research papers in international and national journals and proceedings of international conferences. He is Editor of *Journal of Innovation in Electronics and Communication Engineering (JIECE)* published by Guru Nanak Publishers. He has two books to his credit. He is a lover of innovation and is an advisor for NBA/NAAC accreditation process to many institutions in India and abroad.

Prof. R. K. Singh is Associate Director, Guru Nanak Institutions Technical Campus, who is an alumnus of REC (now MNIT Jaipur) and did his M.Tech. from IIT Bombay, in the field of communication engineering. He has served Indian Army in the core of electronics and mechanical engineering for 20 years before hanging his uniform as Lt Col. He has rich industrial experience as Army Officer managing workshops and has been Teaching Faculty for more than 6 years while in service. He started his career as Teaching Assistant at MNIT Jaipur for 1 year before joining Indian Army. As Professor, he has served for more than 12 years after premature retirement from the army services. He has

served as HOD, Vice Principal, and Principal of various engineering colleges before being approved as Associate Director of this institute.

He had hands-on experience with high-tech electronic equipment and has done many courses on radars and simulators. He has published many papers on microstrip antennas, VLSI, and radar signal processing in national and international conferences.

Dr. Girish Kumar received Ph.D. in electrical engineering from IIT Kanpur in 1983. From 1983 to 1985, he was Research Associate in the Electrical Engineering Department, University of Manitoba, Winnipeg, Canada. From 1985 to 1991, he was Assistant Professor in the Electrical Engineering Department, University of North Dakota, Grand Forks, USA. Since 1991, he is Professor in the Electrical Engineering Department, IIT Bombay. His research interests are cell tower radiation hazards and solutions, microstrip antennas and arrays, broadband antennas, and microwave integrated circuits and systems. He has written more than 200 papers in the international and national journals and conference proceedings. He has written a book entitled "Broadband Microstrip Antennas" published by Artech House, USA.

G. M. Rather joined National Institute of Technology, Srinagar (erstwhile Regional Engineering College, Srinagar) in March 1983 just after the completion of B.E. in the year 1982 from the same institute. He is working as Professor in Electronics and Communication Engineering Department since 2004. He has a teaching experience of 34 years behind him. During his service, he has taught courses at PG and UG levels. He completed his M.E. course in electrical communication engineering and Ph.D. both from Indian Institute of Science, Bangalore, in 1988 and 1994, respectively.

He has published around 20 technical papers in national and international journals and conferences. He has expertise in the fields of microelectronics, analog

and digital communication, data communication, and millimeter wave communication and is presently supervising around eight students in these fields.

He is Member of IEEE and IEEE Microwave Society, Fellow of IETE, and Life Member of ISTE (India).

Dr. K. Santhi completed her B.E. in electrical and electronics engineering from Madras University in the year 2000; M.E. in applied electronics from Anna University, Chennai, in the year 2007; and her Ph.D. from Anna University, Chennai, in the year 2015. She is working as Professor in the Department of Electronics and Communication Engineering at Guru Nanak Institutions Technical Campus, Hyderabad. She has published more than 25 publications in international and national journals and conferences which include IEEE, Springer, Elsevier, etc. She has organized many seminars, workshops, and international conferences. She is Member of IEEE and Life Member of ISTE and IETE. Her research interests are digital image processing, electronics, control system, and image analysis for metallography. She has 20 years of teaching and 10 years of research experience.

Contributors

Aishwarya S. Acharya Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

G. S. Anitha Department of EEE, RV College of Engineering, Bengaluru, India

H. M. Anitha Department of Information Science and Engineering, BMS College of Engineering, Bangalore, India

M. Annalakshmi Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Parveen Arora NEC Technologies India Pvt Ltd, Bengaluru, India

S. Arul Jothi ECE Department, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India

C. Arul Murugan Department of ETE, Karpagam College of Engineering, Coimbatore, India

P. Aruna Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India

S. Aruna Deepthi Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India

B. Banuselvasaraswathy Department of ECE, Sri Krishna College of Technology, Coimbatore, India

M. Parisa Beham Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Kumud Kumar Bhardwaj ECE, CMR College of Engineering & Technology, Medchal, India

Gurjit Singh Bhathal Department of Computer Science and Engineering, Punjabi University, Patiala, Punjab, India

Nilamani Bhoi Department of Electronics and Telecommunication Engineering, VSSUT, Burla, Sambalpur, Odisha, India

P. Chandra Sekhar Electronics and Communication Engineering, Osmania University, Hyderabad, India

K. Chandrasekhar Koritala Department of Electrical and Electronics Engineering, RVR&JC, Guntur, India

Abhishek Choubey Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India;
Electronics and Communication Engineering, Gurunanak Institute of Technology, Hyderabd, Telengana, India

Shruti Bhargava Choubey Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India;
Electronics and Communication Engineering, Gurunanak Institute of Technology, Hyderabd, Telengana, India

Sanjay Deshmukh Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Manoj M. Dongre Department of Electronics and Telecom, Ramrao Adik Institute of Technology, Navi Mumbai, India

Milind Fernandes Goa College of Engineering, Ponda, Goa, India

Pavitra Gandhi Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Vivekananda Ganji Department of Electrical and Electronics Engineering, Acharya Nagarjuna University, Guntur, India

K. Gayathree Department of ECE, Sri Krishna College of Technology, Coimbatore, India

Nimish Ghagare Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

M. N. Giri Prasad Department of ECE, JNTU College of Engineering, Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India

Satish T. Hamde Department of Instrumentation Engineering, SGGSIE&T, Nanded, Maharashtra, India

J. P. Harshitha Department of ECE, RVCE, Bengaluru, India

Palak Hora Department of ECE, RVCE, Bengaluru, India

Namrata V. Jad Department of Instrumentation Engineering, SGGSIE&T, Nanded, Maharashtra, India

P. Jayarekha Department of Information Science and Engineering, BMS College of Engineering, Bangalore, India

Naveen Kumar Kabra Sardar Vallabhbhai National Institute of Technology Surat, Surat, India

Radhika S. Kalagudi Department of ECE, RVCE, Bengaluru, India

P. Kalyani Electronics and Communication Engineering, JNTUH, Hyderabad, India

Sushmita Karkera Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

B. Karunakara Rai Department of ECE, NMIT, Bengaluru, India

Amanpreet Kaur Department of Electronics and Communication Engineering, Thapar Institute of Engineering and Technology, Patiala, India

B. Keerthi Department of ECE, G. Narayanamma Institute of Technology and Science, Hyderabad, India

Rajendra Kokate Department of Instrumentation Engineering, Government College of Engineering, Jalgaon, India

L. Koteswara Rao Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation (A Deemed to-be University), Off-campus Centre, Hyderabad, India

C. S. N. Koushik Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India

Jayendra Kumar Department of Electronics and Communication Engineering, National Institute of Technology Jamshedpur, Jamshedpur, Jharkhand, India

Sandeep Kumar Department of Electronics and Communications, Sreyas Institution of Engineering and Technology, Hyderabad, India

Kumari Laxmi Department of Electronics and Communication Engineering Department, Thapar Institute of Engineering and Technology, Patiala, India

G. Laxminarayana Department of ECE, Anurag College of Engineering, Ranga Reddy, Telangana, India

M. Madhavi Latha Electronics and Communication Engineering, JNTUH, Hyderabad, India

S. Maheswara Reddy ECE, GNITC, Hyderabad, Telangana, India

Md Rashid Mahmood Department of Electronics and Communications, Guru Nanak Institutions Technical Campus, Hyderabad, India

Pratima Manhas ECE Department, FET, Manav Rachna International Institute of Research & Studies, Faridabad, India

S. M. Mansoor Roomi Department of ECE, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

A. Merline Department of ECE, Sethu Institute of Technology, Tamil Nadu, India

Jubin Mitra IIT Kanpur, Kanpur, Uttar Pradesh, India

K. Muralidharan Department of ECE, Coimbatore Institute of Technology, Coimbatore, India

K. Muthukumaran Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Sandeep Varma Nadimpalli Department of Information Science and Engineering, BMS College of Engineering, Bengaluru, Karnataka, India

A. Nagaraj Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Gollu Appala Naidu Department of Electronics and Communication Engineering, National Institute of Technology Jamshedpur, Jamshedpur, Jharkhand, India

Srinath R. Naidu Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

Manju Nanda Aerospace Electronics and System Division, CSIR-NAL, Bangalore, India

M. Narayanan Department of Computer Science and Engineering, Malla Reddy College of Engineering, Secunderabad, Hyderabad, Telangana, India

P. Naveen Kumar Department of ECE, University College of Engineering, Osmania University, Hyderabad, India

Tunirani Nayak Department of Electronics and Telecommunication Engineering, VSSUT, Burla, Sambalpur, Odisha, India

S. M. Nazia Fathima Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Shubham Nema Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Sagar Oza Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Khushboo Pachori Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India;
Electronics and Communication Engineering, Gurunanak Institute of Technology, Hyderabd, Telengana, India

Gunjan Parihar Goa College of Engineering, Ponda, Goa, India

Akshay Parmar Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

V. Parthasarathy Department of ECE, Sethu Institute of Technology, Tamil Nadu, India

Zuber M. Patel Sardar Vallabhbhai National Institute of Technology Surat, Surat 395007, India

Suman Patil Department of Information Science and Engineering, BMS College of Engineering, Bengaluru, Karnataka, India

Raj Kumar Patra CMR Technical Campus, Hyderabad, India

M. D. Pawar Department of Electronics and Telecommunication Engineering, Maharashtra Institute of Technology, Aurangabad, India

Bhakti V. Pawar Department of Electronics and Telecom, Ramrao Adik Institute of Technology, Navi Mumbai, India

Anju S. Pillai Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Binod Kumar Prasad ECE Department, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

L. Pratap Reddy Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University Hyderabad College of Engineering, Hyderabad, India

B. S. Priyanka Chowdary Department of ECE, RVCE, Bengaluru, India

G. Prutha Department of EEE, RV College of Engineering, Bengaluru, India

N. Puviarasan Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

S. Radharani Department of CSE, University College of Engineering, Osmania University, Hyderabad, India

K. Ragini Department of ECE, G. Narayanamma Institute of Technology and Science, Hyderabad, India

Rahul EC Department, Delhi Technological University, Delhi, India

Rohit Raja Sreyas Institute of Engineering and Technology, Hyderabad, India

M. Rajan Babu Department of ECE, Lendi Institute of Engineering and Technology, Vizianagaram, India

Kapoor Rajiv EC Department, Delhi Technological University, Delhi, India

P. Rajshekhar Rao Department of Avionics, IST, JNTUK, Kakinada, Andhra Pradesh, India

P. Ramana Reddy Department of ECE, JNTUACEA, Anantapuram, Andhra Pradesh, India

M. Ramkumar Raja ECE Department, Coimbatore Institute of Engineering and Technology, Coimbatore, India

Md. Rashid Mahmood Department of Electronics and Communications, Guru Nanak Institutions Technical Campus, Hyderabad, India

B. V. Ravi Kumar CBIT, Hyderabad, India;
Research Centre Imarat (RCI), Hyderabad, India

G. Reddy Hemantha Department of ECE, Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India

P. Rohini Department of Computer Science and Engineering, Faculty of Science & Technology, ICFAI Foundation for Higher Education (A Deemed-to be University), Hyderabad, India;
Department of CSE, JNTUA, Anantpuramu, India

M. Rohitha ECE, SPMVV, Tirupathi, Andhra Pradesh, India

T. Ruba Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

B. Sahana Department of ECE, RVCE, Bengaluru, India

B. Sanjana Department of ECE, G. Narayanamma Institute of Technology and Science, Hyderabad, India

Sripadam Satish NIT, Tiruchirappalli, Tamil Nadu, India

Dhaval Shah Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

V. Shavali Department of ECE, JNTUA, Anantapuramu, Andhra Pradesh, India

C. Shoba Bindu Department of CSE, JNTUACEA, Anantpuramu, India

Sandeep Shukla IIT Kanpur, Kanpur, Uttar Pradesh, India

M. Shyama Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

K. Sindhu Priya CBIT, Hyderabad, India

Amardeep Singh Department of Computer Science and Engineering, Punjabi University, Patiala, Punjab, India

Chamandeep Singh NIT, Tiruchirappalli, Tamil Nadu, India

Niharika Singh Department of Instrumentation Engineering, SGGSIE&T, Nanded, Maharashtra, India

Suryansh Singh Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

G. R. Sinha Myanmar Institute of Information Technology, Mandalay, Myanmar

M. Sinthuja Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India

E. Sreenivasa Rao Vasavi College of Engineering, Hyderabad, India

G. M. Sreerama Reddy Department of ECE, CBIT, Kolar, Karnataka, India

E. Sri Harshitha ECE, SPMVV, Tirupathi, Andhra Pradesh, India

K. Sridevi Department of ECE, CIT Sandwich Polytechnic College, Coimbatore, India

Kone Srikrishnaswetha Department of Electronics and Communications, Sreyas Institution of Engineering and Technology, Hyderabad, India

S. P. V. Subba Rao Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India

K. Sukanya Department of ECE, TKR College of Engineering and Technology, Ranga Reddy, Telangana, India

D. Suresh Department of Electrical and Electronics Engineering, Vignan Institute of Technology and Science, Hyderabad, India

B. Suresh Kumar CBIT, Hyderabad, India

K. Susheela ECE, SPMVV, Tirupathi, Andhra Pradesh, India

T. Swapna Rani ECE, CMR College of Engineering & Technology, Medchal, India

R. Tamilselvi Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India

Shaveta Thakral ECE Department, FET, Manav Rachna International Institute of Research & Studies, Faridabad, India

Gadhamsetty Ravi Theja Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru, India;
Amrita Vishwa Vidyapeetham, Coimbatore, India

M. M. Tripathi EE Department, Delhi Technological University, Delhi, India

C. Vani Department of CSE, University College of Engineering, Osmania University, Hyderabad, India

S. Varadarajan Department of ECE, SVU College of Engineering, S V University, Tirupati, India

Pavithra K. Yadav Department of Information Science and Engineering, BMS College of Engineering, Bengaluru, Karnataka, India

Yamini Devi Ykuntam Department of ECE, Lendi Institute of Engineering and Technology, Vizianagaram, India

Part I
Signal and Image Processing

Detection and Classification of Exudates and Non-exudates in Retinal Images

R. Tamilselvi, M. Parisa Beham, A. Merline and V. Parthasarathy

Abstract The retina of human eye plays a key function in the vision, and it is a light-sensitive layer. The optics of eye produces an image figure in the retina. The various eye diseases like diabetic retinopathy, myopia, macular pucker, and macular hole have an effect on the retina. The retina is affected by these diseases which are vascular disease and cause vision mutilation and blindness. These diseases happen due to diabetics, aging, and nearsightedness. Exudates are the pathological condition of the retina. So the early detection of these is very important. In the paper, an efficient methodology like Otsu thresholding method and the *K*-means clustering method is proposed for the recognition of exudates. After detecting the exudates, various texture feature extraction processes are involved. Finally, the classification method is performed using Backpropagation Neural Networks (BPN). The main spotlight of the projected work is to develop algorithms for exudates recognition and categorization of retinal images in pathological or non-pathological, convalescing investigation of the fundus images. The experimental results acquired from the projected method of extracting the features and classification method exposed that non-diseased cases were recognized with 90% exactness while temperate and severe cases were 99% accurate.

Keywords Retinal images · Diabetic retinopathy · Neural networks · Exudates · Detection and classification

1 Introduction

Digital images of the fundus image provide valuable information regarding pathological changes caused by the human eye. Diabetic associated eye disease is a major origin of avertable blindness. Diabetes can cause flagging in the body's blood vessels. The slim layer of nerve tissue is the retina covering the reverse of the eye that

R. Tamilselvi (✉) · M. Parisa Beham · A. Merline · V. Parthasarathy
Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India
e-mail: rts.ece@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_1

identifies light and expresses information to the brain. The macula is the principal region of the retina and is used for fine detail. Formation of hole in the macula affects vision, mostly when sensing and achieving other visually challenging tasks, but it does not damage total blindness. Macular hole commonly affects the people in aged 55 and common in women. Myopia is another eye disorder referred to as "short sight" or "nearsightedness". Myopia initiates in early days and distance vision may happen to blur so gradually, so that the child is unconscious he or she does not see as clearly as other groups do. Some people practice tardy arrival of myopia, promising as a problem in middle age even they have not experienced vision difficulties in their childhood. A lot of technology is growing in daily life to analyze the various pathological conditions of the eye. Among those, the commonly occurring diseases are Diabetic Retinopathy (DR), macular hole, myopia, and macular pucker [1].

The aim of our task is to develop a system that senses and identifies exudates and non-exudates using digital fundus retinal images. Morphological methods recognize exudates and classified using BPN. In previous research, the accuracy in finding a disease using image segmentation is difficult. So Otsu thresholding is used and *k*-means clustering is proposed to detect the abnormal retina and BPN for classification with more accuracy with the Gray-Level Co-occurrence Matrix (GLCM) in feature extraction.

2 Related Works

There are a lot of researches going on in detection of exudates and non-exudates. Some of the research related to the same is reviewed here. In [1], the authors developed a method using Probabilistic Neural Network (PNN) classifier to automatically detect exudates and non-exudates. It provides an accuracy of 96.9% and improved memory consumption. In [2], the authors describe the plan of detection system for retinal images using neural network. The most secure method for the recognition of an identity is the retina-based recognition. The author discussed retina image acquisition, extraction of some important features, and classification. Extracted features are given as input to the various neural network classifiers. The performance comparison is done for the images from DRIVE, STARE, and MESSIDOR, and the segmentation results of DRIVE and MESSIDOR databases are shown and the performance measures are estimated to evaluate their proposed method.

In [3], the author addressed the trouble of distinguishing true vessels as an after processing action to vascular structure categorization. The analyzed solutions revealed that the BPN classifier is approximately more effective when compared with the SVM classifiers. The efficiency achieved is based on the parameters used and the feature set. In [4], the author described one of the neural network algorithms, BPN. The hint behind the projected method is quite easy, and the output is experienced against the desired output. If results are not acceptable, then the weights between layers are adjusted, and the process is constantly processed frequently until the error becomes very less. In [5], the authors described the segmentation task as

a multi-label inference task and observed the hidden advantages of the grouping of convolution neural networks and planned prediction. In this paper, the convolution neural network-based replicas accomplish the best performance and significantly adapt the advanced method for automatically segmenting the retinal blood vessels on DRIVE dataset with the accuracy of 95.33%. In [6], the author explained the BP algorithm based on a graphical approach.

Motivated by the issues observed from the related works, the main offerings of our proposed work are as follows:

- Otsu thresholding and *k*-means clustering are proposed to detect the exudates.
- Gray-Level Co-occurrence Matrix (GLCM) in feature extraction and selection.
- BPN classification for the classification of pathological conditions such as diabetic retinopathy, myopia, macular hole, and pucker.

3 Proposed Methodology

The proposed methodology is shown in Fig. 1. The proposed method consists of two phases. Phase I comprises acquisition of image, preprocessing, and segmentation. Phase II consists of extraction of important features and classification process.

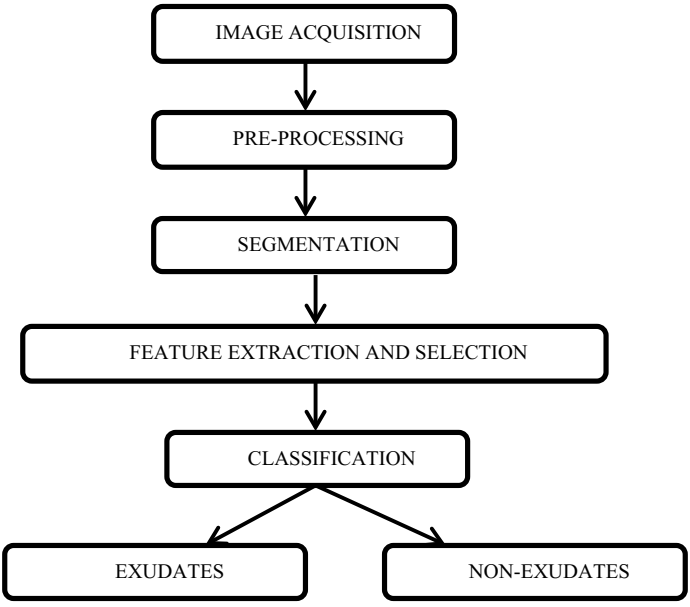


Fig. 1 Flow diagram of the proposed work

3.1 Image Acquisition

The various retinal images are retrieved from the publically available datasets such as DIARETDB1 and STARE. The four kinds of diseases are considered for the analyses such as are diabetic retinopathy, macular hole, myopia, and macular pucker.

3.2 Preprocessing

Denoising of image is an imperative component of image reconstruction process. Various noises get added during acquirement, transmission, reception and storage, and recovery processes. Without any compromise in the image, denoising operation is performed to eliminate noises with preserving edges. Peak Signal-to-Noise Ratio (PSNR) and Mean Square Error (MSE) are the parameters to describe the image quality [7].

3.2.1 Parametric Description

The performance parameters are most chief criteria to validate the output simulation results. The PSNR and MSE are considered as evaluation parameters. The MSE is explained in Eq. 1,

$$\mathrm{MSE} = 1/m \times n \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, J) - K(i, j)]^2 \tag{1}$$

where $m \times n$—a noise-free monochrome image I, and K—its noisy approximation. The PSNR (in dB) is characterized in Eq. 2,

$$\mathrm{PSNR} = 10 \log_{10}\left(\frac{\mathrm{MAX}_\mathrm{I}^2}{\mathrm{MSE}}\right) \tag{2}$$

where MAX_I is the maximum probable pixel value of the image. Three filters such as median filter, averaging filter, and Wiener filter are used for denoising the images. Results can be exposed through assessment among them [7].

3.3 Retinal Image Segmentation

Image segmentation is the most imperative and hard crisis in image analysis. In computer vision, image segmentation is the method of separating an image into useful parts or objects [8].

3.3.1 Otsu Thresholding

Otsu thresholding is a significant method to select enough threshold of gray level for detaching the objects from their background. Otsu is a usual entry selection region based segmentation method. Otsu method is a worldwide accepted thresholding method which depends only on the gray value of the image [9].

3.3.2 *K*-means Clustering

Another efficient segmentation method is a grouping of clusters. *K*-means clustering care for each thing as having a position in space. The method features the partitions, so that objects within each cluster are as close to each other as probable and as far from objects in other clusters as probable. *K*-means clustering describes the total number of clusters and a distance metric to quantify how two objects are very much intimate to each other [10].

3.4 Feature Selection and Extraction

The segmented images are classified into exudates and non-exudates by analysis of their features to make it effective. Image textures are extracted using GLCM. The textural characteristics are correlation, contrast, energy, homogeneity, and entropy [11].

3.5 Classification Using BPN

Backpropagation is a training method used for a multilayer neural network. It is a gradient descent method which minimizes the total squared error of the output computed by the net [3]. Any neural network is anticipated to react properly to the input patterns. The training of a neural network by backpropagation takes place in three stages. After the training of neural network, the neural network has to compute the feed forward phase only. Even if the training is slow, the trained net can produce its output immediately. For classification new data, known input values and output values are enough.

4 Results and Discussion

From the preprocessing comparison results, best filter is used for image segmentation to analyze the affected area in retina [8]. The comparison of filters is performed with

the help of MSE and PSNR. The simulation result of preprocessing is shown with the parameters MSE and PSNR considered. The proposed method is evaluated using publically available DIARETDB1 database [11].

4.1 Results of Filtering Techniques

The Filtered outputs of the Diabetic Retinopathy Image are shown in Fig. 2. The above results showed that the average filter works well for pepper noise with the PSNR of 51.94 dB. For the same image, median filter holds good for the salt-and-pepper noise and Poisson noise with the PSNR values of 52.07 and 52.33 dB, respectively, whereas Wiener does not showed much improved performance for all types of noises. The above results showed that the average filter showed the performance with the PSNR of 42.65 dB for Poisson noise. For the same image, median filter holds good for the salt-and-pepper noise and Poisson noise with the PSNR values of 43.85 and 44.24 dB, respectively, whereas Wiener does not showed much improved performance for all types of noises (Fig. 3).

The Filter output of macular hole image is shown in Fig. 3. For the macular pucker affected eye, the filtered output results are shown in Fig. 4. The above results

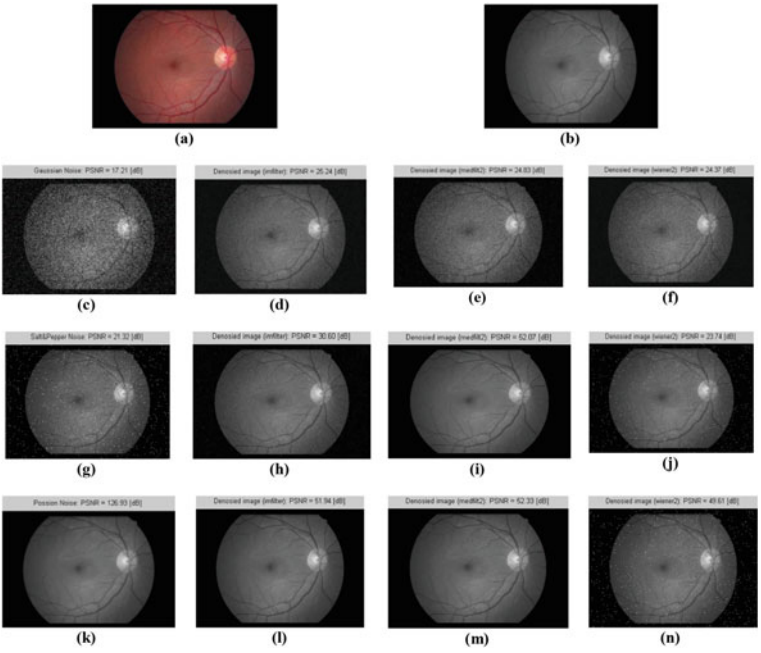


Fig. 2 Filter outputs of diabetic retinopathy Image1

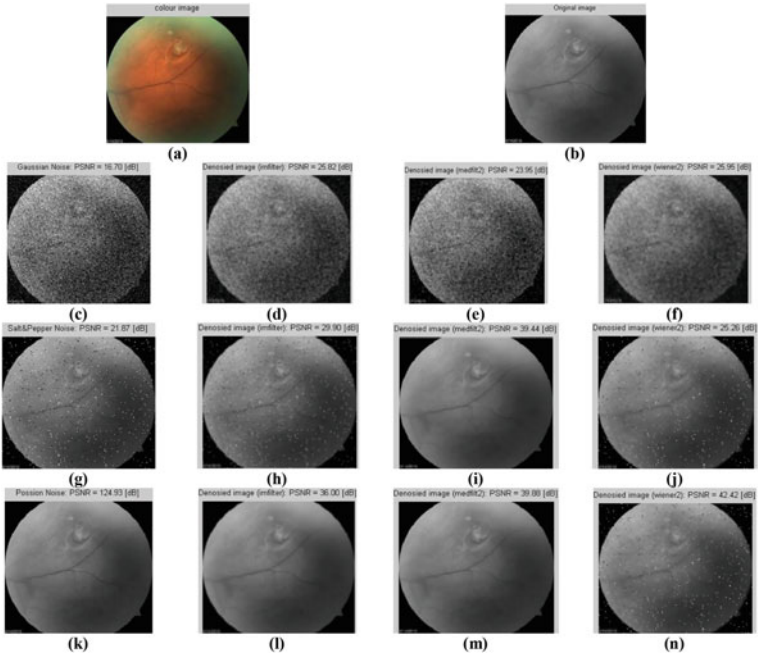


Fig. 3 Filter output of macular hole image

showed that the average filter works well only for the Poisson noise with the PSNR of 44.35 dB.

For the same image, median filter showed the performance for the salt-and-pepper noise and Poisson noise with the PSNR values of 46.81 and 47.70 dB, respectively, whereas Wiener showed better results with PSNR value of 42.95 dB for the Poisson noise. The above results showed that the average filter works well only for the Poisson noise with the PSNR of 43.35 dB. The PSNR values of the diseased eye images are tabulated and shown in Table 1.

4.2 Results of Segmentation

This section describes the segmented results of various eye images. As stated, two methods are used for segmentation purposes such as Otsu thresholding and *k*-means clustering method. The preprocessed output of various images is given as input for the segmentation process. The segmentation output of retinopathy image is shown in Fig. 5a, b using Otsu thresholding method and *k*-means clustering method respectively. Similarly, the segmented result of Macular Hole is shown in Fig. 6a, b.

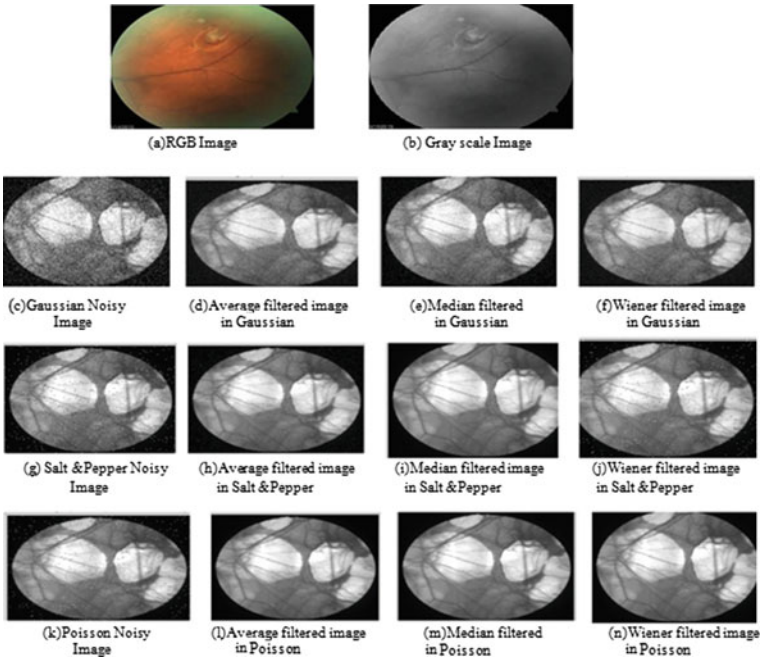
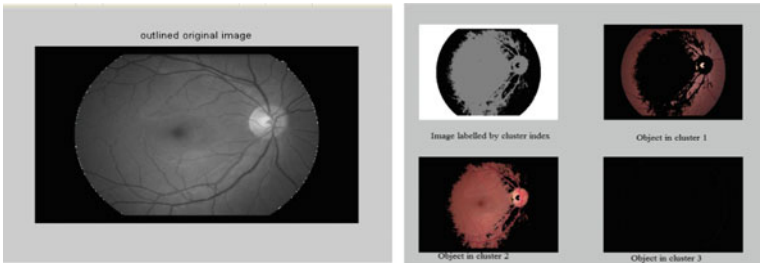


Fig. 4 Filter output of macular pucker image



(a) Otsu Thresholding (b) K-means clustering

Fig. 5 Segmentation of Diabetic Retinopathy Image1

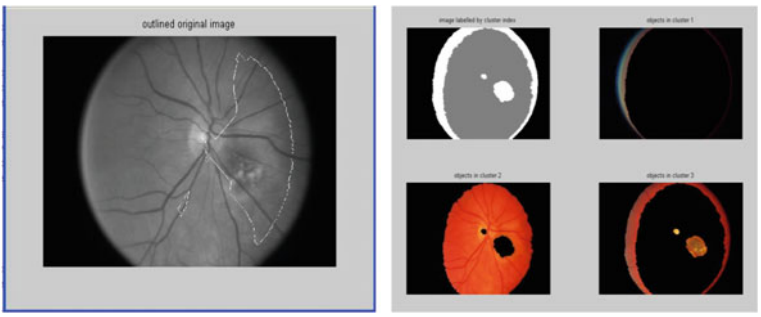
The segmentation result using Otsu threshold method is shown in Fig. 7a for Myopia affected eye image. Figure 7b shows cluster index image of Myopia and the objects in three clusters using *k*-means clustering method sequentially. The segmentation result using Otsu threshold method is shown in Fig. 7c for Macular pucker image. Figure 7d, shows cluster index image of macular pucker and the objects in three clusters using *k*-means clustering method sequentially.

Table 1 PSNR (dB) values of filtering methods

Images	Noise type	Filtering techniques		
		Average	Median	Wiener
Diabetic retinopathy Image1	Gaussian noise	25.24	24.83	24.37
	Salt-and-pepper noise	30.60	52.07	23.74
	Poisson noise	51.94	52.33	49.61
Myopia	Gaussian noise	26.52	23.80	26.75
	Salt-and-pepper noise	31.17	43.85	25.69
	Poisson noise	42.65	44.24	38.77
Macular hole	Gaussian noise	25.82	23.95	25.95
	Salt-and-pepper noise	29.90	39.44	25.26
	Poisson noise	36.00	39.88	42.42
Macular pucker	Gaussian noise	26.24	24.36	25.63
	Salt-and-pepper noise	30.84	46.81	24.40
	Poisson noise	44.35	47.70	42.95
Diabetic retinopathy Image2	Gaussian noise	25.93	24.49	25.26
	Salt-and-pepper noise	30.90	43.85	24.46
	Poisson noise	43.35	44.02	42.76

4.3 Results of Feature Extraction

As we discussed above, the feature extraction process using GLCM is extracted for five eyeball images. The basis feature values are tabulated and shown in Tables 2 and 3 for Otsu thresholding and *k*-means clustering technique, respectively. These features are given as input to the BPN.



(a) Otsu Thresholding method **(b) K-means clustering method**

Fig. 6 Segmentation of Image

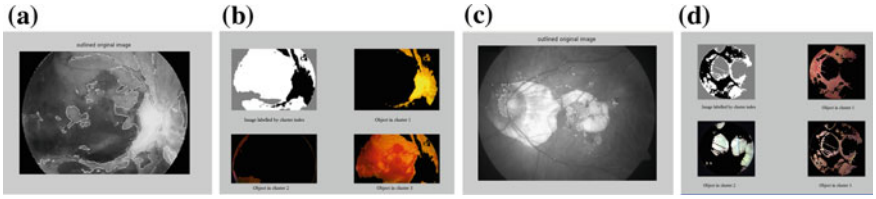


Fig. 7 **a**, **b** Otsu thresholding and *k*-means clustering results of myopia; **c, d** Otsu thresholding and *k*-means clustering results of macular pucker image

Table 2 GLCM features for retinal images—Otsu thresholding

Images	Contrast	Correlation	Energy	Entropy	Homogeneity
Diabetic retino Image1	0.1018	0.9691	0.3019	4.9425	0.9803
Myopia	0.7966	0.8636	0.1254	7.2502	0.9063
Macular hole	0.2342	0.9374	0.1694	6.7684	0.9500
Macular pucker	0.2978	0.9679	0.1513	7.2301	0.9435
Diabetic retino Image2	0.1792	0.9555	0.2774	5.9906	0.9556

Table 3 GLCM features for retinal images—*K*-means thresholding

Images	Contrast	Correlation	Energy	Entropy	Homogeneity
Diabetic retino Image1	0.2453	0.9359	0.5298	2.3744	0.9742
Myopia	0.3243	0.8667	0.4834	2.8521	0.9550
Macular hole	0.1357	0.9410	0.5225	2.6317	0.9781
Macular pucker	0.9328	0.9290	0.3897	3.3187	0.9558
Diabetic retino Image2	0.2550	0.9794	0.3515	4.1742	0.9668

Based on Tables 2 and 3, it is seen that all the features are almost the same for Otsu thresholding and *K*-means clustering method. So average of the features can be given as input to BPN.

4.4 Results of Classification

Output nodes obtain better results by modifying the weighted connection through the training data in the learning phase. Here the classification is done for five images. 100 to 50 epochs are taken for iteration. The network parameters are gradient descent and error tolerance. In the classification phase, size of the gradient descent and error tolerance is chosen as 0.1 and 0.01 respectively. Test data includes the calculation

of true positive, true negative, false positive, and false negative. True Positive (TP) relates to the quantity of exactly that is expected by the classifier. False Negative (FN) relates to the quantity of wrongly predicted samples as destructive by the classifier. False Positive (FP) resembles the quantity of nonpositive samples erroneously expected as positive by the classifier. True Negative (TN) is the amount of negative samples correctly expected by the classifier. TP Rate = TP/(TP + FN) and FP Rate = FP/(TN + FP). Precision is the section of samples that essentially turns out to be positive in the group the classifier has acknowledged as a positive class. Precision = TP/(TP + FP). The precision or accuracy calculated for the normal cases is 90%, and for the pathological images such as diabetic retinopathy, myopia, macular hole, and macular pucker, it is 99%.

5 Conclusion

Early detection and timely treatment of retinal diseases can slow down the progression of the disease and avert blindness. The preprocessing is performed to get a better quality of image. The denoised image is further segmented using Otsu method and the *k*-means clustering method. Finally, the segmented images are classified using BPN algorithm. After exudates detection process, we realize the retinal image classification in pathological and non-pathological. The pathological conditions include any one of the diseases such as DR, macular hole, myopia, and pucker. The previous classification results showed the accuracy of 85–95%. The results obtained from the proposed method exposed that normal images were classified with 90% accuracy and moderate or severe cases were detected with 99% accurate.

Ethical Approval The database "DIARETDB1—Standard Diabetic Retinopathy Database" used in this paper is a public database for benchmarking diabetic retinopathy detection from digital images. The main objective of the design has been to unambiguously define a database and a testing protocol which can be used to benchmark diabetic retinopathy detection methods. The database can be freely downloaded and used for scientific research purposes, and it was ethically approved.

References

1. Vijaya Kumar HS, Bharathi PT, Madhuri R (20161) A novel method for image analysis and exudates detection in retinal images. Int J Adv Res Innov 4(1)
2. Manoj S, Muralidharan SPM (2013) Neural network based classifier for retinal blood vessel segmentation. Int J Recent Trends Electr Electron Eng
3. Nithya KA, Rajini A (2014) Classification of normal and abnormal retinal images using neural networks. Int J Adv Res Comput Eng Technol 3(9)
4. Cilimkovic M. Neural networks and back propagation algorithm, institute of technology blanchardstown
5. Dasgupta A, Singh S. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation

6. Singh P, Shree R (2016) A comparative study to noise models and image restoration technique. Int J Comput Appl 149(1)
7. Tania S, Rowaida R (2016) A comparative study of various image filtering techniques for removing various noisy pixels in aerial image. Int J Signal Process Image Process Pattern Recognit 9(3)
8. Chahar PS, Thakare VV (2015) Performance comparison of various filters for removing gaussian and poisson noises. Int Res J Eng Technol 02(05)
9. Patil AB, shaikh JA (2016) OTSU thresholding method for flower image segmentation. Int J Comput Eng Res 06(05)
10. Date MK (2013) Brain image segmentation algorithm using K-Means clustering. Int J Comput Sci Appl 6(2)
11. Kauppi T, Kalesnykiene V, Kamarainen J-K, Lensu L, Sorri I, Raninen A, Voutilainen R, Uusitalo H, Kälviäinen H, Pietilä J. DIARETDB1 diabetic retinopathy database and evaluation protocol. Technical report

Performance Analysis of Nanoparticles in Healthcare and Biomedical Applications

T. Ruba, R. Tamilselvi, M. Parisa Beham and K. Muthukumaran

Abstract Applications of nanoparticles in health care have recently turned into the most emerging research platforms, since the nanomaterials exhibit novel and superior properties than the conventional materials, which make it possible for new medical applications. Extensive biomedical applications are explored by the nanoparticles domains which are principally either purely organic or inorganic materials. Inspired by the unique characteristics, novelty, and wide applications of nanoparticles, this work provides a comprehensive survey on performance analysis of nanomaterials in biomedical applications. Analysis of new results, techniques, and characteristics of nanomaterials mainly in the area of biomedicine are well described.

Keywords Biomedical applications · Health Care · Nanomaterials · Nanoparticles · Nanostructures · Nanosensitizers · Nanoprotectors

1 Introduction

The advancement of nanoscience and nanotechnology has much engrossed the attention of modern researchers in the current scenario. In the last period, a wide range of investigations have been completed on nanomaterials in biomedical applications, involving hallucination, exploitation, and fabrication of materials on the range of 100 nm down to 1 nm. Nanoscale particles are progressively providing major influence on humanoid health. So they are extensively used in the applications of diagnostic and therapeutic. The exploit of nanomaterials in the arena of biomedicine opens up many opportunities in the brawl against each and every kind of cancer, tumor and cardiac detection, bone marrow generation, and other diseases. Generally, the nanomaterials are categorized into three types based on the dimension: (i) quantum dots, zero-dimensional materials, having variations in shape and diameter; (ii) nanorods

T. Ruba (✉) · R. Tamilselvi · M. Parisa Beham · K. Muthukumaran
Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India
e-mail: Ruba.Thanapandiaraj@gmail.com; rubasuresh20@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_2

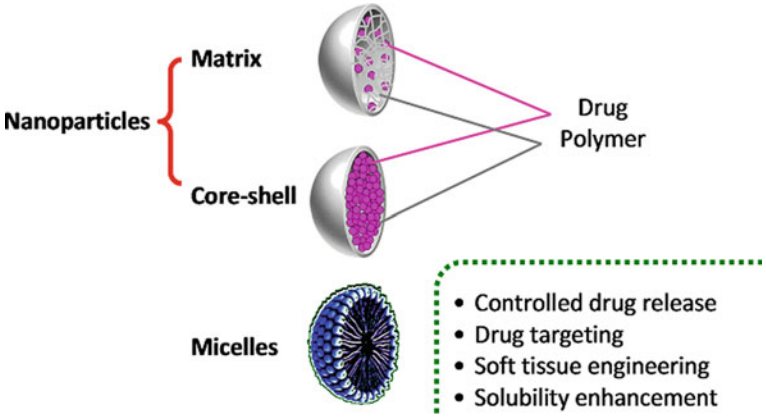


Fig. 1 Polymeric nanoparticles-based nanocomposite hydrogels

and nanowires are of one-dimensional materials; and (iii) nanobelts, nanodisks, and nanosheets are in the variety of two-dimensional materials.

In this paper, our focus is on advanced nanoparticles (NPs) used in different biomedical applications. Their success rate is also very high compared to other existing materials used for biomedicine [1]. The properties of NPs such as ultra-small size (refer Fig. 1), tunable optical emission, enhanced mechanical properties, extremely high surface area-to-volume ratio, and super-paramagnetic behavior make them turn into popular in biomedical applications. Multifunctional NPs are also used in diagnostic and therapeutic applications. There are lots of nanomaterials presents particularly for clinical applications. Among those, high-quality nanomaterials with controlled size and shape will lead to tuned responses. The selection of nanomaterials based on their properties should be suited for any application. Most common properties of nanomaterials needed for the applications of biomedicine are extremely smaller size, appropriate particle surface, controlled shape, ease of chemical modification, and good biocompatibility. Though a multiplicity of nanoparticles has already faced the biomedical applications problem, it is yet a developing field depending on the application scenario and biomedicine issues. Such a proliferating field requires review of nanoparticles and materials available as on date to enable researchers to choose from.

This paper is organized as follows: Sect. 2 describes the analysis and synthesis of nanoparticles. Section 3 explains how the Nanoparticles are involved in MRI imaging. Nanoparticles as radiosensitizers and radioprotectors for radiotherapy treatment are detailed in Sect. 4. Section 5 accomplishes the work with detailed inference from the related reviews.

2 Analysis and Synthesis of Nanoparticles

In the recent decade, more and more NPs have been found, experimented, and analyzed by the researchers, particularly in biomedical applications. Applications of nanomaterials in biomedical and healthcare research are shown in Fig. 2. Maier-Hauff et al. [2] used and analyzed the NPs for hyperthermia treatment. Nano auroshell is proposed by Leung et al. [1] that destroys the target tumor cells by releasing heat upon absorption of near-infrared (NIR) light. In [3], the silver and copper nanoparticles are produced on glass slides by using magnetron sputtering. In magnetron sputtering with the sputter time and plasma power, the size and concentration of the nanoparticles can be easily controlled. Higher plasma power is also required for the synthesis of silver nanoparticles to cultivate much earlier than copper nanoparticles.

Zamperini et al. [4] synthesized and characterized the pure hydroxyapatite (HA) and hydroxyapatite decorated with silver (HA@Ag) nanoparticles to analyze the antifungal effects. Hu et al. [5] described that the (C-dots) carbon nanodots with <10 nm of typical size are newly exposed fluorescent carbon nanoparticles. All the physical and chemical characteristics of these C-dots are excellent. Cancerous tissue detection of systemic emission of chemotherapeutic agents ruins a major issue in cancer treatment. To address this issue, Brann et al. [6] proposed a near-infrared responsive oligonucleotide-coated (AS1411, hairpin, or both) gold nanoplate loaded with doxorubicin (DOX). This becomes nontoxic without the trigger release. After 5 min of exposure to laser, these particles become intensely toxic to cells in order to activate DOX discharge. They also proved that untriggered DOX-loaded conjugate with no laser exposure consequences in slight to no toxicity for the exposure of the

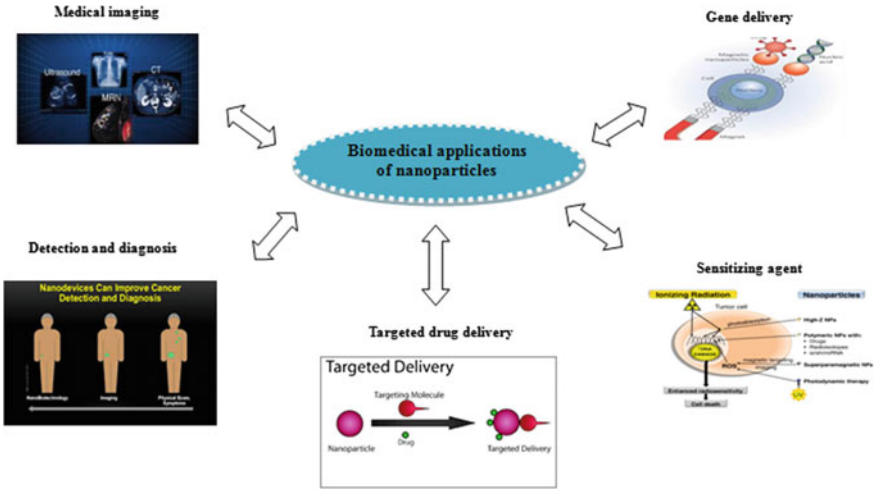


Fig. 2 Applications of nanomaterials in biomedical field of research

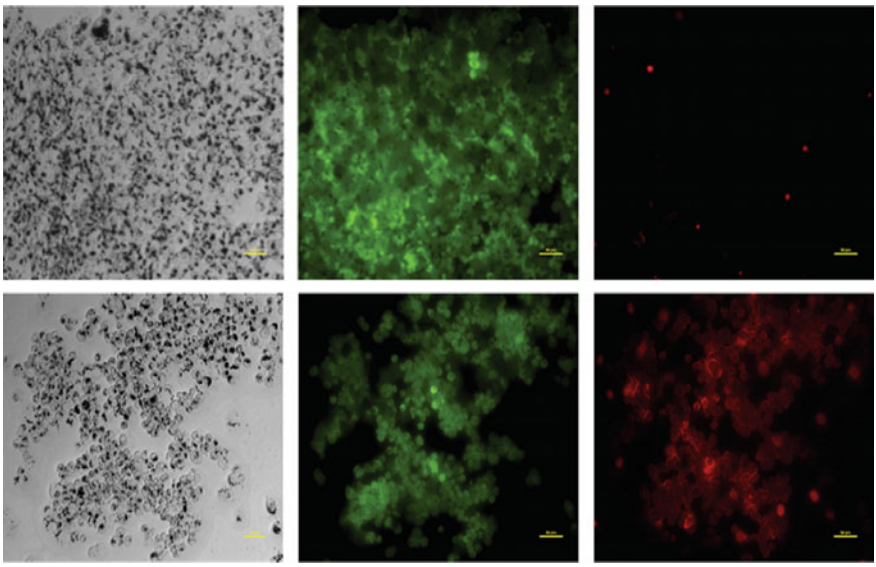


Fig. 3 Bright field (left), green fluorescence (live cells), and red fluorescence (dead cells). Top row: GNP-Hairpin-DOX (1 OD), no laser. Bottom row: GNP-Hairpin-DOX (1 OD), with laser treatment

cells, while laser-triggered release of DOX causes substantial cell death as shown in Fig. 3.

In recent years, colonic-targeted drug delivery system is widely used to combat colon cancer.

Dicer-substrate small interfering RNA (DsiRNA) had used for cancer therapy owing to its strength in aiming particular gene of interest, though its usage is controlled due to poor cellular uptake and rapid degradation. To tackle this, Katas et al. [7] applied chitosan–graphene oxide (CSGO) nanocomposite to distribute DsiRNA efficiently into cells. Furthermore, pectin was used as compatibilization agent to permit specific release to the colon and save the nanocomposites from the insensitive surroundings in the stomach and small intestine.

Also, more works have been going on to obtain higher therapeutic efficiency, and the inventive combination of phototherapies with other existing treatments might reduce the intolerable toxicity in normal tissues while sustaining the desired tumor clampdown effect.

3 Nanoparticles for MRI Imaging

There are various conventional imaging techniques for diagnosing cancer. Among those, magnetic resonance imaging (MRI) delivers the finest resolution than other techniques and is noninvasive or at least minimally invasive. In addition to that, usage

of contrast agent (CA) in MRI further reduces its contrast problem and increases its sensitivity. In order to enhance and improves the quality of MRI images, superior nanoparticles are used as a contrast agent. This type of imaging technique is called nano-MRI. Nowadays, iron oxide nanoparticles have a major role in improving the contrast of medical imaging.

Lawaczeck et al. [8] conclude that small and ultrasmall superparamagnetic iron oxide nanoparticles (SPION and USPION) are the foremost to be applied in MRI contrast enhancement. SPIONs are usually monocrystalline composed of magnetite (Fe_3O4) or maghemite ($-Fe_2O_4$). Parkes et al. [9] proposed that gold (Au@Fe), cobalt (Co), and platinum (Pt@Fe) have much greater saturation magnetization value than the iron. They have a greater extended effect on proton relaxation ($r1 = 7.4\ mM^{-1}\ s^{-1}$, $r = 88\ mM^{-1}\ s^{-1}$ for copolymer at 1.5 T, 3.9 nm core diameter, 28 nm particle diameter) providing better contrast compared to iron oxide with equal concentration and permit minor particle to be employed without compromising MRI sensitivity. Kim et al. [10] employed super-paramagnetic metal oxide nanoparticles with high magnetization values for improving the contrast of T2 (threshold weighted) images. Zhou et al. [11] suggested that MRI is a clinical imaging technique for effective functional imaging of defected soft tissues, and also for imaging solid tumors. Li et al. [12] constructed uniform, monodisperse, and core–shell structured nanospheres composed of mesoporous silica-coated individual upconversion luminescence/magnetic nanoparticle. This nanosphere is applied for MRI dual modal imaging and for anticancer drug delivery. Rosa et al. [13] dealt with the problems raised in molecular imaging and nano-MRI techniques. It is presented that the monitoring of biological processes and biochemistry via early identification through cell and molecule visualization are two applications of molecular imaging (MI). After the MI era, medical diagnosis tremendously grew on imaging technologies such as MRI which visualize tissues in vivo and also the activity of the brain called as functional MRI (fMRI). Even though this functionality is enough for imaging tissue arrangements, it cannot readily visualize molecules where the resolution required is in the scale of a nanometer. Recently, nano-MRI observes structures prospectively which cannot be envisioned by molecular imaging [14]. Thus, the authors accomplished that the nano-MRI technique which has sensitivities many times superior than MRI technique. This is the domain in which nano-MRI is emerging with a variety of techniques to support molecular biology in attaining single-molecule resolution.

4 Nanoparticles as Radiosensitizers and Radioprotectors for Radiotherapy Treatment

The accurate detection of cancer cell in any imaging modality will lead to a successful treatment for cancer patients. Radiotherapy places a vigorous role in diagnosing cancer. The objective of successful radiotherapy is to rises the sensitivity of cancer cells. Similarly, it minimizes the sensitivity of normal cells. This may be attained

either by enhanced localization of dose or by using radioprotectors for healthier cells and/or radiosensitizers for cancer cells. Radioprotectors are intended to minimize the destruction in healthy tissues during irradiation. Radiosensitizers are used to increase the sensitivity of the cancer cells during irradiation. There are several conventional chemical agents that are used as radioprotectors and radiosensitizers, e.g., amifostine, carbogen, nicotinamide, and betulinic acid. Although they have been potentially used for effective treatment, there are some limitations regarding size, biocompatibility, and chemical modification.

In recent years, due to the development in the arena of nanotechnology and superior properties of nanoparticles, novel nanoparticles have been used as radioprotector and radiosensitizers. Chandra et al. [15] described the methods for synthesizing gold nanoparticles (AuNPs). They also described the applications of those for detecting various diseases using electrochemical detection methods. In [16], they use hydrothermal reaction method for synthesizing superparamagnetic zinc ferrite spinel ($ZnFe_2O_4$) nanoparticles. The synthesized material is then used in cancer therapy as a radiosensitizers. The result of this work indicated that these $ZnFe_2O_4$ nanoparticles increased the radiotherapy effectiveness, as well as it can be easily alienated from the cell environment through an external magnetic field after the radiotherapy. Werner et al. [17] demonstrated that Genexol-PM is better than Taxol as a radiosensitizer and holds high potential for clinical application. Popov et al. [18] addressed that the administration of CeO_2 nanoparticles increases the time of endurance of the mice which is exposed to X-ray radiation. Feliciano et al. [19] established the nanoparticles which are conjugated with nitroxide radicals (redox nanoparticles, RNP^N). This effectively scavenged the radiation-induced reactive oxygen species (ROS) that the conventional low molecular weight (LMW) antioxidants. All of these radioprotectors and radiosensitizers show its potential towards the successful treatment. Performance of various nanoparticles for different biomedical applications is shown in Table 1.

5 Conclusion

In the recent scenario, advanced nanomaterials have an imperative role in medical imaging, such as sensitizing agents in diagnosis, targeted drug delivery, and gene delivery. The use of nanomaterials is not restricted to these issues in biomedical field. It may extend in various research fields like health and safety, environmental and industrial applications, and societal impacts. It substantiates that in future, almost all technologies will come under nanotechnology due to its superior properties like smaller size, increased power storage, reduced power consumption, etc. In a nutshell, various nanoparticles were effectively developed and still more can be developed, and they have a great potential in all therapy treatments in healthcare applications.

Table 1 General performance of nanomaterial-based biomedical applications

S. No.	Author	Year	Used nanoparticles	Agent	Application	Remarks
1	Katas et al. [7]	2017	Chitosan–graphene oxide (CSGO) nanocomposite	As therapeutic agent	Colon cancer	CSGO-DsiRNA nanocomposites with pectin coating particularly destroy the cancer cells. It is a promising therapeutic agent for cancer treatment
2	Popov et al. [18]	2016	CeO_2 nanoparticles (200 nM)	As therapeutic agent	Mice tumor	Cerium oxide nanoparticles reduce the toxicity
3	Brann et al. [6]	2016	Oligonucleotide-coated gold nanoplate with doxorubicin loading	As chemotherapeutic agents	Cancerous tissue detection	Nontoxic to cells.
4	Meidanchi et al. [16]	2015	Synthesized a superparamagnetic zinc ferrite spinel ($ZnFe_2O_4$) nanoparticles	As radiosensitizing agent	Cancer therapy	This will increase the effectiveness of the radiotherapy treatment
5	Li et al. [12]	2013	Uniform, monodisperse, and core–shell structured nanospheres composed of mesoporous silica-coated individual upconversion luminescence/magnetic nanoparticle	As contrast agent and drug delivery agent	MRI dual modal imaging and for anticancer drug delivery	Magnetic nanoparticle having high magnetization value
6	Werner et al. [17]	2013	Genexol-PM	As radiosensitizing agent	Cancer therapy	Genexol-PM provides better results when compared to others
7	Kim et al. [10]	2011	Super-paramagnetic metal oxide nanoparticles	As contrast agent	MRI image enhancement	Super-paramagnetic metal oxide nanoparticles have high magnetization values for improving the contrast of T2 images
8	Parkes et al. [9]	2008	Cobalt (Co), gold(Au@Fe), and platinum (Pt@Fe) nanoparticle	As contrast agent	MRI image enhancement	Gold (Au@Fe), cobalt (Co), and platinum (Pt@Fe) have much greater saturation magnetization value than the iron

References

1. Leung JP, Wu S, Chou KC, Signorell R (2013) Investigation of sub-100 nm gold nanoparticles for laser-induced thermotherapy of cancer. J Nanomater 3:86–106
2. Maier-Hauff K, Ulrich F, Nestler D, Niehoff H, Wust P, Thiesen B, Orawa H, Budach V, Jordan A (2011) Efficacy and safety of intratumoral thermotherapy using magnetic iron-oxide nanoparticles combined with external beam radiotherapy on patients with recurrent glioblastoma multiforme. J Neurooncol 103(2):317–324
3. Kumari P, Majewski P (2013) Adsorption of albumin on silica surfaces modified by silver and copper nanoparticles, J Nanomater 2013(839016):7 pages
4. Zamperini CA, André RS, Longo VM, Mima EG, Vergani CE, Machado AL, Varela JA, Longo1 E (2013) Antifungal applications of ag-decorated hydroxyapatite nanoparticles. J Nanomater 2013(174398):9 pages
5. Hu Q, Wujcik EK, Kelarakis A, Cyriac J, Gong X (2017) Carbon-based nanomaterials as novel nanosensors. J Nanomater 2017(3643517):2 pages
6. Brann T, Patel D, Chauhan R, James KT, Bates PJ, Malik MT, Keynton RS, Toole MG (2016) Gold nanoplates as cancer-targeted photothermal actuators for drug delivery and triggered release. J Nanomater (2036029):11 pages
7. Katas H, Mohd Amin MCI, Moideen N, Ng LY, Adhwa Megat Baharudin PA (2017) Cell growth inhibition effect of DsiRNA vectorised by pectin-coated chitosan-graphene oxide nanocomposites as potential therapy for colon cancer. J Nanomater (4298218):12 pages
8. Lawaczeck R, Menzel M, Pietsch H (2004) Superparamagnetic iron oxide particles: contrast media for magnetic resonance imaging. Appl Organomet Chem 18(10):506–513
9. Parkes LM, Hodgson R, Lu LT et al (2008) Cobalt nanoparticles as a novel magnetic resonance contrast agent—relaxes at 1.5 and 3 Tesla. Contrast Media Mol Imaging 3(4):150–156
10. Kim T, Cho E-J, Chae Y, Kim M, Oh A, Ji J (2011) Urchin-shaped manganese oxide nanoparticles as pH-responsive activatable T_1 contrast agents for magnetic resonance imaging. Angew Chem Int Ed 50(45):10589–10593
11. Zhou Z, Lu Z-R (2013) Gadolinium-based contrast agents for MR cancer imaging. Wiley Interdiscip Rev Nano Med Nanobiotechnolgy 5(1):1–18
12. Li C, Li X, Hou Z et al (2013) Multifunctional NaYF4: Yb/Er/Gd nanocrystal decorated SiO2 nanotubes for anti-cancer drug delivery and dual modal imaging. RSC Adv 3:8517–8526
13. Rosa L, Blackledge J, Boretti A (2017) Nano-magnetic resonance imaging (nano-MRI) gives personalized medicine a new perspective. Biomedicines 5(1) (Franc B (ed.))
14. Das S, Mitra S, Paul Khurana SM, Debnath N (2013) Nanomaterials for biomedical applications. Front Life Sci 7(3–4):90–98
15. Chandra P, Singh J, Singh A, Srivastava A, Goyal RN, Shim YB (2013) Gold nanoparticles and nanocomposites in clinical diagnostics using electrochemical methods. J Nanopart
16. Meidanchi A, Akhavan O, Khoei S, Shokri AA, Hajikarimi Z, Khansar N (2015) $ZnFe_2O_4$ nanoparticles as radiosensitizers in radiotherapy of human prostate cancer cells. Mater Sci Eng 46:394–399 (Elsevier)
17. Werner ME, Cummings BS ND, et al (2013) Preclinical evaluation of genexol-PM, a nanoparticle formulation of paclitaxel, as a novel radiosensitizer for the treatment of non-small cell lung cancer. Int J f Radiat Oncol Biol, Phys 86(3):463–468 (Elsevier)
18. Popov AL, Zaichkina SI, Popova NR, Rozanova OM, Romanchenko SP, Ivanova OS, Smirnov AA, Mironova EV, Selezneva II, Ivanov VK (2016) Radioprotective effects of ultra-small citrate-stabilized cerium oxide nanoparticles in vitro and in vivo. RSC Adv 6(108):106141–106149
19. Feliciano CP, Tsuboi K, Suzuki K, Kimura H, Nagasaki Y (2017) Long-term bioavailability of redox nanoparticles effectively reduces organ dysfunctions and death in whole-body irradiated mice. Biomaterials 129:68–82 (Elsevier)

Comparative Analysis of Different Clustering Techniques for Video Segmentation

Tunirani Nayak and Nilamani Bhoi

Abstract Video segmentation is an extremely challenging and active area in the field of video processing and computer vision. Video segmentation techniques can be classified basically into two approaches: one approach for which there are preassigned thresholds and another clustering approach for which the number of clusters has been used, which is known. Here, we have studied and analyzed the cluster-based techniques such as mean-shift, K-means, and fuzzy C-means segmentation algorithms. We have evaluated and compared the performances of segmentation methods qualitatively and also quantitatively. To calculate the different quantitative metrics, the images and ground truth of the CDnet 2014 database have been used.

Keywords Video segmentation · Mean-shift · K-means · Fuzzy C-means

1 Introduction

The video segmentation refers to decomposing a video data into meaningful elementary parts that have strong application and correlation with the real world contained in the video data. The collective numbers of segments resulted from video segmentation cover the entire real-time video data. The video signal keeps temporal information, which introduces the object motion and includes camera motion concept. Therefore, video has both temporal nature and spatial (static) nature [1]. Segmentation of video can be set with respect to temporal, spatial, or spatiotemporal [2–5]. In a video, the static image can be obtained by segmenting a frame in spatial domain. In temporal segmentation, the sequence of video frames is segmented.

T. Nayak (✉) · N. Bhoi
Department of Electronics and Telecommunication Engineering, VSSUT, Burla, Sambalpur, Odisha, India
e-mail: tuniraninayak@gmail.com

N. Bhoi
e-mail: nilamanib@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_3

The segmentation criteria may be motion-based or color/texture-based which can be used both individually and combine form. In motion-based segmentation methods [6], there is a group of pixels with similar motion, and they are separated by multiple layers. In most of the methods [7–9], the optical flow estimation is preprocessed, and then the pixels are divided based on the pattern of motion. Some methods [6, 10, 11] are transformed and together have an effect on motion and performance evaluation.

In the last few years, many ways of segmentation have been analyzed for different images as normalized cuts [12], mean-shift [13], segmentation based on lossy compression [14], weighted aggregation (SWA) [15], and the fuzzy C-means technique [16]. Many works have been done in application of skin tumor border in medical images using segmentation which is unsupervised [17]. In these methods [16, 17], adaptive thresholding has been implemented by using scale-space filter for the extraction of the required number of clusters using histograms. Another method is proposed for fuzzy C-means algorithm which is based on histogram [18, 19] and also purposed for simplification of mean-shift filter and K-means clustering both by using background modeling. One of the commonly used models is mixture of Gaussian (MOG), kernel density estimation (KDE), etc., for background estimation.

Here, we have studied and analyzed the mean-shift algorithm [13], for the segmentation of each frame independently with the given predefined parameters. Fuzzy K-means (also called fuzzy C-means) and K-means are two simple clustering methods. K-means finds complex clusters (fusion of single clusters), and fuzzy C-means is the more formalized statistical method and finds a soft cluster with more than one block with a certain probability.

The paper is prepared as follows: Sect. 2 deals with methodology of video segmentation methods, i.e., mean-shift, K-means, and fuzzy C-means; Sect. 3 gives the performance evaluation metrics. The experimental and discussion results are presented in Sect. 4, and the conclusions of this study are given in Sect. 5.

2 Study of Cluster-Based Video Segmentation Methods

We have analyzed and studied three video segmentation algorithms such as mean-shift, K-means, and fuzzy C-means that are discussed in the following sections.

2.1 Video Segmentation Using Mean-Shift

Mean-shift is a simplest procedure where each data point moves to the center of data in neighboring countries, introduced by Fukunaga and Hostetler [13]. The kernel density function was introduced by Cheng's research into the mean-shift procedure.

In video segmentation, mean-shift algorithm is a direct resumption to protect the smoothness of interruptions. Each pixel is in its neighborhood. Then $\{x_1, \ldots, x_n\}$ is a

set of data points that are based on geometric data that measure the dimension whose function dimension is d, and then the mean-shift vector of x is defined as follows:

$$M(x) = \frac{\sum_{i=1}^{n} x_i K_w(x_i - x)}{\sum_{i=1}^{n} K_w(x_i - x)} \tag{1}$$

where x is the center of the kernel and $K_w(x) = |W|^{-1/2} K\left(W^{-1/2}x\right)$. W is the approximate density of the dimension $d \times d$, and matrix values determine the solution, the detection, the mode, and the smallest number of consecutive pixels. W is usually a diagonal matrix and W is a symmetric positive bandwidth matrix in which the values determine the smallest number of contiguous pixels. $K(x)$ is a d-variate nucleus and is regarded as nonnegative free radical symmetric functions centered at zero.

The normal kernel defined below was used in our segmentation algorithm:

$$K(x) = (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2}\|x\|^2\right) \tag{2}$$

Because of the mean-shift, the $M(x)$ axis always shows the direction of the maximal density, and the average intermediate procedure is obtained by calculating the core change and changing the vectors of the vector $K(x)$. At the end of the approximate procedure, it is guaranteed at the nearest point where the estimate is zero gradients. Define as

$$y_{j+1} = \frac{\sum_{i=1}^{n} x_i K_w\left(y_j - x_i\right)}{\sum_{i=1}^{n} K_w\left(y_j - x_i\right)} \quad j = 1 \tag{3}$$

Equation (2) can be modified as

$$y_{j+1} = \frac{\sum_{i=1}^{n} x_i \exp\left(-\frac{1}{2}\sum_{l=1}^{d}\left(\left(y_j^l - x_i^l\right)/w_l\right)^2\right)}{\sum_{i=1}^{n} \exp\left(-\frac{1}{2}\sum_{l=1}^{d}\left(\left(y_j^l - x_i^l\right)/w_l\right)^2\right)} \tag{4}$$

where y_j^l and x_i^l are the lth component of y_j and x_i, respectively. Thus, mean-shift algorithm is an update procedure for y_j.

The mean-shift segmentation algorithm is as follows:

Step 1: Two-dimensional image size is a map of the two-dimensional space and looks at each pixel by $x_i \in R^d$.
Step 2: Initialize the value $j = 1$ and $y_{i,j} = x_i$.
Step 3: Calculate the value of $y_{i,j+1}$ according to Eq. (2) when it is approaching a certain point and then set $z_i = y_{i,c}$, where $y_{i,c}$ is the convergence value of $y_{i,j}$.
Step 4: Imagine domain sharing $\{C_p\}_{p=1,\ldots,m}$ by grouping together all z_i than w_1 to the lth function ($l = 1, 2, \ldots,$ d).

Step 5: For each $i = 1, 2, \ldots, n$, assign $L_i = \{p | z_i \in C_p\}$.
Step 6: Neglect the spatial regions having less than N pixels.

2.2 *Video Segmentation Using* K-*Means*

The simplest algorithm for learning without follow-up are the K-means, which can solve cluster problems. This is one of the easiest ways to divide datasets provided by a number of clusters (assume k clusters) that are supposed to fix priority. The main purpose is to set up a wider group for all (k centroids). These centroids are placed in the way different locations lead to different results. Therefore, everything is put as far as possible to get a good result. So each point in the dataset is selected at one time and connects it to the nearest center. When each point is processed, the first step is complete. So, at this point, the centroid is recalculated by clusters that came from the previous result. When we have these new centroids, we have to create a new connection between the same data and the nearest new center. Since it is a method, a loop is created and uses this cycle; we can notice that centroids k change their position step by step to further changes. Finally, we minimized an objective function. The objective function is given as

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \tag{5}$$

$\sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$ is the objective function within group I. x_i is a vector and c_j the corresponding cluster center.

The K-means segmentation algorithm is as follows:

Step 1: Consider the first vector of the K function in clusters.
Step 2: Sample vector is calculated using the minimum distance principle.
Step 3: New average is a new cluster for each cluster.
Step 4: Repeat step if the center is moved to Step 2, or stop.

2.3 *Video Segmentation Using Fuzzy* C-*Means*

Clear compilation plays an important role in problem-solving in recognizing patterns and sample identities and organizing video images. Different separation methods are provided by clusters, and most of them are based on distance criteria. The widely used algorithm is the C-means (FCM) delaying algorithm. This is useful when the cluster count is predetermined. In this way, the algorithms try to place each data point on a cluster. What makes FCM different is that it cannot resolve the absolute membership in the data points provided in the cluster. Instead, it calculates the probability

(membership) that the data point given will belong to a cluster. So depending on the accuracy of grouping required to exercise proper tolerance, it may be advisable. Since absolute membership is not calculated, FCM can be fast because of the number of changes required to achieve special components.

In the following FCM algorithm iteration, the following goals (objective function) are minimized:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \left\| x_i - c_j \right\|^2 \tag{6}$$

Here N is the number of data points, C is the necessary number of clusters, c_j is the central vector j and δ_{ij} is the membership level for the data point i in the j group. The $\|x_i - c_i\|$ standard measures the (or near) approximation of the x_i data point with the central vector c_j of j clusters. Note that in each iteration, the rules maintain a vector center for each cluster. These data are calculated on the average weight of the weighted data displayed by membership.

The fuzzy C-means segmentation algorithm is as follows:

Step 1: Start the membership-matrix U with values between 0 and 1.
Step 2: Calculate the center cluster c.
Step 3: Calculate the function that costs according to the equation. Stop if either it is below a certain tolerance value or its improvement, than the previous repetition is below certain threshold levels.
Step 4: Calculate the new U. Please go to Step 2

3 Performance Evaluation Metrics

The quality rating of the video segmentation algorithm is by comparing the result of the segment with part of the ground truth and obtaining the following statistics:

True Positive (TP): Intersection of pixels in the foreground of the segmented output and ground truth.

False Positive (FP): Intersection of pixels in the foreground of the segmented output and pixels in the background of ground truth.

True Negative (TN): Intersection of pixels in the background of the segmented output and ground truth.

False Negative (FN): Intersection of pixels in the background of the segmented output and pixels in the foreground of ground truth.

Using these statistics, assessments including FPR, FNR, Precision (Pr), Recall (Re), Accuracy (Ac), and Specificity (Sp) are calculated.

$$\text{FPR} = \frac{\text{number of FP}}{(\text{number of FP} + \text{number of TN})}$$

$$\text{FNR} = \frac{\text{number of FN}}{(\text{number of FN} + \text{number of TP})}$$

$$\text{Precision(Pr)} = \frac{\text{number of TP}}{(\text{number of TP} + \text{number of FP})}$$

$$\text{Recall(Re)} = \frac{\text{number of TP}}{(\text{number of TP} + \text{number of FN})}$$

$$\text{Specificity(Sp)} = \frac{\text{number of TN}}{(\text{number of FP} + \text{number of TN})}$$

$$\text{Accuracy(Ac)} = \frac{\text{number of TP} + \text{number of TN}}{(\text{number of TP} + \text{TN} + \text{FP} + \text{FN})}$$

These above parameters describe the performance [20–22] of the moving object.

4 Experimental Results and Discussion

Analysis of video segment algorithms has been tested and implemented using Matlab. We have selected some of the testing segments, the standard video clips called CDnet 2014 [23] in the section of the video. These sets of videos differ in size with corresponding ground truth. The mean-shift, K-means, and fuzzy C-means segmentation algorithms are applied over the above dataset and got the segmentation result.

In mean-shift algorithm, we have taken the spatial_bandwidth = 40, color_bandwidth = 3 and convergence factor = 3. Different values of bandwidth W can be taken for the performance measurement.

In K-means algorithm, we have taken the cluster number = 8 and we choose the cluster center randomly.

In fuzzy C-means algorithm, we choose the cluster center randomly (Fig. 1).

To measure the quality of the segmentation algorithms, we calculate different metric measurements [20–22]. The metrics were reported for CDnet for 2014 [23]. We have provided a comparative analysis of mean-shift, K-means, and fuzzy C-means clustering-based segmentation. We can see that for baseline category, the accuracy is more for K-means method. The baseline and shadow categories consist of simple sequences where the normal order of pedestrian and car is the focus. With different types of fraud and variation related to Interm.Object motion, fuzzy C-means method gives very good recall compared to other. In shadow category, the accuracy is more than 90% for mean-shift segmentation. But the recall is high for fuzzy C-means method. In dynamic background category, the accuracy is more than 90% both in mean-shift and K-means methods but recall is more in fuzzy C-means method (Tables 1, 2, 3 and 4).

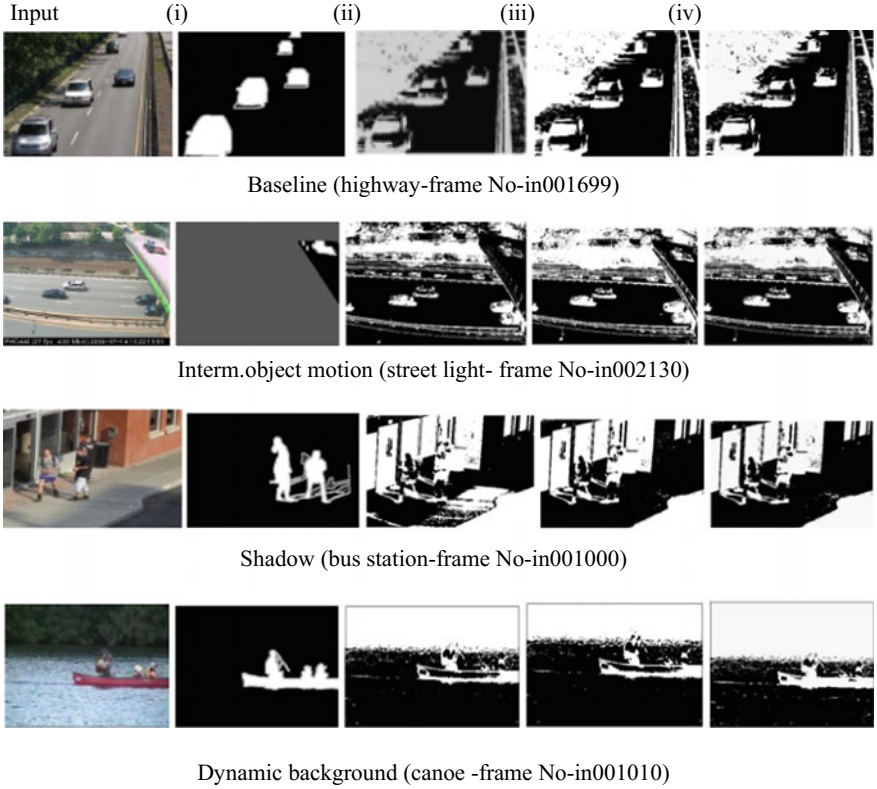


Fig. 1 Segmentation results obtained for various sequences of the CDnet dataset: (i) ground truth maps; (ii)–(iv) output results of mean-shift, *K*-means, and fuzzy *C*-means methods, respectively

Table 1 Performance measures (FPR, FNR, Pr, Re, Ac, Sp) of different methods for baseline image

Methods	FPR	FNR	Pr	Re	Ac	Sp
Mean-shift	0.0562	0.9389	0.0161	0.0611	0.9307	0.9438
K-means	0.0168	0.9864	0.0123	0.0136	0.9685	0.9832
Fuzzy *C*-means	0.8484	0.1469	0.0145	0.8531	0.1618	0.1516

Table 2 Performance measures (FPR, FNR, Pr, Re, Ac, Sp) of different methods for Interm. Object motion image

Methods	FPR	FNR	Pr	Re	Ac	Sp
Mean-shift	0.1642	0.8929	0.1200	0.1071	0.7099	0.8356
K-means	0.1481	0.7692	0.2000	0.2308	0.7660	0.8519
Fuzzy *C*-means	0.3523	0.3333	0.2051	0.6667	0.6500	0.6477

Table 3 Performance measures (FPR, FNR, Pr, Re, Ac, Sp) of different methods for shadow image

Methods	FPR	FNR	Pr	Re	Ac	Sp
Mean-shift	0.0508	0.9067	0.0505	0.0933	0.9251	0.9492
K-means	0.0535	0.9385	0.0307	0.0615	0.9227	0.9465
Fuzzy *C*-means	0.8768	0.1095	0.0264	0.8905	0.1431	0.1232

Table 4 Performance measures (FPR, FNR, Pr, Re, Ac, Sp) of different methods for dynamic background image

Methods	FPR	FNR	Pr	Re	Ac	Sp
Mean-shift	0.0531	0.9620	0.0131	0.0380	0.9304	0.9469
K-means	0.0566	0.9447	0.0182	0.0553	0.9269	0.9434
Fuzzy *C*-means	0.8599	0.1554	0.0157	0.8446	0.1514	0.1401

5 Conclusion

In the process of visualizing a video and a computer, one of the most difficult and most active areas of the research is the part of the video material. One of the most important issues for successful use of successful video sequences is the video segment that focuses on dividing video files into essential video frames. Various techniques are being used by the developers in reviewed. We analyzed three video segmentation algorithms based on the discovery of the changes in the video clip. We have used mean-shift, *K*-means, and fuzzy *C*-means clustering-based segmentation algorithm for video segmentation and measured the performance evaluation metrics. Mean-shift and *K*-means methods provide good accuracy in comparison to fuzzy *C*-means. But the recall is generally more in case of fuzzy *C*-means method.

References

1. Jiang H, Zhang G, Wang H, Bao H (2015) Spatio-temporal video segmentation of static scenes and its applications. IEEE Trans Multimed 17(1)
2. Wu GK, Reed TR (1999) Image sequence processing using spatiotemporal segmentation. IEEE Trans Circ Syst Video Technol 9(5):798–807
3. Kim EY, Hwang SW, Park SH, Kim HJ (2001) Spatiotemporal segmentation using genetic algorithms. Pattern Recognit 34(10):2063–2066
4. Koprinska I, Carrato S (2001) Temporal video segmentation: a survey. Signal Process Image Commun 16(5):477–500
5. Megret R, Dementhon D (2002) A survey of spatio-temporal grouping techniques. In: Language and media process, University of Maryland, College Park, MD, USA, Tech. Rep. LAMP-TR-094/CS-TR-4403
6. Kumar MP, Torr PHS, Zisserman A (2008) Learning layered motion segmentations of video. Int J Comput Vis 76(3):301–319
7. Shi J, Malik J (1998) Motion segmentation and tracking using normalized cuts. In: Proceedings of the ICCV, pp 1154–1160

8. Fowlkes C, Belongie S, Malik J (2001) Efficient spatio temporal grouping using the nystrom method. In Proceedings of the CVPR, pp 231–238
9. Khan S, Shah M (2004) Object based segmentation of video using color, motion and spatial information. In: Proceedings of the CVPR, vol 2, pp 746–750
10. Remers D, Oatto S (2003) Variational space-time motion segmentation. In: Proceedings of the ICCV, pp 886–893
11. Itnick CL, Jojic N, Kang SB (2005) Consistent segmentation for optical flow estimation. In Proceedings of the ICCV, vol 2, pp 1308–1315
12. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905
13. Comaniciu D, Meer P, Member S (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24(5):603–619
14. Mobahi H, Rao S, Yang AY, Sastry SS, Ma Y (2011) Segmentation of natural images by texture and boundary compression. Int J Comput Vis 95(1):86–98
15. Sharon E, Galun M, Sharon D, Basri R, Brandt A (2006) Hierarchy and adaptively in segmenting visual scenes. Nature 442(7104):719–846
16. Lim YW, Lee SU (1990) On the color image segmentation algorithm based on the thresholding and the fuzzy *c*-means techniques. Pattern Recognit 23(9):935–952
17. Hance GA, Umbaugh SE, Moss RH, Stoecker WV (1996) Unsupervised color image segmentation with application to skin borders. IEEE Eng Med Biol, 104–111
18. Devikar MM, Jhac MK (2013) Segmentation of images using histogram based FCM clustering algorithm and spatial probability, Department of Telecommunication Engineering, CMRIT, Bangalore, India. Int J Adv Eng Technol
19. Ali SM, Abood LK, Abdoon RS (2013) Clustering and enhancement methods for extracting 3D brain tumor of MRI images. Remote Sensing Research Unit, Department of Computer Science, University of Baghdad, Department of Physics, University of Babylon, Volume 3, Issue 9
20. Senior A, Hampapur A, Tian Y, Brown L, Pankanti S, Bolle R (2000) Appearance models for occlusion handling. In: Proceedings of the 2nd IEEE workshop performance evaluation of tracking and surveillance
21. Erdemand CE, Sankur B (2000) Performance evaluation metrics for object based video segmentation. In: Proceedings of 10th European Signal Processing Conference, vol 2, pp 917–920
22. Marichal X, Villegas P (2000) Objective evaluation of segmentation masks in video sequences. In: Proceedings of 10th European Signal Processing Conference, vol 4
23. Wang Y, Jodoin P-M, Porikli F, Konrad J, Benezeth Y, Ishwar P (2014) CDnet 2014: an expanded change detection benchmark dataset. In: Proceedings of the IEEE conference on workshops of computer vision and pattern recognition (CVPR), pp 387–394

Fingerprint Identification with Combined Texture Features

Namrata V. Jad and Satish T. Hamde

Abstract The recognition using biometric fingerprint is low-cost and superior automated technique of verifying best match among two human fingerprints. In this paper, texture features are used for fingerprint matching. The matching methodology uses local scale and rotation invariant information. This paper proposes a novel method for fingerprint verification and identification by using combination of features of LBP, HOG and SIFT. The fingerprint image is first enhanced, segmented and rotated/scaled for different angles/values. The SVM and multi-SVM are used for verification and identification, respectively. The methodology is applied on database created by fingerprint sensor H520J00122 and also tested on FVC2004, FVC2002 database. The analysis points such as accuracy, time, TAR, FAR and FRR are measured. The proposed method also ensures the highest accuracy with moderate time of execution.

Keywords Fingerprint identification LBP · HOG · SIFT · SVM · Multi-SVM

1 Introduction

Authentication of person plays a vital role if we want to restrict fraud person to enter the facility, or to check whether the person is entitled to access the privileged information and also to examine whether the person had previously applied for the job. The biometric system is being tremendously deployed in the government sector, national ID cards, issuing of driver's license, mobile, bank account and security criminal identification [1, 2]. Biometrics are human characteristics. It can be characteristics of behaviour and physiology. In this paper, a physiological aspect (Fingerprint) is used to authenticate a person. It has its own challenges too. The drawbacks observed as

N. V. Jad (✉) · S. T. Hamde
Department of Instrumentation Engineering, SGGSIE&T, Nanded, Maharashtra, India
e-mail: 2016min010@sggs.ac.in

S. T. Hamde
e-mail: sthamde@sggs.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_4

$$\text{Total Success Rate} = \frac{\text{Number of true fingerprint matched}}{\text{Total no. offingerprints taken in database}}$$

$$\text{False Acceptance Rate} = \frac{\text{No. of false fingerprint accepted as true}}{\text{Total no. of fingerprints not taken in database}}$$

$$\text{False Rejection Rate} = \frac{\text{Rejected no. of true fingerprints as false}}{\text{Total no. of fingerprint taken in database}}$$

These drawbacks are recovered by implementing an efficient fingerprint authentication methodology which is discussed in this paper. This paper is separated into four sections: Sect. 2 consists of literature survey, Sect. 3 includes proposed fingerprint texture-based method and Sect. 4 concludes the analysis of comparison of single and group feature descriptor. The conclusion is mentioned in Sect. 5.

2 Literature Survey

The conventional imaging strategy in the field of person authentication by fingerprints has been gifted with a variety of advances in some basic and advanced work over the years. In this section, we revisit the approaches proposed by researchers during the last 25 years. The overall focus is concentrated on features-based matching. The fingerprint feature extraction methods can be introduced into the following four ways: (1) filter bank-based, (2) minutiae-based, (3) minutiae and texture, and (4) texture-based.

Filter Bank-Based Methods
S. Prabhakar and Jain, Lin with Pankanti in [3] worked on a set of Gabor filters, which uses [4] algorithm of a filter bank-based, that extract features, which are local and also global one. However, for matching, Euclidean distance is measured between the features of two corresponding fingerprints. This method is too fast, but it is observed that the accuracy obtained by this method is not enough. In [3], Jain et al. show that for complex oriented textures like fingerprint, local and global information are needed to be combined. The approach is based on core point extraction.

Minutiae-Based
The authors of [5] introduce minutia feature descriptor named Minutiae Cylinder Code (MCC). The MCC matches two fingerprints by comparing with to minutiae structure that is local. Matching is unaffected even if fingerprint undergoes translation, rotation of minutiae descriptor. This needs alignment of fingerprint at time of registration. F. Chen, X. Huang, J. Zhou [6] used the hierarchical strategy. This technique's key point is it does separation of matching step into many required stages and rejects a number of wrong fingerprints. Their proposed strategy speeds up identification and can almost save 50% of searching time computed with conventional methods.

Minutiae- and Texture-Based
A fingerprint minutiae detail is clubbed with texture details, which gives rise to a hybrid matching method [7]. Jain et al. in [7] deal with the hybrid matching system. Hybrid matching system clubs the details of minutiae as well as texture. It gives rise to more improved matching performance. Here point decides region of interest; therefore, the system senses only a limited portion of fingerprint for getting texture details. Thus, the number of impression of same person may have a small overlapping region.

Texture-Based Methods
R. Haralick, K. Shanmugam, Hak D. [8, 9] use a texture image and describe its second-order statistics. This statistical approach uses gray-level co-occurrence matrix (GLCM) and two-dimensional histogram. Here, textural features computation is quicker. Identification accuracy obtained is 89%. Yazdi and Gheysari [10] obtain texture features using co-occurrence matrices. They have used an ANN artificial neural network (feedforward) with two hidden layers for classification. Its obtained results show superior performance. Hence, the further part of paper presents implemented strategies to improve recognition rate and reduce the maximum possibility of complex way of computation and processing, along with fulfilling the aim of methodology.

3 Proposed Texture-Based Approach

Person recognition undergoes three steps mainly: enrolment, verification and/or identification. Figure 1 shows the proposed texture-based identification methodology. In enrolment step, a set of biometric data is registered, and its features are retrieved after each image is segmented. The segmentation is included to get prominent features. The biometric data includes authorised and some unauthorised fingerprints. The feature set is obtained by Local form of Binary Pattern (LBP), separate Histogram of Oriented Gradient (HOG) and Scale Invariant Feature Transform (SIFT) method from train and test images with five different angular rotations and scales. The feature set so obtained is used to train classifier-I (SVM) and classifier-II (multi-SVM).

Classifier-II is trained by a set of only authorised person's fingerprints. In the second verification step, input fingerprint is classified by classifier-I, i.e. Support Vector Machine (SVM). Its result shows whether input fingerprint is authorised or unauthorised. If it is authorised, then only identification is done. The classifier-II used is multi-SVM with more than one hyper-plane. The no. of hyper-planes of multi-SVM is decided by number of enrolled classes of data. Its result will be the best matched database fingerprint.

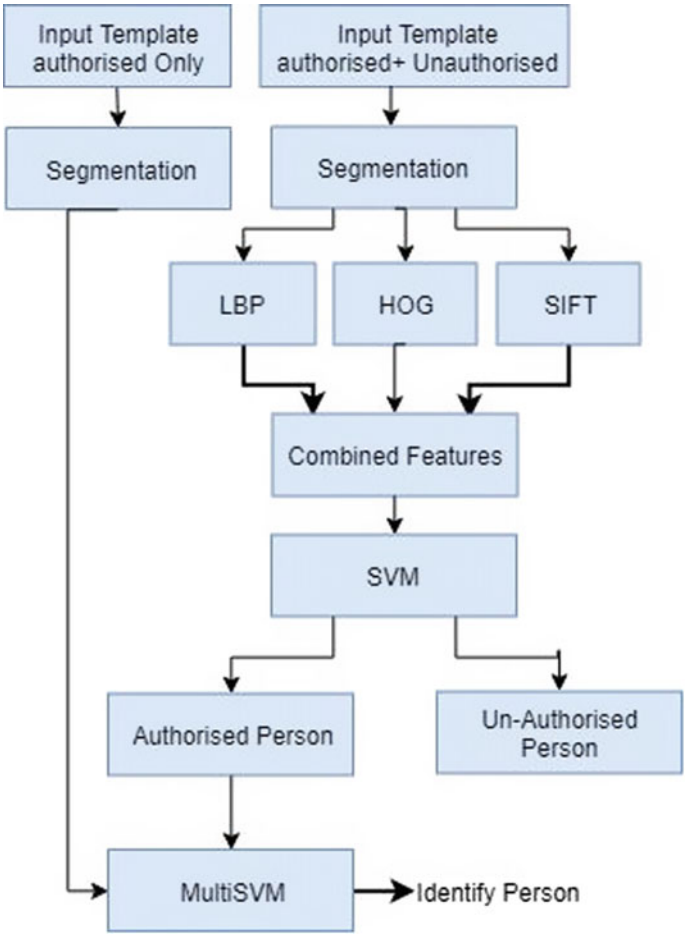


Fig. 1 Block diagram of proposed texture-based method

3.1 Enhancement and Segmentation

A contrast enhancement by histogram equalisation improves the perceptibility of objects [11]. Then fingerprint region in an image is segmented by mean and variance method [12]. The foreground fingerprint and background region are separated so that the texture feature of only fingerprint area is retrieved, thus providing features of only interested fingerprint region.

3.2 Feature Extraction

The separate and combined feature descriptor is applied to extract features. In this work, descriptors used are HOG, SIFT and LBP.

Local Binary Pattern (LBP) Descriptor
It is the most popular in texture classification [13, 14]. The steps to create LBP feature vector are given in the following text.

The images are received from database are first divided into blocks (3 × 3) so here we considered its eight neighbourhood pixels. Now according to pixel intensities of each pixel (let g_c) and its each neighbourhood(g_p), the binary value of neighbourhood is decided. Here replaced neighbour pixel by zero, if centre pixel's value is more than the neighbour's value else replace by one. It is formulated in Eq. (1).

$$s(x) = \begin{cases} 1 \; x \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

where $x = (g_p - g_c)$. Now each pixel is surrounded by binary values as shown in Fig. 2.

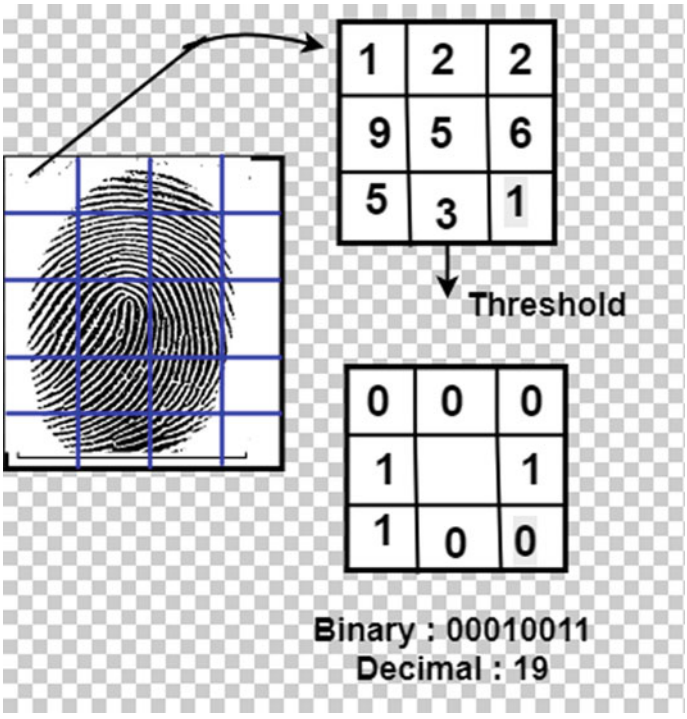


Fig. 2 LBP representation

The binary sequence is ordered in a clockwise direction to get an eight-digit binary number and is then converted to decimal equivalent using the following equation:

$$\mathrm{LBP}_{\mathrm{PR}} = \sum_{P=0}^{P} S\left(g_{\mathrm{p}} - g_{\mathrm{c}}\right)2^{P} \tag{2}$$

where g_{p} stands for the central pixel's gray value, g_{c} represents the value of its neighbours, P stands for the all involved neighbours and R is the radius of the neighbour. Thus, equation depicts that given image is converted to decimal equivalent image.

Then, the histogram is calculated and it can be a 256-dimensional feature vector. Suppose image size is $I * J$, and once LBP patterns of every pixel are computed. The texture is represented by its histogram as represented in Eq. (3) as

$$H(K) = \sum_{i=1}^{I}\sum_{j=1}^{J} f(\mathrm{LBP}_{\mathrm{PR}}(i, j), k \in [0, k]) \tag{3}$$

where

$$f(x, y) = \begin{cases} 1 & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

Normalise the histogram, which is LBP feature vector for that image. The LBP feature vector can pass to the classifier (SVM).

SIFT Descriptor

Some details may change if the scale of an image is changed. Thus, SIFT plays an important role in those case [15, 16]. SIFT algorithm depicts that it first calculates Gaussian filter for a number of scales and subtracts Gaussian (DOG) of one scales from other. It does a comparison of each and every pixel obtained in DoG image with its nearby eight pixels' intensity then each pixel of DoG images is Compared to its nine pixels at neighbourhood. It is followed for each nearby scales. If it is observed the pixels' intensity value is the more or less than neighbouring pixels then it is treated as a candidate key point. Then store key point for all training data. Compare the location of key point of the given image with the key points of training data.

HOG Descriptor

The use of HOG descriptor can be called the characterization of the orientation and magnitude of the pixels. A group of local histograms of an image is computed. These histograms have information of the different orientations of the gradients in a local region of an image [17]. HOG descriptors steps include the computation of the image gradient, orientation histograms for corresponding locations and normalisation of the histograms.

3.3 Classifier

The classification task is distributed to two classifiers. In order to classify the person as authorised or not, a primary classifier support vector machine is used. The SVM provides the final result in case of unauthorised person. The Identification is done at secondary classifier, in case of authorised person.

Primary Classifier (SVM)
For primary classification task, supervised learning algorithm that is support vector machine is used [3]. It does binary classification in which a known training dataset contains authorised and some unauthorised dataset to make predictions. Its crucial step is finding the best hyper-plane [1, 2]. The feature set provided to SVM is mapped into a dimensionality which is higher. It uses $f(x)$ (mapping function) for best provided feature sets. It does a linear separation. It is desired to get the largest distance among two different fingerprint data within feature space. It is needed to solve an optimization problem. The optimization can be solved by using information, input vector and type of kernel function (RBF, linear, polynomial and sigmoid). The selection kernel function plays a vital role in SVM. The accuracy of the model is improved by proper selection of kernel function available in SVM. Here, we use the RBF kernel due to finite and localised response across the total range. A residue x_j which is unlabelled can be classified by the function shown in Eq. (4)

$$f(x_j) = \sum_{i=1}^{l} y_i \alpha_i K\left(x_i x_j\right) + b \tag{4}$$

This labels x_j. If the value of f (x_j) is above zero, the value residue x_j is classified as authorised; otherwise, x_j is unauthorised class. The best prediction is obtained by adjusting the threshold.

Secondary Classifier (Multi-SVM)
Multi-SVM classification model is similar comparison of one with all. Here multi-SVM is trained if query fingerprint is authorised. It is considered for differential identification, in which each person's fingerprint has its own sub-classifier, that is, binary SVM. For database of 80 authorised person, 80 differential classes are generated so that differential identification is done. Multi-SVM training would make use of 80 networks, as each 80 separate SVM would include in the recognition of particular person.

4 Result and Comparative Analysis

The database considered for analysis of result is obtained from the database by sensor H520J00122 and database FVC 2002 and FVC 2004. The result shows the comparative analysis of the effect of features on identification performance, as shown in

Table 1 Comparative analysis

Descriptor	SVM	Multi-SVM	Time
LBP	75	84	12.76
SIFT	75	82	72.89
HOG	75	85	29.69
LBP + HOG	92	96	23.68
LBP SIFT	75	89.05	69.89
LBP + SIFT + HOG	97.50	98.05	434.63

Table 2 Match count FAR, FRR and TSR

Match count	FAR	FRR	TSR
34	4.91	44.44	94.44
26	0	65.38	100
30	4	63.33	93.33
14	1.52	71.42	92.85
10	0	60	100
7	0	28.57	100

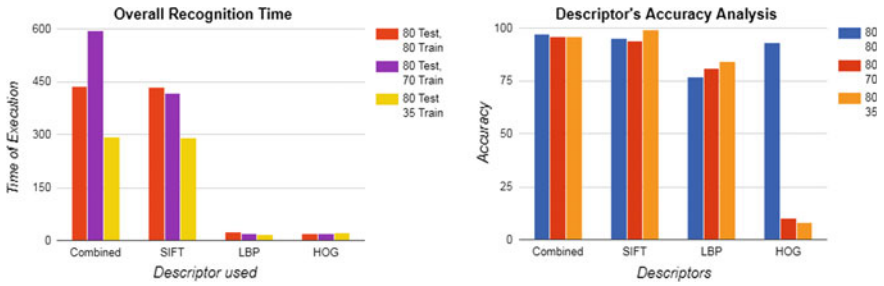


Fig. 3 Analysis chart

Table 1. This analysis depicts that accuracy is improved, and the result of matching and identification is invariant to rotation and scale change, if we combine feature descriptor HOG, LBP and SIFT. The match count is varied for different groups of test data. False/true accepted/rejected rate (FAR value, FRR value and TSR value in percentage) is measured for some test data as shown in Table 2. The false (unauthorised) fingerprint acceptance rate obtained is less. The rejection rate of unauthorised is more. TSR is near to highest. Table 3 depicts the person identification performance, when proposed descriptor used. Figure 3 represents the analysis chart showing accuracy and time of separate and proposed descriptor.

Table 3 Performance evaluation

Database	Descriptor	Accuracy	Time (s)	Train images	Test images
Created DB	HOG	8.57	21.93	80	35
		10	19.32	80	70
		93	20.73	80	80
	LBP	84.12	18.56	80	35
		81.24	20.77	80	70
		77	25.46	80	80
	SIFT	99	290.18	80	35
		94.29	416.02	80	70
		95	434.1	80	80
	HOG + LBP + SIFT	96	293.88	80	35
		96.53	596.68	80	70
		97.31	437.43	80	80

5 Conclusion

In proposed work, the preprocessing is carried out by scaling and rotating a fingerprint image. The mean variance algorithm properly extracts fingerprint portion of the image to which, five rotations, five scales variations are applied. Then, the image is applied to descriptors like LBP, HOG and SIFT combinations. It is observed that the final combined LBP, SIFT and HOG features improve the classification of authorised and unauthorised fingerprints. This reduces execution time if input query fingerprint is unauthorised; hence, no multi-SVM is in picture. Thus, SVM itself can provide the final result for unauthorised person. Performance evaluation is also done by parameters like FAR, FRR and TSR. The proposed work provides the highest accuracy.

References

1. Maltoni D, Jain AK, Maio D, Prabhakar S (2003) Handbook of fingerprint recognition. Springer-Verlag, New York
2. Author F, Zhang D et al (2000) Automated biometrics: technologies and systems. Kluwer, Boston, MA, USA
3. Jain A, Prabhakar S, Hong L, Pankanti S (2000) Filter bank-based fingerprint matching. IEEE Trans Image Process 9(5)
4. Feichtinger HG, Thomas S (2000) Gabor analysis and algorithms. Birkhauser
5. Cappelli R, Ferrara M, Maltoni D (2010) Minutiae cylindrical code: a new representation and matching technique for fingerprint recognition. IEEE Trans Pattern Anal Mach Intell 32(12)
6. Chen F, Huang X (2013) Hierarchical minutiae matching for fingerprint and palm print identification. IEEE Trans Image Process 22(12)

7. Jain A, Ross A, Prabhakar A (2001) Fingerprint matching using minutiae and texture features. In: International conference on image processing, Oct 2001
8. Haralick R, Shanmugan K, Dinstein J (1973) Teture features for image classification. IEEE Trans Syst Man, Cybern SMC-6(6)
9. Haralick R (1979) Statistical and texture approaches to texture. Proc IEEE 67(5)
10. Yazdi R, Gheysari K (2008) A new approach for the fingerprint classification based on gray-level co-occurrence matrix. Int J Comput Inf Eng 2(11)
11. Barnauti N (2016) Fingerprint recognition improvement by using histogram equalization and compression method. Int J Eng Res Gen Sci 4(2)
12. Fenq W, Lin X (2009) An improved fingerprint segmentation algorithm based on mean and variance. In: International Workshop on Intelligent Systems and Applications
13. Monika MK (2000) A novel ngerprint and minutiae Matching using LBP. In: Third international conference on reliability, infocom technology and optimization, May 2000, pp 846–859
14. Guo Z, Zhang L, Zhang D (2010) A completed modelling of local binary pattern operator for texture classification. IEEE Trans Image Process 19:1657–1663
15. Lowe D (2004) Distinctive image from scale invariant keypoints. Int J Comput Vision, 91–110
16. Lowe D (1999) Object recognition from local scale invariant features. In: Proceedings of international conference on computer vision, pp 1150–1157
17. Dalal N, Triggs B (2005) Histogram of oriented gradients for human detection. In: IEEE conference on computer science on computer vision pattern recognition, vol 1, June 2005, pp 886–893

Tuberculosis Detection Using Shape and Texture Features of Chest X-Rays

Niharika Singh and Satish Hamde

Abstract Tuberculosis (TB) is a major life-threatening hazard, globally. Mortality rate increases if the disease remains undiagnosed and untreated. Detection of disease in the early stage is the most promising way to increase the lifespan of patients, especially in the regions with limited resources worldwide. We present an automatic TB detection method which uses conventional digital chest radiographs. The method consists of three main stages. We first extract the lung region from the Chest X-Ray (CXR) image using log Gabor filtering technique followed by morphological methods. A set of texture and shape features of segmented dataset is computed. The feature vector thus computed enables the support vector machine to classify the input CXR into healthy and TB-infected. ROC curve and confusion matrix of classifier show its exceptionally good performance. We attain an AUC of 0.98 and 0.96 on MC and CHN dataset, respectively, with 100% specificity.

Keywords Features · CXR image · TB · Support vector machine (SVM) · Classifier

1 Introduction

Tuberculosis is the second largest killer in the world among all the infectious diseases. In 2016, over 1.7 million TB-infected people died, of which nearly 400,000 people were co-infected with HIV [1]. TB is a challenge globally. Although many progresses are made in diagnosis and treatment of disease, unfortunately, TB remains a major cause of mortality worldwide. TB is prevalent in sub-Saharan Africa where poverty and malnutrition diminish patients' immunity [1]. Moreover, the problem has been aggravated by increased rate of infections in HIV positive population and rapid

N. Singh (✉) · S. Hamde
Department of Instrumentation Engineering, SGGSIE&T, Nanded, Maharashtra, India
e-mail: singhniharika2694@gmail.com

S. Hamde
e-mail: sthamde@sggs.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_5

emergence of multidrug-resistant TB. Immunodeficient people are at higher risk of latent TB getting turned into active.

The emergence of disease and shortcomings of available clinical tests make a room for the development of economical screening system to monitor the recovery during treatment. Chest radiography is an important tool to diagnose pulmonary TB when it is not confirmed bacteriologically [2].

Conventional radiographic techniques have been gifted with many notable advances over the years. Many trendsetting papers based on CAD system have been presented in the previous 15 years. Jaime Melendez [3] deals with Multiple Instance Learning (MIL) based CAD system. Their system handles the problem of TB detection by classifying the pixels in supervised manner. Though this approach gives good results with high accuracy, it is an exhausting and sluggish method. Moreover, if new training data is to be used, complete reprocessing is to be accomplished.

CAD system development for X-ray analysis is really complex. To reduce the complexity of the task, researchers started to concentrate on solving the sub-problems instead of developing a full-edged CAD system. Segmentation of the lung field is one of the typical tasks to proceed in this direction. Candemir [4] introduces the segmentation of lung field using graph cuts optimization approach. Karagyris [5] introduces the lung anatomy segmentation. He has segmented out the ribs and heart. This work is helpful in locating the manifestations of TB.

van Ginneken [6] uses a multiscale filter bank to extract texture feature. They handle the problem by subdividing the lung lobe into overlapping regions of various sizes. Texture features are extracted for each of these subdivided regions. Feature vectors from corresponding regions in left and right lung lobe are subtracted to get some additional features for each region. The separate training set is constructed. Every region is classified by voting among the KNN.

Xu [7] works on the fact that TB is mainly characterised by the presence of cavities in the upper zone and focuses on the cavity detection. However, the presence of cavity is only one of the TB manifestations. Jaeger in [8] segments the CXR image by taking the average of the three different masks: the intensity mask, lung model mask and log Gabor mask. In the proceedings of the same work [9], they introduce a lung model for segmentation. Features are extracted in two sets—features inspired by object detection and contrast-based features. They achieve very low specificity.

The proposed method combines a few object detection inspired features, a few of content-based image retrieval-based image features and a set of GLCM-based features [9, 10, 11] and gives improvised results. The paper is framed in the following manner: Sect. 2 describes the overview of the proposed screening method. It includes all the three steps of processing—segmentation, feature extraction and classification. Experimental results are discussed in Sect. 3. Section 4 concludes the work.

2 Methodology

In this section, the implemented method of lung field extraction, feature extraction and classification are going to be discussed. The flow of work is shown in the flow diagram in Fig. 1.

2.1 Extraction of Lung Field

In chest radiographs, it is difficult to separate out lung fields using pure pixel-based approaches. Thresholding based on visual appearance does not work due to highly mingled intensities. However, wavelets can extract solid regions of our interest with accuracy. The enhanced image can be further binarised easily by thresholding.

Log Gabor is found adequate for the enhancement of image [12]. Log Gabor function is nothing but the Gaussian transfer function when scaled on the logarithmic frequency. In the polar coordinate system, log Gabor filter is composed of two components: radial component, which varies radially and angular component, which shows angular variations. The radial filter is given as

$$R(\omega) = \exp\left(\frac{-\left[\log(\omega/\omega_0\right]^2}{2\log\left(\sigma_r^2\right)}\right) \tag{1}$$

And angular frequency response is given as

$$A(\theta) = \exp\left(\frac{-[\theta - \theta_0]^2}{2\log\left(\sigma_\theta^2\right)}\right) \tag{2}$$

Thus, the overall log Gabor function is given as

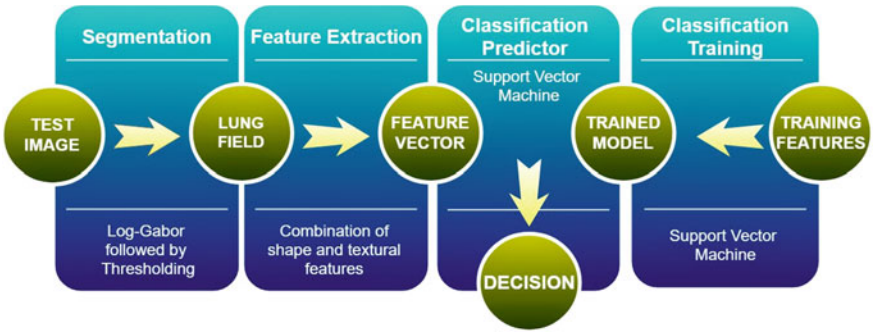


Fig. 1 Flow diagram

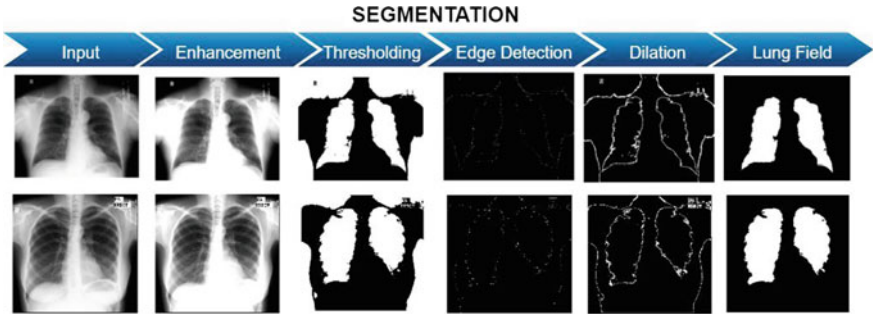


Fig. 2 Flow of segmentation: **a**, **b** segmentation of randomly chosen normal and abnormal CXR image from Montgomery dataset, respectively

$$G(\omega, \theta) = \mathrm{R}(\omega) * A(\theta) \tag{3}$$

where

(ω, θ)	the polar coordinates;
ω_0	centre frequency;
θ_0	angle of orientation;
σ_r	deals with the scale bandwidth and
σ_θ	handles angular bandwidth.

The segmentation is carried out in five steps: First, CXR is enhanced using log Gabor filter in Eq. (3) with $\omega_0 = 1/9$, $\theta_0 = 0$, $\sigma_r = 0.5$, $\sigma_\theta = \pi/9$ [10]. After enhancement, the threshold of 0.5 is applied to the enhanced image [6]. After this, edge detection using Sobel operator finds the edges. The reason for applying the edge detection step after thresholding is to avoid unnecessary large amount of edges [13]. The edges so found are not prominent as shown in Fig. 2. So as to highlight and smoothen the edges, dilation with an appropriate structuring element is done. Finally, the closed-loop enclosure of lung field is filled, and rest all the edges in the image are removed.

2.2 Feature Extraction

Following feature descriptors are used:

- Histogram of Intensity (ID),
- Histogram of Gradient magnitude (GD) and
- Shape Descriptor histogram (SD),

$$\mathrm{SD} = \tan^{-1}\left(\frac{\lambda_x}{\lambda_y}\right) \tag{4}$$

- Curvature Descriptor (CD),

$$\mathrm{CD} = \tan^{-1}\left(\frac{\sqrt{\lambda_x^2 + \lambda_y^2}}{1 + I(x, y)}\right) \tag{5}$$

where $I(x, y)$ denotes intensity of pixel (x, y), and λ_x, λ_y are the eigenvalues of Hessian matrix with $\lambda_x \leq \lambda_y$.

The eigenvalues of Hessian matrix for shape and curvature feature descriptor as given in Eqs. (4) and (5) are computed using multiscale approach by Frangi in [14].

- Histogram of Oriented Gradients (HOG) and
- Local Binary Patterns (LBP).

Each of these descriptors is quantised into 32 bins of histogram.

- Colour Layout Descriptor: CLD gives information regarding the dominant colour in the image using discrete cosine transform coefficients [15].
- Edge Histogram Descriptor (EHD): EHD gives information about the local edge distribution in image [15].
- Tamura Texture Descriptor: Tamura descriptor is actuated by human perception of vision. It is comprised of a set of six features; three of them (contrast, directionality and coarseness) have a strong correlation with human perception [16].
- Texture features by GLCM approach: Statistical texture analyse the spatial distribution of intensity values [10, 11].

2.3 Classification

Support Vector Machine (SVM), which is a supervised learning algorithm, is used as classifier to detect chest radiographs with TB manifestations. SVM generates hyperplane with largest margin to distinguish samples within two classes. In general, larger margin indicates lower generalisation error of classifier. Results are calculated with linear SVM.

3 Experimental Results

Experiments are done over two different datasets. Both of them are publicly available: (1) Montgomery County Hospital (MC) collection and (2) Shenzhen Hospital, China (CHN) collection. We have used 160 images of CHN dataset, out of which 75 are normal and 85 are TB-infected. Among 160, we have taken 115 images for training and rest 45 images (20 and 25—normal and abnormal, respectively) for testing which contributes about 28% of dataset. We use the reports by radiologists, confirmed on the

Table 1 Classification performance analysis on both datasets

Dataset	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)
MC dataset	96	0.98	91	100
CHN dataset	97.8	0.96	95	100

Table 2 Confusion matrix of SVM for Shenzhen dataset

	Normal	Abnormal
Normal	20	0
Abnormal	1	24

Table 3 Confusion matrix of SVM for Montgomery dataset

	Normal	Abnormal
Normal	20	0
Abnormal	2	28

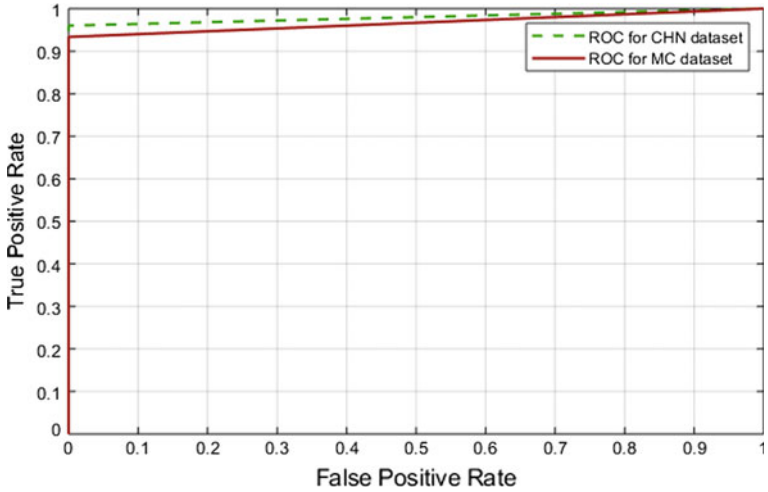


Fig. 3 ROC curve for SVM model efficiency analysis on both datasets

basis of clinical findings and patient history as ground truth. The proposed algorithm is showing very promising results on both the datasets. Table 1 shows the accuracy of the algorithm on the two respective datasets. In experimentation, it has been found that 1 out of 25 TB-infected CXR is incorrectly diagnosed while all the 20 normal CXRs are correctly predicted by the classifier as shown in confusion matrix of SVM in Table 2.

MC dataset consists of 138 CXRs. 58 of those are of infected patients while 80 are healthy images. We take 20 from healthy images and 30 from diseased images for training the classifier.

Classifier predicts all the healthy images perfectly while 2 of the diseased are predicted incorrectly. The confusion matrix of SVM for MC dataset is given in Table 3. Specificity of the algorithm on both the datasets is calculated as 100%, which shows that all the TB negative cases are predicted correctly. However, we achieve a sensitivity of 91 and 95.2% on MC and CHN datasets, respectively. Figure 3 shows the ROC curve for efficiency analysis of the classifier for both the datasets.

4 Conclusion

In this study, the potential of selection of features for classification can be appraised. We input a combination of texture and shape features to support vector machine. This approach is giving a really commendable result on both of the datasets. To check the robustness, we could have been tested on another dataset. But due to the lack of more datasets, it has not currently been possible. We are trying to get more datasets from Indian hospitals. The achievement of this algorithm is an AUC of 0.98 and 0.96 on MC and CHN datasets, respectively. In future experiments, we will experiment on more datasets to check the robustness of the algorithm.

References

1. World Health Org (2017) Global tuberculosis report
2. Daley CL, Gotway MB, Jasmer RM (2011) Radiographic manifestations of tuberculosis. In: A primer for clinicians, Second edn. Francis J. Curry National Tuberculosis Centre
3. Melendez J (2015) A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-rays. IEEE Trans Med Imaging 34(1)
4. Candemir S (2014) Lung segmentation in chest radiographs using anatomical at-lases with nonrigid registration. IEEE Trans Med Imaging 33(2)
5. Karargyris A, Antani S, Thoma G (2011) Segmenting anatomy in chest X-rays for tuberculosis screening. In: 33rd annual international conference of the IEEE EMBS, Boston, Massachusetts USA
6. van Ginneken B, Katsuragawa S, ter Haar Romeny BM, Doi K, Viergever MA (2002) Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans Med Imaging 21(2)
7. Xu T, Cheng I, Senior Member IEEE, Mandal M (2011) Automated cavity detection of infectious pulmonary tuberculosis in chest radiographs. In: 33rd annual international conference of the IEEE EMBS, Boston, Massachusetts USA
8. Jaeger S, Karargyris A, Antani S, Thoma G (2012) Detecting tuberculosis in radiographs using combined lung masks. In: 34th annual international conference of the IEEE EMBS San Diego, California USA
9. Jaeger S (2014) Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging 33(2)
10. Mohanaiah P, Sathyanarayana P, GuruKumar L (2013) Image texture feature extraction using GLCM approach. Int J Sci Res Publ 3(5)
11. Srinivasan GN, Shobha G (2008) Statistical texture analysis. In: Proceedings of world academy of science, engineering and technology, vol.36

12. Arrospide J, Salgado L (2013) Log—gabor filters for image based vehicle verification. IEEE Trans Image Process 22(6)
13. Joykutty B, Samuvel B (2016) A three stage detection of tuberculosis using adaptive thresholding in chest radiographs. Int J f Eng Res Technol 5(08)
14. Frangi AF, Niessen WJ, Vincken KL, Viergever MA (1998) Multiscale vessel enhancement filtering. In: Medical image computing 7 computer assisted interventions, Boston USA. Lecture Notes in Computer Science, vol 1496
15. Nazir A, Ashraf R, Hamdani T (2018) Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor. In: International conference on computing, mathematics and engineering technologies
16. Tamura H, Mori S, Yamawaki T (1978) Textural features corresponding to visual perception. IEEE Trans Syst Man Cybern SMC-8(6)
17. Wang W, Li J, Huang F, Feng H (2008) Design and implementation of Log-Gabor filter in fingerprint image enhancement. Pattern Recognit Lett 29:301–308 (Elsevier)

Role of X-Rays in Assessment of Bone Mineral Density—A Review

S. M. Nazia Fathima, R. Tamilselvi and M. Parisa Beham

Abstract Application of X-rays in healthcare applications is the energetic field involving more research challenges. The physician uses conventional or traditional X-ray images of the bone which is the majority used imaging modality to treat bone diseases, which disclose the potential for new medical applications. Inspired by the special characteristics, novelty and wide applications of X-ray, this work provides a comprehensive survey on the role of X-ray in the medical field. X-ray involves bone mineral density (BMD) measurement which is the conventional method to assess the fracture risk. Even though a lot of papers are describing the methods involved in X-ray, this paper explains the current trends with the advantages and disadvantages of X-ray in the medical field. Principles, fundamentals and characteristics of X-ray in the emerging area of biomedicine are also well described. The various recent works related to the principle, BMD measurement and analysis methods of X-ray are also discussed.

Keywords Medical field · Health care · X-ray images · Bone mineral density (BMD) · Osteoporosis

1 Introduction

Digital early experiments of the quantity of BMD used traditional X-rays which comprise a wedge with a step design made from an aluminium or ivory phantom. At each step of the phantom with the known densities, the minerals or density of the bone is estimated by an image assessment of the concentration of the minerals in

S. M. Nazia Fathima (✉) · R. Tamilselvi · M. P. Beham
Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India
e-mail: naziafathimasm@gmail.com

R. Tamilselvi
e-mail: rts.ece@gmail.com

M. P. Beham
e-mail: parisabeham@sethu.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_6

the bone. The next development in the area of measurement of bone density is the invention of single-photon absorptiometry (SPA) in the year 1963 by Cameron et al. Iodine (I-125) or americium (Am-241) which is a radioactive source, with energies of 27 and 60 keV is used in this method. The principle behind this technique is that the subject positioned their specific region in a water bath to afford the same path extent through which the gamma rays would be passed. Measurement of bone density involves the procedure of estimation of quantity of minerals present in the bone region examined by the procedure of subtraction e photons filtered by the spongy tissue from the photons filtered by both the tissues. This procedure established a good place in the medical field in the condition of quantification of the bone, but it was restricted to a marginal site. But when the density of the bone is to be measured at axial portions such as hip or spine, two different energies from gamma rays are mandatory to differentiate both hard and soft tissue, in which the soft tissue is of inconsistent depth [1].

Dual-energy problem is engaged in dual-photon absorptiometry (DPA), facilitating the concurrent diffusion of two photon energies of 44 and 100 keV from gadolinium-153 from gamma rays [1]. Algebraic equations derivation is used for the estimation of bone and soft tissue. In the late 1980s, single X-ray absorptiometry (SXA) and single-photon absorptiometry (SPA) have been outdated which are superior and expensive radioactive sources.

A dual-energy X-ray absorptiometry, commonly known as densitometry or dual X-ray absorptiometry (DXA) can produce whole body image and discriminate dissimilar body structures. The basic rule of DXA is similar to the DPA that the strength of the X-rays communication, formed from a constant X-ray source, at elevated and small energies. History of various techniques for BMD measurement is shown in Table 1. A lot of papers in X-Ray explained the basic principles and operation of X-ray. No review paper is available to discuss the survey related to the X-ray. This scenario motivated us to present a review paper on X-Ray.

In this paper, our focus is on X-Ray image review that is being used extensively in BMD applications. The main objective is to present a survey on X-ray that appeared in the literature over the past decade under all severe conditions that were not discussed

Table 1 History of BMD measurement

Technique	Principle
X-rays	Measurement of BMD is based on the aluminium or ivory phantom
SPA	Measurement of BMD is based on the gamma rays. Bone density is estimated by the process of decreasing of the photons filtered by the soft tissue from the photons filtered by bone and soft tissue
DPA	Works based on the concurrent communication of two photon energies of 44 and 100 keV from gadolinium-153 from the gamma rays
SXA	Based on the photon source as an X-ray system with solid state detectors
DEXA	Based on the quantity of the X-rays produced during the transmission and derived from an unwavering X-ray source, at huger and small amount of energies

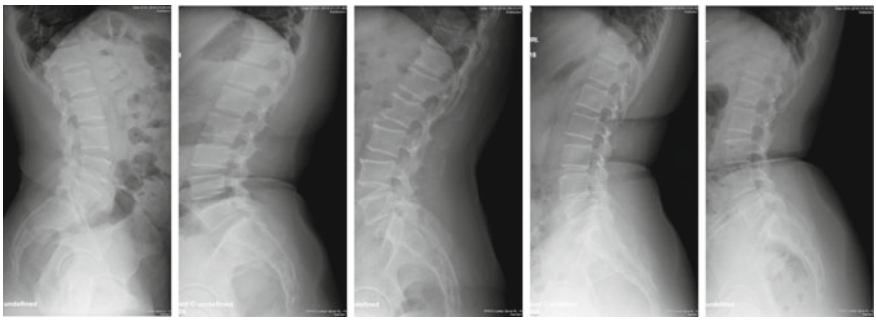


Fig. 1 X-ray images of spine

in the previous survey. The X-Ray image of the spine and femur are depicted in Fig. 1.

Some features of X-ray images are as follows:

1. Fracture diagnosis and treatment,
2. Evaluation of skeletal maturation,
3. Bone densitometry and
4. Hip replacement.

Osteoporosis is a diverse disease that influences diverse parts of the bones with different harshness. In the United States, more than 1.5 million osteoporotic fissures occur every year. This disease occurs commonly in women; in particular, people who are more than 50 years, osteoporosis occurs commonly in men also [1].

The primary obstacle of osteoporosis is fractures happening after minimum trauma. The proximal femur, the vertebral bodies, the distal radius and the proximal humerus are the normally frequent sites of fracture. Hip fractures are associated with increase in the short term death and high morbidity. Hip, vertebral and radius fractures boost the threat of upcoming risk of bone fractures. Common features of ageing are the diminishing minerals present in the bone and alteration of structure of the bone. The densitometry clarity of osteoporosis was traditionally given by consent of a board organized by the World Health Organization (WHO) in the year 1994. Only limited papers are available for describing the BMD measurement in X-Ray images.

This paper is planned as follows: Sect. 2 described the Principles of X-Ray. Section 3 explains the fundamentals of X-ray images in BMD measurement. Different methods in X-ray absorptiometry are described in Sect. 4. BMD measurements and various algorithms in X-rays are detailed in Sect. 5. Section 6 explains the BMD measurements and interpretations for osteoporosis. The work with detailed inference from the related reviews is concluded in Sect. 7.

2 Principles of X-Ray

Roentgen in the year 1896 successfully established X-rays. The first X-ray operation is executed to extort a part of a darning needle entrenched in the hand of a seamstress. Non-invasive diagnostic methods are overwhelmed by X-ray methods. In a simple nutshell, it is known that even today the specialized adjustment of conventional X-ray radiography is the dominated and most regularly used diagnostic method in medicine.

X-ray tube is used for producing X-Rays. Al filter used in the tube removes the minimal energy because as they are unable to break through the entity and involvement to the data on the film, they can simply append needless to the dose observed by the entity. Collimator extracts or filters the radiation of the X-rays external to the area of the image on the film. Reduction and Compton spreading appears at the object. The grid at the bottom allows the photons straight away from the source to the film. Figure 2 explains the basic principle of X-rays. The normal linear attenuation law explaining the reduction of X-rays are given by the equation:

$$I(x) = I_0 \exp(-\mu x) = I_0 xp(-\mu/\rho x \rho) \tag{1}$$

In Eq. 1, x is the distance passed through the material and μ is the linear attenuation coefficient in units of m^{-1}. I_0 is the intensity at the entry to the material ($x = 0$), $I(x)$ is the intensity at distance x, and ρ is the density of the material [2]. Mass attenuation coefficients for some common tissues are given in Table 2.

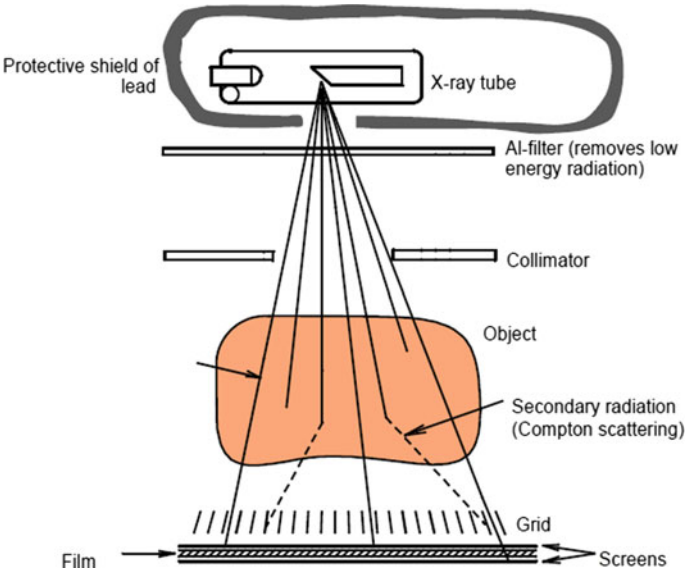


Fig. 2 Tissue density based on the attenuation

Table 2 Mass attenuation coefficients for typical tissues in μ/ρ

μ/ρ given in cm^2/g	50 keV	100 keV	200 keV
Air	0.208	0.154	0.122
Water	0.227	0.171	0.137
Adipose tissue	0.212	0.169	0.136
Muscle	0.226	0.169	0.136
Bone	0.424	0.186	0.131
Lead	8.041	5.549	0.999

3 Fundamentals of X-Ray Images in BMD Measurement

Radiographic absorptiometry (RA) explains the measurement of bone mass from the normal radiographs of marginal spot, regularly the hand or heel. The merits of conventional radiographs are less costly, more specific and perfect. This technique most commonly existing than any other bone mineral density calculating techniques as there is no necessitate for dedicated equipment. The main drawback of RA is that measurement of bone mineral contents is responsive to variations in the background of the soft tissues; the technique is restricted to the appendicular bones. The variables involved in the measurement are voltage setting, exposure time and film variables. An aluminium wedge is included in the radiograph to spot on for the above said variables. But, however, it has been accepted as a selection measure for the BMD measurement. Two dissimilar energies are used for radiographs of the region (hand) and an aluminium reference wedge. In the radiograph, the image is locked electronically and was investigated to establish the mean density of the middle phalanges of the second, third and fourth fingers. The results are stated in aluminium equivalents. Aluminium based on step-wedge captures the hand radiographs positioned on the film and optical densitometer is used for the image analysis which is a very standardized one. There exists a defined density of the aluminium step-wedge for comparing the obtained BMD results. From the survey, it is stated that RA dimensions were used to create or estimate the low bone mass of the lumbar spine and femoral neck with 90 and 82% sensitivity, respectively [3].

4 Different Methods in X-Ray Absorptiometry

Dual-energy problem is engaged in dual-photon absorptiometry (DPA), facilitating the concurrent conduction of rays from the gamma with the use gadolinium-153 of two photon energies of 44 and 100 keV. Algebraic equations derivation is used for the estimation of bone and soft tissue. In the late 1980s, single X-ray absorptiometry (SXA) and dual-energy X-ray absorptiometry (DEXA) have been outdated which are superior and expensive radioactive sources. A dual-energy X-ray absorptiometry, commonly known as densitometry or dual X-ray absorptiometry (DXA) is capable

of producing whole body image and discriminate dissimilar body structures. The basic rule of DXA is similar to the DPA that the depth of the X-rays transmission, formed from a constant X-ray source, at the elevated and small amount of energies. The analysis technique differentiates hard and smooth tissue in a different way. The stationary scan table is used with axial bone densitometry; the modality of option to calculate the BMD and whole body densitometry are the major modalities of DXA in medical techniques to assess body composition [1].

5 BMD Measurements and Various Algorithms in X-Rays

Sheng et al. [4] evaluate the recital of edge detectors used for outline detection of the right and left femur bone in X-ray images. Various edge detectors such as Roberts, Sobel, Prewitt, Canny's and Laplace operators are used for experimentation. The end results show that the Laplace operator could give good accuracy than any other in its application to X-ray images of femur bones, which has importance to medical and forensic experts. The results of Laplace operator can be compared with the results of other edge detection methods. Another method explains the estimation of BMD from the volume of human femur. The first step involves the extraction of femur. The second step involves the finding of reference point. If the centre point is a circle, parameters such as intensity and diameter can be estimated at the midpoint. From the value of the first midpoint, new point is established. The third step is the estimation of the quantity of the bone. Then the last step is the calculation of BMD for the detection of osteoporosis.

6 BMD Measurements and Interpretations for Osteoporosis

Table 3 provides the general performance of the X-Rays in the medical field. The table explains the fundamentals and principles of X-rays and various methods involved in the BMD measurements. BMD measurements have a significant responsibility in the assessment of patients at risk of osteoporosis and in the suitable use of anti-fracture treatment. Genetics is the far most vital feature in both genders, where family history and revise of monozygotic and dizygotic twins have revealed that hereditary feature can make 80% of the unevenness in mass of the bone, where age, heredity, previous fragility fracture, parental history of fragility fracture, female gender and early-stage menopause are the most imperative osteoporosis-related threat factors that cannot be distorted [5, 6]. The most crucial risk factors that can be malformed and are correlated to lifestyle are smoking, inactive lifestyle, low body mass index, diet lacking in calcium, high alcohol consumption and low levels of vitamin D due to shortage of sunlight exposure or malnutrition. Since there is no bone density

Table 3 General performance of X-rays in medical field

SL. No.	Author	Year	Remarks
1	Haidekkera et al. [3]	2004	The author explains the need for animal models, which are widely used to link in vivo modification in bone mineral density (BMD) with alteration in disease state of bone
2	Kheng [10]	2006	In this proposal, medical image segmentation methods are reviewed. According to the review, no perfect segmentation methods are existing in X-rays
3	Lim et al. [12]	2006	In this paper, three different stages are used such as removal of outline of the bone of attention, taking out of features from the X-ray image and organization of the bone based on the extracted features
4	Damilakis et al. [13]	2010	This article describes concise survey of the current X-ray techniques used for the accurate amount of minerals estimation of the skeleton, data on the intensity of radiation revelation linked with these techniques and details relevant to the radiation protection problems
5	Reimer et al. [14]	2011	This article explains the concordance between spine fracture detection on traditional lateral X-ray and lateral X-rays obtained from DXA scans. Vertebral anterior and posterior heights and the anterior/posterior (AP)-ratio were calculated

measurement which gives high percent accuracy results, the BMD test is a chief interpreter of whether an individual will have a fracture in the future [7, 8].

Generally, results of BMD measurement was evaluated with the normal or peak BMD of a healthy 30-year-old adult. Based on this the T-score value is presented. The difference was calculated in the units of standard deviations (SDs). T-score of 0 indicates that the person is a healthy young adult. If there are more standard deviations (below 0, designated as negative numbers), then the lesser your BMD and the greater your threat of fracture. T-score between +1 and −1 is measured as normal or healthy [9]. A T-score between −1 and −2.5 designate that you have small bone mass, which does not represent that the person was diagnosed with osteoporosis. But a T-score of −2.5 or lower indicates that the person was diagnosed with osteoporosis. The more value of negative number, the more cruel the osteoporosis. Correct analysis of a BMD quantity will be inclined by the reason of measurement such as to diagnosis osteoporosis, to evaluate fracture risk [1, 10, 11].

7 Conclusion

In the recent scenario, a lot of advanced scanning systems play the main role in medical imaging, as accuracy is needed in diagnosis and targeted drug delivery. But in the past decade, X-Ray images played a major role in the analysis of various bone fractures. BMD plays a vital role in the detection of osteoporosis. In a nutshell, various scan images were successfully developed and still more can be developed, and they have a great potential in all therapy treatments in healthcare applications. Even though a lot of scan systems are accessible for measurement of BMD, in the present system, DEXA is the acceptable gold standard for BMD measurement. Although handiness of DXA may be different from place to place, this method presently extensively exists in both hospitals and educational medical centres. Even though DEXA has its own advantages, it also has certain limitations. Only two-dimensional capacity of bone mineral content is provided for a three-dimensional bone. Also, it does not provide the information about the bone depth. Another important issue is the failure of bone detection algorithms in paediatric DEXA images. Thus, there is a lot of scope for the researchers to carry out further research in terms of this BMD measurement especially in children.

References

1. Nazia Fathima SM, Tamilselvi R, Parisa Beham M (2018) Role of dual-energy X-ray absorptiometry in assessment of bone mineral density—a review. In: Proceedings of international conference on informatics computing in engineering systems (ICICES, 2018)
2. Jensen M, Wilhjelm JE (2006) X-ray imaging: fundamentals and planar imaging. Report 1:8
3. Haidekkera MA, Stevensb HY, Frangosb Bárez JA (2004) Computerized methods for X-ray-based small bone densitometry. Comput Methods Program Biomed 73:35–42
4. Sheng HP, Huggins RA (1979) A review of body composition studies with emphasis on total body water and fat. Am J Clin Nutr 32(3):630–664
5. Crabtree NJ, Leonard MB, Zemel BS (2004) Dual-energy X-ray absorptiometry. In: Current clinical practice: bone densitometry in growing patients: guidelines for clinical practice, pp 41–57
6. Choplin HR, Lenchik L (2014) A practical approach to interpretation of dual-energy X-ray absorptiometry (DXA) for assessment of bone density. Curr Radiol Report 2:48
7. Lorente-Ramos R, Azpeitia-Armán J, Muñoz-Hernández A, García-Gómez JM, Díez-Martínez P, Grande-Bárez M (2011) Dual-energy X-ray absorptiometry in the diagnosis of osteoporosis: a practical guide, AJR 196:897–904. 0361–803X/11/1964–897
8. Garg MK, Kharb S (2013) Dual energy X-ray absorptiometry: Pitfalls in measurement and interpretation of bone mineral density. Indian J Endrocrinol Metab
9. Roth J, Palm C, Scheunemann I, Ranke MB, Schweizer R, Dannecker GE (2004) Musculoskeletal abnormalities of the forearm in patients with juvenile idiopathic arthritis relate mainly to bone geometry. Arthritis Rheum 50(12):77–85
10. Kheng LW (2006) Segmentation of bone structures in X-ray images. Thesis report, pp 1–78
11. Rabinovich CE (2000) Bone mineral status in juvenile rheumatoid arthritis. J Rheumatol Suppl 58:34–37
12. Lim SE, Xing Y, Chen Y, Leow WK, Howe TS Png MA (2006) Detection of femur and radius fractures in X-Ray images. In: 2nd International Conference on Advances in Medical Signal and Info

13. Damilakis J, Adams JE, Guglielmi G, Link TM (2010) Radiation exposure in x-ray-based imaging techniques used in osteoporosis. Eur Radiol 20(11):2707–2714
14. Reimer E (2011) A comparison of DXA and radiographic morphometry of the spine in osteoporotic women. Thesis Report, University of OSLO

Ms. S. M Nazia Fathima received her B.Tech. (Information Technology) from Sethu Institute of Technology, Virudhunagar, India in 2008, M.E. (Computer Science and Engineering) from Sethu Institute of Technology, Virudhunagar, India, in 2011. She is pursuing her Ph.D. in Anna University, Chennai.

Dr. R. Tamilselvi received her B.E. (Electronics and Communication Engineering) from Bharathidasan University, Trichy, in 2002, and M.Tech. in Advanced Communication Systems from SASTRA University, in the year 2004. She received her Ph.D. from Anna University, Chennai, in the year 2014. Currently, she is a Professor in the Department of Electronics and Communication Engineering in Sethu Institute of Technology, Virudhunagar, India. Her research interests include image processing, signal processing and neural network. She is a life member of ISTE, BMESI and IACSIT.

Dr. M. Parisa Beham received her UG degree in Electronics and Communication Engineering from Institution of Engineers, India, in 2000, and M.E. degree in Applied Electronics from Mohamed Sathak Engineering College, Ramnad, in 2006. She received her Ph.D. from Anna University, Chennai, in the year 2016. Currently, she is working as an Associate Professor in the Department of Electronics and Communication Engineering, Sethu Institute of Technology, Virudhunagar, India. Her areas of research include computer vision, pattern recognition, image analysis and image processing. She is a life member of IET, IE (I), ISTE and IUPRAI.

Estimation of Face Pose Orientation Using Model-Based Approach

M. Annalakshmi, S. M. Mansoor Roomi and M. Parisa Beham

Abstract In the domain of computer vision and pattern recognition, though there are numerous methods for face recognition, it is still remaining as a very challenging problem in real life applications. Face detection and recognition suffer from many problems which are caused by the variations in orientation, size, illumination, expression, and poses. This paper mainly revolves around face detection and oriented pose identification. The state-of-the-art Constrained Local Model (CLM) is applied to detect the face from any wild facial image. The extracted feature points are used to segregate the dominant parts of faces. From the dominant feature points, nose tip and eye points have been identified. Applying the geometrical parameters between the nose tip and eye points, the pose orientation of the wild face has been identified. This method is very simple and accurate. The performance evaluation has been done on unconstrained Essex database and internal wild database collected from internet.

Keywords CLM model · CLM search · Segregation · Pose estimation · Geometrical parameters

1 Introduction

The most significant human sense is vision. Henceforth, the fact that images play vital role in human perception is not surprising. Computer vision is the science that develops algorithmic basis by which useful information can be automatically

M. Annalakshmi · M. Parisa Beham (✉)
Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India
e-mail: parisaphd2011@gmail.com

M. Annalakshmi
e-mail: annam.baluss@gmail.com

S. M. Mansoor Roomi
Department of ECE, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India
e-mail: smmroomi@tce.edu

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_7

Fig. 1 Various pose orientations of wild face images

extracted and analyzed from a set of images. The need to emulate perception makes image understanding a vital ingredient of any computer vision system.

In the literature, so many algorithms are proposed for face detection and recognition. The pose orientation (as shown in Fig. 1) is a major issue in recognizing face which drastically reduces the recognition rate. Asthana et al. [1] proposed an ideal approach to the problem of pose-invariant face recognition which uses a 3D pose normalization method. This outperforms state-of-the-art recognition for datasets with wide pose variation. Though various poses are considered, extreme cases are not taken into account. Asthana et al. [2] proposed a novel discriminative regression-based approach for the CLM frame-work, which is known as Discriminative Response Map Fitting (DRMF) method. This method outperforms the state-of-the-art RLMS fitting method and the tree-based method. Zhu & Ramanan [3] depicted a tree-structured model which is easy to optimize compared to dense graph structures. A novel method for pose robust face recognition towards practical applications is proposed by Yi et al. [4]. An extended 3D Morphable Model (3DMM) which isolates identity variations from those due to facial expressions is proposed by Chu et al. [5]. Some of the facial deformations are not handled well in [4] and [5]. To overcome those problems Sarode & Anuse [6] proposed a model-based approach in which moment based feature extractions (Hu's and Zernike Moments) are implemented. The performance of Hu's moments is not satisfactory and thus the recognition rate is poor. Tai et al. [7] used an Orthogonal Procrustes Regression based approach.

This method shows high efficiency for the misaligned test images with pose variations in spite of the training images (frontal or non-frontal). The primary contributions of this work are as follows:

1. The standard CLM model is used to identify the feature points of the wild face images.
2. The eye points and nose tips have been identified using geometrical formulation by which the face orientation is estimated.

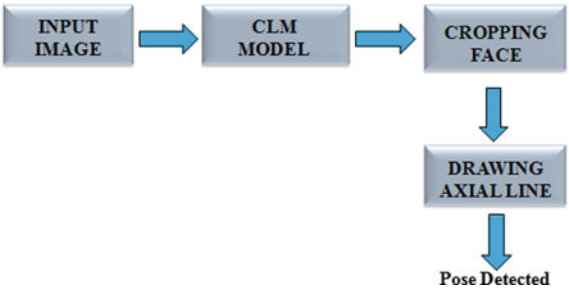


Fig. 2 Overall flow diagram of the proposed method

3. The proposed algorithm is validated with the unconstrained and wild database having high degree of pose variations.

2 Proposed Method

The proposed method is used to detect the orientation of the face captured in an unconstrained environment. The overall flow diagram is shown in Fig. 2. First, the input image is applied to a CLM model where the face is detected and then the face model is fitted. This model gives 68 feature points for the detected face. These points include nose, eyes, mouth, and outer boundary of the face. For the ease of analysis, the face regions are cropped. The feature points obtained are then used for segregating the various parts of the face. Following which an axial line is drawn which passes through the nose tip. Depending on the slope of this line, the orientation of the face is determined. This methodology aids in the automatic categorizing of database in many face detection and recognition algorithms.

3 CLM Model

Face detection and alignment is a two-step process involved in CLM. As a first step, human face is detected using a Viola–Jones face detector [8]. This detector finds face in the given input image and gives rectangles of where the faces are found. Next, the rectangle has to be closely observed to determine the boundary of nose, mouth, etc. Feature points are detected for that parts on face. The basic architecture of CLM model is given in Fig. 3.

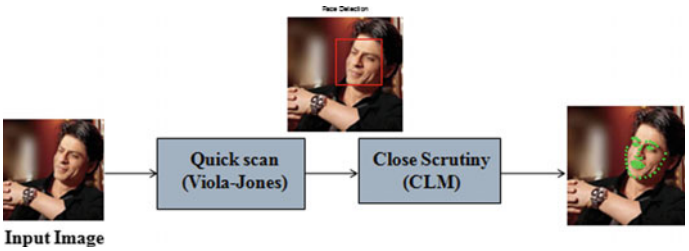


Fig. 3 Basic architecture of CLM

3.1 Model Building

CLM model consists of two parts; one part describes the shape variations of feature points called 'Shape Model', and the other part describes each patch of image around the feature points, called 'Patch Model'. The shape model describes how the face shape can vary and is done with PCA. Average of all shapes is calculated to obtain the mean shape. The mean shape is subtracted from each shape, which gives the information of how each shape varies from the mean face. The Patch model describes how image around each feature point should look like. In this model, Linear Support Vector Machine (SVM) is employed to recognize the local patch around the feature point.

3.2 CLM Implementation

To apply CLM on face alignment, first, a CLM model has to be built which contains the shape information and patch appearance information. This information can be used to search given image for faces.

Building Shape Model with PCA: The building shape model includes two steps: i. Remove scale, rotation, and translation with Procrustes analysis, output aligned shapes. ii. Do PCA on aligned shapes.

Procrustes Analysis: Two shapes of the applied face are obtained using Procrustes. Here, it is mandatory to find rotation 'θ', scale 'a' and translation parameters '(t_x, t_y)' to minimize the mean square error between the first and transformed second shape.

The optical parameter values can be derived analytically by taking partial derivatives of the mean square error function. All the shapes are aligned to the first shape, which removes scaling, translation, and rotation variation leaving only the face variation. Finally PCA can be applied on the aligned shapes.

PCA on Aligned Shape: PCA is used to find the basis of face variations. The basis of variations (Eigenvectors) and a value indicating how much the total variation that basis function accounts for eigenvectors is obtained by considering M images, each

image contains N feature points, and each feature points is given by (x, y) coordinates, representing the feature coordinates of an image into one vector:

$$x = [x_1, x_2, x_3, \ldots, x_N, y_N]^{\mathrm{T}} \tag{1}$$

Since the number of images is M, we have x^i, where $i = 1, 2, 3, \ldots, M$. The mean shape is given by

$$\overline{x} = \frac{1}{M} \sum_{i=1}^{N} x^i \tag{2}$$

Then the mean shape is subtracted from every x^i. Now the resultant will be given as

$$X = \begin{bmatrix} x_1^{(1)} & y_1^{(1)} & x_2^{(1)} & y_2^{(1)} & \cdots & x_N^{(1)} & y_N^{(1)} \\ x_1^{(2)} & y_1^{(2)} & x_2^{(2)} & y_2^{(2)} & \cdots & x_N^{(2)} & y_N^{(2)} \\ \cdots & & & & & & \\ \cdots & & & & & & \\ x_1^{(M)} & y_1^{(M)} & x_1^{(M)} & y_1^{(M)} & \cdots & x_1^{(M)} & y_1^{(M)} \end{bmatrix} \tag{3}$$

where each of X is the vector x^i from one shape. Then PCA is calculated from eigenvector and eigenvalue of $X^{\mathrm{T}}X$. The eigenvectors corresponding to large eigenvalues are alone considered. Now the eigenvectors and eigenvalues are used to form a matrix P.

Shape Decomposition: x can be written as a linear combination of eigenvectors plus the mean shape, where $\bar{x}$ represents the mean shape.

$$x = \overline{x} + PB \tag{4}$$

P is the eigenvector and B is the weight vector given by

$$B = P^{\mathrm{T}}(x - \overline{x}) \tag{5}$$

Thus the shape is reconstructed using Eq. 4.

Building Patch Mode: The most common notation of linear SVM is to express output as inner product of input data and support vectors.

$$y^{(i)} = \sum_{j=1}^{N_s} \alpha_j \langle x_j, x \rangle + b \tag{6}$$

where x_j is a subset of input training sample, called support vectors, and α_j the weight of the support vector, N_s being the number of support vectors, and 'b' is a

bias. For CLM, SVM output is written as a linear combination of elements of input vector:

$$y^{(t)} = w^t \cdot x^{(i)} + \theta \tag{7}$$

where w^t represent the weight for each element of input data, and θ is a constant acting as a bias. In CLM case, the set of training data are the patches extracted from the training images, and the output $y^{(i)}$ is set to 1 if the sample comes from its feature point position, and -1 if the data is randomly sampled away from the feature points. Using these data, linear SVM can be trained by finding w and θ. By allowing output errors, a solution is found for w and θ. Now, that the output is not exactly -1 and 1, then the output is interpreted as an indicator of closeness of the output.

3.3 CLM Search

The process of finding the nose, eyes, and mouth in the rectangle image is called Searching. The search process consists of the following steps: (1) Viola–Jones face detector is used to find faces in the image which draws rectangles around the faces. (2) An initial guess of each feature point position is made using mean shape. (3) For each feature point, a patch of image from its current position is cropped, and linear SVM is employed to find output at each point in the local region. A quadratic function is fitted to the response image thus obtained. (4) The best feature point positions are found by optimizing a function, created by combining the quadratic functions from step 3 and shape constraints from CLM shape model. (5) Each feature point is moved to its new position, and the above steps are repeated, until all feature points reach their best position. For each feature point, a response image is obtained. If high value is obtained in the response image, then the match score is high. This response image is used to determine the position of each feature point. The best position for each point from the response image is obtained by taking into account the allowed shape variation. Since the search is done only in a local region around current position, the result may be the best one locally. To find the global best position, the above steps have to be repeated until all points reach a stable position.

3.4 Segregation

The CLM model provides the final fitting for the face. From this model, a total of 68 feature points are obtained. These feature points are distributed over the face. The feature points are numbered as:

- 1:18 is used to represent the outer boundary
- 19:27 is used to represent the eyebrow

- 28:36 is used to represent the nose
- 31 is used to represent the nose tip
- 37:48 is used to represent the eyes
- 49:68 is used to represent the mouth

The required features points are extracted from the total feature points and can be further used for analysis.

4 Pose Orientation Estimation

The proposed pose orientation estimation is described in this section. The cropped face is used for the analysis. The first step is to draw the axial line which divides the image into two equal halves. The line that connects the eye points is drawn first and then the vertical line is drawn. The vertical line is drawn using midpoint of the horizontal line and the slope computation. Then the vertical and horizontal lines are extended for the full image. This process is further analyzed to find the pose of the image. The pose of the face is estimated using the slope of the drawn axial line. If the slope of the line is negative, then the face is said to be tilted towards the left. On the other hand, if it is positive then the face is said to be tilted towards the right.

5 Results and Discussions

The ultimate aim if this work is to identify the pose orientation of different facial images taken in an unconstrained environment. As manual classification of pose orientated database is tedious, it is necessary for an automatic system. For performance evaluation and validation of our algorithm, the wild face images which are more unconstrained are collected from the Internet. The database consists of 5 different subjects, each constituting 20 sample images. Our proposed algorithm has also been evaluated on Standard Essex unconstrained database [9].

5.1 CLM Fitted Images

In the proposed methodology, the input image is first applied to the CLM model fitting. The sample of CLM fitted images is shown in Fig. 4.

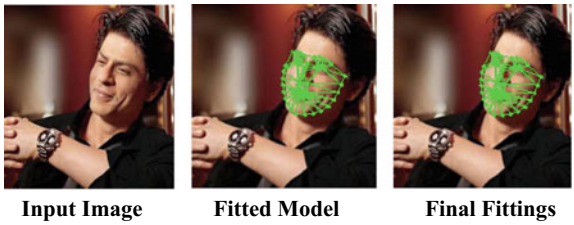


Fig. 4 CLM fitted images

Fig. 5 Eyes and nose points are extracted from the face feature points

Fig. 6 Axial lines drawn between eyes points and nose tips

Fig. 7 Result of pose orientation estimation

5.2 Segregation

From the feature points, the various parts of the input image can be extracted. Figure 5 shows various dominant parts that are extracted from the feature points. After extracting the nose tip, the vertical and horizontal axial lines are drawn that passes through the nose tip. In the next step, a line is drawn which connects the eye points of the detected face. Figure 6 shows the axial lines through eyes and nose tips. In Fig. 7a, both axial lines have been extended for the full image. By calculating the slope value of the vertical axial line, the pose orientation or direction of tilt of the input face image is obtained as shown in Fig. 7b.

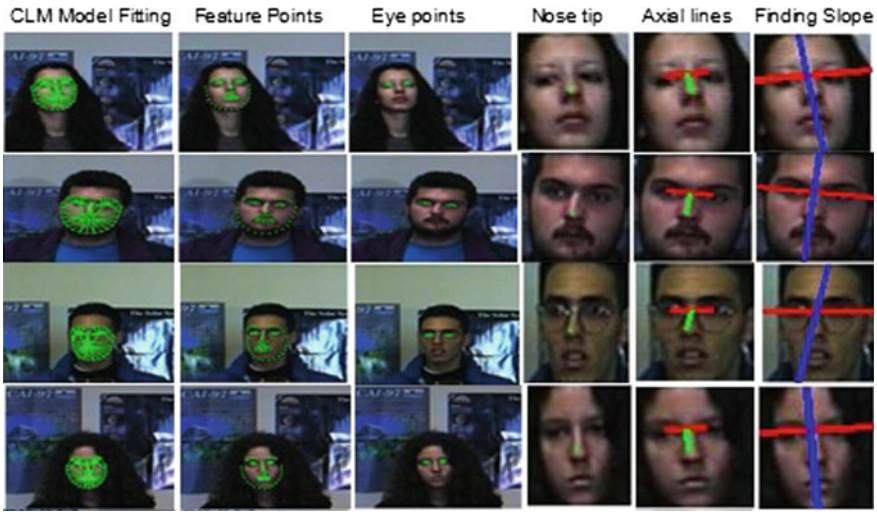


Fig. 8 Qualitative results on Essex database

5.3 Results on Essex Database

The Essex database is taken under extreme illumination condition, wild background, various facial expressions, and pose orientations. From the Essex database, 10 subjects each with 20 samples have been considered for performance evaluation. Figure 8 shows the qualitative results of the proposed algorithm.

6 Conclusion

This study of pose orientation and estimation presents an effective way of fitting a model to the input face image. The CLM model proved to be highly invariant to scaling, translation, and poses. This proposed algorithm has no constraints on the size of the image or the quality of the image. The best way of analyzing the parts of the face is presented.

The various parts of the face are segregated for further analysis. This algorithm is best suited for automatic grouping of images. This can be used in face recognition algorithms for pose orientation normalization prior to matching. Our qualitative experiments have demonstrated that the proposed method is competitive at recognizing faces with different horizontal wild pose orientations different from those in the gallery. The results have been shown on Essex and Internally collected wild database. The future work revolves around determining extreme poses. The model is obtained by analyzing the patch model individually. This might take more time. Hence, it is

required to build parallel processing for patch analysis. The database samples have to be increased for obtaining higher accuracy.

References

1. Asthana A, Marks TK, Jones MJ, Tieu KHT, Rohith MV (2011) Fully automatic pose-invariant face recognition via 3D pose normalization, 2011, TR2011-074
2. Asthana A, Zafeiriou S, Cheng S, Pantic M (2013) Robust discriminative response map fitting with constrained local model, CVPR, 2013
3. Zhu X, Ramanan D (2013) Face detection, pose estimation and landmark localization in the wild, CVPR, 2013
4. Yi D, Lei Z, Li SZ (2013) Towards pose robust face recognition, CVPR, 2013
5. Chu B, Romdhani S, Chen L (2014) 3D aided face recognition robust to expression and pose variations, 2014
6. Sarode JP, Anuse AD (2014) Face recognition under pose variation, 2014
7. Tai Y, Yang J, Zhang Y, Luo L, Qian J, Chen Y (2016) Face recognition with pose variation and misalignment via orthogonal procrustes regression. IEEE Trans Image Process 25(6)
8. Jones MJ, Viola PA (2003) Fast multi-view face detection. Technical Report TR2003-96, 2003, Mitsubishi Electric Research Laboratories, Cambridge, MA
9. The Essex Genealogist. Online database: American Ancestors.org. New England Historic Genealogical Society, 2011

Accurate Classification of Cancer in Mammogram Images

M. Parisa Beham, R. Tamilselvi, S. M. Mansoor Roomi and A. Nagaraj

Abstract In the last decade, machine learning plays a vital role in the detection of breast cancer. Mammography is a proficient tool for early stage detection of breast cancer. In this work, a simple technique for breast cancer image classification in l mammogram images is proposed. Highly discriminant local binary patterns are extracted from the wavelet normalized mammogram images. K-nearest neighbor classifier is used to categorize the abnormal cancer cell images. A mammogram database is created to evaluate the efficacy of our algorithm. From the experimental results, the performance of our algorithms is comparatively good with very less computational time.

Keywords Mammogram database · Cancer cell detection · Benign and malignant · LBP · K-NN classifier

1 Introduction

In biomedical engineering, the concepts are related to biotechnology that is used for various healthcare purposes. It solves problem in engineering field that is related to biotechnology and it is used for advanced treatment. Biomedical filed focuses on human health and healthcare at various levels. It includes the development of various diagnosis and therapy medical devices that range from clinical equipment to micro-

M. Parisa Beham (✉) · R. Tamilselvi · A. Nagaraj
Department of ECE, Sethu Institute of Technology, Virudhunagar, Tamil Nadu, India
e-mail: parisaphd2011@gmail.com

R. Tamilselvi
e-mail: rts.ece@gmail.com

A. Nagaraj
e-mail: nagaraj.sa@gmail.com

S. M. Mansoor Roomi
Department of ECE, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India
e-mail: smmroomi@tce.edu

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_8

implants. A biomedical engineering application includes the design and development of active medical devices, orthopedic and medical imaging. In the biomedical imaging, mammogram is an X-ray scan of the breast to detect and experiment any changes in the breast. Medicolegal implications and high volume variability in interpretation are an issue in mammography. Also, double reading is not feasible in mammogram images. More modalities/procedures can be helpful in patient management but are time consuming. Positioning was one of the main challenges identified by the radiographers. First, we do the positioning with a phantom, then with their peers and finally on patients. Image quality assessment and control in mammography are significant due to the impact on Breast Pathology detection (BPD).

Mammogram (MGM) image analysis has been an active research area and much research has been carried out by early researchers. A comprehensive literature survey on various techniques to identify cancer cells in MGM images is collected. Analyzing the entire image and by referring many papers, many attempts were made to improve its rituality of the X-ray images and accuracy of the calculating relative parameters. Sapate and Talbar [1] focus on the efforts made in the direction of solving the pectoral muscle extraction issue in the preprocessing task for detecting breast cancer in the early stage using MGMs. On the other side, there are very rare cases where the computational complexity has been calculated with a due importance.

Sivakumar et al. [2] used an enhanced artificial bee colony optimization algorithm to detect the borders and to find the nipple spot. Their proposed EABCO algorithm is used to resolve the extraction of suspicious region from the MGM image using bilateral subtraction. The experimental results and comparison based on the bilateral approach produce better result. Maitra et al. [3] used more number of algorithms such as binary homogeneity enhancement algorithm, standard convolution methods, image segmentation method, seeded region growing algorithm, pectoral muscle detection algorithm, and breast boundary detection algorithm. The combination of these methods has provided magnificent detection of mass in MGMs images like cyst, calcification, or fatty masses. Nithya et al. [4] proposed region growing algorithm with top-down methodology. This process isolates the region of interest (ROI) from the background image. From the segmented images, they are using genetic algorithm features. In this survey paper, the mammography is regarded as one of the optimal ways to find breast cancer in the early stage. Enhancement of the mammography image is done by median filter. The overall segmentation is handled by region growing algorithm.

Percha et al. [5] presented methods such as plasmon coupling effects, back propagation neural network, and detection method. They used micro-calcification algorithm, computer-aided detection (CAD) algorithm, and clustering algorithm to detect the micro-calcification clusters and show the result of 84.3%. Rubin et al. [6] describe how to differentiate the speculated masses and cyst. The most frequently occurred abnormalities to indicate breast cancer are masses and calcifications. Gray level co-occurrence matrices (GLCM), a texture feature is used to differentiate the masses and normal breast tissues from the input image. Angayarkanni et al. [7] use edge detection and segmentation method for the detection of tumor in the breast from mammogram

images. By applying various image processing algorithms, the demarcation of cyst and tumor in the mammogram image is acquired.

From the literature survey, many researchers used different algorithms to detect the cancer cells, but those methods could not obtain high accuracy and the computation time is also high. Motivated by this fact, the aim of our paper is to detect abnormality of mammogram images in a simple and efficient manner. Filtering techniques are used to remove the noise, normalization is used for improving the quality of MGM image and feature extraction techniques, and K-NN classifier is used to increase the classification accuracy. Our experimental analysis proves that the cancer cell images are detected more accurately.

2 Proposed Methodology

In this proposed methodology, the complete steps involved to diagnose the breast cancer using digital mammogram images are given in Fig. 1. The first step is database of mammogram images are collected and grouped for further process. Next, in the preprocessing step, discrete cosine transform (DCT), single-scale retinex (SSR), gradient and wavelet normalization techniques are used for the enhancement of the mammogram images. Then the discriminant features are determined using standard local binary pattern (LBP) algorithm. Finally, the abnormality detection has been obtained using simple k-nearest neighbor classifier.

2.1 Database Collection

Internally, the mammogram images have been collected from Pixels Scan Centre, Trichy. Totally, 149 mammogram images are collected from different age group persons. The total number of images in normal mammogram is 122 and the abnormal

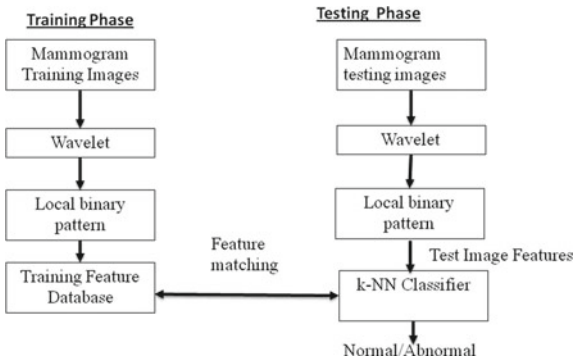


Fig. 1 Step by step process of the proposed methodology

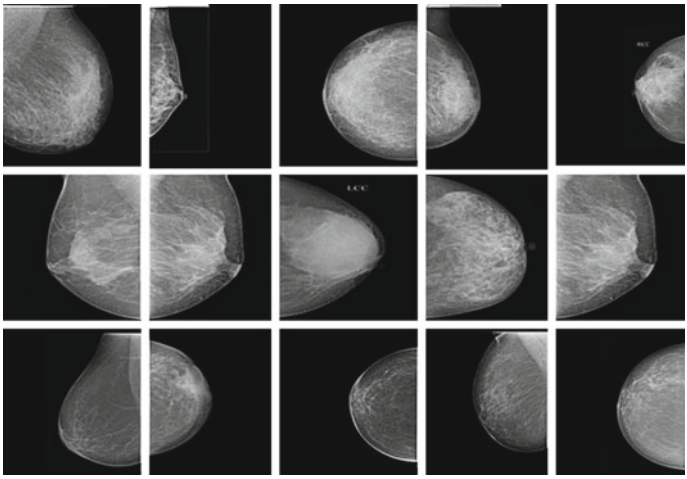


Fig. 2 Sample MGM images from our database top: normal, center: malignant, bottom: benign

mammogram is 27 images. From that normal image, we further separated the benign and malignant images. Thus, these images are further classified into three groups, namely normal images, benign images, and malignant or abnormal images. All the 122 normal images and 27 abnormal images are used for our experimentation and evaluation. The dimension of all the images is 3540 $\times$ 4740. So as to reduce the complexity, all the images are resized to 128 $\times$ 128. An example of images from mammogram database is shown in Fig. 2.

2.2 Preprocessing

In image processing, normalization is a technique that alters the dynamic range of image pixel intensity values. Normalization is also called as contrast stretching or histogram stretching. Wavelet normalization is frequently applied for image enhancement because of lower computational complexity. Wavelet normalized images are shown for normal, malignant, and benign mammogram images in Fig. 3.

2.3 Feature Extraction

Feature extraction technique is the one, which is used to extract discriminant features from the normalized images. These features can be applied to standard classifier to classify the abnormal mammogram images accurately. Local binary pattern (LBP) [8], a standard feature extraction algorithm is used in this work. The main objective

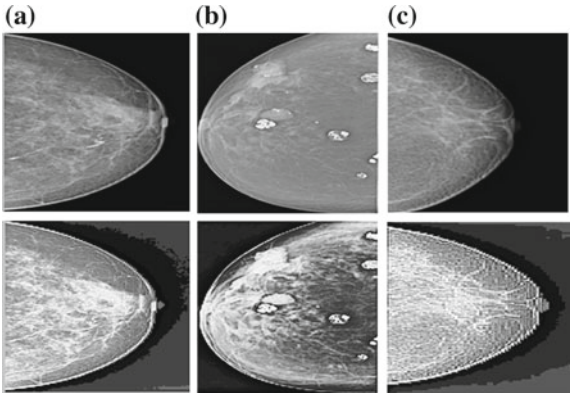


Fig. 3 Top row: **a** original image, **b** benign image, **c** malignant image. Bottom row: respective wavelet normalized images

of LBP is used to sum up the local formation of an image by comparing each pixel with its neighborhood pixel.

The LBP is found by

$$\mathrm{LBP}_{P,R} = \sum_{P=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \tag{2}$$

2.4 K-NN Classifier

K-NN classifier is used for binary or multiclass classification or regression problems. In training phase, a model using the training images is created from the training instances for given data, and the classification algorithm is used to find the relationships between predictors and targets objects in an image. In testing phase, a test sample class labels is known, but that class label is not involved in the training model. The K-NN classifier is comparing the objects by converting all features of the instances into numerical value for different types of data that is numbers colors [9]; geolocation, Booleans, etc., and it represents instances as vectors of features in an n-dimensional space. The k value will be chosen. If k is too small then it is called as over-fitting. The algorithm performs too well on the training set, compared to its true performance on unseen test data.

3 Experimental Results

The internal mammogram database is randomly divided into training and testing data. The data can be divided such that all classes have balanced representation in both testing and training. Feature vectors which are extracted from the normalized images are nothing but a collection of features of an MGM image. In case of MGM images, geometric features or texture features can be derived.

In this work, after wavelet normalization, texture features are extracted using LBP and then the feature vectors are applied to the K-NN classifier for accurate classification. Using K-NN classifier, by varying the k values, the classification accuracy has been obtained under various normalization techniques. From Table 1, it is observed that compared to all the other normalization techniques [10], and LBP features extracted from the wavelet normalized images have produced better classification rate of 73.46% when the k value is one. The experimentation has also been done by varying the block sizes of LBP as 3, 8, 9 and 12. From Fig. 4, it is inferred that the higher classification accuracy of 71.14% is obtained by the proposed method when the block size is 12.

Table 1 Classification rate of various normalization techniques

Normalization method	Classification rate				
	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
DCT	63.26	67.34	57.06	59.18	55.10
SSR	73.46	61.22	59.18	63.26	67.34
Gradient	69.38	61.22	57.14	59.18	55.1
Wavelet	**73.46**	**63.26**	**63.26**	**63.26**	**53.06**

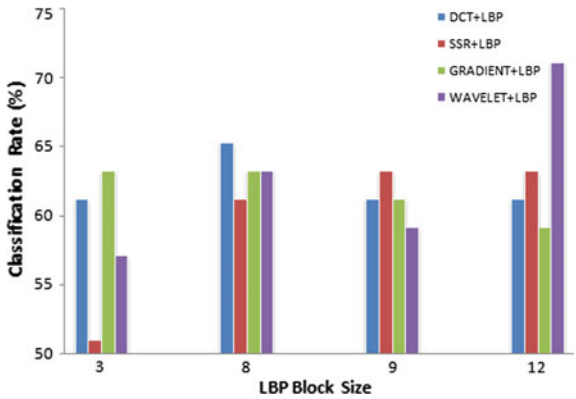


Fig. 4 Classification rate is shown for various LBP block size

4 Conclusion

It is mandatory to detect the breast cancer in its earlier stage to save/extend the life of women. In this work, wavelet normalization technique is applied to enhance the image quality and to improve the edge and line points. LBP is used to extract the highly discriminant features. The classification results have been shown for various normalization techniques by varying the k values and LBP block sizes. Our proposed method performs well compared to other normalization techniques. As a future direction, still more prominent feature extraction algorithms and fusion of normalization techniques can be applied to improve the overall classification accuracy of abnormal mammogram images.

Ethical Approval The mammogram database used in this paper is provided by Pixel scans, Trichy. The ethical committee of Pixel scans has reviewed and approved to conduct research using this mammogram database and publish papers based on the results using that biomedical images.

References

1. Sapate S, Talbar S et al. (2005) Pectoral muscle extraction algorithms applied to digital mammograms. Institution Politecnico Nacional, Centro de Innovation y Desarrollo Technology en Computer, Mexico 7(2)3:5–6
2. Sivakumar R et al. (2007) Diagnose breast cancer through mammogram enhancement method and nipple position. Inf Commun Technol Res Mumbai 8(3)
3. Maitra IK et al. (2007) Automated digital mammogram segmentation for detection of abnormal masses. Indian J Comput Sci Eng (IJCSE) 1(1). ISSN 2007-3689
4. Nithya R et al. (2011) Improving performance of breast cancer detection for mammography image. Int J Adv Comput Sci Appl 25(5)
5. Percha B et al. (2012) Computer aided detection algorithm for digital mammogram images. Int J Comput Trends Technol (IJCTT) 2(2)
6. Rubin D et al. (2013) Identification of abnormal masses in digital mammography images. In: 13th International Arab conference on information technology ACIT 2013
7. Angayarkanni N et al (2014) The application of image processing techniques for detection and classification of cancerous tissue in digital mammograms. J Pharm Sci Res 8(10):1179–1183
8. Kuzmiak et al. CM (2000) Automatic classification of mammography report by bi-rad breast tissue composition class. University of North Carolina, USA, 4(6)
9. Gonzalez-Patino D, Villuendas-Rey Y, Argüelles-Cruz AJ (2015) Mammogram image segmentation using bioinspired novel bat swarm clustering. Instituto Politecnico Nacional, Centro de Investigation en Computation, Mexico City, Mexico 67(2):12–16
10. Lipson J (2016) An efficient image processing methods for mammogram breast cancer detection. J Theor Appl Inf Technol 69(1)

Low-Power Extended Binary Pattern Image Feature Extraction

S. Arul Jothi and M. Ramkumar Raja

Abstract The blindness in the people is due to various reasons such as age-related macular degeneration (AMD) and diabetic retinopathy (DR) [1]. This work deals with identification of affected and healthy images based on the discrimination capabilities in fundus image textures. For this purpose, texture descriptor algorithm extended binary pattern (EBP) is used for retinal images and the area and time consumption has been reduced by the means of it. The main aim of the work is to reduce time consumption and also categorize the retinal diseases with the retina background texture.

Keywords Extended binary patterns · Retinal disease

1 Introduction

The retinal disease destroys the sharp, central vision which in turn increases the cases of visual problems. It is found that most of the visual problems are curable if they are diagnosed in the early stage. The age and the diabetics are the two main causes of retinal disease. The early diagnosis of the disease, through appropriate method, will reduce the cost of treatment than that of the disease become chronic (Fig. 1).

S. Arul Jothi (✉)
ECE Department, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India
e-mail: arulbe2005@gmail.com

M. Ramkumar Raja
ECE Department, Coimbatore Institute of Engineering and Technology, Coimbatore, India
e-mail: ramkumarrajavlsi@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_9

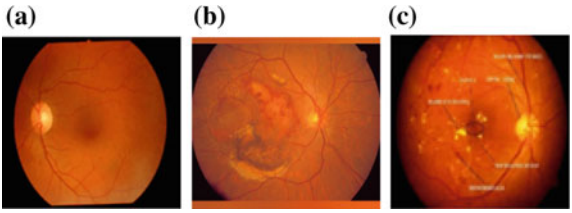


Fig. 1 Affected retina images. **a** Healthy, **b** DR and **c** AMD

2 Literature Survey

In grayscale, the variations are a very robust approach. In local image, the texture of spatial configuration has been differentiated by the operator. The tools used for rotation invariant texture analysis are the joint distribution [2]. In the facial representation that has been proposed here is LBP, it has been used in many applications like the classifications of textures and retrieval of images areas. Here, the proposed methods are assessed using the task, face recognition [3]. Myocardial infarction disease affected persons are been mentioned in this work. Here, the difference in texture of scarred myocardium tissue is obtained using LBP [4]. In the brain, lesions of white matters are the signals of hypertense in weighted-T2 MRI. The idea is to find the patients with dementia and healthy elderly can be distinguished using classification of framework [5]. Treatment of different eye diseases the retinal fundus images are used. Big screening of DR has been done with the help of its main resources. Abnormal signs have been obtained by the processing of fundus images [6]. Multi-resolution analysis is used for detection of age-related macular degeneration. In the wavelet coefficient, the textural pattern distributions are used to describe an image [7]. There are few operators that are defined to extract the local gray level of the image along with the magnitude and sign features of local difference [8]. Small spatial support area of it is LBP. For a 3×3 neighborhood, the features are calculated [9]. LBP features the robust to rotation of image, and it is very less sensitive to histogram equalization and also for noise [10]. To handle the noises of image, the directed and undirected difference masks of Noise-robust descriptor are designed. These patterns obtained as results are consecutively encoded to form texture of final representation [11].

3 Materials and Method

The database consists of images of retina that are affected. The databases used are structured analysis of the retina (STARE) and E-OPHTHA [12].

Table 1 Comparison between LBP and EBP

Parameters	Feature	TPR	TNR
AMD	LBP	1.000	0.990
	EBP	1.000	0.992
DR	LBP	0.856	0.897
	EBP	0.924	0.941

4 Existing Method

An effective local binary pattern (LBP) feature extraction is used for retinal disease classification [13]. Local binary pattern is used to calculate texture pattern probability that is summarized into histogram, and LBP values for all the pixels of the image are calculated and the feature is extracted to identify the retinal disease. After the features extraction, the data set has to be preprocessed before the classification of the retinal disease.

5 Proposed Method

An automatic method without prior segmentation for retinal image classification has been developed. Extended binary pattern (EBP) algorithm based on the texture analysis of the retina is used for classification. Figure 2 shows the flow diagram of the proposed method where the input image is rescaled, masking is done with neighborhood pixels, filtering process to remove the noise, angular difference and radial difference are two features extracted, principal component analysis is used for optic disc detection. Statistical information obtained from these histograms are used as feature set in the classification stage. After the features are extracted, the data must be preprocessed using data normalization and data resampling before the classification stage. The true positive and the true negative value are calculated for the statistical analysis. Cross validation of the data is carried out for the extended binary pattern algorithm to validate the data sets as in Fig. 3. VLSI implementation of extended binary pattern algorithm is carried out and the area, power, and device utilization of the algorithm are calculated and compared with the local binary pattern algorithm. Identification and classification of the retinal disease by the extended binary pattern algorithm based automatic system are carried out with less time and low power when compared with the local binary pattern algorithm as shown in Table 1.

Formula:

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$
$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

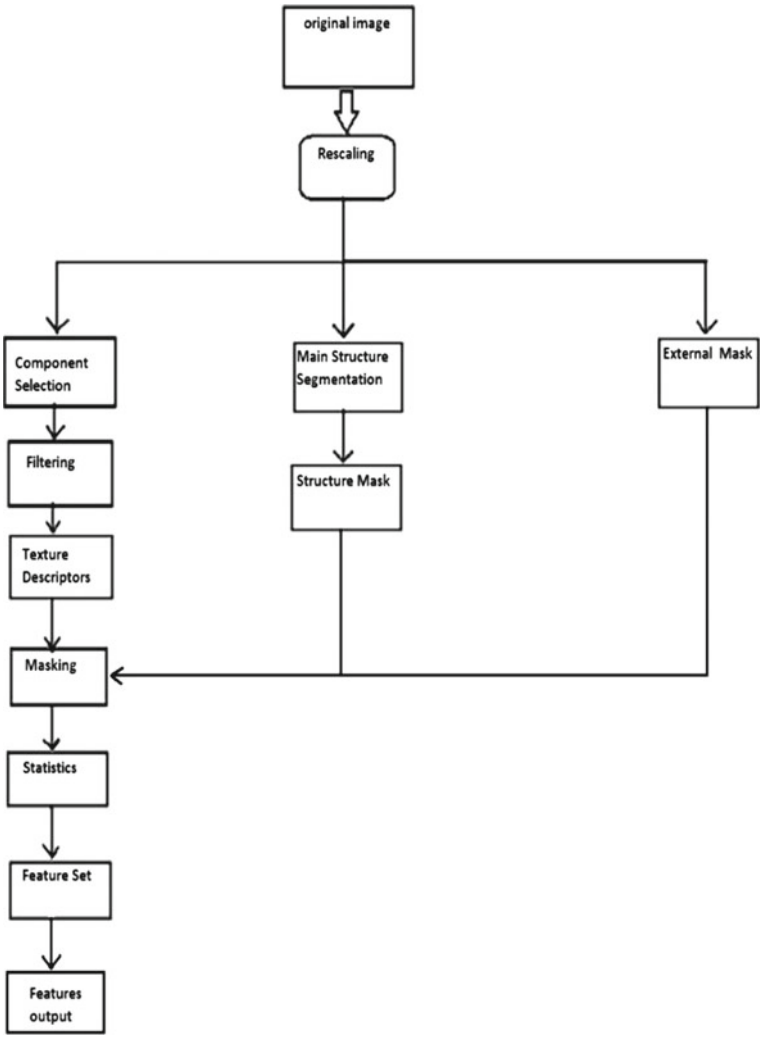


Fig. 2 Flow diagram of the proposed method

6 Result

Classification of the retinal disease was experimented with proposed procedure. The performances of the algorithms were measured using two parameters: true positive rate and true negative rate which in turn are used to calculate sensitivity and specificity of the retinal images. The values of the parameters are highlighted in tables. The performance of EBP was compared with LPB (Figs. 4, 5, 6, 7 and Tables 2, 3).

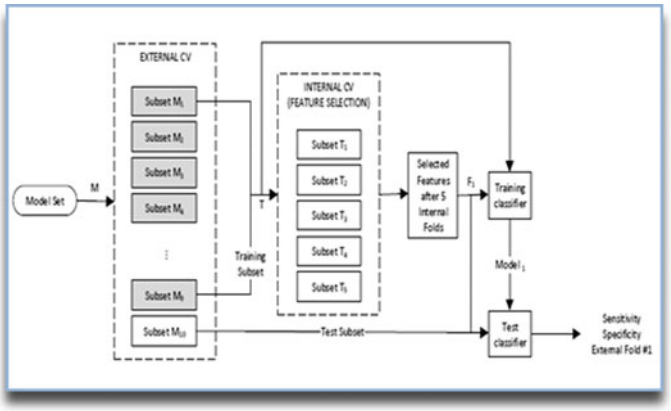


Fig. 3 Flowchart of cross validation

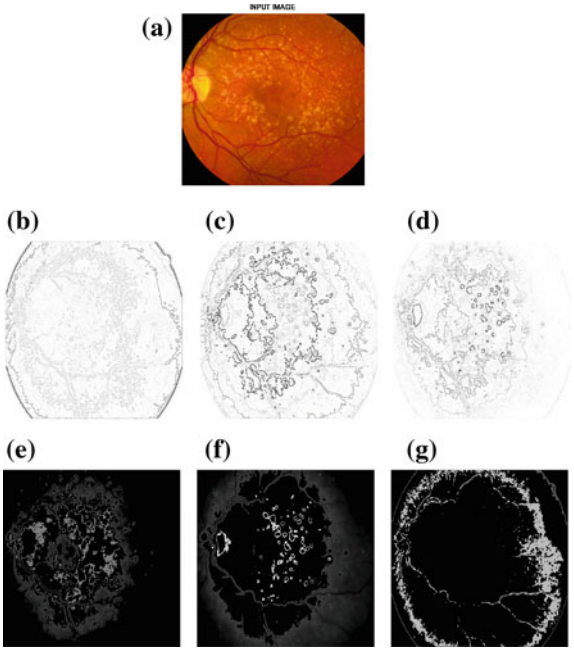


Fig. 4 Feature extraction using **a** AMD fundus image, **b**–**d** EBP images calculated on R, G and B components, respectively. Optic disc and vessel segments are removed. **e**–**g** VAR images calculated on R, G and B components, respectively. Optic disc and vessel segments are removed

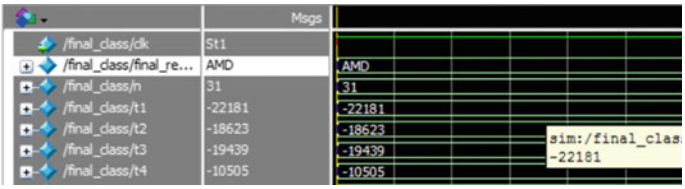


Fig. 5 Final identification of the eye disease

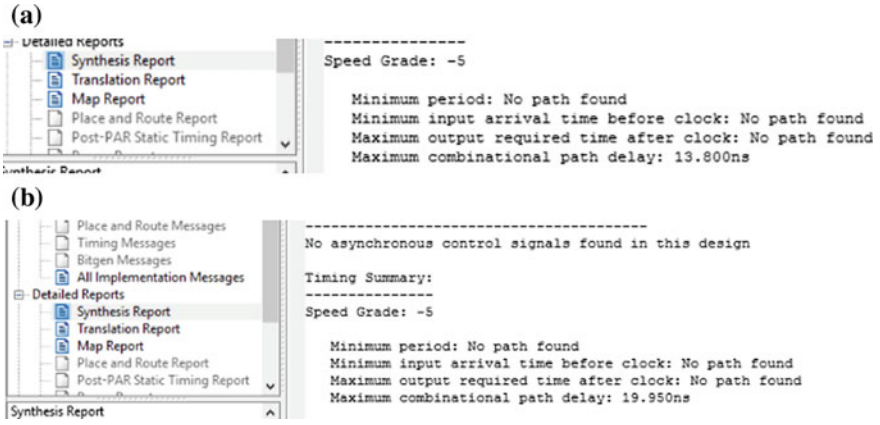


Fig. 6 Here **a** and **b** are time comparisons

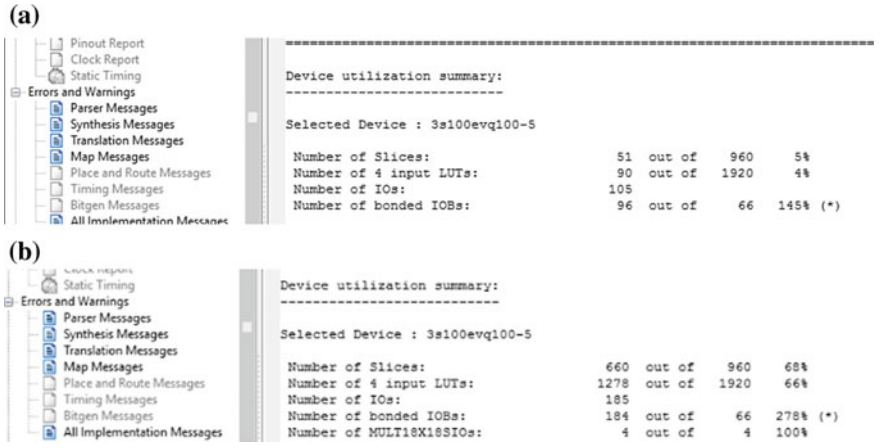


Fig. 7 Here **a** and **b** are comparison between device utilization

Table 2 Device utilization between LBP and EBP

Parameter	LBP	EBP
Number of Slices	660	51
Number of 4 input LUTs	1278	90
Number of IOs	185	105
Number of bonded IOBs	184	96

Table 3 Time summary and power consumption between LBP and EBP

Parameter	LBP	EBP
Maximum combinational path delay	19.950 ns	13.800 ns
Power consumption	0.707 W	0.582 W

7 Conclusion

In this paper, a new approach for retinal disease identification and classification was proposed. To differentiate disease-free patients from visually impaired patients, texture discrimination was used. The performance of EBP was compared with LBP. The obtained results of EBP provide valuable features for eye disease screening and reduction in the size and time consumption when compared with LBP.

References

1. Morales S, Engan K, Naranjo V, Colomer A (2015) Retinal disease screening through local binary patterns
2. Ojala T, Pietikinen M, Menp T (2001) A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. In: 2nd international conference on advances in pattern recognition, pp 397–406
3. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 28(12)
4. Kotu L, Engan K, Eftestol T, Woie L, Orn S, Katsaggelos A (2012) Local binary patterns used on cardiac MRI to classify high and low risk patient groups. In: Proceedings of the 20th European signal processing conference (EUSIPCO), pp 2586–2590
5. Oppedal K, Engan K, Aarsland D, Beyer M, Tysnes OB, Eftestol T (2012) Using local binary pattern to classify dementia in MRI. In: 9th IEEE international symposium on biomedical imaging (ISBI), May 2012, pp 594–597
6. Mookiah M, Acharya UR, Martis RJ, Chua CK, Lim C, Ng E, Laude A (2013) Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: a hybrid feature extraction approach. Knowl-Based Syst 39:9–22
7. Garnier M, Hurtut T, Ben Tahar H, Cheriet F (2014) Automatic multiresolution age-related macular degeneration detection from fundus images. In: Proceedings of SPIE, vol 9035, pp 903532–903532–7
8. Guo Z, Zhang D, Zhang D (2010) A completed modeling of local binary pattern operator for texture classification. IEEE Trans Image Process 19(6):1657–1663
9. Maenpaa T, Pietikainen M (2003) Multi-scale binary patterns for texture analysis. In: Bigun J, Gustavsson T (eds) Image analysis. Lecture notes in computer science, vol 2749, pp 885–892

10. Liao S, Law M, Chung A (2009) Dominant local binary patterns for texture classification. IEEE Trans Image Process 18(5):1107–1118
11. Song T, Li H, Meng F, Wu Q, Luo B, Zeng B, Gabbouj M (2014) Noise-robust texture description using local contrast patterns via global measures. IEEE Signal Process Lett 21(1):93–96
12. World Health Organization (WHO) (2013) Universal eye health: a global action plan 2014–2019
13. World Health Organization (WHO) (2010) Action plan for the prevention of avoidable blindness and visual impairment 2009–2013

A Study on Smart Electronics Voting Machine Using Face Recognition and Aadhar Verification with IOT

Kone Srikrishnaswetha, Sandeep Kumar and Md. Rashid Mahmood

Abstract Voting is an important process in which people can choose their own leader for the government. The device which we use for the voting process is an electronic voting machine with highly secured steps such as having unique number aadhar card, biometric of face recognition with IOT. This had a secured database saved system. Voting was an important role in our democratic system as we have the right to select our leader for the government.

Keywords Smart voting machine · IOT · Embedded systems · Security systems · Biometric of the face · Aadhar card

1 Introduction

Voting for a leader is the most important process which carries the essential result of the opinion of the people in selecting their leader for the government [1]. Of course, the voting machine is a mechanical device but more commonly designed by an electronic substance so it is an electronic voting machine. The process of voting is proposed that was highly secured and a technology of face recognition. It was designed for different uses for long usage, for security purpose, for high efficiency and accuracy of votes and many. This may be useful in more or less to all voters. This electronic machine is also called as e-voting, which mean it can do two different

K. Srikrishnaswetha · S. Kumar
Department of Electronics and Communications,
Sreyas Institution of Engineering and Technology, Hyderabad, India
e-mail: Srikrishnaswetha24@gmail.com

S. Kumar
e-mail: drsandeep@sreyas.ac

Md. Rashid Mahmood (✉)
Department of Electronics and Communications, Guru Nanak Institutions Technical Campus, Hyderabad, India
e-mail: er.mrashid@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_10

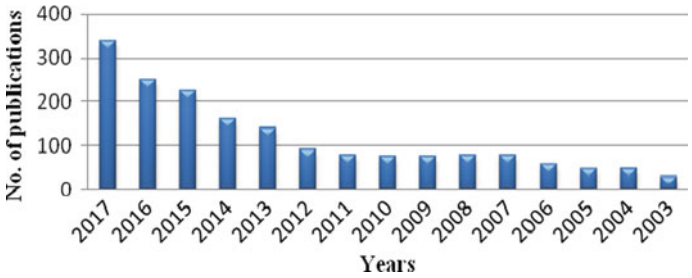


Fig. 1 Smart voting machine related field for the last 15 years without IOT

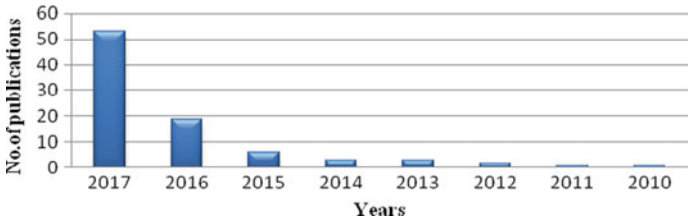


Fig. 2 Smart voting machine related field for the last 8 years with IOT

types of voting such as counting a number of votes and casting which is meant that selecting of their right leader [2, 3]. Some of the examples of the electronic voting technologies are optical scan voting, punch card, and voting kiosks. It also includes some of the different types of networks such as the usage of the telephonic mobile network, private use of a computer network, and Internet social media network. As for how the e-voting system involves DRE voting system (Direct Recording Electronic voting system) [4]. Now in our world of different technology, many old methods of this voting system are hence being changed to match the present situations in other fields of life. The voting involves the secret listing of votes and secret voting boxes to these electronic machine methods. The research community work on the technology of the voting machine without IOT from the last 15 years was increasing in every year as shown in Fig. 1 and data were taken for this graph from the well-known site from Science Direct.

The research community work on the technology of the voting machine without IOT from the last 10 years was increasing in every year as shown in Fig. 2 and data were taken for this graph from the well-known site from Science Direct.

2 Literature Analysis

"A Novel Electronic Voting Machine Design with Voter Information Facility Using Microcontroller" was proposed by "D. Ashok, in 2011" in this paper and the votes are calculated by using the keys given and the result is shown on the LCD screen.

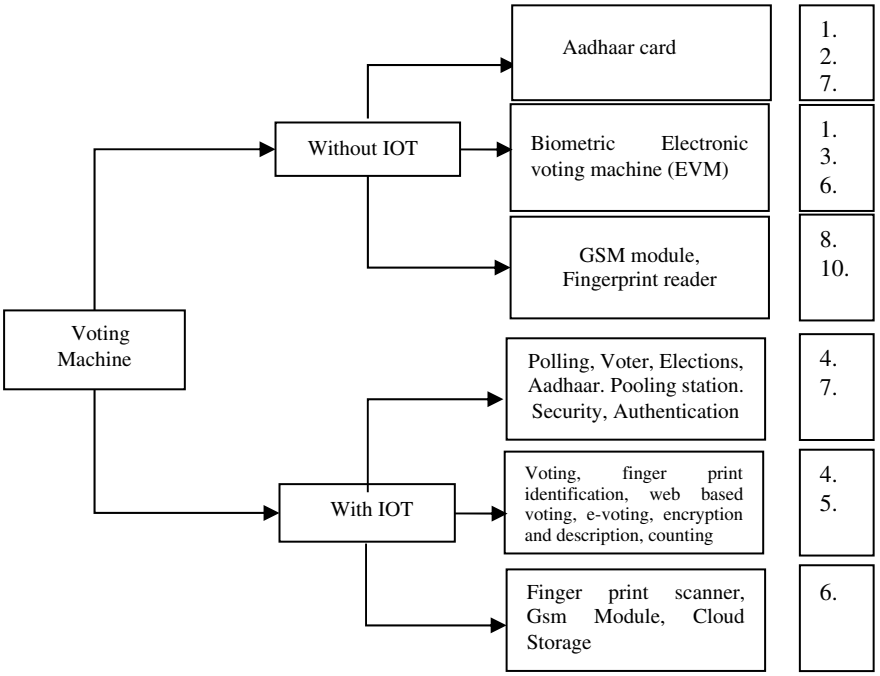


Fig. 3 Some of the techniques used the voting machine

The design of electronic voting machine explained in this paper is secured and accurate and it can be improved in the future for the power savings. The voting machine design explained in this system is accurate, clearly displays the message and highly secured [1].

"Development of Electronic Voting Machine with the Inclusion of Near Field Communication ID Cards and Biometric Fingerprint Identifier (Near Field Communication), fingerprinter Arduino" was proposed by "Syed Mahmud in 2014" as we are using NFC cards as the voter's ID card and it allows each one by one card to carry some data which can be used to link and identify the owner of a card. It is also simple technology and quick technologies that can identify the correct owner. The advantage of this technology is by using fingerprint so they cannot cheat by duplicate voting [5]. "Electronic Voting Machine EVM system" was proposed by "Kumar in 2012" that has to be studied further and should reach all the levels of different communities, so that the confidence of the voter will be increased and official's election will make more involvement in using the secured EVMs for conducting faster, secure, accurate voting elections as shown in Fig. 3 [6].

"Online Voting System for India based on AADHAAR ID" was proposed by Agarwal in 2013 and it is more secure and highly efficient than the regular voting system. Invalid votes and time are taken in announcing the results can be removed easily. A common AADHAAR identity card is an important component used in this

Table 1 Comparison of the proposed techniques for voting machine

S. No.	Author name	Year	Title of paper	Methodology	Remarks
1	Kumar et al. [6]	2012	Electronic voting machine	EVM (electronic voting machine), biometric EVM	It was an automated system. By this, the vote will be accurate, democracy, privacy, and verifiability
2	Agarwal et al. [3]	2013	Online voting system for India based on AADHAAR ID	AADHAAR ID, electronic voting machine, the online election process	Identify the details of the voting person. It was the easy and correct verification of the voters as well as candidates
3	Kumar DA, Ummal Sariba Begum T [1]	2011	A novel design of an electronic voting system using a fingerprint	LCD, microcontroller, sensors	Votes are accurate and clear display by the LCD
4	Hasan et al. [5]	2014	Development of electronic voting machine with the inclusion of near field communication ID cards and biometric fingerprint identifier	EVM (electronic voting machine), NFC (near field communication)	Identification of a person was correct so that the same person cannot be repeated for voting. Secure voting will be done
5	Ali et al. [4]	2014	Microcontroller based smart electronic voting machine system	(EVM), direct recording electronic (DRE), control unit (CU), and smart electronic systems	Automatic count of votes. It reduces the work
6	Anandaraj et al. [7]	2015	Secured electronic voting machine using biometric	GSM module, fingerprint reader, Arm 7 cortex	GSM is used for sending and receiving data
7	Kalaiselvi, Fabio Trindade VK, de Carvalho JPA [8]	2017	Smart voting	EVM, Aadhar card, fingerprints, voter id	Transparent voting machine. Reduces manual work

Table 2 Comparison of proposed techniques for voting machine with IOT

S. No.	Author name	Year	Title of paper	Methodology	Remarks
1	Santhosh et al. [12]	2016	Electronic voting machine using internet	Voting, fingerprint identification, web-based voting, e-voting, encryption and description, counting	It is a method of voting through the Internet without going to the voting booth. It is fast to access, high insecurity. The online voting process does not create any error. Even soldiers from abroad can participate in elections
2	Nithya et al. [10]	2016	Advanced secure voting system with IoT	Fingerprint scanner, PIC 16F877A, GSM module, cloud storage	Due to this biometric process, the recognition of the person's fingerprint will be easier and the duplicate person can be identified
3	Gawhale et al. [9]	2017	IOT based E-voting system	Raspberry Pi B+ Model, Atmega8 microcontroller, LCD display, matrix keyboard, USB to TTL converter, router	In this, the system is used for long distance. It saves time, money, and effort to reach the polling booth. Display of result will be easy and quick. It gives high data security
4	Devgirikar SJ, Deshmukh SA, Paithankar VP, Bhawarkar NB [11]	2017	Advanced electronic voting machine using the internet of things (IoT)	Fingerprint, voting system, web technology, security	Here scanner is used to reduce or remove the unwanted human error. It is capable to handle multiple modules in various centers
5	Channakeshava et al. [13]	2017	Distributed voting system using IOT	Terminal, IoT, biometric, polling, voter, candidate, elections, Aadhaar. Pooling station. Security, authentication	It increases voting percentage. It reduces error of duplicate voting. Reduces manpower. Reduces cost

(continued)

Table 2 (continued)

S. No.	Author name	Year	Title of paper	Methodology	Remarks
6	Khan et al. [14]	2018	IoT security: Review, blockchain solutions, and open challenges	IoT security Blockchain IoT protocols network security data security	IOT security having mainly three issues depending on high level, low level, and intermediate level layers
7	Coulter, Pan et al. [15]	2018	Intelligent agents defending for an IoT world: a review	Intrusion detection, artificial intelligence, machine learning, internet of things, cybersecurity, data driven	Development of IOT requires flexible agents they should be adjusted, heal, and trust in actions

paper as we can identify the duplicate voting persons. It is used for the identifying and helps in the verification of both voters and candidates [3]. "Secured Electronic Voting Machine using Biometric" is highly secured voting machine and which reduces manpower efficiently makes the work completed on time. In this machine, they introduced some new topics and that is implemented by ARM processor. Due to adding the implementation of aadhar card system, it can be improved in the future by the adding of identifying of votes in the system for more secured polling [7]. "Smart voting" can identify the people trying to vote the second time is not being allowed as once the fingerprint is scanned, authentication the login is being locked for the user. The casted vote is being updated at each instance of time in the database. The election results can be published on the same day with high accuracy and efficiency [8].

"Microcontroller Based Smart Electronic Voting Machine System" is the fast method of conducting the elections and accurate counting. It reduces the strength of frauds on large scale. Because its code is not accessible and cannot be changed once it is scanned [4]. "Electronic Voting Machine Using Internet" is a method of voting through the Internet without going to the voting booth. It is fast to access, high insecurity. The online voting process does not create any error. Even soldiers from abroad can participate in elections [9]. "Advanced Secure Voting System with IoT" explained about the biometric process and recognition of the person will be easier and the duplicate person can be identified [10].

"IOT Based E-Voting System" is in this and the system is used for long distance. It saves time, money, and effort to reach the polling booth. Display of result will be easy and quick. It gives high data security [11]. "Advanced Electronic Voting Machine using the Internet of Things (IoT)" is here scanner and is used to reduce or remove the unwanted human error. It is capable to handle multiple modules in various centers [12]. "Distributed Voting System Using IOT" is in this and it increases voting

percentage. It reduces the errors of duplicate voting. It reduces manpower. It reduces cost [13] (Tables 1 and 2).

3 Proposed Methodology

According to our literature analysis, we have found many problems in last 15 years of voting process as shown in Fig. 4.

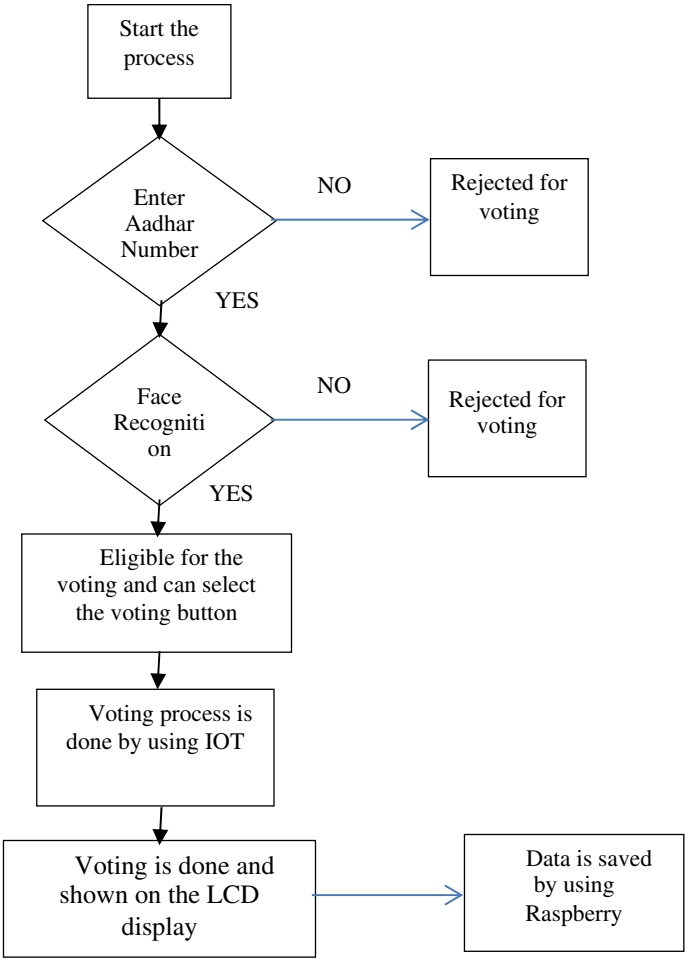


Fig. 4 The voting process of proposed methodology

1. In the process of the online voting system, a person can repeat the voting process and can create duplicate votes.
2. In the process of the secured electronic voting machine, we cannot send the data for longer distance and on time and cannot save the data.
3. In the process of the advanced secure voting system, the problem was through wireless communication and they may have Illegal voting due to this number of votes may increase.
4. In the process of EVM using the Internet, they will be a secured system but the people who are staying in villages cannot vote through the internet.
5. In the process of distributed voting of using IOT, it can be done using only one stage of election procedure at same polling stations.

They are a number of problems since few years but we giving a solution for few of the above problems with the new proposed methodology in Fig. 3. The above flowchart was explained by the following steps:

1. The process was started by entering the unique aadhar number.
2. If the unique number was wrong, then the person is rejected for voting. If the unique number was correct, the person can enter into the next step.
3. The person's face will be verified in this step and if he/she fails in this step, they are rejected for the further process.
4. If the face of the person matches, they are eligible for the voting of their leader and they can select the voting button.
5. After selecting the voting button, the data is saved automatically by using the Raspberry concept. So, that the same person cannot be eligible for voting for the second time.
6. Finally, on the LCD display, your voting is done will be displayed.

4 Conclusions

In India, the electronic voting machine was not secure at present days. They are many security problems. So, this design of the electronic voting machine was completely based on security purpose to show the original result to the people. This system is a small contribution to a fair election. The review proposed in this paper was voting machine with highly secured. We can secure the votes using some technologies by using IOT in the voting machine and it can develop a smart system of detecting an incorrect match of biometric at booth and biometric in aadhar card, it can trigger alertness and can communicate to the authorized person. A country with less voting percentage will struggle to develop their country by choosing the right leader. This is mainly due to the failure of the security level in the present voting system.

References

1. Kumar DA, Ummal Sariba Begum T (2011) A novel design of an electronic voting system using a fingerprint. Int J Innov Technol Creative Eng 1(1):12–19
2. Jacobi RP, Trindade F, de Carvalho JPA, Cantanhede R (2000) JPEG decoding in an electronic voting machine. In: IEEE 13th symposium on integrated circuits and systems design, pp 177–182
3. Agarwal H, Pandey GN (2013) Online voting system for India based on AADHAAR ID. In: 11th international conference on ICT and knowledge engineering (ICT&KE), pp 1–4
4. Ali SM, Mehmood CA, Khawja A, Nasim R, Jawad M, Usman S, Khan S, Salahuddin S, Ihsan MA (2014) Micro-controller based smart electronic voting machine system. In: IEEE international conference on electro/information technology (EIT), pp 438–442
5. Hasan SM, Anis AM, Rahman H, Alam JS, Nabil SI, Khalilur Rhaman M (2014) Development of electronic voting machine with the inclusion of near field communication ID cards and biometric fingerprint identifier. In IEEE 17th international conference on computer and information technology (ICCIT), pp 383–387
6. Kumar DA, Ummal Sariba Begum T (2012) Electronic voting machine—a review. In: IEEE international conference on pattern recognition, informatics and medical engineering (PRIME), pp 41–48
7. Anandaraj S, Anish R, Devakumar PV (2015) Secured electronic voting machine using biometric. In: IEEE international conference on innovations in information, embedded and communication systems (ICIIECS), pp 1–5
8. Kalaiselvi, Fabio Trindade VK, de Carvalho JPA, Cantanhede R (2017) Smart voting. In: Second international conference on computing and communications technologies (ICCCT'17), pp 1–10
9. Gawhale S, Mulik V, Patil P, Raut N (2017) IOT based E-voting system. Int J Res Appl Sci Eng Technol (IJRASET) 5(5)
10. Nithya S, Ashwin C, Karthikeyan C, Ajith Kumar M (2016) Advanced secure voting system with IoT. Int J Eng Comput Sci 5(3):33–37
11. Devgirikar SJ, Deshmukh SA, Paithankar VP, Bhawarkar NB (2017) Advanced electronic voting machine using internet of things (IOT). In: 2nd national conference recent innovations in science and engineering (NC-RISE 17), vol 5, no 9, pp 83–85
12. Santhosh M, Kavitha S, Keerthana R, Suganya L, Krishnakumar S (2016) Electronic voting machine using internet. Int J Commun Comput Technol 4(2)
13. Channakeshava RN (2017) Distributed voting system using IOT. Int J Innov Res Comput Commun Eng 5(2)
14. Khan MA, Salah K (2018) IoT security: review, blockchain solutions, and open challenges. Future Gener Comput Syst 82:395–411
15. Coulter R, Pan L (2018) Intelligent agents defending for an IoT world: a review. Comput Secur 73:439–458

Exhaustive Analysis of Image Enhancement Using Point-to-Point Transformation

Pratima Manhas, Shaveta Thakral and Parveen Arora

Abstract Image can be processed via variety of transformation techniques. Transformations of signals not only extract useful information from any real signal but also process it as per demanding application. Image transformation can be histogram equalization, discrete wavelet transform, morphology or point transformation. Point transformation just improves visual appearance of an image without fetching any new information. It is just required for detection or extraction of some information as per demanding application. Point transformation can be point-to-point transformation, local-to-point transformation, or global-to-point transformation. Transformation technique can be chosen based on complexity of transformation technique. The aim of this paper is to discuss point-to-point transformation technique which is advantageous in terms of constant complexity compared to local to point and global-to-point techniques as it is calculated per pixel and it is point to point. Different types of image enhancement techniques such as image negative, separation of image into RGB parts, gamma correction and log transformation are implemented using MATLAB.

Keywords Pixel · Point transformation · Image · Signal · Local · Global · Enhancement

P. Manhas (✉) · S. Thakral
ECE Department, FET, Manav Rachna International Institute of Research & Studies, Faridabad, India
e-mail: pratima.fet@mriu.edu.in; pratimamehak@gmail.com

S. Thakral
e-mail: shaveta.fet@mriu.edu.in

P. Arora
NEC Technologies India Pvt Ltd, Bengaluru, India
e-mail: parveen.arora@india.nec.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_11

1 Introduction

Transformations of signals not only extract useful information from any real signal but also process it as per demanding application. Natural signal in raw form is taken and processed in one form or other depending upon application. The transformed signal after processing can be used further to retrieve useful information. Image can be processed via various types of transformation techniques. It can be point-to-point transformation, global-to-point transformation, or global-to-point transformation.

In point-to-point transformation technique, the value of pixel chosen in a specific coordinate in processed image is only dependent on the value of pixel falling in the same coordinate in the input image. In local-to-point transformation, a window of pixels usually rectangular shape is chosen in input image and the value of one element in a specific coordinate of processed image depends on pixels taken within a window surrounding same specific coordinate in the input image. In global-to-point transformation, the value of the output at specific coordinates in processed image is dependent on all the pixels in the input image. The example of global-to-point transformation is Discrete Fourier Transform (DFT).

Transformation technique can be chosen based on the complexity of transformation technique. Point-to-point transformation technique is advantageous in terms of constant complexity as compared to local-to-point and global-to-point techniques as it is calculated per pixel and it is point to point [1]. In local-to-point transformation, complexity per pixel is dependent on window size. Complexity tends to be w^2 for window sixe of $w * w$. In global-to-point transformation, complexity per pixel is dependent on complete image size. Complexity tends to be I^2 for image size of $I * I$.

Point transformation just improves the visual appearance of an image without fetching any new information. It is just required for detection or extraction of some information as per demanding application. One of the applications of point transformation is to enhance the contrast of an image or another may be to convert grayscale image to colored image. Point transformation can improve visual appearance or make targets easier to detect/extract.

2 Point-to-Point Transformation Types

Image enhancement techniques majorly improve contrast and helpful in exhaustive analysis useful for wide range of applications [2]. Image enhancement proves to be useful in medical image processing, remote sensing, satellite communication, etc. Image enhancement works on intensity transformations to increase the contrast between certain intensity values so that image can be analyzed by various angles from deeper regions. Various types of image enhancement techniques are:

(1) Inversion/Photo Negative
(2) Gamma Correction
(3) Logarithmic Transformation

Inversion

It is one of the basic techniques which found enormous applications in medical field of digital image processing popular as medical image processing. Basically, negative of an image is created through inversion technique where values of gray pixels are inverted. One of the most useful applications of inversion can be seen with digital mammogram image. Digital mammogram image when transformed into its negative after applying Inversion transformation may be helpful to detect white patch which may be an indication of cancerous region.

Gamma Correction

Gamma correction is another image enhancement technique which works on the principle of controlling the overall brightness of an image by reproducing colors. Sometimes images look too dark or other time seems bleached out. Gamma parameter is varied in such a way that brightness or RGB ratio is controlled and image looks superior [3].

Technically every image can be seen as a group of pixels where each and every pixel represent intensity level. Usually, grayscale images have pixel values varied between 0 and 1. As pixel value changes from 0 to 1; brightness also varies means 0 indicating no brightness and 1 indicating full brightness. No brightness means darkness and can be seen as black color and full brightness means no darkness and can be seen as white color. Brightness level basically indicates luminance. Sometimes while capturing images, nonlinear luminance effect occur and sometimes displaying them also causes nonlinearity in luminance levels and pictures seen are very annoying. Hence, Gamma correction technique basically compensates the nonlinear luminance effects to enhance display.

It is given as

$$\text{Gamma Correction Function } [G]: = G^{\wedge}\gamma$$

where γ is a constant and "^" is called power operator. The value γ is called as gamma.

$$\text{gamma correction} = 1/\text{gamma}$$

Gamma correction function works luminance of image. The luminance value lies between 0 and 1. A default gamma value of 2.2 is usually popular and gives good results on most monitors [4].

Logarithmic Transformation

The log transformation is given as:

$$s = c \log(1 + r)$$

where c is a constant and it is assumed that $r \geq 0$. Log transformation is working on mapping principle where a narrow range of low-level grayscale intensities are mapped into a wider range of output values and wide range of high-level grayscale intensities are mapped into a narrow range of high-level output values. Inverse-log transform works in an opposite manner as compared to logarithmic transformation. This transform is used to expand values of dark pixels and compress values of bright pixels.

3 Matlab Code and Results

Following are the Matlab codes to visualize the image into three RGB planes, negative/complement function, gamma correction, and logarithm transformation.

3.1 Matlab Code to Visualize Three RGB Planes and Image Negative

```
J=imread('download.jpg');
I=rgb2gray(J);
imshow(I);title('Original
Image'); vector=0:1/255:1;
r=vector;
g=vector*0;
b=vector*0; lutr=[r' g' b'];
r=vector*0; g=vector;
b=vector*0; lutg=[r' g' b'];
r=vector*0; g=vector*0;
b=vector; lutb=[r' g' b']; figure;
image(I(:,:,1));
colormap(lutr);title('Red
Plane'); figure;
image(I(:,:,1));
colormap(lutg);title('Green
Plane'); figure;
image(I(:,:,1));
colormap(lutb);title('Blue Plane');
k=imcomplement(I);
figure;imshow(k);title('Negative Image');
```

The R, G, and B indicates the red, green, and blue color plane. I have taken my own image for performing various operations. It is shown in Figs. 1, 2, 3, 4, and 5.

Fig. 1 Original image

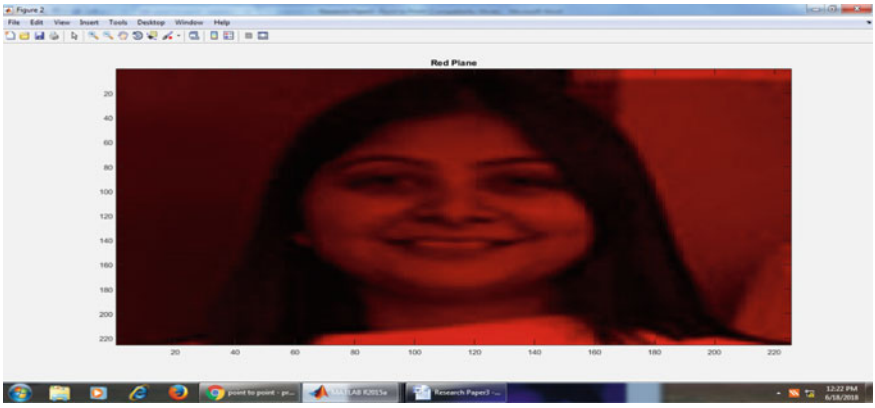


Fig. 2 R Plane of image

3.2 Matlab Code for Gamma Correction

Matlab code for gamma correction is given below:

```
p = imread('download.jpg'); I
= rgb2gray(p);
J = imadjust(I,[],[],0.5);
imshow(I)
figure, imshow(J); title('gamma correction');
```

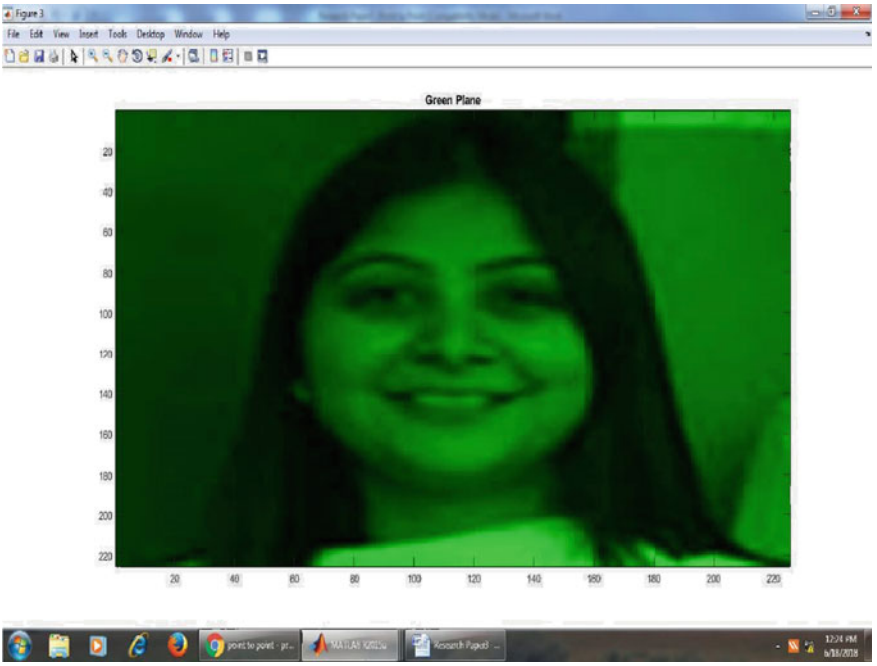


Fig. 3 G Plane of image

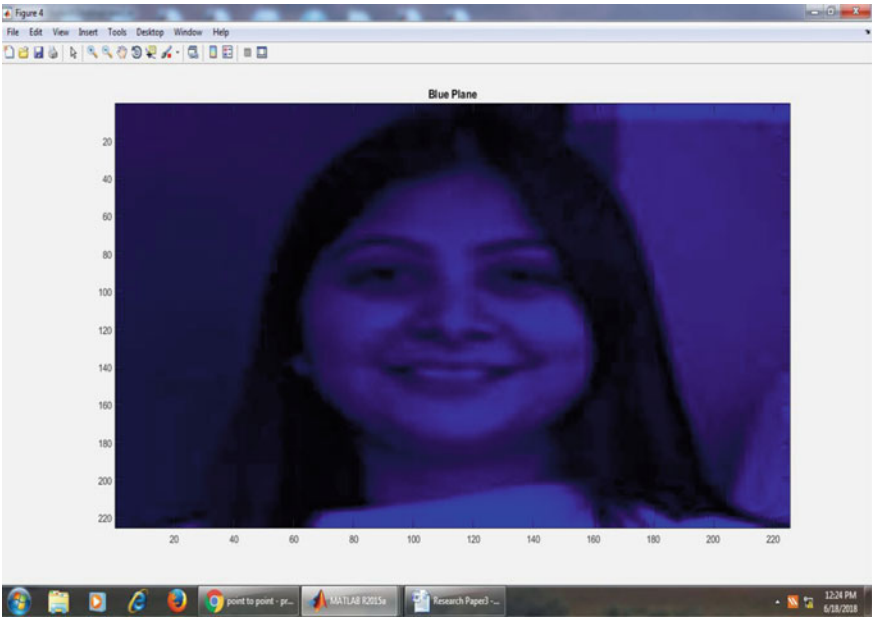


Fig. 4 B Plane of image

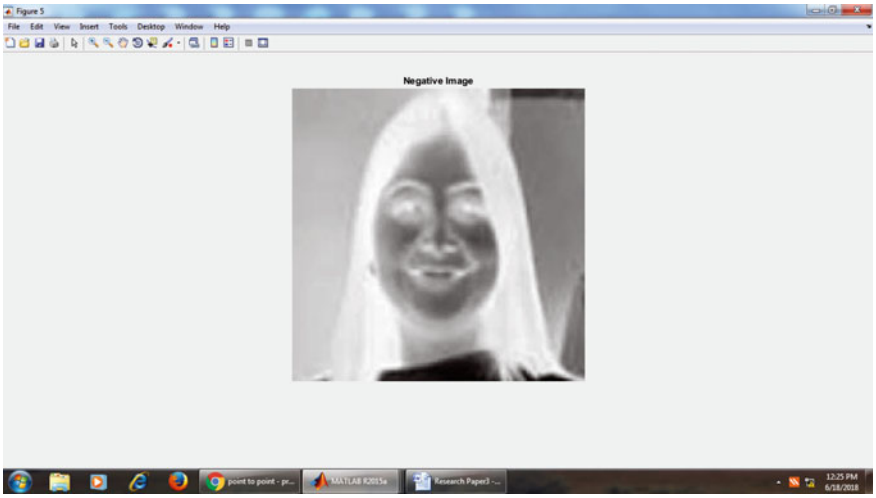


Fig. 5 Negative of image

3.3 Matlab Code for Logarithm Transformation

The Matlab code for gamma correction is given below:

```
a1 = imread('log.jpg'); % Read the image
a = double(a1)/255; % Normalized Image
c = 2; % Constant
f = c*log(1+(a)); % Log Transform
subplot(1,2,1),imshow(a1),title('Original Image');
subplot(1,2,2),imshow((f)),title('Log Transformation Image');
```

4 Conclusion

The different types of transformation techniques are applied on the image. The results show the image enhancement by using different techniques. Image negative is used to convert the darker region of the image into brighter region which will identify the image more clearly. So, it is used in medical image processing. Similarly, gamma correction and logarithm transformation also lead to enhance the image brightness. The work can be further modified by using various frequency transformation method and arithmetic operation on image which helps in reducing the noise (Figs. 6 and 7).

Fig. 6 Gamma correction of image

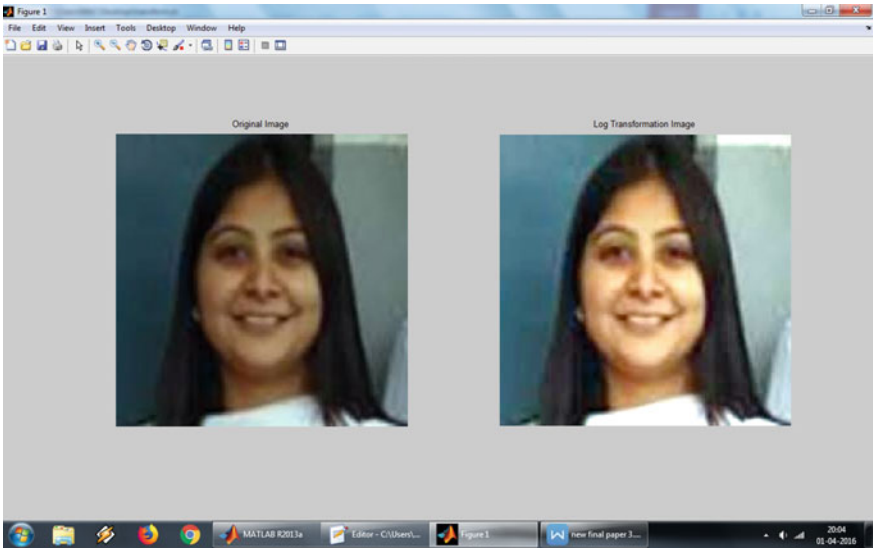


Fig. 7 Log transformation output

References

1. Beucher S, Meyer F (1992) The morphological approach to segmentation: the watershed transformation. In: Optical engineering, vol 34. Marcel Dekker Incorporated, New York, pp 433–433
2. Chang Y, Saito S, Nakajima M (2007) Example-based color transformation of image and video using basic color categories. IEEE Trans Image Process 16(2):329–336
3. Manhas P, Thakral S (2018) Comparative analysis of different wavelet filters image enhancement using histogram equalization. In: International conference on materials, applied physics and engineering (ICMAE)
4. Lee MC, Chen WG (1999) Microsoft Corp.: image compression and affine transformation for image motion compensation. U.S. Patent 5,970,173
5. Prucnal PR, Saleh BE (1981) Transformation of image-signal-dependent noise into image-signal-independent noise. Opt Lett 6(7):316–318
6. Zhang, X, Wandell BA (1997) A spatial extension of CIELAB for digital color-image reproduction. J Soc Inf Disp 5(1):61–63
7. Umeyama S (1991) Least-squares estimation of transformation parameters between two point patterns. IEEE Trans Pattern Anal Mach Intell 13(4):376–380
8. Qi D, Zou J, Han X (2000) A new class of scrambling transformation and its application in the image information covering. Sci China Ser E Technol Sci 43(3):304–312
9. Lin H, Du P, Zhao W, Zhang L, Sun H (2010) Image registration based on corner detection and affine transformation. In: 2010 3rd international congress on image and signal processing (CISP), vol 5. IEEE, pp 2184–2188
10. Beier T, Neely S (1992) Feature-based image metamorphosis. In: ACM SIGGRAPH computer graphics, vol 26, no 2. ACM, pp 35–42

Quantized Local Trio Patterns for Multimedia Image Retrieval System

P. Rohini and C. Shoba Bindu

Abstract In this paper, we propose a novel feature descriptor named quantized local trio pattern (QLTP) for multimedia image retrieval application. The QLTP extracts quantized edge information from the pixels in a specified neighborhood. QLTP integrates the quantization and trio patterns for image retrieval. Performance of the QLTP is evaluated by conducting experiments on Corel-10,000 databases. Experimental results exhibit an improvement in terms of avg. retrieval precision (ARP) and avg. retrieval rate (ARR) as compared to the other related methods.

Keywords Retrieval · Local binary patterns · Avg. retrieval precision · Avg. retrieval rate · Database

1 Introduction

Rapid advancements in the Internet and some allied technologies have made the way for the generation of number of images every moment by medical, academic, scientific, industrial, and many more applications. Huge databases make the user's job tedious to browse and manage. Content-based image retrieval describes the image according to the explored parameters. It extracts the features from visual part of the image such as texture, shape, and color. Contributions made for the improvement in the efficacy of retrieval systems are given in [1–4]. Texture is one of the primary features used in CBIR. Researchers proposed many retrieval methods by using the texture as the primary element [5–7]. Local binary pattern (LBP) [8] became the foundation for a few retrieval algorithms. Takala [9] introduced Block-based LBP.

P. Rohini (✉)
Department of CSE, Faculty of Science and Tech., IFHE, Hyderabad, India
e-mail: rohini10@ifheindia.org

P. Rohini
Department of CSE, JNTUA, Anantpuramu, India

C. Shoba Bindu
Department of CSE, JNTUACEA, Anantpuramu, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_12

Yao and Chen proposed LEPSEG, LEPINV for texture description [10]. Concepts of local trio patterns [11] and quantized patterns [12] motivated us to propose a new feature descriptor named quantized local trio patterns binary pattern (QLTP) for image retrieval. Major contributions of the proposed work are as follows. (1) The new feature descriptor (QLTP) integrates the quantization and local trio patterns for image retrieval. (2) The proposed method collects the texture information by extracting the relationship between pixels. (3) Retrieval performance of the proposed descriptor is tested on benchmark Corel-5k database. The retrieval results show significant improvement as compared to the state-of-the-art feature for image retrieval. Organization of the paper is as follows: Sect. 1 provides an introduction of CBIR. Section 2 presents a review of local patterns. The proposed descriptor is given in Sect. 3. Experimental results using different similarity measures are discussed in Sect. 4. Section 5 covers the summery of the work.

2 A Review of Local Patterns

2.1 Local Binary Patterns

LBP depicts the connectivity between the pixels in a specified neighborhood. A 3×3 window is considered to derive the pattern. When the intensity of the center pixel is equal to or more than that of a neighboring pixel, a binary 1 is coded, else 0 is coded as mentioned in the equations below.

$$\mathrm{LBP}_{Y,Z} = \sum_{p=1}^{Y} 2^{(p-1)} \times f(M(x_p) - M(x_c)) \tag{1}$$

$$f(y) = \begin{cases} 0 \text{ otherwise} \\ 1 \ y \geq 0 \end{cases} \tag{2}$$

$M(x_c)$ and $M(x_p)$ represent the intensity values of central pixel and neighbor, respectively. The number of neighbors is represented by Y and the radius is shown by Z.

2.2 Local Extreme Sign Trio Pattern

Vipparthy [12] proposed the local extreme sign trio patterns for an image retrieval system. LESTP extracts the extreme sign information from an image using trio values. First extreme edge is obtained by considering sign of local difference between the reference pixel and its eight surrounding neighbors at radius one. A detailed presentation on LESTP is available in [12].

3 Proposed Method

Motivated by Local trio patterns, we propose a novel feature descriptor by applying the quantization to the image as specified in Eqs. (3)–(6). Trio patterns are obtained according to Eq. (7).

$$\mathrm{DQE}(I(p_c))|_{0^\circ} = \{g(I(p_1), I(p_4), I(p_C)); g(I(p_2), I(p_5), I(p_C)); g(I(p_3), I(p_6), I(p_C))\} \quad (3)$$

$$\mathrm{DQE}(I(p_c))|_{45^\circ} = \{g(I(p_{13}), I(p_{16}), I(p_c)); g(I(p_{14}), I(p_{17}), I(p_c)); g(I(p_{15}), I(p_{18}), I(p_c))\} \quad (4)$$

$$\mathrm{DQE}(I(p_c))|_{90^\circ} = \{g(I(p_7), I(p_{10}), I(p_C)); f_2(I(p_8), I(p_{11}), I(p_C)); f_2(I(p_9), I(p_{12}), I(p_C))\} \quad (5)$$

$$\mathrm{DQE}(I(p_c))|_{135^\circ} = \{g(I(p_{19}), I(p_{22}), I(p_C)); g(I(p_{20}), I(p_{23}), I(p_C)); g(I(p_{21}), I(p_{24}), I(p_C))\} \quad (6)$$

where

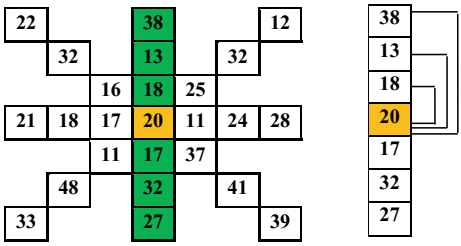
$$g(t1, t2, p) = \begin{cases} 1 & \text{if } (p >= t1) \\ -1 & \text{if } (p1 <= t2) \\ 0 & \text{if } (t1 > p < t2) \end{cases} \quad (7)$$

The process of quantization is depicted in Fig. 1. Calculation of QLTP is shown in Fig. 2 for the pixel at the center. Local difference between the center pixel and the neighbor pixels on both the sides is "18, 12, 7, −7, −3, −2" which are arranged n descending order according to the magnitudes. The first value refers to extreme edge

Fig. 1 A 7 × 7 image and the quantized structure

22	46	41	38	42	43	12
44	32	45	13	46	32	47
47	33	16	17	25	30	46
21	18	17	20	11	24	28
34	39	11	14	37	54	59
45	48	37	32	42	41	38
33	60	61	27	62	63	39

22			38			12
	32		13		32	
		16	18	25		
21	18	17	20	11	24	28
		11	17	37		
	48		32		41	
33			27			39



18	12	7	-7	-3	-2
1	1	1	-1	-1	0
1	1	1	0	0	0
0	0	0	1	1	0

Threshold (t1) =3, Threshold (t2) =-3

Fig. 2 Calculation of the proposed pattern from the quantized structure

of higher pixel. These edge values are converted into trio, $(-1, 0, 1)$, values using two threshold limits $\tau 1$ and $\tau 2$. In this paper, we consider thresholds as $\tau 1 = -3$ and $\tau 2 = 3$. According to (7), the conversion QLTP is represented as 1, 1, 1, −1, −1, 0. Similarly, the procedure is followed in case of all neighboring pixels. Finally, the QLTP is converted into one upper and one lower pattern. This pattern shows a significant improvement than other state-of-the-art techniques. QLTP extracts edge information as compared to the popular LBP.

In our proposed method, the given image is quantized to get HVDA structure as shown in Fig. 1. The HVDA reflects the stretch of pixels in horizontal, vertical, diagonal, and anti-diagonal directions.

Upon quantizing the image in different directions, the trio patterns are extracted in all the four directions of the quantized structure. For instance, extraction of trio patterns for the vertical strip of pixels is shown in Fig. 2

3.1 Algorithm

Input: Image Output: Retrieval result

(1) Convert the color image to grayscale image
(2) Extract the quantized patterns for the grayscale image
(3) Collect the trio patterns by setting any two threshold values
(4) Derive the ULESTP and LLESTP from the strips of pixels
(5) Construct a histogram to create the final feature vector

(6) Compare query image with database images
(7) Retrieve the images according to the best matches.

4 Experimental Results and Discussions

The efficiency of the proposed method is determined by means of the experiments on standard databases. During the experimentation, each database image becomes a query image. Accordingly, the proposed work identifies n images $Y = (y_1, y_2, \ldots, y_n)$ according to the shortest distance. Performance is evaluated in terms of avg. precision (ARP), avg. recall (ARR) as indicated below: Evaluation metrics for a query image Im_q are calculated according to the equation given below.

$$\text{Pre: } P(\mathrm{Im}_q) = \frac{\text{No. of related images retrieved}}{\text{Total images retrieved}} \tag{8}$$

$$\text{Avg. Retrieval Prec: ARP} = \frac{1}{|\mathrm{DBS}|}\sum_{k=1}^{|\mathrm{DBS}|} \mathrm{Pr}(\mathrm{Im}_k)\Bigg| \tag{9}$$

$$\text{Recall: Re}(\mathrm{Im}_q) = \frac{\text{No. of related images retrieved}}{\text{Total relevant images in database}} \tag{10}$$

$$\text{Avg. Retreival Rate: ARR} = \frac{1}{|\mathrm{DBS}|}\sum_{k=1}^{|\mathrm{DBS}|} \mathrm{Re}(I_k)\Bigg| \tag{11}$$

4.1 Corel-10k

Corel-10k [13] repository is an agglomeration of 10,000 images classified into different categories. Performance of QLTP is determined by means of precision, recall, ARR, and ARP. Retrieval results in terms of precision and recall are given in Table 1.

Comparisons of the results are shown in Fig. 3; The improvement in performance is mainly due to the discriminating ability of the proposed method against any other method mentioned.

Table 1 Precision and recall values of various methods (Corel-10k)

Image repository	Evaluation metric	CS_ LBP	LEP SEG	LEP INV	BLK_LBP	DLEP	LECTP	QLTP
Corel-10k	%Prec	22.1	24.3	24.8	25.5	26.4	28.3	30.1
	%Recall	11.2	12.4	13.4	14.1	14.9	16.6	18.2

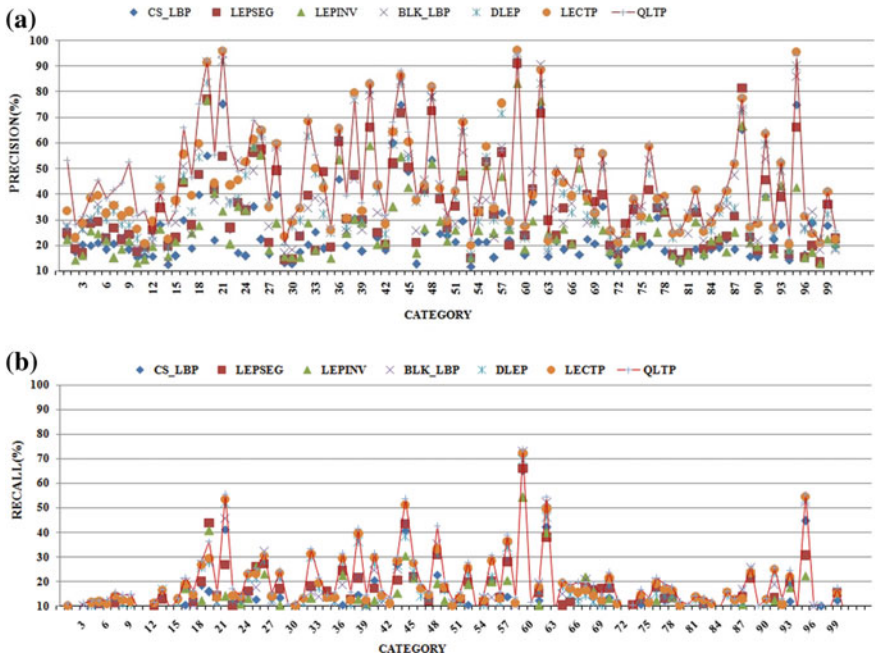


Fig. 3 Comparison of various methods in terms of a ARP b ARR

5 Conclusions and Future Scope

A new method for image retrieval is introduced in this paper. It explores the quantized trio patterns for image indexing and retrieval. Corel-10k repository is used to test the effectiveness of the method. Results after investigation exhibit a substantial improvement in the performance metrics such as precision and recall.

By extracting the facial features of a given image, the proposed approach can be extended in the field of face recognition.

References

1. Jégou H, Perronnin F, Douze M, Sanchez J, Perez P, Schmid C (2012) Aggregating local image descriptors into compact codes. IEEE Trans Pattern Anal Mach Intell 34(9):1704–1716
2. Tsikrika T, Popescu A, Kludas J (2011) Overview of the Wikipedia image retrieval task at ImageCLEF 2011. In: The working notes for the CLEF 2011 labs and workshop, Amsterdam, The Netherlands, 19–22 Sept 2011
3. Rahimi, M, Moghaddam E (2015) A content based image retrieval system based on color ton distribution descriptors. SIViP 9:691. https://doi.org/10.1007/s11760-013-0506-6
4. Simou N, Athanasiadis T, Stoilos G et al (2008) Image indexing and retrieval using expressive fuzzy description logics. SIViP 2:321. https://doi.org/10.1007/s11760-008-0084-1

5. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
6. Everingham M, Eslami SMA, Van Gool L, Williams CKI, Winn J, Zisserman A (2015) The PASCAL visual object classes challenge: a retrospective. Int J Comput Vis 111(1):98–136
7. Kokare M, Biswas PK, Chatterji BN (2007) Texture image retrieval using rotated wavelet filters. J Pattern Recogn Lett 28:1240–1249
8. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
9. Takala V, Ahonen T, Pietikainen M (2005) Block-based methods for image retrieval using local binary patterns. In: SCIA 2005, LNCS, vol 3450, pp 882–891
10. Yao C-H, Chen S-Y (2003) Retrieval of translated, rotated and scaled color textures. Pattern Recogn 36:913–929
11. Vipparthi SK, Nagar SK (2016) Local extreme complete trio pattern for multimedia image retrieval system. Int J Autom Comput 13:457. https://doi.org/10.1007/s11633-016-0978-2
12. Koteswara Rao L, Venkata Rao D (2015) Local quantized extrema patterns for content-based natural and texture image retrieval. Hum Cent Comput Inf Sci 5:26. https://doi.org/10.1186/s13673-015-0044-z
13. Corel 1000 and Corel 10000 image database [Online]. Available: http://wang.ist.psu.edu/docs/related.shtml

Multiple Color Channel Local Extrema Patterns for Image Retrieval

L. Koteswara Rao, P. Rohini and L. Pratap Reddy

Abstract A novel feature descriptor, multiple color channel local extremal pattern (MCLEP) is proposed in this manuscript. MCLEP combines the key features of local binary and local quantized extrema patterns in a specified neighborhood. Multi-distance information computed by the MCLEP aids in robust extraction of the texture arrangement. Further, MCLEP features are extracted for each color channel of an image. Retrieval performance of the MDLP is evaluated on benchmark datasets for CBIR, namely Corel-5000, Corel-10000, and MIT-Color Vistex, respectively. Proposed technique exhibits substantial improvement as compared to other recent feature descriptors in terms of ARP and ARR on standard databases.

Keywords Color · Texture · LBP, LQEP, and retrieval

1 Introduction

Rapid advancements in the Internet and the allied technologies have paved the way for the generation of large number of images every moment by medical, academic, scientific, industrial, and other applications. It becomes burdensome for a user to browse through the large databases. Content-based image retrieval describes an image according to some explored features. It resolves the key issues in traditional annotation method by extracting the features from visual data of the image such as texture,

L. Koteswara Rao (✉)
Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation (A Deemed to-be University), Off-campus Centre, Hyderabad, India
e-mail: koteswararao@klh.edu.in

P. Rohini
Department of Computer Science and Engineering, Faculty of Science &Technology, ICFAI Foundation for Higher Education (A Deemed-to be University), Hyderabad, India

L. Pratap Reddy
Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University Hyderabad College of Engineering, Hyderabad, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_13

shape, and color. Some contributions made for the improvement in the efficacy of retrieval systems are discussed in [1–4]. Texture is one among the primary features used in the automated retrieval system such as CBIR. Many researchers proposed several retrieval methods by using the texture as the key element [5–8].

A feature vector, local binary pattern (LBP) [9] became the foundation for a few retrieval algorithms. Takala [10] introduced Block-based LBP. Yao and Chen proposed LEPSEG, LEPINV for texture description [9]. S Murala et al. introduced a novel frame work, DLEP for retrieval [10]. Koteswara Rao et al. [11] introduced LQMeP for image retrieval. Koteswara Rao et al. proposed local color Oppugnant quantized patterns [12]. Koteswara Rao and Venkata Rao have proposed the local quantized Extrema patterns [13].

In the present work, the texture from the image is derived by using a new feature descriptor. Performance is evaluated on Corel image repository.

Contents are arranged as follows: Sect. 1 covers a review of retrieval approaches. Two primary ways of extracting the information are given in Sect. 2. Section 3 provides a detailed framework of the proposed method explains the algorithm of the proposed MCLEP method. In the next section, a summary of results is given. Conclusions and future scope are provided in Sect. 5.

2 Relevant Work

2.1 Local Binary Patterns (LBP)

LBP [9] explores the connectivity between a pixel at the centre and its immediate neighbors from a 3 × 3 window. When the intensity of the centre pixel is equal to or more than that of a neighboring pixel, a binary 1 is coded, else 0 is coded (see in Eqs. 1 and 2).

$$\mathrm{LBP}_{Y,Z} = \sum_{p=1}^{Y} 2^{(p-1)} \times f(M(x_p) - M(x_c)) \tag{1}$$

$$f(y) = \begin{cases} 0 \text{ otherwise} \\ 1 \quad y \geq 0 \end{cases} \tag{2}$$

$M(x_c)$ and $M(x_p)$ represents intensity values of central pixel and neighbor respectively, the no. of neighbors are represented by Y and radius is shown by Z.

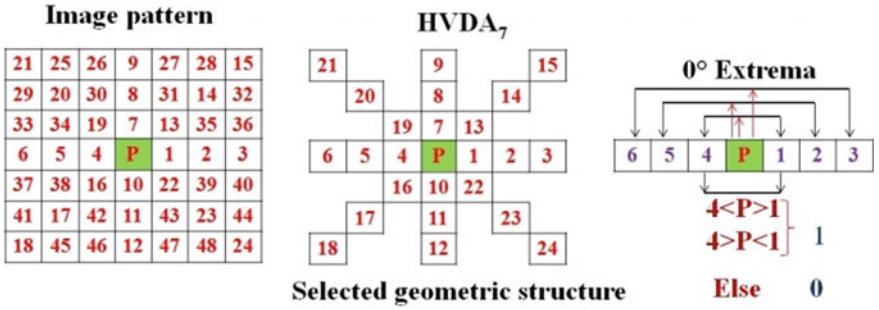


Fig. 1 LQEP calculation for a given 7 × 7 pattern

2.2 Local Quantized Extrema Patterns (LQEPs)

By applying the quantization to DLEP, Koteswara Rao [13] proposed a feature descriptor named LQEP. In the first step, geometric structures are separated from an image in specific directions. Extremal operation is performed on the derived structures. Figure 3 shows the calculation of the pattern described in [13]. For instance, computation of pattern in horizontal direction is done using the equation. Then, the four directional extremas (DE) in 0°, 45°, 90°, and 135° directions are obtained as follows (Fig. 1).

$$\mathrm{DE}(I(g_c))|_{0^\circ} = \left\{ y_2((g_1), I(g_4), I(g_p)); y_2(I(g_2), I(g_5), I(g_p)); y_2(I(g_3), I(g_6), I(g_p)) \right\} \quad (3)$$

here,

$$y_2(j, k, p) = \begin{cases} 1 \text{ if } (j > p) \text{ or } (k > p) \\ 1 \text{ if } (j < p) \text{ or } (k < p) \\ 0 \text{ otherwise} \end{cases} \quad (4)$$

LQEP is computed according to Eq. (5).

$$\mathrm{LQEP} = \left[\mathrm{DE}(I(g_c))|_{0^\circ}, \mathrm{DE}(I(g_c))|_{45^\circ}, \mathrm{DE}(I(g_c))|_{90^\circ}, \mathrm{DE}(I(g_c))|_{135^\circ} \right] \quad (5)$$

At the end, image is converted to LQEP maps with range of pixel intensities from 0 to 4095. A histogram is built using Eq. (6).

$$\mathrm{Hstg}_{\mathrm{LQEP}}(s) = \sum_{b=1}^{Q_1} \sum_{c=1}^{Q_2} y_2(\mathrm{LQEP}(b, c), s); s \in [0, 4095]; \quad (6)$$

3 Proposed Method

Proposed MCLEP inherits the salient properties of LQEP and LBP, i.e., the relative difference of information between a central pixel and its neighbors in a local neighborhood. MCLEP is capable of capturing the texture information at various places for any color channels. Let M^x is the xth channel of an image M, with X, Y, C as size. Here, $x \in [1, N]$ and N denotes the total number of color channels. Let the N neighbors at the radius D of any pixel with location coordinates $J^t(m, n)$ for $m \in [1, P]$ and $n \in [1, Q]$ are defined as $J_t^U(x, y)$ where $s \in [1, R]$.

$$\text{MCLEP}_{N,D}^{t,d}(x, y) = \left\{\text{MCLEP}^1(x, y); \text{MCLEP}^2(x, y)\right\} \tag{7}$$

$$\text{MCLEP}^2(x, y) = \text{LQEP}_{N,D}^{t,d}(x, y) \tag{8}$$

$$\text{MCLEP}^1(x, y) = \text{LBP}_{N,D}^{'}(x, y) \tag{9}$$

Here, $d = 1, 2, 3, 4$ represents various distances at which the MCLEP is computed. LBP and LQEP are computed using the equations above in (1)–(5). MCLEP for tth space contains one LBP and four LQEPs.

3.1 MCLEP Feature Vector

Two major components of pattern recognition applications are feature representation and matching/classification techniques. For delivering competitive performance, it is essential to ensure robust implementation of both parts. MCLEP histogram HSTG is generated as the feature vector for robust image feature representation where HSTG $= \{\text{HSTG}_x\}, x \in [1, N]$. Let $J(m, n)$ be a color image of size X, Y, C with W intensity levels, i.e., $I \in \{0,1, \ldots, W-1\}$, where (m, n) is the location coordinate for $m \in [1, X]$ and $n \in [1, Y]$. The MCLEP histogram HSTG(.) is computed as given in Eq. (10),

$$\text{HSTG}_t(\text{MDLP}) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f(\text{MCLEP}_{Nb,R}^{t,d}(x, y), l) \tag{10}$$

$$f(x, y) = \begin{cases} 1 & x = y \\ 0 & \text{else} \end{cases} \tag{11}$$

Algorithm
i/p: *image*; *o/p*: *retrieved image*

1. Load the image
2. Compute the MCLEP

3. Match the query image with the database image
4. Retrieve the similar images

3.2 *Query Matching*

A distance metric such as given in Eq. (12) is used to compute the distance between the images under consideration. Here, $s_R = (s_{R_1}, s_{R_2}, \ldots s_{R_{Lg}})$ denotes the feature vector of query image. Similarity matching is carried out using d_1 distance metric as mentioned in Eq. (12). Each image is denoted by feature vector $s_{DBS_j} = (s_{DBS_{j1}}, s_{DBS_{j2}}, \ldots s_{DBS_{jLg}})$; $j = 1, 2, \ldots, |DBS|$.

$$d_1 \text{ distance: } d(R, I_1) = \sum_{k=1}^{Lg} \left| \frac{s_{\text{DBS}_{jk}} - s_{R,k}}{1 + s_{\text{DBS}_{jk}} + s_{R,k}} \right| \tag{12}$$

where $s_{\text{DBS}_{jk}}$ represents kth feature of jth image of database |DBS|.

4 Experimental Results and Discussions

Efficiency of the devised method is calculated by means of the experiments on standard databases. Corel image repository is used for this purpose. In every experiment, each database image becomes a query image. Accordingly, the proposed framework identifies n images $X = (x_1, x_2, \ldots, x_n)$ as per the shortest distance. Performance is evaluated in terms of avg. precision/retrieval precision (ARP), avg. recall/retrieval rate (ARR) as indicated below: Evaluation metrics for a query image I_q are calculated according to the equation given below.

$$\text{Prec: } P(I_q) = \frac{\text{Number of related images retrieved}}{\text{Total images retrieved}} \tag{13}$$

$$\text{Avg. Retrieval Prec: ARP} = \frac{1}{|\text{DBS}|} \left| \sum_{k=1}^{|\text{DBS}|} P(I_k) \right| \tag{14}$$

$$\text{Recall: Re}(I_q) = \frac{\text{Number of related images retrieved}}{\text{Total relevant images in database}} \tag{15}$$

$$\text{Avg. Retreival Rate: ARR} = \frac{1}{|\text{DBS}|} \left| \sum_{k=1}^{|\text{DBS}|} \text{Re}(I_k) \right| \tag{16}$$

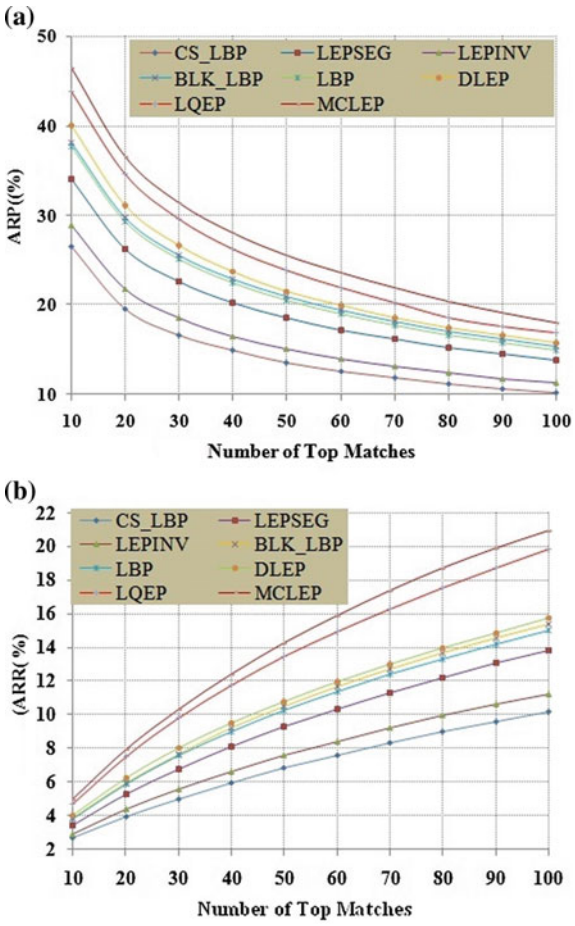


Fig. 2 Comparison of various methods in terms of **a** ARP **b** ARR on Corel-10k

a. **Corel-10k**
Corel-10k [14] is a collection of 10,000 images classified into different categories. Results of precision and recall are provided in Table 1. Comparisons of the results are shown in Fig. 2 whereas the retrieved images are provided in Fig. 3. The improvement in performance is mainly due to the discriminating ability of the proposed method against any other method mentioned in Table 1, Figs. 2 and 3.

b. **ImageNet-25k**
ImageNet-25k is a composition of 25,000 images. These are classified into different categories. Performance of MLCEP is determined by means of ARR and ARP (Fig 4).

Table 1 Results of various methods on corel database

Image repository	Evaluation metric	CS_LBP	LEP SEG	LEP INV	BLK_LBP	LBP	DLEP	LQEP	MCLEP
Corel-10k	%Prec	25.4	26.2	27.6	29.7	22.2	30.1	31.9	33.5
	%Recall	18.0	19.6	20.8	21.6	19.7	23.8	24.9	25.6

Fig. 3 Retrieved images using the proposed method

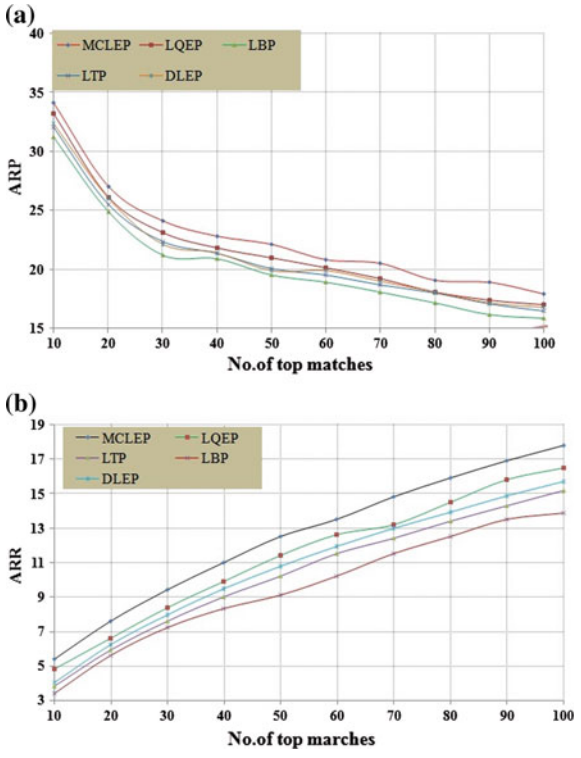


Fig. 4 Comparison of various methods in terms of **a** ARP **b** ARR on ImageNet-25k

5 Conclusions and Future Scope

A new method for image retrieval is introduced in this paper. It explores the directional data using multiple color channel local extrema pattern from color planes. Two types of repositories are used to test the effectiveness of the method. Results after investigation exhibit a substantial improvement in the performance metrics such as Precision and Recall.

By extracting the facial properties of any image, the proposed approach can be extended in the field of face recognition.

References

1. Walia E, Goyal A, Brar YS (2014) Zernike moments and LDP-weighted patches for content based image retrieval. SIViP 8:577. https://doi.org/10.1007/s11760-013-0561-z
2. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
3. Jégou H, Perronnin F, Douze M, Sanchez J, Perez P, Schmid C (2012) Aggregating local image descriptors into compact codes. IEEE Trans Pattern Anal Mach Intell 34(9):1704–1716
4. Tsikrika T, Popescu A, Kludas J (2011) Overview of the Wikipedia image retrieval task at ImageCLEF 2011. In: The working notes for the CLEF 2011 labs and workshop, Amsterdam, The Netherlands, 19–22 Sept 2011
5. Everingham M, Eslami SMA, Van Gool L, Williams CKI, Winn J, Zisserman A (2015) The PASCAL visual object classes challenge: a retrospective. Int J Comput Vis 111(1):98–136
6. Kokare M, Biswas PK, Chatterji BN (2007) Texture image retrieval using rotated wavelet filters. J Pattern Recogn Lett 28:1240–1249
7. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
8. Takala V, Ahonen T, Pietikainen M (2005) Block-based methods for image retrieval using local binary patterns. In: SCIA 2005, LNCS, vol 3450, pp 882–891
9. Yao C-H, Chen S-Y (2003) Retrieval of translated, rotated and scaled color textures. Pattern Recogn 36:913–929
10. Murala S, Maheshwari RP, Balasubramanian R (2012) Directional local extrema patterns: a new descriptor for content based image retrieval. Int J Multimedia Inf Retrieval 1(3):191–203
11. Koteswara Rao L, Venkata Rao D, Pratap Reddy L (2016) Local mesh quantized extrema patterns for image retrieval. SpringerPlus 5:976. https://doi.org/10.1186/s40064-016-2664-9
12. Koteswara Rao L et al (2018) Local color oppugnant quantized extrema patterns for content based natural and texture image retrieval. In: Multidimensional systems and signal processing. Springer (accepted)
13. Koteswara Rao L, Venkata Rao D (2015) Local quantized extrema patterns for content-based natural and texture image retrieval. Hum Centric Comput Inf Sci. https://doi.org/10.1186/s13673-015-0044-z
14. http://wang.ist.psu.edu/docs/related.shtml

Part II
Embedded Systems

Home Automation and E-Monitoring Over ThingSpeak and Android App

K. Susheela, E. Sri Harshitha, M. Rohitha and S. Maheswara Reddy

Abstract The paper discusses the usage of Internet of Things utilized for remote checking of the encompassing physical parameters and different stuffs with the utilization of sensors that for remote detecting of continuous information and move them into the coveted shape and forward the detected information over the system cloud by means of 'Web Connection' [1]. Here, the venture work manages the IoT 'ThingSpeak' web server which is a liberal open API server that goes about as a host for the assortment of sensors to screen the detected information at cloud. And furthermore introduces an adaptable home control and observing framework utilizing an installed small scale web server, with IP network for getting to and controlling gadgets and apparatuses remotely utilizing Android-based smart telephone application called Blynk. This process utilizes Wi-Fi Modules that procedures and exchange the detected information to the ThingSpeak Cloud and controlling the gadgets through App. The model was tried to produce constant data graphically over thing talk website page and on Blynk App as well.

Keywords IOT · Wi-Fi module (NODE MCU ESP8266) · API (application program interface) · Android app (blynk)

1 Introduction

The improvements in web advances and remote sensor networks (WSN), a fresh out of the plastic new example are being figured it out. Gigantic expand in clients of web and changes on the internetworking connected sciences empower systems administration of day by day protests [1]. "Internet of Things (IoT)" is about physical articles addressing each other, machine-to-machine interchanges and individual-to-

K. Susheela · E. Sri Harshitha · M. Rohitha
ECE, SPMVV, Tirupathi, Andhra Pradesh, India

S. Maheswara Reddy (✉)
ECE, GNITC, Hyderabad, Telangana, India
e-mail: mr.mahesh1515@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_14

pc correspondences will presumably be enhanced to "things" [2]. Key connected sciences with the goal that it self-discipline the more drawn out term IoT will respect insightful sensor connected sciences including WSN, Nanotechnology and Miniaturization. Individuals when in doubt inside their home interface with nature settings like light, air, and numerous others and control hence. In the event that the settings of the environment will likewise be made to answer to human conduct mechanically, there are a few points of interest.

2 Internet of Things: Home Automation and E-Monitoring Over ThingSpeak and Android App

Usage for Internet of Things utilized for checking standard home stipulations by methods for minimal effort universal detecting process. The depiction about the incorporated system structure and the interconnecting instruments for the solid estimation of parameters with the shrewd sensors and transmission of information by means of Wi-Fi [2] through web is being exhibited. It gives an adaptable home control and observing framework utilizing an implanted microweb server, with IP availability for getting to and controlling gadgets and apparatuses remotely utilizing Android-based smart telephone application.

2.1 Wireless Home Automation Technology (What) Using Internet of Things

IoT (web of things) is snappy rising science which involves communication among issues by methods for web without human impedance. It has made human presence simpler and quiet. Now-a-days advanced gadgets in home are becoming expediently

because of which there's a need of approaching and controlling the contraptions remotely. This paper demonstrates to an ease and adaptable abiding control approach using an Arduino, web server with IP availability for interfacing with devices and home gear remotely utilizing Android-focused shrewd portable application [3].

It exhibits the value of the strategy making utilization of contraptions much the same as delicate switches, temperature sensors, and water-level sensors.

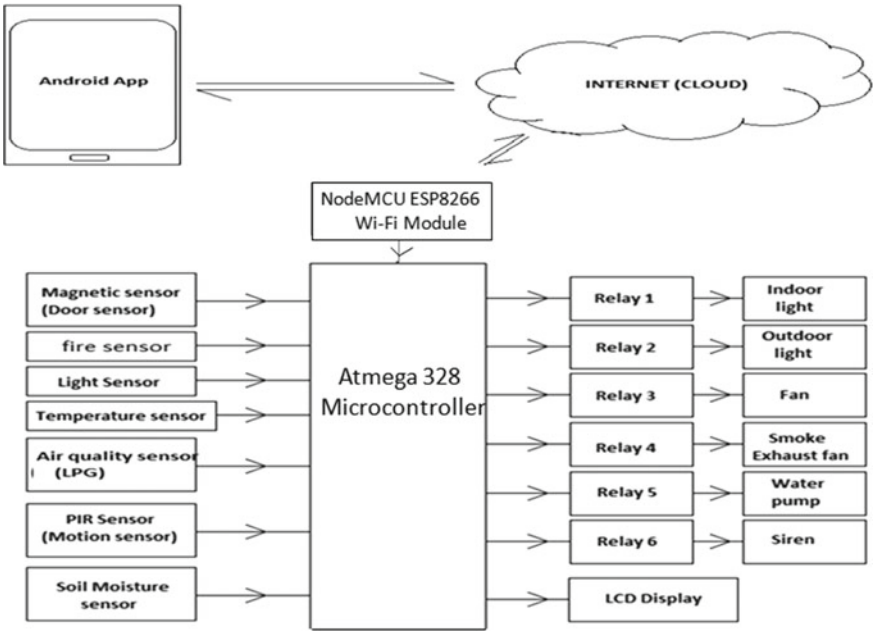
2.2 Towards the Implementation of IoT for Environmental Condition Monitoring in Homes

On this a strong implementation for Internet of things used for monitoring regular home stipulations by the use of low price ubiquitous sensing system. The description about the integrated community structure and the interconnecting mechanisms for the riskless measurement of parameters by wise sensors and transmission of knowledge through Internet is being offered [4]. The longitudinal learning system used to be ready to provide a self-control mechanism for higher operation of the instruments in monitoring stage.

3 System Overview

IoT (web of matters) is rapid technology which includes interaction amongst matters through Internet without human interference. It has made human life less complicated and relaxed. Nowadays, digital instruments in house are growing speedily due to which there is a need for accessing and controlling the gadgets remotely.

The fundamental aim of the mission support a strategy that interface the issues mechanically by means of sensors and the detected learning will screen or envision over web cloud or net server and in like manner outfit an economical and adaptable to control framework making utilization of a microcontroller, net server with IP network for associating with apparatuses remotely using Android established android app and in like manner give security.



3.1 Atmega328 Microcontroller

Atmega 328P is a 28 pin DIP. The Atmel eight-bit AVR RISC headquartered microcontroller combines 32 kB ISP flash reminiscence with learn-even as-write capabilities, 1 kb EEPROM, 2 kB SRAM, 2 normal rationale I/O traces, 32 normal cause working registers, three flexible timer/counters with examine modes, inside and outside interrupts, serial programmable USART, a byte-oriented 2-wire serial interface and a SPI serial port, 6-channel 10-bit A/D converter having eight channels in TQFP and QFN/MLF programs, programmable watchdog timer with internal oscillator and five application-selectable power saving modes. The device operates between 1.8 and 5.5 V. The gadget achieves throughput drawing near 1 MIPS.

3.2 Wi-Fi (Wireless Fidelity)

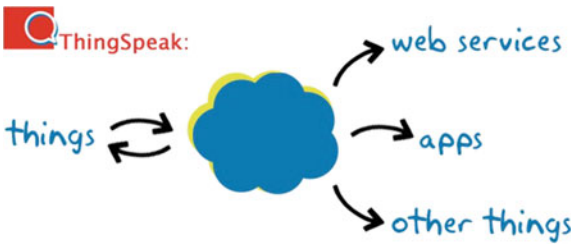
Wi-Fi is the name popular with wireless network technology that uses radio waves to provide wireless high-speed Internet and network connections. Wi-Fi is short for wireless fidelity means which reproduces exactly the same. Wi-Fi is used to cover the distance up to 10–100 m with a frequency band of 2.4 and 5 GHz. Wi-Fi is based upon 802.11 standard protocols and it is one of the fastest communication protocols.

One of the best Wi-Fi module that suits this project is Nodemcu because of its following features

- NodeMCU is Wi-Fi for Internet of things.
- NodeMCU contains a single-chip ESP8266 WiFi SoC.
- The NodeMCU board adds a USB/UART converter chip.
- It is becoming more popular because of its integrated wifi-PCB antenna.

3.3 ThingSpeak Web Server

ThingSpeak is an "open-source information stage and application program interface (API) for Internet of Things". It empowers constant information accumulation, investigate, picture, and information put in channels. Each channel permits to put away to 8 fields of information, utilizing something like 255 alphanumeric characters each. ThingSpeak is an API which tunes into approaching information, timestamps it, and yields it for both human clients (through visual diagrams) and machines. As indicated by the Things crest site, 'Things' are objects that are offered sensors to gather information. Information is sent and obtained by means of basic "Hypertext Transfer Protocol" (HTTP). "Things peak" speak with the assistance of web office which goes about as an 'information bundle' transporter between the associated 'things' and the "Things top" cloud. The information is then transferred to the cloud and from that point it can be utilized for an assortment of purposes.



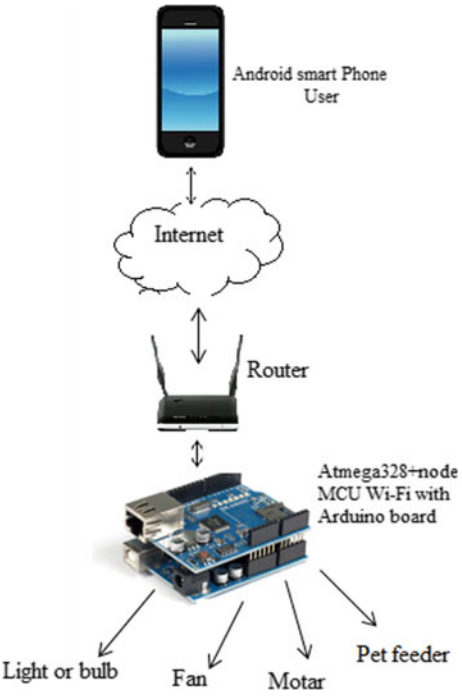
3.4 Android App (Blynk)

Blynk is an application which is easy to understand stages for controlling equipment over the Internet. Blynk is Platform for Makers and electronic specialists to enable them to attach their Smartphone to any associated equipment and assemble a visual UI for their venture in minutes.

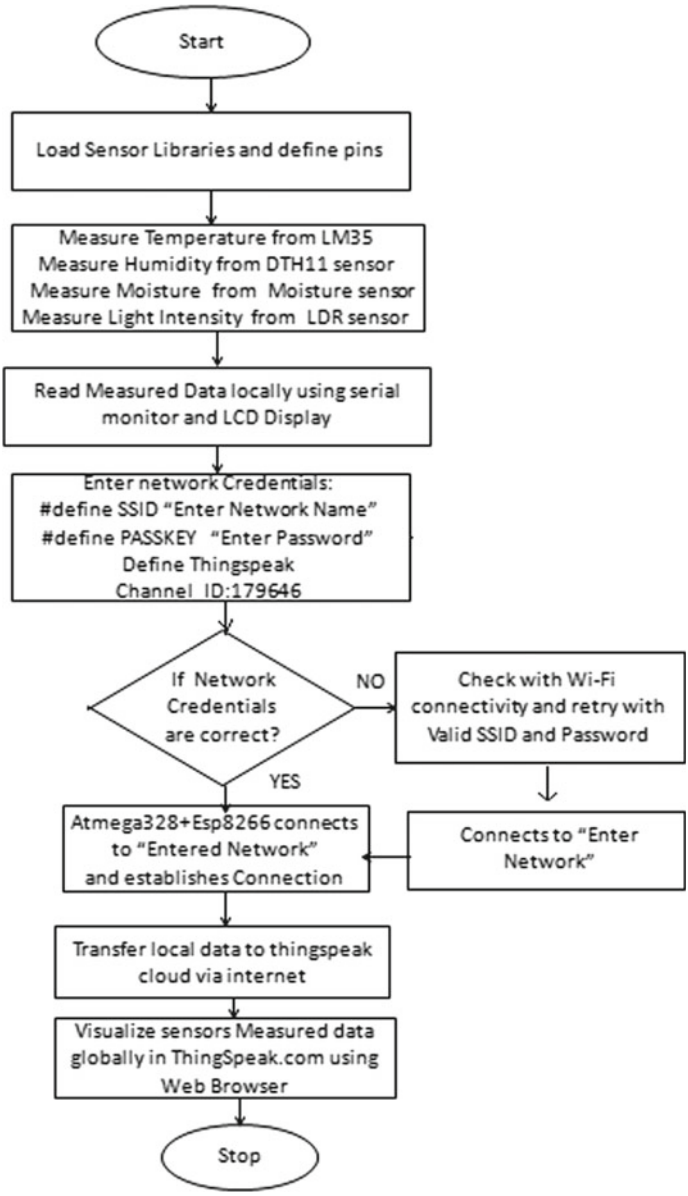
1. Blynk App permits to you make astonishing interfaces for your activities utilizing different gadgets we give.
2. Blynk Server is in charge of all correspondences between advanced mobile phone and equipment. You can utilize blynk cloud or run your private blynk server locally. Its open-source cloud effortlessly handles a huge number of gadgets.

4 Architecture

This system consists of one Atmega 328 microcontroller and two node MCU Wi-Fi modules. All the sensors and relay switches are connected to microcontroller. Microcontroller continuously monitors the sensors data and sends them to two Wi-Fi modules, one transfers the data to "ThinkSpeak" cloud and other transfers the sensors data to "Blynk" android app. The Node MCU which transfers the sensors data to android app also takes the commands from the user who is monitoring and controlling the system. Here, there are four buttons for applications like motor controlling for Pet feeding, AC socket for external electrical devices controlling, fan control and door opening and closing motor controls which are placed at door. There is one intensity slider to increase or decrease the light intensity in a room. A graph is also there to know the air quality levels in a room.



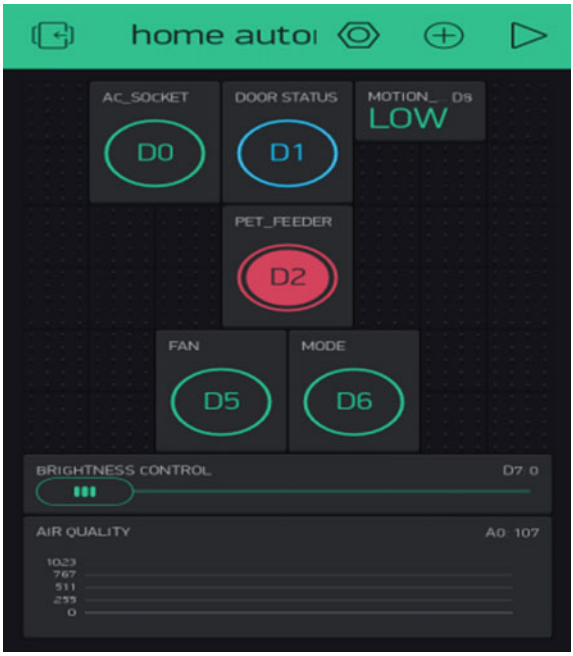
5 System Flow Chart



6 ThingSpeak Output Result

7 Blynk IoT Output Results

"ThinkSpeak" cloud and other transfers the sensors data to "Blynk" android app. The Node MCU which transfers the sensors data to Android app also takes the commands from the person who is monitoring and controlling the system.



8 Future Work

The future of MATLAB in ThingSpeak and vice versa provides a deep study and analysis of sensed data at critical level that is to manage the surrounding environment where the parameters are important to measure.

9 Conclusion

Hence, the Internet of Things facilitates numerous benefits to the society and from my project, can provide, and prove the strength of IoT using the ThingSpeak-based sensing and monitoring devices over Internet. ThingSpeak API is capable to contribute the services for the purpose of building vast number of IoT applications and any Android-based Smartphone with built-in support for Wi-Fi that can be used to access and control the devices at home.

References

1. Surie D, Laguionie O, Pederson T (2008) Wireless sensor networking of everyday objects in a smart home environment. In: Proceedings of international conference on intelligent sensors, sensor networks and information processing, pp 189–194
2. Chen C-H, Gao C-C, Chen J-J (2011) Intelligent home energy conservation system based on WSN. Presented at the international conference on electrical, electronics and civil engineering, Pattaya
3. ElShafee A, Hamed KA (2012) Design and implementation of a Wifi based home automation system. In: World academy of science, engineering and technology, pp 2177–2180
4. Al-Qutayri MA, Jeedella JS (2010) Integrated wireless technologies for smart homes applications. In: Al-Qutayri MA (ed) Smart home systems. InTech

Smart Automation to Robot

G. Prutha and G. S. Anitha

Abstract In this paper, the concept of the Smart Automation to Robotics is discussed. The main constraint of this paper is to obtain the power supply and to adjust the angle for the robotic arm. The power supply from the grid supplies to the power cord of the robots. The power is converted from the AC-DC bridge converter using the bridge rectifier converter. Then, the DC–DC conversion is made to reduce according to the required amount of voltage (Liu et al. in Robot assisted smart firefighting and interdisciplinary perspectives, 2016, [1]). The buck converter is used in this paper to reduce the voltage and to supply the power constantly to robotics. It supplies to brushless DC motor (Aamir et al. in Express Briefs 62(8):816–820, 2015, [2]). There will be six motors used in this paper. The angle of each motor difference is maintained such that the pallet from each station is picked and placed easily at the respective station. For this process, even position sensor is used (Chinag et al. in 40th annual conference of the IEEE industrial electronics society (IECON-2014), Dallas, TX, USA, 2014, [3]). By this, the robot simulation is done in the MATLAB/SIMULINK in 2017a version. The same process is carried in the hardware but the controller used in it is PLC programming. PLC is used for the robot to control the move to the required station in the store.

Keywords PLC · PI · Robotic arm · Load · Buck converter · PI controller

1 Introduction

Now in the world is stepping ahead to go with the new scheme of work in under automation process. So even in this paper, the automation to the robotics with the smart and intelligent process is done in this work. The main intention of this work

G. Prutha (✉) · G. S. Anitha
Department of EEE, RV College of Engineering, Bengaluru, India
e-mail: pruthags@gmail.com

G. S. Anitha
e-mail: anithags@rvce.edu.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_15

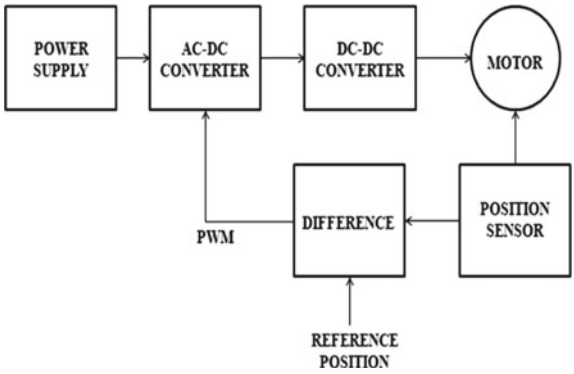


Fig. 1 Block diagram of smart automation to robotics

is to automate each thing such that the manpower applied to it is made down. The robots are the one where the work of man can be replaced by this machine. This in turn will do its work more appropriate than man. The mistake consumption will be lesser compared to the human [4] in the industries. So everywhere in the industry, they are going for the robotic automation [5].

The smarter way in improving the timing to be covered for all six stations is one of the smartness to be improvised in this work.

The robot used in this paper has six-axis motor where in that they have six motors present in it each motor has their own strategy to rotate at this angle as the program is made in that format and fed to it. The work of robot in this paper is to place and pick the pallet at the required station to continue with the operation it is to undergo. Figure 1 shows the block of smart automation to robotics with the working conditions [6–8]. The output of bridge rectifier is provided to DC–DC converter where the converter buck is preferred as the motor requirement is lesser voltage compare to the supply so the converter buck is considered [9]. Then the output of is given to the motor. The motor axis and motion is sensed with the position sensor is considered such that to provide the power flow continuously such that the robot work continuously [10–12]. The drive controller is considered to drive motor accordingly to the mechanical load requirement.

Figure 2 shows the circuit of the smart automation to robotics. The working is done by converting the AC supply to DC. Using bridge rectifier the, in this paper, robot arm is used to do its operation the pallet for all six station and the continuous loop operation as to be performed for production of glow plugs in the industry [13–17]. As the company requirement of the glow plug is very high rated, they require 24 h operation as to be performed. So, this was a challenging aspect to run the robot 24 h without a stop. The continuous loop operation was must in this peak area. This paper was selected at this area to get good amount of production as well as a good quality of glow plug.

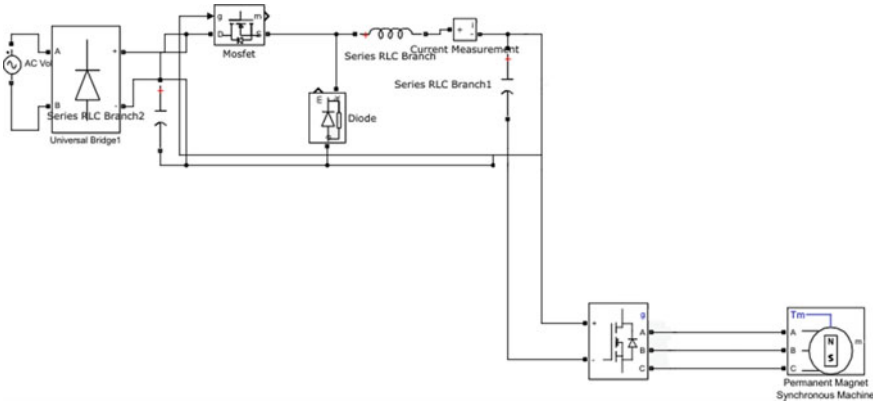


Fig. 2 Circuit diagram of smart automation to robotics

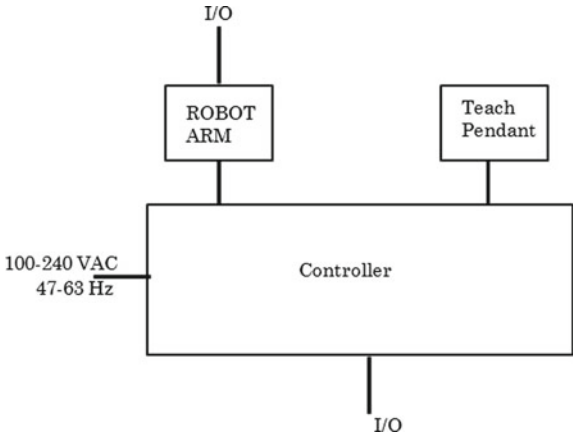


Fig. 3 Block diagram of robotic arm with controller

2 Robotic Arm with Controller

Figure 3 shows the block of arm of robotics with controller. The controller is mainly used will sense the positions of each motor. The work of the robotic arm is divided with six divisions with each motor [18]. Each motor is placed accordingly with the portion the rotation undergoes in the robot. The work is considered from supply of the power till the load will be completed. The teaching to the robot is one with teach pendant in the hardware to each motor. Each motor is considered as base, shoulder, elbow, wrist 1, wrist 2, and wrist 3. The angle of each motor will be fed with the teach pendant according to the required work for all the station in the glow plug heating unit line.

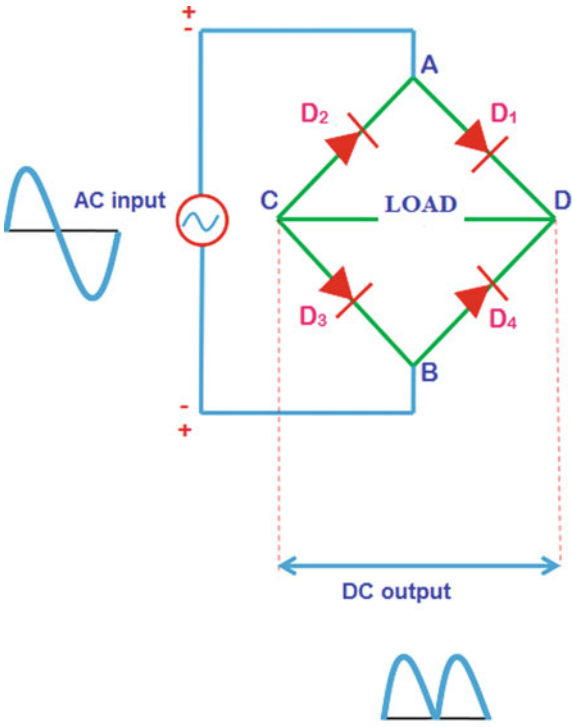


Fig. 4 Bridge rectifier

3 Concepts

The intention of this paper is robotic applications for the designing of power supply to give the continuous power to the robot. The second concept is to maintain the position of each motor using position sensor [19].

3.1 Power Supply

The AC-DC conversion is required as the power supply from the grid is 240 V, 50 Hz that is to be converted to 24 V. So first the Bridge Rectifier is used to convert from AC to DC (Fig. 4).

In Fig. 4 during the positive cycle, the diodes $D1$ and Diode 3 will be forward bias and the flow of current will be in forward direction, i.e., ADCB. Then, the diodes $D2$ and $D4$ will be verse bias and the flow of current will be in reverse direction. During the negative cycle, the diodes $D2$ and $D4$ will be forward bias and the current flows in forward direction, i.e., ACDB. Then the diodes $D1$ and $D3$ will be reverse bias and the current flows in reverse direction [20].

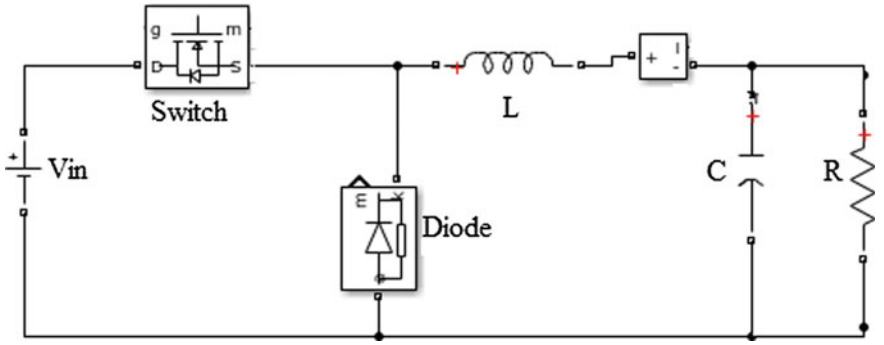


Fig. 5 Circuit of buck converter

Table 1 Angle inclined for the motors

Position of motors	Angle inclined	Torque (Nm)
Base	360°	25
Shoulder	180°	20
Elbow	120°	20
Wrist 1	120°	15
Wrist 2	90°	15
Wrist 3	90°	10

The buck converter is to convert from 240 to 24 V. The buck converter is used to maintain constant voltage and send to the motors present in the robot [21].

In Fig. 5, the circuit of buck converter is shown. Its operation is to lesser the voltage according to the requirement of the work.

- Switch will be open (off-state), current is zero.
- Switch will be closed (on-state), current increases, and inductor produces opposing voltage across the terminals is used for changing current.

This is the working procedure of the buck converter in this work [3].

3.2 Position Sensor

The position sensor used in motors where the Hall effect sensor will be used. The Hall Effect sensor senses the position of the magnet present in the motors [4].

The motor magnet is positioned with the effect of hall sensor. The torque load is weighed from base to wrist 3 the load varies from one motor to the other motor.

In this Table 1, the position of each motor is given with the angle inclined. The load carried by the pallet is mentioned in this table. The angle is set accordingly for required movement in the Robot to move to the required position to all the station.

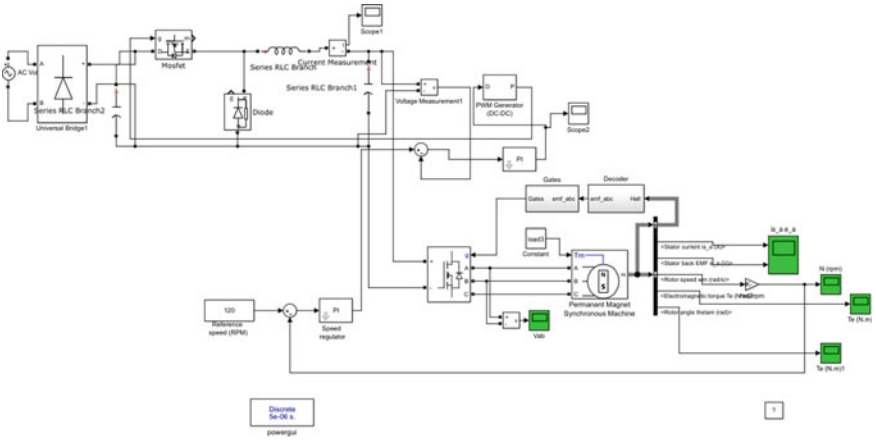


Fig. 6 Simulation of smart automation to robotics

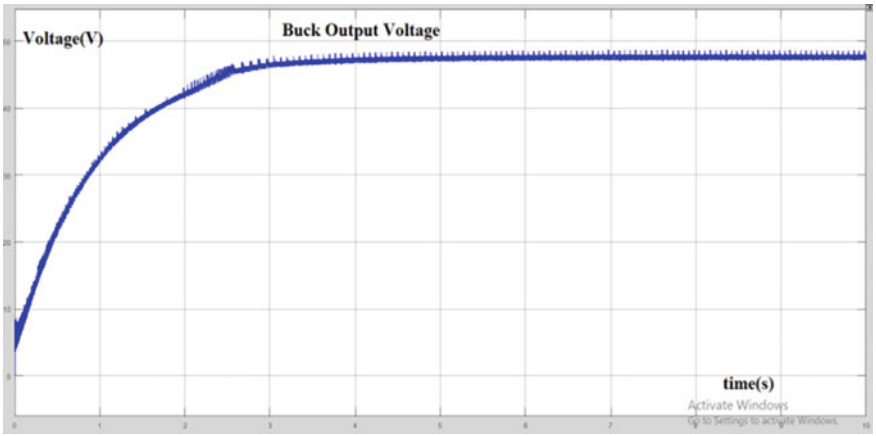


Fig. 7 Buck output voltage

4 Simulation Results

The Robot is designed in this paper to the high-risk area in the BOSCH Company in Glow Plug area. The robot is installed in the heating unit line to pick the pallet and place the pallet to store the plugs after the work is done. In this paper, our main intention is to reduce the consumption of human energy where the person was working hard in this area. So, the robot plan was made to execute to this work. The robot's main concern of work in this paper is to implement and to design the power supply and to teach the position angle to robot. The robot is thought to move the required station to place and pick the pallet to that station (Figs. 6, 7, 8, 9, 10, 11, 12, 13 and 14).

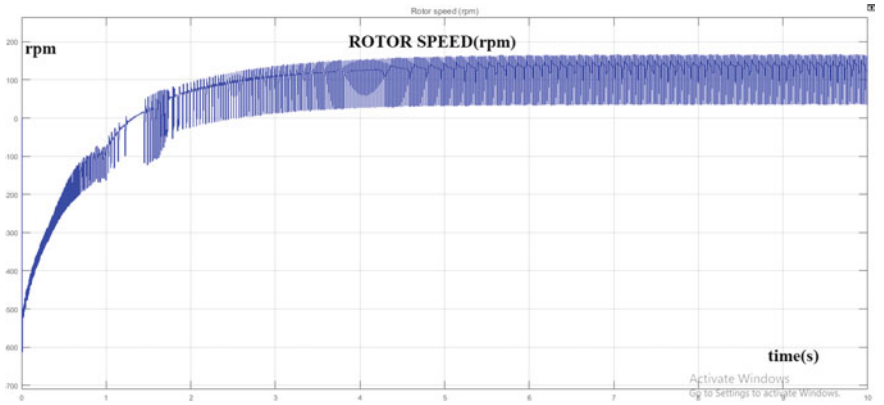


Fig. 8 Rotor speed

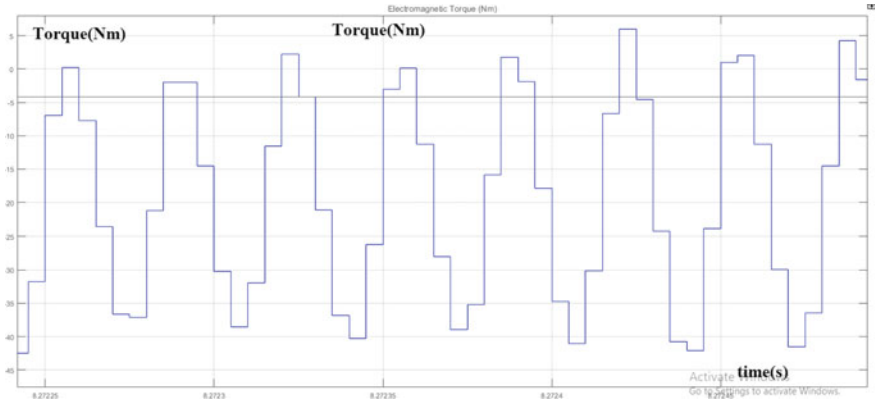


Fig. 9 Torque at base

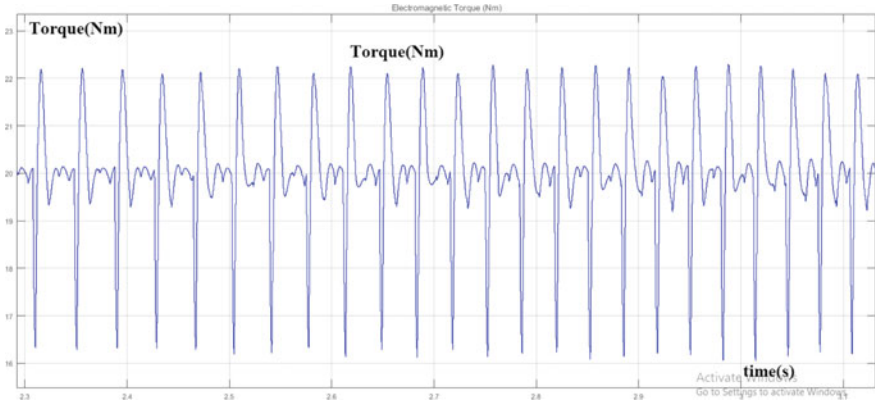


Fig. 10 Torque at shoulder

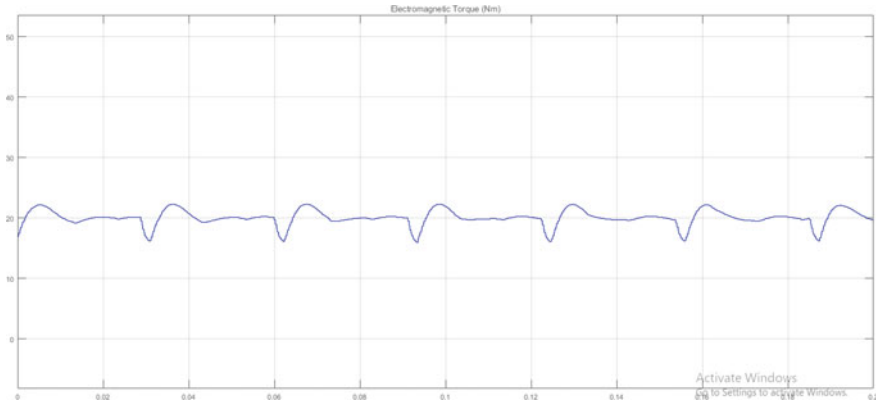


Fig. 11 Torque at elbow

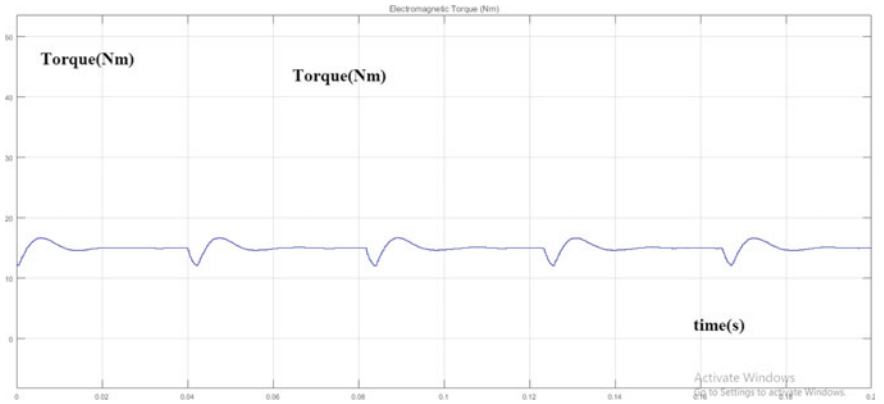


Fig. 12 Torque at wrist 1

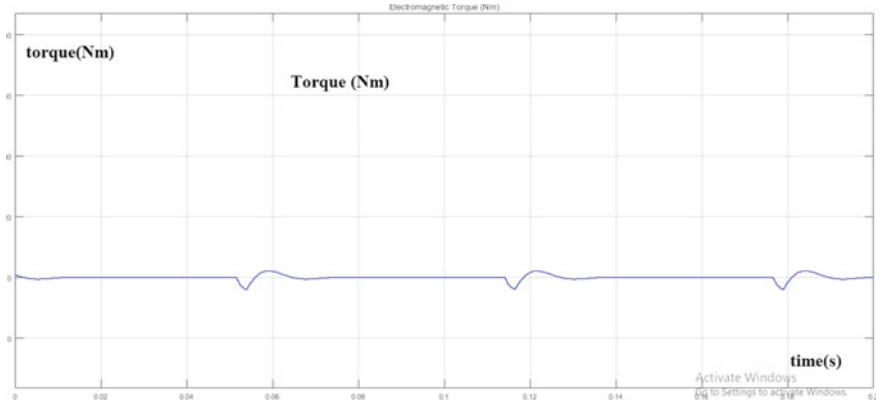


Fig. 13 Torque at wrist 2

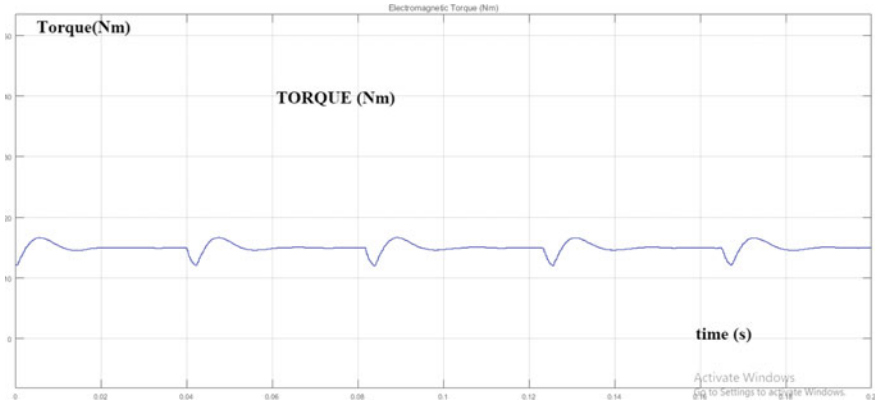


Fig. 14 Torque at wrist 3

5 Hardware Implementation

The working of hardware is implemented in Bosch company. Hardware components and their implementations. Later it discusses the method to create a dc–dc converter. The battery will charge and discharge through dc–dc converter. The designing of the converter of buck is done with this value. Here: $D = 0.4$, $V_{\text{in}} = 240$ V, $V_{\text{out}} = 24$ V, $f_{\text{S}} = 50$ kHz, $I_{\text{out}} = 0.5$ A, $L = 2.4$ mH, $C = 2.5$ μF (Figs. 15, 16 and 17). The motors rotation format is mentioned in Fig. 16.

6 Results and Discussion

Using SIMULINK, the model of Robot is simulated to support the hardware. The simulation consists the process of robot simulation in this project which intends to support the hardware with the power supply and position angle of all six motors. The power supply from 240 to 24 V is achieved using DC–DC Buck converter is given to motor. The position angle is set in the teach pendant. The robot is made to run with continuous cycle for 24 h a day in the company to achieve good production.

7 Conclusion

The hardware implementation and test results of the converters are presented. A detailed analysis of the power supply applied and position angle obtained for robot is presented. Power supply and position sensor are designed. Switching sequences are developed to sense each motor angle. Simulation of the model is carried out and

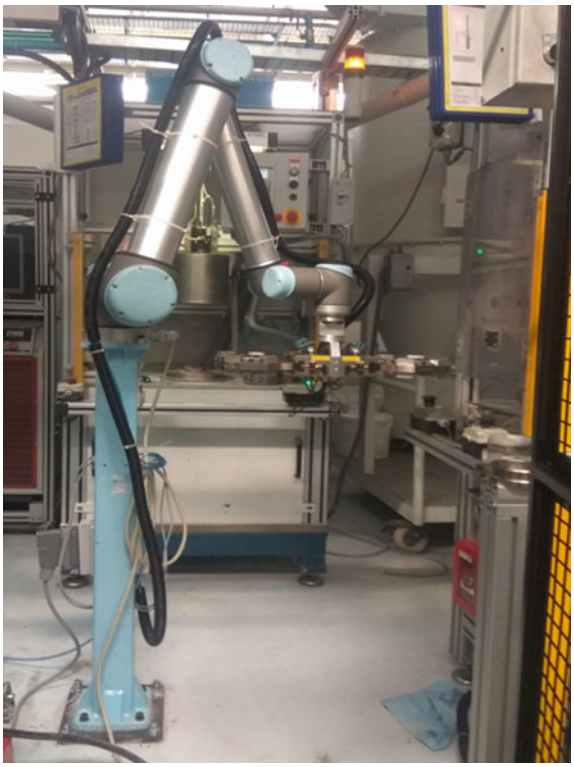


Fig. 15 Hardware setup of smart automation to robotics

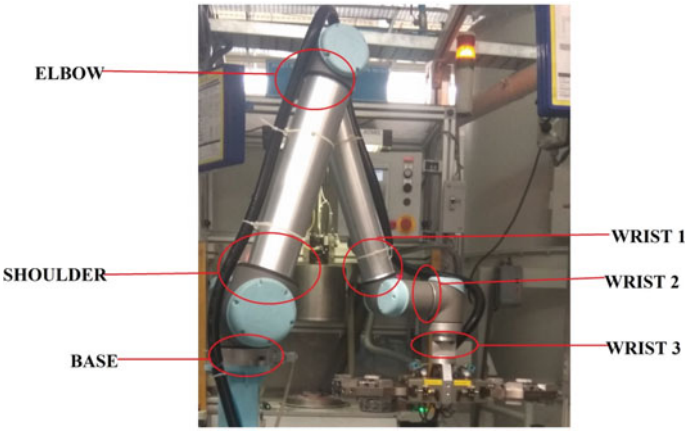


Fig. 16 Different position of motors present in robotic arm

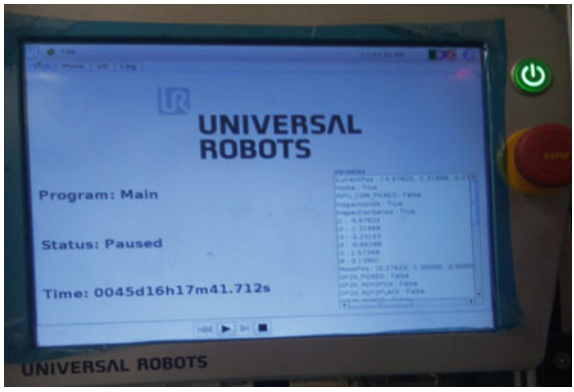


Fig. 17 Teach pendant of robot

comparison between hardware and software is made. The voltage is converted from 240 to 24 V by using buck converter. The base motor angle 360° is achieved by providing load of 17 Nm. Similarly for all motors present in robot is calculated:

- At Base: 360° is achieved for 17 Nm.
- At Shoulder: 180° is achieved for 16 Nm.
- At Elbow: 120° is achieved for 16 Nm.
- At Wrist 1: 120° is achieved for 15 Nm.
- At Wrist 2: 90° is achieved for 15 Nm.
- At Wrist 3: 90° is achieved for 10 Nm.

Considering rotation angle and load, robot is designed to pick and place the pallet.

8 Future Scope

The intelligence automation could be used to make simpler programs of PLC. Instead of PLC, we can go with microprocessor where the cost cutting can be done to this project.

Acknowledgements I like to thank the institute and BOSCH Company to make our work in progress with all their support and wish and would like to thank our family and friends for all their contribution for doing this project successful.

References

1. Liu P, Yu H, Cang S (2016) Robot assisted smart firefighting and interdisciplinary perspectives
2. Aamir M, Mekhilef S, Kim H-J (2015) High-gain zero-voltage switching bidirectional converter with a reduce number of switches. Express Briefs 62(8):816–820. https://doi.org/10.1109/tcsii.2015.2433351

3. Chinag S J, Shieh H-J, Chen M-C (2014) Design optimization for the adoption of bonded magnets in PM BLDC motors. In: 40th annual conference of the IEEE industrial electronics society (IECON-2014), Dallas, TX, USA. ISBN: 978-4799-4032
4. Boese C, Kirchoff MR, Feldmann M, Guttler J, Buttgenbach S (2009) New generation of integrated position sensor system for parallel robotic applications. Sensors (2009-IEEE)
5. Waheed A, Aslam B, Awais M (2016) Design of indoor intelligent agent of outdoor Combi-Forklift robot as effector
6. Chang H-F, Chang C-P, Lin W-T, Li YG (2016) Study of application of assessment modes in development of smart automatic production systems in Industry 4.0
7. Anupama A, Seema PN Srikanth V (2010) Variable step-size algorithm implemented two phase soft-switched interleaved boost converter. In: IEEE conference 2012 vol 29, no 8, pp 753–46
8. Asaduzzaman A, Moniruzzaman M, Chidella KK, Tamtam P, Simulation methods using VisualSim to assess autonomous power systems
9. Anusuya V, Vishnupraba G (2015) Controller for buck converter and three phase inverter, vol 4, no 4, Apr 2015
10. Chinag SJ, Shieh H-J, Chen M-C (2009) Modeling and control of PV charger system with SEPIC converter. IEEE Trans Power Electron 56(11):4344, 4353
11. Hintz A, Prasanna UR, Rajashekara K (2014) Novel modular multiple-input bidirectional DC-DC power converter (MIPC), pp 2343–2350. https://doi.org/10.1109/ipec.2014.6869917
12. Lai C-M, Lin Y-J, Hsieh M-H, Li J-T (2016) A newly-designed multiport bidirectional power converter with battery fuel-cell vehicle system, pp 163–166. https://doi.org/10.1109/itec-ap.2016.7512941
13. Asaduzzaman A, Moniruzzaman M, Chidella KK, Tamtam P (2010) An efficient simulation method using VisualSim to assess autonomous power systems. In: International power electronics conference, Dubai, vol 23, no 4, pp 213–295
14. Aamir M, Mekhilef S, Kim H-J (2015) High-gain zero-voltage switching bidirectional converter with a reduced number of switches. IEEE Trans Circuits Syst II Dhaka 62(8):816–820
15. Cultura AB, Ziyad M (2012) Design and analysis of a 24 Vdc to 48 Vdc bidirectional DC-DC converter specifically for a distributed energy application. IEEE J (Jaipur) 10:2654–5426
16. Broday GR, Nascimento CB, Agostini E, Lopes LAC (2015) A tri-state bidirectional buck-boost converter for a battery/supercapacitor hybrid energy storage system in electric vehicle applications. In: IEEE vehicle power and propulsion conference (VPPC), Ahmadabad, pp 1–6
17. Fardoun AA, Ismail EH, Sabzali AJ, Al-Saffar MA (2012) New efficient bridgeless Cuk rectifiers for PFC applications. IEEE Trans Power Electron Egypt 27(7):3292–3301
18. Chae J, Cha H, Kim H-G (2014) Analysis and design of two-phase zero-voltage switching bidirectional dc-dc converter using coupled inductor, pp 301–306. https://doi.org/10.1109/icit.2014.6894884
19. Aquila D, Liserre M, Monopoli V, Rotondo P (2008) Overview of PI based solutions for the control of DC buses of a single-phase H-bridge multilevel active rectifier. IEEE Trans Ind Appl 44(3):857–866
20. Hintz A, Prasanna UR, Rajashekara K (2014) Novel modular multiple-input bidirectional DC-DC power converter (MIPC). In: International power electronics conference IPEC, Hiroshima, ECCE ASIA, pp 2343–2350. https://doi.org/10.1109/ipec.2014.6869917
21. Loianno G, Cross G, Qu C, Mulgaonkar Y, Hesch JA, Kumar V (2015) Flying smart phones

Posture Monitoring and Back Pain Relief Using ATmega328, Android and Unity

Aishwarya S. Acharya, Pavitra Gandhi, Sushmita Karkera, Nimish Ghagare and Sanjay Deshmukh

Abstract Neck and back pain is one of the most common medical problems that are faced by 80% of individuals at some point in their lives. Bad posture decreases the flexibility of the joints making them prone to injuries. It can also lead to hyper-kyphosis, an extremely hunched back. This leads to numerous visits to physicians and physiotherapists. These sessions are expensive, making affordability a major concern. In this paper, we propose and implement a cost-effective approach that proposes a device that detects bad posture and helps the user correct it using ATmega380 and Inertial Measurement Unit (IMU) which is equipped with various accelerometers and gyroscopes. Once the patient is familiar with using the device, he/she will be able to improve his/her posture and reduce muscle pain from home itself, eliminating costs arising from visits to physiotherapists and orthopaedists. Our proposed method will monitor the upper body movement alerting the user through vibrations in the device and notifications in an Android application. It will visually show the correct movements to go through day-to-day activities using animations. Also, it will provide pain relief exercises using Virtual Reality (VR) for improved posture and a healthier lifestyle.

Keywords ATmega380 · Cost effective · IMU · Posture monitoring · Biofeedback · Back muscle pain relief · Spondylitis · Forward head posture · Texting neck

1 Introduction

The average human head weighs about 5 kg. However, based on the degree of tilt of the neck, the weight on the cervical spine can be as much as 27 kg [1]. Short-term effects of bad posture are not very significant; however, poor posture has shown to have severe long-lasting effects. The long-term effects are spondylitis which is a

A. S. Acharya (✉) · P. Gandhi · S. Karkera · N. Ghagare · S. Deshmukh
Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: karkerasoni@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_16

Table 1 Angles of the head and their corresponding weights in pounds and kilograms

Angle of head (degrees)	0	15	30	45	60
Weight on spine (lbs)	10	27	40	49	60
Weight on spine (kg)	4.5	12	13.6	22	27

common age-related problem that affects both the joints and the discs present in the human body, hyperkyphosis which is an extremely hunched back, to name a few. Back pain has also proved to increase the risk of early death by 13% [2]. These severities could have had been avoided simply by more care and attention to the body. This paper aims to build a device that users can attach around their neck and back to encourages healthy and an improved lifestyle. This device will keep track of user's head tilting angles along the X and Y axes using an IMU chip, thus alarming the user using sound or vibration and a notification on the smartphone. The data collated will be sent to the Android application through a Bluetooth Module, HC-05 which makes communication possible between the microcontroller and the mobile application.

Accelerometers and gyroscopes are commonly incorporated in an IMU. They detect the position change, measuring inclination and angles of rotations from two sensors that are combined to obtain the positions. When attached on the body, the sensors attached will monitor the force of the head subtended on the backbone [3]. Studies have revealed that people having improper posture have more severe neck pain than others. The application also provides games in which virtual reality is incorporated. This game will not only encourage users to be in an upright position but also teach them how to use these exercising models in their day-to-day life. The biofeedback mechanism will ensure correct posture. The motion animations are made on Akeytsu version 1.0.3.0. Akeytsu is a powerful tool used to create 3D animations for game development. Unity software was used for the Virtual Reality games. Unity is a cross-platform game engine that is used to develop simulations for computers, smartphones and consoles.

2 Literature Survey

A survey was conducted for the Indian population where 41% of men and women between 30 and 60, in urban areas had neck and back pain. On doing further research, we found that the average cost of one session with a physiotherapist is about Rs. 800–1500 per session, and the physiotherapist recommends somewhere around 8–10 sessions for the patient. The cost of therapy combined with the lack of awareness leads to high negligence regarding back pain. This posture negligence came out to be among 55% of people. For the user to experience minimum pain in the neck and back, the posture should be such that the spine faces minimum weight (Table 1).

A group of students in Mountain View, California, manufactured a device called Lumo Lift [4]. It monitors the body posture and helps the body to maintain correct

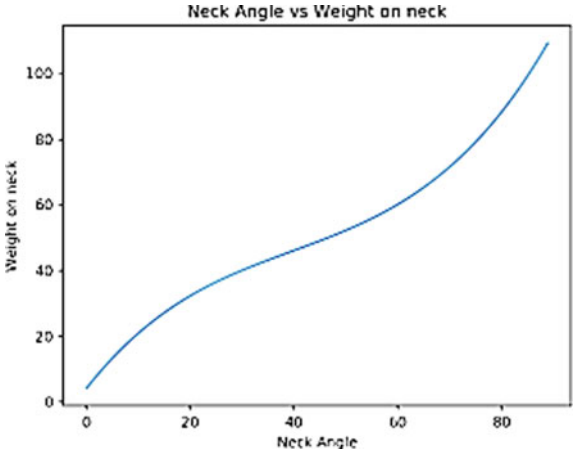


Fig. 1 Graph of neck angle versus weight on neck

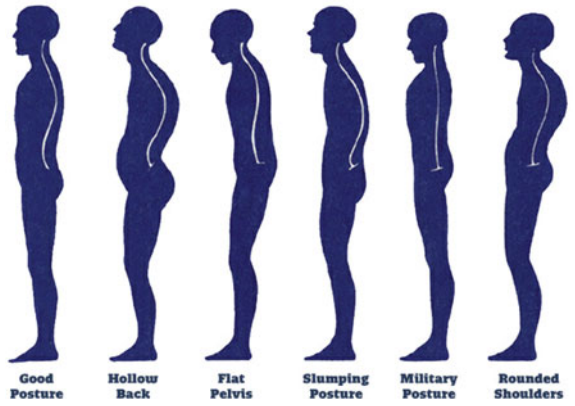


Fig. 2 Comparison of various back postures

posture, much like Upright Go [5]. These products' prime market reach is in the United States. University of Illinois, Urbana-Champaign has also created a similar device [6]. In India, the major population is ignorant about their health, which encouraged us to make a device for posture control along with an inbuilt coach using interesting Virtual Reality Exercises which makes you exercise. The solution proposed in the paper is a simplistic, low-cost one which not only monitors posture but also makes people exercise daily, making it different than the existing solutions (Figs. 1 and 2).

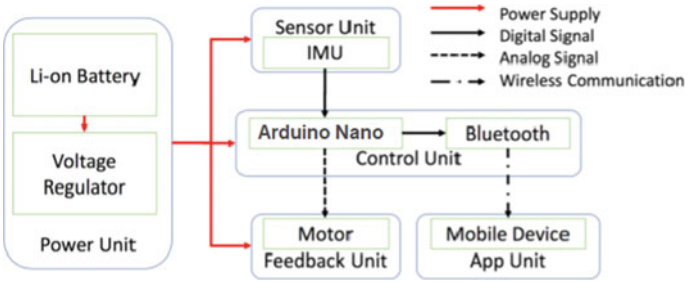


Fig. 3 System block diagram

3 System Overview

The proposed posture monitoring system contains a wearable device will have three major components: control unit, sensors feedback and virtual reality. The control unit consists of a Central Processing Unit (CPU). We are using Arduino Nano board with an ATmega328 chip as the processing system. It consists of an ATmega380 which is a high-performance 8-bit microchip AVR RISC based microcontroller. It can operate between 1.8 and 5.5 V [7]. The Arduino Nano chip has 8 analogue input pins and 14 digital I/O pins so interfacing sensors is easy [8].

We have used two Inertial Measurement Units (IMU) which contain gyroscopes for angular velocity and the accelerometer measures the linear acceleration, which would measure the user's neck angle as compared with the ideal position. The IMU provides us with three values: roll, pitch and yaw. Using this data, we have calculated the angles with the X, Y and Z axis.

The user's posture details will be communicated with the help of a Bluetooth module and an Android application (Fig. 3).

4 Implementation

We interfaced IMU6050 with Arduino and the connections were debugged for proper functioning [9]. The sensor was calibrated to get stable X, Y and Z values, based on yaw, pitch and roll with minimum error, displayed on Serial Monitor of Arduino IDE. The communication will take place via Bluetooth which is a reliable and ordered connection [10]. Bluetooth Module HC-05 was paired with the Android device on which developed application was installed. Android Studio IDE was used to develop the application. The mobile application will real-time track posture of the user indicating whether posture is healthy and provide Virtual Reality games.

For incorporating exercises, Unity software and Akeytsu software are being used with the mobile application. Figure 4 shows the flowchart of the working of the system. The sensors initially will be calibrated with the threshold values of the

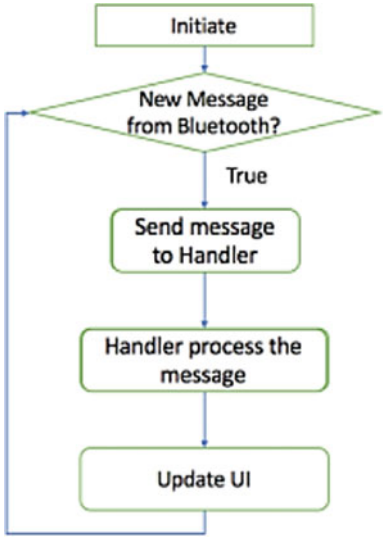


Fig. 4 The flow of the Android application

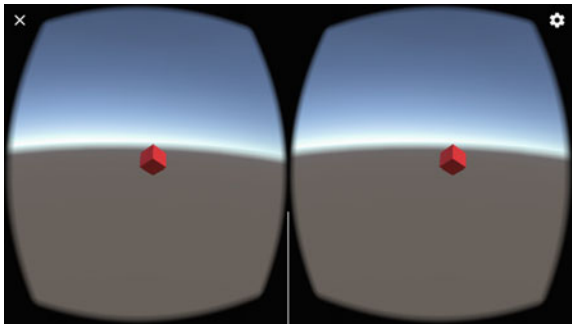


Fig. 5 VR game to be viewed through VR headset

angles as suggested by the physiotherapist. These values will be constantly checked again the current values and if greater, the user will be alerted. All the mentioned exercises and games will be created in accordance with health regulations and have been verified by physiotherapists for safety and effectiveness (Figs. 5, 6 and 7).

5 Results and Discussion

The objective of the project is to assist the users in practicing good posture. We aim in providing users confident and healthier lifestyle.

Table 2 compares values obtain from the protractor versus values measured by the IMU (Table 3).

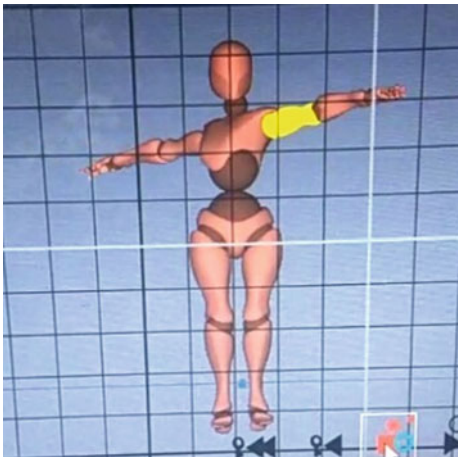


Fig. 6 Human animated model created using Akeytsu software used for demonstrating muscle movements

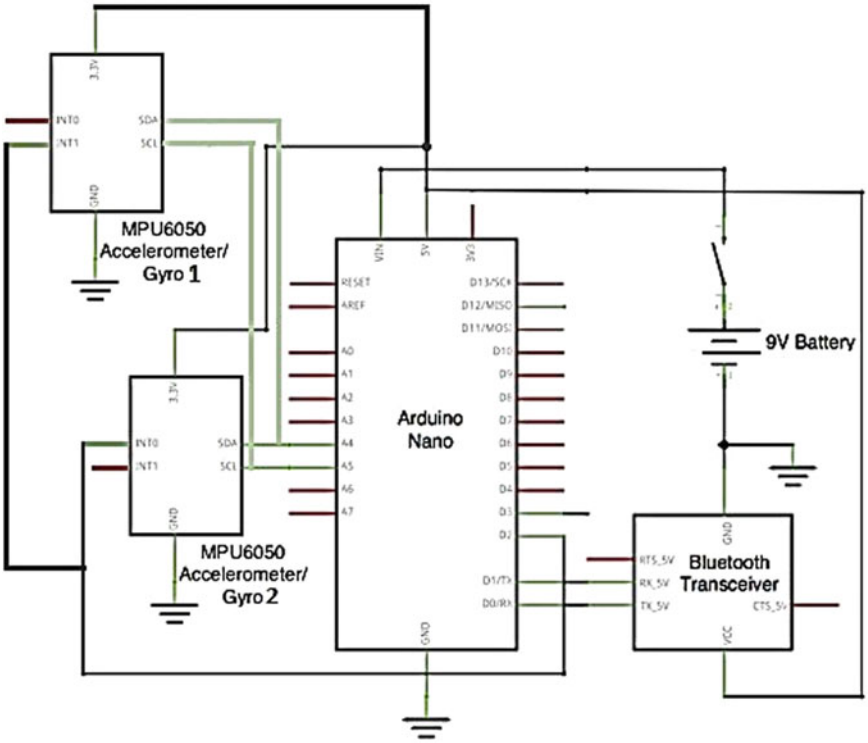


Fig. 7 Circuit diagram of all connections with individual units

Table 2 Comparison of real angles with measured angles

Real angle from protractor (degrees)	10	30	50	70	95
IMU measured angle (degrees)	11	31	49	70	94

Table 3 Sensors and their description

Sensors	Output	Results obtained
IMU 6050	VDD analogue voltage range of 2.1–3.6 V	Angle values obtained are approximately equal to the angles measured by protractor with $\pm 1^{\circ}$ error
Bluetooth module HC-05	Power: +3.3 V DC 50 mA	Successful serial communication of angle values obtained from IMU to the mobile application
Vibration sensor	Digital output: 3.3–5 V	Vibrating successfully when the angle exceeds calibrated threshold

6 Conclusion and Future Scope

This paper has described a wearable neck brace to monitor the upper body movements ensuring a healthy posture, and interactive VR exercises for relief from pain that would help in the prevention as well as quick recovery from back muscles pain at home. This approach is novel in the sense that it provides an entertaining way to exercise and improve body posture. This method is likely to gain response because a significant percentage of urban adult human population suffers from back pain. The mentioned approach offers over 10–15 times reduction in cost in comparison with the package cost of painkiller medicines, massage therapy, relieving sprays, transcutaneous electrical nerve stimulation and cost of physiotherapy. This one-time cost of the device will prove to be a blessing to the working population who do not have the time to take care of their health and later have to face back pain problems because of their negligence. We aim to expand this product in the future such that the entire body posture can be improved and add various technologies for communication with a doctor while using the device. Finally, a more compact design can be made using smaller components and more efficient design and placement.

Acknowledgements The authors would like to thank Prof. Sanjay Deshmukh for his continuous guidance. We would also like to express our utmost gratitude to Dr. Anand Walawalkar for his efforts in helping us through our research in back and neck exercising techniques.

References

1. Gleiber M, Conceirge Spibe surgery. Available: https://michaelgleibermd.com/news/heads-text-neck-ruin-posture/
2. Indian Express Article, Back pain ups risk of early death, Feb 2017
3. De Rossi D (2007) Wearable kinesthetic sensors for body posture and movement analysis. J Biomech 40(S2):S425–S427
4. Lumo lift posture activity and coach monitor. Available: https://www.lumobodytech.com/lumo-lift/
5. Upright. Smart wearable posture trainer. Available: https://www.uprightpose.com/
6. Capozzo JA, GonG J, Wen J (2017) "Health-care neckband" UIUC design review, senior design spring
7. ATMEL ATMega 380/P datasheet, Oct 2016. arduino nano
8. Arduino Nano (V2.3). User manual released under the creative commons attribution share-alike 2.5
9. Winer K (2016) Arduino sketch for MPU-9250 9DoF with AHRS sensor fusion [online]. Available: https://github.com/kriswiner/MPU-9250
10. Google bluetooth socket [Online]. Available: developer.android.com/reference/android/bluetooth/BluetoothSocket.html
11. Business Wire India (2017) Report on people with neck and back pain on world spine day
12. IMU-3000—SparkFun electronics (2010) [Online]
13. HC-05 bluetooth module (master/slave): rhydoLABZ India [Online]
14. Bartalesi R, Lorussi F, Tognetti A, Tesconi TM, Zupone G, Carbonaro, Smart vest for posture monitoring in rehabilitation exercises

Adaptive Neural Network-Based LMS for DSTATCOM

Vivekananda Ganji, D. Suresh and K. Chandrasekhar Koritala

Abstract This paper presents distributed static compensator (DSTATCOM) to eliminate harmonics and reactive power of the nonlinear load. The DSTATCOM is realized using voltage source inverter with DC bus capacitor. The DSTATCOM acts as a harmonics current source and inject reactive compensation current. The control schemes for determining the reference compensating currents of the three-phase DSTATCOM based on least mean square (LMS) algorithms are presented. The performance of LMS algorithm with PI and artificial neural network controller (ANN) is studied. The algorithm used for training ANN controller is Levenberg–Marquardt backpropagation (LMBP). The training data for ANN controller is generated offline. The firing pulses for DSTATCOM are obtained with hysteresis current controller. An extensive simulation study is carried out to test the performance of ANN controller and compared with PI controller.

Keywords PI controller · LMS · Active power filter · Harmonic

1 Introduction

Nowadays, there has been advancement in semiconductor technology, power electronics devices with high voltage, high current and high switching frequency devices. Power electronics devices employed in the electrical apparatus such as uninterruptible power supply, switched mode power supplies, electric drives, battery

V. Ganji (✉)
Department of Electrical and Electronics Engineering, Acharya Nagarjuna University, Guntur, India
e-mail: vivek.nanda20@gmail.com

D. Suresh
Department of Electrical and Electronics Engineering, Vignan Institute of Technology and Science, Deshmukhi, Hyderabad, India

K. Chandrasekhar Koritala
Department of Electrical and Electronics Engineering, RVR&JC, Guntur, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_17

charger, etc. The most of these electric apparatus may generate the harmonics and reactive power due to the nonlinear characteristics of such loads. The existence of the harmonic current and reactive current leads to increases power losses. In addition, it will also reduce the active power flow capability of the distribution system. The existence of unbalanced load on the distribution system causes heating of transformers and generators [1]. The DSTATCOM was also known as active filter and can be used for compensation of harmonics, reactive power, voltage regulation, power factor correction and load balancing in the distribution system, etc. [2]. The compensation performance characteristic of DSTATCOM depends on reference current assessment and the control process used for drawing out of same reference current. In literature, many control schemes are reported for efficient estimation. Some of these reported schemes are instantaneous reactive power (IRP) theory, synchronous reference frame (SRF) theory, neural networks-based approaches and per phase dc voltage regulation methods [3–10]. Among these projected control methods, instantaneous reactive power and synchronous reference frame theories are commonly used. But these control algorithms suffer from various transformations, addition and multiplications. This paper presents DSTATCOM which is controlled using LMS algorithm integrated with PI and ANN controller algorithm for compensation of harmonics current, and reactive power demand and performance characteristics of these controllers are compared. This LMS-based control algorithm is simple to implement and needs less computational efforts [10–11]. A fast least mean square-based reference current estimator extracts reference current without any phase shift. The assessment of reference currents was carried with neural network-based LMS adaptive online estimator algorithm for the computation of weights [10]. An extensive MATLAB simulation is presented for these two control techniques of DSTATCOM. Simulation results demonstrate the effectiveness of these two control algorithms of DSTATCOM for the elimination of harmonics and oscillatory reactive power demand of the nonlinear load.

2 System Configuration

Figure 1 depicts the topological circuit diagram of DSTATCOM with nonlinear load connected to electrical distribution system.

3 Control Algorithms

The performance characteristics of the DSTATCOM depend mainly on the control scheme implemented to generate the compensating reference currents. This paper presented two controller schemes to get the reference current, which is significant matter in the control of the DSTATCOM.

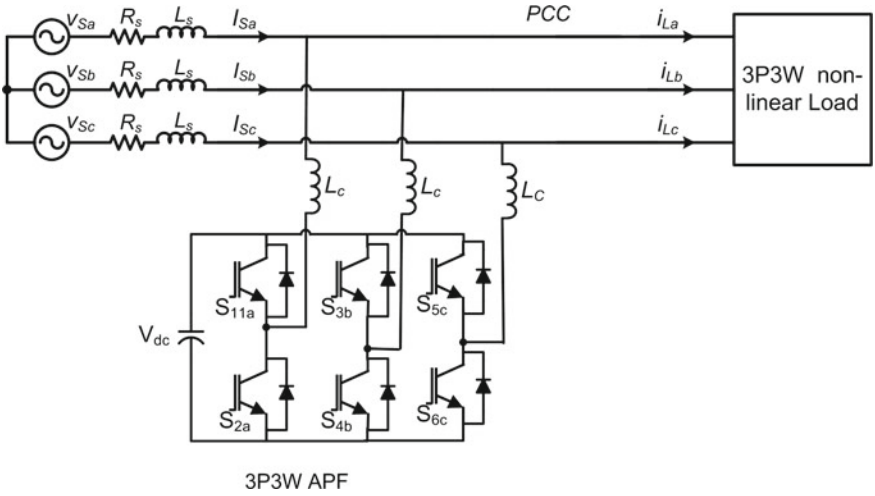


Fig. 1 DSTATCOM topology

4 ANN Controller

In order to improve the dynamic performance of the DSTATCOM ANN controller is designed and integrated with LMS adaptive estimator algorithm for reference current extraction. In this paper, multilayered feedforward network with three layers is designed. The three layers, namely, input layer with two neurons, hidden layer with 20 neurons and output layer with one neuron, respectively, are designed. The ANN controller is tuned with the conventional method. The activation functions are sigmoidal for input and hidden layers and pure linear in the output layer, respectively. The ANN controller topology is shown in Fig. 2.

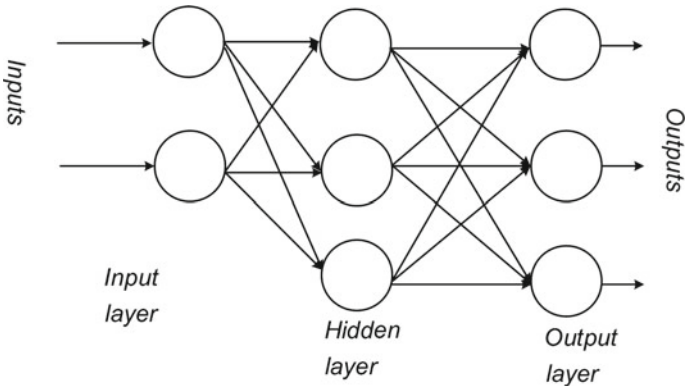


Fig. 2 Adaptive neural network controller

The training algorithm used is Levenberg–Marquardt backpropagation (LMBP). The MATLAB programming of ANN training is given as follows:

```
net = newff(tric,troc,10,{'tansig','tansig'},'traingd');
net.trainParam.show=50;
net.trainParam.epochs=10000;
net.trainParam.goal=0.0001;
net=train(net, trinc, troutc);
y=sim(net,trinc);
gensim (net,-1);
```

5 Reference Current Generation

The performance of the DSTATCOM depends on its reference current generation. The instantaneous source voltage is

$$v_{s_n}(t) = V_{m_n} \sin(\omega t) \tag{1}$$

When the nonlinear load connected to the source, then it injects the current harmonics in the source. The load current consist of both fundamental and harmonics components, which can be expressed as follows:

$$\begin{aligned} i_{L_n}(t) &= \sum_{n=1}^{\infty} I_{n_n} \sin\left(n\omega t + \phi_{n_n}\right) \\ &= I_{1_n} \sin(\omega t + \phi_1) + \sum_{n=2}^{\infty} I_{n_n} \sin\left(n\omega t + \phi_{n_n}\right) \end{aligned} \tag{2}$$

Thus, the instantaneous load current can be expressed as follows:

$$\begin{aligned} i_{L_n}(t) &= I_{1_n} \sin(\omega t) * \cos\left(\phi_{1_n}\right) \\ &\quad + I_{1_n} \sin\left(\phi_{1_n}\right) * \cos(\omega t) + \sum_{n=2}^{\infty} I_{n_n} \sin\left(n\omega t + \phi_{n_n}\right) \end{aligned} \tag{3}$$

The load current is consists of active, reactive and harmonics component and can be expressed as

$$i_{L_n}(t) = i_{f_n}(t) + i_{r_n}(t) + i_{h_n}(t) \tag{4}$$

where $i_{h_n}(t)$, $i_{r_n}(t)$ and $i_{f_n}(t)$ represent the sum of the harmonics, reactive component and represent active component of current, respectively. If DSTATCOM injects harmonics and reactive component of current, under this condition source needs to

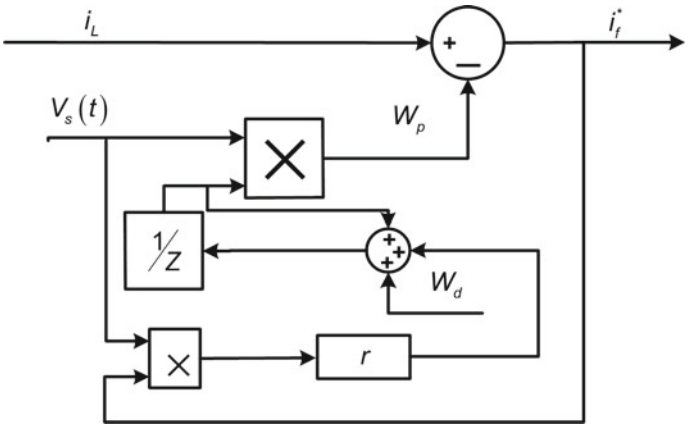


Fig. 3 Reference current generation using LMS algorithm

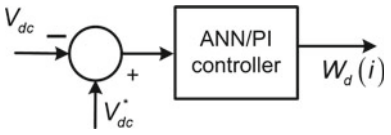


Fig. 4 DC voltage regulation loop

supply only active component of the current. The current injected, which is injected by DSTATCOM, can be expressed as

$$i_{c_n}(t) = i_{L_n}(t) - i_{f_n}(t) = i_{L_n}(t) - I_{1_n}\cos(\phi_n)\sin(\omega t) \tag{5}$$

From (3), after compensation, current supply by DSTATCOM

$$i_{c_n}(t) = i_{L_n}(t) - I_{sm_n}\sin(\omega t) \tag{6}$$

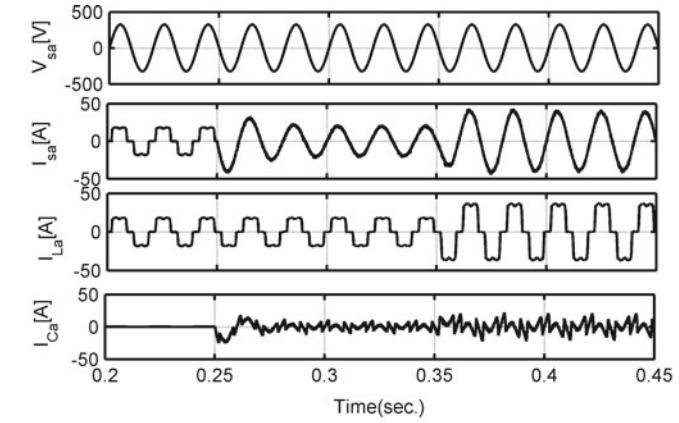
where $I_{sm_n} = I_{1_n}\cos\phi_{1_n}$.

The voltage across the dc side capacitors varies with the load demand. In order to maintain dc capacitors voltage constant, the PI/ANN controller is used Fig. 3.

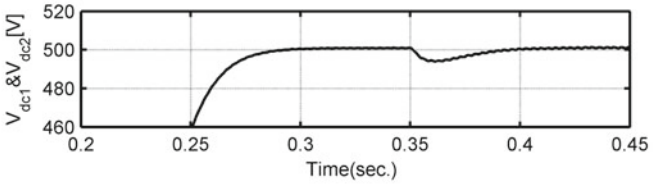
The dc link voltage control loop of DSTATCOM with ANN/PI controller is shown in Fig. 4.

6 Simulation Results

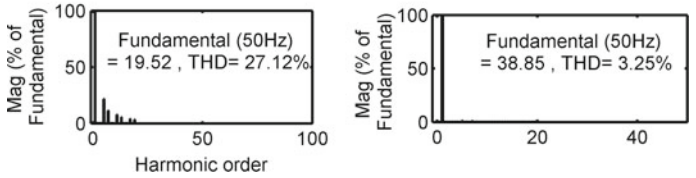
An uncontrolled rectifier with RL load on dc side is used as nonlinear load. Figure 5a–d shows the simulated response of currents, dc side capacitor voltages and harmonics spectrum, respectively, of DSTATCOM using PI controller. Figure 6a–b



(a) Source current waveforms



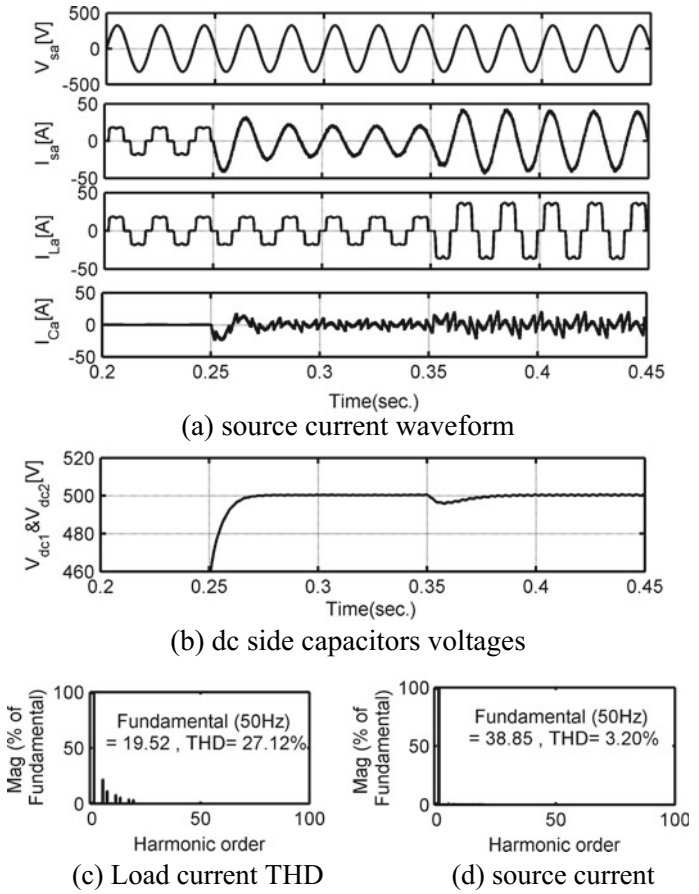
(b)



(c) Load current THD (d) Source current THD

Fig. 5 Compensation characteristics of DSTATCOM with PI controller

shows the simulated response of currents, dc side capacitor voltages and harmonics spectrum, respectively, of DSTATCOM with ANN controller. After compensation with DSTATCOM, in both the cases the source current is changed from stepped waveform to sinusoidal waveform as seen in Figs. 5a and 6a. Source current is in phase with source voltage to conform the power factor becomes almost unity. After compensation with DSTATCOM, the load current THD is 26.7% but the source current THD is reduced to 3.41 and 3.40% with PI controller and ANN controller as shown in Figs. 5d and 6d. In both the cases dc side capacitor voltage is controlled at reference level but the settling time required for dc side capacitor to attain reference value is 0.06 and 0.05 s for PI controller and ANN controller, respectively, as observed from Figs. 5b and 6b. At instant $t = 0.3$ s, the load current is increased from 20.2 to 40 A. The equivalent variation in the source current, compensating current



(a) source current waveform

(b) dc side capacitors voltages

(c) Load current THD

(d) source current

Fig. 6 Compensation characteristics of DSTATCOM with ANN controller

dc side capacitors voltages has been observed to be compensating the increased load current. However, the increase of compensating current observed to be smooth in case of ANN-LMS than PI-LMS controller-based control scheme.

The increased load current also causes the decrease in the dc side capacitors voltage from it reference value to compensate the increase in load current. The settling time of the dc side capacitors voltage is 20 and 10 ms for PI controller-based controller algorithm and adjustable step-size algorithm, respectively. The dc voltage regulation is smooth in both the cases but restoration time is more in case of PI controller than that of the ANN controller.

7 Conclusion

A control scheme for high-performance three-phase DSTATCOM with improved power factor and displacement factor has been proposed. The compensating performance of controller has been investigated with different controllers. The current control based on hysteresis controller has been implemented. The simulation results of DSTATCOM have been carried in MATLAB/Simulink power system block sets with proposed control schemes. The simulated response shows that the proposed control schemes with ANN controller have a good performance in the harmonic compensation and improving the power factor, displacement factor and dynamic performance of the system.

References

1. Singh B, Verma V (2008) Selective compensation of power-quality problems through active power filter by current decomposition, power delivery. IEEE Trans Power Deliv 23(2):792–799
2. Gayadhar P, Ray PK, Puhan PS, Dash SK (2013) Novel schemes used for estimation of power system harmonics and their elimination in a three-phase distribution system. Int J Electr Power Energy Syst 53:842–856
3. Saad S, Zellouma L (2009) Fuzzy logic controller for three-level shunt active filter compensating harmonics and reactive power. Electric Power Syst Res 79(10):1337–1341
4. Premalatha S, Dash SS, Babu PC (2013) Power quality improvement features for a distributed generation system using shunt active power filter. In Procedia Eng, ICONDM 2013 64:265–274
5. Singh GK, Singh AK, Mitra RA (2007) Simple fuzzy logic based robust active power filter for harmonics minimization under random load variation. Electric Power Syst Res 77(8):101–111
6. Panda AK, Mikkili S (2013) FLC based shunt active filter (p–q and Id–Iq) control strategies for mitigation of harmonics with different fuzzy MFs using MATLAB and real-time digital simulator. Int J Electr Power Energy Syst 47:313–336
7. Singh B, Solanki J (2006) A comparative study of control algorithms for DSTATCOM for load compensation. In Proc IEEE ICIT, 15–17 Dec 2006, pp 1492–1497
8. Singh B, Verma V, Solanki J (2007) Neural network-based selective compensation current quality problems in distribution system. IEEE Trans Ind Electron 54(1):53–60
9. Widrow B, McCool JM, Ball M (1975) The complex LMS algorithm. Proc IEEE 63(4):719–720
10. Suresh D, Singh SP (2016) Design of single input fuzzy logic controller for shunt active power filter. IETE J Res 61(5):500–509

Modeling and Simulation of Dual Redundant Power Inverter Stage to BLDCM for MEA Application

B. Suresh Kumar, B. V. Ravi Kumar and K. Sindhu Priya

Abstract The increasing demand for electrification functions on control surfaces of aircraft leads to a new concept of new advancement "more electric aircraft (MEA)". In the aviation industry, an electromechanical actuator (EMA) is used to maintain the orientation of aircraft, landing gears, and braking systems. As the electrical components such as inverter and BLDC motor are key components in EMA, the designing of these components became a critical issue. To enhance the reliability in actuation system in aircraft, this paper proposes dual redundant power inverter system to run BLDC motor with redundancy management and logic, fault tolerant, and fault diagnosis by using MATLAB/SIMULINK with the results.

Keywords Electromechanical actuator (EMA) · Dual redundant · Inverter · BLDC drive · Fault diagnosis

1 Introduction

More electric aircraft (MEA) [1] is a new concept which involves the maximum utilization of electrical technology in control surfaces with expertise mastered control architectures. In such flight control surfaces, an electromechanical actuator (EMA) [2] plays a major role to control the direction of aircraft. A subsystem of an EMA includes power electronic circuitry, controller, a BLDC drive, ball screw mechanism for position detection, and its associated components. As the fundamental structural

B. Suresh Kumar · B. V. Ravi Kumar · K. Sindhu Priya (✉)
CBIT, Gandipet, Hyderabad, India
e-mail: kalvasindhupriyareddy@reddifmail.com

B. Suresh Kumar
e-mail: bskbus@gmail.com

B. V. Ravi Kumar
e-mail: venkataravi.kumar@rcilab.in

B. V. Ravi Kumar
Research Centre Imarat (RCI), Hyderabad, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_18

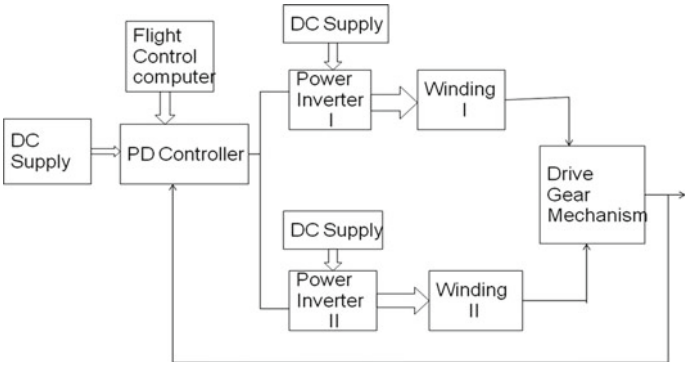


Fig. 1 Electromechanical actuator block diagram

blocks in EMA are power converters and BLDC drive, designing of these components became a challenge. A simple fault occurrence in any subsystem of electromechanical actuator leads to whole outage in the system and many lives would be at risk. So, aircraft applications require safety and accessibility. So "redundancy management" is necessary.

Dual redundant [3] logic consists of two modules consisting of power inverter stage and BLDC motor windings as shown in Fig. 1. If any fault occurs, the other alternative module will drive the whole actuation system maintaining the reliability. But how rather it switches to other alternative is discussed in this paper.

This paper presents the MATLAB simulation of mathematical model of BLDC drive, power inverter stage to EMA.

2 Redundancy Management

The redundancy logic [4] to this system is limited for only one fault in each module. In Fig. 2, Primary chain is always in operation under normal conditions. If any fault occurs, then only redundant chain takes the service. If there is a fault in PI-1, the system can run without interruption either through PI-2, W2 and PI-2, W1. Of course, the latter is preferred because if the second fault occurred in W2, and then the whole system will be out of redundancy.

3 Mathematical Modeling of BLDC Motor

The BLDC motor can be modeled in the same way as conventional dc motor. The detailed approximated equations d-q transform is discussed in the paper [5]. The equivalent electrical circuit is shown in Fig. 3.

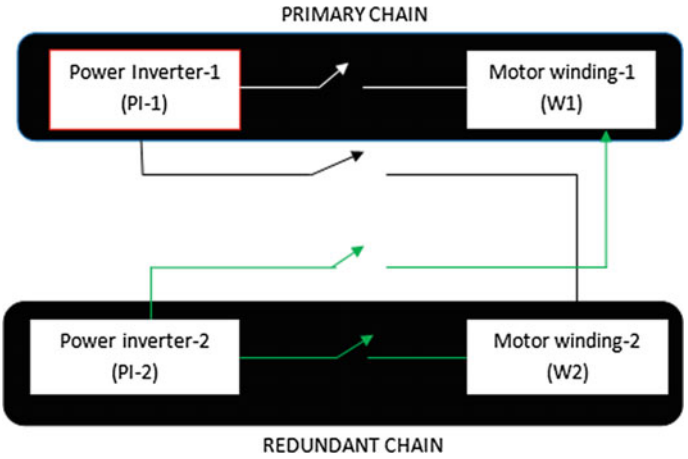


Fig. 2 Redundancy management

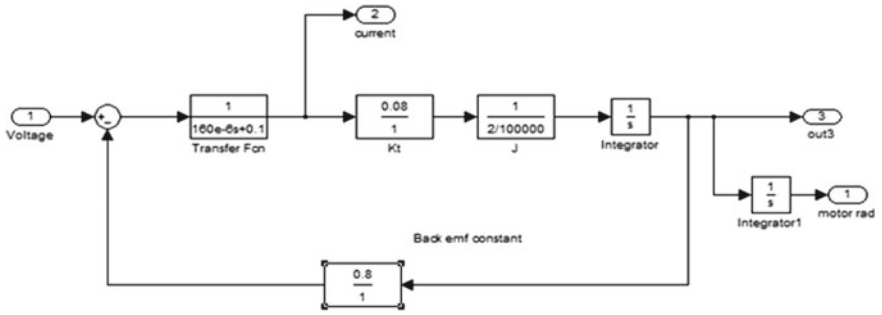


Fig. 3 Mathematical model of BLDC motor

The electrical equivalent of the armature coil can be described by an inductance (L_a) in series with a resistance (R_a) in series with an induced voltage (V_a) which opposes the voltage source.

A differential equation for the equivalent circuit can be derived by using Kirchhoff's voltage law around the electrical loop.

$$V_a = I_a R_a + L_a \mathrm{d}i/\mathrm{d}t + e_b$$

By using Laplace transform of the above equation,

$$I(s) = (I_a(s) - e_b(s))/(Ls + R_a)$$

The electromagnetic torque is proportional to the current and can be written as

Table 1 Possible switching cases

Healthy/section	Faulty section	Recommended path
Case-I		
PI-1 → W1	If Fault at W1	PI-1 → W2
PI-1 → W2	If fault at PI-1	PI-2 → W2
Case-II		
PI-1 → W1	If Fault at PI-1	PI-2 → W1
PI-2 → W1	If Fault at W1	PI-2 → W2
Case-III		
PI-1 → W1	If Fault at both PI-1 and W1	PI-2 → W2

$$T_e = K_{\mathrm{t}} I_{\mathrm{a}} \text{ and } T_e = J\,\mathrm{d}w/\mathrm{d}t + B\omega;$$
$$\text{where } \omega = \mathrm{d}\theta/\mathrm{d}t$$

On Laplace transform $\omega(s) = T_e/(Js + B)$.
Here K_t is torque constant and J is inertia of motor.

3.1 Power Inverter Design

A various number of inverter topologies with the fault diagnosis and fault treatment have been presented in the previous papers [6]. On comparison and analysis, two three-phase standard six switch power inverters are used.

4 Fault Diagnosis and Fault Management

With the aim to improve the reliability of the system, dual redundancy brushless dc motor [7] is designed. Fault diagnosis and fault treatment are discussed in previous research papers [8] by various techniques. In this paper, the probability to occur faults in various available subsections is presented in the table for different cases (Table 1).

5 SIMULINK Diagram

The whole actuation system with the redundancy is shown in Fig. 4. It consists of primary chain and redundant chain, an LC filter to get smoother current waveform,

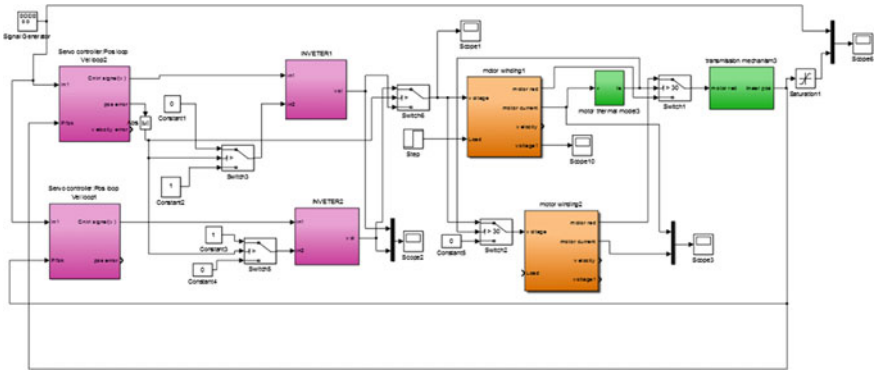


Fig. 4 Simulation diagram of EMA with dual redundancy logic

Table 2 Simulation parameters

Rating of BLDC motor	10 Kw
Voltage	100 V
Current	50 A
Inductance	160 μH
Armature resistance	0.1 Ω
Torque constant	0.08
Torque	8 Nm
Inertia *J*	2×10^{-5} kg m^2

and a mechanical transmission system where the electrical motor output is converted into mechanical system as position, and this position is taken as feedback signal.

The chosen parameters which are used in the simulation of this paper have been presented in the following tabular column (Table 2).

5.1 Simulation Results

In Fig. 5, the output of the actuator position follows the input signal. Due to the velocity feedback loop, the phase delay occurs between the input commands and output produced, which is shown in Fig. 6. When an overload fault occurs in motor Winding1, due to high rise in current, the motor could not drive the load and deviates from the output and comes out of scope which is shown in Fig. 7. And the respective fault currents and the power inverter voltages switching are shown in Figs. 8 and 9, respectively.

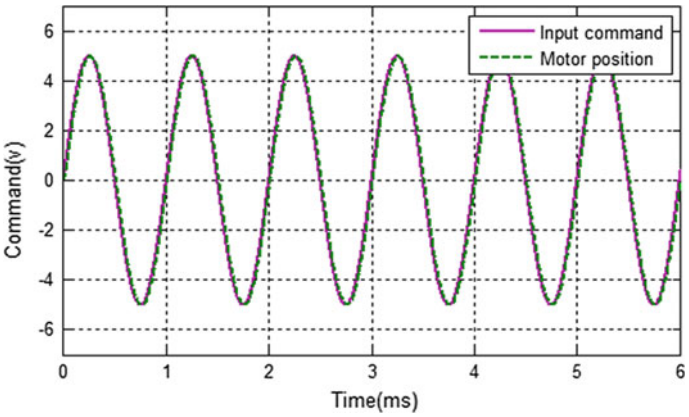


Fig. 5 Electromechanical actuator position under normal condition

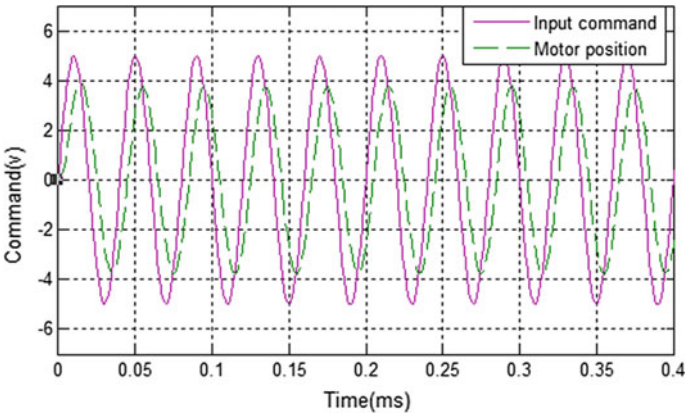


Fig. 6 Phase lag due to velocity loop

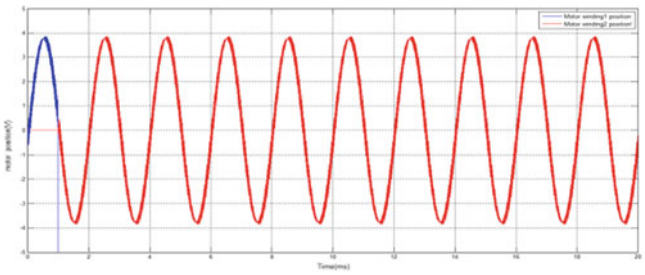


Fig. 7 When fault occurs in Winding1 due to overload

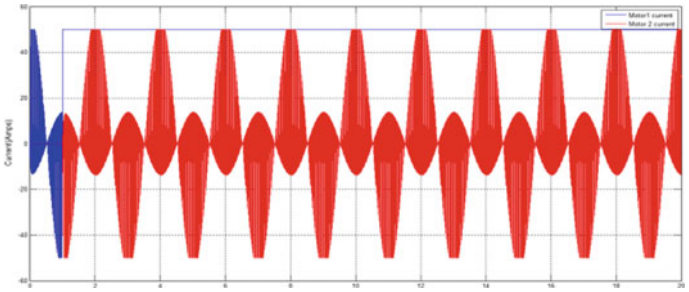


Fig. 8 Motor currents when W1 fails

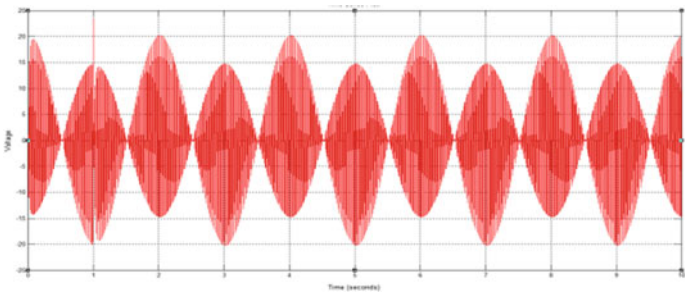


Fig. 9 Power inverter voltages

6 Conclusion

An electromechanical actuator comprised of position control system with the dual redundancy technique applied to both BLDC motor and power inverter has been designed and simulated, and results are presented. The fault diagnosis and treatment within the stipulated time are well explained through the presented output results.

References

1. Abd Elhafez AA, Forsyth AJ, A review of More Electric Aircraft, Aerospace sciences and Aviation technology. ASAT-13-EP-01
2. Garcia A, Cusido J, Rosero JA, Ortega JA, Romeral L (2008) Reliable electro-mechanical actuators in aircraft. IEEE Trans A&E Syst, Aug 2008
3. Jaiso H, Gopinath A (2011) A dual redundancy technique for electromechanical actuation system applied to launch vehicles. IEEE Trans, vol 11
4. Tuo Z, Liang D, Dynamic modeling and characteristic analysis for dual redundancy PM Brushless dc servo system
5. Ravikumar BV, Sivakumar K, Design of new switched stator BLDC drive to improve the efficiency of an electric vehicle

6. Wlchenko BA (2004) Fault tolerant of three phase AC motor drive topologies: a comparison of features, cost and limitations. IEEE Trans Power Electron 19(4):1108–1116
7. Jing L, Shanfeng L (2016) The balanced control of dual redundancy BLDC motor servo system. IEEE Trans, vol 978
8. Zhaoyang F, Jinglin L (2014) Fault diagnosis of Dual redundancy BLDC motor. In: 19th international conference on electrical machines and systems, Aug 2014

Part III
Communications

Comparative Analysis of Unstacked and Stacked Bow-Tie Antenna Structures for Wireless Applications

Kumari Laxmi and Amanpreet Kaur

Abstract The research article presents a comparative analysis of stacked and unstacked fractal antenna geometries. Four fractal antenna geometries, namely, an aperture coupled bow-tie antenna, a scaled aperture coupled bow tie with the same dimension on upper and lower patch, a stacked bow-tie geometry with smaller upper patch and larger lower patch, and a stacked geometry with larger upper patch and smaller lower patch with bow-tie antennas for both stacked layers are studied and analysed. The overall size of antenna is 57.84×57.84 mm^2 for all the four geometries. The ground used for an unstacked geometry is DGS, i.e. defected ground slot, and ground used for stacked geometry is partial ground. The effect of both the grounds has been shown and studied. The submitted antenna results for all the four geometries are analysed and studied in terms of impedance bandwidth, broadband gain and VSWR.

Keywords DGS · Partial ground · Sierpinski gasket fractal antenna

1 Introduction

Among all the technologies, wireless communication is the biggest contribution to mankind. Without the use of cables, wires or any other forms of electrical conductors information can be transmitted over a long distance. Antennas are a medium between transmitter and free space or between free space and receiver at transmitter or receiver side. In the last two decades, microstrip antennas (MSA) have been proposed as the best antenna because of their low profile and good output. Fractal microstrip antennas are small, efficient and reliable antennas that offer multiband characteristics. These

K. Laxmi (✉) · A. Kaur
Department of Electronics and Communication Engineering Department, Thapar Institute of Engineering and Technology, Patiala, India
e-mail: laxmi.punani93@gmail.com

A. Kaur
e-mail: amanpreet.kaur@thapar.edu

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_19

antennas show lower frequency band resonances with small sized. To increase the bandwidth with single fed structure microstrip antenna, an approach called stacking of multiple antenna layers is used. The stacked patch antenna works at multi-bands; the reason for this operation is that they offer multi-resonance properties [1].

In the current research paper, an analysis of improvement in the output in terms of impedance bandwidth when stacked microstrip antenna structures are used is presented. Four different geometries are chosen. First, an unstacked bow-tie antenna is presented and its results in terms of impedance bandwidth are presented. Then a stacked antenna structure with same patches on two stacked substrate layers is chosen and outputs are analysed. A comparative analysis is done by taking a 3/4th smaller patch above the standard patch and smaller patch below standard patch. Antenna geometries and outputs are discussed in the next section.

2 Antenna Geometry

(1) *Simple unstacked bow-tie geometry*:

The antenna geometry has three layers. The topmost substrate layer has a metal patch on in, i.e. of a bow-tie pattern formed by joining two Sierpinski gasket fractals at the apex. The middle layer is metallic ground which is 0.035 mm in height having cross-shaped and triangular-shaped defects in it [2]. The lower layer is a substrate at its bottom a feed line printed up to the centre of the patch that has a stub of length λ/4 attached to it (Fig. 1 and Table 1).

The bow-tie patch has two triangles joined at the apex. Each triangle is subjected to two iterations so as to make a Sierpinski gasket individually of second order. The design equation used to get a resonance of 4.31 GHz from the antenna input is mentioned in Eq. (1) (Fig. 2 and Table 2).

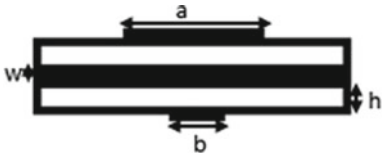


Fig. 1 Front view of unstacked antenna

Table 1 Dimensions of unstacked antenna

Parameters	Dimensions (in mm)
a	19.5
b	3
h	1.57
w	0.035

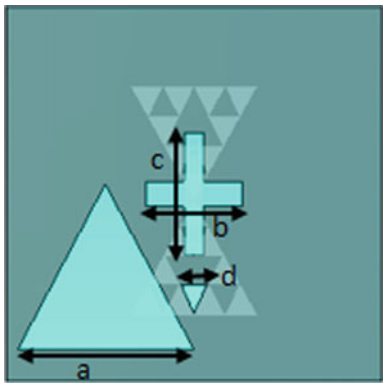


Fig. 2 Defective ground of unstacked bow-tie antenna

Table 2 Dimensions of defected ground

Parameters	Dimensions (in mm)
a	27.5
b	15
c	19
d	4

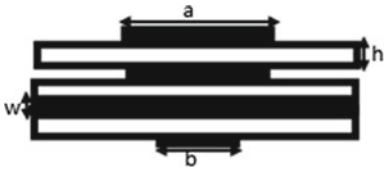


Fig. 3 Front view of stacked antenna

Table 3 Dimensions of stacked antenna

Parameters	Dimensions (in mm)
a	19.5
b	3
h	1.57
w	0.035

$$f_{m,n,1} = \frac{2c}{3a_{\text{eff}}(\varepsilon_{r\text{eff}})^{\frac{1}{2}}}\left(m^2 + mn + n^2\right)^{\frac{1}{2}} \tag{1}$$

$$a_{\text{eff}} = a + h(\varepsilon_r)^{-\frac{1}{2}} \tag{2}$$

$$\varepsilon_{\text{eff}} = \frac{1}{2}(\varepsilon_r + 1) + \frac{1}{4}(\varepsilon_r - 1)\left(1 + \frac{12h}{a}\right)^{-\frac{1}{2}} \tag{3}$$

(2) *Stacked structure with identical patches on active and parasitic layers* (Fig. 3 and Table 3):

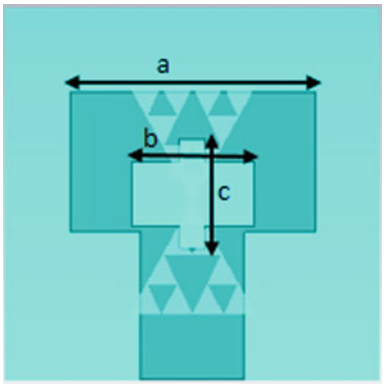


Fig. 4 Partial ground of stacked structure with same patch above and below

Table 4 Dimensions of Fig. 4

Parameters	Dimension (in mm)
a	37.84
b	19
c	17

Table 5 Dimensions of Fig. 5

Parameters	Dimensions (in mm)
a	37.84
d	21
e	10

The stacked microstrip antenna presented here has three layers of FR4 substrate. Above the existing geometry as mentioned in Sect. 1, an FR4 substrate with identical bow-tie pattern as earlier is printed. In order to increase the bandwidth of the two new resonances that arise from this configuration the ground is made partial with plus shape aperture slot modified to achieve the desired results [3] (Fig. 4 and Table 4).

(3) *Stacked structure with a parasitic patch 3/4th of the size of the standard patch below*:

This geometry is the same as the stacked structure mentioned previously. The parasitic patch is made 3/4th of the size of the standard patch below. Due to the scaling of the parasitic patch, the resonances get excited on higher frequency. The partial ground is same but the plus slot has been optimised as to get a good bandwidth (Fig. 5, Table 5).

(4) *Stacked structure with an active patch 3/4th of the size of the standard patch above*:

In this geometry, the active patch is made 3/4th of the size of the standard parasitic patch. Due to this higher frequency, resonances get excited. And a bandwidth of 1250 MHz is achieved (Fig. 6 and Table 6).

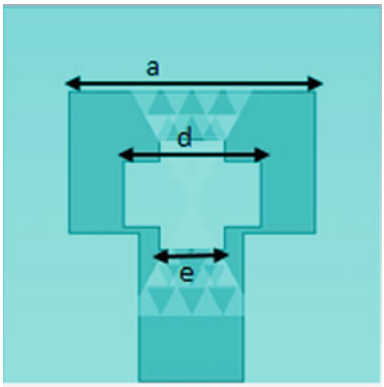


Fig. 5 Partial ground with optimised plus slot of stacked structure with smaller patch above

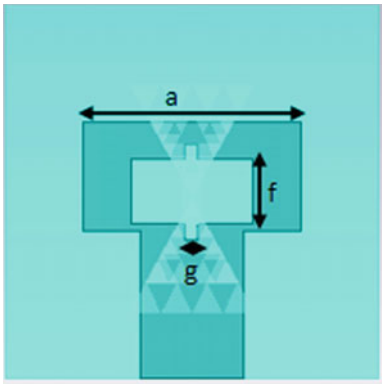


Fig. 6 Partial ground with optimised plus slot of stacked structure with smaller patch below

Table 6 Dimensions of Fig. 6

Parameters	Dimension (in mm)
a	37.84
f	10
g	2

3 Results and Discussions

This section presents the simulated results of the antenna in terms of return loss, radiation pattern and broadband gain.

3.1 Return Loss

Figures 7 and 8 show the return loss plot for all four geometries presented in the previous section. It can be observed from Fig. 7 that unstacked structure covers resonance

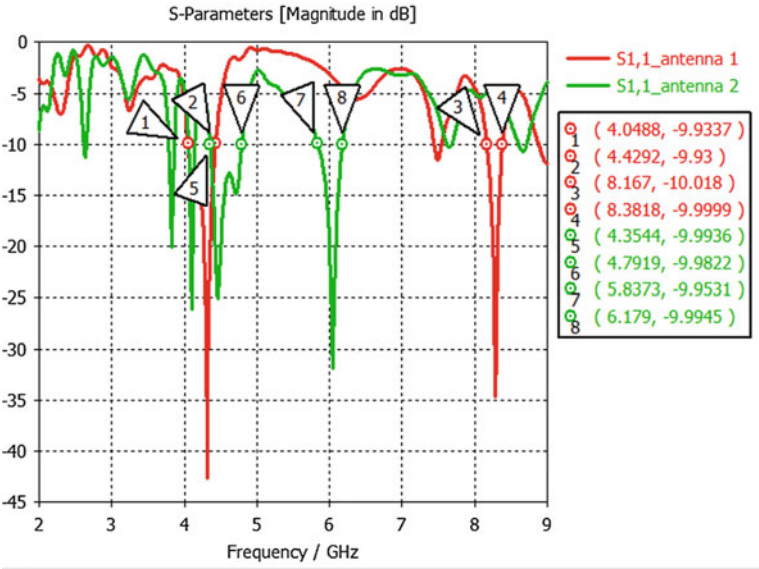


Fig. 7 Simulated values of S11 for first two geometries

of 4.31 and 8.28 GHz with bandwidths of 380 and 270 MHz. A stacked structure with the same dimension on the above and below patch shows four resonances at 3.82, 4.1, 4.46 and 6.04 GHz with bandwidths of 440, 90, 80 and 330 MHz, respectively. Figure 8 shows the stacked structure with patch 3/4th of the size of the standard patch below covers resonances of 3.846, 4.418, 4.937 and 7.928 GHz with bandwidths of 1120, 160 and 420 MHz. Fourth structure with patch below 3/4th of the size of the standard patch above shows resonances at 5.25, 6.00, 7.394 and 8.057 GHz with bandwidths of 190, 160 and 1250 MHz, respectively.

3.2 Broadband Gain

Figure 9 shows the combined plot of broadband gain of two geometries. As seen, the unstacked geometry has a peak gain of 2 dBi. The stacked structure with identical patches on active and parasitic layers has a peak gain of 5.3 dBi. The stacked structure with a parasitic patch 3/4th of the size of the standard patch below has a peak gain of 2.5 dBi. The fourth geometry with an active patch 3/4th of the size of the standard patch above has a peak gain of 4.0 dBi as shown in Fig. 10.

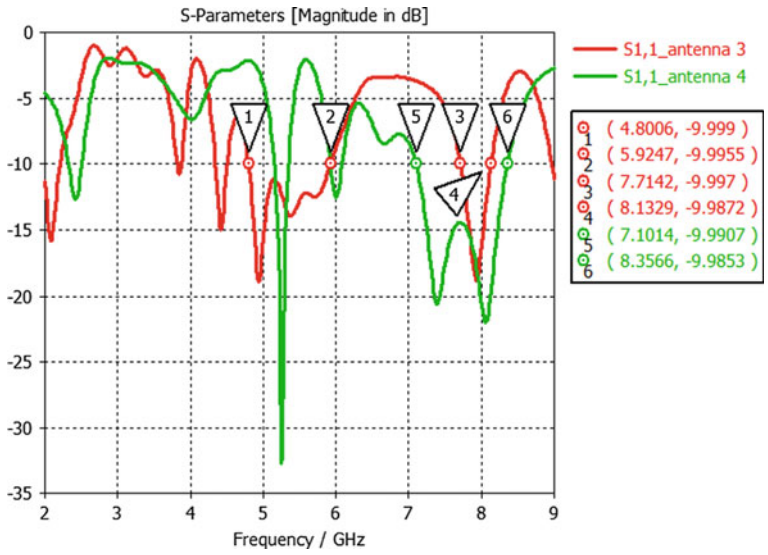


Fig. 8 Simulated values of S11 for next two geometries

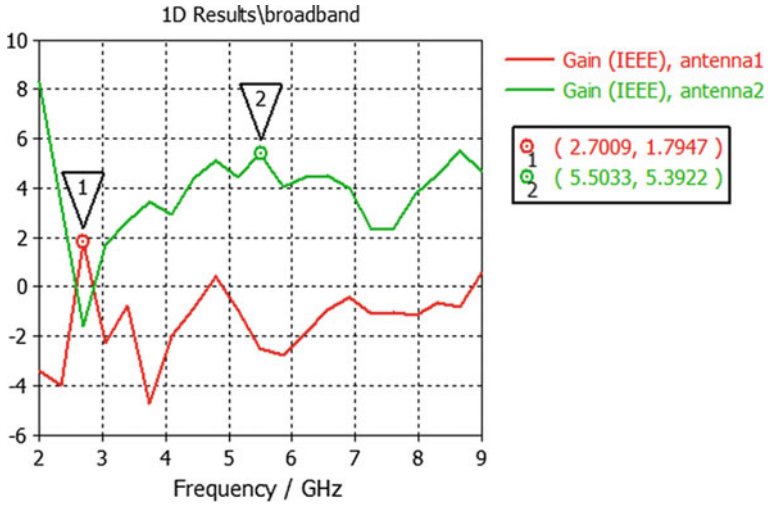


Fig. 9 Broadband gain for first two geometries

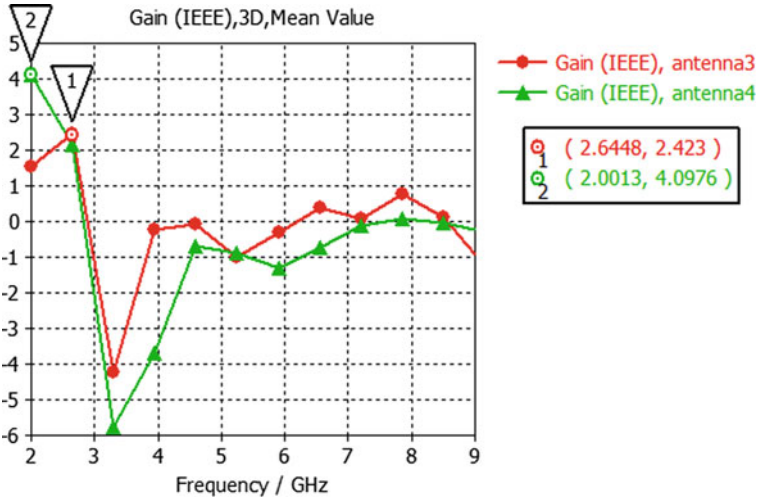


Fig. 10 Broadband gain for next two geometries

4 Conclusion

The paper presents a comparative analysis of four aperture coupled fractal microstrip antenna geometries. It is observed that unstacked geometry proposed above shows two resonances at 4.317 and 8.28 GHz with bandwidths of 380 and 270 MHz, respectively. A stacked structure with the same dimension on the above and below patch shows four resonances at 3.82, 4.1, 4.46 and 6.04 GHz with bandwidths of 440, 90, 80 and 330 MHz, respectively, because of partial ground. Stacked structure with patch 3/4th of the size of the standard patch below covers resonances of 3.846, 4.418, 4.937 and 7.928 GHz with bandwidth of 1120, 160 and 420 MHz, respectively. Fourth structure with patch below 3/4th of the size of the standard patch above shows resonances at 5.25, 6.00, 7.394 and 8.057 GHz with bandwidths of 190, 160 and 1250 MHz. It is also seen that a stacked structure with same sized antenna on both the driven and parasitic layers shows the highest gain and directional characteristics. The unstacked structure is suitable for WLAN and ITU satellite communication band. The stacked structures cover E, G and C bands, which are used for short range Wi-Fi, wireless broadband and some cordless telephone.

References

1. Yon H, Awang AH, Ali MT, Subahir S, Kamaruddin SN (2016) Comparative analysis for multilayer stacked substrates microstrip patch antenna. In: 2016 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE), Langkawi, 2016, pp 34–37

2. Goyal A, Kumar R, Gupta K, Singh DK (2015) A triple band stacked patch antenna with slotted ground structure. In: 2015 international conference on futuristic trends on computational analysis and knowledge management (ABLAZE), Noida, 2015, pp 556–560
3. Kumar J, Shirgan SS (2014) Compact partial ground plane 1 × 2 Patch Antennas. In: 2014 international conference on computational intelligence and communication networks, Bhopal, 2014, pp 33–37. https://doi.org/10.1109/cicn.2014.19
4. Wang J, Fralich R, Wu C, Litva J (1990) Multifunctional aperture coupled stack patch antenna. Electron Lett 26(25): 2067–2068. https://doi.org/10.1049/el:19901333
5. Kaur A, Khanna R (2015) A stacked sierpinski gasket fractal microstrip antenna for IEEE802.11a/b/UNII-1 and C band applications. In: 2015 IEEE 4th Asia-Pacific Conference on Antennas and Propagation (APCAP), Kuta, 2015, pp 321–323

An Efficient Method to Classify the Peer-to-Peer Network Videos and Video Servers Over Video on Demand Services

M. Narayanan

Abstract As one of the wildest emerging technologies, P2P has attracted attention on live streaming and VoD. Video plays an energetic part in communication and any kind of relaxing activity for entertainment. In order to provide reasonable service across all seasons, two machine learning techniques are used where the availability of server depends on the hits across the season. Popular videos are sorted out based on the greatest number of hits primarily and the recovery phase selects solitary or many similar cases from the preceding popularity videos that are stored. The updated/modified video records are reused as per query. In the revised phase, the present popularity record is updated. Finally, the updated popularity records are preserved in the retaining phase. Application of AODE algorithm results in grouping video server as seasonal and nonseasonal. The content of the video is categorized on the basis of the hearer's test at the commencing remains additionally scrutinized emerging in 90% clarity of classification.

Keywords Average estimators · Reasoning based on case · Categorizing video server · Peer-to-Peer networks · Seasonal-based videos · Video on Demand

1 Introduction

Video is believed as a kind of enjoyment that gives the vision of steady motion of a report across the representation of an arrangement of pictures and is recognized as a film. Until the nineteenth century, the world could not unravel its mystery, though the origin dates back to the second century in China. Gradually, television became the major means of entertainment, substituting newspaper and radio. One good reason for its fame above newspaper and wireless is the discernible charm that enables viewers to investigate deeper and be engrossed in the involvement.

M. Narayanan (✉)
Department of Computer Science and Engineering, Malla Reddy College of Engineering, Maisamaguda, Dhulapally Post, Secunderabad, Hyderabad, Telangana, India
e-mail: narayanan_baba@yahoo.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_20

At the end of the twentieth century, advancing movies appeared alongside the progress of the Internet and WWW. The foremost facilities, endowed through the Internet are browsing the web and transferring files. Nevertheless, the needs of the clients are not met with these services providing data about transcript, pictures, and exchange of documents.

The motivation for this study to find ways for transmitting live broadcast over the Internet to a PC is the accomplishment of radio and television. Subsequently, multimedia information like audio and video above the web are communicated. The entire multimedia data are distributed by using protocols meant for file download.

Case-Based Reasoning (CBR) algorithm is used for classifying popular videos and Averaged One-Dependence Estimators (AODE) algorithm is used for categorizing video servers into seasonal video server and non-seasonal video server. A literature review of P2P networks is included in Sect. 2, followed by proposed work in Sect. 3. Implementation is explained in Sect. 4 and finally, Sect. 5 comprises the conclusion.

2 Literature Survey

P2P (Peer-to-Peer) computing is the most desired overlay network distribution setup. In this system, super-peer network portrays a contemporary centralized topology pattern. Therefore, this overlay aids in improving the presentation of P2P requests, especially live streaming. Using Firefly algorithm, the problem of node failure is addressed in this work by the application of gossip protocol. From a collection of peers, it chooses the failover node to sustain the entire network performance [1, 2].

P2P Networks offer video services based on the seasons. Case-Based Reasoning algorithm is used for categorizing acquainted videos and Averaged One-Dependence Estimators (AODE) algorithm is used to categorize video servers on the seasonal and nonseasonal basis. Mature and local domains are classified as nonseasonal with maximum hits in every season. As a result, the audiovisual service engages specific server for providing services. Therefore, the P2P network provides quality service during all periods and depending upon the hits of the season the server is availed [3].

Management of video contents in P2P computing is discussed in this paper. Usually accepted data blocks are allocated in the cache and algorithms are used for replacing it. There is less data on account of nonavailability or less availability of old data. Multilevel cache solves this problem by saving the old data segments and therefore there is an increase in speed. Connectivity problems are addressed by bandwidth-based clusters [4].

Video streaming is considered to be one among the foremost common application of P2P computing and close peers facilitate for accessing information segments. Cache replacement algorithm replaces old data blocks with new data blocks. Multilevel cache is used to retain both new and old data segments. The success rate is increased as a result and the data segments are also maintained efficiently [5].

To examine P2P traffic and to find a competent resolution to tackle the alike employing Unsupervised Machine Learning way is the target of this paper. An expec-

tation–maximization (EM) algorithm is counseled to become the maximum probable estimation of peers. This is a type of artificial intelligence technique. The EM successive approximation will trade between an expectation (E) step and a maximization (M) step. Using the current estimation for the factors, expectation (E) step will create a function for the expectation of the log possibility calculation. Alternatively, maximization (M) step estimates the factors maximizing the feasible log-possibility found on the E step [6].

P2P computing has its own challenges, like offering quality service mainly due to network traffic issues. For reducing the traffic congestion, the traffic is analyzed by engaging adaptive routing and design observation. The traffic depends upon hidden listening peers on traffic statistics. To get the highest likelihood approximation of peers, an iterative technique called adaptive routing scheduling algorithm is suggested. It consists of two parts; the first part gives the next peer location and the second part explains about the exact response given to the node by the peer in the explicit adaption request (EAR) [7].

This research discloses the outline of mesh-based P2P Networks to evade overloading issues with the servers. This type of network addresses the problem of data download, as there exists neither master nor slave relationship between the clients and is suitable for live streams. The requested data is stored in the peers' cache. The header contains the latest data blocks and the tail contains the old [8]. The part selection subjects that external after broadcasting video contents above P2P webs are introduced. This algorithm periodically decides the pieces that are to be requested from other peers for downloading [9].

The protocol and algorithm suggested by Lloret et al. (2012) are methodical to enable selecting the best relevant video characteristics and highest coding methods to hold the streaming of videos. A server for controlling videos will choose the finest video server and also selects the algorithm and the web presentation [10]. The results are judged in contrast to those existing in the journalism and give a recommendation for effort on network supervision [11]

To tackle the P2P layered video streaming, Hua et al. (2013) give an illustration of a new block arrangement mechanism. The main function is about striking a good balance between the different aspects. The block arranging subject is next adjusted to a subject that uses the significant sum of the distributed video segments. Simulation is grasped out to assess the presence of the algorithms. Analysis proves that the suggested algorithm is successful in utilizing the bandwidth and by providing quality audiovisuals [12].

The suggested method is not probable to be utilized for green certification or legislation/regulation satisfaction because of its nature as an estimate. Further, the system is not meant for developing new products. However, the process can boost up monotonous product development techniques by providing immediate and satisfactory accurate evaluation method [13].

Context-aware peer invention and peer selection are considered on the basis of the request context information. The benefits of CA-P2P in accomplishing rapid detection and alliance have been demonstrated via experiments. CA-P2P is capable

of supporting current and evolving services like the commercial ad and catastrophe-related communication [14].

3 Proposed Model

3.1 Categorizing the Popular Video

Case-based reasoning is an area of artificial intelligence, which comprises of the hypothetical fundamental analysis, system development and practical application built on experience-based issue resolving. At present, the CBR is believed as one of the most enthusing and core scrutiny areas. It has the scope of generating diverse methods and applications and then effectively apply them for determining the problems in the specific area.

One or several appropriate cases (made earlier/stored experiences) are chosen to rectify the new complications that arise. The answers/responses from the cases are perused here. The key assumption in this approach is that the same issues have the same type of remedy. Subsequent to selecting the similar kind of issues, the responses from the preceding cases are adjusted to spiral into a solution for the current issue.

3.2 Categorizing Video Server Using Averaged One-Dependence Estimators

It is a classification learning probability-based algorithm used for the categorizing purpose. During the training sequence, a three-dimensional combined frequency chart is created where the benefits of the class indexes single dimension and the supplementary binary dimensions are accompanied by the benefits of the features. This system was built for dealing with characteristic independence obstruction of the well-known inexperienced Bayes classifier. This algorithm differs strikingly from the naive Bayes because of the fact that it builds more precise classifiers by slightly enhancing the amount of calculation.

Averaged one-dependence estimators (AODE) possess a greater interest in the enhanced efficiency in training and classification. The outcome and conversation segment encompasses the various categories of video server and the categorization of video server into seasonal- and nonseasonal-based video server in terms of their high categorizing effectiveness.

Each seasonal comprehensive one-dependence estimators in AODE is learnt and also the kind of video demand obtained from the video server. The kind of demand could be BASICS OF COMPUTER, C++, ORACLE, NS2, VB, MATLAB, VB.NET, LINUX, JAVA, or others. And the forecasting is created by forecasting averages of the whole season using one-dependence estimator. A one-dependence estimator is

crafted for every single attribute for cutting its intricacy and averting construction learning. The attribute is the basis for all features. The AODE classifier is stated below.

$$P(S|F) = \frac{\sum(\text{Seasonal}(Ai) \geq mP(S, F) \prod_{j=1}^{n} P(Fj|S, Fi)}{\sum(S \in F \sum(\text{Seasonal}(Ai) \geq mP(S, F) \prod_{j=1}^{n} P(Fj|S, Fi)}$$

where

P	Probability
Ai	Set of Training Instances
m	Threshold
Seasonal	Defined as the seasonal BASICS OF COMPUTER, C++, ORACLE, NS2, VB, MATLAB, VB.NET, LINUX, JAVA, or others
S	Video Server
F	Features ∈ Seasonal

The limited possibility is P ($S|F$), the grade of belief in S, assumed that the proposal F is correct. This is the concept that determines the possibility of an event assumed that another event has occurred either by theory, assertion or confirmation. The video server finds the plethora of regular videos which might be BASICS OF COMPUTER, C++, ORACLE, NS2, VB, MATLAB, VB.NET, LINUX, JAVA, or others. Compared to representative discovering algorithms, the examinations clarify that AODE achieves improved possibility assessment. For enhanced performance of probability estimation and managing a bigger list, the AODE must be improved. The drawbacks in a feature collection of AODE for possibility assessment are reviewed and examined in order to suggest an effective algorithm.

4 Result and Implementation

The development handles the categorization of video server and its details. AODE classifies Video Server as seasonal and nonseasonal. First, the video details are classified on the basis of the hypothesis which could be BASICS OF COMPUTER, C++, ORACLE, NS2, VB, MATLAB, VB.NET, LINUX, and JAVA.

By and large, the number of servers for classification is 200 with 272 videos in a plethora of arrangements like FFmpeg, mp4, DVD, avi, and mjpeg. Based on the number of hits for a particular video on particular days, the server is labeled as seasonal and nonseasonal. First, classification is carried out depending on the listener's test and by applying AODE it is again classified which yields higher clarity close to 90%.

The subsequent parameters are supposed for executing with Java programming. Declaration of parameters is embodied in Table 1 and the parameter terms are remarked as Seeds, Peers, Cache, Fragment time, Sendport, Receiveport, Capacity, Packets/sec. Each parameter contains the lowest and the maximum values that

Table 1 Declaring parameters

Parameter name	Min	Max	Threshold value
Seeds	2	50	25
Peers	2	100	50
Cache	1 mb	3 mb	3 mb
Fragment time	1 s	2 s	2 s
Sendport	1022	–	–
Receiveport	1900	–	–
Capacity	100	500	300
Packets per sec	1	5	3

Table 2 Listener test assessment

Sl. no	Video categories	No. of videos	Occurrence in no. of server
1	BASICS OF COMPUTER	33	11
2	C++	14	24
3	ORACLE	35	14
4	NS2	27	13
5	VB	27	14
6	MATLAB	14	27
7	VB.NET	6	5
8	LINUX	46	27
9	JAVA	70	25

can be realized by a Java program. For each parameter, the threshold rate is requested and larger aftermath attained alongside this value.

Table 2 denotes primarily 3 columns. Column 1 displays video categories according to the client request. Column 2 displays the quantity of video hits made by and column 3 denotes the subsequent videos accessible in servers.

Figure 1 has been ready on the premise of the hits generated throughout every season. For example, BASIC OF COMPUTER video gets 3273 hits in the odd semester and gets 101 hits in even semester.

Embodies an association upheld the server hit for e period and additionally the servers' hits are computed for every single period and drawn in Fig. 3.

Table 3 and Fig. 2 display the amount of hits for every video with the server in both semesters. Here is an attempt to classify servers as odd semester and even semester. Based on the semester result, the hit rate will vary. The subjects in the corresponding semester always have a solid hit rate. Along with corresponding student hit rate, the failed candidate hit rate will be combined to give a higher hit rate in the corresponding semester. The representations of acronyms utilized in Table 3 and Fig. 3 are as follows.

B BASICS OF COMPUTER; C-C++; O-ORACLE; N-NS2; V-VB

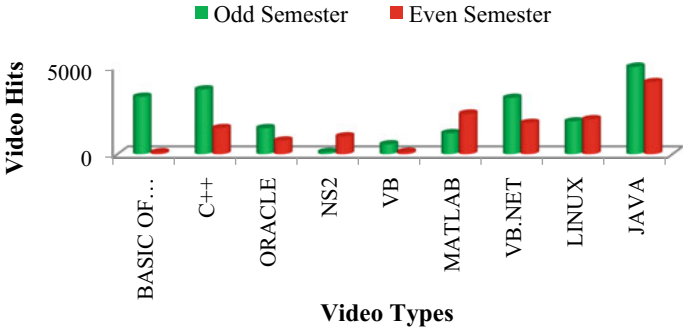


Fig. 1 Total no. of hit videos in complete season

Table 3 Total no. of hit videos in every season

Sl. no	B	C	O	N	V	M	VN	L	J
1	Odd	120	140	15	110	70	25	200	300
2	Even	170	190	166	160	120	75	250	350

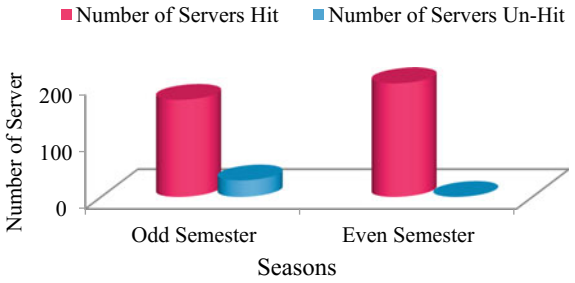


Fig. 2 Sorting on the basis of server hit on every season

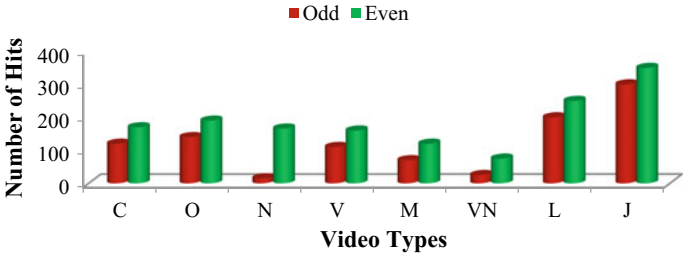


Fig. 3 Total no. of hit videos in entire season

M MATLAB; VN-VB.NET; L-LINUX; J-JAVA

5 Conclusion

By using the method of Retrieve, Reuse, Revise and Retain, case-based reasoning algorithm categorizes the popular videos. Averaged one-dependence estimators algorithm classifies the video server according to the period. The algorithm shows that the JAVA and LINUX domain has the greatest or stable hit in each and every season. On this basis, the video service allocates a certain server and that specific server provides service to customers. So, by applying these algorithms, the P2P network offers improved service during the various seasons and the servers are made available depending on the hits at that particular period.

References

1. Madeshan N, Chokkalingam A (2016) An efficient super-peer selection for Peer-to-Peer live streaming networks over video-on demand service. J Comput Theor Nanos 13(7):4606–4613
2. Madeshan Narayanan, Chokkalingam Arun (2014) An efficient super peer selection algorithm for Peer-to-Peer (P2P) live streaming network. J Theor Appl Inf Tech 70(1):1–8
3. Narayanan M, Arun C (2015) Categorize the video server in P2P networks based on seasonal and normal popularity videos using machine learning approach. In: Electronics and Communication Systems (ICECS), 2015 2nd international conference on, 1220–1228, IEEE, Coimbatore 2015
4. Narayanan M, Arun C (2014) An efficient technique for video content managing in Peer-to-Peer computing using multilevel cache and bandwidth based cluster. In: Signal Propagation and Computer Technology (ICSPCT), 2014 international conference on, pp 317–322, IEEE, Ajmer 2014
5. Narayanan M, Arun C (2012) An efficient method for handling data segment with multi-level caching over Video-on-Demand using P2P computing. Eur J Sci Rec 93(2):206–213
6. Narayanan M, Arun C (2015) To manage traffic in P2P networks using expectation-maximization approaches over Video-on Demand services. Int J Appl Eng Res 10(4):2958–2966
7. Narayanan M, Arun C (2015) Manage traffic in P2P live streaming using adaptive routing scheduling over Video-on Demand service. ARPN J Eng Appl Sci 10(13):5581–5587
8. Narayanan M, Arun C (2014) Outliers based caching of data segment with synchronization over Video-on Demand using P2P computing. Res J Appl Sci Eng Technol 7(21):4559–4564
9. Szkaliczki T, Eberhard M, Hellwagner H, Szobonya L (2014) Piece selection algorithms for layered video streaming in P2P networks. Discrete Appl Math 167:269–279
10. Lloret J, Canovas A, Tomas J, Atenas M (2012) A network management algorithm and protocol for improving QoE in mobile IPTV. Comput Commun 35(15):1855–1870
11. Kettig O, Kolbe HJ (2011) Monitoring the impact of P2P users on a broadband operator's network over time. IEEE Trans Network Serv Manage 8(2):116–127

12. Hua KL, Chiu GM, Pao HK, Cheng YC (2013) An efficient scheduling algorithm for scalable video streaming over P2P networks. Comput Networks 57(14):2856–2868
13. Jeong MG, Morrison JR, Suh HW (2015) Approximate life cycle assessment via case-based reasoning for eco-design. IEEE Trans Autom Sci Eng 12(2):716–728
14. Li H, Zhao C, Wu W (2014) Research on case-based reasoning search method. In: the 26th chinese control and decision conference (2014 CCDC) IEEE, 2360–2364

A Robust Wavelet Based Decomposition and Multilayer Neural Network for Speaker Identification

M. D. Pawar and Rajendra Kokate

Abstract The major aim of this research paper is to recognize and figure out the problem of recognizing a speaker from its voice, we propose a new methodology for feature extraction based on speakers pitch, stationary Wavelet, and multilayered Neural Networks. In this exercise, we designed a methodology to study for Text-Dependent Speaker Identification. Wavelet analysis comprises Stationary wavelet analysis, Continuous wavelet analysis, and discrete wavelet analysis, the classification module comprises an artificial neural network, General regression forming the decision through majority test/train result scheme. A performance test is conducted using the recorded database for text dependent and text independent. Stationary wavelet with the multilayered neural network has shown better accuracy and faster identification time compared with traditional MFCC, discrete, and continuous wavelet transform approaches.

Keywords Speaker identification and recognition · Stationary wavelet · Neural networks methodology

1 Introduction

Acoustic Communication is one of the fundamentals required as a prior condition for the existence of human society, nowadays application based on information technologies are growing faster with an increase in these applications for secure transaction. There are many areas where the secure transaction is needed for remote access ter-

M. D. Pawar (✉)
Department of Electronics and Telecommunication Engineering, Maharashtra Institute of Technology, Aurangabad 431005, India
e-mail: manjupawar2583@gmail.com

R. Kokate
Department of Instrumentation Engineering, Government College of Engineering, Jalgaon 425001, India
e-mail: rdkokate3394@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_21

minals, phone banking, and biometric identification techniques for a secure access to many applications. From speech samples taking alone could be an accurate conclusion made whether the speaker is male or female, adult or child. Speaker recognition divides person verification and identification [1]. The earlier work or published data exists with different terms for speaker identification including voice verification and recognition system [2].

To execute the planning of speaker identification/verification system requires the human voice samples to convey meaning to the machine, the human voice consists of sound samples which are characterized by behavior and physiology of the individual. For instance, utterance produced by individuals are from the same vocal tract [2, 3] to have a typical pitch range along with characteristics associated with gender. This results in a highly correlated speech signals for any particular speaker identification system that can be categorized into text independent and text dependent [4], pitch frequency is the one kind of quality that makes easy to identify sound range, which will be calculated by F0 = 1/T0 which also varies person wise, pitch frequency of child is as extreme as 400 Hz, an adult male's pitch frequency is as low as 100 Hz, [5, 6]. The speaker identification system uses the following techniques or combinations, like linear Predictive system (LPC), LPCC, Mel-frequency cepstral coefficient (MFCC) [5, 7] Wavelet transform method, Vector quantization method, Dynamic time wrapping (DTW), Hidden Markov model (HMM), Gaussian mixture model (GMM) [8], etc. According to the feature, the pitch frequency varies in different contexts and time, mel-scale algorithm dynamically construct mel filter bank based speaker's pitch frequency by detecting the variations of the pitch [8, 9]. The training time and equal error rate between LS-SVM and conventional SVM is compared, to improve HMM performance for speaker identification/verification system used a combination of (HMM), (VQ). Basically, vector quantization is used for data compression purpose and distance measures which are vector mapping process as classifiers in speech recognition system [10–12]. The motivation is to develop generalized, accurate, and less recognition time creating the most efficient system which can work on a large dataset, using wavelet transform for feature extraction and multilayered neural network as classification purpose, the primary goal is to develop an understanding of the acoustic features to be used in speaker identification/verification system. For speech and speaker identification system, artificial neural network (ANN) model have been effectively worked, radial basis network, general regression (GRNN) models, multilayered perceptron, ADALINE and developed neural network techniques with cascade neural network [13–15]. This paper's setup is as follows:

Section 1: describes the general introduction for the speaker identification/verification methodology.

Section 2: discusses the proposed methodology with feature extraction using wavelet and its types.

Section 3: Database.

Section 4: describes the multilayered neural networks and its types.

Section 5: Discussed experimental results and conclusion.

2 Proposed Method

The proposed and developed method in this paper, Stationary wavelet transform with neural network is explained in this section. The introduced methodology based on numbers of the block as shown in Fig. 1 which is into four stages: Recording the input speech/voice samples or dataset, feature extraction, classification, and final stage is correct speaker identification (Fig. 2).

2.1 Feature Extraction by Wavelet

Theoretical Overview

In the wavelet analysis, the basic wavelet transform is a generalized wavelet decomposes and gives richer signal analysis. Wavelet transform provides the solution to represents a signal in time as well as frequency domain, or wavelet transform is computed when the signal analyzed by multiplying function to signal [16, 17] mother

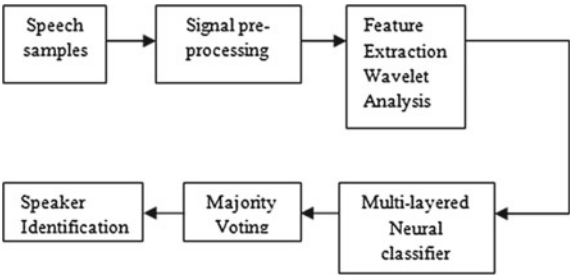


Fig. 1 Proposed system

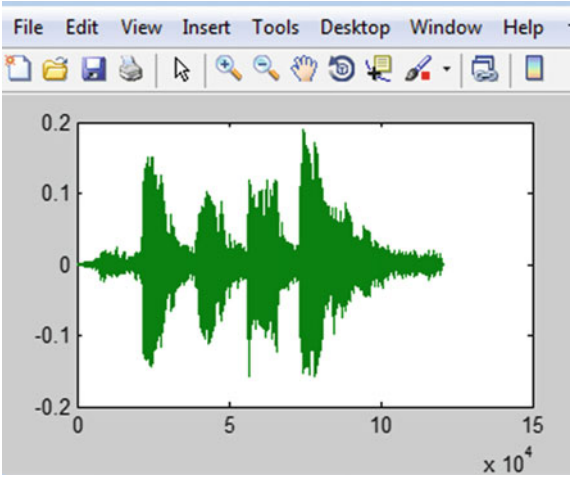


Fig. 2 Speech sample

wavelet function is used for signal transformation which can be further used in translation and scaling for transformation, continuous wavelet and discrete wavelet transformations are the basic two types, continuous wavelet transform signal equation can be represented by the following equation:

$$XWT(\tau, S) = \frac{1}{\sqrt{s}\int x(t)}.\varphi\left(t - \frac{\tau}{s}\right)\mathrm{d}t \tag{1}$$

where $x(t)$ is signal to be analyzed and (t) is mother wavelet function, a role related to the location of the window as the window is shifted across the signal and scaling parameter relates with frequency.

2.2 Discrete Wavelet Transform

Wavelet transform decompose a signal in high as well as low frequency components by passing the signal in a high bandpass filter and a low bandpass filter successively [18, 19]. High frequency represents global information of the signal whereas low frequency represents local information or frequency information, capturing this low-frequency information is important here because it contains the acoustic characteristics of speakers shown by following equation:

$$\varphi j, k(t) = \frac{i}{\sqrt{s}}\varphi\left(t - k\tau 0s\frac{0}{s0}\right)$$

τ Translation and s is scale parameter in wavelet transform for signal translation parameter.

$$\varphi j, k(t) = \frac{i}{\sqrt{s}}\varphi\left(t - k\tau 0s\frac{0}{s0}\right) \tag{2}$$

The decomposition shown by

$$Y \text{ high } (k) = \sum\nolimits_n x(n).g(-n + 2k) \tag{2.1}$$

$$y \text{ low } (k) = \sum\nolimits_n x(n).h(-n + 2k) \tag{2.2}$$

where y high$[k]$ and y low$[k]$ are the output of the low-pass and high-pass filter respectively.

2.3 Continuous Wavelet Transform

Continuous wavelet transform is basically used as a continuous wave processing tool for calculating and showing stationary and nonstationary signals in the time-varying frequency spectrum, mathematically represented by

$$F(w) = \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-jwt}\,\mathrm{d}t \tag{2.3}$$

Multiplied by a sinusoidal frequency, and constituent transform resulted in Fourier coefficients, [20], whereas in CWT 'C' is scale,

$$C(\text{Scale, Position}) = \int f(t)\varphi(\text{Scale, Position}, t)\mathrm{d}t. \tag{2.4}$$

2.4 Stationary Wavelet Transform

The discrete wavelet transform method is frequency-invariant transform and to restore translation invariance is slightly different to discrete-level transform, called decimated discrete wavelet transform also known as stationary wavelet transform (SWT), which performs possible decomposition at different levels of signals [21, 22]. Figure 3 shows the basic structure of stationary wavelet transform (SWT) having different decomposition levels, for a given maximum level j, denoted by $j = 0$ or 1 as per the choice of even or odd indexed elements at j step, at every decomposition level labeled by a sequence of 1 and 0's. All calculated transformation is known as decimated discrete wavelet which is again possible to calculate all decimated values for a given speech signal without filtering, by computing detail and approximation coefficients for every sequence. The stationary Wavelet transform is very simple and close to the discrete wavelet transform, in stationary wavelet given signal is obtained by convolving with the appropriate filters downsampling, which is done at level 1 only, whereas detail coefficients and approximation coefficients are both of size N at level 1. The basic step j involves approximation coefficient at level $j - 1$, with upsamples can be seen in the following figure [23, 24].

Features used in the proposed system

Wavelet coefficient provides information about the signal in frequency and time, the following are features used in this work:

1. Mean of absolute value for the approximation and each level of detail coefficients with frequency distribution information provided,
2. The standard deviation at each level and amount of change in frequency distribution information provided by these features.

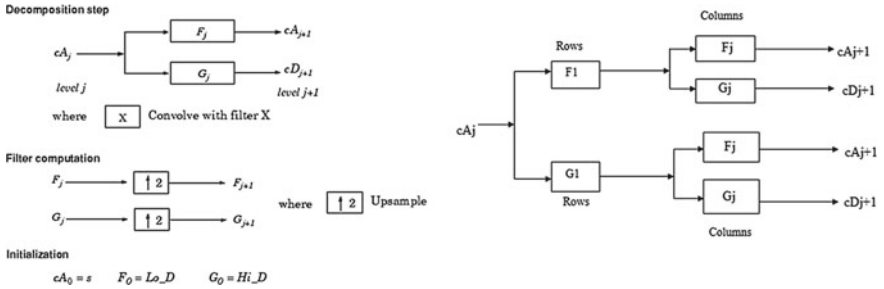


Fig. 3 Stationary wavelet transform

3 Database

For speaker identification system, the function of feature extraction and classification is performed using wavelet and its types with multi-neural network expert system. To implement this system speech samples for text-dependent contain 10 speech samples per 20 speakers, a PC microphone is used to record speakers voice. The word "Jordon Kingdom" from 40 speakers (20 Male and 20 Female) of 400 utterances and each speech was recorded 10 times by the speaker, To record the clips, used a microphone and recording software called Sound Forge to converted into.wav files with 44 kHz sampling frequency. In-house dataset contains speech samples of institute's students which contains 10 speech samples per speaker for 40 speakers and Sound Forge software was used to take the speaker's voice in.sfg file and converted into.wav file.

4 Classification

Figure 4 shows the architecture of a multilayered network using FFBPNN which consists of three layers namely the input layer, hidden layer, and output layer and the neurons are arranged along these layers [25]; more than two neurons can be combined in one layer and appropriate network could contain two or more layers, in the proposed work each input vector x1 is allied with each single neuron input through weight matrix w1, weighted inputs and bias function is used to form its owner scale output, bias function is used as a threshold which is associated with neuron in the input layer, various input neurons are taken together from *s*-element net input vector for effective mapping of input values to output values.

Standard multilayered network is backpropagation which is a gradient descent algorithm, as this is the Widrow-Hoff learning rule, and the backpropagation algorithm is as follows [19, 26 and 27].

Step 1: Initialize the weight to small random values.
Step 2: do the following steps for each training pair

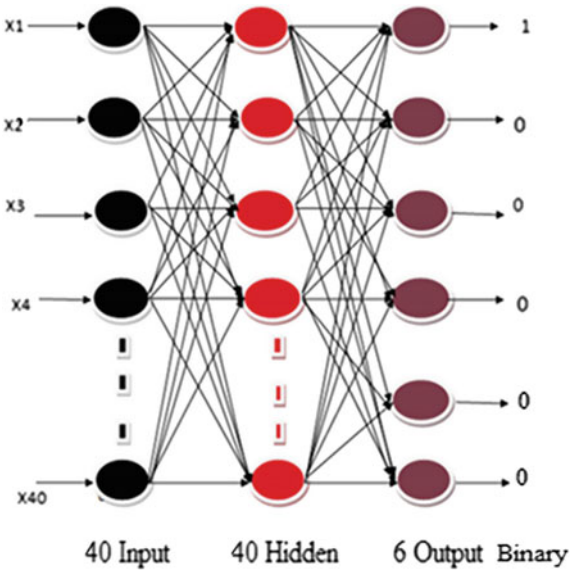


Fig. 4 Architecture of multilayered network

Step 3: Feed-forward, each input units receive input signals and transmit these signals to all layers,
Step 4: Each hidden unit (z_j, $j = 1, \ldots, p$) sums its weighted input units, apply activation function and calculate the output.
For training, we need to define training rule and use delta learning rate,

$$\Delta W_i = \eta(D - Y)X_i \tag{4.1}$$

where η is learning rate, 'D' is desired output and Y is actual output.
Step 5: At the last stage, test the stopping condition, where error value become zero with numbers of epochs.

Next term is GRNN, Generalized regression, which is introduced by Specht in 1991 [27] widely used in most of the identification/verification tasks, as the basic principle of general regression neural network needs training data to train itself by containing input–output mapping, in this calculated output using weighted average output training database where weight is calculated using the Euclidean distance, if distance or weight is very large then weigh will be very low and if it gives more weight to output means the distance is very small. Generalized regression neural network is a learning algorithm used for deriving continuous variables like transient samples in the sound sample [28–30]. Extracted feature vectors from the test signal are fed to all the ANNs in parallel and three classification outputs are calculated corresponding to the three classifiers used here. Figure 6 shows the linear regression of input and output performance values, the majority voting scheme is employed for the classification results of the ANNs (Fig. 5).

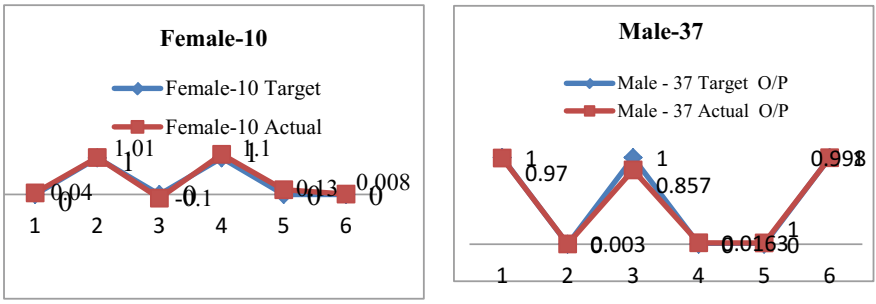


Fig. 5 Performance of neural network classification female-20 and male-3

5 Result and Discussions

The experimental setup was as follows: In-house dataset, a laboratory in the institute contains speech samples of institute's students. This contains 10 speech samples of 40 speakers. Each saved sound clip was a partially spoken word present for a particular English word (Jordon Kingdom) speakers age is varied between 19 and 24 years and included 20 males and 20 females, Sound Forge software was used to record speaker's voice in.sfg file and converted into.wav file. As the system is "Text-Dependent" system, 10 same utterances were used to record a speaker's sound sample. Speech signals were recorded via PC-sound card with a spectral frequency of 44 kHz and sampling frequency of 8000 Hz, bit rate of 256 Kbps. Table 1 shows all the related parameters for the best performance result and their description in the neural network.

Table 1 Parameters used under neural network

Function	Description
Type of network	Feed-forward
No. of epochs	50
Performance function (MSE)	10–5
Differential transfer function	Tan-sigmoid, Sigmoid, Linear
Number of neurons in layers	40—inputs, 40—hidden, and 6—outputs
Training function	Levenberg–Marquardt
Minimum performance gradient	10–10
Goal_error	10e-5
Performance function	MSE

5.1 Performance of Stationary Wavelet Transform with FFBPNN and Regression on Dataset

In the current study, feed forward backpropagation multilayered network is utilized. For the proposed system 40 input sound signals in that 20 male sound signal and 20 female sound signals of 400 utterances which are extracted by using Stationary wavelet and these extracted features designed with target (T) are designed as six binary digits for each feature vector and fed to FFBPNN as a classifier. Table 2 shows the six-digit binary target output for 40 features in that 1–20 are binary set target output for female and from 21 to 40 are male target output. Decided way to

Table 2 Target binary matrix

target <6x40 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
2	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
3	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0
4	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
7																																								

Table 3 Output binary actual output for female

y2 <6x40 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.9944	0.4555	0.6708	0.4864	1.0274	0.3882	0.9962	0.0391	0.7152	0.0404	0.4969	0.3951	0.5254	0.4137	0.6431	0.2467	0.8272	0.6796	1.0396	0.0098
2	-0.0038	0.6264	0.8319	0.4625	-0.0397	0.5744	0.9977	0.4802	0.0174	1.0505	0.7696	0.2831	0.4348	0.6681	0.6659	0.1560	0.2580	0.8975	0.9181	0.0072
3	-0.0051	0.0415	0.3421	0.6729	0.9971	0.4943	0.9969	0.5117	0.4873	-0.1382	0.3320	0.9496	0.6317	0.3473	0.5475	0.4288	0.5579	0.2604	-0.0710	0.9689
4	-0.0171	0.1744	0.0127	0.2580	0.1706	0.1091	-0.0107	0.3865	1.1355	1.1328	0.7683	0.9741	0.6633	0.8905	0.7375	0.3752	0.2115	-0.0274	0.0420	0.0289
5	-0.0142	0.3635	0.3685	0.3167	-0.0275	-0.0288	-0.0083	0.0351	-0.2336	0.1330	0.5791	0.2558	0.0106	0.1818	0.2841	0.8187	0.6320	0.2807	1.0074	1.0276
6	-0.0017	0.1191	-0.0028	0.2370	-0.0217	0.5786	-3.3486...	0.4679	0.0162	0.0083	0.0053	0.0046	0.5700	0.0618	-0.0192	-0.0048	0.0943	0.0077	0.0254	0.0027
7																				

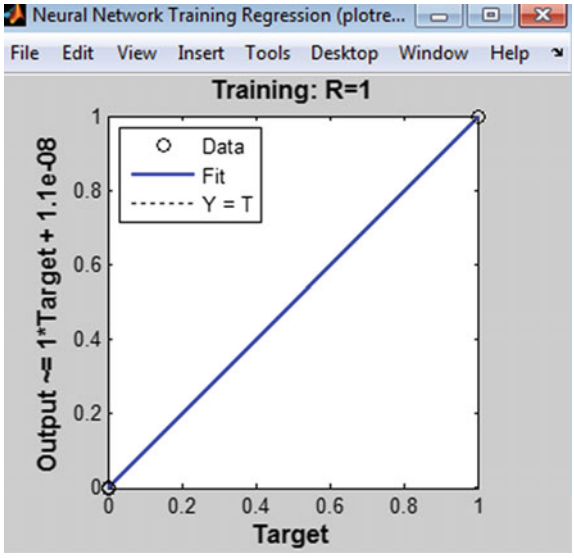


Fig. 6 Generalized regression

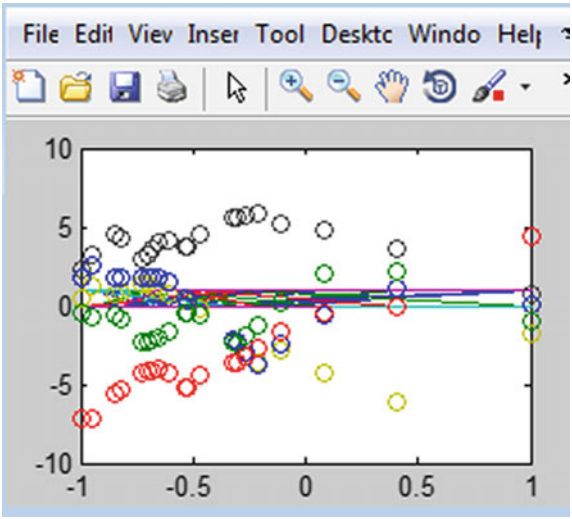


Fig. 7 Speech samples at training

arranged desired output matrix is to use the number of columns which is equal to the number of speakers, binary-decoded columns were used where a binary value of each column shows speakers number as an identity, Fig. 7 shows training of epochs with input speech signals for 390 samples. For the present system, 20 input sound samples for female and 20 input voice samples for male with target (T) is designed as a six binary digits for each input features vector. Table 3 shows the table of the actual output value of the female feature vector of six binary digits and Fig. 5 represent the performance output result of Female-10 and Male-37 with target input and actual output values. During supervised training, a set of weights are described to find out to converge on which the value is needed to produce the correct training response called network training convergence (Figs. 6 and 7).

6 Conclusion

In the current methodology, different types of wavelet transforms and multilayered neural network based approach have been used for speaker identification/recognition system. SWT is successfully used to extract feature in order to build a dynamic identification system.

The main purpose of the multilayered neural network is to design a model that strongly map input with output using an in-house created database so that the model can be used to produce the actual output when the suitable output is unknown. Tables 4 and 5 shows the comparative result with other methodology. Feature extraction was the main task for speaker identification/recognition system, which involved speaker-specific features extraction from a speech signal in the stationary wavelet transform

Table 4 Output result comparison

Class	Database	DWT at level 1	CWT	SWT at level 1	Std as threshold	% Accuracy using DWT	% Accuracy using SWT
Male−Female	40 (400 utterances)	34	30	38	0.09/0.04	85	95
Male	20 (200)	17	13	19	0.0982	85	95
Female	20 (2000	17	17	20	0.0412	85	100

Table 5 Comparative result

Parameters	Class	Database (voice samples)	Identified	% Accuracy
Standard dev.	Male–Female	40	38	95
Mean		40	28	70
Max.		40	33	82
Min.		40	32	80

domain. The accuracy of the proposed method depends upon the feature vector was which needed to build the speaker's model for classification purpose.

The speech dataset used for analysis contains ten samples per of twenty male and twenty female voice profiles. Proposed method stationary wavelet transform—Neural network, compared to the number of introduced methods like MFCC-NN, CWT-NN, DWT-NN using text-dependent recording from a combination of 40 speakers under different noise levels is shown in Table 5, Experimental analysis shows that the proposed method had a high recognition rate for verification and identification, in comparison to other published methods results indicated that the proposed method is superior with the highest recognition with the identification system used 40 feature vectors containing 400 utterances of 40 people in that 20 male and 20 female speakers. The presented speech corpus reached 98% accuracy.

References

1. Kumar P, Lahudkar SL (2015) International Journal on Recent and Innovation Trends in Computing and Communication 3(4):2106–2109. ISSN: 2321-81692106, IJRITCC, Apr 2015
2. Sarikaya R, Pellom BL, Hansen JHL (1998) Wavelet packet transform features with application to speaker identification. In: Proceedings of the IEEE Nordic Signal Processing Symposium, pp 81–84
3. Dan Z, Zheng S, Sun S, Dong R (2008) Speaker recognition based on LS-SVM. In: The 3rd international conference on Innovative computing Information and Control ICICIC, 2008, pp 01–04
4. Wang P (2005) Feature extraction bases on Mel-Scale wavelet transform for heart sound analysis. In: Proceedings of the engineering in medicine and annual conference, Shanghai, China, pp 01–05
5. Abdalla MI, Ali S (2010) Wavelet-based Mel-frequency cepstral coefficients for speaker identification using hidden Markov models. J Telecomm 1(2):16–21

6. Kekre HB, Kulkarni V (2010) Speaker identification by using vector quantization. Int J Eng Sci Technol 2(5):1325–1331
7. Wu J-D, Lin B-F (2009) Speaker identification using discrete wavelet packet trans-form technique with irregular position. Expert Syst Appl 36:3136–3143
8. Suvarna Kumar G, Prasad Raju KA, Rao M et al (2010) Speaker recognition using GMM'. Int J Eng Sci Technol 2(6):2428–2436
9. Speaker Identification based on GFCC using GMM. Int J Innov Res Adv Eng (IJIRAE) 1(8):224. ISSN: 2349-2163. www.ijirae.com (2014), IJIRAE—All Rights Reserved
10. Srinivasan A (2012) Speaker identification and verification using vector quantization and Mel frequency cepstral coefficients journal of applied sciences. Eng Technol 4(1):33–40. ISSN: 2040-7467 © Maxwell Scientific Organization, 2011 Published: 01 Jan 2012, pp 110–114
11. Kabir A, Ahsan SMM (2007) Vector quantization in text dependent automatic speaker recognition using Mel-frequency Cepstrum Coefficient. In: 6th WSEAS international conference on circuit systems, electronics control and signal processing, Cairo, Egypt, 29–31 Dec 2007, pp 352–355
12. Sunitha C (2015) Speaker recognition using MFCC and improved weighted vector quantization algorithm. Int J Eng Technol (IJET) 7(5):1685–1692. ISSN: 0975-4024
13. Dhonde SB (2015) Speaker recognition system using Gaussian mixture model. Int J Comput Appl 130(14):0975–8887, 38 Nov 2015
14. Pawar RV, Kajave PP, Mali SN (2005) Speaker identification using neural network. World Academy of Sci Eng Technol 12:31–35
15. Revada LKV, Rambatla VK, Ande KVN (2011) A novel approach to speech recognition by using generalized regression neural networks. IJCSI Int J Compute Sci Issues 1:483–489
16. Lu W, Sun W, Lu H (2009) Robust watermarking based on DWT and non-negative matrix factorization'. Comput Electr Eng 35:183–188
17. Wu JD, Lin BF (2009) Speaker identification using discrete wavelet packet transform technique with irregular decomposition. Expert Syst Appl 36(2):3136–3143
18. Vetterli M, Kovacevic J (1995) Wavelets and sub band coding. Prentice-Hall, New Jersey
19. Shukla A, Tiwari R, Hemant Kumar M, Kala R (2009) Speaker identification using wavelet analysis and modular neural networks. J Acoust Soc India (JASI) 36(1):14–19
20. Daqrouq K, Abu Hilal T, Sherif M, El-Hajjar S, Al-Qawasmi A (2009) Speaker identification system using wavelet transform and neural network. IEEE Xplore, 24 Oct 2009, pp 559–564
21. Hariharan M, Yaacob S, Hasrul MN, Wei OQ, Speech emotion recognition using stationary wavelet transform and timbral texture features, School of Mechatronic Engineering, University Malaysia Perlis (UniMAP)
22. Hariharan M, et al. (2013) Objective evaluation of speech dysfluencies using wavelet packet transform with sample entropy. Dig Sign Process A Rev J 23:952–959
23. Sindhu R, Siew-Chin Neoh, Hariharana M (2014) A hybrid expert system approach for telemonitoring of vocal fold pathology. Int J Medical Engineering and Informatics 6(3):218–237.
24. Keskes H, Braham A, Lachiri Z (2013) Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM. Electr Power Syst Res 97:151–157
25. Ge Z, Iyer AN, Cheluvaraja S, Sundaram R, Ganapathiraju A (2017) Neural network based speaker classification and verification systems with enhanced features interactive intelligence. In: Intelligent systems conference, 7–8 Sept 2017, London, UK Inc., Indianapolis, Indiana
26. Ye J (2004) Speech recognition using time domain features from phase space reconstructions. Ph.D thesis. Marquette University Milwaukee, Wisconsin
27. Specht DF (1991) A general regression neural network'. IEEE Trans Neural Network 2(6):568–576
28. Amrouche A, Rouvaen J (2006) Efficient system for speech recognition using general regression neural network. Int J Intell Technol 1(2):183–189
29. Almaade N, Aggoun A, Amira A (2015) Speaker identification using multimodal neural networks and wavelet analysis. IET Biom 4(1):18–28

30. Ganchev T, Tasoulis D, Vrahatis M, Fakotakis D (2007) Generalized locally recurrent probabilistic neural networks with application to text-independent speaker verification. Neuro Comput 70:1424–1438
31. Wang JC, Yang CH, Wang JF, Lee HP (2007) Robust speaker identification and verification. Taiwan IEEE Comput Intell Mag 2:52–59
32. Nijhawan G, Soni MK (2014) Speaker recognition using MFCC and vector quantisation. Int J Recent Trends Eng Technol 11(1):211–218

A Novel Multiband Star-Shaped Dipole Antenna for GSM and LTE Application

Gunjan Parihar and Milind Fernandes

Abstract A novel multiband dipole antenna for GSM and LTE application is presented in this paper. The antenna is comprised of two radiation dipoles which are orthogonal to each other, a pair of microstrip stubs and a pair of coaxial cable. The antenna is designed for dual-band operation (i.e., at 1.45 and 2.6 GHz) and multiband operation (i.e., 0.9, 1.4, 2.1 and 2.6 GHz). Also, an additional spike is introduced between each half of the dipole antenna which further achieves multiband operation at 0.94, 1.42, 2.19 and 2.76 GHz. The design of the antenna element has been completed using IE3D software. The antenna is placed upon and under the FR4 substrate, having a dielectric constant value of 4.4 and a thickness of 0.8 mm.

Keywords Star-shape · Dual-band · Multiband · Dipole antenna

1 Introduction

In recent years, there has been a rapid and diverse development of communication services. Also, it has become necessary to have a multiband antenna because every wireless carrier or nation uses different frequency bands. These multiband antennas are required to support wireless communication as shown in [1–5]. Global System of Mobile Communication (GSM) has 900 MHz as its working frequency and Long-Term Evolution (LTE) has 2600 MHz as its working frequency. Both GSM and LTE are part of these communication services and technologies.

The antenna used in [6–15] are all dual-polarised antennas. Their structure differs only in the basic design and in the manner in which they are fed. The structure in [7] is in shape of bow-tie which is orthogonally placed. It is excited by a microstrip stub. Thus, it has planar configuration and is lightweight. Also, obtains an impedance

G. Parihar (✉) · M. Fernandes
Goa College of Engineering, Ponda, Goa, India
e-mail: gunjanparihar141994@gmail.com

M. Fernandes
e-mail: mf@gec.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_22

bandwidth of 45%. The different parts in antenna structure of [8] are orthogonal dipoles which is part of the main radiator, dipoles are excited with the help of baluns, reflector and a bedframe. The bandwidth achieved is 48%. *Y*-shaped structures in [9] are used to feed the dipoles. Both of them are printed on FR4 substrate. A wide 45% bandwidth is obtained. Antenna configuration in [10] comprises of the folded dipoles along with strip lines as the part of the main radiator, feeding is through *Y*-shaped structures and the coaxial cables. Stable radiation pattern with a high cross-polarisation discrimination (XPD) over its frequency band is achieved. The major difficulty with the butterfly-shaped antenna in [11] is its narrow bandwidth of 21% and its size which restricts its usage.

In this paper, a novel structure for GSM and LTE operation is proposed. The antenna is designed for dual-band and multiband operation. Also, the characteristics of the antenna are observed after the introduction of an additional spike between each half of the dipole antenna. Moreover, this structure is simple and planar and hence easy to be manufactured. Completion of the dipole antenna design is done using IE3D software.

2 Star-Shaped Dipole Antenna Element

Figure 1a shows the basic configuration of the proposed star-shaped antenna. The antenna is comprised of two radiation dipoles, a pair of microstrip stubs and a pair of coaxial cable. Here the radiation dipoles and the feeding microstrip stubs are placed upon and under the FR4 substrate, having a dielectric constant value of 4.4. Thickness of FR4 substrate is $t = 0.8$ mm. Each pair of dipoles consists of one element on top of FR4 substrate and the other element on the bottom of the substrate. Microstrip stub extends from one element of a dipole to the ports connected on the opposite element in the individual pair. The height of substrate from the ground is H. At port 1, the feed extends from height $h1$ to height $h2$ and at port 2, the feed extends from height $h2$ to $h1$. The stubs are at a height of $h2$ and $h1$ for port 1 and port 2 respectively to avoid intersection.

3 Simulation Results for Dual-Band Operation

With the proper adjustment of the angle between the two spikes in each half of the dipole, a second resonant mode is obtained. For angle greater than 90 degrees we obtain return loss of -14.71 and -14.86 dB at 1.45 and 2.6 GHz respectively as shown in Fig. 1b. The bandwidth achieved at each of these frequencies is 160 and 126 MHz respectively. This is the dual-band operation of the star-shaped dipole antenna. Figure 2 shows the radiation patterns of the star-shaped dipole antenna for dual-band operation simulated at (a) 1.45 GHz, (b) 2.6 GHz.

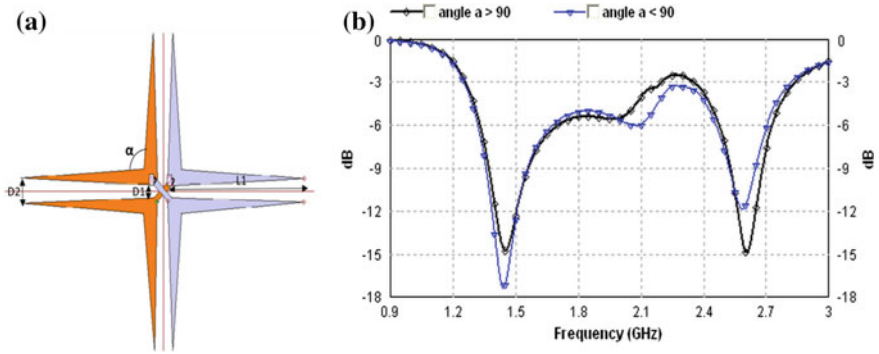


Fig. 1 **a** configuration of the star-shaped antenna; $L1 = 88$, $D1 = 3.5$, $D2 = 9.7$, $t = 0.8$, $w = 3$, $l = 15.5$, $H = 36$ (parameter = mm) **b** Effect of variation of angle between the 2 spikes

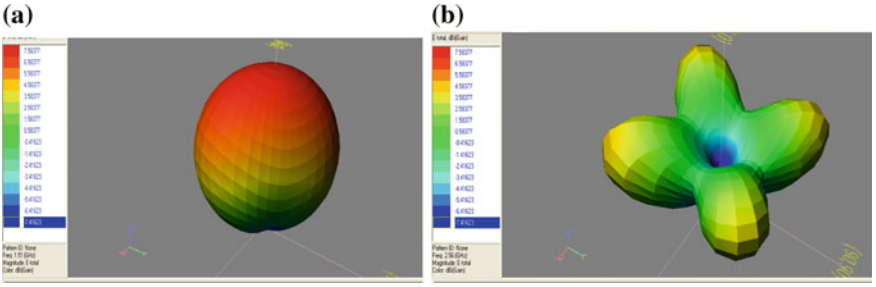


Fig. 2 Radiation patterns of the star-shaped dipole antenna for dual-band operation simulated at **a** 1.45 GHz, **b** 2.6 GHz

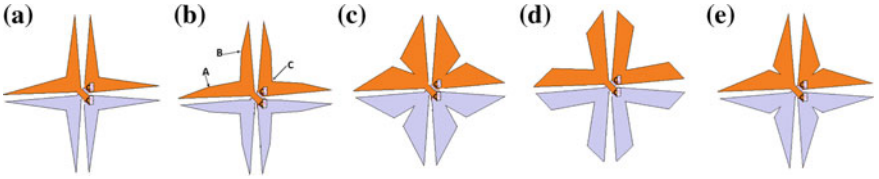


Fig. 3 Configuration of the star-shaped antenna as points A and B are moved from its original position **a** original shape, **b** movement vertically upwards, **c** moving A and B inside, **d** moving A and B outside, **e** moving A and B more inside as compared to (**c**)

The centre point of each side of the dipoles is marked as A and B as shown in Fig. 3 and the deviation of it from its original position is also shown. As we increase the centre point A and B return loss of lower resonant frequency increases and that of higher resonant frequency decreases. Also the movement of these points inside and outside shifts the resonant frequency towards left.

4 Simulation Results for Multiband Operation

Point C is marked in Fig. 3. The antenna is shown to have high return loss value as the distance of point C increases inside the antenna. The antenna is shown to have high return loss value as the distance of point C increases inside the antenna. As shown in Fig. 4a Return loss of −16.63, −18.17, −20.37, −19.64 dB at 0.998, 1.5, 2.32, 2.82 GHz, respectively, is obtained.

Further, increasing and decreasing the length $L1$ also has an effect on the movement of resonant frequencies as seen in Fig. 4b. The optimum value of $L1$ is found to be 80 mm so as to get the multiband operation at 0.9, 1.4, 2.1, 2.6 GHz. The radiation patterns of the star-shaped dipole antenna for multiband operation is shown in Fig. 5.

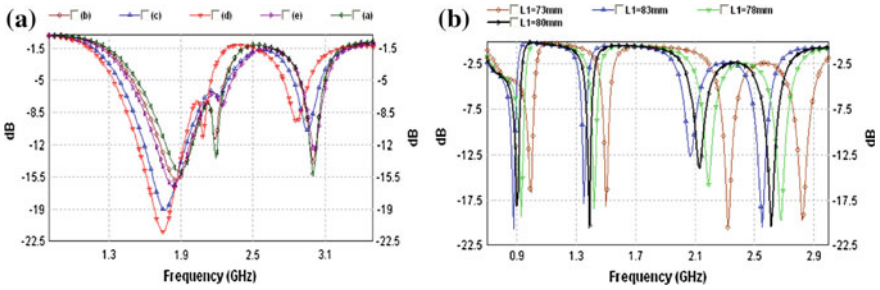


Fig. 4 **a** Effect of movement of mid-point A and B on $S11$ parameter. **b** Multiband operation as length $L1$ is optimised

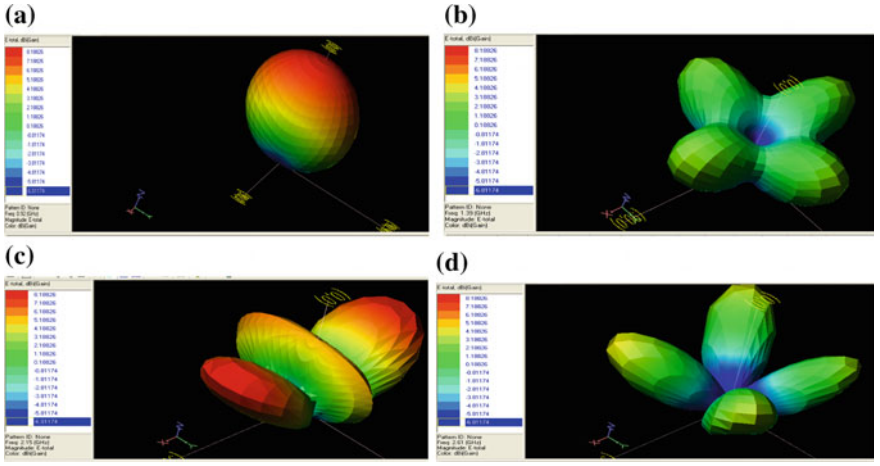


Fig. 5 Radiation patterns of the star-shaped dipole antenna for multiband operation simulated at **a** 0.9 GHz, **b** 1.4 GHz, **c** 2.1 GHz, **d** 2.6 GHz

5 Dipole Antenna Element with Additional Spike

In this section, the discussion is made on the proposed antenna with the introduction of an additional spike of length L between each half of the dipole antenna. Thus the change in the antenna structure is as shown in Fig. 6a.

6 Simulation Results for Dipole Antenna Element with Additional Spike

Now, the effect of various important parameters on this modified structure of star-shaped dipole antenna is studied. Distances $D1$ and $D2$ are important parameters in the antenna design. As the distances $D1$ and $D2$ increase resonant frequencies shifts to 0.922, 1.309, 2.053, 2.412 GHz having return loss values of −11.96, −18.03, −7.73, −8.82 dB, respectively. But if we decrease distances $D1$ and $D2$ resonant frequencies shifts to 0.876, 1.42, 2.206, 2.769 GHz having return loss values of −7.576, −9.838, −13.056, −21.878 dB, respectively. Thus, it is evident that as the distance between dipoles is increased the lower two resonant frequency becomes significant and is shifted towards the right and left respectively. Whereas with the decrease in distance between dipoles the higher two resonant frequency becomes significant and these two frequencies is shifted towards the right simultaneously. This is shown in Fig. 6b.

The next important parameter is the length of edges of the dipole antenna $L1$. The effect of increasing and decreasing the size of $L1$ is as shown in Fig. 7a. For $L1 = 80$ mm the corresponding return loss values are −21.632, −14.453, −12.3, −12.361 dB at 0.949, 1.42, 2.188, 2.76 GHz, respectively.

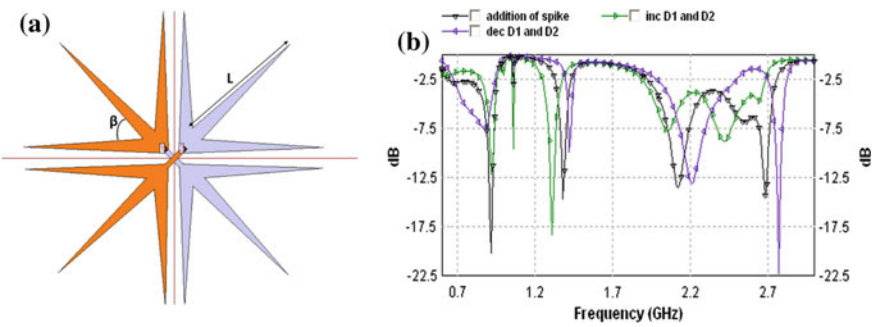


Fig. 6 **a** Geometry of the star-shaped antenna with additional spike. **b** Effect of variation of distance between dipoles $D1$ and $D2$ on $S11$ parameter

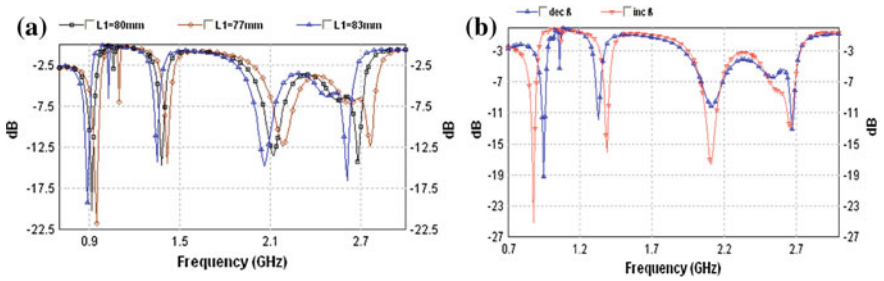


Fig. 7 **a** Effect of variation of length of edges $L1$ on $S11$ parameter **b** Effect of change of angle β on $S11$ parameter

The effect of variation of the angle between the two spikes β of individual half of dipole on the return loss is shown in Fig. 7b. The higher value of return loss is obtained as the angle is increased. Also, with the increase in the angle the first resonant frequency shifts towards lower frequency and second resonant frequency shifts towards higher frequency whereas third and fourth resonant frequency remains constant. The optimum values of return loss obtained after increasing the angle are −24.239 dB at 0.879 GHz, −15.887 dB at 1.389 GHz, −17.57 dB at 2.108 GHz and −12.717 dB at 2.658 GHz respectively.

The length of the spike L also plays an important role in the performance of the antenna. The values of return loss as we decrease the length of the edge by 10 mm are −11.49 dB at 0.888 GHz, −14.434 dB at 1.4 GHz, −24.07 dB at 2.16 GHz and −20.36 dB at 2.628 GHz respectively. The values of return loss with the increase in $L3$ by 10 mm are −8.4 dB at 0.89 GHz, −10.409 dB at 1.432 GHz, −25.952 dB at 2.197 GHz and −22.957 dB at 2.631 GHz, respectively.

7 Conclusion

A multiband planar antenna has been developed for GSM and LTE application. For the dual-band configuration the two resonant frequencies are 1.45 and 2.6 GHz having a bandwidth of 160 and 126 MHz respectively. For multiband configuration, the four resonant frequencies are 0.99, 1.4, 2.1 and 2.6 GHz. After the introduction of an additional spike between each half of the dipole antenna resonant frequencies are obtained at 0.94, 1.42, 2.19 and 2.76 GHz. Hence, the proposed configuration is suitable for GSM and LTE communications.

References

1. Wong K-L (2002) Compact and broadband microstrip antennas. Wiley and Sons, Inc
2. Ludwig AC (1973) The definition of cross polarization. IEEE Trans Antennas Propag 21:116–119
3. Fujimoto K (2008) Antennas for base stations in mobile antenna systems handbook, 3rd edn. Artech House Inc, Boston, pp 141–211
4. Kumar G, Ray KP (2003) Broadband Microstrip Antennas. Artech House, Inc, Boston
5. Balanis CA (1997) Antenna theory: analysis and design, 2nd edn. Wiley and Sons
6. Chiou TW, Wong KL (2002) Broadband dual-polarised single microstrip patch antenna with high isolation and low cross polarisation. IEEE Trans Antennas Propag 50:399–401
7. Cui YH, Li RL, Fu HZ (2014) A broadband dual-polarized planar antenna for 2G/3G/LTE base stations. IEEE Trans Antennas Propag 62(9):4836
8. Huang H, Liu Y, Gong S (2016) A broadband dual-polarized base station antenna with sturdy construction. IEEE Antennas Wireless Propag Lett 1536 1225
9. Chu Q-X, Wen DL, Luo Yu (2015) A broadband ±45 dual-polarized antenna with Y-shaped feeding structure. IEEE Trans Antennas Propag 63(2):483–490
10. Wen D-L, Zheng D-Z, Chu Q-X (2016) A dual-polarized planar antenna using four folded dipoles and its array for base stations. IEEE Trans Antennas Propag, 2016 IEEE
11. Wu R, Chu QX (2016) A broadband dual-polarized buttery antenna for base stations. IEEE (2016)
12. Zheng D-Z, Chu Q-X (2016) The design of dual-polarized antenna for base station applications. In: IEEE European Conference on Antennas and Propagation (EuCAP) 2016
13. Cui YH, Li RL, Wang P (2013) A novel broadband planar antenna for 2G/3G/LTE base stations. IEEE Trans Antennas Propag 61(5):2767–2774
14. Huang Y-H, Wu Q, Liu Q-Z (2009) Broadband dual-polarised antenna with high isolation for wireless communication. Electron Lett 45(14):714715
15. Li B, Yin Y-Z, Hu W, Ding Y, Zhao Y (2012) Wideband dual-polarized patch antenna with low cross polarization and high isolation. IEEE Antennas Wirel Propag Lett 11:427430

Performance Enhancement for VANET Security Using Attack-Resistant Trust (ART)

Bhakti V. Pawar and Manoj M. Dongre

Abstract The significant research work in vehicular ad hoc networks (VANETs) has managed to change the way of traveling and transportation with help of wireless network, in order to improve the standards of intelligent transport system (ITS). There are several challenges in VANET technologies where QoS and security are key points. Many security-related challenges are posed by VANETs such as data reliability, non-repudiation, availability, privacy protection, access control, network overhead etc. So, in this paper, an attack-resistant trust (ART) scheme for VANET communication is used and compared it with baseline method. The proposed technique is capable of exhibiting superior performance with respect to the baseline method in terms of QoS metrics like throughput and communication overhead.

Keywords Malicious · Precision · Security · Trust · Vehicular ad hoc network (VANET)

1 Introduction

The VANET is a non-centralized form of communication which, unlike MANET, does not consist of any centralized administration. Here, nodes are vehicles; they are needed to be in contact with each other to deliver the important data using multi-hop paths. The VANETs have potential to share data same way as human being making communication in VANET more practical and safe [1].

There are three fundamental parts of VANET correspondence i.e., onboard unit (OBU), roadside unit (RSU), and application unit (AU). There is communication between the vehicles and roadside units (RSUs). RSUs are processed with many

B. V. Pawar (✉) · M. M. Dongre
Department of Electronics and Telecom, Ramrao Adik Institute of Technology, Navi Mumbai, India
e-mail: pawarbhaktiv34@gmail.com

M. M. Dongre
e-mail: mmdongre04@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_23

capabilities such as to sense any wrong thing happening nearby, to process the data received, and to share data with other vehicles. The OBUs are in car with the driver; it communicates with the RSU it came across [2].

The AU contacts with both OBU and RSU. Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications are two types of communications present in VANETs. These communications have capabilities to provide messages or if required warnings of accidents, traffic situations such as crowd, mishap, natural damage, and other transport related situations. With the increase in dependence of communication, processing, and maintenance technologies, the security and privacy difficulties in VANET have occurred which includes trust, security, control of access, requirement, privacy, and availability. The use of wireless communication links in VANET leads to attacks such as eavesdropping, active tampering as well as other types of sophisticated attacks. These attacks are difficult to detect and cope with the malicious attacks. This makes the security of driver, vehicles as well as passengers to be in danger causing trouble in transportation. We trust that the reliability of VANETs could be enhanced by addressing the data trust and node trust individually. Vehicular ad hoc network is a communication network for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications which are formed to develop the intelligent transport system (ITS) in road traffic activity for the avoidance of accidents, post-accident examination, mitigation of traffic jams, and other non-security applications [2, 3].

2 Security Trust Schemes

2.1 CTMR

CTMR stands for categorized trust-based message reporting scheme for VANET. It is a decentralized trust management scheme. There are two trust schemes such as role-based and experiment-based, combined to form a trust scheme; they also used piggybacking process whenever required. The nodes involved are individually evaluated during interactions. The trust value of nodes helps to assign its category level, and the degree of trust is assigned by measuring the confidence. The data related to drivers has been saved. Hence, CTMR presents a node trust management for VANETs. In this decentralized scheme, the system monitors malicious nodes and records their actions with respect to their consumed time. V2I and V2V communication are the two forms of communication with advisory network used to message efficiently using RSU [4].

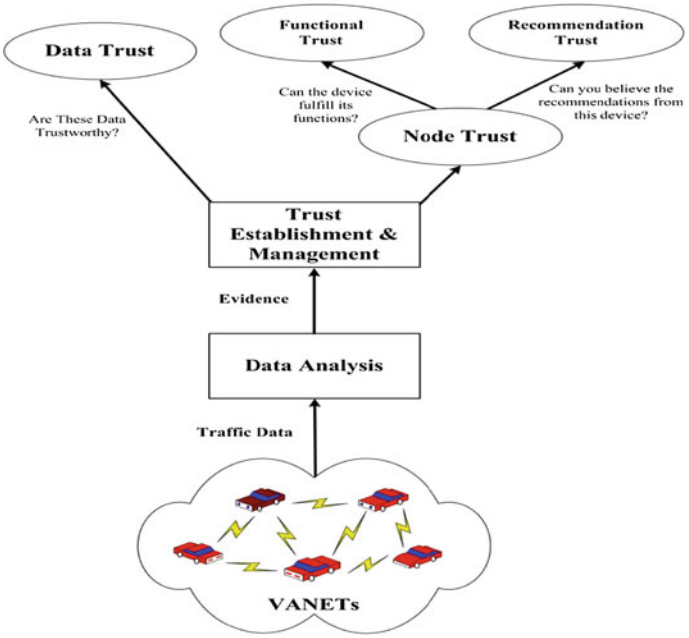


Fig. 1 Overview of the ART scheme [5]

2.2 Attack Resistance Trust (ART)

In VANETs, data is transferred among the nodes, RSUs, OBUs, and AUs as well and stored for further use. The stored data is gathered and verified whenever required. This gathered data is stored in compressed form, by properly analyzing it. The reliability on the trust of the administration depends on this stored data; the data trust and nodes trust for their progress depend on the same. Collaborative filtering is used for recommending the trust [5] (Fig. 1).

To trust any node blindly is not so good for the security of the network. The data sent by nodes can be true or false or partially right or wrong. The trusts in ART are classified into two different types: data trust and node trust. Data trust is how true the information is and node trust is about the nodes' loyalty during exchange of data.

2.3 Evidence Combination Using DST

Evidence is proof used to verify the trust, in ART scheme. In order to gather data, the trustworthy and non-trustworthy nodes send data which is verified but it is not that easy to search proof and combine these data. But it is important to combine different bits of proofs, so that the information trust and important data trust can be

appropriately evaluated. Bayesian interference can also be used instead of Dempster Shafer theory (DST). Bayesian has a drawback that it completely ignores the refused event, whereas DST not actually refuses the event, but this makes DST more appropriate in case of uncertainty or less availability of information. In this work, in Dempster Shafer theory of evidence (DST), whether it is right or wrong, all the proofs or evidences must be used in fused form. In Dempster Shafer theory of evidence (DST), uncertainty is replaced by probability and the level of uncertainty is bounded by belief (bel) and plausibility (pls). The supporting evidence is represented by belief and it is a lower bound of the interval. The non-refusing evidences are the upper bound and are termed as plausibility. For example, if node N_k observes that any of its neighborhood node, such as N_J, has q probability of dropped packets, then N_k will have belief of q degree and no or zero degree belief for its absence. The value of belief is considered as α_i and the observations by node N_k can be considered as follows [5]:

$$\mathrm{bel}_{N_k}(\alpha_i) = \sum\nolimits_{e:\alpha_i \in \alpha_i} m_{N_k}(\alpha_e) \tag{1}$$

Here, α_e are all basic events that compose of the event, α_i and $mN_k(\alpha_e)$ stands for the view of event α_e by node N_k. In this case, node N_k merely gets one single report of node N_j from itself, i.e., $\alpha_i \in \alpha_i$. Therefore, we can derive that $\mathrm{bel}N_k(\alpha_i) = mN_k(\alpha_i)$. Note that $\bar{\alpha}_i$ denotes the nonoccurrence of the event $\bar{\alpha}_i$. Since the equation $\mathrm{pls}(\bar{\alpha}_i) = 1 - \mathrm{bel}(\bar{\alpha}_i)$ holds for belief and plausibility, we can further derive the following [5]:

$$\mathrm{bel}_{N_k}\left(N_j\right) = m_{N_k}\left(N_j\right) = q$$

and

$$\mathrm{pls}_{N_j} = 1 - \mathrm{bel}_{N_k}\left(\bar{N}_j\right) = 1 - q$$

Given that belief indicates the lower bound of the uncertainty interval and represents supportive evidence, we define the combined packet dropping level of node N_i as follows [5]:

$$pd_{N_j} = \mathrm{bel}\left(N_j\right) = m\left(N_j\right) = \oplus_{K=1}^{K} m_{N_k}\left(N_j\right)$$

Here, $m_{N_k}\left(N_j\right)$ denotes the view of node N_k on another node N_j. We can combine reports from different nodes by applying the Dempster's rule which is defined as follows [5]:

$$m1\left(N_j\right) \oplus m2\left(N_j\right) = \frac{\sum_{q,r:\alpha_q \cap \alpha_r = N_j} m1\left(\alpha_q\right) m2(\alpha)}{1 - \sum_{q_1 r:\alpha_q \cap \alpha_r = \phi} m1\left(\alpha_q\right) m2(\alpha_r)}$$

The evidences received by the mobile node in local area and other proofs shared by nonlocal nodes are used together in the fused form in dumpster's rule of combination.

3 Simulation and Evaluations

It is not easy for two nodes to communicate directly and without any difficulties. So, the node has to depend on other node, that is, vehicle for communication. However, a node may deny to share data due to some reasons, and then it is required to compromise by adverse. Therefore, it is very critical to trust any vehicle and interact with it. But if the vehicle has never been interacted with other vehicles, then it has to depend on the trust factor of other nodes, to evaluate the trustworthiness of the neighboring nodes. Assume that $P = [P1,\ P2 \ldots Pq]$ signifies the arrangement of q nodes in the VANETs. The vector $VX = \left[\overline{V_{X_1}}, \overline{V_{X_2}} \ldots \overline{V_{X_q}}\right]$ means the suggestion trust ratings that node X makes for each P_i in P. So also, the proposal trust evaluations that node Y keeps for every node can be meant as $VY = \left[\_\overline{V_{Y_1}}, \overline{V_{Y_2}} \ldots \overline{V_{Y_q}}\right]$. The validity of proposals of node Y can be figured by the closeness of the trust rating data between node X and node Y. In this report, the cosine-based likeness metric is used to measure how comparable the two vectors are [5].

$$\cos\left(\bar{i} \cdot \bar{j}\right) = \frac{\bar{i} \cdot \bar{j}}{\|\bar{i}\| * \|\bar{j}\|}$$

Here, the user-based collaborative filtering is used to help determine the recommendation trust of other nodes. More specifically, the value of the unknown trust rating τ_A; B for node A and another node B is usually computed as an aggregate of the ratings of some other users for the same node B, which is shown as follows [5]:

$$\tau_{A,B} = \mathrm{agg}\tau_{N_i \in \hat{N}} \tau N_i, B$$

where $\hat{N}$ denotes the arrangement of nodes that have most comparable suggestion trust ratings with regard to the trustworthiness of node B. At the end of the day, nodes which have comparable trust preferences on a few nodes may likewise have comparable preferences on others. Consequently, this technique gives suggestions or expectations to the objective node based on the feelings of other similarly invested nodes. In specific, the proposal trust is determined utilizing the following steps [5].

- Trust rate formation: The trust rating of each node N_i for each individual node N_j is written in the form of $q \times q$ matrix R.
- Selection of trustful nodes: The topmost K node will be selected only after computing all the similarities in all nodes. The functional trust of every selected node is also considered, so that only recommendation of the nodes can be considered which can complete the given task.

- Trust calculation to be predicted: On basis of Resnick's standard prediction formula, $T_i k$ is calculated as follows, where the predicted trust rating of node i on node k, $T_i k$ is calculated. Let S_i be the set of most similar nodes for node i. $\bar{R}_i = \sum_k R(i,k)$ and $\bar{R}_j = \sum_k R(j,k)$ stand for the overall trust ratings of node i and node j, respectively [5].

$$T_{ik} = \bar{R}_i + \frac{\sum_{j \in S_i} \cos(i,\, j) * \left(R_{(i,j)} - \bar{R}_j\right)}{\sum_{j \in S_i} |\cos(i,\, j)|}$$

3.1 Abbreviations and Acronyms

- P = Precision
- R = Recall
- DST = Dempster Shafer theory
- ART = Attack-resistant trust
- RSU = Roadside unit
- OBU = Onboard unit
- AU = Application unit
- CTMR = Categorized trust-based message report
- ITS = Intelligent transport system
- VANET = Vehicular ad hoc network

3.2 Performance Evaluation

We use network simulation 2 as the simulation platform, and Table 1 lists the parameters used in the simulation scenarios. We utilize the weighted voting strategy as the baseline technique when we assess the execution of the ART method, on the basis of the fact that the weighted voting technique has been broadly utilized as a part of numerous past trust methods for wireless systems [6, 7]. We utilize the accompanying two parameters to assess the exactness of the ART conspire: Precision (P) and Recall (R), which are both generally utilized as a part of machine learning and data recovery to evaluate the exactness [5, 8]. Here, we utilize both P and R esteems to assess how exact the proposed ART plot is the point at which it is utilized to recognize fake nodes in VANETs. The parameters considered for simulation are in Table 1.

$$P = \frac{\text{number of truly malicious nodes found}}{\text{total number of untrustworthy nodes found}}$$

$$R = \frac{\text{number of truly malicious nodes found}}{\text{total number of truly malicious nodes}}$$

Table 1 Simulation parameters

Parameters	Value
Simulation area	600 × 600
Number of nodes	50, 100, 200
Transmission range	120 m
Number of malicious nodes	5, 10, 15, 20, 25, 30, 35, 40
Node motion speed	5 m/s, 10 m/s, 20 m/s
Simulation speed	900 s
Type of network	Mobile
Routing protocol	AODV
Node movement	Random
Tool	NS2

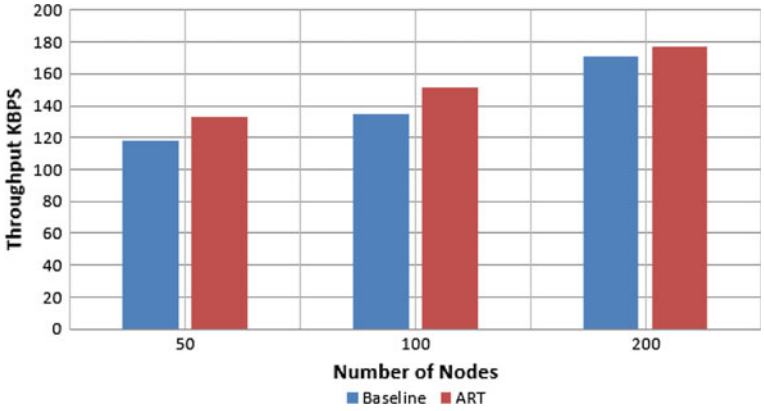


Fig. 2 Throughput rate of 50, 100, 200 nodes

The ART scheme always manages to show a higher accuracy than the baseline strategy when the nodes' number changes. In addition, when the number of nodes is higher, both the techniques show superior accuracy. This is genuine on the grounds that it will probably get true information from others when there are a higher number of all-around acted nodes. Network simulation 2 is used as the simulation platform, and Table 1 lists the parameters used in the simulation scenarios. On the basis of the fact, the baseline weighted voting technique has been broadly utilized as a part of numerous past trust methods for wireless systems [6, 7].

The two parameters that assess the accuracy of the ART are Precision (P) and Recall (R); both are generally used as a part of machine learning and data recovery to evaluate the exactness [5, 8]. Both P and R tell approximately how accurate the proposed ART plot is the point at which it is used to find the fake nodes in VANETs. In Fig. 2, the graph shows the comparative study of simulation between the ART and baseline for throughput parameter.

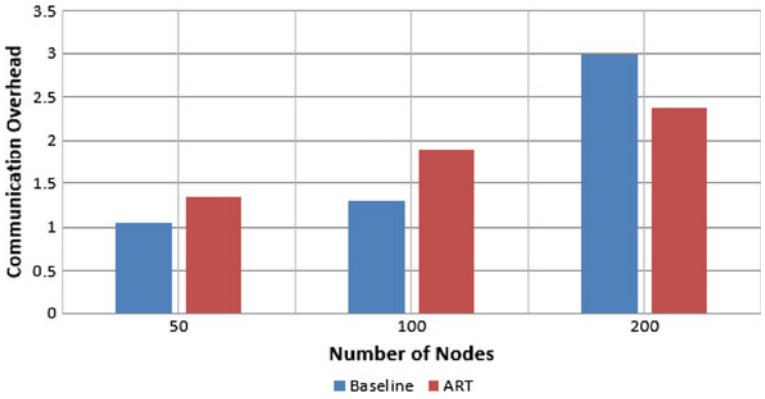


Fig. 3 Communication overhead for 50, 100, 200 nodes

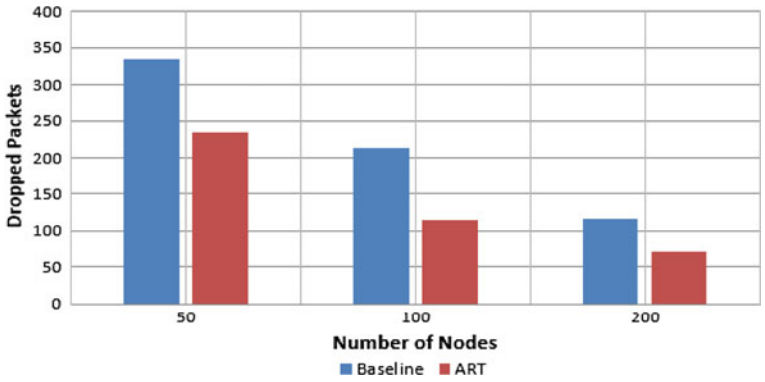


Fig. 4 Packets dropped for 50, 100, 200 nodes

- Throughput: It is defined as the amount of data transferred over the period of time expressed. The throughput measures accuracy of the network that can constantly provide information to the sink. The throughput is typically measured in bits every second.

$$\text{Throughput} = \text{Received data}/\text{Data transmission period}$$

- Communication overhead: It can be considered as the ratio of the control information sent to the actual data received at each node in VANET. It can also be explained as processing time required by node to transmit a data (Fig. 3).
- Packets dropped: Packets dropped is the ratio of the number of packets that never reached the destination to the number of packets originated by the source.

Here is the comparative study of baseline and ART for the number of packets dropped in the VANET communication. These packets are dropped because of the fake or malicious nodes. The graphical representation is shown in Fig. 4.

4 Conclusion

In ART, we studied the two separate metrics such as data trust and node trust. The simulations are conducted and results are compared for both ART and baseline schemes for observing throughput, packets drop and the communication overhead. The results show that there is improvement in throughput, packet drop and overhead readings in ART scheme compared to the baseline scheme.

References

1. Al-Sultan S, Al-Doori MM, Al-Bayatti AH, Zedan H (2014) A comprehensive survey on vehicular ad hoc network. J Netw Comput Appl 37:380–392 (Elsevier)
2. Kakkasageri MS, Manvi SS (2014) Information management in vehicular ad hoc networks: a review. J Netw Comput Appl 39:334–350 (Elsevier)
3. Sharef BT, Alsaqour RA, Ismail M (2014) Vehicular communication ad hoc routing protocols: a survey. J Netw Comput Appl 40:363–396 (Elsevier)
4. Monir MB, Aziz MHAE, Hamid AAA (2013) A trust-based message reporting scheme for vanet. Int J Emerg Technol Adv Eng 5:376–394 (Citeseer)
5. Li W, Song H (2016) ART: an attack-resistant trust management scheme for securing vehicular ad hoc networks. J IEEE Trans Intell Transp Syst 17:4960–969 (IEEE)
6. Nait-Abdesselam F, Bensaou B, Taleb T (2008) Detecting and avoiding wormhole attacks in wireless ad hoc networks. IEEE Commun Mag 46.4:127–133 (IEEE)
7. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the art and possible extensions. IEEE Trans Knowl Data Eng 17:734–749 (IEEE)
8. Raya M, Papadimitratos P, Gligor VD, Hubaux J-P (2008) On data-centric trust establishment in ephemeral ad hoc networks. In: INFOCOM 2008, the 27th conference on computer communications. IEEE, pp 1238–1246
9. Maheswari D, Sekar S, Manju Barkavi V, Kamalam GK (2017) Misbehavior detection in VANET using an attack resistant trust management scheme. Int J Adv Res Trends Eng Technol (IJARTET) 4(Special Issue 1)
10. Pophali M, Mohod S, Yengantiwar TS (2014) Trust based opportunistic routing protocol for VANET communication, TS. Int J Eng Comput Sci 3.8:7408–7414
11. Chen R, Bao F, Chang MJ, Cho J-H (2014) Dynamic trust management for delay tolerant networks and its application to secure routing. IEEE Trans Parallel Distrib Syst 25:1200–1210 (IEEE)
12. La Vinh H, Cavalli AR (2014) Security attacks and solutions in vehicular ad hoc networks: a survey. Int J AdHoc Netw Syst (IJANS) 4
13. Rebolledo-Mendez G, Reyes A, Paszkowicz S, Domingo MC, Skrypchuk L (2014) Developing a body sensor network to detect emotions during driving. IEEE Trans Intell Transport Syst 15:1850–1854 (IEEE)

14. Yeh L-Y, Lin Y-C (2014) A proxy-based authentication and billing scheme with incentive-aware multihop forwarding for vehicular networks. IEEE Trans Intell Transport Syst 15:1607–1621 (IEEE)
15. Taha S, Shen X (2013) A physical-layer location privacy-preserving scheme for mobile public hotspots in NEMO based VANETs. IEEE Trans Intell Transport Syst 14.4:1665–1680 (IEEE)
16. Li Z, Liu C, Chigan C (2013) On secure VANET-based ad dissemination with pragmatic cost and effect control. IEEE Trans Intell Transport Syst 14:124–135 (IEEE)
17. Zeadally S, Hunt R, Chen Y-S, Irwin A, Hassan A (2012) Vehicular ad hoc networks (VANETs): status, results, and challenges. J Telecommun Syst 50.4:217–241 (Springer)
18. Lu, R, Lin X, Liang X, Shen X (2012) A dynamic privacy-preserving key management scheme for location-based services in VANETs. IEEE Trans Intell Transport Syst 13:127–139 (IEEE)

Wireless Protocols: Wi-Fi SON, Bluetooth, ZigBee, Z-Wave, and Wi-Fi

Gollu Appala Naidu and Jayendra Kumar

Abstract The 4G/5G/fast growing world wireless technology plays a vital role as wired connections nowhere using because of its installation complexity. Wireless accessing of networks, wireless services, and IoT controlling are the current trend in technology for providing quickness, easiness, mobility, feasibility, and/or flexibility. Apart from that, other merits include wireless networks that can dynamically from the network in the real world and it is easy deployment than wired. Wireless protocols such as Bluetooth (IEEE 802.15.1), ZigBee (IEEE 802.15.4), Wi-Fi (Wireless Fidelity) (IEEE 802.11), and Z-Wave (proprietary-based standard) are widely used technologies, and Wi-Fi SON (Wi-Fi Self-organizing/Optimizing network) is a newly raising technology. Wi-Fi SON is an automated intelligent network. In this research paper, we are describing the wireless technologies importance, features, working, and comparison among them and mainly focused on Self-organizing/Optimization Networks (SONs); one such network is Wi-Fi SON.

Keywords Wireless protocols · SON · Wi-Fi SON · Bluetooth · ZigBee · Z-Wave · Wi-Fi

1 Introduction

Nowadays, diverse wireless devices are implemented, such as smart hubs for home automation, by supporting short-range wireless or Internet of Things (IoT) protocols as shown in Fig. 1. Homes are becoming automated through smart hubs and mobile applications using cloud storage (Virtual Storage). Smart home providers give user interfacing a mobile application for easy controlling of smart home

G. A. Naidu (✉) · J. Kumar
Department of Electronics and Communication Engineering, National Institute of Technology Jamshedpur, Jamshedpur 831014, Jharkhand, India
e-mail: ganaidu707@gmail.com

J. Kumar
e-mail: jkumar.ece@nitjsr.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_24

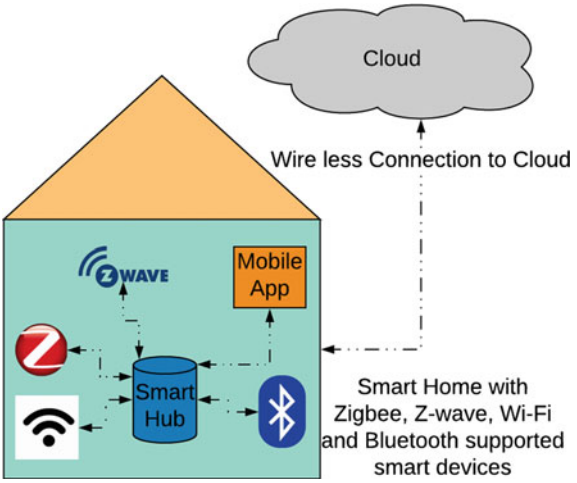


Fig. 1 Smart home with wireless protocols ZigBee, Z-Wave, Wi-Fi, and Bluetooth-supported smart hub and devices

appliances (that are installed at user's home). The mobile application communicates with the smart hub placed at home via the cloud of that vendor provides. Smart home appliances such as smart bulbs, door locks, switchboards, sensors, etc. are supporting one of the short-range protocols among ZigBee, Z-Wave, Wi-Fi, and Bluetooth. These smart home appliances can be controlled via smart hub by pairing with it. In recent world, smart hubs evolve with intelligence and interact with humans by supporting voice commands. From past few years, automation in every field has been increased throughout the world in a greater growth. It expanded the scope for disciplines including telecommunication, information theory, artificial intelligence, computer networks, IoT control, sensors, and actuators technology in a consolidated way, the route to better solutions, new implementations, and optimized performance.

In automation industry, the preferences increase to employ wireless technologies for greater mobility, scalability, low cost, and making easy in smart home and smart city applications.

Wireless service provided should ensure users a guaranteed quality of experience and speed every time, at any place through all devices. At the same time, among the different wireless protocols, they have their own merits and demerits. To find right applications of protocol or technology at an exact place, it is mandatory to get rid of unwanted loss and high cost, and acquires best range or coverage of it. As a fact, the short-range wireless industry is currently lead by four protocols: the Bluetooth [1], ZigBee [1], Wi-Fi [1], and Z-Wave which are corresponding to the IEEE 802.15.1, 802.15.4, 802.11a/b/g, and proprietary-based standards, respectively.

For accessing Internet, Wi-Fi has become the most popular network preferences for network providers and for network operators because of its easiness in establishing the connection. This brings new challenges and requirements for upcoming generations Wi-Fi service to give ideal or picture perfect and uniform network/service quality. Due to that automation, wireless network has become crucial. The challenges

can be eliminated by Wi-Fi Self-organizing/Optimizing Networks (Wi-Fi SON) [2]. This paper provides the solution to a seamless wireless connection.

2 Wireless Protocols

In this section, the Bluetooth, ZigBee, Wi-Fi, and Z-Wave protocols are introduced which correspond to the IEEE 802.15.1, 802.15.4, 802.11a/b/g, and proprietary-based standards, respectively. The IEEE organization is only responsible for defining protocol standards in Mac and Physical layers, for worldwide usage. For the implementation of each protocol feature, specifications covering the network, security, and applications areas are maintained by a group of companies called alliance. Such alliance examples are ZigBee alliance, Z-Wave Alliance, etc. Basic comparison among the protocols is shown in Fig. 2.

2.1 Bluetooth

Based on the wireless radio system, Bluetooth comes under IEEE 802.15.1 standard, initially implemented by Ericsson. Bluetooth's most common application is "cordless computer" [3] having several devices including PCs, laptops, mouse, keyboard, scanner, printer, etc. All them are equipped with Bluetooth card and connect wirelessly. It supports short range and uses personal network application known as Wireless Personal Area Network (WPAN). It is also a solution for data exchange between mobiles, laptops, digital cameras, etc. Bluetooth comes in with wireless headset; no need to hold mobile or laptop. Bluetooth has two types of connectivity topologies: one is piconet and another is scatternet. Any Bluetooth device can be a slave or master in the network. Within the network, the one which establishes a network or the one which form WPAN can be the piconet. Further added devices

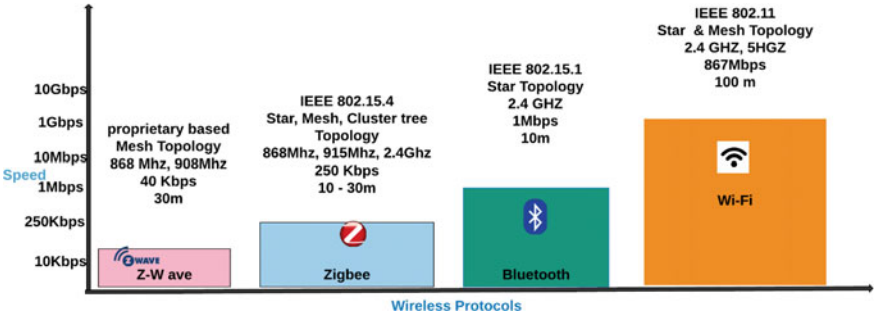


Fig. 2 Wireless protocols comparison (Bluetooth, Wi-Fi, ZigBee, and Z-Wave)

such as discovered by piconet are called slaves. The main function of the master is to synchronize the network with its clock. Slaves are allowed to communicate only with master in a point-to-point (P2P) manner. The master communicates may be either point-to-point or point-to-multipoint.

Bluetooth low energy (BLE) [4] is a distinctive attribute of Bluetooth 4.0 [4] specifications. It has implemented for ultralow power devices, and BLE makes reuse of classic Bluetooth devices. In Bluetooth network, there can be maximum of eight nodes with master and operate frequency range is from 2400 to 2483.5 MHz. It supports data speed up to 1 Mbps, and Line of Sight (LoS) range is 10 m. Bluetooth is not only replacement of cable technology but also used for the creation of ad hoc Personal Area Networks (PAN) with the good security level. Bluetooth has low power consumption and low cost, which makes it an attractive solution for controlling the smart home appliances using smart mobile devices.

2.2 ZigBee

ZigBee [4–7] protocol is low-power, low-cost, and low data rate wireless personal area network, supported range up to 10–100 ms, especially designed and implemented for smart home applications, as an IEEE standard 802.15.4 and maintained by ZigBee Alliance (a group of companies). This works in MAC layer of the network stack and operates at different frequencies such as 2.4 GHz, 915 MHz, and 868 MHz. Its data rates are very low as they are used for short range and in personal areas like homes so its data rates are as follows: 250, 40, and 20 Kbps, respectively. To ensure reliability in data exchange or communication, it uses a dynamic routing protocol. ZigBee supports star, mesh, and cluster tree network topologies for maximizing its range of the network by interconnecting with nodes throughout the network, and is shown in Fig. 3. In mesh network, there exist at least two pathways among the nodes and all nodes in the network are self-routed and able to communicate with other nodes and exchanges data.

ZigBee devices are self-organized in network, support multiple hops, and create mesh networks of more reliable with long-life battery performance. ZigBee has two different types of devices for forming ZigBee network: A coordinator or router or Full Function Device (FFD), and end device or Reduced Function Device (RFD). FFD can be used as a coordinator or router or end device, and RFD can only be used as end device. Coordinator has the capability of initiating a network and can add other coordinators or routers or end devices into the network. Routers are devices added by FFDs and have routing ability but cannot initiate the network, while end device can only respond to routers or coordinators. ZigBee device-type RFDs (end devices) are actually implemented for small applications such as ZigBee bulb, switchboards, smart plug, sensors, and door locks. They can only transfer small piece of data, attached to single coordinator at a time. The RFDs are implemented with minimum capacity or minimal resources and memory because of its limitations in usage. So these devices are called resource-constrained devices. In general, RFD sends its

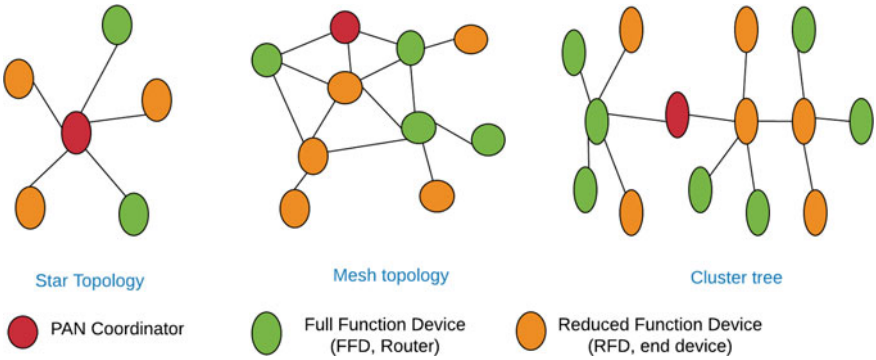


Fig. 3 ZigBee network topologies (Star, mesh, and cluster Tree)

NODE identifier, end point number, IEEE address, and supported ZigBee cluster identifier. Supported ZigBee clusters and cluster classification are specified by ZigBee alliance. End device or RFD manufacturers have to follow the ZigBee Cluster Library (ZCL) standard in order to get certification from ZigBee alliance. After an FFD or coordinator is activated, network up and running for the first time, it can create an own ZigBee network to become the network coordinator. Each coordinator can form different networks side by side without inferring other networks in operation. This is possible by choosing a unique PAN identifier (it is 16-bit address in Hexadecimal format), which is not in use by other ZigBee networks. Once the PAN identifier is chosen, scan for the strong channel is based on its frequency band support. The PAN coordinator can allow or add ZigBee end devices by going into discovery mode for few seconds to add in its network. An RFD may connect to coordinator network as an end node at the end of a branch, because it can attach to one coordinator at a time. For providing synchronizing services in network, any FFD may act as a coordinator or router. Among different coordinators, the one PAN coordinator which initiated the network acts an overall PAN coordinator, having greater computational resources, routing tables, and device tables.

ZigBee provides security [8] such as AES encryption for the secure transfer of the data at the high level, and no security option is also supported.

Silicon Labs is a major and prominent vendor of ZigBee modules, followed by Texas Instruments, NXP Semiconductors, and Marvell which are other major vendors. ZigBee end devices are supplied for several companies.

2.3 Z-Wave

It is an IoT [9] protocol designed for controlling home appliances and lightweight industrial applications remotely. It has steady speed up to 40 kbps and 30 m range of LoS and indoor may reduce. It has a security level of 128-bit AES encryption and

avoids interference with other leading protocols such as Wi-Fi, Bluetooth, and other systems that operate on more crowded 2.4 GHz because it operates at 908.42 MHz in the US and 868.42 MHz [4] in Europe. It was initially developed by Zensys, and later that was acquired by sigma designs in 2008.

Z-Wave is a proprietary standard protocol, and have controller and slave device types for forming a network. Slave nodes reply and perform commands issued by controlling devices and low-cost devices, and unable to initiate messages. Depending on the functionality, end device or slave devices may have several forms. Z-Wave supports mesh network topology. There is always one controller that has overall control of Z-Wave network. Since it was designed for smart home applications, it supports maximum 232 devices in the network. The Z-Wave devices classified by segregating them into classes such as general class, base class, and command classes.

2.4 Wi-Fi

Wireless Fidelity (Wi-Fi) [10, 11] is the most popular wireless network that follows IEEE standard 802.11n/a/b/g for Wireless Local Area Network (WLAN). Users are allowed to browse the Internet and connect to cloud at their broadband speeds (provided by network vendor) when connected to an Access Point (AP) or in ad hoc mode. IEEE 802.11b/g uses 2.4 GHz band and IEEE 802.11a uses 5 GHz band. IEEE 802.11n utilizes both 2.4 and 5 GHz bands with Multiple-Input Multiple-Output (MIMO) mechanisms. Wi-Fi infrastructure has multiple elements that collaborate to provide a wireless LAN that supports station mobility transparently to the above layers. In infrastructure setup, access points (Router) and wireless stations are associated. The group of devices (Access Points + Wireless Stations) is a Base Service Set (BSS) [4], where each wireless station connects to the Internet through its associated access point. Each BSS consists of one Service Set ID (SSID) for identifying BSS.

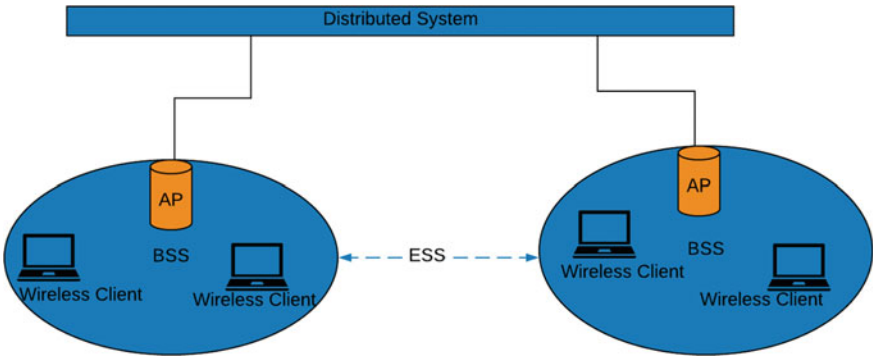


Fig. 4 Wi-Fi infrastructure

If several APs were connected through wired distribution system, different BSSs are named as Extended Service Set (ESS). It is shown in Fig. 4.

Wi-Fi usage has become mandatory at almost all places, in-home networks, industries, and public places. The data rate is up to 11–867 Mbps [4].

Wi-Fi has security at all levels, and at most high-security mode is Wi-Fi Protected Access 2 (WPA2) 802.11i and also has no security option (open network). As compared to other protocols, Wi-Fi has higher power consumption (~116 mA) due to its high data rate. Wi-Fi has good range LoS up to 100 m and indoor around 35 m.

3 Wi-Fi Self-organizing Network (Wi-Fi SON)

Automation of their own network dynamically based on traffic in the real-world network is one of the main new technologies in the current generation. These types of automated networks are called Self-organizing/ Optimization Networks (SONs). SON research and implementation have become faster after switching from 3G to 4G to sustain the huge amount of traffic. This can simply be called introducing intelligence in wireless networks and this can be achieved by changing certain network parameters in the SON algorithms. These techniques can be performed efficiently due to the availability of rich models on Key Performance Indicators [2, 12, 13] (KPIs) and their dependencies on each other and their interrelation among them. However, SON implementation of any wireless technology is not an easy task. The function used in SON must be tuned to each individual network. Optimum parameters [14] must be used in order to comply with the existing network. There are several routing protocols for SON. Each and every routing protocol shares certain key properties. The major challenge in these kinds of self-automated networks is that increasing Bandwidth (BW) leads to incorrect routing protocols, which is not able to choose the optimal routing paths. These are some few challenges in SON implementation. There are three types of routing protocols for SON: (i) Proactive routing protocol, (ii) reactive routing protocol, and (iii) hybrid routing protocol.

Network SON has been mostly studied in telecommunication networks, sensor networks, and ad hoc networks to automate operations and management. SONs are for improving Quality of Service (QoS) in mobile networks, added automation algorithms with a set of principles and concepts. So, SONs required less maintenance than traditional networks. The recent SON technology is implemented in Fourth-Generation Long-Term Support (4G LTE) by adding several SON functions. It is not completely possible to apply the same concept in Wi-Fi systems as cellular and Wi-Fi systems have basic differences in the IEEE layer (MAC Layer) and physical layer standards. Use cases appear to be alike but the firsthand application of cellular concepts of SON is not constructive here. Wi-Fi SON has two types of architecture: (i) Distributed SON and (ii) centralized SON. In this article, we are not discussing similarities or challenges of cellular SON and Wi-Fi systems, mainly focused on Wi-Fi SON use cases. The overall concept of Wi-Fi SON is shown in Fig. 6.

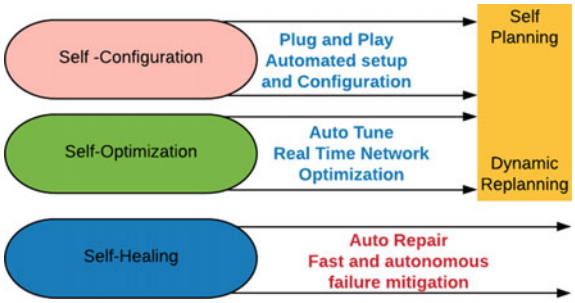


Fig. 5 Wi-Fi SON features (self-configuration, self-optimization, and self-healing)

The Wi-Fi SON main features [2, 11–13] are self-configuration, self-optimization, and self-healing. These features of Wi-Fi network make it able to build or automate the process based on the traffic changes in the network. Sample functioning of key features is shown in Fig. 5.

3.1 Self-configuration

During initial setup of a network at customer site, it requires configuration of APs and network maintenance, and device swap or change in the network environment is mandatory. Generally, all these would be done by a person by changing configurations at APs. But in large network deployments, this should be done by itself, which enables intelligent initial parameter settings. This process is called self-configuration. This is dynamic plug and plays initial settings for newly deployed nodes. This process is generally done by a server called Auto-configuration Server (ACS). Most important use cases of self-configuration are explained below in Fig. 6.

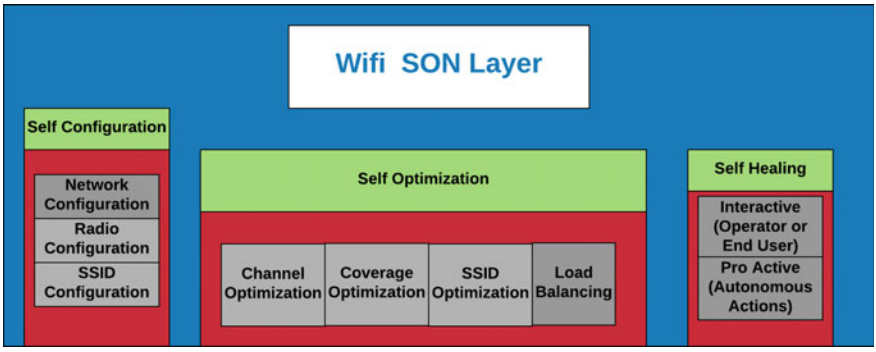


Fig. 6 Wi-Fi SON concept

Network Initial Configuration: For this, SON uses remote zero-touch configuration and initial settings of newly integrated APs such as IP, server address, and authentication. This configuration can be done by Dynamic Host Configuration Protocol (DHCP) or bootstrap protocol agent with help of server ACS.

Radio Initial Configuration: This is basically implemented at vendor-specific setting by predefining radio settings enabled (2.4 GHz or 5 GHz), selecting channel, auto selection of channel, modulation, and coding schemes index, transmission power, etc, while implementing this configuration, Wi-Fi SON has challenges and restrictions.

SSID Initial Configuration: This automated SSID configuration is totally different from cellular network configurations because single AP can have multiple SSIDs as it depends on services supported by the network. This can be achieved by updating newly deployed AP's SSID initial configurations in the network to Wi-Fi SON with help of neighborhood function.

3.2 Self-optimization

Self-optimization is required during operational state to optimize the network and for reducing operational maintenance and expenditure. Self-optimization keeps eye on network performance and maintenance of KPIs at required range. Self-optimization includes channel optimization, coverage optimization, load optimization, SSID optimization, energy optimization, etc. Some of them are discussed below.

Coverage Optimization: Abrupt signal drop and traffic counters are clear indications of the poor coverage area due to weak signal strength. Generally, testing of the network strengths and coverages required technician interventions, and due to this, maintenance cost increases and time consumes. The typical operation of this optimization is increasing large coverage area and capacity in an automated way. Wi-Fi SON function automatically detects the weak strengths and poor coverages by continuous monitoring of APs servers. Dynamic Transmit Power Control (TPC) handles the poor coverage by stabilizing the power among the APs. Coverage optimization function allocates power dynamically at APs, so that traffic and encroaching on adjoining APs reduces without disruption of device performance attached to that particular AP.

SSID Optimization: This is required because most of the network operators enable multiple SSIDs in their integrations in it to the creation of high density of networks. Because of that reason, Wi-Fi SON faces some challenges in SSID optimization. Each time, neighboring networks adding or removing of APs required updating of the table of SSIDs. This optimization can be achieved by setting at initial SSID configurations as discussed above. Wi-Fi SON introduces ideal mode management for improving energy efficiency by deactivating APs that do not create contention in the network as power saving is a primary issue to deal with it the concern in Wi-Fi-based networks.

Load Optimization: In general, Wi-Fi network load depends on channel and/or band balancing.

Channel Balancing: For optimum channel balancing, Wi-Fi SON uses the below criteria. The selection criteria depend on the number of devices connected per AP, capitalization of channel per AP, throughput of channel per AP, bandwidth, and so on. The efficiency of the channel selection function may reduce, whose primary metric in most cases is defined to reduce hidden or exposed node and due to this criterion so care must be taken.

Band Balancing: Wi-Fi has two available bands: 2.4, and 5 GHz. It has to amend network employment and throughput per Dual-Band Dual-Concurrent (DBDC) APs. Wi-Fi SON can handle concurrent requests in the network. This is practically implemented at two levels: (i) AP level and (ii) device level. At AP level, distributed SON function optimizes and at device level centralized SON function optimizes by observing band.

3.3 Self-healing

Self-healing mechanism is to mitigate or repair network faults by itself by triggering appropriate recovery actions. In a fault identification management system, for each type of fault detected, the separate recovery mechanism is there. Based on identified fault, alarm triggers regardless of whether automatically detected or automatically corrected or automatically detected and automatically corrected the fault. This is done by continuous monitoring and diagnosis of the network to identifying root causes. Self-healing in cellular SON and Wi-Fi is totally different. In Wi-Fi SON, sometimes customer interaction is required. This is due to differing architectures, multiple bands, coverage overlapping, etc.

Interactive Healing: The naming convention itself describes that automated fault recovery is not possible, so additional data may require, which leads to a field technician or end user has to provide some data to recover the network back. This fault correction process is time-consuming due to human interventions and additional processing. Example of this use case is an instance of poor coverage area detected and TPC is unable to eliminate the problem by balancing power at APs. Here, end user is guided by a technician, like proper placing location of APs and recovery procedure if the same fault occurs again.

Proactive Healing: This type of self-healing method overcomes drawbacks of interactive healing. It diagnoses the system and recovers in an automated way. This healing requires careful design and implementation; otherwise, this leads to network failure. It depends on other optimizations like self-optimization function output is resistant to the healing method triggered by self-healing lead to deadlocks. Self-healing functions should implement in the way of resolving network faults. Some of the basic remedies for software faults are system reboot, reconfigurations, system initializations, and reloading of back up data, and the hardware faults are resetting faulty resource, isolation of faulty resources, etc.

4 Conclusion

In this paper, we present a review of primary and prominent, leading wireless protocols ZigBee, Z-Wave, Bluetooth, Wi-Fi's requirements, importance in-home automation, basic features, use, challenges, comparison among them as shown in wireless protocols comparison graph, and smart devices available in the market for a smart home. Smart home devices integrate with a smart hub which supports the above IoT wireless protocols that have been described. Newly emerging technology Wi-Fi SON's working, characters, importance, and requirement of the protocol in the current generation have been described. Wi-Fi SON is a guaranteed network to provide quality of service offering, by eliminating human interventions and reducing other costs involved. We further investigate SON algorithms for improving its performance.

References

1. Lee J-S, Su Y-W, Shen C-C (2007) A comparative study of wireless protocols: bluetooth, UWB, ZigBee, and Wi-Fi. In: The 33rd annual conference of the IEEE industrial electronics society (IECON), 5–8 Nov 2007
2. Gacanin H, Ligata A (2017) Wi-Fi self-organizing networks. Challenges and use cases. IEEE Communications Magazine, July 2017
3. Singh P, Sharma D, Agrawal S (2011) A modern study of bluetooth wireless technology. Int J Comput Sci Eng Inf Technol (IJCSEIT) 1(3)
4. Rahman ABA (2015) Comparison of internet of things (IoT) data link protocols
5. Sharma H, Sharma S (2014) A review of sensor networks: technologies and applications. In: Proceedings of 2014 RAECS UIET Panjab University Chandigarh, 06–08 Mar 2014
6. Saad C, Mostafa B, Cheikh EA, Abderrahmane H (2014) Comparative performance analysis of wireless communication protocols for intelligent sensors and their applications. (IJACSA) Int J Adv Comput Sci Appl 5(4)
7. Al-Sarawi S, Anbar M, Alieyan K, Alzubaidi M (2017) Internet of things (IoT) communication protocols: review. In: 8th international conference on information technology (ICIT)
8. Krejci R, Hujnak O, Svepes M (2017) Security survey of the IoT wireless protocols. IEEE
9. Naidu GA, Kumar J, Garudachedu V, Ramesh PR (2018) 6LoWPAN border router implementation for IoT devices on Raspberry Pi. In: 3rd international conference on internet of things and connected technologies (ICIoTCT)
10. Zhu H, Chang ASF, Kalawsky RS, Tsang KF, Hancke GP, Bello LL, Ling WK (2017) Review of state-of-the-art wireless technologies and applications in smart cities. IEEE
11. Kim SM, Ishida S, Wang S, He T (2017) Free side-channel cross-technology communication in wireless networks. IEEE/ACM Transactions on Networking
12. Sridhar R (2008) Self organising network, study research paper, GmbH, Munich, Germany, May 2008
13. Marwangi MMS, Fisal N, Yusof SKS, Rashid RA, Ghafar ASA, Saparudin FA, Katiran N (2011) Challenges and practical implementation of self-organizing networks in LTE/LTE-advanced systems. In: Proceedings of the 5th international conference on IT & multimedia at UNITEN (ICIMU 2011), Malaysia, November 2011
14. Kumar A (2016) A survey of self-organizing networks. European Research Project Paper

Implementation of Wavelength Diversity Technique in Free-Space Optical Link

Shubham Nema, Sagar Oza, Akshay Parmar, Dhaval Shah and Suryansh Singh

Abstract Free-space optical link for terrestrial communication greatly depends on environmental conditions. Varying climatic conditions across regions cause turbulence on the deployed FSO link, which inherently affect its performance. Various techniques have been introduced to mitigate these effects. This paper provides the study on wavelength diversity technique, in which FSO link is established with two optical channels of wavelength 532 and 1350 nm, to understand its performance under varying climatic conditions. Either channel gets selected dynamically based on attenuation levels. Observations are recorded under fog and clear atmosphere. Also, experiments have been performed to understand the effect of distance variation on FSO link under clear atmosphere.

Keywords Free-space optical communication · Atmospheric turbulence · LASER · Wavelength · Diversity · Bit error rate · Performance

1 Introduction

Free-space optical (FSO) communication has emerged as a better substitute to RF communication in the wireless domain around the globe. It has license-free operation and overcomes the bandwidth limitation of radio communication. Multiple applications where line of sight communication can be established use FSO link. Satellite communication, spacecrafts, military jets, and WAN–WAN network are few such use cases of FSO communication.

S. Nema · S. Oza · A. Parmar · D. Shah (✉) · S. Singh
Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, S.G. Highway, Room no. PG-201, Ahmedabad 382481, Gujarat, India
e-mail: dhaval.shah@nirmauni.ac.in

S. Nema
e-mail: 12bec095@nirmauni.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_25

1.1 Background and Previous Work

Performance of free-space optical link largely depends on environmental conditions around the region where it is deployed. Climatic changes, turbulences, fog, rain, smoke scatters away light, and induces noise and attenuation in the signal which degrades signal integrity. There have been various regimes introduced to mitigate the effects of these atmospheric changes on FSO link.

One of the methodologies is aperture averaging. It states that the amount of measured radiation may be increased by increasing the size of the receiver lens aperture. In effect, any irradiance fluctuations across the collecting lens are averaged by the size of the lens. However, study suggests that aperture averaging provides good results with weak turbulence conditions but could not improve performance much under moderate to strong turbulence [1].

Various coding schemes had been introduced to reduce the BER of received signal. Methods including phase shift keying and quadrature amplitude modulation are implemented but do not guarantee enough improvement to be commercially viable. Forward Error Correction (FEC) inserts check bits into the data stream, which contributes to an additional power and bandwidth overhead on the system to send extra bits [1, 2].

Diversity techniques are considered as promising solution, in which the optical wavefront propagates in at least two distinctive ways, increasing the probability that either of one is expected to perform better. Spatial diversity (requires multiple transmitter and receivers to improve performance) consumes significant power and still depends on climatic conditions. Temporal diversity (requiring a signal to be transmitted twice, separated by a time delay) increases the maximum achievable bit rate.

Wavelength Diversity (WD) is a technique which does not pose a limit on maximum achievable bit rate but still leads to significant power reduction compared to spatial diversity. Also, the overhead due to error correction coding schemes is reduced which effectively increases the performance.

So, WD technique is one of the effective ways to mitigate the effect of varying climatic conditions and still maintaining same or even higher performance.

1.2 Establishing Free-Space Optical Communication Link

A simple implementation of free-space optical communication link for serial data transmission can be modeled as shown in Fig. 1. At transmitter end, the USB data from computer is converted to serial data stream. The data stream is passed on to the LASER driver circuit which operates the LASER in the desired operating range and controls its switching. LASER converts electrical signals to optical signals producing optical beam.

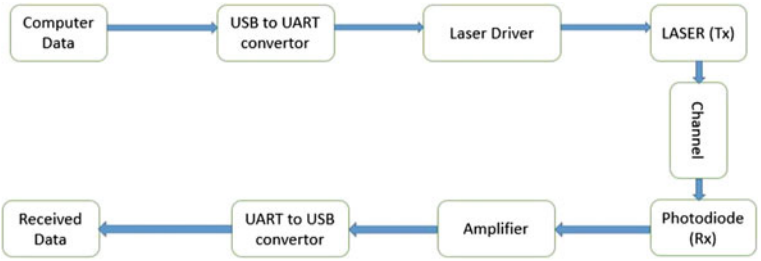


Fig. 1 Free-space optical link block diagram

At the receiver end, the photodiode produces current based on the intensity of optical signals incident on it. However, free-space channel induces various noises which gets added to the signal. The signal is passed to the transimpedance amplifier, which amplifies the signal and improves Signal-to-Noise Ratio (SNR). Also, it converts the current signal into voltage signals. Signal is then converted from serial data to the computer-readable binary format through USB.

2 Wavelength Diversity Technique

As mentioned earlier in the introduction, in the places where atmospheric conditions keep varying, simple LASER communication setup mentioned above will become ineffective. Wavelength diversity technique can be helpful at such places. It is a Multiple-Input Signal-Output (MISO)- or Multiple-Input Multiple-Output (MIMO)-based technique in which multiple LASERs with different wavelengths transmit the same signal. Selective gain combining technique is used to choose the either channel based on whichever is providing maximum gain [3]. Varying atmospheric conditions will have different effects on the optical beam having different wavelengths. At the receiver's end, based on selective gain combining technique, the signal with least attenuation will be chosen. As a result, chances of receiving signal with better signal

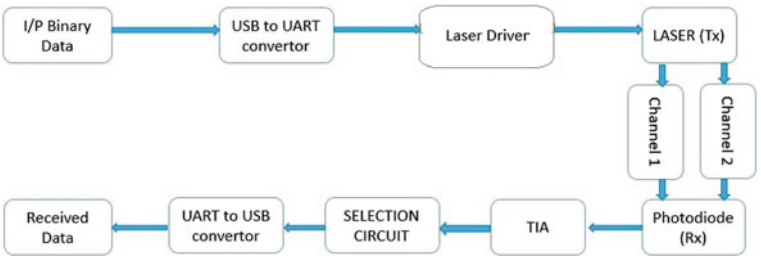


Fig. 2 Block diagram of FSO with WD technique

strength increase. Here, for experimental purpose, two LASERs having wavelength of near extremes in optical radiation is used, i.e., 532 and 1350 nm [4]. The block diagram in Fig. 2 shows the FSO link with WD technique having two channels differing in carrier wavelength used.

2.1 Experimental Setup

To demonstrate the wavelength diversity technique, a free-space optical link is established between two PCs using LASERs with two wavelengths 532 and 1350 nm. The two carrier wavelengths are picked from either extremes of visible spectra to understand the behavior of channel with these wavelengths and observe the system performance.

The link consisted of the following components:

1. USB–TTL converter,
2. LASER driver circuit,
3. LASER,
4. Photodiode,
5. Transimpedance amplifier, and
6. Switching circuit.

2.2 Setup Explanation

When the transmitter transmits the data, the USB-to-TTL converter converts the data into 10-bit UART format. This data is passed to the LASER driver circuit which

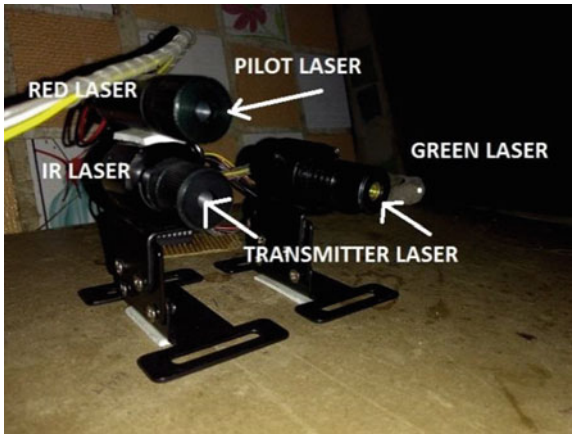


Fig. 3 Setup at transmitter's end

operates the LASER in a desired operating range. As shown in Fig. 3, three LASERs are used to establish the link. The two LASERs having wavelength 532 and 1350 nm modulates and carries the data. However, 1350 nm is not visible to the human eyes. So, to locate the IR LASER's beam at receiver's end to establish the setup, a pilot red LASER in visible region is mounted just above it. The green LASER with 532 nm wavelength is affixed on the platform along with the IR LASER. So, the path of both the LASERs remains parallel to each other and can be easily located on the receiver's end. Figure 3 shows the LASER setup at the transmitter's end.

At receiver's end, two photodiodes are placed at the focal point of optical lenses. These lenses collect all the light coming from the LASERs and focus to the photodiodes. These photodiodes are kept apart from each other at a distance same as that of the LASER. The receiver's setup is shown in Fig. 4.

Current is produced based on the intensity of the light received. Current signals are fed to the transimpedance amplifier to convert it to voltage which goes to multiplexer as two data lines. Simultaneously, these data signals are averaged through low-pass filter and fed to input of comparator: One input at inverting end and one at non-inverting end. Output of the comparator is going to the selection line of 2 × 1 Multiplexer. Now, comparator output would be either logic 0 or 1 based on the

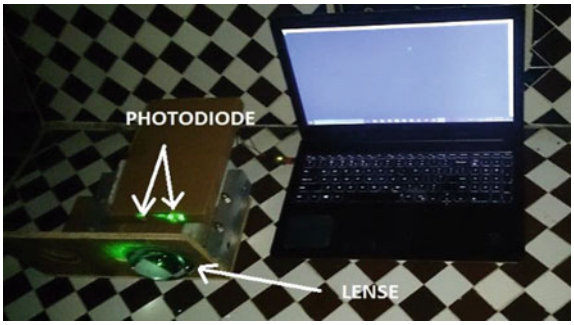


Fig. 4 Setup at receiver's end

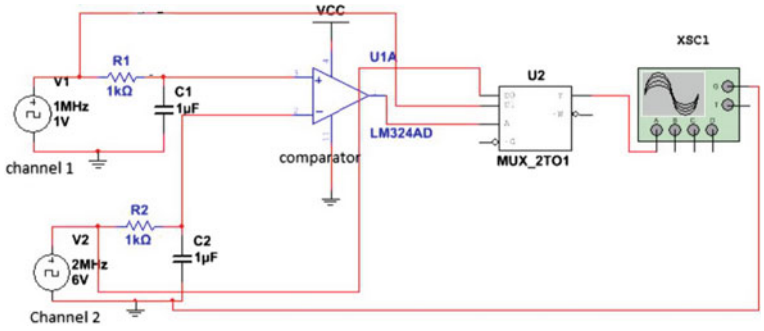


Fig. 5 Schematic of the switching circuit

voltage level of the two data paths, which will set the select line of MUX and the channel having higher voltage would pass to further stages.

The data received is passed to the USB-to-TTL converter and displayed on the serial monitor. Moreover, as the multiplexer circuit works on 5 V voltage level, the signals obtained at the output will be either +5 V or 0 V. Hence, all the unnecessary noises will be removed.

Thus, the selection circuit model dynamically selects channel which performs better based on amplitude of signal. The schematic of the selection circuit can be as shown in Fig. 5.

3 Results and Discussion

3.1 Simulation Theory

The practical results were obtained under normal weather conditions where the attenuation remains nearly around 2–3 dB/km. However, to test fog weather conditions, simulations were performed on OptiSystem 15.1.1 software setting the channel characteristics as that of foggy weather [5]. Channel modeling follows Eq. (1) mentioned. Also, Gamma–Gamma scintillation models to account for the atmospheric fading.

$$P_{\text{Received}} = P_{\text{Transmitted}} \frac{d_{\text{R}}^2}{(d_{\text{T}} + \theta R)^2} 10^{-\alpha \frac{R}{10}} \tag{1}$$

where d_{R}: Receiver aperture diameter (m), d_{T}: Transmitter aperture diameter (m), θ: Beam divergence (mrad), R: Range (km), and α: Atmospheric attenuation (dB/km).

Based on the wavelength of optical source and atmospheric conditions, beam divergence (θ) and attenuation constant (α) were selected. Both beam divergence and attenuation constant highly depend on the size of the particulate matter present in the atmosphere. It leads to two phenomena: scattering and absorption. Under fog conditions, hydrated particles suspended in atmosphere having size of 1um which leads to Mie scattering for light of visible wavelength region. It contributes on an average ~41–68% of total light scattering during fog periods. Hence, light having wavelength in IR region (1350–1550 nm) is preferred.

3.2 Simulation Results

Simulation is performed with the optical source of wavelengths 532 and 1350 nm same as that of actual experiment.

Figures 6, 7, 8, and 9 show the effect on the performance of FSO link with the two wavelengths used under different atmospheric conditions. The comparison of

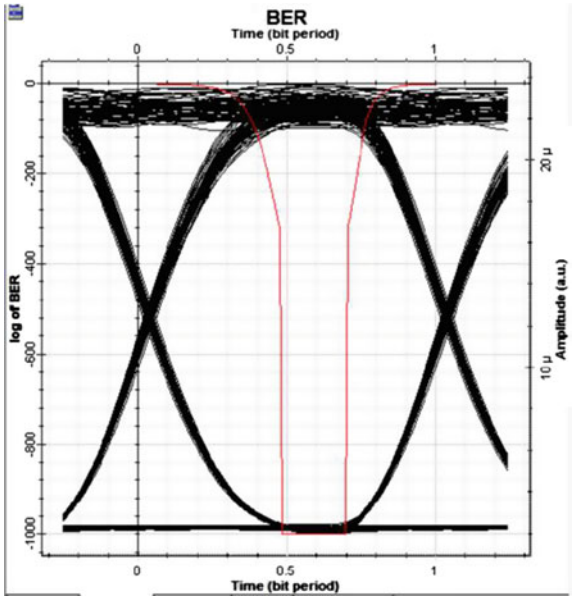


Fig. 6 532 nm_clear_weather

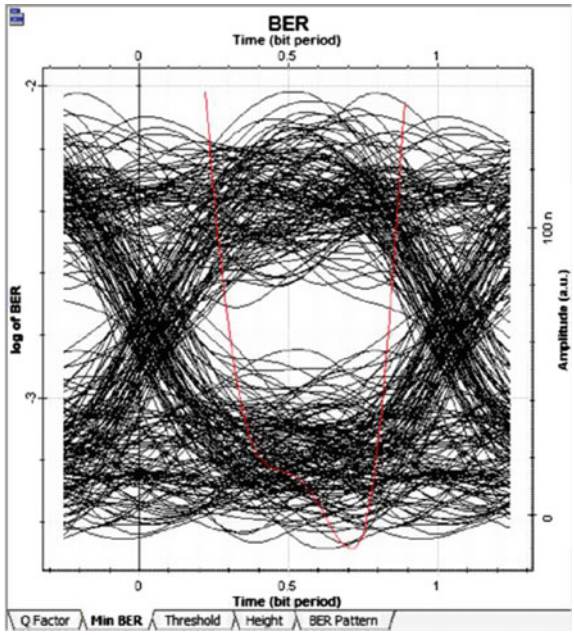


Fig. 7 532 nm_light_fog

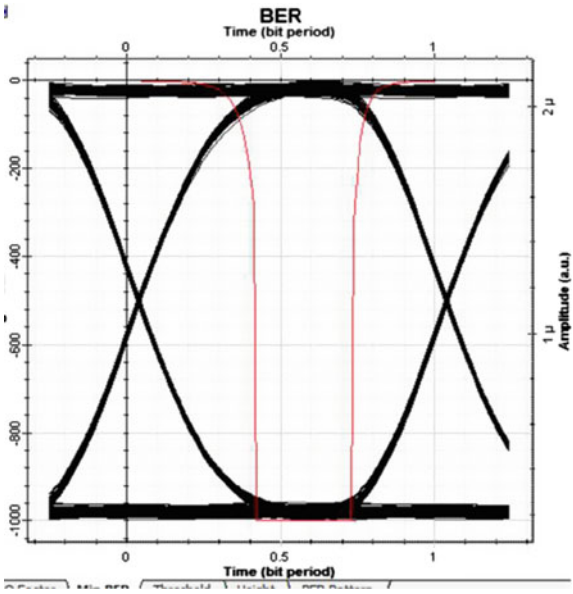


Fig. 8 1350 nm_clear_weather

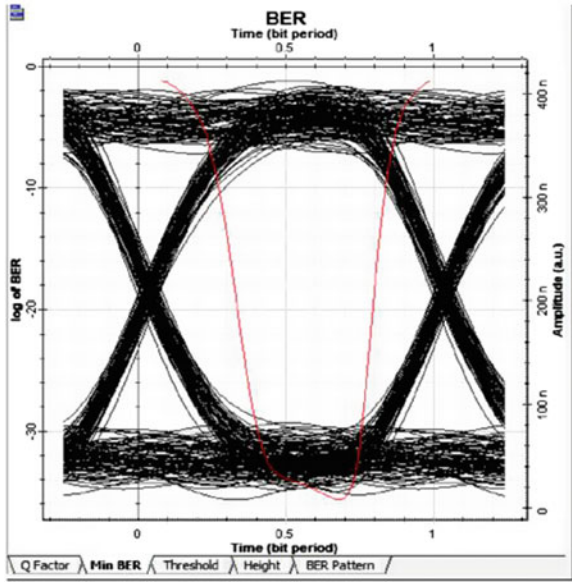


Fig. 9 1350 nm_light_fog

Table 1 Practical results showing amount of the signal received at receiver's end

Baud rate	Received voltage at 500 m (%)	Received voltage at 1 km (%)
4800	85	83
9600	98	95
19200	92.5	86

eye diagram and BER is measured at the receiver for the two wavelengths used to implement wavelength diversity, i.e., 532 and 1350 nm under clear ($\alpha = 3$ dB/km, $\theta = 0.25$ mRad) and light fog ($\alpha = 50$ dB/km, $\theta = 0.5$ mRad) weather conditions. The data rate of FSO link is 1 Mbps. It shows that during light fog the optical source with wavelength 532 nm is giving BER of only 10^{-3} and almost closed eye which degrades the performance by huge margin. However, if under same fog condition if 1350 nm light source is used, it gives BER of 10^{-30} with around 260 mV of eye height at same data rate. So, it provides much better performance under foggy conditions. Even 1350 nm can be used for clear weather, but a study suggests that higher wavelength light source needs higher power to operate [6]. To perform experiments, BER can be measured using Bit Error Rate Tester (BERT) instrument [7].

Hence, 1350 nm wavelength source will consume extra power for comparable performance as that of 532 nm light source. So, these criteria clearly indicate the need for implementation of wavelength selection technique having a careful optimization of operating wavelength to achieve better performance with minimum possible power consumption under varying weather conditions.

3.3 Experimental Results

FSO link was established with the WD technique implemented. Distance between the transmitter and the receiver was nearly 500 m. Later, it was increased to approximately 1 km. To check the working of link, string of characters was sent from the transmitter's end. By the number of characters received at receiver, the results were noted. The link was tested on different baud rates having maximum baud rate at 19,200. However, the best results were obtained with 9600 baud. Also with increase in the distance, the results got depleted by nearly 4%. These results can be shown in Table 1.

All the practical results were obtained under normal weather conditions as it is difficult to create exact fog condition practically. Here, the attenuation remains nearly around 1–2 dB/km. As the distance increased, the voltage generated by photodiode at the receiver's end decreased. The experiments were repeated with 1350 nm LASER. It was found that results were better with 1350 nm LASER compared to the 532 nm LASER under clear weather. This indicates that simulation and practical results were the same for clear atmospheric conditions.

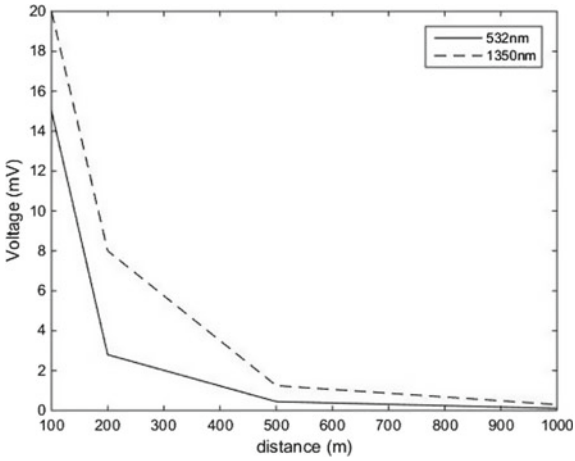


Fig. 10 Voltage versus distance graph for 532 and 1350 nm wavelength LASERs

Figure 10 shows the attenuation in voltage of the received signal at the photodiode. It shows that as the distance increases received voltage decreases due to attenuation. And 1350 nm gives higher voltage margin compared to 532 nm at the same distance. This proves that 1350 nm performs better compared to 532 nm matching the simulation. Also as the wavelength increases, specific molecular absorption (dB/km) reduces, hence higher wavelength is preferred for longer distances [8].

4 Conclusion

In this paper, simulation and implementation of wavelength diversity technique on FSO link to improve the performance under varying atmospheric conditions is reported. It includes the design of receiver circuit for channel selection based on voltage margins and a working setup of WD technique. Simulations were performed to understand the effect on received signal with optical source of wavelengths 532 and 1350 nm under clear and light fog conditions. The obtained results clearly indicate that light source of wavelength 1350 nm provided much lower BER compared to 532 nm source. However, 532 nm consumes lesser power providing comparable performance under clear weather. So, the operating wavelength was carefully chosen to achieve better performance under different climatic conditions. Based on the simulation, FSO link was established with 532 and 1350 nm LASERs, and the results showed better voltage margin with 1350 nm LASER. Results were correlated with the simulated results justifying the need of wavelength diversity.

References

1. Wasiczko LM (2004) Techniques to mitigate the effects of atmospheric turbulence on free space optical communication. PhD thesis
2. Shah D, Kothari DK (2015) BER Performance of FSO link under strong turbulence with different Coding Techniques. IJCSC 8:4–9
3. Hajjarian Z, Fadlullah J (2009) MIMO free space optical communications in turbid and turbulent atmosphere. J Commun 4(8):524–532
4. Wainright E, Refai HH, Sluss JJ (2005) Wavelength diversity in free-space optics to alleviate fog effects. In: SPIE, Bellingham, proceedings of SPIE, vol 5712
5. Tool from Optiwave, OptiSystem_Getting_Started.pdf. Optical Communication System Design Software v12
6. Kaushal H, Kaddoum G (2015) Free space optical communication: challenges and mitigation techniques. Cornell University Library, June 2015
7. Kvcala R, Kvicera V, Grbner M, Fiser O (2009) BER and availability measured on FSO link. Radio Eng 16(3):7–12 (September 2007). Maitra A, Das S, Shukla AK (2009) Joint statistics of rain rate and event duration for a tropical location in India. Ind J Radio Space Phys 38:353–360 (December 2009)
8. Rouissat M, Borsali AR, Chikhbled ME (2012) Free space optical channel characterization and modeling with focus on algeria weather conditions. J Comput Netw Inf Secur 3:17–23

Moment of Inertia-Based Approach to Recognize Arabic Handwritten Numerals

Binod Kumar Prasad

Abstract Simpler methods of feature extraction and better accuracy have always been primary needs of a handwriting recognition system to be a successful real-time system. The novelty of this paper lies in the introduction of two unique methods of feature extraction which are Pixel Moment of Inertia (PMI) and Delta Distance Coding (DDC). PMI absorbs angular variations in the samples and DDC performs précised local curve coding for better recognition accuracy. Multiple Hidden Markov Model (MHMM) has been used to neutralize the effect of two very frequent writing styles of numerals "4" and "7" on their recognition rates. The paper uses MNIST database, and overall recognition accuracy of 99.01% has been achieved.

Keywords Pixel moment of inertia (PMI) · Delta distance coding (DDC) · Multiple hidden Markov model (MHMM)

1 Introduction

The use of digital computers has increased manifold because of the craze for digitization. The digital machines are to recognize the handwritten characters or numerals in postal addresses on envelopes, amount on bank checks or data filled in several forms, etc. The individuality of one's handwriting makes it more challenging to develop a recognition system with maximum efficiency. In [1], authors have proposed a neural network-based character classification system. Features have been derived from star layered histogram. Center of gravity (CoG) is determined from contour of character region, and several equal-distance apart lines are considered from CoG to the contour. The first point of intersection of the line with character gives rise to the first layer of histogram, and so on unless the line touches the boundary of the region containing the character. The system shows an overall accuracy of 97.1% on MNIST database.

B. K. Prasad (✉)
ECE Department, Guru Nanak Institutions Technical Campus,
Ibrahimpatnam, Telangana, India
e-mail: binod.krpr@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_26

A multistage cascaded recognition scheme using wavelet-based multiresolution representations and multilayer perceptron (MLP) classifiers have been used to recognize mixed numerals of Indian script Devanagari and Bangla along with English numerals in [2]. The work has introduced two large databases for the above scripts taken from real life just like MNIST data but for English numerals, MNIST dataset itself is considered. It misclassified 0.81 and 1.01% of training and test MNIST English numeral samples, respectively, to report overall recognition rate of 98.79%. In [3], authors have presented an evolution of well-known effective KNN classifier termed as Li-K nearest neighbor algorithm to recognize handwritten numerals in MNIST dataset. As per the proposed theory, the complex samples are represented into elements in line with the Lie group data structure and then Lie algebra is used to simplify the distance calculations. Classification is done by means of Li-KNN algorithm. They have found minimum and about steady error rate beyond $K = 11$ in case of numerals "1" and "9". Li-KNN method has shown an average error rate of about 21.02% which is lesser than that with LieMeans and LieFisher methods of classification. Authors have used holographic nearest neighbor algorithm (HNN) with features based on moment of inertia about centroid and topological characteristics of image to recognize letters of English alphabet [4]. Wen et al. [5] used a novel shadow removal technique with SVM integration method to read numerals in license plate. In [6], authors have showed the performance of handwritten digit recognition system according to (i) the effect of input dimension, (ii) effect of kernel functions, (iii) comparison of different classifiers (ML, MLP, SOM and LVQ, RBF, SVM), and (iv) comparison of three types of multiclass SVMs (one-to-others, pair-wise, decision tree). Researchers are focused on different directional features especially gradient features to read characters and numerals [7–9]. Though various methods have been reported till now, the offline handwriting recognition is still an active area for research toward exploring the newer techniques to improve recognition accuracy [2]. This paper uses multiple hidden Markov model (MHMM) as classifier where the inputs are feature elements obtained from PMI and DDC. The paper organization is as follows.

Section 2 details out methods of feature extraction; Sect. 3 describes the classifiers, whereas Sect. 4 deals with experiments and results and finally, conclusion has been drawn in the very next section.

2 Feature Extraction

2.1 Pixel Moment of Inertia (PMI)

In order to absorb angular variations in the numeral sample images, moment-based feature, i.e., Pixel Moment of Inertia is extracted. Sample image is of size 50×50. It is divided into 25 blocks (zones) by means of a window of size 10×10. The pixel at the center of each block is taken as the reference element (Fig. 1), and moment of

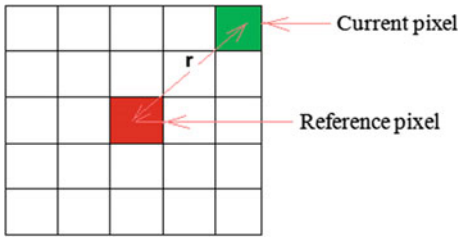


Fig. 1 Frame for calculation of PMI

inertia of all other elements is calculated about that every element with the help of the formula:

$$I_k = \sum_{i=1}^{100} m_i r_i^2; \quad k - 1, 2, 3 \ldots, 25$$

where m_i is taken as pixel value of current element and r_i, the distance of that very element from the reference element. The normalized value of I_k is the actual feature element. The number of feature elements for PMI of each sample is 25.

2.2 Delta Distance Coding (DDC)

The sample image of size 50 × 50 has been divided into four equal sub-images each of size 25 × 25. Twenty-five equally spaced samples, on each sub-image curve, are considered and their distance from upper edge (reference level) is noted (Fig. 2).

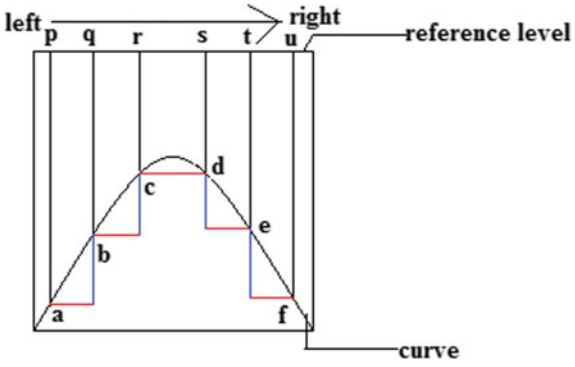


Fig. 2 Schematic diagram of DDC

If there are more than one sample on same vertical line, the average of respective distances is considered. The coding is done based on whether the distance of a sample is greater or smaller than or equal to that of the just previous sample. The reason behind taking 25 samples is to get very adjacent pixels such that the coding could represent the local curve more authentically. Total feature elements out of this feature $= 25 \times 4 = 100$, i.e., F (DDC) $= 100$.

3 Classifiers

The paper uses multiple hidden Markov model (MHMM) as classifier. It is now an established fact in literature that HMM-based systems being segmentation-free outperform other segmentation-based approaches [10–14]. They are effective at boosting the efficiency of inertial-sensing-based online handwritten numeral and character recognition also [15–17]. We have used left-to-right chain model to build hidden Markov model for each of the ten English numerals.

The paper proposes the concept of multiple HMM (Fig. 3) to counter drastic variations in samples. The numerals "4" and "7" are generally written in two different forms. While training, two different HMM sub-models named "HMM 4a" and "HMM 4b" in case of digit "4" and "HMM 7a" and "HMM 7b" in case of digit "7" have been developed from the two different sets of the numerals. For rest of the numerals, only one MHMM model is formed. The outcome of the comparator is the result of the proposed classifier. As per [18], the number of states has been increased gradually from 5 up to 24 when satisfactory results were obtained as shown in the results section. In the process of training, optimization has been done by Baum–Welch algorithm and Viterbi algorithm has helped in the process of recognition by finding the maximum probability.

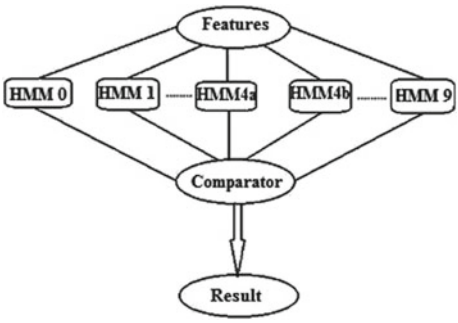


Fig. 3 Proposed Multiple HMM classifier

4 Experiments and Results

This paper uses MNIST database. The training process of the classifier MHMM is executed with the features extracted from training samples of a numeral to form its model. For a testing sample, the maximum probability of matching with any one of the trained models is found by the comparator (Fig. 3) and numeral corresponding to that very model is the result of classification. The whole experiment has been conducted in the following three phases as shown in Table 1—In phase 1, feature elements out of PMI and in phase 2, those out of DDC have only been considered for training and testing of the classifier. In phase 3, both the features have been used together to achieve better accuracy of the system.

The individual accuracy of all the classes in three phases has been tabulated in Table 2. It is obvious that the results in phase 3 are far better than the rest two phases. It is because both the features PMI and DDC when taken together, extract finer details of the respective sample which help to get better accuracy of that class while testing. The results of Table 2 have been shown in Fig. 4 by means of vertical bar diagram for better comparative view. The overall accuracy of the proposed system with PMI and DDC is **99.01%**. In Table 3 and Fig. 5, the class-wise percent of recognition accuracy of the proposed system has been compared with result of another reported work [19] on the same MNIST database. Features from maximum profile distance in four directions, the presence of water reservoirs, fill hole density, etc. have been considered in the paper.

Table 1 Phases of experiment

Phase	Features	Tool
1	PMI	MHMM
2	DDC	MHMM
3	PMI + DDC	MHMM

Table 2 Recognition rate with different feature sets

Digits	Percent recognition rate with		
	PMI	DDC	PMI and DDC
0	88.2	90.3	99.7
1	86.5	91.0	99.8
2	87.8	89.5	99.7
3	79.4	83.8	97.8
4	81.2	83.4	98.6
5	80.6	87.7	98.9
6	89.0	89.6	99.5
7	88.4	88.9	99.6
8	78.7	84.6	97.8
9	88.3	89.7	98.6

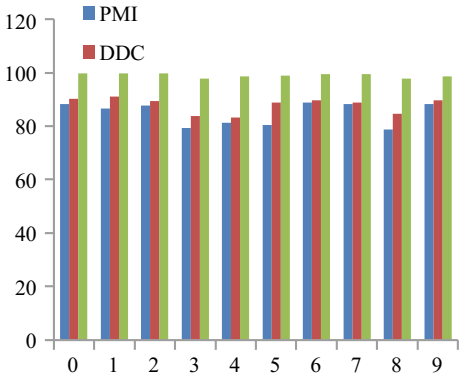


Fig. 4 Bar diagram to compare the efficacy of feature(s)

Table 3 Comparison of results for individual digits

Digits	% Accuracy in	
	Ref. paper [19]	Current paper
0	96.47	99.7
1	97.95	99.8
2	97.26	99.7
3	97.56	97.8
4	97.27	98.6
5	97.47	98.9
6	96.09	99.5
7	95.72	99.6
8	95.71	97.8
9	97.23	98.6

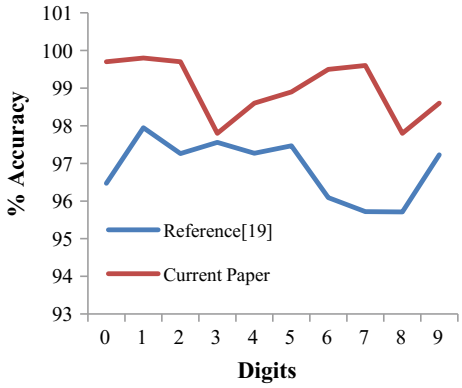


Fig. 5 Line diagrams comparing results of the two systems

Table 4 Comparison of recognition and error rates of different systems

First Author of work	Reference	% Recognition	% Error
Mayraz	[20]	98.30	1.70
Zhao	[21]	91.24	8.76
Karungaru	[1]	97.10	2.90
Babu	[19]	96.94	3.06
Proposed	NA	99.01	0.99

Overall performance of the current system has been compared with other reported methods applied on MNIST database (Table 4), so that better recognition efficiency of the proposed system could be ratified.

5 Conclusion

The current system applies pixel moment of inertia (PMI) and delta distance coding (DDC) to support Multiple Hidden Markov Model (MHMM) in order to recognize handwritten English numerals. PMI tackles angular variations in handwritten samples, whereas DDC encodes a curve precisely. MHMM takes care of intra-class variations of some specific numerals, i.e., "4" and "7". The system has achieved competitive recognition rate of 99.01% by dint of such effective features and classifier. Due to interclass similarity between some handwritten numeral pairs like "3" and "8", their recognition accuracies are, respectively, lower. In future works, some more robust features are to be applied to differentiate such pairs so that overall efficiency of the system could be hiked further.

Acknowledgements MNIST database is a universally accepted benchmark database of English or Arabic numerals available on the link "Y. LeCun [MNIST OCR data]". This database has been used in this proposed work. I am thankful to the authority of Guru Nanak Institutions Technical Campus, Ibrahimpatnam, T.S. India to provide us with Internet and other allied facilities to carry out research works.

References

1. Karungaru S, Terada K, Fukumi M (2013) Hand written character recognition using star-layered histogram features. In: SICE annual conference 2013, Nagoya University, Japan, p 1151
2. Bhattacharya U, Chaudhuri B (2009) Handwritten numeral databases of indian scripts and multistage recognition of mixed numerals. IEEE Trans Pattern Anal Mach Intell 31(3):444–457
3. Wang B, Zhang L, Wang X (2013) A classification algorithm in Li-K nearest neighbor. In: Fourth global congress on intelligent systems. ISBN: 978-1-4799-2886-6/13 (IEEE)
4. Torres LA, Méndez JC, Luis E, Gómez G (2000) Translation, rotation, and scale-invariant object recognition. IEEE Trans Syst Man Cybern Part c Appl Rev 30(1):125–130

5. Wen Y, Lu Y, Yan J, Zhou Z, von Deneen KM, Shi P (2011) An algorithm for license plate recognition applied to intelligent transportation system. IEEE Trans Intell Transport Syst 12(3):830–845
6. Zhao B, Liu Y, Xia SW (2000) Support vector machines and its application in handwritten numerical recognition. In: 15th international conference on pattern recognition, vol 2, pp 720–723
7. Ding K, Liu ZB, Jin LW, Zhu XH (2007) A comparative study of Gabor feature and gradient feature for handwritten Chinese character recognition. In: ICWAPR, Beijing, China, pp 1182–1186
8. Fujisawa H, Liu CL (2003) Directional pattern matching for character recognition. In: 7th ICDAR, Edinburgh, Scotland, pp 794–798
9. Liu CL, Nakashima K, Sako H, Fujisawa H (2004) Handwritten digit recognition: investigation of normalisation and feature extraction techniques. Patt Recogn 37(2):265–279
10. El Abed H, Margner V (2011) ICDAR—Arabic handwriting recognition competition. Int J Doc Anal Recogn 14:3–13
11. El-Hajj R, Likforman-Sulem L, Mokbel C (2005) Arabic hand-writing recognition using baseline dependent features and hidden markov modeling. In: Eighth international conference document analysis and recognition, pp 893–897
12. Mokbel C, Abi Akl H, Greige H (2002) Automatic speech recognition of Arabic digits over Telephone network. In: Research trends in science and technology, RSTS'02, Beyrouth
13. Natarajan P, Saleem S, Prasad R, MacRostie E, Subramanian K (2008) Multilingual off-line handwriting recognition using hidden Markov models: a script independent approach. Springer Book Chapter on Arabic and Chinese Handwriting Recognition, vol 4768, pp 235–250. ISSN: 0302-9743
14. Rabiner LR (1998) A tutorial on Hidden Markov Models and selected applications in speech recognition. In: Proceedings of the IEEE, vol 77, no 2, pp 257–286
15. Kallio S, Kela J, Korpipää P, Mäntyjärvi J (2006) User independent gesture interaction for small handheld devices. Int J Patt Recogn Artif Intell 20(4):505–524
16. Kim S, Park G, Yim S, Choi S, Choi S (2009) Gesture-recognizing hand-held interface with vibrotactile feedback for 3D interaction. IEEE Trans Consum Electron 55(3):1169–1177
17. Hsu YL, Chu CL, Tsai YJ, Wang JS (2015) An inertial pen with dynamic time warping recognizer for handwriting and gesture recognition. IEEE Sens J 15(1):154–163
18. Abou-Moustafa KT, Cheriet M, Suen CY (2004) On the structure of hidden Markov models. Patt Recogn Lett 25:923–931 (Elsevier)
19. Babu U, Venkateswarlu Y, Chintha A (2014) Handwritten digit recognition using K-nearest neighbour classifier. In: Proceedings of world congress on computing and communication technologies
20. Mayraz G, Hinton GE (2002) Recognizing handwritten digits using hierarchical products of experts. IEEE Trans Patt Anal Mach Intell 24:189–197
21. Zhao Z, Liu CL, Zhao M (2013) Handwriting representation and recognition through a sparse projection and low-rank recovery framework. Int Joint Conf Neural Netw (IJCNN), 1–8

Fault-Tolerant Techniques for Wireless Sensor Network—A Comprehensive Survey

M. Shyama and Anju S. Pillai

Abstract Wireless Sensor Network (WSN) has become closer to our day-to-day life and has boundless future potentials. As the demand and popularity increases, the responsibility to make the network more reliable also increases. These lead to the development of fault-tolerant WSN. Usually, WSNs are used for monitoring or controlling an unattended or hazardous area, which needs a highly reliable fault-tolerant system. A fault-tolerant system will continue its service even in the presence of a fault and has to detect the faults and recover the system from the failure. Thus, a fault-tolerant system has a lot of requirements. The focus of this paper is on exploring various fault-tolerant techniques for WSN. This paper examines the exigency of developing a fault-tolerant wireless sensor network and elucidates the mode of attaining a fault-tolerant WSN.

Keywords Fault-tolerant techniques · Fault recovery techniques · Fault detection techniques

1 Introduction

Wireless sensor network is a collection of autonomous sensors connected through a self-configured and self-organized network which is wireless. In this modern era, everything we deal with demands autonomous operation, which in turn needs a lot of sensors and processors. The applications of WSNs are mainly in unattended and hazardous areas where energy saving and stable networking are main constraints. The nodes are deployed in hazardous area which cannot be recharged and replaced manually. Because of these unsolved issues, WSNs are very prone to failure. Faults

M. Shyama (✉) · A. S. Pillai
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: shyamaanish@gmail.com

A. S. Pillai
e-mail: s_anju@cb.amrita.edu

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_27

are inevitable, but careful measures are to be taken to prevent the fault by not transmitting from the lower level to higher level and create a hazardous situation. The only solution to save a WSN from being hazardous is making it a fault-tolerant network. There are lots of methods for fault tolerance. This paper presents a comprehensive survey of fault-tolerant techniques, fault detection, and fault recovery classification.

2 Fault Tolerance in Wireless Sensor Networks

Fault tolerance is one of the most desirous characteristics of a sensor network. Network lifetime is the main design considerations of a WSN so that the sensor nodes are alive as long as possible. It is impossible in most of the cases to troubleshoot a sensor node manually. Design of a fault-tolerant system is the only solution to this. Starting from the design phase, possible faults, environmental disasters, detection of those faults, impact of the faults in the network and its solution, etc. should be foreseen. Even if a node fails, the network should be available to the user without any interruption. A complete fault-tolerant system ensures the availability, dependability, and reliability of a system irrespective of the presence of faults [1–7]. The following section briefs the idea about fault-tolerant system requirements.

2.1 Faults

A fault is an abnormal condition of a component or a system which will lead to an error. Faults can be active or passive. If the faults are continuing to exist in the system, it will lead to an error in the network operation, thereby deviating from the expected services. A fault can occur at different levels and layers [1, 3, 6, 7] as shown in Fig. 1.

Node level. Faults can occur in node components. It can be hardware or software.

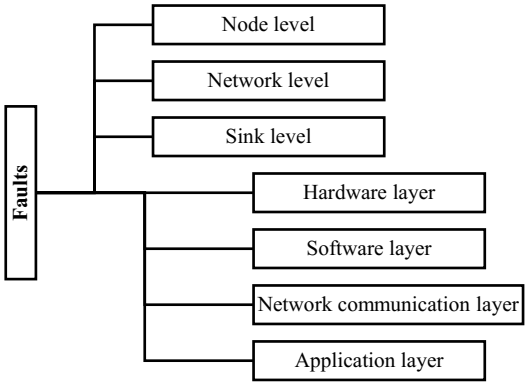


Fig. 1 Classification of faults

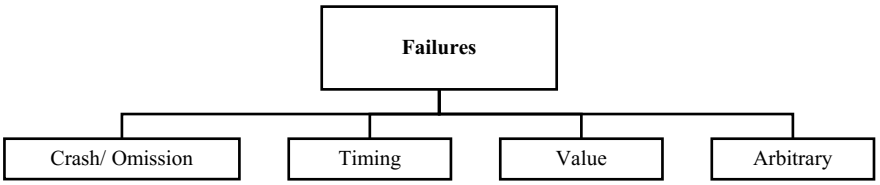


Fig. 2 Classification of failures

Network level. Even if the sensed data are correct, a system failure can occur due to an error in data collection, data routing, or due to an error in a routing protocol.

Sink level. A fault in sink node will cause a disaster in the system.

Hardware layer. This layer mainly includes the component failure of sensing nodes.

Software layer. A software layer includes system OS and middleware.

Network communication layer. This layer includes the wireless part in WSN.

Application layer. As the name implies, it is application specific. The protocols, algorithms, techniques, methods, etc. are completely application oriented.

2.2 Failures

A failure is a deviation from the expected service of a system, thereby deviating from its indented functionality. A fault in any component can lead to a system failure. It can be classified as shown in Fig. 2 [1, 6].

Crash/Omission. If a service is not responding to the request due to any physical damage or interference, then a crash failure has occurred.

Timing. This type of failure is happening in systems which is time stringent.

Value. An error in the generated value is the reason for this type of failures.

Arbitrary. All the failures which are not included in the above three categories are categorized as arbitrary failures. Byzantine failure is an example of this type.

2.3 Fault-Tolerant Techniques

Fault tolerance is an unavoidable part of WSN. There are many ways to achieve fault tolerance in wireless sensor networks [3, 4, 8]. A comprehensive summary is presented in Fig. 3A, B.

Fault-tolerant techniques can be classified according to different criteria.

According to time criteria. This criterion classifies the fault tolerance techniques according to what time the technique is applied. They are preventive and curative [3].

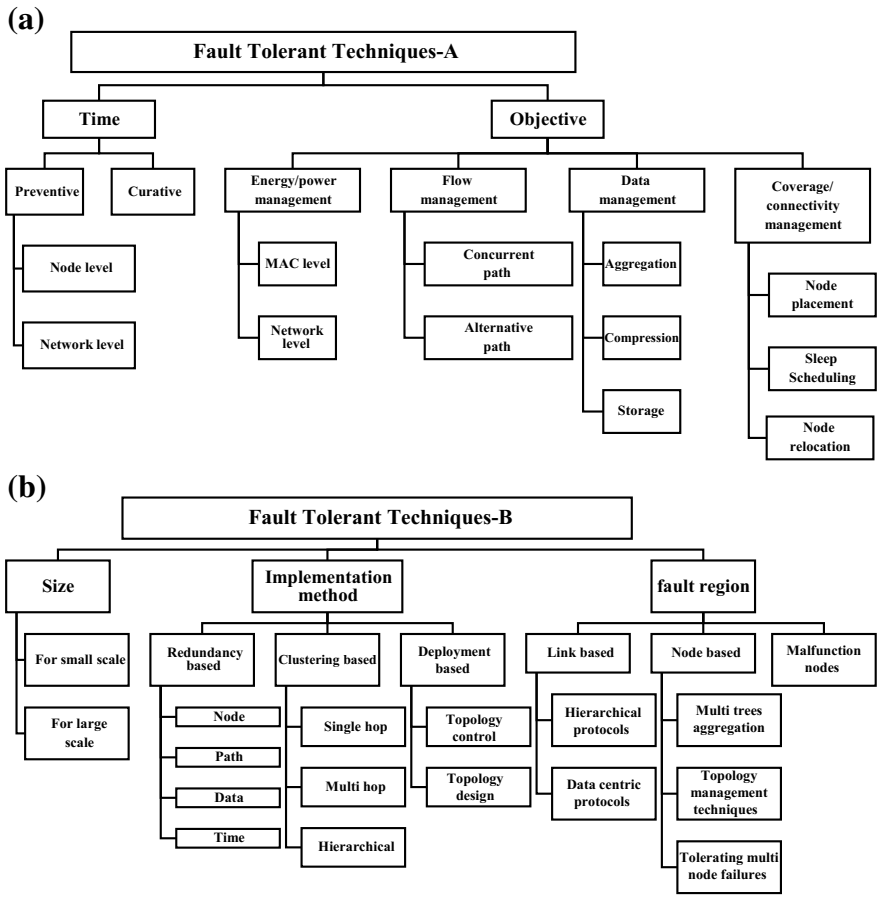


Fig. 3 A, B Summary of fault-tolerant techniques

Preventive. As the name implies, this includes all the techniques that can be used to prevent a failure by implementing an alternative for the failed service. It can be at node level and at network level.

Curative. This includes the techniques that start their duty only after an error is noticed in the functioning of a network. They try to recover the system usually by applying redundancy techniques.

According to main objective. This criterion classifies according to the main objective of a fault tolerance mechanism when it is applied. There can be of four main classifications [3].

Energy/power management. The main goal of this class is to enhance the energy efficiency of a system, reduce energy depletion, and reduce node failure to increase network lifetime. This can be at MAC level and network level.

Flow management. This is to enhance data flow through the network. This can be done by finding reliable routs and base station, thereby establishing the best path between nodes and recovering from path failure.

Data management. Data management techniques take care of data size and information quality.

Coverage/connectivity management. This class of techniques enhances the coverage and connectivity of a network for better service.

According to size. This criterion classifies according to the size of a network [3]. The techniques applied for a small scale may not be suitable for a large scale because of the difference in their network structure.

Small-scale WSN. In a small-scale WSN, each node is very important and has a unique role in producing an output. If a node fails in small-scale WSN, it affects the desired service and the network reliability.

Large-scale WSN. In a large-scale WSN, the entire network may be divided into small clusters. So, if a node fails in any cluster, it will not affect the final service quality and reliability.

According to implementation method. This criterion is according to the implementation method of a fault tolerance technique [4].

Redundancy based. This technique is mainly used for the applications in which energy, security, and reliability are main constraints.

Clustering based. This is the efficient method for fault tolerance compared to the other two in this criterion. This method also is energy efficient.

Deployment based. This technique is based on different strategies of deployment of nodes. The dislocation or connectivity changes of nodes affect the network topology, and a topology control mechanism is essential for improving the network lifetime.

According to the fault region. This classification is according to the area where a fault has occurred [8].

Link based. A link-based fault tolerance has to concentrate on faults happening in the link between nodes.

Node based. This is for faults in nodes. A node-based fault-tolerant system should be energy efficient.

Malfunction nodes. Malfunction nodes may send incorrect data or unnecessary data which in turn can affect the functioning of a system.

2.4 Fault Detection Techniques

Fault detection is an important part of fault tolerance. The fault tolerance can be applied only by identifying the actual fault at correct time. The fault detection techniques can be classified mainly into two: centralized approach and distributed approach [5, 9–12] as shown in Fig. 4.

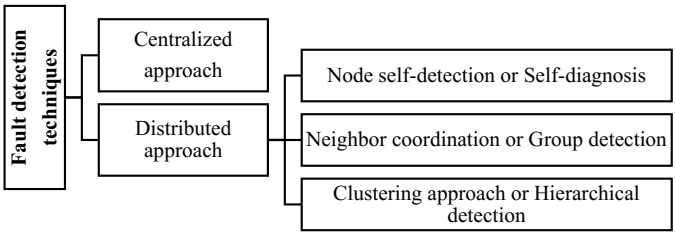


Fig. 4 Various classes of fault detection techniques

Centralized approach. In a centralized approach, usually a base station or a central manager takes the responsibility of fault detection by sending queries periodically.

Distributed approach. In a distributed approach, the detection process is distributed over sensor nodes. A distributed approach can be classified into self-diagnosis, group detection, and hierarchical detection.

Self-diagnosis. Using self-diagnosis, a node can detect faults by itself.

Group detection. In this method, sensor nodes deployed in the same region collaboratively do the detection procedure by assuming similar values are produced.

Hierarchical detection. A detection procedure is used in a hierarchical manner. It starts from the low-level child nodes and forwards to the higher level parent node. The parent node does the aggregation and finds the failure node.

2.5 Fault Recovery

Fault recovery allows a system to continue its service even after the occurrence of a failure. Fault recovery technique depends on the faults detected by the fault detection techniques. So, fault detection and fault recovery are the main procedures of a fault-tolerant system. The various fault recovery methods are shown in Fig. 5 [1, 10–12].

Active replication. This approach is suitable for a service where most of the nodes have the same functionality.

Passive replication. In this approach, a primary node will always do the services by saving a backup in a spare node. The three steps in passive replication are fault detection, primary node selection, and service distribution.

2.6 Fault Management Architecture

The architecture of fault management is mainly classified into three: centralized, distributed, and hierarchical [5, 7, 11] as shown in Fig. 6.

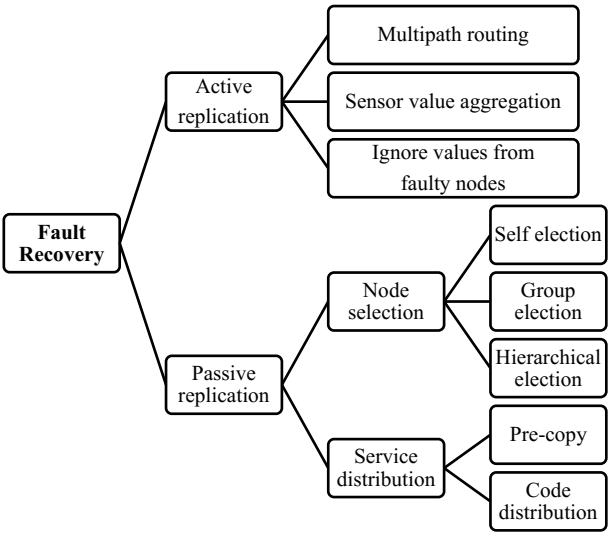


Fig. 5 Various methods for fault recovery

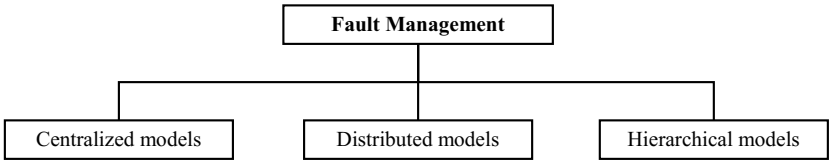


Fig. 6 Fault management architecture

Centralized models. In this approach, fault management is under the control of one centralized controller.

Distributed model. The whole network is divided into small sections with a controller for each part. A central controller also is there to manage the controller of each subsection.

Hierarchical model. This is a combination of centralized and distributed mode.

2.7 Fault Prevention

Fault prevention techniques are for preventing a system from being faulty. It can be done by continuous monitoring of the network [5, 6]. Monitoring of a system can be done by using a base station or by self-monitoring. Figure 7 briefs about the prevention process.

Monitoring node status. The node status includes physical status of the node. It can be any hardware damage, remaining battery status, network status, etc.

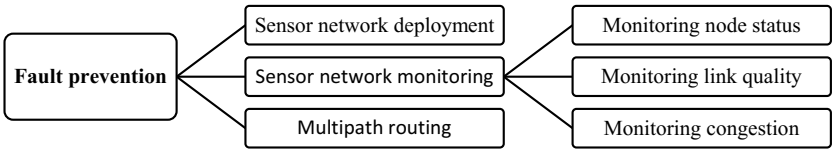


Fig. 7 Fault prevention

Monitoring link quality. The wireless nature is very prone to failure. This sensitive nature makes link monitoring of a WSN more tedious.

Monitoring congestion. Congestion may be occurred due to data overloading. It causes packet loss, packet error, blocking a connection, or delay.

3 Conclusion

Wireless sensor networks made this modern world more autonomous, predictable, and accurate. For making WSN more reliable and dependable, one has to do up-to-date research on each and every aspect of a WSN, mainly in the fault-tolerant property. This paper presents a thorough investigation of mechanisms for attaining fault tolerance in WSN. The main goal of this paper is to understand a fault and a failure and its classifications, and to summarize the techniques for fault tolerance, fault detection, fault prevention, and fault recovery methods which are the crucial aspects of a fault-tolerant system.

References

1. De Souza LMS, Vogt H, Beigl M (2007) A survey on fault tolerance in wireless sensor networks. Sap Research, Braunschweig, Germany
2. Lakshmi CBN, Mohan Rao SK (2016) Enhancing network lifetime of a wireless sensor network using swarm intelligence. Int J Appl Eng Res 11(1):435–439
3. Chouikhi S, El Korbi I, Ghamri-Doudane Y, Azouz Saidane L (2015) A survey on fault tolerance in small and large scale wireless sensor networks. Comput Commun 69:22–37
4. Kakamanshadi G, Gupta S, Singh S (2015) A survey on fault tolerance techniques in wireless sensor networks. In: 2015 international conference on green computing and internet of things (ICGCIoT), pp 168–173. IEEE
5. Paradis L, Han Q (2007) A survey of fault management in wireless sensor networks. J Netw Syst Manage 15(2):171–190
6. Alrajei N, Fu h (2014) A survey on fault tolerance in wireless sensor networks. In: Proceedings of the ASEE North Central Section conference. American Society for Engineering Education
7. Yin MH, Win Z (2014) Fault management using cluster-based protocol in wireless sensor networks. Int J Fut Comput Commun 3(1):36
8. Martinaa M, Vaithiyanadhan V, Raghunadhan A (2015) A survey on categorization of fault tolerance in wireless sensor network. Ind J Sci Technol 8(35)

9. Alhasanat AA, Khitam MA, Abdullah IA, Aws A-Q (2015) Data gathering in wireless sensor networks using intermediate nodes. Int J Comput Netw Commun 7(1):113
10. Kaur EJ, Kaur EP (2015) A survey on fault detection and recovery techniques in wireless sensor networks. Int J Eng Res Gen Sci 3(2) Part 2
11. Kshirsagar RV, Jirapure B (2011) A survey on fault detection and fault tolerance in wireless sensor networks. In: Proceedings of the international conference on benchmarks in engineering science and technology (ICBEST (1)) 3:130–138
12. Bapu BRT, Thanigaivelu K, Rajkumar A (2015) Fault tolerance in wireless sensor networks—a survey. World Academy of Science, Engineering and Technology. Int J Comput Electr Autom Control Inf Eng 9(2):504–509

Preliminary Investigation of Generation of US-Based GPS Signal

S. Radharani, C. Vani and P. Naveen Kumar

Abstract Global Navigation Satellite System (GNSS) generates signals using constellation of satellites from space. The GNSS signals transmit positioning and timing information to GNSS receiver. These receivers use this information to determine their location. There are various GNSSs being developed across the world. They are the United States Global Positioning System, Russia's GLONASS, Europe's GALILEO, China's BeiDou, Japan's QZNSS, and India's IRNSS. Among these, GPS, GLONASS, and GALILEO have the global coverage and BeiDou-I, QZNSS, and IRNSS have the regional coverage. Among all the GNSS, GPS and GLONASS are in operation, whereas all other systems are in the development stage. GPS is a fully developed U.S. Space-based satellite navigation system, which provides its services to the military as well as civilian users. Generally, GPS uses four satellites to find the position of receiver. GPS transmit signals in L-band (1–2 GHz). The two basic GPS signals are L1 (1575.42 MHz) and L2 signals (1227.6 MHz). There are various applications of GPS like location-based services, surveying, mapping, space and military applications, etc. Also, GPS has few limitations such that it cannot work well in underground, underwater, city with tall buildings, indoor regions, etc. In this paper, the generation of GPS signals by the development of software code is presented.

Keywords GNSS · GPS · L1 and L2 signals

S. Radharani (✉) · C. Vani
Department of CSE, University College of Engineering, Osmania University, Hyderabad 500007, India
e-mail: ramsmcaou@gmail.com

P. Naveen Kumar
Department of ECE, University College of Engineering, Osmania University, Hyderabad 500007, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_28

1 Introduction

The world's foremost satellite navigation system is the U.S. Navy's Transit system. The development of this system started in the year 1958. The Transit satellite was launched in the year 1961 and system started working from 1964. The system was opened to civil use from 1967 and was removed from service in the year 1996. Transit comprised between 4 and 7 low altitude (1,100 km) satellites, each broadcasting at 150 and 400 MHz. Transit has many disadvantages such as that it could not provide accurate position continuously. Hence, Global Positioning System (GPS) was developed. The development of the GPS started in the year 1973, and the security of GPS signal is a major concern from issues such as spoofing.

GPS spoofing is measured as a serious threat to the satellite navigation system. In the process of spoofing, the attacker will choose the targeted device and falsify the time and location in GPS receiver. He broadcast the strong signals with spoofed location and time, and it confuses the receiver device to shift from genuine satellite signals to the simulated spoofing signals. Hence, it is necessary to understand the generation of GPS signal and its protection from spoofing. Therefore, the key objective of this paper is the Generation of US-based Global Positioning System signal by developing the software code. This paper helps in design of software-based GPS signal generator which generates the GPS signal code. This can be fed to hardware for directly transmitting the GPS signal via GPS antenna in real time.

2 Theoretical Background

According to Humphreys et al., an attempt of implementing GPS spoof on a digital signal processor has been made to determine the spoofing effects and to develop defences against civilian spoofing. This paper shows that a spoofing attack would overcome generally well-known user-equipment-based spoofing countermeasures [1]. Wang et al. discussed SDR devices with low cost that have been discovered for GPS spoofing. This method can be used to modify the location as well as the time of intended devices, which poses security threats to location-based services. Hence, they found out the positioning spoofing. It is important for developers and vendors to examine position and time data more closely and more frequently and effectively [2]. Naghswander et al. tried to build up a new hardware platform for GPS attacks, and new attacks against GPS infrastructure. In this paper, it is shown that attacks are successful against consumer and professional receivers [3]. Herring et al. discussed how GPS signal is tracked and also concluded how the signals can be generated and which signals can be tracked [4]. Jovanovic et al. proposed different countermeasures and security schemes to facilitate against spoofing attack and showed that the presence of a spoof can be successfully detected [5]. Jafarnia-Jaharomi et al. discussed spoofing, anti-spoofing scenarios, and GPS vulnerabilities that can be exploited potentially. It was shown that industrial GPS receivers are fairly

vulnerable. Spatial processing-based anti-spoofing techniques might be applied as generic and very effectual countermeasure against most spoofing signals currently envisaged [6]. Capparra et al. aimed to assess the effectiveness of symbol-level techniques for the alleviation of spoofing attacks and analyzed the feasibility of spoofing attacks against GNSS signals protected by NMA [7]. Javaid et al. presented a novel GPS spoofing attack detection and mitigation technique for the challenge of secure and accurate navigation [8].

3 Methodology

After understanding the operating principles of GPS and signal structure, the hardware method of generating GPS signal and the process involved in it is studied. Later, using LINUX OS and Google maps application, a particular location is selected. Using this location, software code of about 30 specific functions is generated in "C" and the code is executed which gives the generated GPS signal in the form of the GPS raw data.

The following are the steps involved in the generation of raw data of the GPS signal (Fig. 1).

Step I. By passing the three arguments, i.e., latitude, longitude, and altitude, the main function will check whether it is valid inputs or not and check each option like format of file, mode of the signal, duration navigation, frequency, and ionosphere delays.
Step II. If given input is not static mode, the main function calls dynamic mode function module. This function will read the National Marine Electronics Association data and user motion file from user and calculate the changing user motion parallel with time duration.
Step III. The main function calls the readRinexNav function. This function will read the ephemeris data from the user given Receiver Independent Exchange Format navigation file. It calculates the data, gives output of an array of SV ephemeris data, and returns number of sets of ephemerides in the file.
Step IV. In this step, bit allocations for I/Q buffer are done.
Step V. In this step, satellite visibility, C/A code generation, generation of subframes, initializing channels, generating navigation message, and carrier phase are initialized.
Step VI. In this step, the receiver antenna pattern is gained.
Step VII. In this step, generation of baseband signal is done.
Step VIII. In this step, updating of the code phase, data bit counters, carrier phase, channel allocations navigation message is done for every 30 s. It returns the output of GPS signal in bin format.

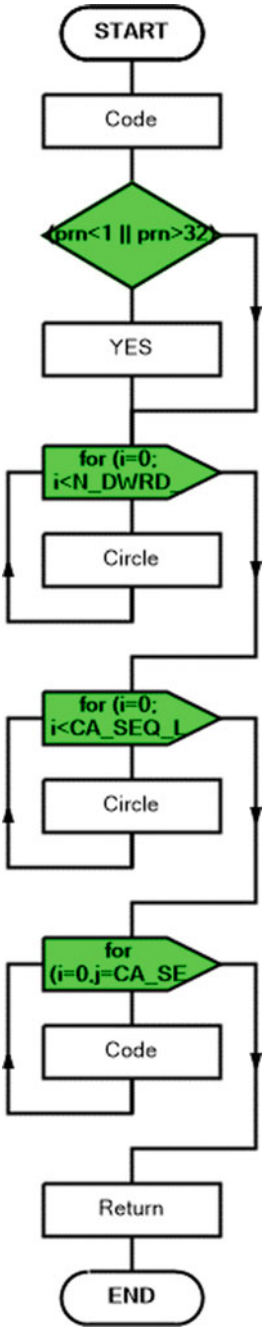


Fig. 1 Flowchart of sample function of generation of code

4 Results and Discussions

In this work, Linux CentOS machine is used to generate the software code for generating GPS signal. A directory has been created with name GPS. After that, the open-source code from GitHub .gpssim.c file has been compiled using $ gccgpssim.c -lm -O3 -o gps-sdr-sim. It compiled productively without any error (Figs. 2 and 3).

To generate a static mode of GPS signal, the latitude and longitude inputs are passed. After that, the Google maps is opened and took random position, i.e., in Cuba (Fig. 4). From there, the location is taken with latitude and longitude parameters as (21.727477°, −78.186467°).

Then, considered these parameters and passed as command line arguments using./gps-sdr-sim -e brdc3540.14n −l 21.727477, −78.186467, 100. Then, terminal started to run time and start creating specified GPS signal for the given location. The process time to generate a.bin file is 401.8 s; after the completion of this process time, a newly created gpssim.bin file is observed (Fig. 5).

gpssim.bin has been keenly studied, observed that the file size is of 2 GB, and also opened it in textpad+ and noticed the raw data in hex format (Figs. 6 and 7).

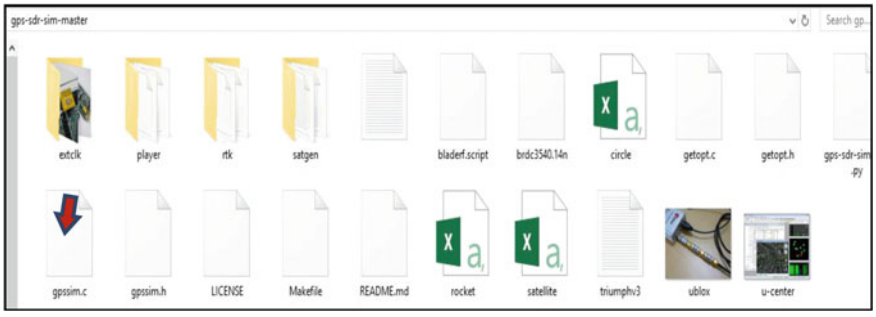


Fig. 2 List of files created with open source

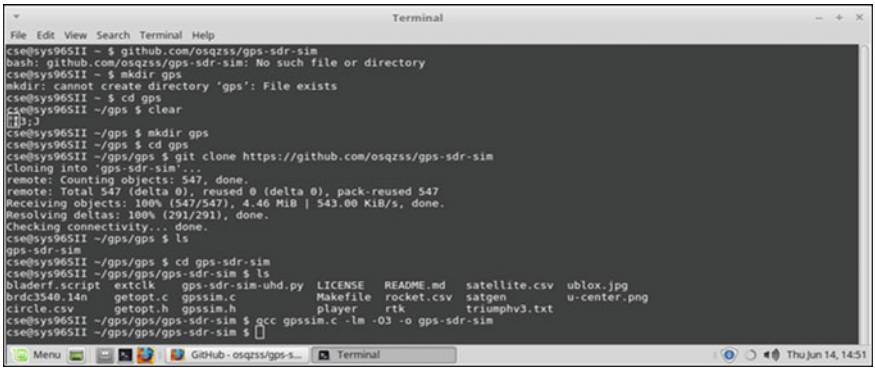


Fig. 3 Output of successful compilation of gpssim.c file

Fig. 4 Cuba location parameters from Google maps

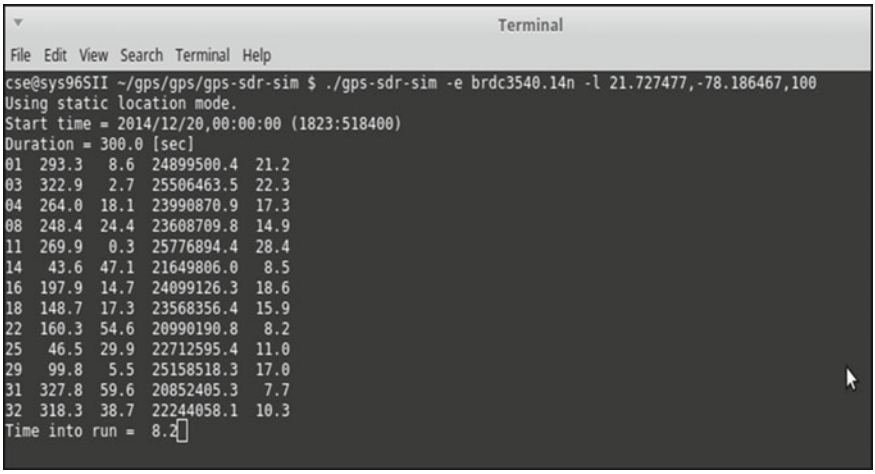


Fig. 5 Process of file generation

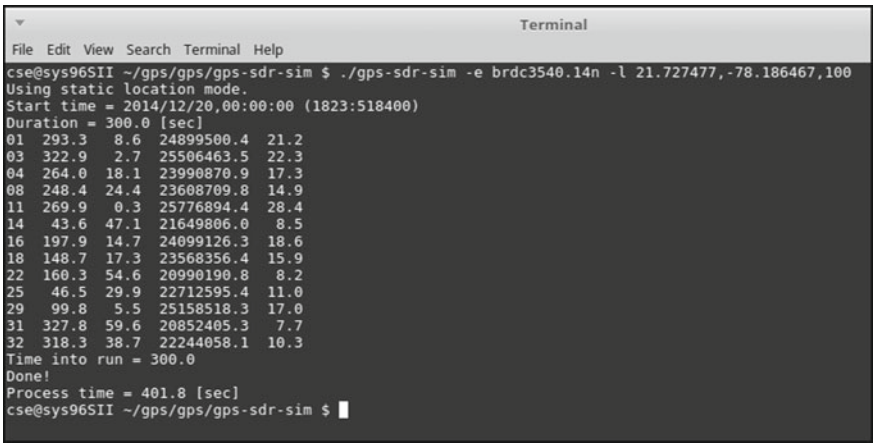


Fig. 6 Generation of GPS signal file

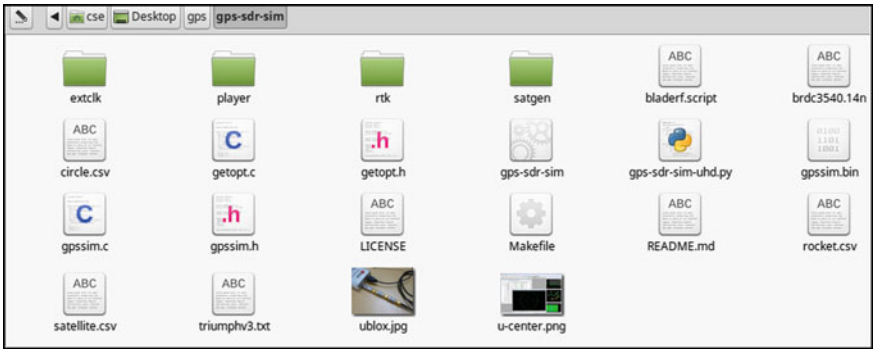


Fig. 7 Newly generated gpssim.bin file

gpssim.bin is file which can be transmitted from the related transmitter, which converts the digital signal to analog signal and transmits. If this RF signal has much signal strength than that of the original satellite GPS signal, receivers will shift from original satellite signal to our simulated signal and the GPS receiver is spoofed.

5 Conclusion

GPS is the oldest satellite navigation system across the world. As of now, simulated GPS signal has been generated that can be transmitted through hardware to spoof the GPS receivers. The same work can be carried for the other GNSS and will be helpful to study spoofing and anti-spoofing techniques.

References

1. Humphreys TE, Ledvina BM, Psiaki ML, O'Hanlon BW, Kintner PM (2008) Assessing the spoofing threat: development of portable GPS civilian spoofer. In: ION GNSS conference, Savanna, Austin, USA, Sept 2008
2. Wang K, Chen S, Pan A (2012) Time and position spoofing with open source project. Research paper, China
3. Naghswander T, Ledvina B, Diamond J, Brumley R, Brumley D (2012) GPS software attacks. Journal paper, North Carolina, USA, Oct 2012
4. Herring T (2012) Principles of GPS, Geoweb.mit.edu, USA, Feb 2012
5. Jovanovic A, Botteron C, Farine P (2014) Multi-test detection and protection algorithm against spoofing attacks on GPS receivers. IEEE paper, Lausanne, Switzerland
6. Jafarnia-Jaharomi A, Broumadan A, Nielsen J, Lachapelle G (2012) GPS vulnerability to spoofing threats and a review of Anti-spoofing techniques. Int J Navig Obs, Canada (July 2012)
7. Capparra G, Laurenti N (2017) Feasibility and limitations of self-spoofing attacks on GNSS signals with message authentication. Conference paper, University of Padova, Italy, Sept 2017
8. Javaid AY, Weiqing FJ (2017) Analysis of GPS—based attacks and a novel GPS spoofing detection algorithm for unmanned aerial vehicle simulation. Journal paper, USA, 2017

Image Compression in Wireless Sensor Networks Using Autoencoder and RBM Method

S. Aruna Deepthi, E. Sreenivasa Rao and M. N. Giri Prasad

Abstract Wireless sensor networks are used in various day-to-day applications. With advanced technology, wireless sensor networks (WSNs) play a key role in networking technologies since it can be expanded without communication infrastructure. There is so much demand for wireless sensor networks that are useful in day-to-day applications. Various techniques like multilayer restricted Boltzmann machine (RBM) network and variational autoencoder (VAE) methods are designed for transmission of images in wireless sensor networks and tested for both gray (2D) and color (3D) images in this paper. The proposed variational autoencoder method is compared with multilayer RBM when the layers are 2, 4, and 8, and the peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR) are compared.

Keywords RBM · Autoencoder · VAE · WSN

1 Introduction

A wireless sensor [1] is equipped with a sensing element and onboard processing, communication, and storage capabilities. Wireless sensor networks have many applications in space study, vehicular movement, and critical movement detection. There are many probabilistic models for deep learning like directed graph models like Bayesian network and undirected graph models like Markov random fields and restricted Boltzmann machine (RBM). It is a shallow, two-layer neural network used

S. Aruna Deepthi (✉)
Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India
e-mail: sadeepthi@gmail.com

E. Sreenivasa Rao
Vasavi College of Engineering, Hyderabad, India

M. N. Giri Prasad
Department of ECE, JNTU College of Engineering, Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India
e-mail: mahendragiri1960@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_29

to create deep models. The first layer is known as the observed layer; the second layer is known as the latent or hidden layer in the RBM. It is a bipartite graph, with no interconnection allowed between variables in the marked layer, and intra-layer communication between the layers. Due to this restriction, the model is termed as a restricted Boltzmann machine. To mathematically formulate the energy function of an RBM, an observed layer that consists of a set of NV binary variables was denoted collectively with the vector v. The hidden or latent layers of NH binary random variables are denoted as h. The RBM is also an energy-based model similar to the Boltzmann machine, where the joint probability distribution is determined by its energy function.

$$p(v, h) = \frac{\exp(-E(v, h))}{z} \tag{1}$$

The model is a symmetric bipartite graph, where each hidden node is connected to every visible node. Hidden units are represented as hi and visible units as vi. The energy function of an RBM with binary visible and latent units is as follows:

$$E(v, h) = -a^T v - b^T h - v^T wh \tag{2}$$

Here, a, b, and W are unconstrained, learnable, real-valued parameters, and the model is split into two groups of variables, v and h. The interaction between the units is described by the matrix W.

2 Proposed Method

The RBM [2, 3] is fed with a dataset from which it learns. For example, for a grayscale image, the lowest level item would be one pixel value of the image, which the visible node would receive. Therefore, if an image dataset has n number of pixels, the neural network processing them must also possess n input nodes on the visible layer.

Now, one single pixel value p is propagated through the two-layer network. At the first node of the hidden layer, [4, 5] p is multiplied by a weight w and added to the bias. The final result is then fed into an activation function that generates the output of the node. The operation produces the outcome, which can be termed as the strength of the signal passing through that node, given an input pixel p. Figure 1 shows the computation involved for a single-input RBM.

Every visible node of the RBM [6] is associated with a separate weight. Inputs from various units get combined at one hidden node. Each p (pixel) from the input is multiplied by a separate weight associated with it. The products are summed up and added to bias. This is sent through an activation function to generate the output of the node. Figure 1 shows the computation involved for multiple inputs to the visible layer of RBMs.

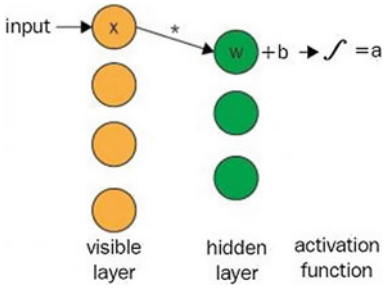


Fig. 1 Computation of an RBM for a one input path

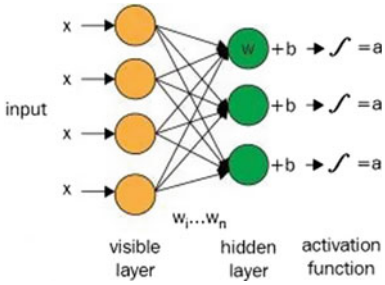


Fig. 2 Figure shows the computation involved with visible units and hidden units

For every hidden node, each input p gets multiplied with its respective weight w. Therefore, for a single-input p and m number of hidden units, the input would have m weights associated with it. In Fig. 2, the input p would have three weights, making a total of 12 weights altogether: four input nodes from the input layer and three hidden nodes in the next layer. All the weights associated with the two layers form a matrix, where the rows represent input nodes, and the columns represent hidden units.

2.1 Variational Autoencoder

A VAE [7] encodes a data sample x to the latent representation z and decodes the latent representation back to data space

$$z \sim \text{Enc}(x) = q(z/x) \tag{3}$$

$$x \sim \text{Dec}(z) = p(x/z) \tag{4}$$

Typically, $z \sim N(0, I)$ is chosen.

$$£_{\text{llike}}^{\text{pixel}} = -E_{q(z|x)}(\log p(x|z)) \tag{5}$$

$$£_{\text{prior}} = D_{kl}(q(z|x)||p(z)) \tag{6}$$

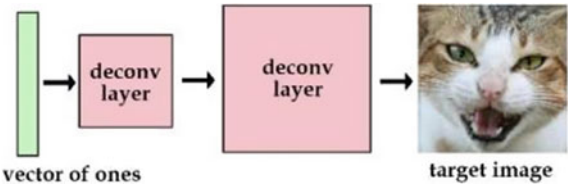


Fig. 3 Image compression using VAE

DKL is the Kullback–Leibler divergence.

A network [8] consists of a few deconvolution layers. The input is set to be a vector of ones. The image compression using VAE is shown in Fig. 3 in which the vector of ones and the de-convolutional layers are considered.

In the example, [1, 0, 0, 0] represents cat image, while [0, 1, 0, 0] represents a dog. In an autoencoder, adding with one component with another component takes the original image and encodes them into vectors. The de-convolutional layers then decode the vectors back to the original images. A simple reparametrized method is applied to optimize the KL divergence. The hidden variable can be taken as a transfer of data. This helps the encoder to be very efficient, creating better hidden variables. This improves the efficiency of the images when they are decoded and the hidden variables are randomly generated.

3 Simulation Results

3.1 Comparison of Improved RBM and VAE

In the performance analysis of RBM, evaluation of reconstructed image is carried out. MATLAB 2017a is used in this paper for simulation. The MNIST [2] and the toy datasets are used in this experiment which are handwritten digital databases.

The graphs shown in Figs. 4 and 5 indicate that the average reconstruction error of the VAE algorithm is always smaller.

3.2 Analysis of the Multilayer RBM Network and Variational Autoencoder

In this section, the 64×64 Lena image is used in 2D grayscale and 3D color image to evaluate the compression performance. The performance of the reconstructed image is evaluated using peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR). Tables 1 and 2 depict the results (Figs 6, 7, and 8).

The compression ratio is in inversely proportional to the number of hidden units. From Table 2, it is inferred that when the number of hidden units is 8, the PSNR

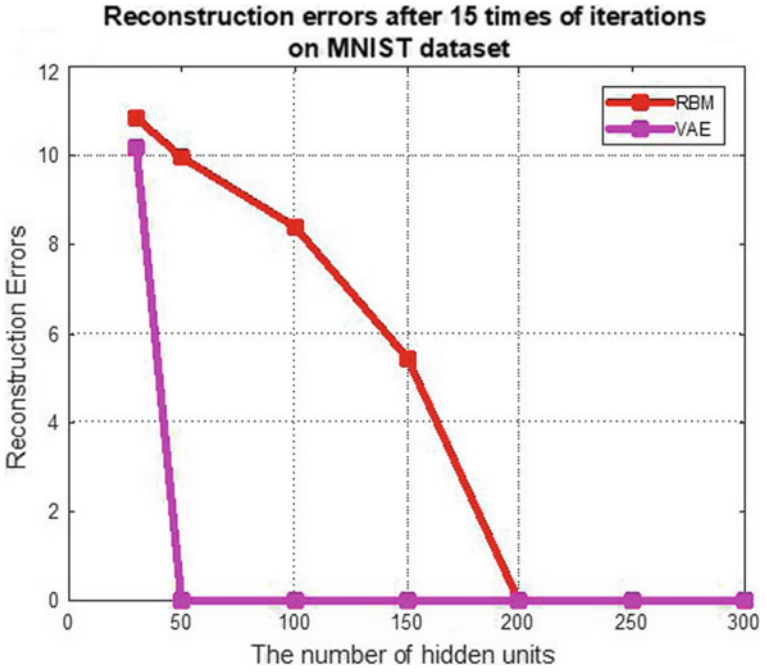


Fig. 4 The reconstruction errors after 2 times of iterations with 20 samples on MNIST

Table 1 The SNR and PSNR of Lena image when using RBM network compression algorithm

Image type	No. of hidden units	SNR (dB)	PSNR (dB)
2D	2	6.2106	60.0957
2D	4	7.3930	61.2781
2D	8	8.4275	62.3126
3D	2	5.5272	58.8745
3D	4	6.5451	59.8924
3D	8	7.1299	60.4772

Table 2 The SNR and PSNR using variational autoencoder

Image type	No. of hidden units	SNR (d)	PSNR (dB)
2D	2	32.4699	134.4858
2D	4	36.3182	138.3341
2D	8	40.4651	142.4810
3D	2	46.6888	148.1670
3D	4	38.8142	140.2923
3D	8	57.5680	159.0461

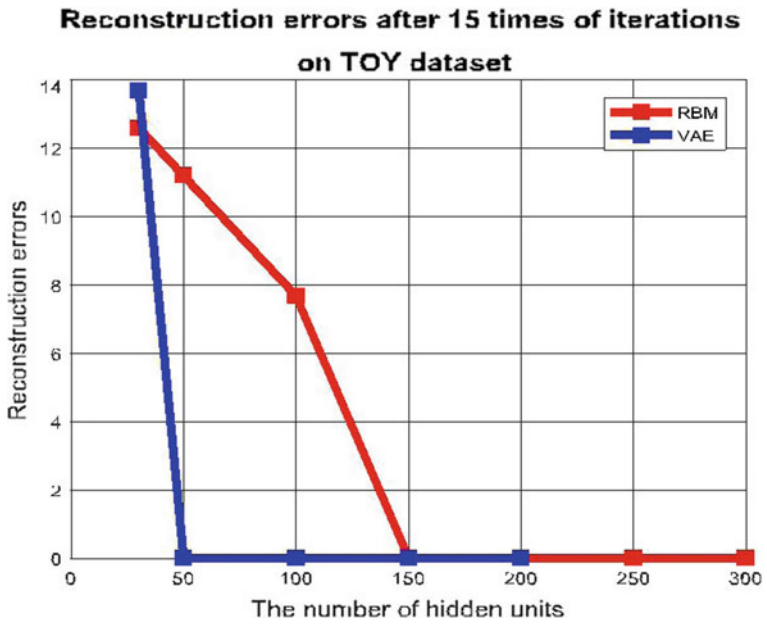


Fig. 5 The reconstruction errors of the two algorithms after 2 times of iterations with 20 samples on the TOY dataset

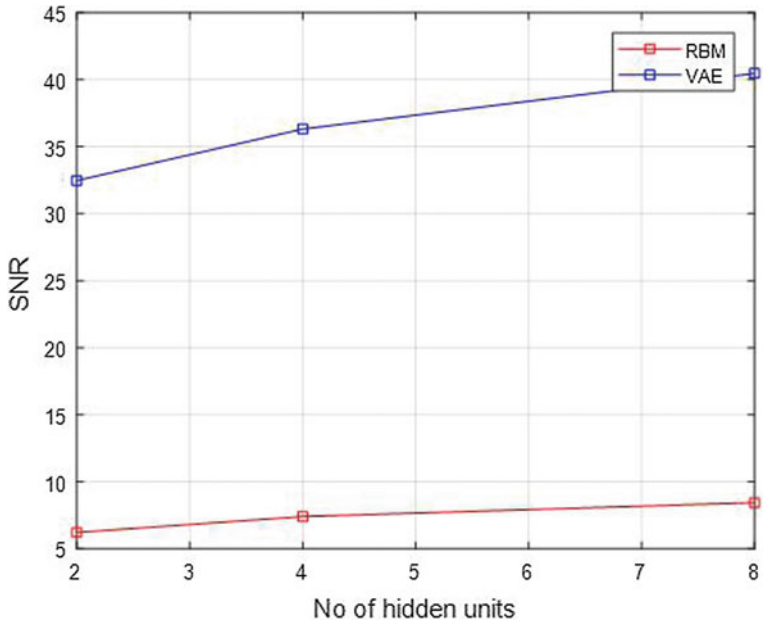


Fig. 6 SNR for gray input image

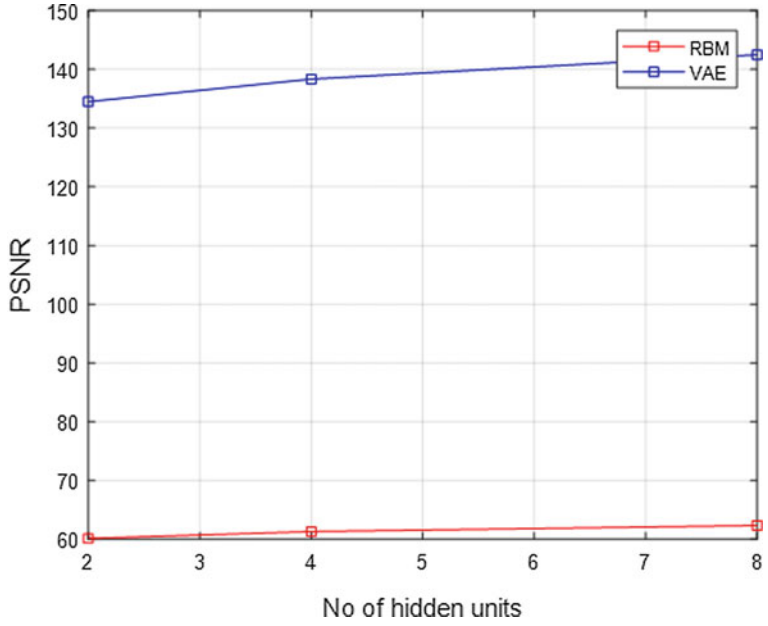


Fig. 7 PSNR for gray input image

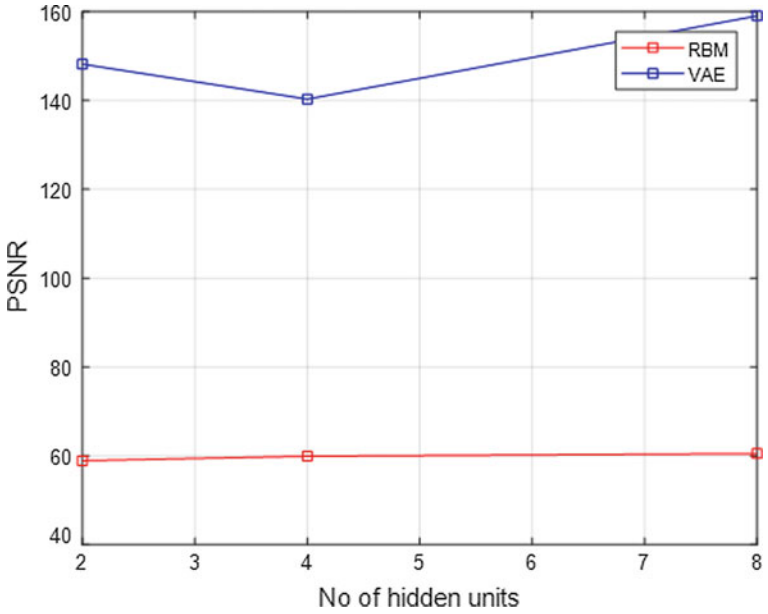


Fig. 8 PSNR for color input image

of the multilayer RBM network compression method is 62.3126 for gray image and 60.4772 for color image.

From the above results, it can be analyzed that the image quality of VAE is superior to that of multilayer RBM networks. When the [3] number of hidden units is 8, the PSNR of the VAE compression method is 142.4810 for gray image and 159.0461 for color image.

4 Conclusion

In WSNs, image transmission and compression plays an important role and many techniques are developed due to rapidly changing developments and applications. Images are compressed and sent over WSN using the variational autoencoder and RBM methods. The performance is evaluated for color and gray images with different hidden layers. An effort is made to find out more transmission and compression methods in WSNs to use them in real-life applications.

References

1. Deepthi SA, Rao ES, Prasad MNG (2018) Image transmission and compression techniques using SPIHT and EZW in WSN. In: 2018 2nd international conference on inventive systems and control (ICISC), Coimbatore, pp 1146–149. https://doi.org/10.1109/icisc.2018.8398984
2. Bengio Y, Courville AC, Vincent P (2012) Unsupervised feature learning and deep learning: a review and new perspectives. Department of Computer Science and Operations Research, University of Montreal, Montreal, Canada
3. Cheng C, Wang S, Chen X, Yang Y (2016) A multilayer improved RBM network based image compression method in wireless sensor networks. Int J Distrib Sens Netw (Hindawi Publishing Corporation)
4. Kim J, Kyung C-M (2010) A lossless embedded compression using significant bit truncation for hd video coding. Circuits Syst Video Technol IEEE Trans 20(6):848–860
5. Vo DT, Solé J, Yin P, Gomila C, Nguyen TQ (2010) Selective data pruning-based compression using high-order edge-directed interpolation. Image Process IEEE Trans 19(2):399–409
6. Dipayan (2017) Deep learning with Hadoop. Packt Publishing
7. Larsen ABL (2016) Autoencoding beyond pixels using a learned similarity metric
8. Frans K (2010) Variational autoencoders explained, vol 9, pp 145–152; Desjardins AC, Bengio Y, Parallel tempering for training of restricted Boltzmann machines. J Mach Learn Res Workshop Conf Proc

A Study on Neutron Star Interior Composition Explorer

C. S. N. Koushik, Shruti Bhargava Choubey and S. P. V. Subba Rao

Abstract It is a NASA Space Science Explorers program under the opportunity mission; it analyzes the electromagnetic, gravitational, and nuclear aspects of pulsars. NICER uses rotation-resolved spectroscopy in thermal and nonthermal emission spectra in the soft X-ray band with good precision and accuracy for detailed analysis of cosmic particle accelerators like in for quarks. It is also be used for X-ray navigation. Its advanced electronics helps it to analyze the data much efficiently (Davis et al. in J Geophys Res Oceans 99(C12):24383–24404) [1].

Keywords NICER · SEXTANT · Pulsar/Neutron star · X-ray navigation

1 Introduction

Neutron star Interior Composition Explorer (NICER) is part of the **Station Explorer for X-ray Timing and Navigation Technology (SEXTANT)** located on International Space Station (ISS) controlled by NASA for research on neutron stars through soft X-ray timing. It uses rotation-resolved thermal and nonthermal spectroscopy of neutron stars in the soft X-ray spectral band. NICER is deployed onto the International Space Station (ISS) [1].

C. S. N. Koushik (✉) · S. B. Choubey · S. P. V. Subba Rao
Sreenidhi Institute of Science and Technology (SNIST), Yamnampet, Ghatkesar, Hyderabad 501301, India
e-mail: koushikcsn@gmail.com

S. B. Choubey
e-mail: shrutibhargava@sreenidhi.edu.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_30

1.1 X-rays

X-rays/Röntgen radiation was named after German scientist Wilhelm Röntgen; these are generally produced when electrons hit any hard or tough metallic surface.

1.1.1 Soft and Hard X-rays

X-rays have photons with high energy of about 5–10 keV with 200–1000 pm wavelength called as hard X-rays which are used in imaging and X-ray crystallography due to their less penetrating capacity. X-rays with less energy and wavelength are called as soft X-rays which easily attenuate and absorbed by the medium in which they propagate [2].

1.1.2 Properties

X-ray is an electromagnetic radiation, traveling at the speed of light, and needs no medium for propagation, with a wavelength of 0.01–10 nm and frequency of 30 petahertz to 30 exahertz (3×10^{16} to 3×10^{19} Hz) having 100 eV to 100 keV of energy. It is used in microscopy for high-resolution images. Hard X-ray photons ionize atoms and interfere with molecular bonds, and hence used in material study by X-ray spectroscopy and crystallography. These can also destroy living cells. Soft X-rays have low energy, causing them to have less penetration capacity and attenuation [3].

1.2 Pulsars

A **pulsar** is a neutron star producing X-rays in all directions due to the rotational energy as it rotates. It has strong magnetic field creating electrical field on star's surface due to its motion, accelerating protons and electrons. X-rays pointing toward the Earth are received as pulses in periodic time intervals. Pulsar retains its angular momentum but with a part of progenitor's radius. The rotation speed decreases over time due to radiation of energy in all directions on the axis on which it rotates; that limit is called death line (Figs. 1 and 2).

2 Review of Literature

NICER was elected as NASA's Science Mission Directorate as in for Mission of Opportunity in 2013, supervised by Goddard Space Flight Center (GSFC), and col-

Fig. 1 Neutron star interior

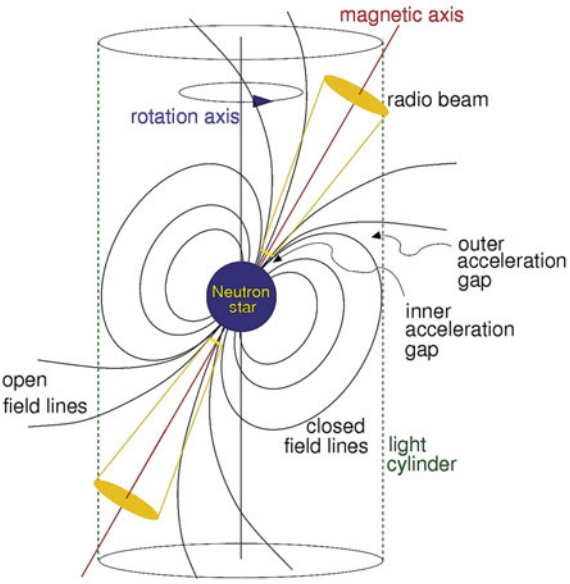
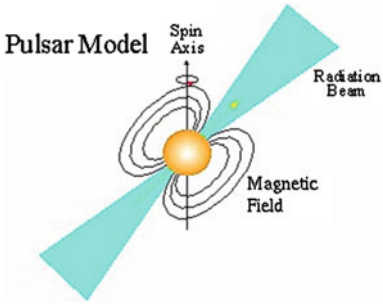


Fig. 2 Pulsar



lects scientific data with advanced navigation technologies at low-cost. In May 2015, a number of tests and design reviews were held but then launched on June 3, 2017. As ISS resupply mission, it was sent to ISS on SpaceX CRS-11 Falcon 9 v1.2 rocket. Since May 2018, NICER discovered about 2000 pulsars. Pulses received are regular like a ticking clock. At high pressure and density, behavior and interactions of quarks are studied with accuracy and precision. NICER provides extension to RXTE; it is 30 times sensitive and has 8 times more energy resolution than RXTE [1].

3 Problem Formulation

The Equation of State (EOS) is used for studying dense matters. It is compared with respect to the pressure in the star where a graph is plotted for mass and radius of the star giving a S-shaped curve (Figs. 3 and 4).

The light curve analysis compares all the proposed EOS models for which 10^{5-6} photons are collected for it. It is done for thermally steady X-ray pulses from the **Millisecond Pulsars (MSPs)** and black holes. Stable in time, it has good signal-to-noise ratio. Far side spot visibility depends on M/R ratio and through gravitational light bending [1].

4 Methodology

4.1 Construction

The four main parts are **Electronics box**, **Flight software**, **Ground Test bed**, **and Ground bed**. It is installed robotically by DEXTRE to exact location. It is done by Space Station Remote Manipulator System (SSRMS) and Special Purpose Dexterous Manipulator (SPDM) altogether. NICER is placed at desired location on EXPRESS Logistics Carrier-2 (ELC-2) which is fixed by Flight Releasable Attachment (FRAM) and Gimbals on ISS to save itself from space junk (Figs. 5, 6, and 7).

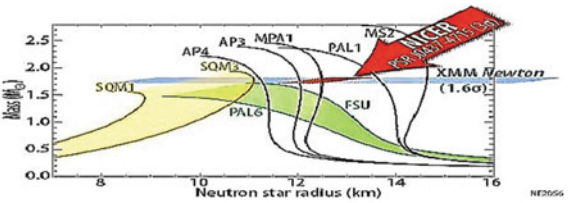


Fig. 3 R/M ratio analysis

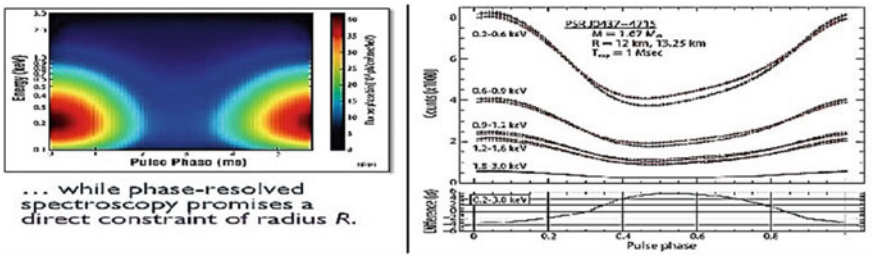


Fig. 4 Light curve analysis

4.1.1 Electronics Box

The 56 X-ray modules have X-ray Concentrators (XRC) and Silicon Drift Detector (SDD) pairs used for X-ray detection. Each XRC intensifies the signal by concentrators and converges the beam by focusing mirrors. The SDD detects individual photons that it receives and records spectral energy resolution and their Time of Arrival (TOA) of pulses. X-ray photons with high immunity to noise are detected having good signal-to-noise power ratio. The processor/controller used is LEON3FT-RTAX processor like any space application for good tolerance of radiation [1] (Figs. 8 and 9).

The photon detectors are Silicon Drift Detectors (SDD) with good quantum efficiency and spectral and time resolution in the photon energy range. It has high

Fig. 5 Deployment status

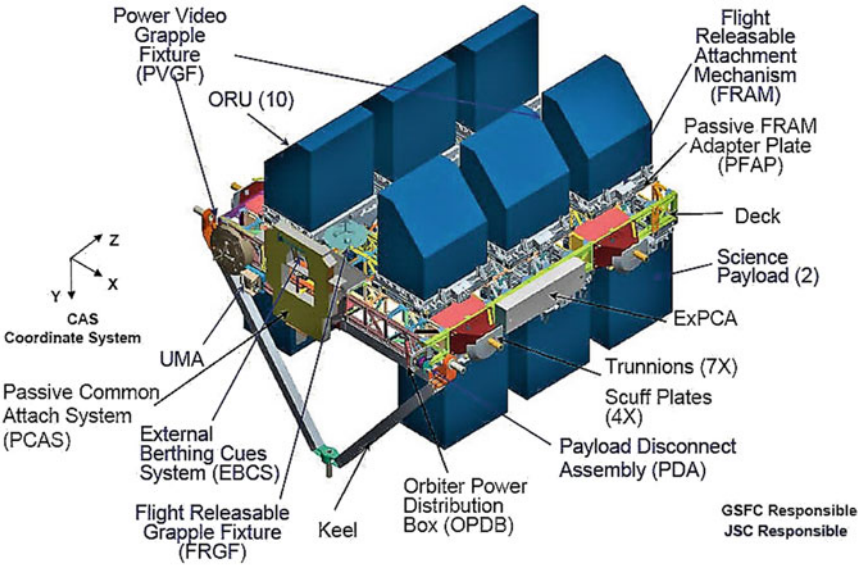


Fig. 6 DEXTRE

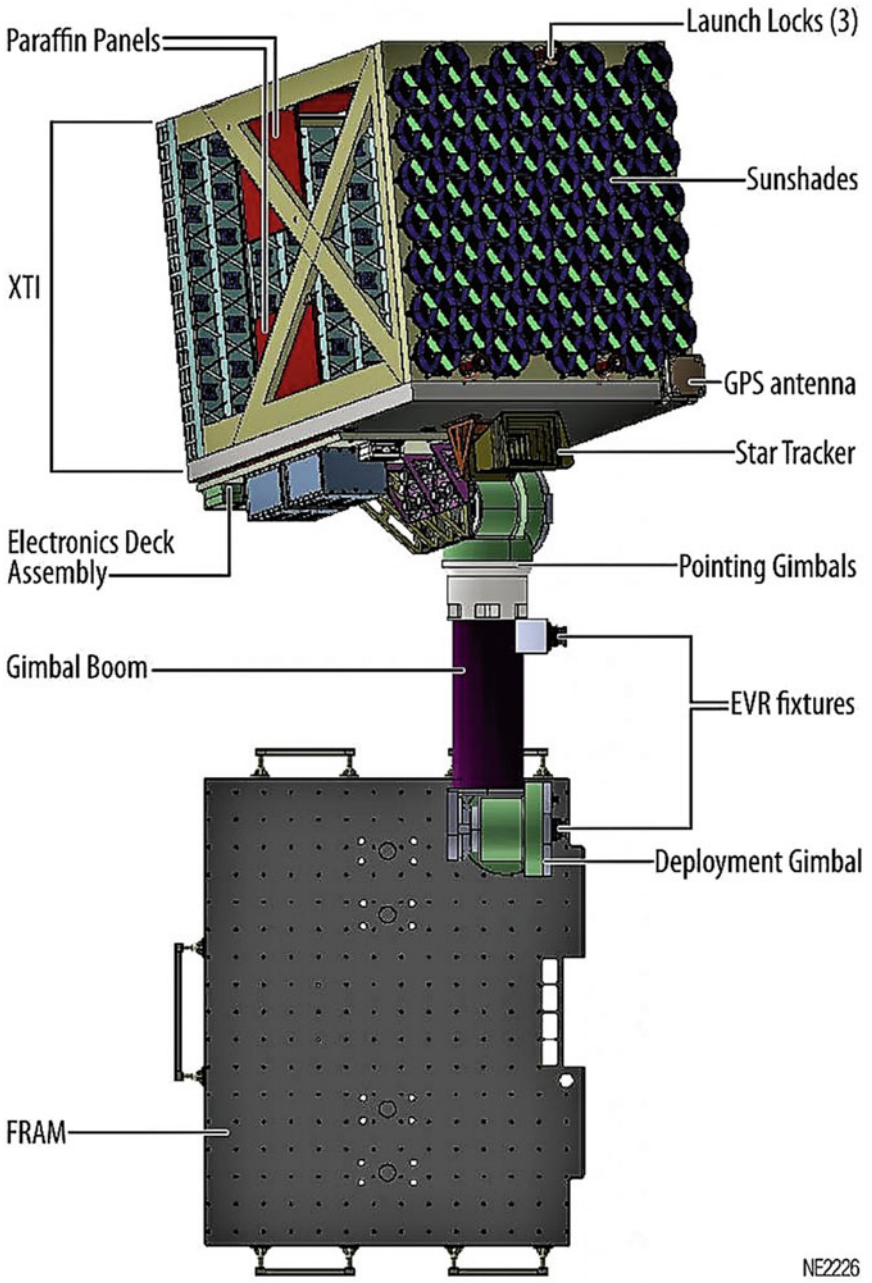

Paraffin Panels
Launch Locks (3)
Sunshades
XTI
GPS antenna
Star Tracker
Electronics Deck Assembly
Pointing Gimbals
Gimbal Boom
EVR fixtures
Deployment Gimbal
FRAM
NE2226

Fig. 7 Parts

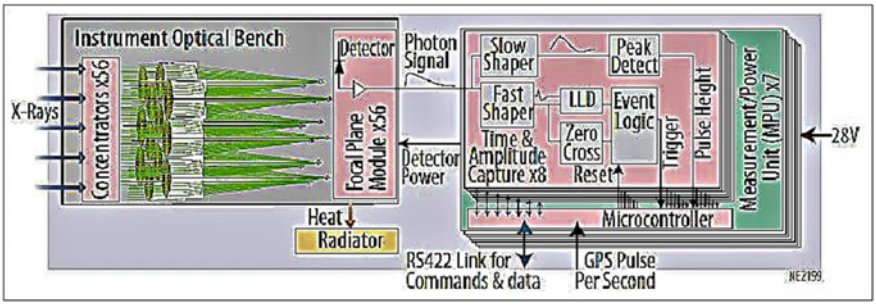


Fig. 8 Electronics box

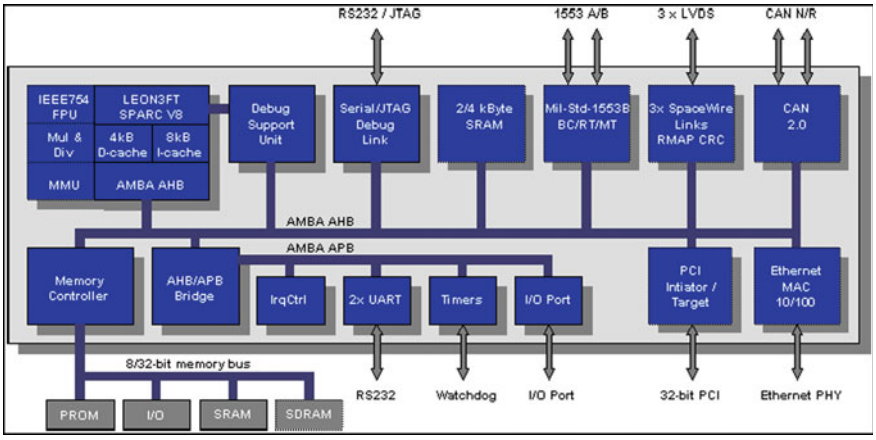


Fig. 9 Microcontroller

precision time tagging for TOA in X-ray spectrum synced to NICER's GPS receiver system [4] (Figs. 10 and 11).

4.1.2 Flight Software and Algorithms (XFSW)

It process photon events related to photon energy which are analyzed with Non-homogeneous Poisson process (NHPP) synced NICER GPS time. It has two main parts: Photon processing algorithms, and GEONS flight software (Goddard Enhanced Onboard Navigation System) which has packages in MATLAB and C language. After enough photon events, it checks for pulse phase and Doppler measurements (Fig. 12).

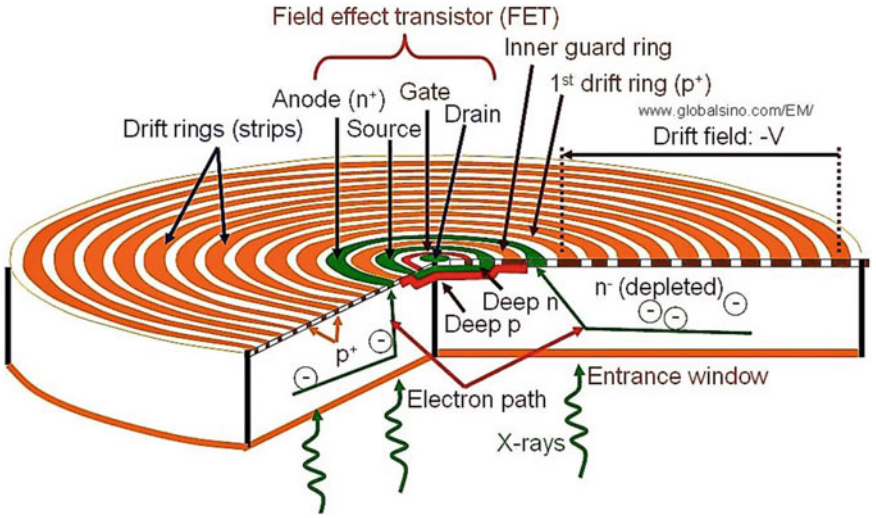


Fig. 10 SDD

Fig. 11 GPS antenna

4.1.3 Ground Test Bed

GSFC X-ray navigation laboratory test bed provides guidance, navigation, etc., which has devices in loop and directives for testing the NICER predictions. Software and hardware give good picturization and navigation outputs in three simulation stages which are as follows:

Stage 0—Phase and Doppler readings are picturized with navigation filter help.
Stage 1—Software achieves the allegiance by virtualizing the readings obtained.
Stage 2—Allegiance is increased more with photon process in Hardware-in-the-Loop (HWIL) process, for greater picturization [5].

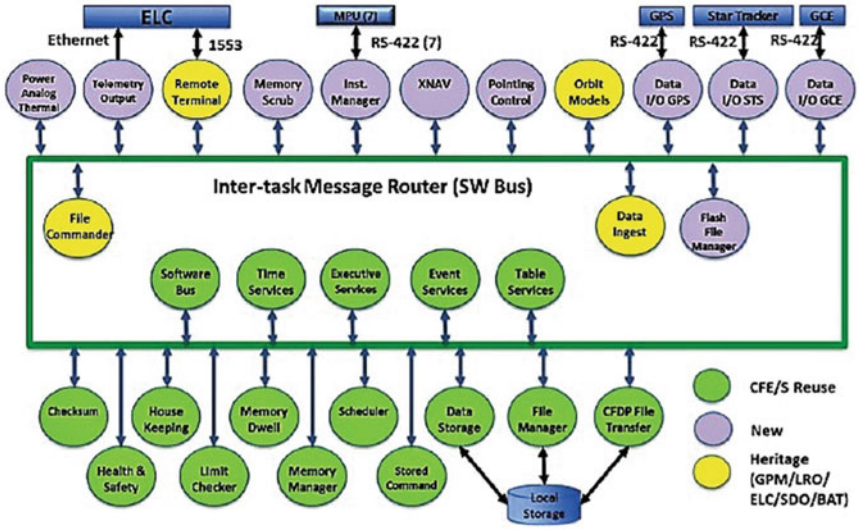


Fig. 12 Flight software

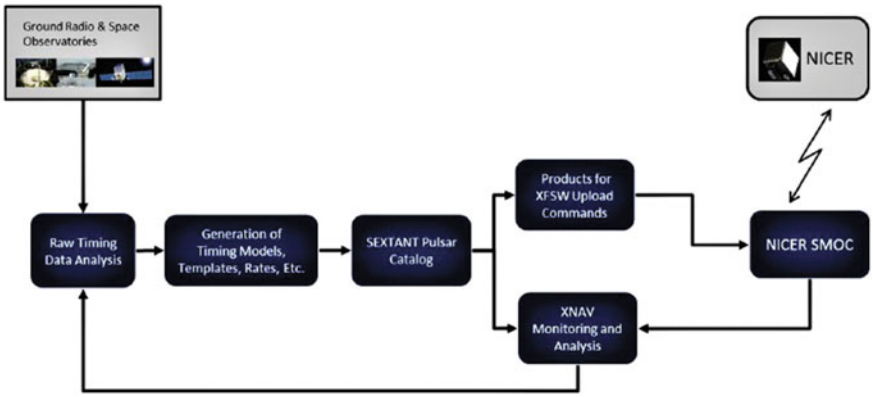


Fig. 13 Ground bed

4.1.4 Ground Bed

It has two main parts: monitoring and trending element in Science Mission Operations Center, for comparison of the pulses received with already stored data. It has pulse timing models—parameterized, and pulse phase predictions—polynomial, used for better comparison and analysis. It is like correlation of the signals being done (Fig. 13).

5 Working

The NICER deployed to the exact position with the help of FRAM and DEXTRE to track the pulsar. The X-ray pulses detected are received from pulsars radiating in direction toward Earth. These are intensified by the concentrators and focused by focusing mirrors to reduce the beam width. For a TOA, the photons are detected at the silicon drift detectors (SDD) which are highly sensitive for every photon producing voltage, which is converted latter into digital format where all the bits are processed and stored in the microcontroller. The TOA is noted with help of GPS antenna. The flight software checks for Doppler measurements and pulse phases filtering out Shapiro delays. In ground test bed, both hardware and software, the data is modeled or simulated for proper visualization. The ground bed correlates the data that it received with the data in pulsar records for better precision and accuracy [4, 5].

6 Conclusion and Future Scope

It is used in the detection of the pulsars, neutron stars, and black holes. It is useful in navigation and study of stars, how space and time are related, and how did universe form from Big Bang and for that it is essential to improvise it further to get data more accurately. It is the updation of both flight and ground software. Using detectors and other hardware for XNAV with great accuracy and sensitivity should be made readily available on future space crafts in space.

References

1. Davis JC, Tapley BD, Ries GW, Eanes RJ, Schutz BE, Shum CK, Watkins MM, Marshall JA, Nerem RS, Putney BH (1994) Precision orbit determination for TOPEX/POSEIDON. J Geophys Res Oceans 99(C12):24383–24404
2. Bamford LMB, Winternitz WA, Heckler GW (2009) A GPS receiver for high-altitude satellite navigation. IEEE J Sel Topics Signal Process 3(4):541–556
3. Real-Time GEO Orbit Determination using TOPSTAR 3000GPS Receiver. Navig J Inst Navig 48(3):169–180
4. Ray PS, Wood KS, Phlips BF (2006) Spacecraft navigation using X-ray pulsars. 2006NRL Rev 95–102
5. Interplanetary navigation using pulsating radio sources. NASA Technical Reports, N74-34150, 1974, pp 1–12

Object Detection Using Higher Quality Optimization Techniques in Video Encoders

Abhishek Choubey, Shruti Bhargava Choubey and Khushboo Pachori

Abstract High Efficient Video Coding (HEVC) is the most recent video coding standard developed for efficient compression of high definition videos. It adapts a quad tree based coding tree structure in which each frame is represented as coding tree blocks (CTB). It uses only half of the bandwidth when compared with H.264/AVC. Though it achieves high coding efficiency, it suffers from problems such as high computational complexity and large computation time. In this paper we propose an edge detection based encoding algorithm using sobel operator for coding unit depth decision. Thus the computational complexity and the computational time are reduced considerably. The experimental results showed that using this method, when compared to HEVC standard the encoding time reduces by 30% without significant loss of coding efficiency and PSNR.

Keywords HEVC · Computation · Compression · Sobel · Coding efficiency

1 Introduction

As the demand for higher visual quality increases the H.264/HEVC encoding standard has become insufficient. To overcome the limitations posed by H.264, a new encoding standard H.265/HEVC has been proposed by ITU-T VCEG and ISO/IEC international organizations to meet the requirements of ultrahigh definition videos. HEVC mainly serves high definition videos [1] H.265 uses quad tree based coding block structure. In this each frame is divided into coding tree blocks (CTB). The different quad tree blocks include Coding Unit (CU), Transform Unit (TU), Prediction

A. Choubey · S. B. Choubey (✉) · K. Pachori
Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology (SNIST), Yamnampet, Ghatkesar, Hyderabad, India
e-mail: shrutibhargava@sreenidhi.edu.in

A. Choubey · S. B. Choubey · K. Pachori
Electronics and Communication Engineering, Gurunanak Institute of Technology, Hyderabd, Telengana, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_31

Unit (PU). The CU block in HEVC is the basic building block similar to macro block in H.264. In HEVC this CU can further split into PU and TU. PU helps to get a good prediction of image blocks. TU is a partition of prediction residue in order to obtain good DCT/DST performance. It explains the advantages of HEVC [2].

Compared with its predecessors the aspects improved in HEVC are (1) the size of Coding Tree Unit can be varied from 8×8 to 64×64 when compared to the fixed size of macro block in H.264/AVC which increases the efficiency in compression of ultrahigh definition videos; (2) flexible coding structure which includes CU (Coding Unit), PU (Prediction Unit) and TU (Transform Unit). Though it achieves high coding efficiency compared to its predecessors it requires high encoding time. Hence the problem of fast mode depth decision has been a matter of research.

In research of HEVC coding depth can be efficiently determined by using motion features based on phase correlation [3] and a Novel deep learning based method for improving efficiency [4]. Bayesian decision algorithm can be used to reduce computational complexity [5]. A model of Scalable-Video-Coding (SVC) uses motion searches to identify the best prediction mode with high coding efficiency.

In the next part of this paper we introduce perceptual video quality measurement. The need to use YUV color space in perceptual video coding is discussed in the Third part. HEVC inter mode coding is introduced in the next part. In the Fifth part a fast coding unit depth decision algorithm based on sobel edge detection is introduced. Experimental results and analysis is presented in final part of this paper.

2 Perceptual Video Quality

Video quality measurement can be divided into two categories: objective video quality metrics and perception based video quality metrics. Objective video quality metrics use standardized metrics such as Mean-Square-Error (MSE) and Peak-Signal-To-Noise-Ratio (PSNR) to measure the quality of the video. On the other hand perceptual metrics take human perception into account in order to calculate the quality of the video.

PSNR metric is commonly used for ease of identification of quality of video but it does not reflect the receiver's perception. For the best quality PSNR must be high. According to metrics results Fig. 1 has best quality. But we can observe that Fig. 2 has better perceptual quality when compared to Fig. 1.

Hence in this paper we obtain the perceptually important regions in each frame of the video using sobel edge detector and use this data to reduce the complexity, number of steps required to decide the Coding Unit (CU) depth such that the time required for encoding can be reduced without significant loss in PSNR.

Fig. 1 Frame with PSNR (29.1) dB

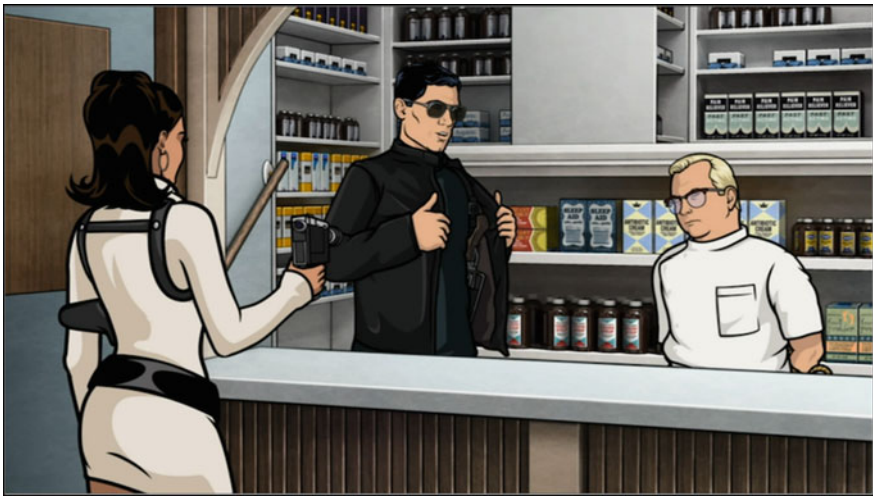


Fig. 2 Frame with PSNR 29.3 (dB)

3 HEVC Inter-Mode Coding Process

In the HEVC HM test model, the default setting supports four different CU sizes i.e. 8×8, 16×16, 32×32, 64×64. The corresponding Coding unit depths are 3, 2, 1, 0 respectively. The Prediction Unit size varies from 4×4 to 64×64, symmetric mode similar with H.264, and new asymmetric mode are provided. In HEVC inter prediction process requires most number of computations compared to other steps. It is also the most memory consuming part in the encoding process.

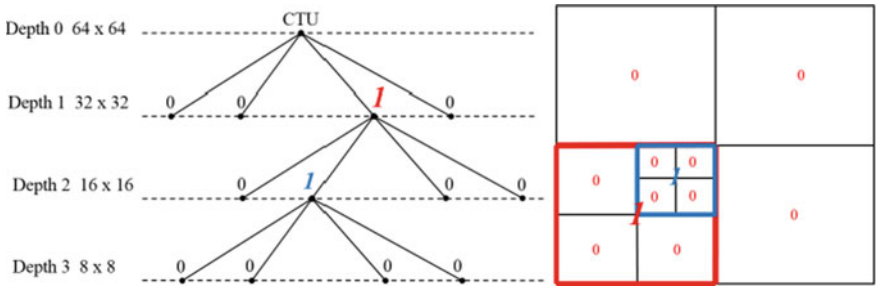


Fig. 3 Coding unit recursive partitioning process

Table 1 Peak-Signal-To-Noise-Ratio (PSNR), Bit rate, Encoding time for a standard video sequence of resolution 352 × 288

Quantization parameter	PSNR (existing) (dB)	PSNR (proposed) (dB)	Bit-rate (existing)	Bit-rate (proposed)	Encoding-time (existing) (s)	Encoding-time (proposed) (s)
30	32.7245	32.2985	494.5896	535.9536	3727.521	1691.740
32	31.6152	31.2053	380.2560	414.8712	3583.194	1498.344
34	30.6457	30.2273	302.5008	330.2640	3284.964	1378.439
36	20.1291	29.1789	238.6920	257.8704	2658.330	1259.486

In the process of HEVC prediction, each incoming frame is divided into LCUs. These LCUs are again partitioned recursively until it reaches the maximum coding depth as shown in Fig. 3.

Now the RD-cost (Rate Distortion-cost) will be calculated for each Coding unit depth and finally the CU size will be determined on the basis of RD-cost. The Coding Unit size will be the size for which the RD-cost is minimum. Even without considering the asymmetric mode a large number of comparisons and calculations are necessary to arrive at the final CU depth. Hence though HEVC increases the encoding efficiency it increase the computational complexity and time. In this paper we propose a texture based algorithm for fast coding unit depth decision in order to reduce the encoding time without significant loss in PSNR.

Table 1 shows Peak-Signal-To-Noise-Ratio (PSNR), Bit rate, Encoding time for a standard video sequence of resolution 352 × 288.

4 Proposed Edge Detection Algorithm Based on Sobel Operator

In inter-mode coding process, for regions of image having smooth texture in video sequence, after calculating and comparing Rate Distortion Cost, these regions select

larger CUs to code; while regions of image have complex texture details in video after calculating RD-cost, these regions select smaller CUs. Based on this we can infer that determination of final CU size is closely related to textural features of each incoming frame in video.

4.1 Determination of Image Edge

The best operator that adapts to changing trend of an image is Image gradient. Numerous differential operators are used to extract image borders using edge determination. The frequently used operators for edge determination are Canny, Roberts, Prewitt, Sobel, etc. In order to obtain good results we choose sobel operator which weighs the effect of pixel positions. It has smooth effect on noise and provides more precise information of edge direction. In this paper, sobel edge operator is used to detect texture features and to finally determine the coding unit depth.

Figure 4 shows two sobel masks, one in x direction and the other in y-direction.

Applying the above sobel masks for a supposed image $f(x, y)$ we obtain two gradients, one in the transversal (horizontal) direction and the other in the longitudinal (vertical) direction. G_x, G_y are the gradients in horizontal and vertical directions respectively.

$$\begin{aligned} G_x = \{&-f(x-1, y-1) + f(x-1, y+1) - 2f(x, y-1) \\ &+2f(x, y+1) - f(x+1, y-1) + f(x+1, y+1)\} \end{aligned} \tag{1}$$

$$\begin{aligned} G_y = \{&f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1) \\ &-f(x+1, y-1) - 2f(x+1, y) - f(x+1, y+1)\} \end{aligned} \tag{2}$$

The sobel value of each pixel can be obtained from the above section.

$$G_{\text{sum}} = \sum_{x=1}^{x<\text{width}} \sum_{y=1}^{y<\text{height}} \sqrt{(G_x)^2 + \left(G_y\right)^2} \tag{3}$$

Here, width refers to the width of current CU size, height refers to the height of current CU size.

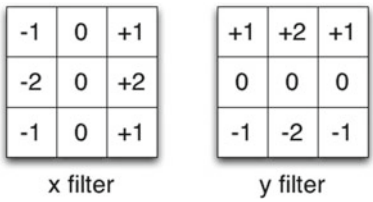


Fig. 4 Sobel masks in x and y direction

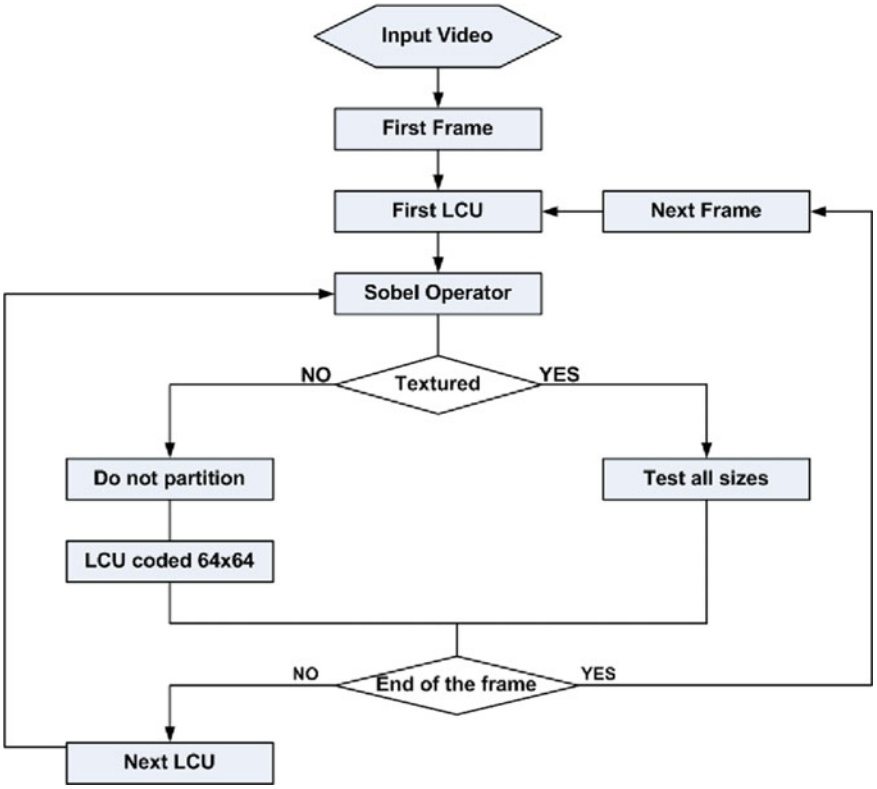


Fig. 5 Flowchart for LCU partitioning

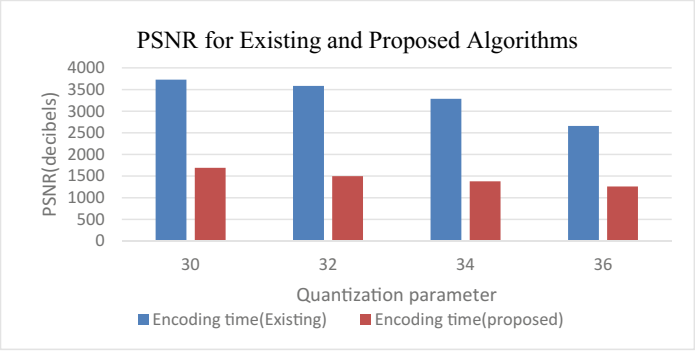
4.2 Proposed Algorithm

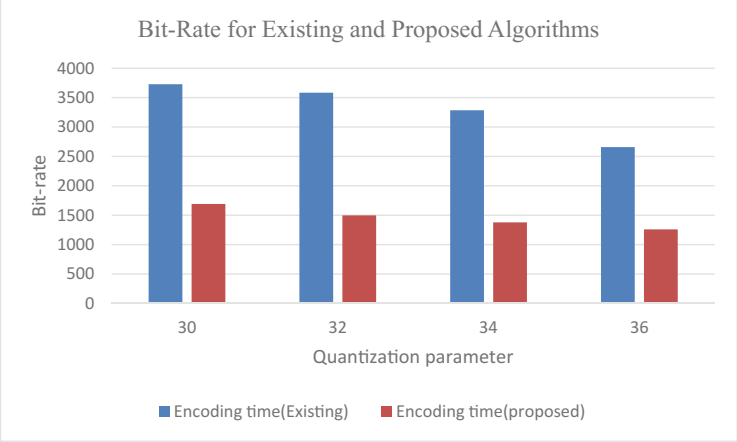
The sobel value obtained for each LCU determines the nature of the texture present. If the texture is smooth and continuous the LCU is not further partitioned. If the texture is complex or discontinuous in nature the LCU is further partitioned based on coding unit depth given in the configuration file. The flowchart is shown in Fig. 5.

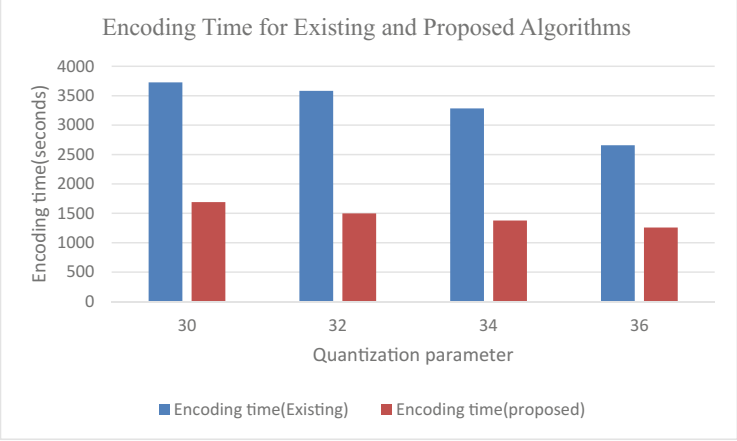
5 Results and Conclusion

Table 1 shows the PSNR, Bit-rate, encoding time for existing and proposed algorithms of a standard yuv sequence of resolution 352 $\times$ 288.

The following graphs compare the PSNR, Bit-rate, encoding time for the existing and proposed algorithms







From the obtained tabular and graphical results, we infer that usage of the proposed algorithm for the video encoding nearly leaves the PSNR unchanged with an

approximate 30% decrease in encoding time and the bit-rate. The video sequences which are encoded using the proposed algorithm are perceptually better when compared to that of the existing one. Hence, the proposed work allows the video encoding to be done optimizes the perceptual quality of the video, reduces bit-rate to about 30% and takes 30% less time for encoding.

References

1. Bross B, Han WJ, Ohm JR, Sullivan GJ, Wang YK, Wieg T (2013) High efficiency video coding (HEVC) text specification draft 10. Document JCTVC-L1003, JCT-VC [Online]. Available http://phenix.int-evry.fr/jct/index.php
2. Shen X (2013) Research on low complexity encoding optimization algorithms of HEVC. Ph.D. Communication and Information System, Zhejiang University, Zhejiang
3. https://www.researchgate.net/profile/Manoranjan_Paul2/publication/265169056_Fast_Intermode_Selection_for_HEVC_Video_Coding_Using_Phase_Correlation/links/54040ca80cf2bba34c1c316a/Fast-Intermode-Selection-for-HEVC-Video-Coding-Using-Phase-Correlation.pdfR. Nicole Title of paper with only first word capitalized. J Name Stand Abbrev (in press)
4. https://www.researchgate.net/publication/320195447_Reducing_Complexity_of_HEVC_A_Deep_Learning_Approach
5. Shen X, Yu L, Chen J (2012) Fast coding unit size selection for HEVC based on Bayesian decision rule. In presented at Picture Coding Symposium (PCS), Krakow, pp 453–456

Part IV
VLSI

Mesochronous Operation Interface to Multicore Processor

K. Sukanya and G. Laxminarayana

Abstract Synchronize processing is a major challenge in multicore processor units. In the designing of multicore processor, as the processing unit work simultaneously, the synchronization of data, instruction and the delay encountered is to be minimized. In the effort for providing synchronous operation, delay synchronization is the main criterion. The delay encountered in bus switching, data fetching, instruction/data allocation, and resource allocation delay is a major concern. In the minimization of delay constraint in resource allocation, the mesochronous operation is suggested. In this work, an interface to mesochronous operation for multicore processor operation is proposed. The presented approach defines the delay application in multiprocessor application with delay synchronization.

Keywords Mesochronous operation · Multicore processor interface · Delay synchronization

1 Introduction

The ever-demanding service compatibility, service rate, and better performance have led to the development of new technologies, new communication mode, and new processing architectures. The evolution in VLSI technologies has resulted in the integration of large distributed processing unit under a signal on chip modelling, where large distributed units are embedded onto a single chip level. This has given the advantage of higher density architecture integrated over a small platform, result-

K. Sukanya (✉)
Department of E.C.E., TKR College of Engineering and Technology, Ranga Reddy 500097, Telangana, India
e-mail: sukanya.addagatla@gmail.com

G. Laxminarayana
Department of E.C.E., Anurag College of Engineering, Ranga Reddy 501301, Telangana, India
e-mail: gln9855@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_32

ing in low area coverage and smaller remote devices. However, with the increase in the integration level, the overhead in data access and the demanded resource for processing also increases. The resource allocation and its utilization leads to an arbitration of data/instruction and the used registers are to be synchronously allocated for the data/instruction operation. In the allocation process, the delay encountered leads to the misalignment of the data/instruction set, which results in fault result generation. It is hence a critical operation in processor design to achieve a higher degree of data/instruction synchronization. In the provisioning of synchronization, among different developed approach, mesochronous operations are developed as an optimal answer. In the design of mesochronous operation for processor operation, in [1] simulation with a sample solution for various processing units at mesoscale has been suggested. The computational tasks in this processor need to determine time breakdown, the location of Lagrangian particles, and the interpolation of those particles. This is observed to be an expensive operation, and the appropriate grid for optimization was needed. In [2] for a sequential code, synchronous programming is defined in the phase of the concurrent programme. Designed to work on single-core processors, an approach to create a multi-thread code using a polychronous data flow model has been suggested [3]. Focus on two different parallel desktop architecture: Multicore processors and GPU Accelerators. The GPU implementation is based on the measurement of a unified device structure, and a multicore implementation was identified by general multiprocessing. A new scheduling algorithm for multicore architecture for pipeline operation is outlined in [4, 5]. This algorithm generates pipeline schedules by forming and solving problems of MILP (mixed integer linear programming). The suggested approach produces schedules that use 71% of the smaller buffers needed to communicate in kernels when compared to traditional approaches. Six-way VLIW digital signal processor (DSP) architecture has been designed and put forward with clustered registry files in [6]. Distinctive length instruction architecture is designed for mathematics and dynamic instruction execution. The single-level memory system architecture is proposed using processor instructions, data caches, instructions, and data on-chip RAM. In [7] a synchronous communications protocol has been proposed in two high-resolution-, low latency conventions for switching to the asynchronous operation is suggested. A network on a chip (NoC) consisting of two hardware components designed to use on a multiprocessor system that communicates on the chip in the entire asynchronous network is suggested. Specific architecture has been generated, which helps to determine the difference between operations and robustness depending on the system. In [8] a runtime environment that supported classes of programming modes and their writing of synchronization in a network processor is presented. This approach implements a single-process multi-thread using a set of specialized inter-process communication (IPC) that exploits the architecture and a set of clusters and I/O subsystem is suggested. This approach provides timeline support for various target programming models in a multiprocess environment. An embedded multiprocessor designing the methodology for enabling Aggressive Programming Properties in System on MPSoC is defined in [9]. It is achieved by integrating the designer-guided code parallelization method using a virtual, generic, and scalable MPSoC unit. To solve data dependen-

cies during task occupation, MPSoC is processed for a communication scheme based on a processor-independent task. In this eight processors implement an application that changes dynamically when running on a continuous, parallel execution. A multi-thread CPU coding is suggested in [10] to perform better performance in parallel processing, along with the number of events used to test different applications. The study also reveals that the use of GPs decreases the number of programmes that help performance and improve the calculation time. An ASAP architecture is defined in [11] which is used as a simple processor architecture with minimal memory to boost energy efficiency. The clocking system and the nearest neighbour communication enhance the scalability and provide opportunities to reduce device fluctuations, global wire controls, and processor failures. In many new design strategies that represent existing design trends to cope with emerging issues in [12] a power consumption, memory bandwidth/ latency, transistor variability, and power/thermal management for multiprocessor design logic for testing is developed. A Total Store Order test is conducted into the memory model of an Intel x86 and Sun SPARC processors for the safety in information flow. Two approaches on the flow of information are provided for these processor architectures. In [14] a synchronous clock of multiple core processors required for synchronous clocking for reliable implementation of parallel computing and complex operation is proposed. Synchronous clocking enables large-scale systems with interconnected stage-lock loops transition systems delays. A brief description of the complex operation, including filter kernels, and signal processing were suggested. The transmission delays sustainable states and PLLs for synchronization were validated. Two regulation improvements based on the criteria for simplifying volume increases with multiple task-to-processor mapping is outlined in [15]. The emphasis is made on dividing the transceiver systems and applying the application specific components to implement an advanced multiprocessor system-on-chip information and control architecture, which can meet the needs of future transceiver operation. A novelty in the Gallows Architect of Platform has been suggested in [16] as a general purpose multiprocessor for the use of hard real-time systems. A clock distribution and synchronization method, suggesting a solution to large amounts of jitter issue is addressed in [17]. A redesigned Data Path processor (RDP) is proposed for the implementation for a number of steps for synchronization. Different developed methodologies in the provision of synchronization were focused with the synchronization in provisioning data and instruction to the processing unit in parallel. However, the delay encountered in the resource allocation and instruction/data transition is to be incorporated to make the multicore operation more efficient. In this paper, the application of mesochronous operation on the multiprocessor SoC design is focused. To present the stated objective, the rest of the paper is outlined in six sections. Section 2 defines the approach of mesochronous operation in processor application. Section 3 outline the integration of mesochronous operation on the application of multicore processor on a SoC Sect. 4 illustrates the simulation result of the proposed approach. Section 5 presents the conclusion derived for the developed work.

2 Mesochronous Operation

In a multicore processor design, clock synchronization is very much needed to perform multiple pipelined operations simultaneously. Wherein, multiple threads are been created for each of the instruction set, a flow of data and instructions are observed inter- and intra-mode in these processors. Wherein, the delay observed in the allocation of these instructions and data were clearly explored in VLSI design. To overcome the delay concern and to achieve a higher level of synchronization, mesochronous operation was suggested. Mesochronous operation, compute the needed delay to be allocated per instruction/delay flow in a register operation, and allocate the delay factor to the master clock pulse to develop a synchronization process. The allocation process of delay factor to a clock allocation is illustrated in Fig. 1.

The delay introduced in the register allocation is compounded with the master clock pulse to derive a new clock called the synch clock which forwards the data to the processing unit at different clock pulses based on the delay derived. This allocation helps in the accurate delivery of data and instruction to the processing unit synchronously. Wherein mesochronous operation in a single core is processor is accountable due to a single line instruction. However, in the case of multicore processor, the designs get dual complex due to data flow synchronization and processor operation synchronization. The approach of integrating a mesochronous operation in a multicore processor is outlined in the next section.

3 Multicore Processor Interface

In the interface of synchronization operation in a multicore processor, delay factors are added as a factor of processor switching and data transition delay in the two cores. The buffer logic is given a delay to hold data from one processor to other for a synchronized operation. In the design of a multicore processor, the delay is computed at per stage basis. In the pipeline operation, the aggregated delay is measured as the time taken per stage of operation. The pipeline stages in a multicore processor are the fetch, decode, execute, memory write, write back stage. It is observed that most of the delay is encountered at the memory interface operation. A generic mode of dual-core architecture is as shown in Fig. 2.

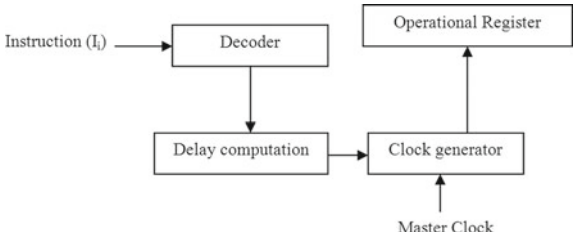


Fig. 1 Clock delay allocation for synchronization in Mesochronous operation

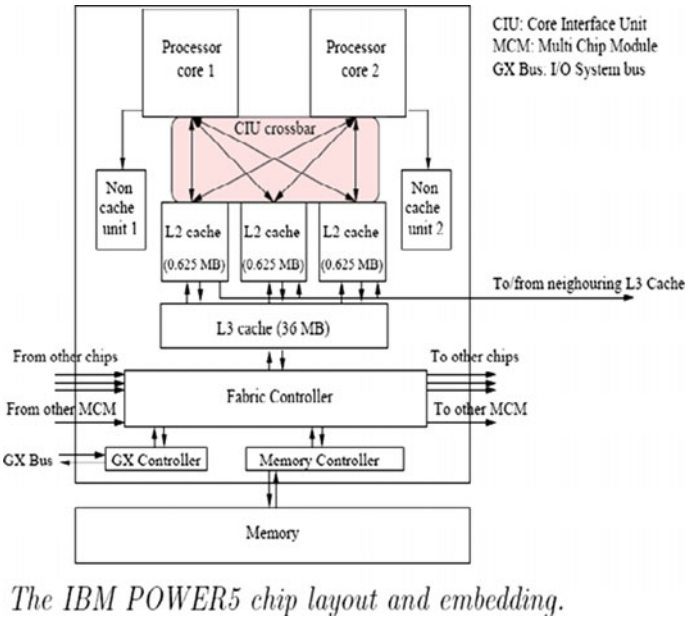


Fig. 2 A two-core processor with cache register interface

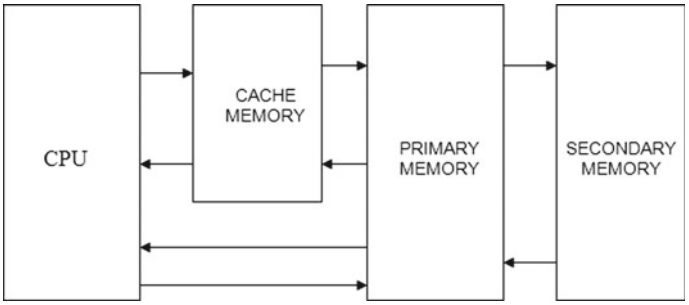


Fig. 3 Data/instruction interface through temporary cache register

The operation of two processors is carried out using a temporary bank of cache register to avoid direct memory accessing delay. The processors could interface and process at a faster rate when used with a temporary register of cache memory. The basic interface of cache register is presented in Fig. 3.

In the allocation of the delay factor in this processor design is the aggregated time for a n-stage operation. The total time period (T_p) is given as

$$T_\mathrm{p} = \sum_{i=1}^{n} t_i \tag{1}$$

where t_i is the time of process for each stage.

Here the clock period is derived as the aggregated delay of the maximum stage period, given as

$$T_{\mathrm{d}} = \max(T_i) \tag{2}$$

The processing clock is then given as

$$T_{\mathrm{pr}} = T_{\mathrm{mstr}} + T_{\mathrm{d}} \tag{3}$$

where, T_{pr} is the processing clock period, T_{mstr} is the master clock period and T_{d} is the time period derive by (2).

Each of the register in the processor unit is then allocated with a register delay (τ_i) for each instruction as derived from the LUT element, given as a predefined computed entry [18]. The offered synchronizing clock to the two processor units is then given as the sum of the register delay with the processing clock (T_{pi}) given as

$$T_{\mathrm{pi}} = (T_{\mathrm{pr}} + \tau_i) \tag{4}$$

The aggregative synchronizing delay (T_0) over a n-core processor is then given by

$$T_0 = T_{\mathrm{pi}} + T_{\mathrm{sw}} \tag{5}$$

Here, T_{sw} is the switching delay for the n-processors. The aggregative delay to each clock signal is given as

$$T_0 = T_{\mathrm{mstr}} + T_{\mathrm{d}} + \tau_i + T_{\mathrm{sw}} \tag{6}$$

Considering the stage delay (T_{d}), switching delay (T_{sw}), constant (K), the allocating delay is defined as

$$T_0 = K * \tau_i \tag{7}$$

where

$$K = T_{\mathrm{mstr}} + T_{\mathrm{d}} + T_{\mathrm{sw}} \tag{8}$$

From Eq. (7), it is observed that for a multicore processor the delay computed per instruction control the delay for the multicore processing, hence the computation overhead of this delay can bring the clock processing overhead delay minimal, hence improving the operational speed. This delay is minimized with a LUT-based approach as outlined in [18].

4 Simulation Result

This proposed solution illustrated a significant approach in the allocation of clock delay in multicore processor design. When the clock delay is to be allocated, the LUT transition reduces the clock delay overhead. The realization of this proposed logic is carried on Aldec's Active HDL 6.3, targeting to Xilinx FPGA Virtex device (xcv300-bg432-6). The operational accuracy of the proposed system is validated by the timing result observed for the developed system.

The instruction decode operation in the processor unit is presented in Fig. 4. The instruction fetched from the instruction cache is read onto 'inst' signal, which is decoded for the opcode and data address. The decoded instruction is given to the processing unit for performing arithmetic and logical operation, and a mapping of the decoded instruction is made to extract the required delay for time synchronization.

The controller unit developed for the proposed unit is presented in Fig. 5. The approach defines the control signal for fetch, decode, execute, and storage operation. The transition of the state for a pipeline operation. The controller controls the read/write operation for the instruction/data in the processor interface.

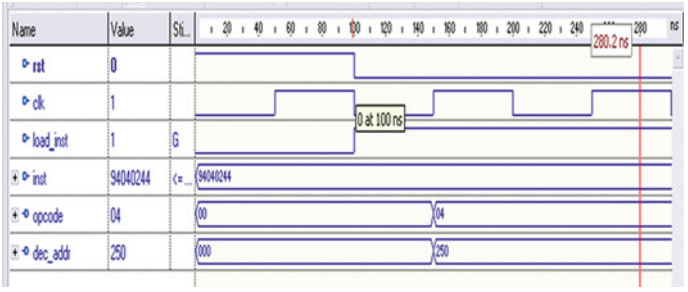


Fig. 4 Instruction decode operation

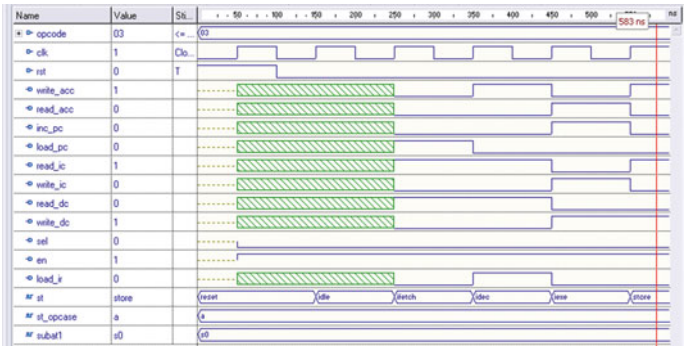


Fig. 5 Control signal interface

The simulation result for the developed approach of multicore processor in a non-mesochronous interface and mesochronous interface is illustrated in Figs. 6 and 7 respectively. The simulation result is illustrated for a instruction of performing a add and accumulate operation for a set of instructions and the simulation result illustrates, the result derived for the proposed approach is at 1750 ns as compared to 3350 ns for a non-mesochronous operation. The delay is encountered in the allocation of operation register and waits period for data/instruction flow.

The implementation of the developed system onto the targeted FPGA device is made using Xilinx ISE 9.1 synthesizer as shown in Fig. 8.

The implementation of the developed approach onto the targeted FPGA is summarized below. The logic requirement is summarized in Table 1.

The operational time period for the proposed approach is obtained as, minimum period: 9.803 ns, and a maximum Frequency: 102.008 MHz) The comparative performance of the multicore design with mesochronous interface and non-mesochronous interface is carried out as listed in Table 2.

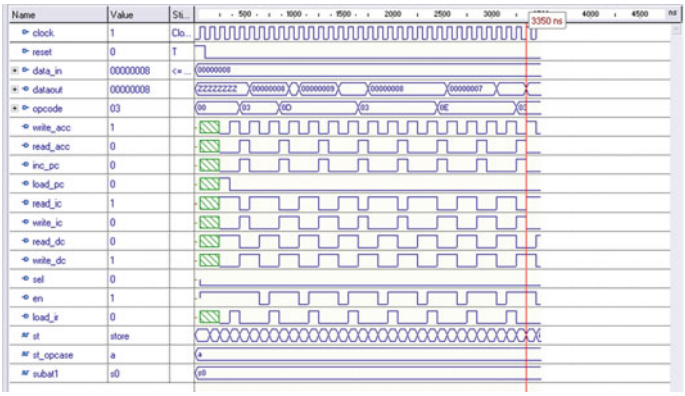


Fig. 6 Simulation result for a multi-core processor without mesochronous interface

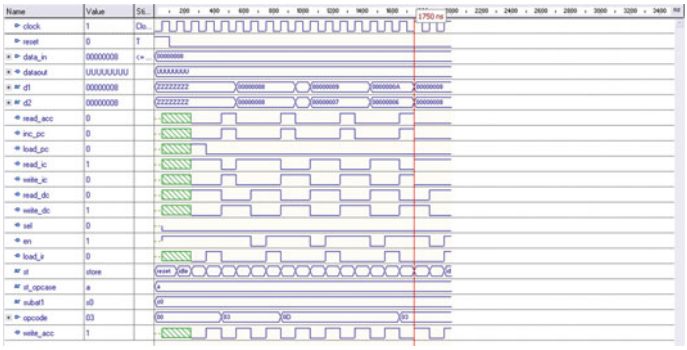


Fig. 7 Simulation result illustrating the processor operation with mesochronous interface

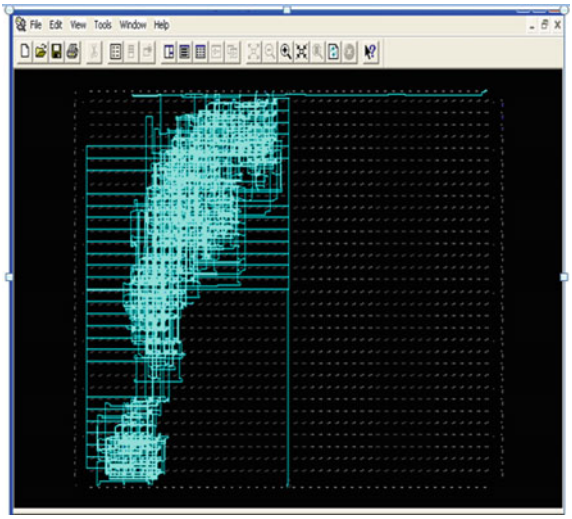


Fig. 8 Implementation of the developed approach over targeted FPGA device

Table 1 Implementation details for the developed approach over XCV300-BG432-6 FPGA device

Target technology	Virtex
Target device	xcv300-bg432-6
Registers	156
IOs	6
BELS	1041
FlipFlops/Latches	192

Table 2 Comparative performance of the developed systems

Design Model	Operational time (ns)	Device speed (MHz)	Logical blocks	Throughput (%)
Multicore design	3350	102.008	1041	55
Integrated Meso-Multicore design	1750	167.105	707	100

The throughput of the developed system is defined as the number of instruction executed in a fixed time period.

5 Conclusion

The synchronization issue for multicore processor interface is been addressed in this paper. The suggested approach defines a new delay computing measure intern of mesochronous approach, where the allocated delay is defined interns of required instruction delay as suggested in a mesochronous operation. The proposed approach, illustrate that under a constant clock delay of stage operation, and processor switching, the delay allocable is directly proportional to the delay required to each instruction operation. In the multicore operation, the instruction delay is derived from the instruction LUT. The function operation is validated, and the system throughput is observed to be higher for the esochronous integrated design in multicore operation.

References

1. Peterson JR, Wight CA, Berzins M (2013) Applying high-performance computing to petascale explosive simulations. In: International conference on computational science. Procedia Comput Sci 18:2259–2268. Elsevier
2. Jose BA, Patelb HD, Shuklaa SK, Talpin J-P (2009) Generating multi-threaded code from polychronous specifications. Electron Notes Theor Comput Sci
3. Liu J, Feld D, Xue Y, Garcke J, Soddemann T (2015) Multicore processors and graphics processing unit accelerators for parallel retrieval of aerosol optical depth from satellite data: implementation, performance, and energy efficiency. IEEE J Sel Top Appl Earth Obs Remote Sens 8(5)
4. Attia KM, El-Hosseini MA, Ali HA (2015) Dynamic power management techniques in multicore architectures: a survey study. Ain Shams Eng J
5. Hatanaka A, Bagherzadeh N (2012) A software pipelining algorithm of streaming applications with low buffer requirements. Comput Sci Eng Electr Eng, Scientia Iranica D 19(3):627–634
6. Shen Z, He H, Yang X, Jia D, Sun Y (2009) Architecture design of a variable length instruction set VLIW DSP. Tsinghua Sci Technol
7. Sheibanyrad A, Greiner A (2006) Two efficient synchronous asynchronous converters well-suited for network on chip in GALS architectures. PATMOS
8. de Dinechina BD, de Massasa PG, Lagera G, Legera C, Orgogozoa B, Reyberta J, Strudel T (2013) A distributed run-time environment for the Kalray MPPAR-256 Integrated Manycore Processor. In: International conference on computational science, ICCS
9. Biedermann A, Huss SA (2013) A methodology for invasive programming on virtualizable embedded MPSoC architectures. In: International conference on computational science, ICCS 2013. Procedia Comput Sci 18:359–368
10. Lewis SJ, Ireland DG, Vanderbauwhede W (2015) Code optimisation in a nested-sampling algorithm. Nucl Instrum Methods Phys Res A 785:105–109
11. Yu Z, Meeuwsen MJ, Apperson RW, Sattari O, Lai M, Webb JW, Work EW, Truong D, Mohsenin T, Baas BM (2008) AsAP: an asynchronous array of simple processors. IEEE J Solid-State Circuits 43(3)

12. Chen Y-K, Kung SY (2008) Trend and challenge on system-on-a-chip designs. J Signal Process Syst 53(1):217–229
13. Vaughan JA, Millstein T (2015) Secure information flow for concurrent programs under total store order. In: Computer security foundations symposium
14. Pollakis A, Wetzel L, Jorg DJ, Rave W, Fettweis G, Julicher F (2014) Synchronization in networks of mutually delay coupled phase-locked loops. New J Phys
15. Brandstätter S, Huemer M (2014) A novel MPSoC interface and control architecture for multistandard RF transceivers. IEEE
16. Kasapaki E, Schoeberl M, Sorensen RB, Muller C, Goossens K, Sparso J (2015) Argo: a real-time network-on-chip architecture with an efficient GALS implementation. IEEE Trans Very Large Scale Integr (VLSI) Syst
17. Kataeva I, Akaike H, Fujimaki A, Yoshikawa N, Takagi N (2012) Experimental demonstration of an operand routing network prototype employing clock control and data synchronization scheme. In: Superconductivity centennial conference
18. Sukanya K, Laxminarayana G. Co-format instruction alignment in pipeline mesochronous operation

A Novel Architecture of High-Speed and Area-Efficient Wallace Tree Multiplier Using Square Root Carry Select Adder with Mirror Adder

Yamini Devi Ykuntam and M. Rajan Babu

Abstract Multipliers are important blocks in designing various digital and high-performance systems, such as processors, signal processing circuits, communication systems. To design high-performance systems, the multiplier used in them should be designed efficiently. Among various multiplier designs, Wallace tree multiplier is fastest in operation. The adder is the main circuit in any multiplier design whose speed of operation affects the performance of multiplier. Among different adder topologies, Square root Carry Select Adder (SQRT CSLA) is good in performance which can be used in the Wallace tree multiplier design as adder block to achieve high performance. In this proposed multiplier design, SQRT CSLA block is modified by replacing Ripple Carry Adder blocks with mirror adder and Binary to Excess-1 Converter (BEC) in order to achieve high performance and less area. Wallace tree multipliers of 4-bit and 8-bit are designed in Verilog. The proposed multiplier is functionally verified using Xilinx ISIM simulator and later synthesized using XST synthesizer in Xilinx ISE design suite. The proposed design of multiplier is compared with at present Wallace tree multiplier designs in terms of number of LUTs and delay (ns).

Keywords Wallace tree multiplier · Ripple carry adder · Square root carry select adder · Binary to excess one converter · Mirror adder

1 Introduction

Noadays, there was a lot of advancement in the electronic industry, which requires the electronic devices to be fast and also smaller in the area. The circuit design with high speed and small area becomes major concerns for the design of digital circuits which

Y. D. Ykuntam (✉) · M. Rajan Babu
Department of ECE, Lendi Institute of Engineering and Technology, Jonnada, Vizianagaram 535005, India
e-mail: yaminieie@gmail.com

M. Rajan Babu
e-mail: rajanbabu.m@lendi.org

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_33

are major building blocks of processors [1, 2]. In processor, speed is an important design requirement. Design of area-efficient and less delay logic systems is the major extensive areas in VLSI circuit design.

The arithmetic unit is the major block of the processor whose speed of performance affects the processor speed of operation. The multiplier is the main building block in the arithmetic unit [3]. There are different types of multiplier architectures in which Wallace tree multiplier topology is good in the speed of operation [4]. In any multiplier topology, the first step of the multiplication process is to produce partial products and add them depending on the topology to generate the final product. The speed of operation of the multiplication process is determined by the addition stage performance.

The first stage of multiplication in Wallace tree multiplier architecture is to generate partial products which are added in later-on stages for getting the final product. From the previous statement, it is obvious that adder block plays an important role in the Wallace tree multiplier to achieve a high speed of operation.

In Wallace tree multiplier architecture, half and full adders are used to add partial products. But carry propagation is slow which is a major negative aspect in this architecture is. Instead of half and full adders, SQRT CSLA is used [5, 6]. But SQRT CSLA consumes more area as it consists of dual Ripple Carry Adder (RCA). To overcome this drawback, Binary to Excess one Converter (BEC) is used in place of RCA with $c_{\text{in}} = 1$ which reduces the area occupied.

But due to BEC delay is increasing in the SQRT CSLA operation. To reduce the area without affecting the performance of SQRT CSLA, RCA with $C_{\text{in}} = 0$ in SQRT CSLA with BEC is replaced with a mirror adder [7]. In this paper, a Wallace tree multiplier designed using SQRT CSLA with mirror adder is proposed and the proposed multiplier design is compared with at-present Wallace tree multiplier architectures in terms of area (number of LUTs) and delay (ns). The multiplier is designed using AOI (AND-OR-INVERT) logic.

2 Basic Wallace Tree Multiplier

A Wallace tree multiplier is very useful hardware design which is used in the digital circuit for multiplication of two integers. In Wallace tree method, the multiplication of two input numbers is done by mainly two steps, i.e., in step 1, partial product rows are generated by using AND gates whereas in step 2 addition process is performed on these partial products in step-by-step process using half and full adders to generate the final output.

Figure 1 shows the 4-bit multiplication process in the Wallace tree multiplication method. Two 4-bit numbers are taken for multiplication. In the first step, each bit of first input number is multiplied with each bit of second input number to generate partial products. As 4-bit numbers are considered, partial products are generated in four rows.

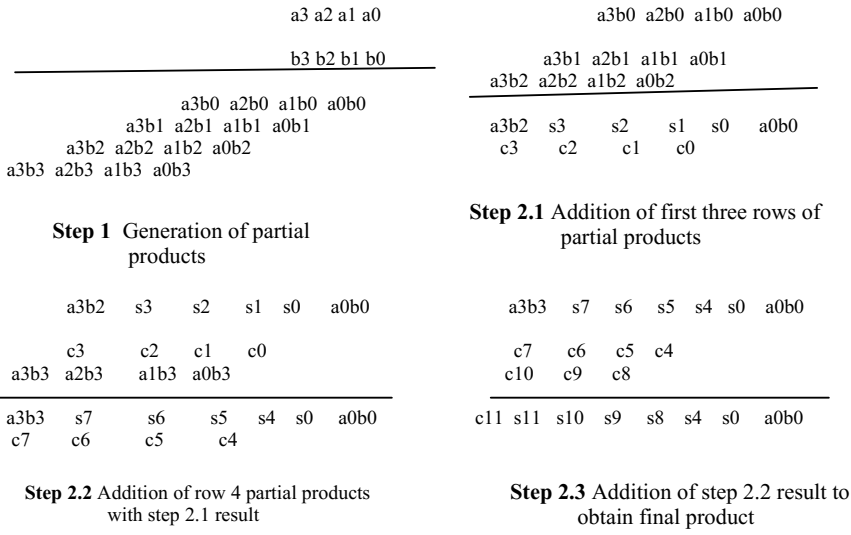


Fig. 1 Steps involved in Wallace tree multiplication method

In the second step, first, the partial products present in the first three rows are added using half and full adders. The result of this addition gives two rows in which the first row consists of sum terms and carry out terms present in the second row. Next, these two rows are added up with the last row of partial product terms and again the same pattern output is resulted. Finally, only two rows are left which are added normally and the resultant is final product output.

The disadvantage of this method to perform the addition of any two bits is that the carry out of pervious two bit addition is required in the final addition process. So there is an increase in delay in obtaining the final product output. To overcome this drawback, half and full adders are replaced by SQRT CSLA for performing addition operation in the final stage of multiplication.

3 Customized Wallace Tree Multiplier

3.1 Wallace Tree Multiplier with SQRT CSLA

To prevail over the drawback of Wallace tree multiplier, a SQRT CSLA [8] is used in the last addition process in stage 2 of Wallace tree multiplier. The addition operation speed is increased due to the usage of SQRT CSLA in the last stage of the multiplier. But SQRT CSLA consists of two RCA blocks for the addition of similar bits due to which multiplier area is increased.

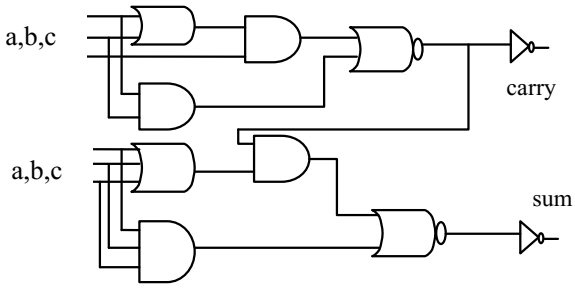


Fig. 2 Mirror adder

3.2 Wallace Tree Multiplier with SQRT CSLA Using BEC

In this multiplier architecture, RCA with $C_{in} = 1$ block in SQRT CSLA is replaced by a BEC block [9–11]. The reason behind is that BEC consists of less number of gates than the RCA. A XOR gate requires five gates to implement in AOI logic. A simple full adder requires 13 gates to implement in AOI logic. So, a 4-bit RCA consists of 52 gates whereas a 4-bit BEC consists of only 18 gates. More than half of the gates are being reduced due to the usage of BEC. Therefore, the problem with the increase in area is solved but there is an increase in delay of SQRT CSLA using BEC operation when compared with normal SQRT CSLA.

4 Wallace Tree Multiplier with SQRT CSLA Using Mirror Adder

To overcome the drawback of Wallace tree multipliers discussed in previous sections, new Wallace tree multiplier architecture is proposed. In this new architecture, RCA block with mirror adder is used in the place of the RCA block with $c_{in} = 0$. The RCA block is formed by connecting full adders in a cascade manner. But here the RCA consists of mirror adders connected in a cascade manner.

There are two reasons to use mirror adder instead of a full adder-

1. Full adder requires 13 gates whereas mirror adder require 10 gates
2. The critical path of full adder consists of six gates, whereas mirror adder critical path has only four gates.

From the above two reasons, it can be said that the mirror adder is best in both area and speed. Figure 2 shows the circuit of mirror adder. And 4-bit BEC shown in Fig. 3 is used in the place of RCA with $c_{in} = 1$ block. In this way, the area and delay of the SQRT CSLA is reduced which in turn reduces the area and delay of the multiplier. A 4-bit Wallace tree multiplier is shown in Fig. 4.

The operation of multiplication is same, i.e., in stage 1 partial products are generated whereas in stage 2 the generated partial products of stage 1 are added up using

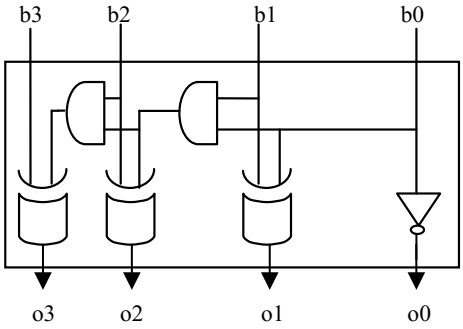


Fig. 3 4-bit BEC

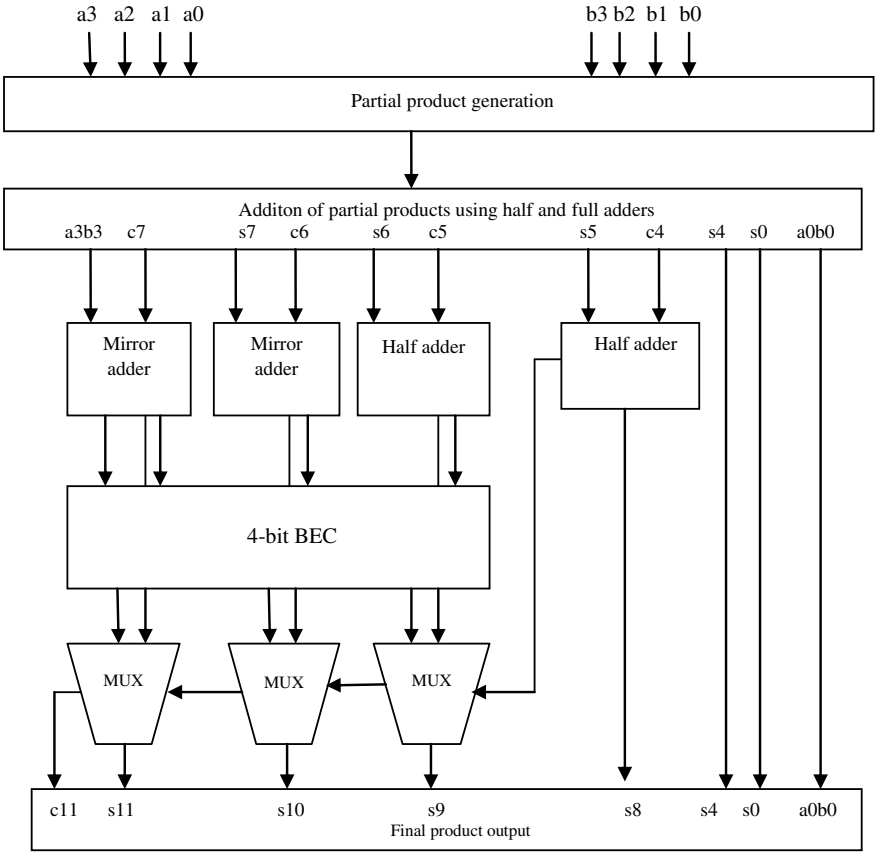


Fig. 4 Wallace tree multiplier using CSLA with mirror adder

half and full adders initially as explained in Sect. 2. The modification is done in the final stage of addition in stage 2, i.e., SQRT CSLA with mirror adder is used instead of half and full adders.

5 Results

All the multiplier architectures are designed for input size of 4-bit and 8-bit using Verilog HDL and functionally simulated using Isim simulator of Xilinx ISE design suite 14.2. The simulation results of 8-bit multipliers are shown in Figs. 5, 6, 7 and 8.

Further, these are synthesized using XST synthesizer. From the synthesis result, area (number of LUTs) and delay (ns) of all multipliers are compared as shown in Table 1. From Table 1, it is noticeable that the proposed multiplier, i.e., Wallace tree

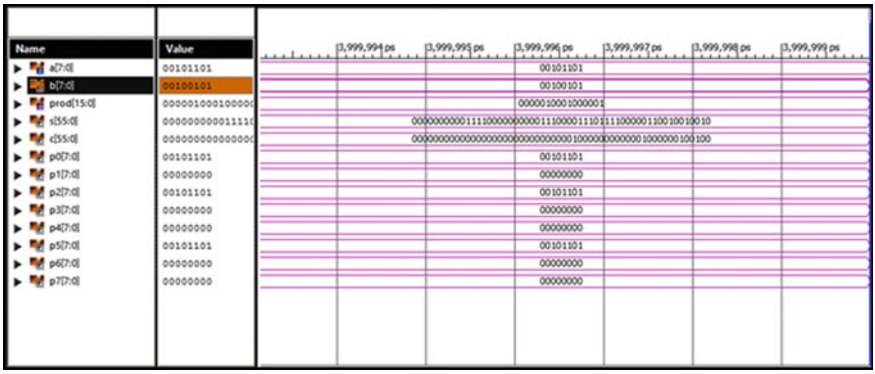


Fig. 5 Simulation result of 8-bit Wallace tree multiplier

Fig. 6 Simulation result of 8-bit Wallace tree multiplier using SQRT CSLA

Fig. 7 Simulation result of 8-bit Wallace tree multiplier using SQRT CSLA with BEC

Fig. 8 Simulation result of 8-bit Wallace tree multiplier using SQRT CSLA with mirror adder

multiplier using SQRT CSLA with mirror adder is occupying less number of LUTs and delay is also less than other three multipliers.

6 Conclusion

A Wallace tree multiplier using SQRT CSLA with mirror adder is designed in this paper. A straightforward method is used to modify the Wallace tree multiplier architecture, i.e., to use mirror adders as a replacement for of full adders in RCA with $c_{in} = 0$ block and BEC in the place of RCA with $c_{in} = 1$ block. Mirror adder has the advantage of less number of gates and low delay compared to full adder and also BEC has less number of gates. With the combination of mirror adder and BEC circuits, an area-efficient and less delay Wallace tree multiplier architecture is designed in this paper. Table 1 in the results section clearly states that the designed multiplier in

Table 1 Area and delay comparison in between the proposed multiplier and other three Wallace multipliers

Word size	Multiplier	Area (number of LUTS)	Delay (ns)
4-bit	Wallace tree multiplier	26	4.33
	Wallace tree multiplier using SQRT CSLA	27	3.03
	Wallace tree multiplier using SQRT CSLA with BEC	24	3.95
	Wallace tree multiplier using SQRT CSLA with mirror adder	19	2.94
8-bit	Wallace tree multiplier	111	8.51
	Wallace tree multiplier using SQRT CSLA	114	6.55
	Wallace tree multiplier using SQRT CSLA with BEC	110	6.78
	Wallace tree multiplier using SQRT CSLA with mirror adder	104	5.10

this paper is occupying less number of LUTs also it has less delay. Therefore, this multiplier design can be used in processors and digital circuits where less area and high speed is prominent.

References

1. Weste N, Harris D (2004) CMOS VLSI design. Addison Wesley, Reading
2. Weste N, Eshragian K (1985–1993) Principles of CMOS VLSI designs: a system perspective, 2nd edn. Addison-Wesley
3. Ercegovac MD, Lang T (2004) Digital arthimetic. Morgan Kaufmann, Elsevier Inc.
4. Bansal H, Sharma KG, Sharma T (2014) Wallace tree multiplier designs: a performance comparison review. Innov Syst Des Eng 5(5)
5. Uma R, Vijayan V, Mohanapriya M, Paul S (2012) Area, delay and power comparison of adder topologies. Int J VLSI Des Commun Syst (VLSICS) 3(1)
6. Rabaey JM (2001) Digtal integrated circuits—a design perspective. Prentice-Hall, Upper Saddle River
7. Ykuntam YD (2016) High speed & area efficient square root carry select adder with mirror adder. NCRCECT
8. He Y, Chang CH, Gu J (2005) An area efficient 64-bit square root carry-select adder for low power applications. In: Proceedings of international symposium on circuits and systems, vol 4, pp 4082–4085
9. Ramkumar B, Kittur HM (2012) Low-power and area-efficient carry select adder. IEEE Trans Very Large Scale Integr (VLSI) Syst 20(2)
10. Ykuntam YD, Nageswara Rao M, Locharla GR (2013) Design of 32-bit carry select adder with reduced area. IJCA 75(2)
11. Bala Sai Kesava R, Lingeswara Rao B, Bala Sindhuri K, Udaya Kumar N (2016) Low power and area efficient Wallace tree multiplier using carry select adder with binary to excess-1 converter. In: 2016 IEEE conference on advances in signal processing (CASP), Cummins College of Engineering for Women

Design of a Highly Reliable and Reconfigurable Pulsed Latch Circuits

B. Keerthi and K. Ragini

Abstract Pulsed latches are the most important storage elements used in many VLSI circuits, due to their low area and power consumption. The pulsed latch circuit consists of a pulse generation circuit, which generates a pulsed clock signal and is given to its corresponding latches. Among different latches namely Modified Hybrid Pulsed Latch (MHLFF), Transmission Gate Pulsed Latch (TGPL), and Static Differential Pulsed Latch (SSASPL), the SSASPL has low power and delay and is used in all designs. The traditional pulse generation circuit used for the generation of the pulsed clock signal does not work well under different process, voltage and temperature (PVT) variations, so the reliability level is very less for these circuits. To enhance the reliability, two novel design approaches namely Header switch based pulsed latch (PL_SW) and MUX-based pulsed latch (PL_MUX) are designed and compared with the traditional pulsed latch in TANNER EDA tool using 180 nm technology.

Keywords Pulsed latches · Flip-flops · PVT variations · Low power

1 Introduction

Flip-flops and pulsed latches are the most considerable storage elements used in many VLSI [1] designs. Nowadays, VLSI designs mostly tend to use pulsed latches due to their less timing overhead, low power, and area consumption. The Master Slave Flip-Flop (MSFF) [2, 3] which uses two latches driven by the opposite clock pulse is replaced with pulsed latch circuits which consists of one latch driven by the pulsed clock signal generated by the pulse generation circuit called as pulser circuit.

B. Keerthi (✉) · K. Ragini
Department of ECE, G. Narayanamma Institute of Technology and Science, Shaikpet, Hyderabad, India
e-mail: keerthi.kittu1595@gmail.com

K. Ragini
e-mail: ragini_kanchimi@yahoo.co.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_34

One drawback with pulsed latch [4] design is that, the traditional pulser circuit does not work well under different process, voltage, and temperature variations. If scaled down voltage is given to the traditional pulser circuit, the pulse becomes too short, so the pulse-width is not sufficient to store the data into the latch. If the pulse is too long, then the latch transparency increases which in turn increases the timing overhead. So at different voltages, the pulser circuit does not work well. Similarly at different temperatures also, the pulser circuit does not work well.

In this article, we are presenting the two novel design approaches namely header switch based pulser circuit and MUX-based pulser circuit which works well under different process, voltage, and temperature variations when compared with the traditional pulser circuit and gains high-reliability level. The reliability level can be judged using the probability of write failure level. The paper is organized as follows, Sect. 2 discusses about different pulsed latch circuit. Section 3 describes about traditional pulse generation circuit. Section 4 gives the detail analysis about the two novel design approaches namely header switch-based and MUX-based pulsed latch circuits. Section 5 gives the conclusion.

2 Different Pulsed Latch Circuits

Sudha Kousalya and G. V. R. Sagar have proposed Modified Hybrid Latch Flip-Flop (MHLFF) [5, 6]. The MHLFF has a small discharge path by only through two transistors (N2-N3). The power consumption of the MHLFF is obtained as 175.23 μW and the delay is obtained as 52 ns. Wael M. Elsharkasy has proposed Transmission Gate Pulsed Latch (TGPL) [7] shown in Fig. 2. The pass transistor logic used in the MHLFF does not provide full swing due to its circuit behavior, so it is replaced by the transmission gate in the TGPL. The power consumption of the TGPL is obtained as 4 μW and delay as 57 ns. The delay is little high in the case of TGPL, when compared with the MHLFF is due to its high clock load (Fig. 1).

Byung Do Yang has proposed Static Differential Sense Amplifier Shared Pulsed Latch (SSASPL) [8] shown in Fig. 3. The drawback of both MHLFF and TGPL circuits is that, they have high delay due to its large discharge path and high clock load. These drawbacks can be overcome using Static Differential Sense Amplifier Shared Pulsed Latch with very low power consumption. The SSASPL circuit is designed with less transistor count and also with simple pulse generation circuit. The power consumption for SSASPL is obtained as 1.98 μW and delay as 51 ns. The power consumption, area, and delay are very much reduced when compared with the other pulsed latch circuits like MHLFF and TGPL. So this SSASPL [9] is used in traditional and two novel pulse generation circuits.

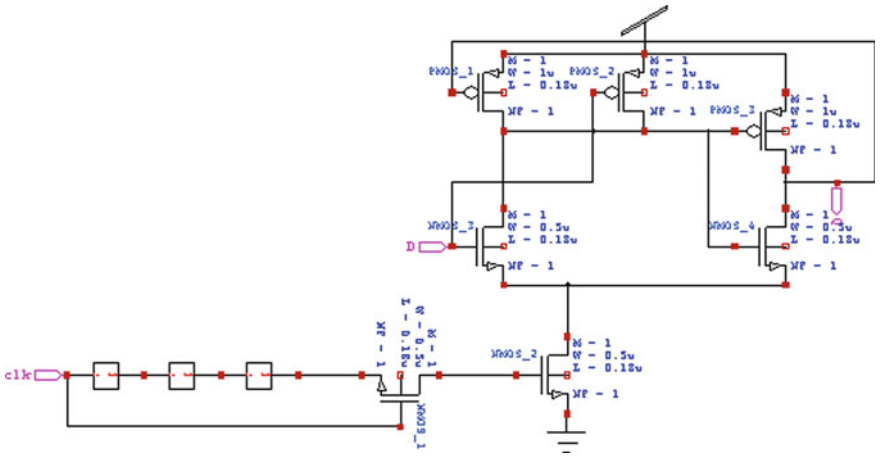


Fig. 1 MHLFF

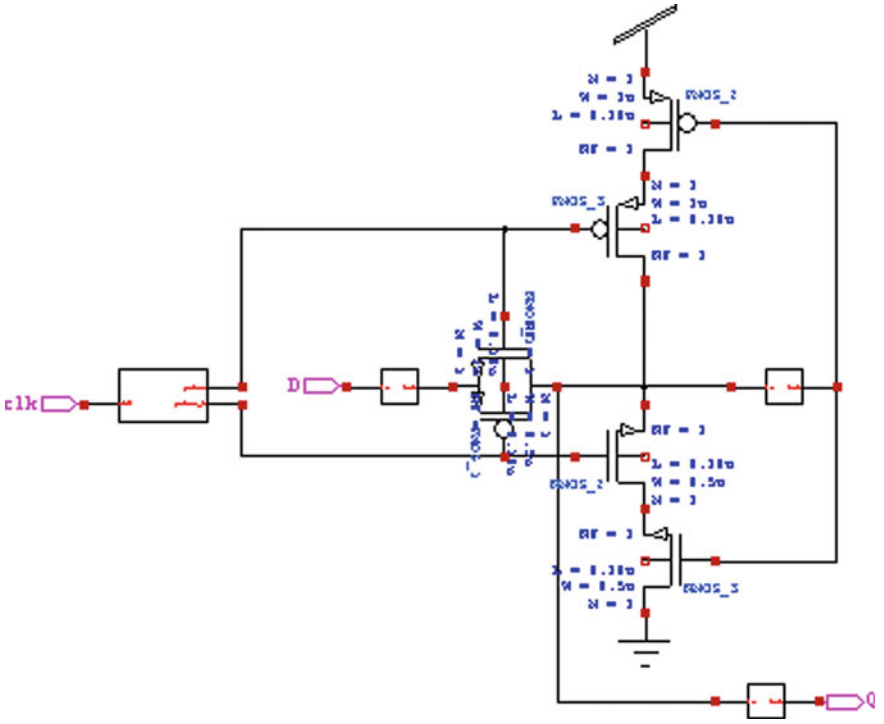


Fig. 2 TGPL

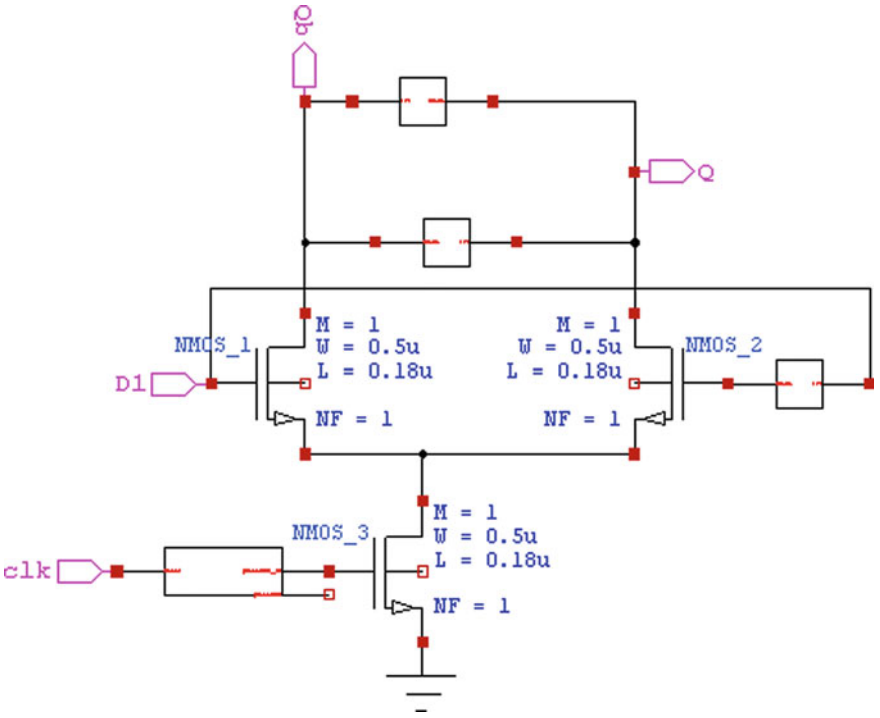


Fig. 3 SSASPL

3 Traditional Pulse Generation Circuit

As shown in Fig. 4, the traditional pulse generation circuit [7] consists of the delay unit, NOT gate, AND gate and a buffer. When the clock pulse is delayed by some unit and inverted, the pulsed clock signal is obtained which is shown in Fig. 5. By using this pulser circuit the clock power very much reduced thereby reducing the overall power consumption. And this pulser circuit is embedded into the Static Differential Sense Amplifier Shared Pulsed Latch (SSASPL) which is the best-pulsed latch among different pulsed latches and using this pulsed latch the 16-bit shift register is implemented and the probability of write failure is calculated for different voltages and temperatures.

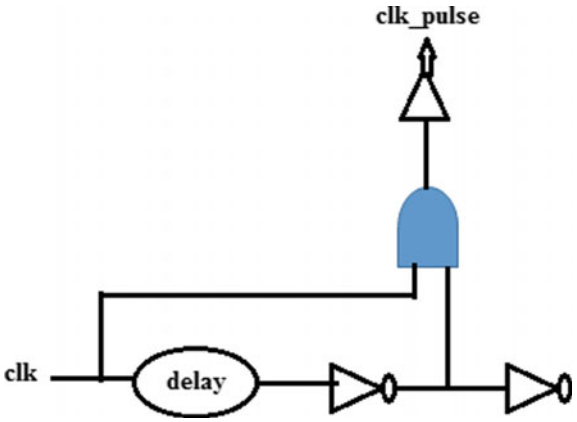


Fig. 4 Traditional pulse generation circuit

Fig. 5 Simulation waveform of traditional pulser

4 Header Switch Based Pulsed Latch (PL_SW) and MUX Based Pulsed Latch (PL_MUX)

4.1 Header Switch Based Pulsed Latch

The header switch-based pulser circuit [7] which works well under different process, voltage, and temperature variations are shown in Fig. 6, uses header PMOS switch which is controlled by the CTRL signal and is given only to the delay unit. We can also use a separate control signal (CTRL) for different PMOS switches. This header switch uses virtual supply voltage which is driven from the main supply voltage. So

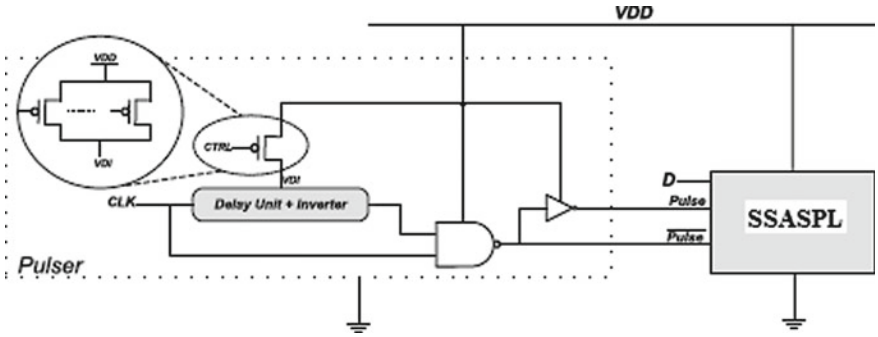


Fig. 6 Header switch based pulsed latch (PL_SW)

using this header switch, the pulse-width of the pulse generation circuit is controlled. This header switch based pulser circuit is similar to the most popular technique called power gating technique.

The CTRL signal used in the header switch based pulser design and MUX-based pulser design structure are same. The CTRL signal is mainly used to control the pulse-width which is required for different voltages. This CTRL signal structure includes three stages namely voltage divider, pseudo-NMOS inverter, and regular CMOS inverters. The header switch based pulser which generates pulsed clock signal, is given to the Static Differential Sense Amplifier Shared Pulsed Latch (SSASPL), which is the best pulsed latch when compared with the MHLFF and TGPL in terms of power and delay. Using this header switch based pulsed latch, the 16-bit shift register is implemented similar to that of the traditional pulsed latch circuit. And also the probability of write failure is calculated for this header switch based pulsed latch and compared with the 16-bit shift register using traditional pulsed latch.

4.2 MUX-Based Pulsed Latch

The MUX-based pulser circuit is another design which is reconfigured using multiplexer to decrease the probability of failures at lower supply voltages and enhances the reliability of the pulse generation circuit. The CTRL signal which is used in the header-based pulser circuit is also used in MUX-based pulser circuit given as the selection line for the multiplexer and also different delays are given as inputs to the multiplexer. So based on the CTRL signal, the delay unit is selected and given to the other logic gates and pulsed clock signal is obtained at the output with different pulse-widths at different voltages. Using this MUX-based pulsed latch, the 16-bit shift register is implemented similar to that of the traditional pulsed latch circuit. And also the probability of write failure is calculated for this MUX based pulsed latch and compared with the 16-bit shift register using traditional pulsed latch (Fig. 7).

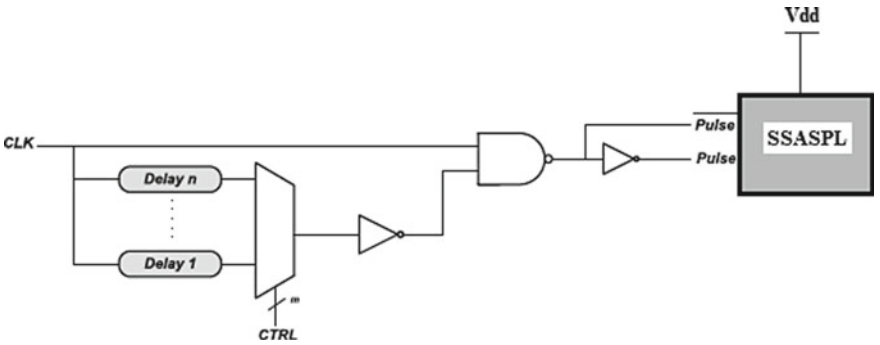


Fig. 7 MUX-based pulsed latch

4.3 Probability of Write Failure Calculation

The probability of write failure indicates the reliability of the particular circuit. If the probability of write failure is less than the target value (1×10^{-8}), then the circuit is highly reliable. The probability of write failure (P_{WF}) for different register implementations using different pulser circuits is calculated as the probability of pulse-width (PW) smaller than the maximum latch write time ($T_{wr\text{-}max}$). The $T_{wr\text{-}max}$ is calculated as the sum of mean of the distribution and standard deviation ($T_{wr\text{-}max} = \mu_{wr} + 6\sigma_{wr}$). The P_{WF} is calculated as

$$P_{WF} = P(PW < T_{wr\text{-}max}) * P(T_{wr} > T_{wr\text{-}max}) \quad (1)$$

From Table 1, it has been shown that the traditional pulser-based shift register undergoes a large number of failures since the values are greater than the target value at scaled down voltage of 0.7 V and the header switch and MUX-based shift registers does not undergo any failures at scaled down voltages. At nominal supply voltage, i.e., 1.05 V, the traditional pulser based shift register, header switch based shift register and MUX-based shift register does not undergo any failures. At different voltage

Table 1 Probability of write failures of different register implementations

T (°C)	$V_{dd} = 1.05$ V			$V_{dd} = 0.7$ V		
	Traditional PL_register	PL_SW register	PL_MUX register	Traditional PL_register	PL_SW register	PL_MUX register
125	0.9×10^{-9}	7×10^{-10}	7×10^{-10}	0.7×10^{-7}	3.1×10^{-9}	1.1×10^{-10}
25	6×10^{-9}	7.9×10^{-10}	11×10^{-9}	2.2×10^{-7}	3.9×10^{-9}	1.1×10^{-9}
−40	3×10^{-9}	2×10^{-9}	6×10^{-9}	3.6×10^{-7}	3.7×10^{-9}	0.6×10^{-9}

scaling levels and temperature levels, the two novel design pulse generation circuits based shift registers do not undergo any failures, so they are highly reliable circuits.

5 Conclusion

The different pulsed latches like MHLFF, TGPL, and SSASPL are compared based on power and delay and choose the best pulsed latch namely SSASPL. The traditional pulser circuit is embedded into the SSASPL and 16-bit shift register is implemented using this and the probability of write failure is calculated for this circuit and similarly both 16-bit shift register using header switch-based pulsed latch and MUX-based pulser are implemented and calculated the probability of failure level and compared with the target value. And later the three shift registers are compared with each other and concluded that the header switch based 16-bit shift register and MUX-based shift register works well under different voltage and temperature variations with less probability of failure level and high level of reliability level. So the two novel approaches are the highly reliable pulse generation circuits.

References

1. Chinnery D, Keutzer K (2002) Closing the Gap between ASIC & Custom: tools and techniques for high-performance ASIC Design. Kluwer, Norwell
2. Stojanovic V, Oklobdzija V (1999) Comparative analysis of master-slave latches and flip-flops for high-performance and low-power systems. IEEE J Solid-State Circuits 34(4):536–548
3. Baumann T et al (2007) Performance improvement of embedded low-power microprocessor cores by selective flip flop replacement. In: Proceedings of 33rd European solid state circuits conference (ESSCIRC), Sept 2007, pp 308–311
4. Paik S, Nam G-J, Shin Y (2011) Implementation of pulsed-latch and pulsed-register circuits to minimize clocking power. In: Proceedings of IEEE/ACM international conference on computer-aided design (ICCAD), Nov 2011, pp 640–646
5. Kousalya S, Sagar GVR (2015) Low power pulsed triggered flip-flop design with conditional pulse enhancement method. IOSR J VLSI Signal Process (IOSR-JVSP) 5(4):49–56
6. Partovi H et al (1996) Flow-through latch and edge-triggered flip-flop hybrid elements. In: IEEE international solid-state circuits conference (ISSCC), Digest of technical papers, Feb 1996, pp 138–139
7. Elsharkasy WM, Khajeh A, Eltawil AM, Kurdahi FJ (2017) Reliability enhancement of low-power sequential circuits using reconfigurable pulsed latches. IEEE Trans Circuits Syst 64(7)
8. Yang B-D (2015) Low-power and area-efficient shift register using pulsed latches. IEEE Trans Circuits Syst 62(6)
9. Keerthi B, Ragini K (2018) Design of a low power and high speed 512-bit shift register using static differential sense amplifier shared pulsed latch circuit. IOSR J VLSI Signal Process (IOSR-JVSP) 8(2):38–43

Design of a Novel High-Speed- and Energy-Efficient 32-Bit Carry-Skip Adder

B. Sanjana and K. Ragini

Abstract In this paper, different 32-bit carry-skip adders, i.e., Fixed Stage Size-Conventional carry-skip adder (FSS-Conv CSKA), Variable Stage Size-Conventional carry-skip adder (VSS-Conv CSKA), Fixed Stage Size-Concatenation and Incrementation carry-skip adder (FSS-CI CSKA) and Variable Stage Size-Concatenation and Incrementation carry-skip adder (VSS-CI CSKA) are designed and compared in terms of power, energy, critical path delay, power-delay product, and energy-delay product using 45 nm static CMOS technology for different range of supply voltages, i.e., 0.7v, 0.9v, 1.1v. The results that are obtained using tanner EDA simulations reveal that the Concatenation and Incrementation carry-skip adder with fixed and variable stage size has 51 and 49% improvement in the critical path delay and energy, compared with those of fixed stage size-conventional carry-skip adder and Variable Stage Size-Conventional carry-skip adder.

Keywords Carry-skip adder (CSKA) · FSS-Conv CSKA · VSS-Conv CSKA · FSS-CI CSKA · VSS-CI CSKA

1 Introduction

Adders are one of the building blocks of arithmetic and logic unit. By increasing the speed of the adders, the overall performance of the arithmetic and logic unit increases. An adder is an elementary unit of microprocessors and it is used in ALU, floating-point unit, memory address generation unit, etc. There are different types of adders like ripple carry adder (RCA), carry select adder (CSLA), carry look ahead adder (CLA), carry-skip adder (CSKA). Among these different adders, carry-skip

B. Sanjana (✉) · K. Ragini
Department of ECE, G. Narayanamma Institute of Technology and Science, Shaikpet, Hyderabad, India
e-mail: boinisanjana@gmail.com

K. Ragini
e-mail: ragini_kanchimi@yahoo.co.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_35

adder is one of the most efficient adders with less delay and energy consumption. In this article, the carry-skip adders are designed using different implementation schemes, i.e., concatenation and incrementation and different stage sizes, i.e., Fixed Stage Size (FSS) and Variable Stage Size (VSS).

The paper is organized as follows, Sect. 2 explains about 32-bit conventional carry-skip adder designs, while Sect. 3 describes the 32-bit Concatenation and Incrementation carry-skip adder designs. Section 4 illustrates about power, critical path delay, power-delay product, energy and energy-delay product of FSS-Conv CSKA, VSS-Conv CSKA, FSS-CI CSKA, and VSS-CI CSKA and Sect. 5 is about conclusion.

2 32-Bit Conventional Carry-Skip Adder Circuits

2.1 32-Bit Conventional Carry-Skip Adder

The 32-bit conventional carry-skip adder [1] consists of RCA blocks [2–5], logics gates, and multiplexers as shown in Fig. 1. Each RCA block has a chain of full adders which generates sum and carry outputs. The inputs from 2-bit numbers are given to XOR gates which produce propagated signals $P_1 = A_1 \oplus B_1$ and so on, where P_1 is the propagation signal. The output from the propagated signal is given to the MUX as selection line. The skip logic is performed using logic gates and multiplexers. The carry output and carry input at each stage of the RCA are given as inputs to the multiplexer. Based on the selection line, the MUX output is produced. The output from the MUX is given as carry input to the next stage and the same process is repeated up to 32-bits. The 32-bit Conventional CSKA and 32-bit Concatenation and incrementation CSKA systems can be implemented using two-stage sizes

1. Fixed stage size
2. Variable stage size

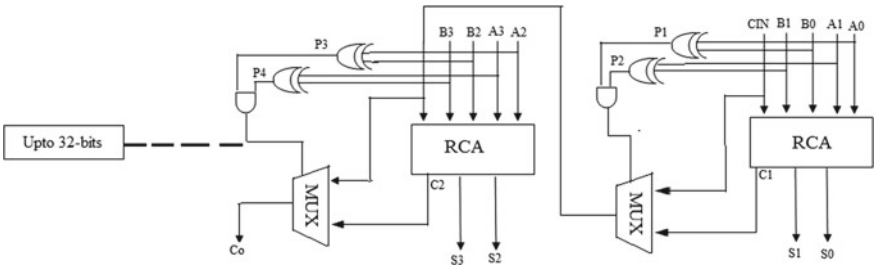


Fig. 1 32-bit conventional carry-skip adder

2.1.1 32-Bit Fixed Stage Size-Conventional Carry-Skip Adder

The 32-bit Fixed Stage Size-Conventional Carry-Skip Adder consists of RCA blocks, multiplexers, and logic gates in each stage. The stage sizes are fixed in FSS-Conv CSKA. Each RCA has four cascaded FA's which generates sum and carry outputs. The summation of A_1 and B_1 gives the propagation signal and so on. The skip operation is performed using the multiplexer and the logic gates in each stage. The propagated signals are given to multiplexer as the selection line and the inputs of the multiplexer are the carry in and carry out of the RCA. The output from the MUX is given as carry input to the next stage and the same process is repeated up to 32-bits. The critical path delay for FSS-Conv CSKA is given by Eq. 1 (Fig. 2).

$$T_D = [M \times T_{\text{CARRY}}] + \left[\left(\frac{N}{M} - 1\right) \times T_{\text{MUX}}\right] + [(M-1) \times T_{\text{CARRY}} + T_{\text{SUM}}] \quad (1)$$

Here M is the FA's of each stage and is calculated by $(0.5N\alpha)^{1/2}(\alpha = T_{\text{MUX}}/T_{\text{CARRY}})$, i.e., 4, N is the total no. of bits, i.e., 32 bits and $Q = (N/M)$, i.e., eight stages. T_{CARRY} and T_{SUM} are the delays of sum and carry produced by the RCA's, T_{MUX} is the delay of MUX output.

2.1.2 32-Bit Variable Stage Size-Conventional Carry-Skip Adder

The 32-bit Variable Stage Size-Conventional Carry-Skip Adder [6] consists of RCA blocks, multiplexers, and logic gates in each stage. Since it is variable stage size, the RCA blocks are designed using different sizes. The stage size of VSS-Conv CSKA is {1,1,1,2,2,2,3,3,4,3,3,2,2,1,1,1}. Each RCA has {1,1,1,2,2,2,3,3,4,3,3,2,2,1,1,1} cascaded FA's which generates sum and carry outputs and the skip operation is performed using the multiplexer and the gates in each stage. The summation of A_1 and B_1 gives the propagation signal and so on. The propagated signals are given to multiplexer as the selection line and the input of the multiplexer is the carry in and carry out of the RCA. The output from the MUX is given as carry input to the next stage and the same process is repeated up to 32-bits. The critical path delay for VSS-Conv CSKA is given by Eq. 2 (Fig. 3).

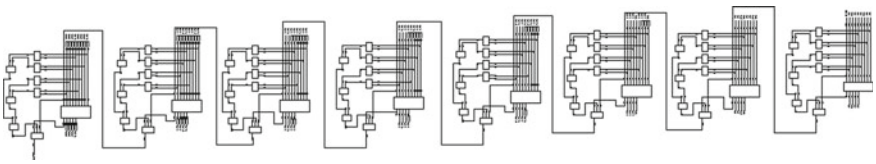


Fig. 2 32-bit fixed stage size-conventional carry-skip adder

Stage Size	M16	M15	M14	M13	M12	M11	M10	M9	M8	M7	M6	M5	M4	M3	M2	M1
	1	1	1	2	2	3	3	4	3	3	2	2	2	1	1	1

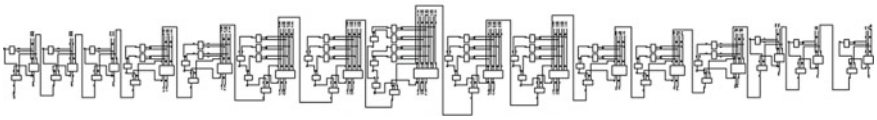


Fig. 3 32-bit variable stage size-conventional carry-skip adder

$$T_{D,\text{opt}} = \left(2\left\lceil\frac{\alpha}{2}\right\rceil - 1\right)T_{\text{CARRY}} + \left(2\sqrt{\frac{N}{\alpha}} - 1\right)T_{\text{MUX}} + T_{\text{SUM}}, \tag{2}$$

where

$$\alpha = T_{\text{SKIP}}/T_{\text{CARRY}}$$

where T_{CARRY} and T_{SUM} are the delays of the sum and carry produced by the RCA's, T_{MUX} is the delay of MUX output, where T_{MUX} is replaced with T_{SKIP} and N is the total number of bits. Disadvantages of 32-bit conventional carry-skip adder circuits: Due to the increase of gates in conventional carry-skip adder, the delay increases and thereby the speed of the circuit decreases.

3 32-Bit Concatenation and Incrementation Carry-Skip Adder Circuits

3.1 32-Bit Concatenation and Incrementation Carry-Skip Adder

The 32-bit concatenation and incrementation of carry-skip adder [1] are shown in Fig. 4, where AND-OR-Inverter (AOI) and OR-AND-Inverter (OAI) gates are used as skip logic instead of 2:1 multiplexer because these gates have a lower delay when compared with the 2:1 multiplexer. The first stage has only one RCA block and the remaining stages, i.e., from 2 to 7 have RCA blocks and incrementation block with skip logics. The incrementation block is shown in Fig. 5, which contains a chain of half adders, where the inputs to the incrementation block are taken from the intermediate results of the RCA and the previous RCA carry. The output of the incrementation block produces sum outputs. The inputs of the skip logics contain the intermediate results produced by the RCA, the carry out of the present RCA and the carry out of the previous RCA and the same process is repeated up to 32-bits.

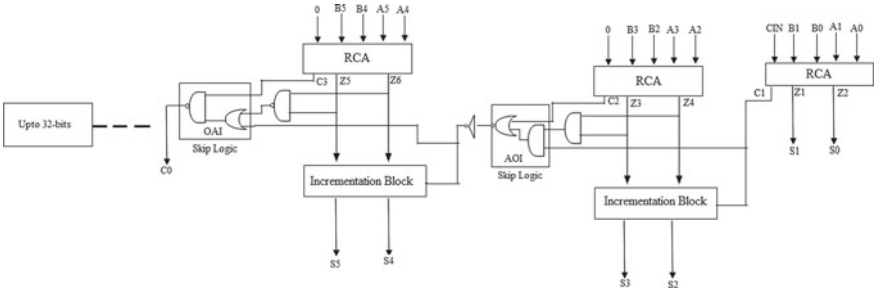


Fig. 4 32-bit concatenation and incrementation carry-skip adder

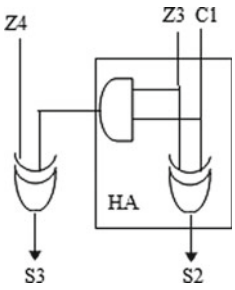


Fig. 5 Incrementation block

3.1.1 32-Bit Fixed Stage Size-Concatenation and Incrementation Carry-Skip Adder

The 32-bit Fixed Stage Size-concatenation and incrementation Carry-Skip Adder consists of RCA blocks, skip logics, and incrementation blocks in each stage. The stage sizes are equal in FSS-CI CSKA. Each RCA has four cascaded FA's and the skip operation is performed using AOI (AND-OR-Inverter) and OAI (OR-AND-Inverter) gates. The inputs of the skip logics contain the intermediate results produced by the RCA's, the carry out of the present RCA and the carry out of the previous RCA and the same process is repeated up to 32-bits. For FSS-CI CSKA, the critical path delay is given by Eq. 3 (Fig. 6).

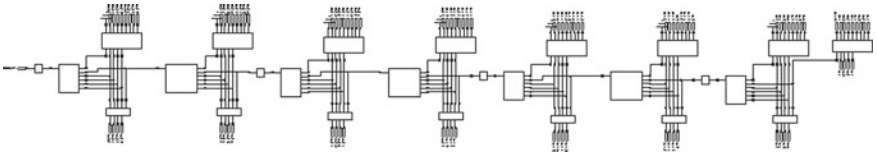


Fig. 6 32-bit fixed stage size-concatenation and incrementation carry-skip adder

$$T_D = [M_1 T_{\text{CARRY}}] + [(Q-2)(T_{\text{OAI}} + T_{\text{OAI}}/2)] + \left[(M_{Q-1})T_{\text{AND}} + T_{\text{XOR}}\right], \quad (3)$$

where

$$M = \frac{\sqrt{N(T_{\text{AOI}} + T_{\text{OAI}})}}{\sqrt{2(T_{\text{CARRY}} + T_{\text{AND}})}},$$

where M is the number of full adders in each stage, Q is the total number of stages, N is the total number of bits, T_{CARRY} is the delays of carry produced by the RCA's, T_{AOI} and T_{OAI} are the delays of AND-OR-Inverter and OR-AND Inverter logics and T_{AND} and T_{XOR} are the delays of AND and XOR gates of incremention block.

3.1.2 32-Bit Variable Stage Size-Concatenation and Incrementation Carry-Skip Adder

The 32-bit Variable Stage Size-concatenation and incrementation Carry-Skip Adder consists of RCA blocks, skip logics, and incrementation blocks in each stage. Since it is variable stage size, the RCA blocks are designed using different sizes (Fig. 7).

The stage sizes of the VSS-CI CSKA is {1,1,1,2,2,3,3,4,5,4,3,2,1}. Each RCA has {1,1,1,2,2,3,3,4,5,4,3,2,1} cascaded FA's and the skip operation is performed using AOI (AND-OR-Inverter) and OAI (OR-AND-Inverter) gates. The inputs of the skip logics contain the intermediate results produced by the RCA's, the carry out of the present RCA and the carry out of the previous RCA and the same process is repeated up to 32-bits. For FSS-CI CSKA, the critical path delay is given by Eq. 4.

$$T_{D,\text{opt}} = \left(2\left\lceil\frac{\alpha}{2}\right\rceil - 1\right)T_{\text{CARRY}} + \left(2\sqrt{\frac{N}{\alpha}} - 1\right)T_{\text{SKIP}} + T_{\text{SUM}}, \quad (4)$$

where $\alpha = (T_{\text{AOI}} + T_{\text{OAI}})/(2 \times T_{\text{CARRY}})$, N is the total number of bits, T_{CARRY} and T_{SUM} are the delays of sum and carry produced by the RCA's, T_{SKIP} is the delay

Stage Size	M13	M12	M11	M10	M9	M8	M7	M6	M5	M4	M3	M2	M1
	1	2	3	4	5	4	3	3	2	2	1	1	1

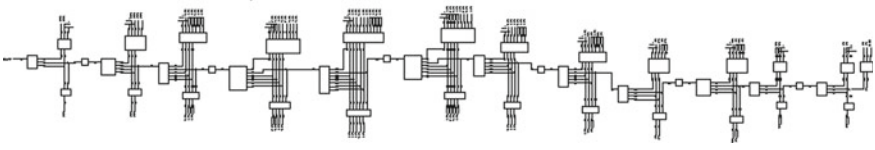


Fig. 7 32-bit variable stage size-concatenation and incrementation carry-skip adder

of Skip output, T_{AOI} and T_{OAI} are the delays of AND-OR-Inverter and OR-AND Inverter logics.

4 Results and Discussion

Figure 8 shows the critical path delay versus different supply voltages. From Fig. 8, the VSS-CI CSKA has less delay compared with other structure and the delay of the CI CSKA is reduced to 51% in case of FSS and 55% in case of VSS when compared with the FSS-Conv CSKA and VSS-Conv CSKA. The power versus different supply voltages is shown in Fig. 9. From Fig. 9, the power of the VSS-CI CSKA has slightly increased when compared with other structure due to the increase of stages. The PDP versus different supply voltages is shown in Fig. 10. From Fig. 10, the VSS-CI CSKA has less PDP when compared with FSS-Conv CSKA and VSS-Conv CSKA, FSS-CI CSKA structure. Figure 11 shows the EDP versus different supply voltages. From Fig. 11, the VSS-CI CSKA has less EDP when compared with FSS-Conv CSKA and VSS-Conv CSKA, FSS-CI CSKA. The EDP of the FSS-CI CSKA and VSS-CI CSKA is ~75–80% less compared with the FSS-Conv CSKA and VSS-Conv

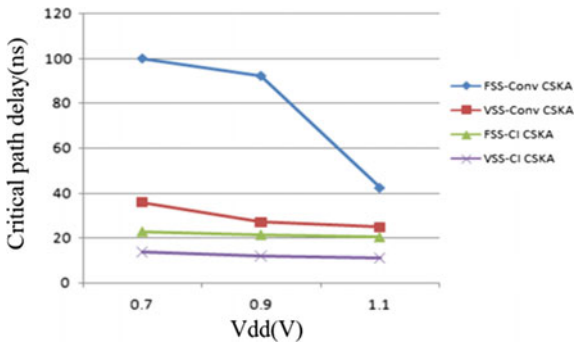


Fig. 8 Critical path delay versus the supply voltages

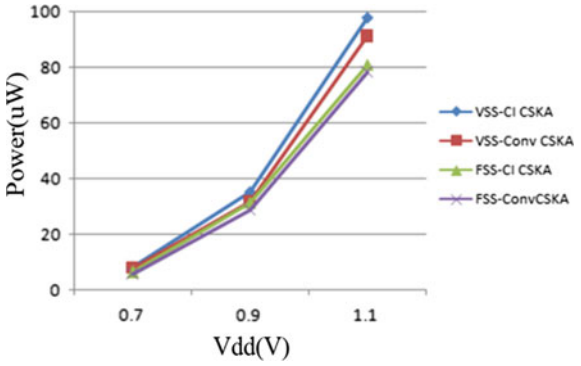


Fig. 9 Power versus the supply voltages

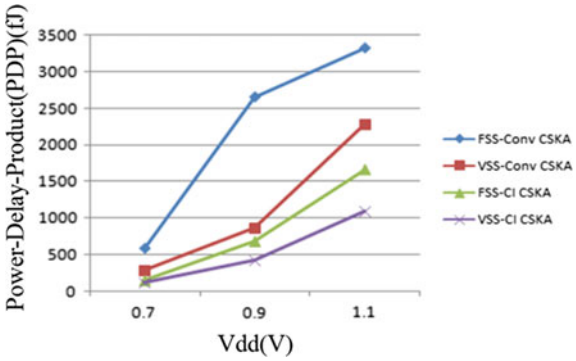


Fig. 10 PDP versus the supply voltages

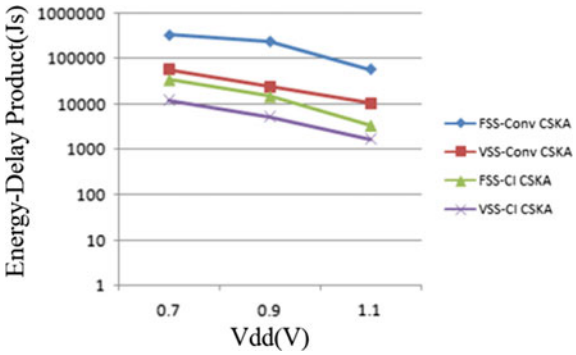


Fig. 11 EDP versus the supply voltages

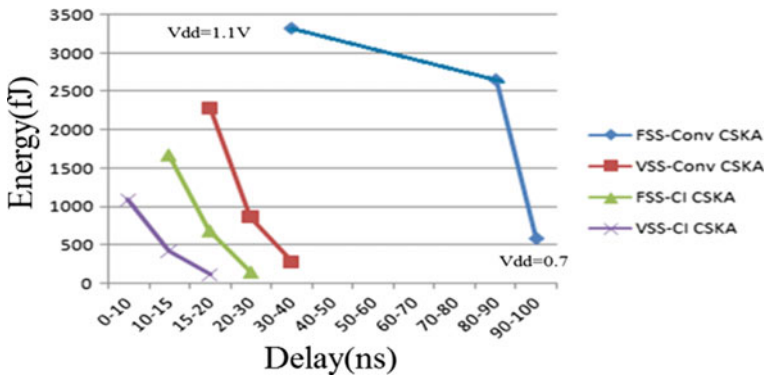


Fig. 12 Energy versus delay

CSKA. Figure 12 shows the energy versus delay. From Fig. 12, the VSS-CI CSKA has less energy compared with FSS-CI CSKA, FSS-Conv CSKA, and VSS-Conv CSKA techniques.

5 Conclusion

The FSS-Conv CSKA, VSS-Conv CSKA, FSS-CI CSKA, and VSS-CI CSKA are designed. These designs are simulated using tanner EDA software with 45 nm technology and the length of the PMOS and NMOS transistor are 0.045u, width of PMOS and NMOS are 0.24u and 0.12u. The FSS-CI CSKA and VSS-CI CSKA techniques have less critical path delay, energy, PDP, and EDP when compared with the FSS-Conv CSKA and VSS-Conv CSKA techniques. The FSS-CI CSKA and VSS-CI CSKA have ~51 and ~49% less critical path delay and energy, compared with those of the FSS-Conv CSKA, VSS-Conv CSKA.

References

1. Bahadori M, Kamal M, Afzali-Kusha A, IEEE, Pedram M, Fellow, IEEE (2016) High-speed and energy-efficient carry skip adder operating under a wide range of supply voltage levels. IEEE Trans Very Large Scale Integr (VLSI) Syst 24(2):421–433
2. Bhattacharyya P, IEEE, Kundu B, Ghosh S, Kumar V, Member, IEEE, Dandapat A, Member, IEEE (2015) Performance analysis of a low-power high-speed hybrid 1-bit full adder circuit. IEEE Trans Very Large Scale Integr (VLSI) Syst 23(10):2001–2007
3. Goel S, Kumar A, Bayoumi MA (2006) Design of robust, energy efficient full adders for deep-submicrometer design using hybrid-CMOS logic style. IEEE Trans Very Large Scale Integr (VLSI) Syst 14(12):1309–1321
4. Zhang M, Gu J, Chang C-H (2003) A novel hybrid pass logic with static CMOS output drive full-adder cell. In: Proceedings of the international symposium on circuits systems, May 2003, pp 317–320
5. Tung C-K, Hung Y-C, Shieh S-H, Huang G-S (2007) A low-power high-speed hybrid CMOS full adder for embedded system. In: Proceedings of IEEE conference on design diagnostics electronic circuits and systems, Apr 2007, vol 13, pp 1–4
6. Alioto M, Palumbo G (2003) A simple strategy for optimized design of one-level carry-skip adders. IEEE Trans Circuits Syst I: Fundam Theory Appl 50(1):141–148

Reduction of Coupling Transition by Using Multiple Encoding Technique in Data Bus and Its Power Analysis

V. Shavali, G. M. Sreerama Reddy and P. Ramana Reddy

Abstract In Deep submicron technology the on chip and off chip buses plays an Important Role in the total amount of power consumed by chip and embedded chip memory. The capacitance plays significant role in the power analysis and maximum amount of power is utilized by chip of capacitance and less power is utilized by the devices. So the power has to be Reduced this can be done by encoding technique which in turn reduces switching activities. There are so many techniques to reduce energy consumption. In this paper a new encoding scheme which saves power. That reduce data bus energy consumption. This method produce an energy saving method based on Similar values and multi-coding which can further reduce on-chip data bus. In this results of simulation are reduce the ratio of Switching Activity (SA) on data bus lines, maximum ratio of Energy saving (ES) and the average ratio with toggle state Bus values.

Keywords Switching values · Similar · Multiple values

1 Introduction

The power consumption plays significant role in the design of Integrated circuits. The methods like scaling, Interconnections reduces the power when the circuits is LSI But for the component level it is better to use encoding techniques. Hence power is saved by using encoding techniques. There are many techniques like Invert method, even method, odd Method. Even and odd method, half width method but these method

V. Shavali (✉)
Department of ECE, JNTUA, Anantapuramu, Andhra Pradesh, India
e-mail: v.shavali@gmail.com

G. M. Sreerama Reddy
Department of ECE, CBIT, Kolar, Karnataka, India

P. Ramana Reddy
Department of ECE, JNTUACEA, Anantapuram, Andhra Pradesh, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_36

does not deals with dynamic power and with self switching activities [1–5]. So the proposed method is used. This method is also used for memory based circuit design where the data can be stored, deleted and manipulated the values against the other Values. Energy consumption has on-chip multi-core circuits. Basically cache has two conditions. They are (1) True: It means hit the data (it can be stored). (2) False: It means fails the data (it cannot be stored). Its area and power consumption for on-chip bus. SVE (similar value encoding) has data bus energy saving. We may implement new technology is FV-MSB. SVE has perform single-core frequent value, SV-MSB has perform multi-core frequent value cache. It has not sufficient energy consumption. New technique may introduce new technology HFVCMC for exactly energy consumption.

2 Research Work (Existing)

In Adaptive bus encoding scheme may compress the data. We may loss the bits and coupling capacitance will reduce. Energy saving technique has different types. They are (1) Physical: Its modifies the layer. (2) Shield: It is a dummy layer. (3) Repeater: It has multiple time data transmission. (4) DVS: Its reduces the voltage. We may Introduce new technique is data link level and It has SA is reduced and Data transmission is fast for chip area increase. It has two categories. (1) Lower self SA (2) Inverting encoding technique reduce the SA. It has easy dynamic energy consumption are odd/even/Gray code encoding techniques. Energy consumption may reduce for odd number is 111 and even is 1111. If it is odd we can add (or) remove 1 and If it is even we can add (or) remove 1 Its drawback is extra control lines and also complexity of circuit is increase.

3 Energy Model (Proposed)

(a) **Single-core and multi-core platform**

Its L1 has 32 KB size where L1 = I1 and D1 and has 4-way set associative first data (D1) and first instruction (I1) cache. Second level (L2) shared cache with size (Figs. 1 and 2).

Its L1 has 32 KB size and L2 has 1 MB size. Five FVMC (Frequent value multiple code) modules-one at each of cores and one at L2 end. Each of private L1 split cache is write-through with a store buffer. The shared L2 cache is write-back and maintains inclusion with respect to L1 cache. In multi-core has cores and L2 cache are connected by bus. Its coupling capacitance of energy consumption will be increase for bit storage purpose.

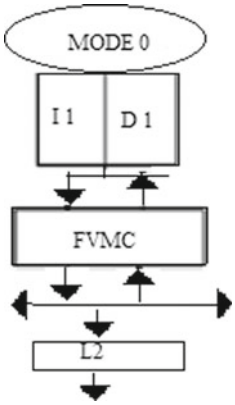


Fig. 1 Single-core plat platform

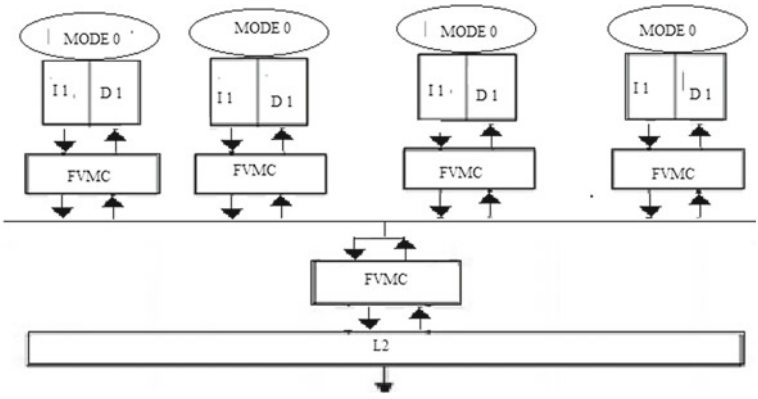


Fig. 2 Multicore platform

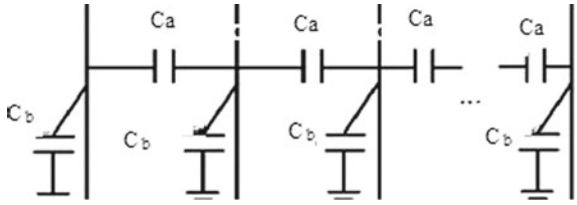


Fig. 3 Capacitance in DSM model

(b) **Energy model platform**

c_a has coupling capacitance between adjacent lines and C_b has self capacitance. The coupling factor is defined as the ratio of coupling capacitance and self capacitance (Fig. 3).

$$K = C_a / C_b$$

Table 1 The values of i, $i + 1$ for state of capacitance

$Y_{(i,i+1),j}$	Data line i at a time i to $i + 1$				
Data line $i + 1$ at a time j to $j + 1$		$0 \rightarrow 0$	$0 \rightarrow 1$	$1 \rightarrow 0$	$1 \rightarrow 1$
	$0 \rightarrow 0$	0	1	0	0
	$0 \rightarrow 1$	1	0	2	0
	$1 \rightarrow 0$	0	2	0	1
	$1 \rightarrow 1$	0	0	1	0

K is high when technology shrinks and it will become larger in future. K is 0 only in self-capacitance for traditional buses. K is about 5 for easy to calculate SA makes the bus line discharge and recharge between 0 and 1.

$$p = p_s + p_i \tag{1}$$

where p—Power consumption, p_s—Power consumption of self-SA, p_i—Power consumption of coupling capacitance

$$P_S = \mathrm{CL} * V_{DD}^2 * \sum_{i=1}^{n} \sum_{j=1}^{t-1} X_{I,j} \tag{2}$$

$$P_I = \mathrm{CI} * V_{\mathrm{DD}}^2 * \sum_{i=0}^{n-1} \sum_{j=0}^{t-1} (Y)_{(i,i+1),j} \tag{3}$$

Put $X = \sum_{i=1}^{n} \sum_{j=1}^{t-1} X_{I,j}$ and
$Y = \sum_{i=0}^{n-1} \sum_{j=0}^{t-1} (Y)_{(i,i+1),j}$ into (2) and (3) Respectively to (1)
P is given by

$$P = (X * \mathrm{CL} + Y * \mathrm{CI}) * V^2\mathrm{DD} \tag{4}$$

Put $K = \mathrm{CI}/\mathrm{CL}$ into (4) we can get

$$P = (X + K * Y) * \mathrm{CI} * V^2\mathrm{DD} \tag{5}$$

For simplicity we denote E_{enc} is the total SA distance and E_{enc} is given by

$$E_{\mathrm{enc}} = X + K * Y \tag{6}$$

From (5) and (6), to obtained a minimum value of P, to make the value of E_{enc} minimum (Table 1).

The method reduces number of bus SA for data transfer by HFVC at each side with frequent values and using MC with other values. When value needs to transfer over bus, the first check whether it is HFVC of sender. If it is, only index value instead of actual value is transmitted and number of active lines is reduced. For non-frequent value and loader bits of partial frequent values, which were not stored in cache. We introduce MC-module to further reduce SA and energy consumption may decrease.

4 Design Approach Encoding

There are HFVC compress FVC and FHBC-FVC stands for frequent value cache and FHBC stands for frequent higher bit cache. HFVC has same bit-width which maximum 32-bit value requires a 32-bit mask and m-bit value requires *m*-bit mask. In encoding scheme performs odd inverting, even inverting, inverting and we can do this data transmission. Its SA and power may reduce.

4.1 Encoder

FHBC discuss two types. They are (1) indication line and (2) Select size of *m*. (1) **Indication line**: It use one bit line to indicate sending value is frequent (or) non-request value. Signal is 1 then sending value is frequent value stored in FVC or FHBC. Signal is 0 indicate an non-frequent value. We should not add an extra line to indicate a frequent value coming from FVC or FHBC. (2**) Select size of m**: *m*-bit brings large number of hits in FHBC. Large number of hits need not necessarily imply a large reduction in SA. If *m* is large, the hit rate will decrease but effect of hit on SA reduction will increases.

A. **Encoder**:

HFVC has to follow the four rules. They are (1) **A hit only in FVC**: It will sent actual value to selector. (2) **A hit in SVC and HFVC at same time**: SVC has higher priority and SVC will sent actual value to selector and reduce SA. (3) **A hit only in HFVC**: Higher order bit in HFVC is sent to OR gate and lower order bit sent to encoder. (4) **A miss in SVC and HFVC**: value sent in directly entered encoder. Encoder calls Algorithm to select encoding scheme and produce encoded value and coding number. The encoder value is sent to following selector and coding number is sent to receiver (Fig. 4).

Algorithm

(1) Initialize the value of *n*, where *n* is the length of Bus and the Data is assigned (2) Divide the length into two parts and find out Same occurrence value and High frequently occurs value (3) The same occurrence value is transmitted into the Bus

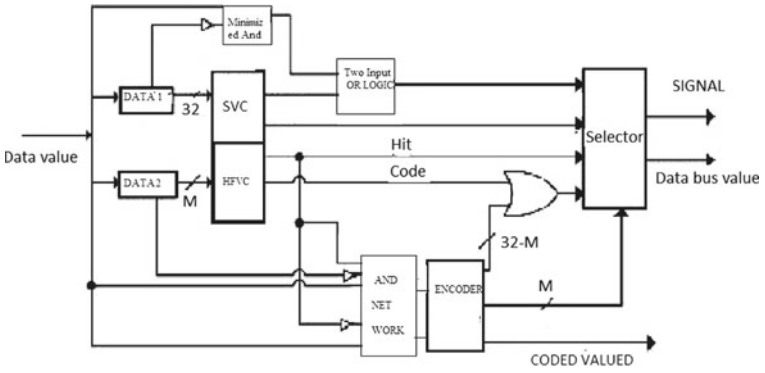


Fig. 4 Basic level encoder

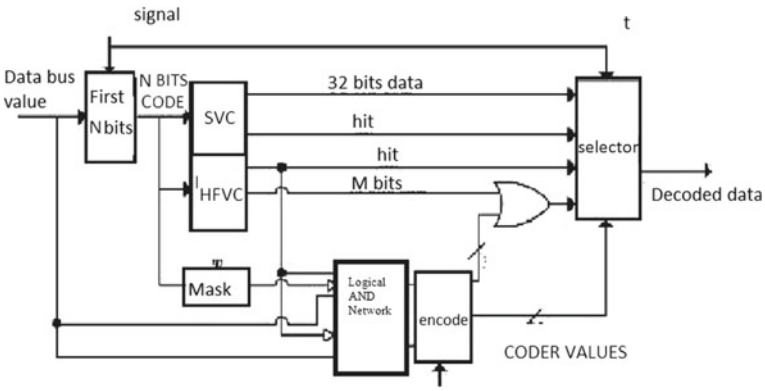


Fig. 5 Basic level decoder

(4) For the high occurrence value divide the length into two parts as even and odd position and is passed into the encoder (5) find out the coupling transitions of even positions and odd positions (6) If odd position is greater invert odd position, if even is greater invert even number of bits. If both odd and even are greater send data as it is. Invert all the bits if high frequently value is greater than half of the length of bus (Fig. 5).

B. **Decoder**:

5 Evaluation

In this FVC store one frequent value (F1), FVC store four frequent value (F4), Multi-coding with four encoding scheme (MC). It uses combination of F1 and MC (F1MC), combination of F4 and MC (F4MC), FV-MSB method (MSB) and HFVCMC where

Table 2 Comparison table

Constrains	Existing method	Proposed method
Registers	20,433	20,383
FlipFlop	20,433	20,383
Bels	44,401	54,554
Latches	20,433	20,383
Clock buffers	3	3
IO buffers	95	95
I buffers	63	63
O buffers	32	32

Table 3 Comparison of power analysis of single core and multicore

Constrains	Single core	Multi core	Single core	Multi core	Single core	Multi core	Single core	Multi core
	50 MHz		100 MHz		200 MHz		300 MHz	
Delay report (ns)	3.377	3.362	3.377	3.362	3.377	3.362	3.377	3.362
Power report (mW)	0.749	0.855	0.948	1.161	1.347	1.774	1.747	2.389
Energy consumption	2.529	2.874	3.201	3.903	4.548	5.964	5.899	8.031

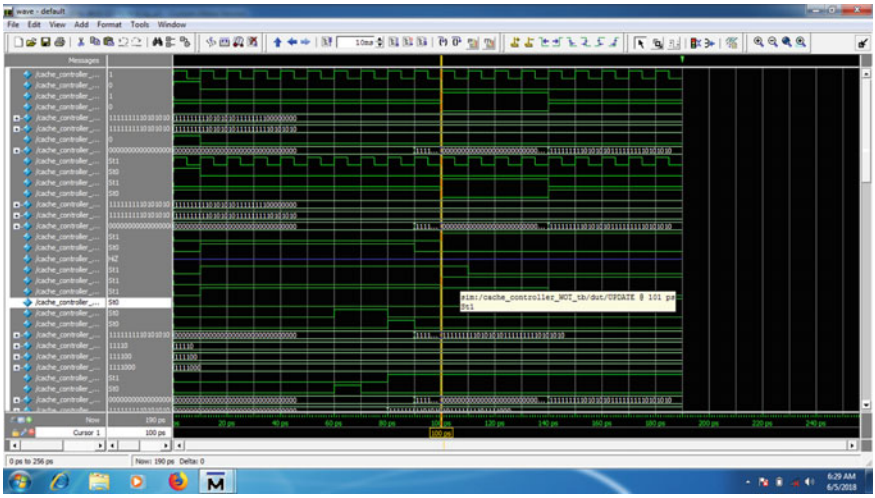
FVC and FHBC store four values of HFVCMC Many of these parameters have been varied to study their efficiency on total SA reduction.

Parameters and Comparisons of Existing and proposed Method

Power Analysis:

Table 2 shows that the proposed method uses Less number of constrains when compared to Existing method as a Result Less amount of power is consumed and memory is also Saved. Literature survey and Reference [6–9] Gives the some Important parameters of Encoder and decoder circuitry for large and complex circuit. The Results were simulated and work was Good But we simulated the work for Small circuits and Focused on minute parameters where power is wasted and not utilized. And we focus the Existing result based on fraction of wastage power. In Table 3 we focus on Delay Report and Power Report which gives power or Energy consumption with Reference papers like [8, 10–19] Gives Delay and [6, 8, 9, 20] Gives power Report In This work we Represent Power and Delay Based on Different Freuency. The Existing Report is Based on Internal parameters and fraction of wastage of power which is not Discuss on Reference papers.

Wave Forms:



6 Conclusions

In this Work Transitions count for the Data has Reduced such that it consumes less power. We Analyze the power with Different Frequency The analysis parameter like Delay Report Decreases with increasing frequency and the Power is also Decreases with increasing frequency. The Internal parameters are also in small quantity which is shown in Table 2 Hence Automatically power utilization will decrease. The energy utilization is High as the Frequency increases thus energy consumption value increases as frequency increases which indicates that maximum power is utilized with out any External disturbance like heat, light, wiring, Vibrations, and Bulk interconnections.

References

1. Zhou H, Wong DF (1998) Global routing with crosstalk constraints. In: Proceedings of 35th design automation conference, pp 374–377
2. Jafarzadeh N, Palesi M (2014) Data encoding technique for reducing energy consumption in NOC. IEEE Trans Very Large Scale Integr (VLSI) Syst
3. Naroska, Lai (2005) Low power dynamic bus encoding for deep sub-micron design. IEEE
4. Micro, Burl (1997) Low-power encoding for global communication in CMOS VLSI. IEEE
5. Wakerly. Text book Digital Ic Applications
6. Sarkar S, Rahul (2017) Adaptive bus encoding for transition reduction on off-chips buses with dynamically varying switching characteristics. IEEE Trans VLSI
7. Mingquan, Jizan, Zhihua (2016) Joint hybrid frequent value cache and multi-coding for data bus energy saving. IEEE
8. Tsai K-L, Ruan S-J (2005) Low power dynamic encoding for deep sub-micron design. IEEE

9. Chiu C-T, Huang W-C (2013) Embedded transition inversion coding with low switching activity for serial links. IEEE Trans VLSI
10. Alamgir M, Basith II, Supon T, Rashidzadeh R (2015) Improved bus shift coding for low power I/O. IEEE
11. Saini S, Mandalika SB (2011) A new bus coding technique to minimize crosstalk in VLSI. IEEE
12. Mahajan A, Mandalika SB (2010) Implementation of low power FFT structure using a method based on conditionally coded blocks. IEEE
13. Krishnan, Preethy AP. A novel techniques for low energy adaptive bus
14. Hirose K, Yasuura H (2000) A bus delay reduction technique considering crosstalk. In: Proceedings of design automation and test in Europe conference and exhibition
15. Kautz WH (1965) Fibonacci codes for synchronization control. IEEE Trans Inf Theory 11:284–292
16. Sakurai T (1993) Closed-form expressions for interconnection delay, coupling, and crosstalk in VLSI's. IEEE Trans Electron Devices 40(12):118–124
17. Tseng H-P, Scheffer L, Sechen C (1998) Timing and crosstalk driven area routing. In: Proceedings of 35th design automation conference, pp 378–381
18. Xue T, Kuh E, Wang D (1997) Post global routing crosstalk synthesis. IEEE Trans Comput-Aided Des 16(12):1418–1430
19. Abinesh, Bhargava, Suresh (2010) Transition inversion based low power data coding scheme for buffered data transfer. IEEE
20. Sravya G (2016) VLSI design of low power data encoding techniques for NOC. IEEE
21. Ojashri A, Sharma A, Saini A (2016) Quadri coding Technique to reduce self transition in VLSI interconnects. IEEE

Low-Power and High-Speed Configurable Arithmetic and Logic Unit

Naveen Kumar Kabra and Zuber M. Patel

Abstract Low power and high speed are critical design issues in the field of microprocessor design. Arithmetic and logic unit (ALU) is one of the most power and delay consuming elements of microprocessor. Conventional approach to design an ALU of microprocessor uses two different units to perform arithmetic and logical operation, respectively. Arithmetic unit is designed with adders while logical unit with logic gates. Adders are selected as per the application requirement based on VLSI matrices. In this paper, we present configurable ALU that improves performance in terms of speed and power. To achieve the objective, configurable ALU uses two adders to perform the same task. We also present a method to increase the number of logical operations. In the proposed design, eight arithmetic and eight logical operations are performed with 4-bit binary data. The proposed design is verified using Xilinx ISE 14.7 design suite and synthesized by genus synthesis solution of cadence at GDPK-45 nm technology. The proposed work offers saving up to 21.196%, up to 20.312%, and up to 6.288% in power, area, and delay, respectively.

Keywords Configurable ALU · Low power · High speed · CLA · RCA · Multiplexer

1 Introduction

An arithmetic and logic unit is combinational functional logic unit, which mainly consist of a multiplexers, adders, and logic gates [1]. Multiplexers are used to select a specific arithmetic and logic operation using select lines. Adders are used to perform arithmetic operations while logical operations are performed by logic gates.

In conventional approach of ALU design, two multiplexers is used, one for selection of arithmetic and other for logical operation. Configurable ALU based on conventional approach takes more area and power. In our design, we propose an archi-

N. K. Kabra (✉) · Z. M. Patel
Sardar Vallabhbhai National Institute of Technology Surat, Surat 395007, India
e-mail: nkkabrasvnit@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_37

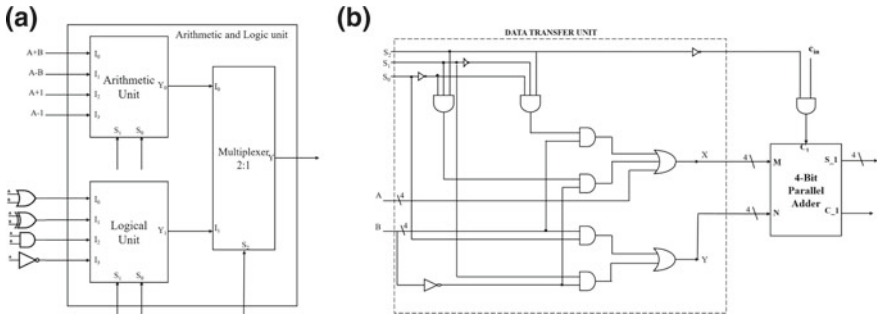


Fig. 1 **a** Block diagram representation of basic ALU architecture, **b** logical diagram representation of conventional ALU architecture [2]

tecture which uses single unit to perform the same and significantly improved all above VLSI matrices.

Furthermore, the paper is arranged as follows: the short review over basic ALU architecture along with conventional ALU architecture proposed in [2] is presented in Sect. 2. The proposed ALU architecture is presented in Sect. 3. The implementation analysis and result verifying the effectiveness of the proposed architecture are presented in Sect. 4. Finally, the paper is concluded in Sect. 5.

2 Basic ALU Architecture

Figure 1a shows block diagram representation of the basic design of ALU which consist of three multiplexers. It performs four arithmetic and four logical operations. First and second multiplexer is used to select any one of the arithmetic and logical operations with the help of select line S_1–S_0, respectively, while the third multiplexer selects the type of operation to be performed with the help of select line S_2.

To avoid the use of two different logic units, an idea was presented in [2] to perform logical operation with the help of arithmetic unit. It is done by inhabiting all input carries into the full adder circuits of parallel adder [2]. To realize logical operation with help of arithmetic unit, select line $\overline{S}_2$ is ANDed with carry generated by each stage of adder as shown in Fig. 1b.

To perform ALU operation either ripple carry adder (RCA) or carry-lookahead adder (CLA) is used as per requirement of the application. If speed is a constrain, then CLA is preferred while for area and power, RCA is used [3–6].

3 Proposed Configurable ALU Architecture

To process different signals in microprocessor and embedded systems, variety of efficient algorithms such as FFT and FIR has been developed. These algorithms are complex and characterized by data-intensive computation. For implementing all such applications, two approaches are used: software running on a general-purpose processor (GPP) and hardware in the form of application-specific integrated circuit (ASIC) [7]. General purpose processor is flexible enough to support the simple application with high performance but it is not capable to handle complexity of the application while ASIC approach can optimize the implementation in terms of power and performance but it becomes application dependent. Its direct architecture–algorithm mapping restricts the range and flexibility of applicability. Thus, numerous applications are developed such as Remarc [8], MorphoSys [9], and configurable multiplier for adaptive computing system [10] have advantages of both approaches. In spite of the above advantages, the deployment of such applications is prohibitive due to its significant area, power consumption and clock [11]. This is due to the fact that maximum of these application are composed of many computational resources such as arithmetic and logic unit (ALU) and other executing modules to perform frequent operations. ALU unit in most microprocessor is designed by using a specific adder structure such as RCA or CLA as per application requirement. Thus, ALU design using single type of adder limits its use. But if we design an ALU with both types of adder structures, we have flexibility to configure them, and then it is possible to achieve more flexibility in performance matrices with same ALU structure. Hence, the idea of configurable ALU is presented. A configurable ALU takes more area compared to ALU architecture with single adder. Therefore, the reducing area consumption must be key focus for reliability, power, and complexity [11].

The configurable ALU is a combinational logic containing both types of adder (RCA and CLA) along with a decoder and a data transfer unit. Thus, configurable ALU allows the user to dynamically configure ALU either for high-speed operation or low-speed (low power) operation. This flexibility optimizes the dynamic requirement of application where switching from low-speed/low-power to high-speed/high-power mode becomes easy.

3.1 Data Transfer Unit

Data transfer unit generates input require by an adder stage according to truth Table 2. Multiplexer is used to design the logic as shown in Fig. 2a which consist of two 2:1 multiplexers (M1 and M2), two 4:1 (M3 and M4), and one OR gate.

3.2 Decoder

Decoder is used to activate any one adder at a time. It is achieved by ANDing output of data transfer unit with output of decoder as shown in Fig. 2a. Truth Table 1 describes the operation of decoder in configurable ALU.

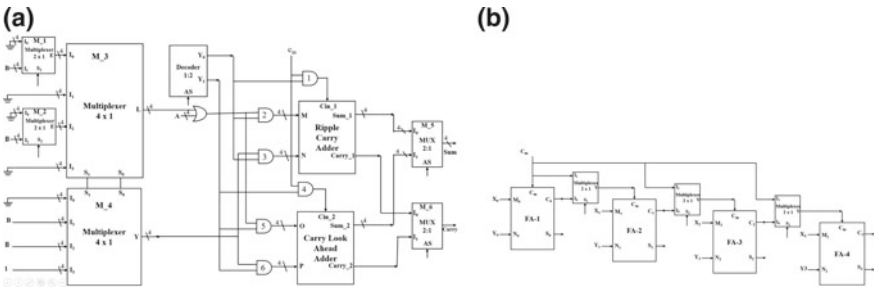


Fig. 2 **a** Architecture of the proposed configurable ALU, **b** block diagram representation of proposed 4-bit RCA

Table 1 Truth table of decoder

AS	Y_0	Y_1	Selected adder
0	1	0	RCA
1	0	1	CLA

1. When add select (AS) input of decoder is at logic 0, then the output of decoder will activate AND gate 1, 2, and 3 which allows to transfer input generated by data transfer unit along with carry input to ripple carry adder. Thus, the operation selected by select line according to Table 2 will be performed by RCA.
2. When add select (AS) input of decoder is at logic 1, then the output of decoder will activate AND gate 4, 5, and 6 which allows to transfer input generated by data transfer unit along with carry input to CLA. Thus, the operation selected by select line according to Table 2 will be performed by CLA.

3.3 Novel Approach to Increase the Number of Logical Operation

In this section, we present a novel approach to increase the number of logical operations. The idea proposed in [2] as described in Sect. 3 which performs only four logical operations. We extended this idea to include four additional logical operations such as NAND, NOR, XNOR, and transfer of A. Truth Table 2 shows arithmetic and logic operation performed by the proposed ALU architecture.

To understand condition of carry to be propagate between adder stages to perform additional four logical operations along with already available operations, we consider inputs $A = 1111[(15)_{10}]$ and $B = 1100[(12)_{10}]$ with required condition to perform them, as illustrated in Table 3. The solution is shown in Table 4 that shows required carry to be propagate between two stages to perform true operation.

Table 2 Functional table of the proposed ALU

S_2	S_1	S_0	C_{in}	Operation	Function
0	0	0	0	A	Transfer A
0	0	0	1	$A + 1$	Increment A
0	0	1	0	$A + B$	Addition
0	0	1	1	$A + B + 1$	Add with carry
0	1	0	0	$A - B - 1$	Subtract with borrow
0	1	0	1	$A - B$	Subtraction
0	1	1	0	$A - 1$	Decrement A
0	1	1	1	A	Transfer A
1	0	0	0	$A \vee B$	OR
1	0	0	1	$A \veebar B$	NOR
1	0	1	0	$A \oplus B$	XOR
1	0	1	1	$A \odot B$	XNOR
1	1	0	0	$A \wedge B$	AND
1	1	0	1	$A \barwedge B$	NAND
1	1	1	0	$\overline{A}$	Complement A
1	1	1	1	A	Transfer A

It is clear from Tables 3 and 4 that to perform the defined function, the following three key points must be taken care:

1. For arithmetic operations, carry generated by each adder stage must be propagated to next stage.
2. For logical operations (OR, XOR, AND, and Complement), carry to each stage must be equal to 0.
3. For four additional logical operations (NOR, XNOR, NAND, and Transfer A), carry to each stage must be equal to 1.

Now with the help of truth Table 2 and point mentioned, it can be seen that condition to perform the logical operation is similar to initial input carry provided to the first stage of 4-bit adder along with select line $S_2 = 1$ while for arithmetic operation, the condition is same as presented in [2] along with $S_2 = 0$. After considering all facts, the logic required to perform defined task is presented in truth Table 5 and it can be implemented by connecting a 2:1 multiplexer between two adder stages. Thus, adder unit of the proposed ALU architecture is designed by introducing a 2:1 multiplexer between two full adder units. The first input to each multiplexer unit is carry generated by the previous stage while the second input will fixed to external carry input or initial carry input as shown in Fig. 2b.

To compare conventional ALU architecture with the proposed configurable ALU architecture in terms of VLSI matrices, conventional ALU architecture [2] described in Sect. 2 is extended to configurable level as shown in Fig. 3a. As conventional ALU

Table 3 Illustration of different condition for arithmetic and logical operation

Operation	Condition of select line	Condition of input carry (C_{in})	Required input		Excepted answer
			X	Y_i	
Add with carry	001	1	A_i	B_i	11100
			1111	1100	
Subtract with borrow	010	0	A_i	$\overline{B}_i$	10010
			1111	0011	
OR	100	0	$A_i + Bi$	0	01111
			1111	0000	
NOR	100	1	$A_i + B_i$	0	10000
			1111	0000	
XNOR	101	1	A_i	B_i	11100
			1111	1100	
NAND	110	1	$A_i + \overline{B}_i$	$\overline{B}_i$	10011
			1111	0011	
Transfer A	111	1	A_i	1	11111
			1111	1111	

Table 4 Illustration of different condition for arithmetic and logical operation

	Add with carry	Subtract with borrow	OR	NOR	XNOR	NAND	Transfer A
X_i	1111	1111	1111	1111	1111	1111	1111
Y_i	1100	0011	0000	0000	1100	0011	1111
C_{in}	1	0	0	1	1	1	1
Carry to be propagate	111–	111–	000–	111–	111–	111–	111–
Final output	11100	10010	01111	10000	10011	10011	11111
Final carry	1	1	0	1	1	1	1

Table 5 Truth table of logic to introduce additional logical operation

S_2	Input carry to each adder stage
0	Propagate carry generated
1	C_{in}

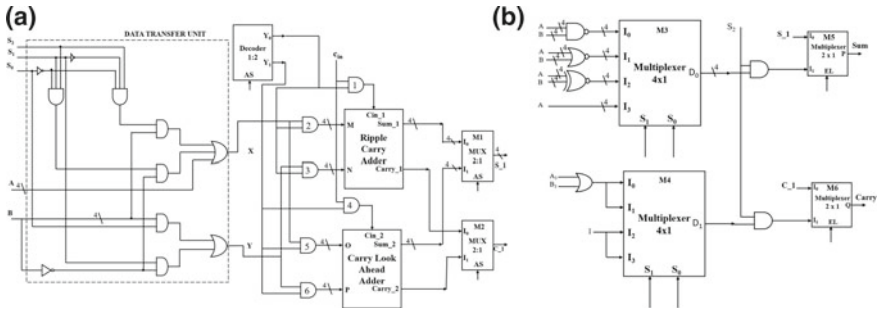


Fig. 3 **a** Configurable ALU based on conventional architecture, **b** extra logic for implementing additional four functions

Table 6 Truth table to define addition four logic functions in conventional configurable ALU architecture

S_2	S_1	S_0	C_{in}	Logical operation	Carry to be generated
1	0	0	0	NOR	1 when $A3$ or $B3 = 1$
1	0	1	0	XNOR	1 when $A3$ or $B3 = 1$
1	1	0	0	NAND	1
1	1	1	0	Transfer A	1

supports only four logical operations while the proposed configurable ALU support eight logical operations, thus an extra logic unit has been designed by considering same logical condition used for the proposed configurable ALU architecture as presented in Table 6 and implemented using two 4:1 multiplexers as shown in Fig. 3b. Multiplexer (M3) selects logical operation while multiplexer (M4) selects respective carry condition with help of select line S_1 and S_0. Output of these multiplexers is ANDed with S_2 to propagate them only when S_2 is equal to 1. Now, the output available from multiplexers M1, M2 in Fig. 3a and output generated from M3, M4 will be connected to 2:1 multiplexers M5 and M6, respectively, to generate final sum and carry outputs using select line named "EL". When EL is at logic 0, then final sum and carry output will be generated by configurable conventional ALU unit according to Table 1. When EL is at logic 1, then final sum and carry outputs is generated by extra logic unit according to Table 6.

4 ASIC Implementation Results

The performance parameter analysis of proposed configurable ALU architecture has been presented and compares it with configurable ALU architecture based on [2]. VHDL code of both configurable ALU architectures are verified on Xilinx ISE 14.7 design suite and synthesized using genus synthesis solution of cadence at GDPK

Table 7 ASIC performance parameter of configurable ALU

Parameter	Proposed ALU architecture	Configurable ALU based on [2]
Area (no. of cells)	153	192
Delay (ps)	1848	1972
Static power (nW)	3.069	3.572
Dynamic power (nW)	2421.948	3073.713
Total power (nW)	2425.017	3077.284

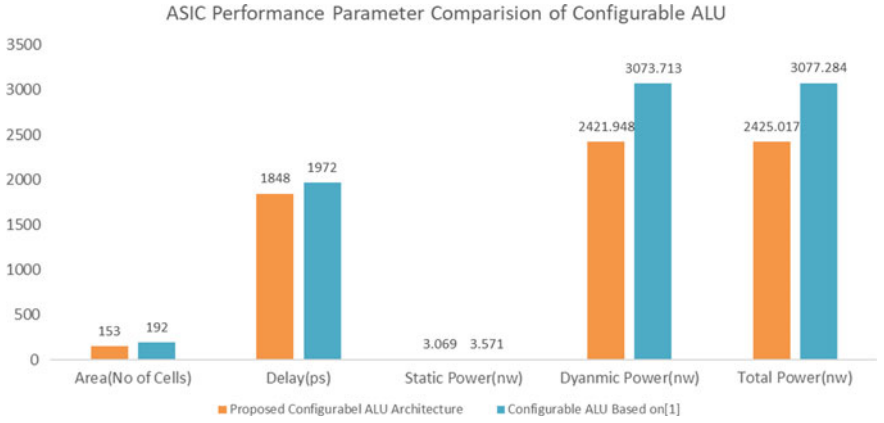


Fig. 4 Graphical representation of performance parameter of configurable ALU

45 nm technology with slow mode operation at operating temperature of 0°–125° and supply voltage of 0.9 V. The synthesized result is presented in Table 7 while Fig. 4 shows its graphical representation. It can be seen from Table 7 that the proposed configurable ALU offers saving up to 21.196%, up to 20.312%, and up to 6.288% in power, area, and delay, respectively.

5 Conclusion

A novel approach has been presented for designing configurable ALU to perform logical operations using adder. The design also increases the number of logical operations as compared to conventional architecture. The proposed ALU is designed with RCA and CLA adders that give flexibility in choosing the speed of operation. The proposed configurable ALU consume 652.267 nW less power, uses 39 number of less standard cells for design, and takes 1848 ps worst delay which is 124 ps less with respect to conventional ALU presented. The proposed ALU can be used in both low-power and high-speed applications.

References

1. Parhami B (2010) Computer arithmetic: algorithms and hardware designs. Oxford series in electrical and computer engineering, 2nd edn. Oxford University Press
2. Mano MM (1995) Digital logic and computer design. PHI
3. Brent RP, Kung HT (1982) IEEE Trans Comput C-31(3):260. https://doi.org/10.1109/tc.1982.1675982
4. Ngai TF, Irwin MJ (1985) In: 1985 IEEE 7th symposium on computer arithmetic (ARITH), pp 9–15. https://doi.org/10.1109/arith.1985.6158967
5. Wei BWY, Thompson CD (1990) IEEE Trans Comput 39(5):666. https://doi.org/10.1109/12.53579
6. Tyagi A (1993) IEEE Trans Comput 42(10):1163. https://doi.org/10.1109/12.257703
7. Feng C, Yang L (2013) In: Proceedings 2013 international conference on mechatronic sciences, electric engineering and computer (MEC), pp 2286–2290. https://doi.org/10.1109/mec.2013.6885424
8. Miyamori T, Olukotun U (1998) In: Proceedings. IEEE symposium on FPGAs for custom computing machines (Cat. No. 98TB100251), pp 2–11. https://doi.org/10.1109/fpga.1998.707876
9. Singh H, Lee MH, Lu G, Kurdahi FJ, Bagherzadeh N, Filho EMC (2000) IEEE Trans Comput 49(5):465. https://doi.org/10.1109/12.859540
10. Pfänder OA, Pfleiderer HJ, Lachowicz SW (2006) Adv Radio Sci 4:231. https://doi.org/10.5194/ars-4-231-2006. https://www.adv-radio-sci.net/4/231/2006/
11. Ramachandran Vaidyanathan JT (2004) Dynamic reconfiguration: architectures and algorithms. Series in Computer science, 1st edn. Springer

FPGA Implementation of Speculative Prefix Accumulation-Driven RNS for High-Performance FIR Filter

G. Reddy Hemantha, S. Varadarajan and M. N. Giri Prasad

Abstract In this paper, we present speculation-driven prefix topology-based finite impulse response (FIR) filters design using residue number system (RNS) arithmetic. The core objective of our proposed framework is to maximize the prefix accumulation in the application of RNS to the design high-performance FIR filter design. To achieve this, we propose a RAM-based reverse conversion model followed by accumulation to produce the modular multiplication. The proposed RNS design makes use of block RAMs available in FPGA devices and appropriate moduli sets in order to accommodate FIR convolution results. The proposed approach is formulated to design precomputed reverse converters for different moduli sets and to implement FPGA as target devices. As a result, we propose speculative parallel prefix topology-based post accumulation technique for RNS-based multiplication, along with a high-performance FIR filter architecture that employs independent modulo channel RNS arithmetic. Experiment results of RNS-FIR design over different number of FIR taps and input operand word lengths alongside with appropriate moduli sets that suits to accommodate FIR end results.

Keywords Parallel prefix adder · FIR filter · RNS system · Speculation

G. Reddy Hemantha (✉)
Department of ECE, Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India
e-mail: hemanthag75@gmail.com

S. Varadarajan
Department of ECE, SVU College of Engineering, S V University, Tirupati, India
e-mail: varadasouri@gmail.com

M. N. Giri Prasad
Department of ECE, JNTU College of Engineering, Jawaharlal Nehru Technological University Ananthapur, Ananthapuramu, India
e-mail: mahendragiri1960@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_38

1 Introduction

In recent year's, path delay optimization in digital signal processing is emerged due to the inventions made with mobile devices and its integration in compatible digital systems demands are efficient functional units [1]. In order to accomplish both area efficiency and path delay optimization, residue number system have been investigated in many works since input operand bit size decomposed several independent short integers to carry out arithmetic operations. Critical path delay optimization can be easily tackled by reducing carry propagation and bit size reduction is directly related to the complexity reduction. In particular, the RNS has several advantageous while implementing it in high-speed FIR filter design.

Though RNS arithmetic is potentially useful for multiply and accumulation (MAC)-dominated applications such as fast Fourier transform [2], digital image processing [3], and wireless communications [4] still, it has limitations in terms of computations complexity in conversion process [5].

Several research has proved that implementing parallel prefix computation in residue number system (RNS) can give considerable path delay and complexity reduction.

2 Related Works

In many existing works RNS based arithemtic unit offers potential metrics for DSP applications such as high speed and area efficiency. FIR filter design is among the multiply intensive DSP blocks and is most widely used as effective methodologies in speech and signal processing. Due to its inherent advantageous like parallel modulo channel and carry propagation less accumulation, RNS algorithm has been investigated in many high-speed FIR filter design. In most cases, delay is considered as a critical parameter that affects the overall system performance of the FIR filter. Though several techniques have been proposed for delay optimization in FIR filter design arithmetic level modifications only give better performances in most cases. It has been incorporated in many ways still, it has the inability in terms of latency problems associated with pipeline stages.

To overcome the limitations caused by pipelining stages sublevel parallel processing driven digital system architecture using RNS algorithm is incorporated which can sustain with significant critical path reductions. While to maintain sustainability's over FIR tap length, prefix topology are formulated according to the delay of the carry chain. Besides, the prefix-based accumulation over other methods, unified models are required to suppress dominant latency cause. In [6], power dissipation is greatly reduced by taking advantage of the high performance due to the prominent parallelism of the RNS system. In [7] developed the extension of the same methodology for higher order FIR filter to mitigate the path delay problems associated with FIR filter length. To overcome the limitations of dynamic range constrain related to

filter length hybrid, moduli sets is presented to implement the modular arithmetic. In [8] used LUT design for RNS modular multiplication and its trade-off over operating speed is measured since in most cases, reverse conversion process degrade the overall RNS system. Improved FIR architecture is presented in [9] by using small moduli sets and appropriate selection of word length sizes. In [10], RNS system is used for variable FIR filter design using least mean squares (LMS) algorithm and coefficient updation is accomplished by binary serial implementation of the moduli set conversion. Common sub expression elimination (CSE) based distributed arithmetic technique is presented [11] and its limitations over critical path rise is solved using RNS arithmetic to achieve both low cost and high performance. Here, both hardware sharing and minimization of adder requirements among taps results with explicit control over critical path depth. In [12], modular, computational parallelism, and distributive properties of RNS is investigated in cryptographic algorithms and its data path and word length optimization are also proved.

In this paper, we propose RNS system with speculative prefix adder for FIR filter design which has the following advantages.

(1) Direct-coupled single-ended accumulation models are used for FIR filter convolution.
(2) It can be easily scaled down to configure any moduli sets to solve dynamic range constrains related to RNS system.
(3) High hardware utilization rate.
(4) Speculative-driven prefix computation can improve the speed of accumulation during reverse conversion.

In addition, to latency reduction, worst-case time-bound analysis-driven assertion mode activation techniques are exploited for appropriate high-speed filter convolution process. This paper exploits additional onboard block RAMs available on FPGA devices to implement reverse conversion process. Here, FIR-RNS design architecture is framed based on the target FPGA technology, speculative prefix accumulation, and moduli set required to accommodate FIR filter output dynamic ranges.

3 Proposed System

3.1 Prefix Topology

Here, the core objective of high-speed parallel architectures is accomplished using prefix adders with reduced critical path which employs a Han-Carlson prefix topology with carry generation using single Brent–Kung in both pre- and post processing unit. As we discussed in the literature survey, the overall RNS system performance can be improved using appropriate reverse converter design. Here, hierarchical prefix accumulation is forwarded to carry out RNS to binary conversion with reduced propagation delay.

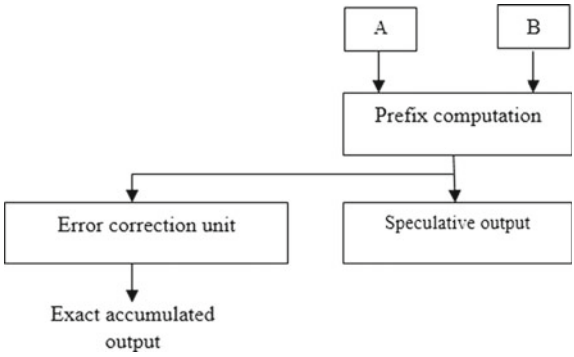


Fig. 1 Proposed speculative parallel architecture

3.1.1 Speculative Prefix Processing

The speculative computation in prefix accumulation model has advantages in optimizing the latency and accumulation speed compared to conventional prefix architectures. Here, approximated values are generated and propagated to the most significant subblock regions; in the post processing if any of the approximation is not matched with end results which can be asserted through carries obtained from LSB block side, appropriate error correction unit is triggered to produce corrected results. This speculation causing significant improvement in critical path delay reduction since the error correction unit is constructed with hierarchical levels of the XOR gates as shown in Fig. 1.

3.2 RNS System

Residue number system is evaluated by a moduli set, which consists of some pairwise relatively prime integers $\{m0, m1, \ldots, mn-1\}$ as moduli's. The dynamic range that can be covered by RNS number system entirely depends on moduli set range and number of moduli's used, which is defined by the product of all moduli in the moduli set. A residue number system with dynamic range M can be used to represent the numbers in the range of $[0, M-1]$. Consequently, the data path propagation of RNS system is accomplished as series of additions in L parallel and independent channels according to the residues. The proposed prefix topology based RNS system is shown in Fig. 2.

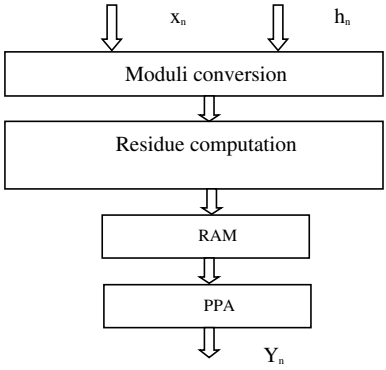


Fig. 2 RNS computation model

3.2.1 Moduli Conversion

Given a set of pair-wise relatively prime moduli $\{m1, m2, \ldots mn\}$ and a residue representation $\{r1, r2, \ldots rn\}$ in that system of some number X, i.e., $ri = |X|mi$ that number and its residues are related by the following equation:

$$|X|M = |\sum_{i=1}^{n} ri|Mi - 1|mi\ Mi|M \tag{1}$$

where M is the product of the mi's, and $Mi = M/mi$. If the values involved are constrained so that the final value X of is within the dynamic range, then the modular reduction on the left-hand side can be omitted.

To understand the formulation of equation, we rewrite X as

$$\begin{aligned} X &\triangleq \{r1, r2, \ldots, rn\} \triangleq \{r1, 0, \ldots 0\} + \{0, r2, \ldots, 0\} + \{0, 0, \ldots, rn\} \\ &\triangleq X1 + X2 + \cdots + Xn \end{aligned} \tag{2}$$

Hence, the reverse conversion process requires finding Xi's. The operation of obtaining each Xi is a reverse conversion process by itself. However, it is much easier than obtaining X.

Consider now that we want to obtain Xi from $\{0, 0 \ldots, ri, \ldots, 0, 0\}$. Since the residues of Xi are zeros except for ri. This dictates that Xi is a multiple of mj where $j \neq i$. Therefore, Xi can be expressed as

$$Xi \triangleq r1 * \{0, 0, \ldots 1, \ldots, 0, 0\} \triangleq ri * Xi \tag{3}$$

where X_i is found such that $|X_i|_{mi} = 1$. We recall from equation that the relation between the number ri and its inverse $ri - 1$ is as follows:

$$(ri\ X\ ri - 1) \bmod mi = 1 \tag{4}$$

We define Mi as M/Mi, where $M = \prod_{i=1}^{k} pi$. Then

$$||Mi - 1|miMi|mi = 1 \tag{5}$$

Since all mi's are relatively prime, the inverses exist

$$Xi = \left|\overline{Mi} - 1\right|\overline{mi}Mi \tag{6}$$

$$Xi = riXi = ri|Mi - 1|miMi \tag{7}$$

$$X = \sum_{i=1}^{n} Xi = \sum_{i=1}^{n} ri|Mi - 1|miMi \tag{8}$$

To ensure that the final value is within the dynamic range, modulo reduction has to be added to both sides of the equation.

3.2.2 Block RAM Based Reverse Conversion

In general reverse converter design, it depends on the bit size of the moduli set to implement the RNS system. For each moduli sets, dedicated reverse converter designs are formulated that best fit the required dynamic ranges. Here, precomputed end results are stored in memory and used as a basic building block for reverse conversion process. During hardware synthesis, this design methodology can be easily optimized using onboard block RAMs or LUTs which is presented in any FPGA devices. The performance of ROM-based reverse converter implementations in RNS modulo arithmetic improve the system performance with the minimized critical path as shown in Fig. 3.

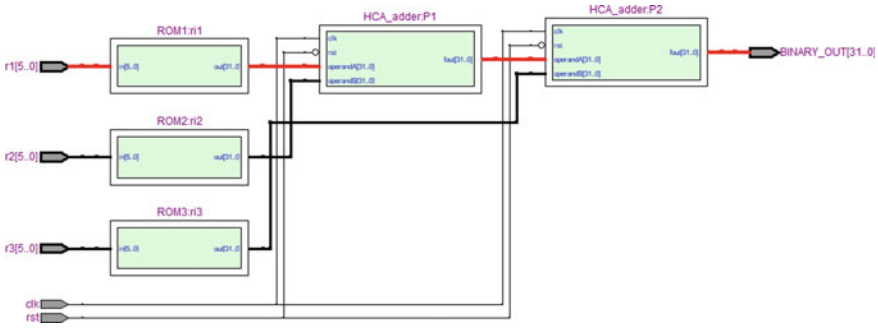


Fig. 3 LUT-based RNS reverse converter design

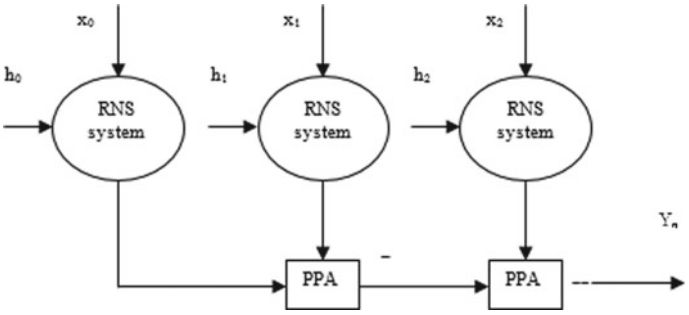


Fig. 4 Prefix topology-driven RNS-FIR architecture

3.3 FIR Filter Based on RNS

Filter realizations using residue number system (RNS) is suitable for implementation of high-speed digital signal processing due to their inherent parallelism, modularity, fault toleration, and local carry propagation properties. Arithmetic operations like multiplication and addition are carried out more efficiently as RNS with localized carry propagation properties of prefix adders as shown in Fig. 4. Though RNS is particularly more suitable for implementing FIR low cost filters, inclusion of speculation driven topology make this system beneficial for perfomance as well, particularly, when large word length and high-throughput rate are demands.

4 Experimental Results

Experimental evaluation was performed for the proposed RNS-FIR design with memory-enabled reverse converters for moduli sets $\{2^n - 1, 2^n + 1, 2^n - 1\}$ [5] with $n = 3, 4$, and 5. All the design modules were described in Verilog HDL and its functionality is verified using ModelSim through exhaustive test bench input stimulus. RNS design specifications and FIR filter designs were synthesized FPGA QUARTUS II EDA tool tailored for the ALTERA CYCLONE III logic family devices. Hardware complexity is evaluated as the number of logical element utilised for device modeling as shown in Fig. 5 and its post map view is shown in Fig. 6.

4.1 Trade-off Measure Moduli Sets Dynamic Range Versus RNS Reverse Conversion

The architecture explained shown in Sect. 3 is consistent with the dynamic word length variations. However, to optimize operand, bit size each of the RAM sizes is

Flow Summary

Flow Status	Successful - Sat Jun 16 02:47:54 2018
Quartus II Version	9.0 Build 132 02/25/2009 SJ Web Edition
Revision Name	TOP
Top-level Entity Name	FIRFILTER
Family	Cyclone III
Device	EP3C16F484C6
Timing Models	Final
Met timing requirements	N/A
Total logic elements	4,281 / 15,408 (28 %)
Total combinational functions	4,231 / 15,408 (27 %)
Dedicated logic registers	522 / 15,408 (3 %)
Total registers	522
Total pins	26 / 347 (7 %)
Total virtual pins	0
Total memory bits	0 / 516,096 (0 %)
Embedded Multiplier 9-bit elements	0 / 112 (0 %)
Total PLLs	0 / 4 (0 %)

Fig. 5 Resource utilization summary

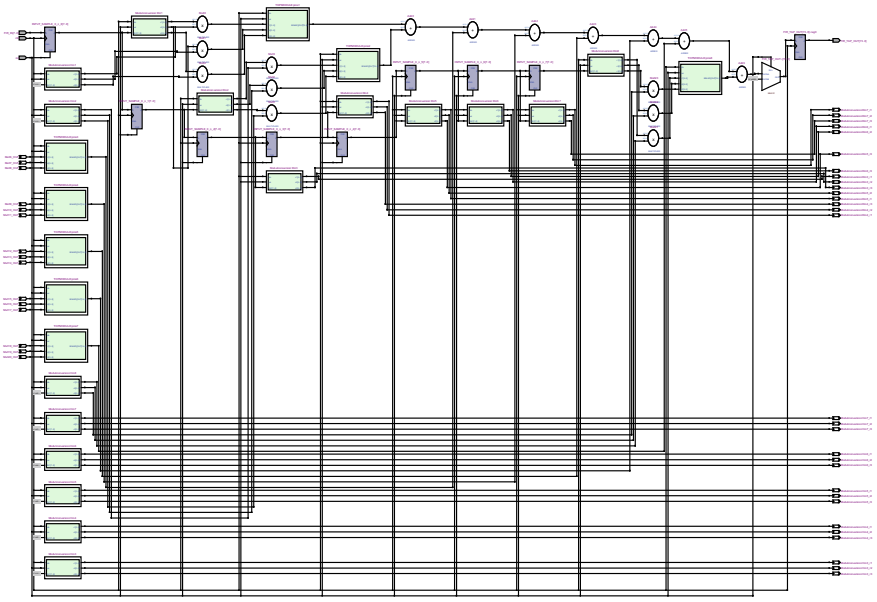


Fig. 6 RTL map viewer summary

resized based on their hierarchical moduli information and linearization is performed over aspect ratios. The aspect ratios are calculated analytically as described in the previous section to model reverse conversion process. The aspect ratios of given moduli sets are selected to be close to the moduli set values. As shown in Table 1, the proposed method is outperformed existing booth encoded RNS model both in 8 bits and 16 bits word length.

The growth of the RAM size with moduli set bit size is exponential, which is convenient with respect to the attainable frequency response as shown in Table 2. This table shows the FPGA synthesis results in terms of number of LEs and delay measures for different values of 4, 5, and 6 bits moduli with the parallel modulo FIR implementation. The maximum frequency of the 8 bits word length is obtained as 57.3 MHz, while for the 12 and 16 bits word length, the maximum frequency of the filter is limited by the accumulation of reverse conversion results.

4.2 Critical Path Retention Performance Measure

Here at the expense of least complexity overhead and performance loss measures, the possible higher order FIR filter design is achieved as compared to other conventional RNS-FIR models. The path delay retention in FIR MAC networks during filter convolution is accomplished using assertion driven prefix networks which act as delay optimization model. In this way, the overall time required to formulate the convolution output is optimized. Note that, filter length is needed to be large enough to preserve most of the finite filter coefficients. Thus, the computational complexity to execute the higher order length is not high. After the parameter evaluations, it is tested on various FIR tap length and its efficiency in path delay reduction is proved as shown in Table 3.

Table 1 State-of-the-art comparison booth encoded RNS versus speculative prefix accumulated RNS multiplier

Input word length size	Bit size	Area (CLBs)	*F*max (MHz)
Booth encoded RNS [13]	8 bit	82	60.7
		68	79.2
Proposed method	16 bit	150	48.49
		118	56.23

Table 2 Trade-off measures of word length optimization over performance efficiency

Input word length size	Moduli set ($2^n + 1$, 2^n, $2^n - 1$)	Area (LEs)	*F*max (MHz)
8 bit	(7, 8, 9)	4281	57.3
12 bit	(15, 16, 17)	4492	40.29
16 bit	(31, 32, 33)	12,165	29.78

Table 3 Performance analyzes of speculative prefix RNS-FIR design

FIR length	Area (LEs)	*F*max (MHz)
4 tap	2096	63.46
8 tap	4281	57.6
16 tap	8623	57.3

5 Conclusion

In this paper, we presented a new speculative adder component to reduce the critical path and RAM-based reverse converter structures for RNS system which can be able to provide prominent end results with FPGA device as target technology. This prefix adder reduces the latency and can be easily extended to any moduli sets, simply by placing appropriate end modulo values for reverse conversion. Experimental results proved that the proposed speculative prefix topology-driven RNS-FIR filter design imposes negligible hardware complexity and performance degradations to the input word length constrains and FIR tap length, making it quite prominent FIR design to any FPGA-based signal processing applications. This work explores significant performance retention with increased FIR tap length by bringing the use of RNSs and memory efficient reverse conversion design.

References

1. Athas WC, Svensson LJ, Koller JG, Tzartzanis N, Chou EYC (1994) Low-power digital systems based on adiabatic-switching principles. IEEE Trans Very Large Scale Integr (VLSI) Syst 2(4):398–407
2. Nussbaumer HJ (2012) Fast Fourier transform and convolution algorithms, vol 2. Springer Science & Business Media
3. Wang W, Swamy MNS, Ahmad MO (2004) RNS application for digital image processing. In: Proceedings of 4th IEEE international workshop on system-on-chip for real-time applications. IEEE, pp 77–80
4. Chen J, Lv T, Zheng H (2004) Cross-layer design for QoS wireless communications. In: Proceedings of the 2004 international symposium on circuits and systems, ISCAS'04, vol 2. IEEE, pp II-217
5. Mohan PVA (2007) RNS-to-binary converter for a new three-moduli set $\{2^{n+1} - 1, 2^n, 2^n - 1\}$. IEEE Trans Circuits Syst II: Express Briefs 54(9):775–779
6. Bhardwaj M, Balaram A (1998) Low power signal processing architectures using residue arithmetic. In: Proceedings of IEEE international conference on acoustics, speech and signal processing (ASSP'98), vol 5, pp 3017–3020
7. Cardarilli GC, Nannarelli A, Re M (2000) Reducing power dissipation in FIR filters using the residue number system. In: Proceedings of the 43rd IEEE Midwest symposium on circuits and systems, vol 1. IEEE, pp 320–323
8. Liu Y, Lai EMK (2004) Moduli set selection and cost estimation for RNS-based FIR filter and filter bank design. Des Autom Embed Syst 9(2):123–139
9. Conway R, Nelson J (2004) Improved RNS FIR filter architectures. IEEE Trans Circuits Syst II Express Briefs 51(1):26–28

10. Bernocchi GL, Cardarilli GC, Del Re A, Nannarelli A, Re M (2007) Low-power adaptive filter based on RNS components. In: IEEE international symposium on circuits and systems, ISCAS 2007. IEEE, pp 3211–3214
11. Patronik P, Berezowski K, Piestrak SJ, Biernat J, Shrivastava A (2011) Fast and energy-efficient constant-coefficient FIR filters using residue number system. In: 2011 international symposium on low power electronics and design (ISLPED). IEEE, pp 385–390
12. Chang CH, Molahosseini AS, Zarandi AAE, Tay TF (2015) Residue number systems: a new paradigm to datapath optimization for low-power and high-performance digital signal processing applications. IEEE Circuits Syst Mag 15(4):26–44
13. Goel S, Bajaj SG, Kaur AG (2016) FPGA Implementation of booth encoded multi-modulus $\{2^n - 1, 2n, 2n + 1\}$ RNS multiplier

Energy-Efficient SRAM Cell Design with Body Biasing

P. Kalyani, M. Madhavi Latha and P. Chandra Sekhar

Abstract This paper presents the need of high-speed SRAM memory operation for minimum supply voltage. The reduction of power dissipation in memories is becoming primary importance in subthreshold region. There are several power reduction techniques which can be applied to SRAM memory cell to design low-power and energy-efficient memory. SRAM cell is designed with body biasing technique and used in the memory array to improve performance. The 32×32 SRAM memory array is implemented using cadence 45 nm technology. Simulation and analysis results are compared with conventional array for better performance.

Keywords SRAM · Subthreshold memory · Body biasing · PVT · Static noise margin · Leakage current · Peripherals

1 Introduction

In system-on-chip memory is the main submodule which faces process variation problem as its technology is single digit nanometer regime and works in subthreshold region [1]. Technology and supply voltage scales in proportion to meet design criteria. Operating SRAM cell in subthreshold region [2] reduces power consumption but has to resolve problems like PVT variations, stability, and large variations in static noise margin [3]. Operating currents are smaller for lower voltage which causes delay problems and in addition to that process variations still degrades the circuit performs. Circuits operating at high temperature have more leakage. Several techniques are

P. Kalyani (✉) · M. Madhavi Latha
Electronics and Communication Engineering, JNTUH, Kukatpally, Hyderabad, India
e-mail: kalyani.pala06@gmail.com

M. Madhavi Latha
e-mail: mlmakkena@yahoo.com

P. Chandra Sekhar
Electronics and Communication Engineering, Osmania University, Hyderabad, India
e-mail: sekharpaidimarry@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_39

introduced to reduce leakage current [4]. Power gating is effective technique to reduce logic core standby leakage. Body biasing is an effective technique to solve the problems when supply voltage is below the threshold voltage [5, 6].

SRAM cells are designed with different transistors but six transistor structures are most suggested to design large memories. Six transistor SRAM cells consist of two cross-coupled inverters and two NMOS access transistors. Every memory cell performs three basic operations like write, read, and holding of single bit data. In Fig. 1, M1, M2 and M3, M4 are two inverter pairs forming bistable latch, M5 and M6 are two NMOS access transistors [7, 8].

Process variation reduces memory cell read and write stability in subthreshold region. Various techniques are there in literature to improve cell performance and stability. Transistor sizing is also one method to improve stability of operation. The 8T SRAM cell has separate write and read paths as shown in Fig. 2 to improve performance and stability for minimum supply voltage [9–11].

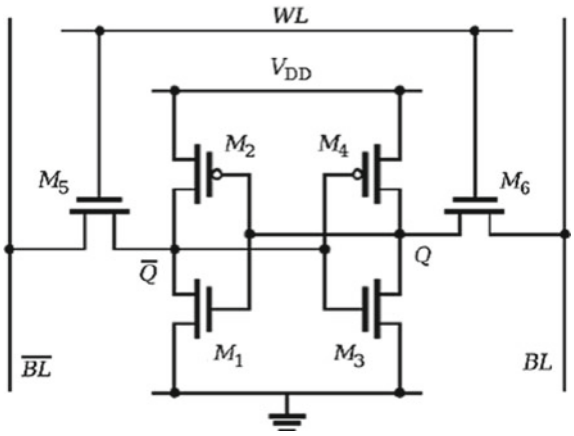


Fig. 1 6T SRAM cell design

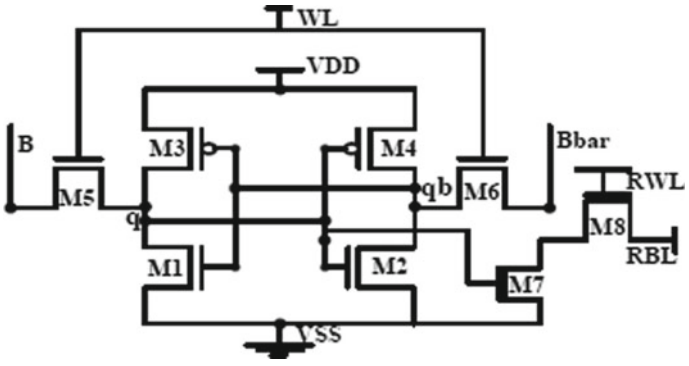


Fig. 2 8T SRAM cell design

2 SRAM Cell Design Using Cadence Tool

A conventional SRAM cell to store single bit is designed with supply voltage 1 and 0.4 V using cadence 45 nm technology and tabulated the power values for three different operations in Tables 1 [12–15] (Figs. 3, 4 and 5).

3 SRAM Cell with Body Biasing

Threshold voltage can be controlled using substrate voltages. It is decreased by decreasing substrate voltage and vice versa. Forward body bias reduces the threshold

Table 1 Power analysis results of conventional 6T SRAM cell with $Vdd = 1$ V

Technology	Mode	Conventional 6T cell
45 nm	Write	85.97 nW
	Read	86.80 nW
	Hold	14.20 pW

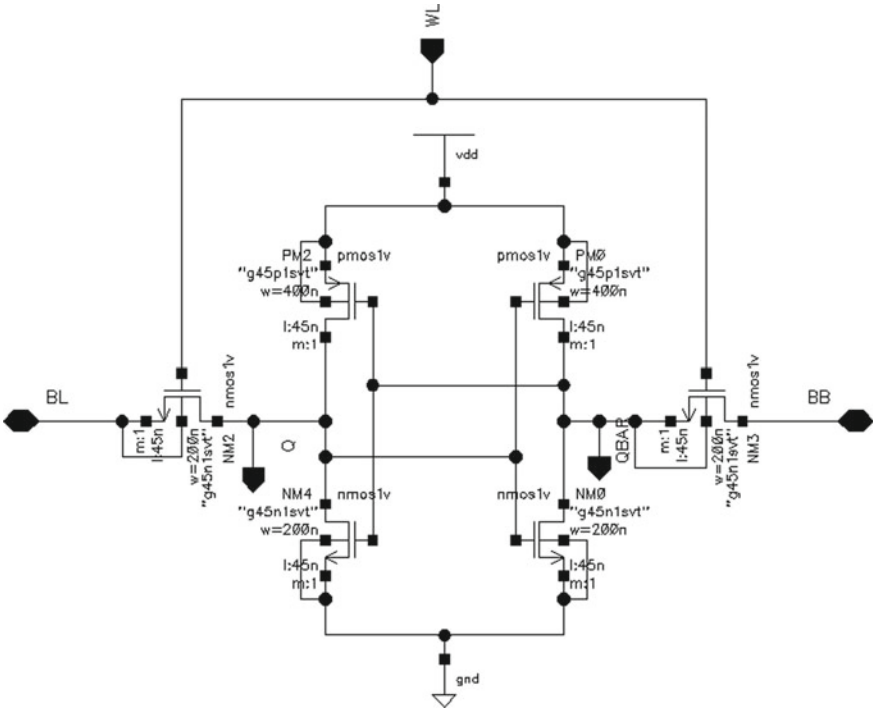


Fig. 3 SRAM 6T Schematic

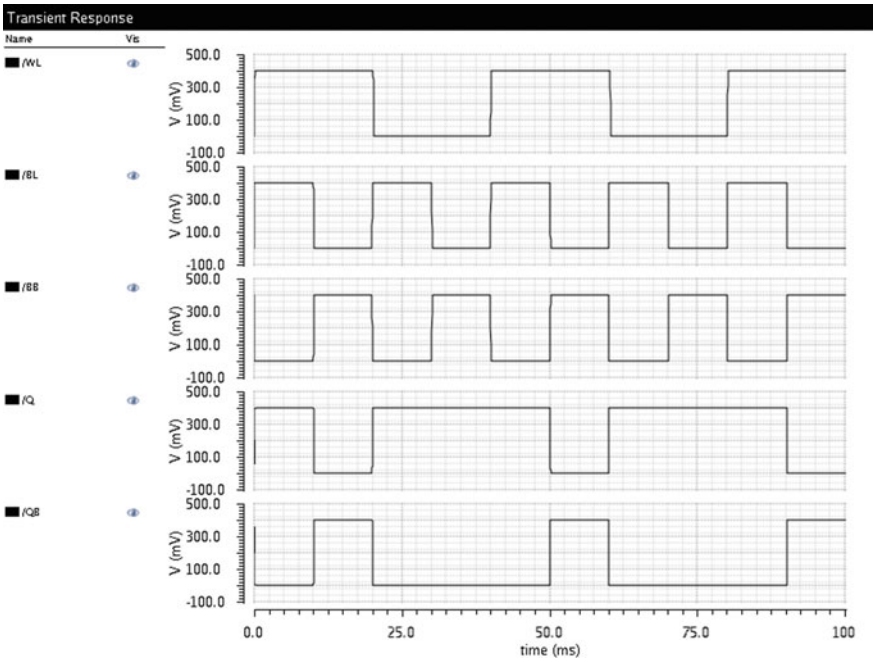


Fig. 4 Simulation waveforms of write operation

Table 2 Device substrate voltage with respect to bias

Device	Substrate bias type	
	FBB	RBB
PMOS	Negative	Positive
NMOS	Positive	Negative

voltage, whereas reverse body bias increases. As discussed in Table 2 type of body voltage depends on biasing method [16, 17].

Leakage current is combination of total six currents but in SRAM cell subthreshold current, gate leakage, and junction band-to-band tunneling are three main sources of leakage. It can be influenced by many factors threshold voltage variation is one among them. Leakage current and forward biasing of drain to substrate determines FBB voltage, whereas junction leakage current and breakdown point decide RBB voltage.

As in Table 3, NMOS (PMOS) transistors operate in ON state, their threshold voltage is decreased by applying the $Vss + \Delta V$ ($Vdd - \Delta V$) voltage to their bodies (FBB). Conversely, when NMOS (PMOS) transistors operate in OFF state, applying the $Vss - \Delta V$ ($Vdd + \Delta V$) voltage to their bodies would increase the threshold voltage (RBB). Bias control generator circuit adjusts the threshold voltage based on the device operating condition.

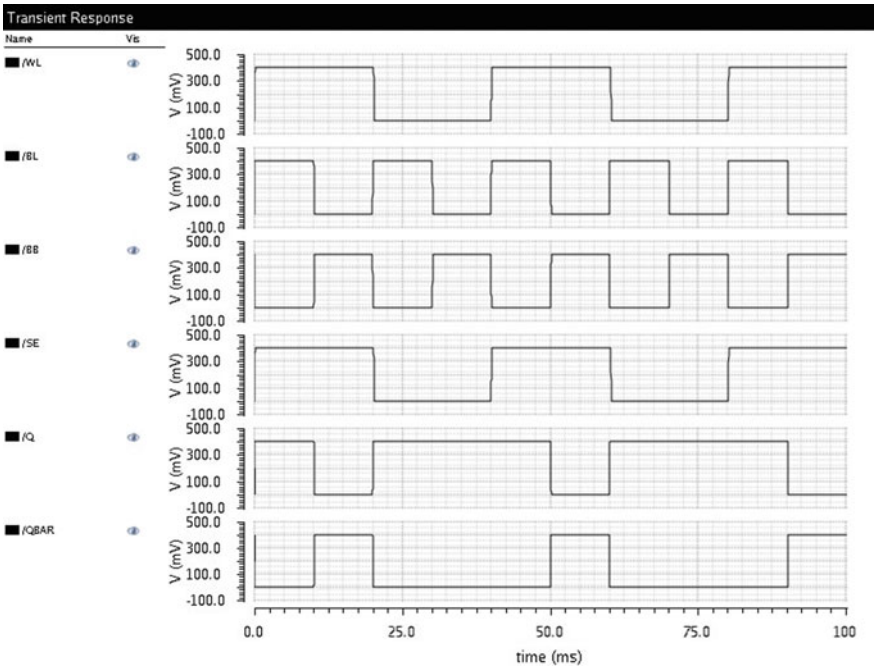


Fig. 5 Simulation waveforms of read operation

Table 3 Substrate voltage with respect to input

Device	Substrate bias type	
	Low	High
PMOS	$Vdd + \Delta V$	$Vdd\text{-}\Delta V$
NMOS	$Vss + \Delta V$	$Vss\text{-}\Delta V$

Table 4 Power results analysis of SRAM 6T with and without body biasing for $Vdd = 0.4$ V

SRAM 6T	Without body bias	With body bias
Avg. power (pW)	3.57	2.797

As shown in Fig. 6, SRAM cell is designed with fixed bias voltages and bias such that ON device is FBB and OFF device is RBB to improve overall performance (Fig. 7 and Table 4).

$$\text{WSNM for } L = 45 \text{ nm}, \quad Vdd = 400 \text{ mv WSNM} = 60 \text{ mv}$$

All the required memory peripherals are designed to implement 32×32 SRAM memory array as shown in Fig. 8 with and without body-biased individual memory bit cells [18, 19].

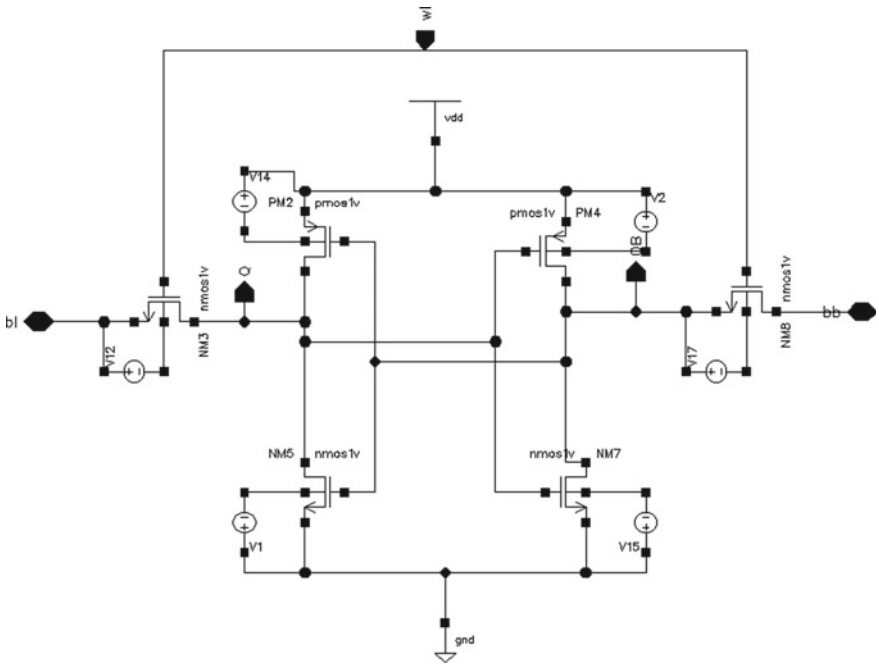


Fig. 6 SRAM 6T with body biasing

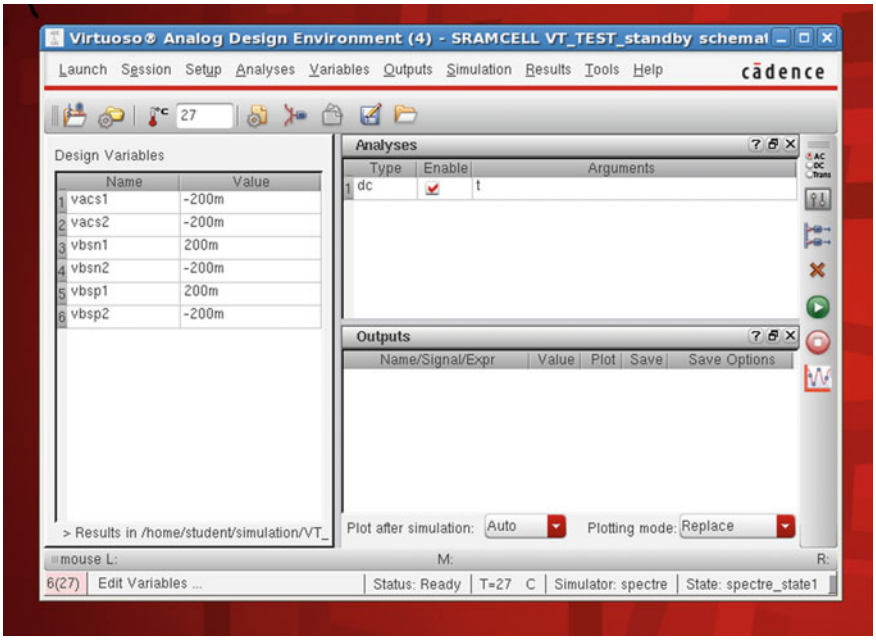


Fig. 7 Design simulation setup

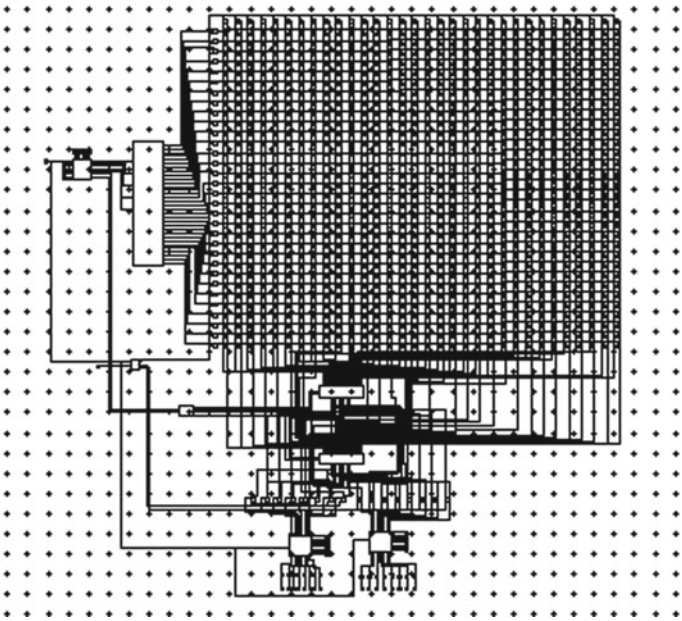


Fig. 8 32 × 32 SRAM array

4 SRAM Cell Design Metrics

a. Static Noise Margin (SNM)
SNM is maximum square nested between two Voltage Transfer Characteristics (VTC). It is caused by the reduction in the trip point of one of two cross-coupled inverters due to the increase in the threshold voltage of one pull-up transistor and one pull-down transistor on the other side [20, 21].
b. Read stability: It is a measure of how much faster cell can be read. RSNM is used to check read stability. Read SNM is more affected by NBTI than Hold SNM [22, 23].
c. Write stability: It indicates how difficult and easy to write memory cell. WSNM is to ensure write stability. Write stability of SRAM cell is measured with write margin. WM is improved with NBTI aging effect. WSNM reduces with supply voltage scaling [24].

5 Conclusion

Body biasing is an effective technique for designing ultra-low power circuits. In active mode FBB, body biasing is used to reduce the threshold voltage to improve

the performance by increasing drive current. Whereas, in standby mode RBB is used to increases the threshold voltage to have low static currents. FBB or RBB is a particular device based on the requirement to improve the SRAM cell design metrics.

References

1. Mostafa H, Anis M, Elmasry M (2011) Adaptive body bias for reducing the impacts of NBTI and process variations on 6T SRAM cells. IEEE Trans Circuits Syst I 58:2859–2871
2. Chen G, Sylvester D, Blaauw D, Mudge T (2010) Yield-driven near-threshold SRAM design. IEEE Trans Very Large Scale Integr (VLSI) Syst 18(11):1590
3. Kauppila AV, Bhuva BL, Kauppila JS, Massengill LW, Holman WT (2011) Impact of process variations on SRAM single event upsets. IEEE Trans Nucl Sci 58(3):834
4. Kim TT-H, Ba NL (2014) Design of a temperature-aware low-voltage SRAM with self-adjustable sensing margin enhancement for high-temperature applications up to 300 °C. IEEE J Solid-State Circuits 49(11):2534
5. Suthar R, Pande KS, Murty NS (2017) Leakage reduction in DT8T SRAM cell using body biasing technique. In: IEEE international symposium on nanoelectronic and information systems, pp 1–5
6. Shakir T, Sachdev M (2014) A body-bias based current sense amplifier for high-speed low-power embedded SRAMS, pp 444–448
7. Hokimoto S, Ishihara T, Onodera H (2016) Minimum energy point tracking using combined dynamic voltage scaling and adaptive body biasing, pp 1–6
8. Zheng N, Mazumder P (2017) Modeling and mitigation of static noise margin variation is subthreshold SRAM cells. IEEE Trans Circuits Syst I 64:2726–2736
9. Khayatzadeh M, Lian Y (2013) Average-8T differential-sensing subthreshold SRAM with bit interleaving and 1 k bits per bitline. IEEE Trans Very Large Scale Integr (VLSI) Syst, 1–12
10. Jiang C, Zhang D, Zhang S, Wang H, Zhuang Z (2017) A yield-driven near-threshold 8-T negative bit-line scheme, pp 315–318
11. Kim TT, Le Ba N (2013) A low voltage 8-T SRAM with PVT-tracking bitline sensing margin enhancement for high operating temperature (up to 300 °C), pp 233–236
12. Calimera A, Macii A, Macii E, Poncino M (2012) Design techniques and architectures for low-leakage SRAMS. IEEE Trans Circuits Syst I 59:1992–2007
13. Gupta SK, Raychowdhury A, Roy K (2010) Digital computation in subthreshold region for ultralow-power operation: a device–circuit–architecture codesign perspective. IEEE 98(2):160
14. Sharma CK, Chandel R (2014) Analysis of SRAM cell designs for low power applications. In: International conference for convergence of technology, p 1
15. Indumathi G, Aarthi alias Ananthakirupa VPMB (2014) Energy optimization techniques on SRAM: a survey. In: 2014 international conference on communication and network technologies (ICCNT), p 216
16. Sharan S, Chandra A, Goel N, Kumar A (2016) Comparison of 6T-SRAM cell designs using DTMOS and VTMOS for low power applications. In: IEEE international conference on power electronics, intelligent control and energy systems (ICPEICES-2016), pp 1–5
17. Torrens G, Alorda B, Carmona C, Malagón-Periánez D, Segura J, Bota SA (2017) A 65-nm reliable 6T CMOS SRAM cell with minimum size transistors. IEEE Trans Emerg Top Comput 1–9
18. Toh SO, Guo Z, Liu T-JK, Nikolić B (2011) Characterization of dynamic SRAM stability in 45 nm CMOS. IEEE J Solid-State Circuits 46(11):2702
19. Vashisht G, Pahuja H, Singh B, Panday S (2016) Design and comparative analysis of low power 64 bit SRAM and its peripherals using power reduction techniques

20. Kumar SS, Soni G (2015) Power, energy and SNM optimization of 6TSRAM cell using power gating technique. In: 2015 fifth international conference on communication systems and network technologies, p 889
21. Zheng N, Mazumder P (2017) Modeling and mitigation of static noise margin variation in subthreshold SRAM cells. IEEE Trans Circuits Syst I: Regul Pap 64(10):2726–2736
22. Saeidi R, Sharifkhani M, Hajsadeghi K (2014) Statistical analysis of read static noise margin for near/sub-threshold SRAM cell. IEEE Trans Circuits Syst I: Regul Pap 61(12):3386
23. Saeidi R, Sharifkhani M, Hajsadeghi K (2014) A subthreshold symmetric SRAM cell with high read stability. IEEE Trans Circuits Syst II: Express Briefs 61(1):26
24. Ruchi Sudeb Dasgupta (2018) Compact analytical model to extract write static noise margin (WSNM) for SRAM cell at 45 nm and 65 nm nodes. IEEE Trans Semicond Manuf 31(1):1–8

Increasing the Verification Analysis Using Tool Assessment as Per DO-254

Manju Nanda and P. Rajshekhar Rao

Abstract Programming under simulation-based testing remains the essential methods for utilitarian approval for HDL (hardware description language) plans. Code coverage which ensures to utilize simulation assets and a measure of test design. In this way, more target techniques, which utilize some all around characterized practical code measurements to play out a quantitative examination of simulation fulfillment, are proposed and quickly getting a response. For this reason, numerous utilitarian code measurements are proposed to confirm the design written in HDL. Keeping in mind the end goal to screen the code variation during simulation, a committed apparatus is required other than the test system. Finally, this paper includes the proposed well-known methodologies to achieve an accuracy of a test system using Code coverage.

Keywords HDL · RTL · Coverage analysis · VHDL · Verification tool

1 Introduction

Advance informal confirmation strategies has somewhat eased this issue. In any case, mechanized strategies constantly include a thorough investigation of a substantial state space and are consequently compelled to little segments of a design. In other case, applying these techniques to complex circuits with numerous functions is troublesome. To address this deficiency, simulation-based testing strategies have been produced. These strategies apply better control flow statement by utilizing different components to create input and assess recreation comes about. Approval scope, anyway ambiguous the idea, is basic for assessing and controlling

M. Nanda (✉)
Aerospace Electronics and System Division, CSIR-NAL, Bangalore, India
e-mail: manjun@nal.res.in

P. Rajshekhar Rao
Department of Avionics, IST, JNTUK, Kakinada, Andhra Pradesh, India
e-mail: rajshekhar7566246322@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_40

such blends of simulation-based and formal strategies. The perfect is to accomplish extensive approval without repetitive exertion. Code measurement's help estimated this perfectly by going about as heuristic measures that evaluate check fulfillment and distinguishing insufficiently practiced design perspectives. Scope investigation can be instrumental in dispensing computational assets and organizing diverse approval techniques. Although coverage-based procedures are routinely connected effectively to extensive modern designs [1].

DAL (design assurance levels) Levels A and B airborne equipment is basic to the well-being of the aircraft and its travelers. Code coverage analysis is one of them and it is the most ordinarily utilized strategy for design affirmation for FPGAs/PLDs. Primarily as a result of the way that there is as of now an outstanding instrument that can help—it is called Code Coverage using assertion and it is a standard component of Active-HDL [2].

2 Overview of Approach

To assess the applicant's plans in relation to the identified hardware level, safety features, and safety-related hardware requirements and to ensure that plans meet objectives identified in RTCA/DO254, and that the hardware will comply with other applicable hardware policy and guidance, shortly after completion of the hardware planning process, or any other necessary point. Plans may not be fully completed; however, they should be mature prior to the review and under configuration. Finally, tool qualification plans as applicable, findings/observations/issues from pre-review activities (highly recommended) by designees/applicants/developers, safety assessment, system architecture, hardware level justification, safety features, company policy and work instructions referenced in the plans and standards, and other data deemed necessary.

In the different useful scope measurements, code scope is the most well-known class. In this classification, a few measurements are the code scope measurements utilized as a part of programming testing, for example, assertion coverage and decision coverage, and a few measurements are proposed for HDL. As a rule, this sort of measurements is more straightforward to utilize, yet some outline blunders may not be distinguished in the test. FSM scope is another well-known class of useful scope measurements. It is considered to discover all bugs in a limited state machine (FSM) outline, yet its computational intricacy is substantially higher. Other than these measurements, there are as yet numerous measurements proposed for checking HDL outlines. The detailed design clarification about those practical scope measurements is given in the accompanying segments.

Code scope measurements, to a great extent, got from measurements utilized as a part of programming testing, recognize which structure classes in the HDL code to practice during simulation. The least complex such structure is a line or a square of code (an arrangement of lines with no control branches). More modern code-based measurements utilized as a part of equipment confirmation are branch, articulation,

and way scope. These measurements include the control course through the HDL code during requirement-based testing and are best portrayed by the control stream diagram (CFG) comparing to that code [3, 4, 5].

- Code Coverage Analysis contributes to the viability and culmination of necessities-based confirmation. It is a technique in which a framework is isolated into top-down level components and afterward examined at that level. On account of FPGAs/PLDs, the meaning of a component can be lines of HDL code or yields of combinatorial rationales. The activity of the candidate is to demonstrate that the necessities-based testbench covers the HDL configuration as characterized in the prerequisite and the confirmation design.
- Code Coverage examination gives a few unique kinds of data identified with the confirmation scope of the design data. Dynamic HDL offers the accompanying scope apparatuses:
- Report Coverage gives a point-by-point data on articulations that are executed during the design simulation to requirement-based testing. It analyzes each executable explanation and checks how frequently it has been executed. This data gives input on which parts of the plan were checked and which are untested. It likewise finds dead code. Branch Coverage inspects branches of the if or case explanation and checks how frequently a genuine or false condition was met by each branch amid the reproduction. Branch Coverage likewise gathers measurements for VHDL chose and restrictive flag task proclamations.
- Way Coverage is like Branch Coverage, anyway this sort of examination permits checking not just whether a condition was met and in result a specific branch of an announcement was inspected, however it additionally gathers data about back-to-back proclamations which have been executed, branches which have been analyzed, and how they assessed (to genuine or false) during reproduction. Arrangements of explanation executions and condition assessments shape interesting mixes of program execution alluded to as ways. A way is the subject to examine by Path Coverage.
- Practical Coverage demonstrates measurements for OVA, PSL, and SystemVerilog affirmations and cover proclamations. It indicates what number of affirmations began to assess during reproduction, what number of cases, and what number of succeeded. For the cover explanations, the measurements demonstrate the quantity of "matches" (Fig. 1).

Validation and Verification Process for the code coverage conducts the Hardware Design Description and Hardware Block Diagram review to determine whether the captured Hardware Design Description and Hardware Block Diagram satisfy the objectives of RTCA/DO254. This review also reduces the risk of designer producing a Hardware Design Description and Hardware Block Diagram that does not meet RTCA/DO254 objectives or other certification criteria. The Hardware Design Description and Hardware Block Diagram review should take place after the initial completion of the conceptual design process.

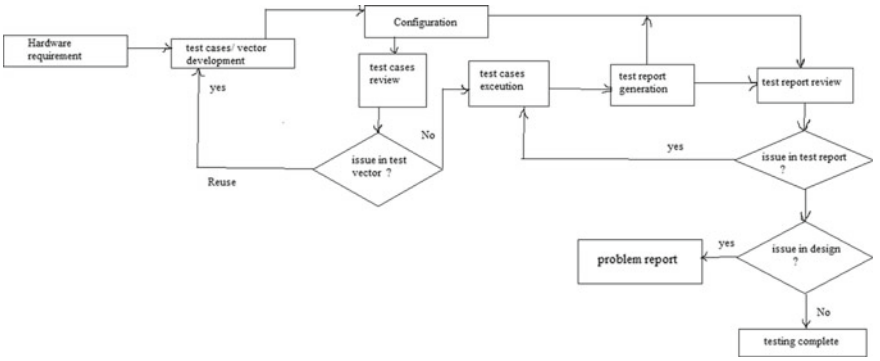


Fig. 1 Verification analysis as per Hardware requirement

Table 1 Method of verification

Design	Verification methods	Reports
RTL code	Behavioral simulation at unit level and integration level	Code coverage, behavioral simulation and traceability reports
Post translate verification model	Functional simulation at integration level	Functional simulation results
Post place and route verification model	Static timing analysis, functional and timing simulation at integration level	Static timing analysis, functional and timing simulation reports

Verification of design description includes correctness of requirements, conceptual design data in order to ensure that conceptual design data correctly and completely represent the behavior specified in the requirements.

Development tools used during the hardware development and verification processes are listed in Table 1. The information provided includes the tool name, tool vendor, and the main function of the tool. Qualification of a hardware development or verification tool is required when the outputs of the tool are not independently verified and the tool eliminates, reduces, or automates required lifecycle processes. The qualification status of each tool is also indicated [6, 7, 8] (Table 2).

3 Proposed Methodology

The reason for confirmation in a DO-254 program is to guarantee that a plan plays out the capacity indicated by its prerequisites. Confirmation is viewed as one of DO-254's "supporting procedures": It is certifiably not a particular period of improvement, yet rather happens all through advancement. A key part of the plan

Table 2 Verification analysis using tool assessment

Tool name	Tool vendor	Tool version	Tool functions
Microsoft Office	Microsoft	2016	Word Processor, Spreadsheets, Power point etc.
MS Project	Microsoft	2010	Project Management
ReqTrace	Mentor Graphics	2015.1a	Capture/Record, Trace requirements
Xilinx ISE	Xilinx	14.7	Synthesis, Place and Route
HDL Designer	Mentor Graphics	2015.1	RTL Development, Linting, Code to Requirement Trace, Coding Guidelines, Rule Checking
Questa Core	Mentor Graphics	10.4	Functional Simulation, Post synthesis simulation, Post Place and Route Simulation. Code coverage analysis
Clear Quest	IBM Rational	8.0.1	Bug tracking
Clear Case	IBM Rational	8.0.1	Version Management/tracking of all changes of source and design data

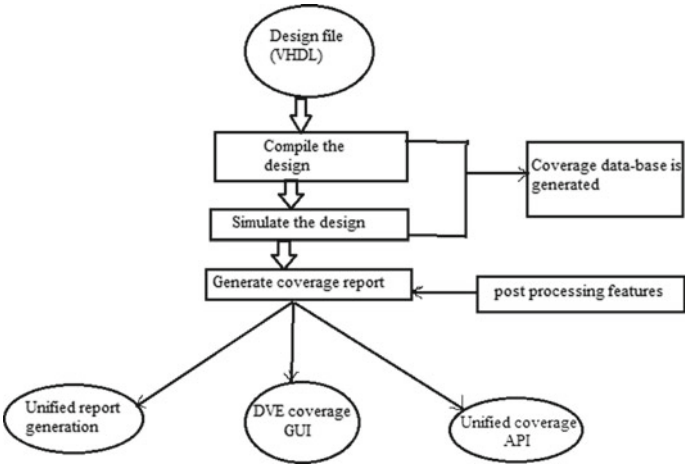


Fig. 2 Verification analysis as per coverage metrics

affirmation DO-254 depends on the confirmation being performed at each key stage, from the most particular models to the last testing of the segment in the framework, with the goal that it fulfills settled upon finish criteria (Fig. 2).

In the most security basic DAL A and B devices, the check must be completed autonomously of the design. The check work commonly performed on, for instance, a FPGA configuration would incorporate recreation testing, and conceivably different examination (e.g., linting, clock-space crossing investigation, and so forth) of the

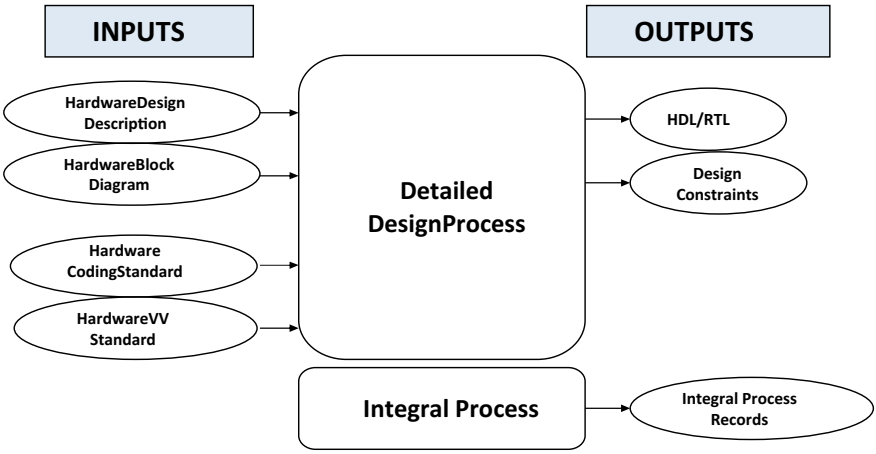


Fig. 3 Overall analysis for an critical module under coverage analysis

HDL code, static planning examination of the blended netlist, back-explained timing reproduction of the netlist, lab testing, and last framework testing as listed below.

- Generates the detailed design data for the hardware item includes interconnection data, component data, HDL and test methods.
- Test features should be designed in, where necessary, to facilitate verification.
- Constraints on the design should be identified.
- Derived requirements produced during the detailed design process should be fed back to the conceptual design or other appropriate processes.
- Requirement omissions and errors discovered during the detailed design process should be provided to the appropriate process for resolution (Fig. 3).

These recreation tests are produced from prerequisites. They should in this way be completely investigated to guarantee they check that those necessities are met. Be that as it may, in the help of plan affirmation, another inquiry must be inquired. Do these prerequisites-based tests adequately test all the usefulness inside the execution? This is of central significance in light of the fact that a plan could precisely execute an arrangement of necessities, yet at the same time contain other usefulness that could in actuality show unintended and conceivably ruinous practices. Some kind of extra action or metric is important to guarantee that the necessities-based confirmation (RBV) exercises adequately test the majority of the usefulness executed in the outline.

This metric is normally offered by running code scope during HDL reproduction [7] (Fig. 4).

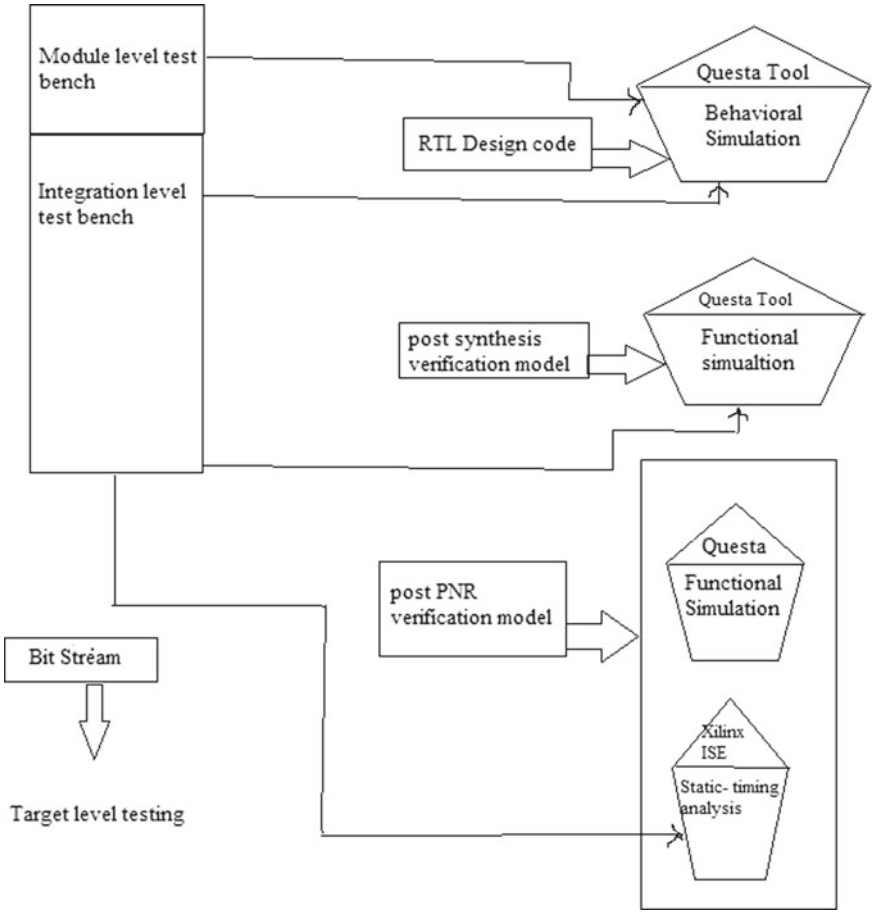


Fig. 4 Special consideration of tool assessment

4 Conclusion

The scope-driven useful confirmation is quickly getting well-known. Numerous practical scope measurements have been proposed to confirm the plans written in HDL. An overview on a few well-known utilitarian scope measurements has been introduced in this paper. In any case, albeit such huge numbers of various useful scope measurements have been proposed, there is as yet not a solitary metric like the stuck to blame demonstrate in the assembling test being prominently acknowledged as total and solid. A ton of endeavors is as yet required in growing better utilitarian scope tests.

References

1. Ip CN (2000) Simulation coverage enhancement using test stimulus transformation. In: Proceedings of IEEE/ACM international conference on computer-aided design, digest of technical papers. IEEE Press, Piscataway, N.J., 2000, pp 127–133
2. Ho P-H et al (2000) Smart simulation using collaborative formal and simulation engines. In: Proceedings of IEEE/ACM international conference computer-aided design, digest of technical papers. IEEE Press, Piscataway, N.J., 2000, pp 120–126
3. Fournier L, Koyfman A, Levinger M (1999) Developing an architecture validation suite: application to the power PC architecture. In: Proceedings of 36th design automation conference. ACM Press, New York, 1999, pp 189–194
4. Malka Y, Ziv A (1998) Design reliability-estimation through statistical analysis of bug discovery data. In: Proceedings of 35th design automation conference. ACM Press, New York, 1998, pp 644–649
5. Arditi L, Clave G (2000) A semi-formal methodology for the functional validation of an industrial DSP system. In: Proceedings of IEEE international symposium on circuits and systems. IEEE Press, Piscataway, N.J., 2000, pp 205–208
6. Taylor S et al (1998) Functional verification of a multiple issue, out-of-order, superscalar Alpha processor—the DEC Alpha 21264 Microprocessor. In: Proceedings 35th design automation conference. ACM Press, New York, 1998, pp 638–643
7. Kantrowitz M, Noack LM (1996) I'm done simulating; now what? Verification coverage analysis and correctness checking of the DEC chip 21164 Alpha Microprocessor. In: Proceedings of 33rd design automation conference. ACM Press, New York, 1996, pp 325–330
8. Zhu H, Hall PV, May JR (1997) Software unit test coverage and adequacy. ACM Comput Surv 29(4):366-427
9. Grinwald R et al (1998) User defined coverage—a tool supported methodology for design verification. In: Proceedings of 35th design automation conference. ACM Press, New York, 1998, pp 158–163
10. Vemuri R, Kalyanaraman R (1995) Generation of design verification tests from behavioral VHDL programs using path enumeration and constraint programming. IEEE Trans Very Large Scale Integr (VLSI) Syst 3(2):201–214

Reducing Power in Register Files for CAM- and SRAM-Based Processor Units

K. Muralidharan and K. Sridevi

Abstract Power consumption reduction in out-of-order superscalar processor has become very important in recent era due to the utilization of superscalar processor in all portable systems. Register Files is the one of the power-hungry source in the out-of-order processor design. The proposed design of Register File using power gating technique reduces both static power and dynamic power dissipation. The new design of Register also includes with a minimal amount of redesign and verification efforts, the minimum level of design risk and least amount of hardware overhead and without any significant impact on the performance of the out-of-order superscalar processor.

Keywords Out-of-order super scalar processor · Register Files · Power gating · Low power

1 Introduction

Nowadays embedded applications in the imaging, automation, multimedia, networking, IoT, and different types higher end consumer applications domains are increased due to high-performance requirements. So we need a more complex out-of-order superscalar processor in order to meet their performance goals. Due to its complex operation that is a high-performance on-chip integration is also increases. If on-chip integration increases, the power consumed by that processor also increases. So that necessity for low power design in the battery-powered applications increases sharply. Because of its expensive packaging and cooling costs and less reliability, Low power design is a serious design consideration in the higher end applications [1].

K. Muralidharan (✉)
Department of ECE, Coimbatore Institute of Technology, Coimbatore 641014, India
e-mail: muralidharan.k@cit.edu.in

K. Sridevi
Department of ECE, CIT Sandwich Polytechnic College, Coimbatore 641014, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_41

A Register File (RF) is an array of memory structure which is used for temporary data storage in the superscalar processor. RF used to store an all kinds of data, including the operands, Arithmetic Logic Unit (ALU) computation results, execution unit cache information and memory addresses allocation of data retrieved by the out-of-order superscalar processor units [2]. In order to perform specialized hardware operations also Register Files were used. In modern out-of-order superscalar processors, multiple Register Files to accommodate the different types of logic and execution units which means that integer instructions executed by integer executions only. Floating point execution units manage floating point instructions only. It needs larger Register Files in order to store floating point data and corresponding instructions. Based on the architectural design, processor specifications and their speed may be limited in the RF which is calculated by data access time [3]. Hence Register Files is an important component in order to determine the superscalar processor performance. Due to its frequent access and corresponding operations performed by the processor Register Files becomes a high-power hungry unit in the out-of-order superscalar processor.

Homayoun, et al. discussed a dynamic and optimized Register file adaptive resizing technique in order to improve the performance and energy-delay reduction in Register Files. Heuristics problem was created in this design while mapping instructions in different clusters [4].

2 Register Files

Register File is the central core of an out-of-order superscalar microprocessor. It provides communication of register values between producer and consumer instructions. If issue width increases, the number of ports and the registers required will also increase [5]. Due to this area of a multi-ported Register Files to grow more quadratically based on their issue width.

2.1 Conventional Design of Register Files

All the bit lines should be pre-charged high to read or write the data in each and every cycle in memory. During the write operation, bit lines were set to be high. To read the data content of any entry, any one of the corresponding bit line would be temporarily discharged. The sense amplifier is used to detect such type of differences. But the bit lines must run across the entire ROB and Register File height.

The write bit lines are the prominent sources of power dissipation Register File structures due to multiple accesses to the same cell in each and every cycle [6]. More leakage current dissipation is also occurred due to the leakage currents of the memory

cells, the bit lines which flow through the two pass transistors in an off condition. In order to eliminate the leakage current dissipations in the memory cells, the portion of power consumption would be to turn off for the unused entries and their associated word line drivers using such a gated-V_{dd} or gated-V_{ss} power gating technique (Sleep mode).

Figure 1 shows the existing circuit diagram of Register Files. It is necessary to detect when the lower partitions of RF becomes empty after the end of the cache miss period. It can be realized using an additional bit in each row of the lower partition. While using a lower partition, an additional bit is set and then reset when the entry is released during the commit operation. Bit detects whether the partition is empty or not. The logic to ORing all bits in the lower partition of RF is not on the critical path. Then the downsizing decision is done parallel to accessing Register file. All the bit lines should be pre-charged high to read or write an entry cycle. To read the content of data entry, one of bit line would be conditionally discharged. To detect such a multiple access, the sense amplifier is used. Even though this structure is effective in reduction of power without significant loss of performance, further power reductions are required.

3 Proposed Register Files

In order to improve the performance of the register file, 9T SRAM with is used in deep submicron technology based microprocessor design. The new design in the register file is implemented to reduce the power consumption.

The proposed design reduces power dissipation by minimizing the all type parasitic capacitance, short circuit and leakage current. The proposed design of Register Files is shown in Fig. 2. Since many leaky paths are in extensive fan in gates, noise immunity and EMI of the gate is also reduced. The sense amplifier is used for sensing the voltage difference on the bitline and their corresponding complement while performing read operation. In order to have separate pre-charge circuitry, it has to been duplicated for both the upper and lower segments in the Register File configuration.

4 Results

Using BSIM 3V3 65, 45 and 32 nm, technology on Tanner EDA tool, the conventional Register Files and proposed Register Files in Out-of-Order superscalar processer have been simulated. The testing condition is prepared and tested through the identical input values on room temperature with supply voltage ranging from 0.6 to 1.0 V. Length and Width of channel for PMOS and NMOS transistors are 650 nm and 65 nm, 450 nm and 45 nm, 320 nm and 32 nm for 65 nm, 45 nm and 32 nm technology transistor, respectively. Transient analysis is used to calculate conventional and proposed design of Register File power consumptions.

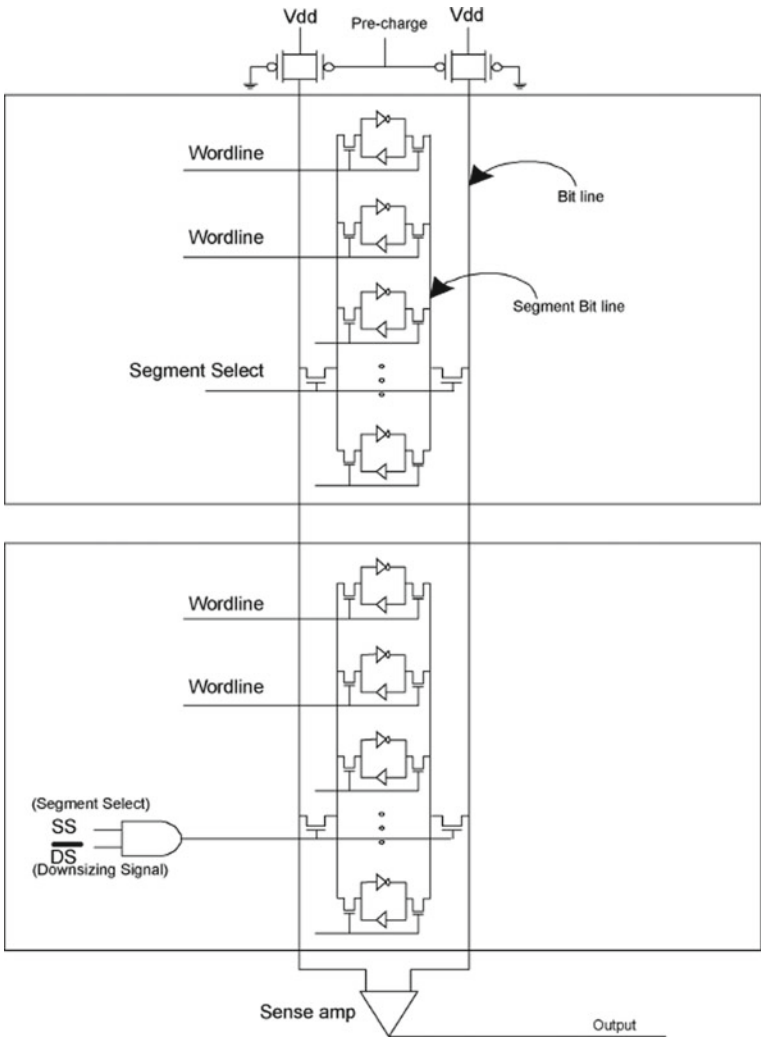


Fig. 1 Circuit diagram of existing RF

The loaded data and a stored data output for the data is 11012 in the memory. The corresponding output waveform of a proposed Register files is shown in Fig. 3.

The average power consumed by the proposed Register File design is compared with conventional Design of Register Files for various input voltages in different technologies applied in order to check process variation check in the out of superscalar processor.

The power reduction is observed in the proposed design of Register Files of out-of-order superscalar processor shown in Figs. 4, 5, and 6 for 65 nm technology, 45 nm

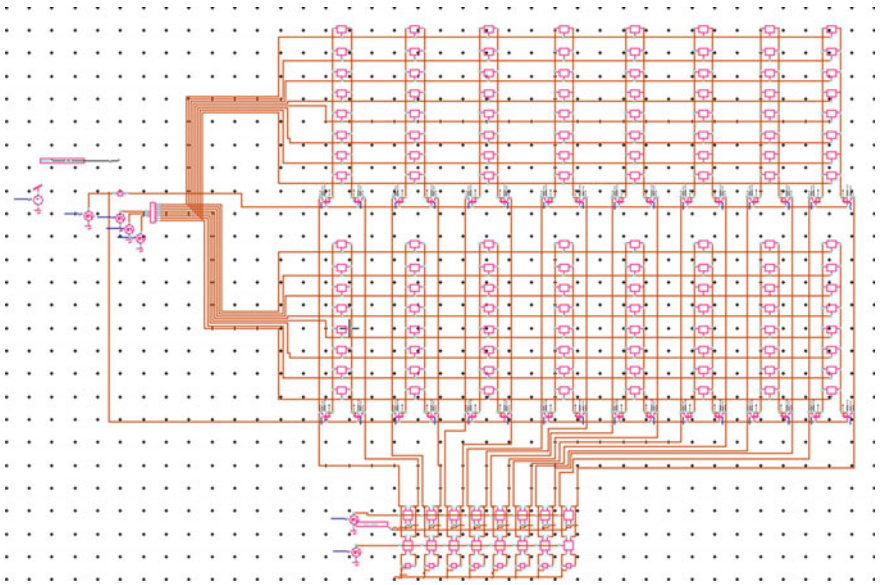


Fig. 2 Simulation diagram of proposed Register Files

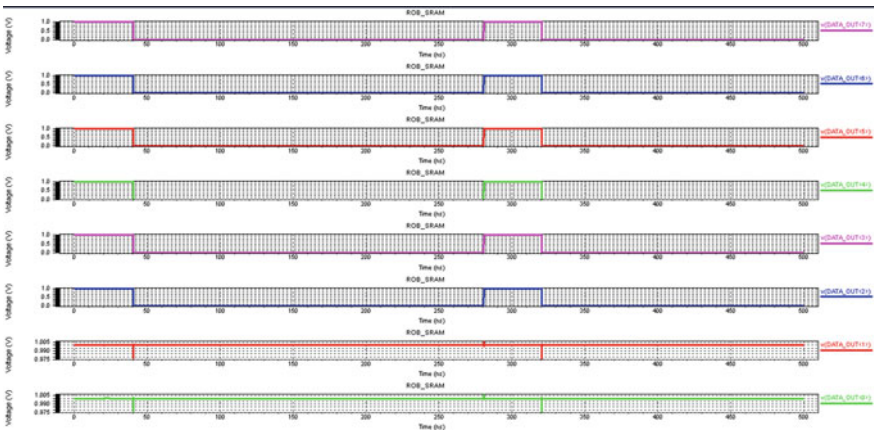


Fig. 3 Output waveform of the new design Register Files

technology, and 32 nm technology design of register Files in the out of superscalar processor respectively.

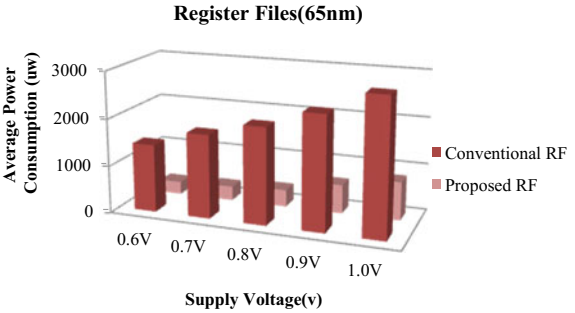


Fig. 4 Power analysis of Register File (65 nm)

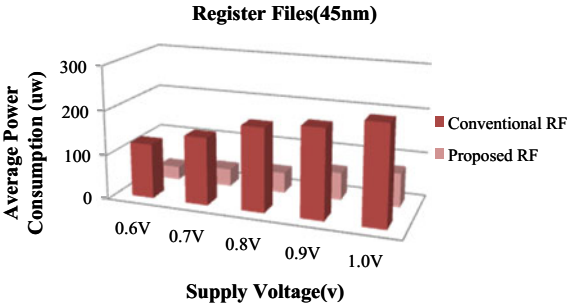


Fig. 5 Power analysis of Register File (45 nm)

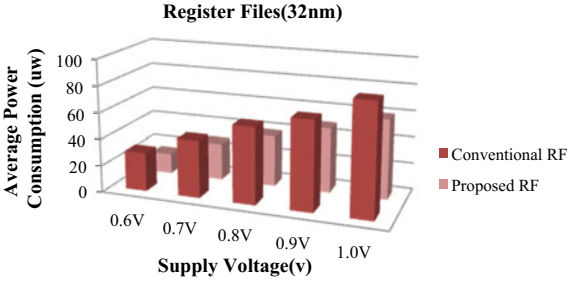


Fig. 6 Power analysis of Register File (32 nm)

5 Conclusion

The design structure of new Register Files was implemented using power gating techniques. The observed average power consumption in the proposed design of Register Files from power analysis at 65, 45, and 32 nm fabrication technology shows power reduction compared with conventional Register File. The reduction of power prolongs the lifetime of the superscalar processor.

References

1. Homayoun H (2011) Reducing power in all major CAM and SRAM based processor units via centralized, dynamic resource size management. IEEE Trans Very Large Scale Integr Syst 19(11)
2. Kucuk G, Ponomarev D, Ghose K (2002) Low-complexity reorder buffer architecture. In: 16th ACM international conference on supercomputing (ICS'02), New York, June 2002, pp 57–66
3. Hu JS, Vijaykrishnan N, Irwin MJ (2004) Exploring wakeup-free instruction scheduling. In: Proceedings of the 10th international conference on high-performance computer architecture (HPCA-10 2004), 14–18 February 2004, Madrid, Spain
4. Hu Z, Buyuktosunoglu A, Srinivasan V, Zyuban V, Jacobson H, Bose P (2004) Microarchitectural techniques for power gating of execution units. In: International symposium on low power electronics and design, 2004
5. Stark J, Brown MD, Patt YN (2000) On pipelining dynamic instruction scheduling logic. In: Proceedings of the international symposium on microarchitecture, December 2000
6. Do A-T, Chen S, Kong Z-H, Yeo KS (2013) A high speed low power CAM with a parity bit and power-gated ML sensing. IEEE Trans Very Large Scale Integr (VLSI) Syst 21(1)

Comparative Analysis of 8-Bit ALU in 90 and 45 nm Technologies Using GDI Technique

T. Swapna Rani and Kumud Kumar Bhardwaj

Abstract Arithmetic and Logic Unit is the fundamental building block of almost all the digital devices. ALU is the heart of the Central Processing Unit (CPU) of the computer. It is generally used to perform the arithmetic and logical operations in the device. ALU can perform arithmetic operations like addition, subtraction, multiplication, increment, decrement, shift operations, and logical operations like AND, OR, XOR, NAND, NOR, etc. Any digital device has its characteristics; based on its characteristics ALU is configured. The parameters like the speed of operation, power, accuracy, etc., are taken care during the design. This paper presents the design of 8-bit ALU which performs 16 operations in 45 and 90 nm technology and compares parameters like power and delay.

Keywords Arithmetical logical unit · GDI technique · Low-power VLSI

1 Introduction

GDI Technology is also called as GATE DIFFUSION INPUT, it is the current trend going on in VLSI industry. In this design a GDI Cell which uses PMOS and NMOS transistors [1–3]. AND, OR, XOR, etc., can be implemented using GDI technique. The basic disadvantage of CMOS technology is that as its name suggests it is used to design the complementary functions and also uses number of transistors to implement any logic gates. When the number of transistors increases in the circuit power and delay gets increased which are much undesired for any VLSI Engineer. In order to reduce the number of transistors, power and delay of the circuit GDI technique are preferred. Figure 1 shows the basic gate implementation.

T. Swapna Rani (✉) · K. K. Bhardwaj (✉)
ECE, CMR College of Engineering & Technology, Medchal, India
e-mail: tswapnarani@cmrcet.org

K. K. Bhardwaj
e-mail: kumudbhardwaj12@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_42

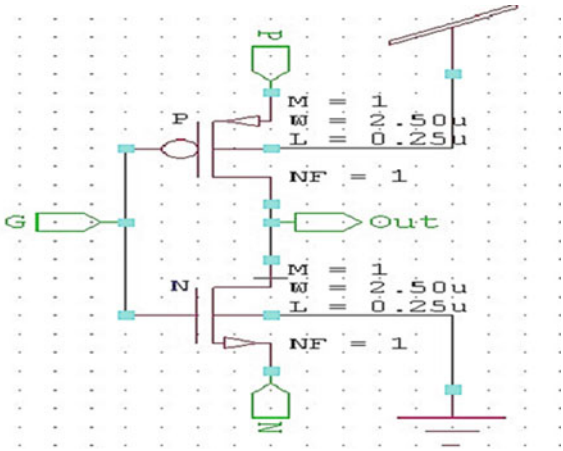


Fig. 1 Basic GDI cell

1.1 Working Procedure of Basic GDI Cell

In a GDI cell, three inputs can be given instead of giving only one input. The first input *G* is given to the gate terminal, input *N* is given to the source of NMOS and the third input *P* is applied to the drain of PMOS, only the substrate connection of PMOS and NMOS are given to power supply and ground respectively which is shown in Fig. 1.

1.2 Power Consumption in VLSI

There are mainly three regions in MOSFET

1. Linear region
2. Saturation region
3. Cut-off region

For $V_{GS} > V_T$, drain current will become

$$I_D = k'\left(\frac{W}{L}\right)(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2. \tag{1}$$

This equation is for the linear region. Here drain to source voltage V_{DS} is less than $V_{GS} - V_T$.

$$I_D = \frac{1}{2}k'\left(\frac{W}{L}\right)(V_{GS} - V_T)^2 \tag{2}$$

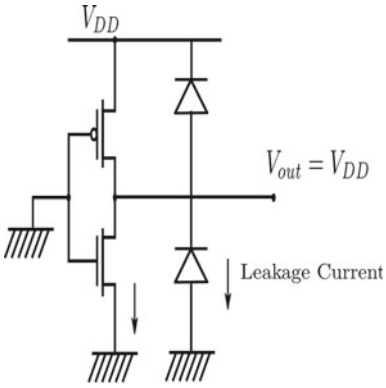


Fig. 2 Leakage current

This is the case for saturation region. Here V_{DS} is greater than or equal to $(V_{GS} - V_T)$.

$$I_D = 0 \tag{3}$$

This is the case for cut-off region. Here V_T is threshold voltage at which MOSFET [4, 5] conducts, V_{GS} is the gate to source voltage, W/L defines the geometry of the device can be combined with the factor k'. Here V_{GS} is less than the threshold voltage V_T.

Power loss on the bus can be caused by three sources

1. Leakage current occurs through the n-p junction diode, which is shown in Fig. 2.
2. Short-circuit takes place when there is another path going to ground.
3. Charging and discharging of parasitic capacitance takes place. Leakage Component of power dissipation can be shown by Fig. 1, that is, the drain to source current leakage current is independent of the drain to source voltage V_{DS}. The reduction in the losses gives better result to the system [6–8].

The rest of the paper is organized as follows. In Sect. 2, the block diagram of ALU has been described. Section 3 shows the result and analysis of the ALU with different technologies. Section 4 deals with the conclusion.

2 Block Diagram of ALU

An 8-bit ALU which performs 16 operations is designed. Two 8 × 1 Mux and one 2:1 Mux have been used to design ALU as shown in Fig. 3 [9–11]. Based on the operation to be performed few modules have only single 8-bit inputs and the rest have two 8-bit inputs. There are totally four select lines S0, S1, S2, S3, where the S0 represents the MSB and the S3 represents the LSB. For different combinations of

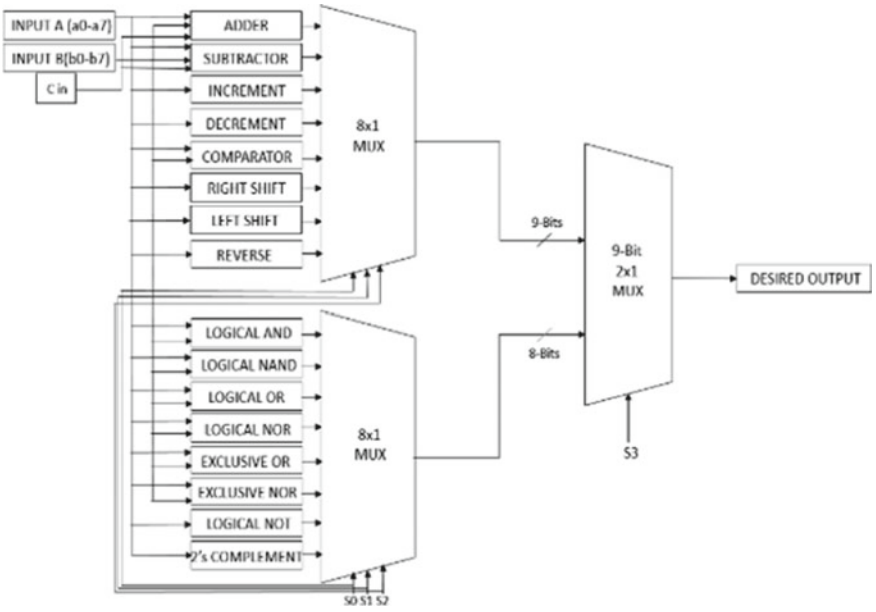


Fig. 3 Block diagram of ALU

inputs show different operations, e.g., 0000 = Addition, 0001 = Logical AND, 0010 = Subtraction, 0011 = Logical NAND, 0100 = Increment, 0101 = Logical OR, 0110 = Decrement, 0111 = Logical NOR, 1000 = Comparator, 1001 = EX-OR, 1010 = Arithmetic Right Shift, 1011 = Ex-NOR, 1100 = Arithmetic Left Shift, 1101 = Logical Not, 1110 = Reverse, 1111 = 2'S complement.

2.1 Adder and Subtractor

In this block diagram, modified carry-select adder is implemented. Instead of using the three ripple carry adders as in conventional carry-select adder, one RCA is replaced with a Binary to Excess-1 converter (BEC). This would reduce the size and makes the design more efficient in terms of speed than the conventional carry-select adder.

Subtractor is the direct implementation of the logic equation of Difference = A xor B xor Bin and the Borrow = A'B + A'C + BC for one bit the same is cascaded to form for 8-bit input.

Table 1 Comparative analysis of different technologies

Parameters	45 nm	90 nm
Delay	17.44 ps	170 ps
Static power	0.009428 mW	9.905 mW
Dynamic power	0.007252 mW	0.325 mW
Total power	0.01668 mW	10.23 mW

2.2 Incrementer and Decrementer

In this module, the applied input value is increased by 1. This has been achieved by adding binary 1 on the LSB, which has been implemented by using half adders and cascading them to form for 8-bit inputs.

In decrement block applied input value has been decreased by 1. This has been achieved by subtracting binary 1 on the LSB, which has been implemented by using half subtractions and cascading them to form for 8-bit inputs.

2.3 Arithmetic Right Shift and Left Shift

A module which can perform one-bit right shift on the applied data is implemented. The right shifter has been implemented using the concept of the barrel shifter, here the 2×1 MUX is used to build a right shift operation, the inputs are connected based on the requirement and if the select line is high the outputs generated would be a right shifted data.

In the left shift, a module which can perform one-bit left shift on the applied data is designed. Left shifter has been implemented using the concept of barrel shifter [12–14], here the 2×1 MUX is used to build a left shift operation, the inputs are connected based on the requirement and if the select line is high the outputs generated would be a left shifted data. Apart from this, comparator [4, 15, 16] has been designed in order to generate the output data which is greater among the two applied inputs. The comparator used here is based on one's complement logic. With the reverse module, bits get reversed.

3 Result and Analysis

Table 1 shows the comparative analysis of the 45 and 90 nm of ALU using GDI technique. From delay point of view, 45 nm technology is the best approach. Also from the static and dynamic point of view 45 nm is the best approach in designing of ALU and at last total power also get reduced in 45 nm technology.

4 Conclusion

The ALU which can perform 16 operations is designed which will give one output, which is the desired operation. This design is implemented by using GDI technique using both technologies 90 and 45 nm. The results prove that 45 nm technology is better than 90 nm as it uses less power and the delay. As the design uses GDI technology, the area and transistor count get reduced which in turn gives an advantage in terms of cost, size, power consumption, and also the delay.

Acknowledgements Behind every achievement, there is unfathomable sea of gratitude to those who supported it and without whom it would ever have been a success one. We are thankful to Principal and head of department for their support and encouragement.

References

1. Koppad D, Hiremath S. Low power 1-bit full adder circuit using modified gate diffusion input (GDI). In: Conference on first international micro and nano technologies
2. Uma R, Davachelvan P. Low power and high speed adders in modified gate diffusion input technique. In: Computer networks & communications (NetCom)
3. Esther Rani T, Asha Rani M, Rao R (2011) Area optimized low power arithmetic and logic unit. 978-1-4244-8679-3/11/$26.00 ©2011, IEEE
4. Wolf W (1998) Modern VLSI design—systems on silicon. Prentice Hall
5. Tsividis Y (1999) Operation and modeling of the MOS transistor. Mc Graw-Hill
6. Anitha D, Chari KM (2014, Oct) Low power ALU design considering PVT variations. Int J Comput Appl 104(17)
7. Kamaraju M, Kishore KL, Tilak AVN (2010, Dec) Power optimized ALU for efficient datapath. Int J Comput Appl 11(11)
8. Swetha S (2016, July) Design of low power and area efficient full adder using modified gate diffusion input. Int J Comput Appl 145(8)
9. Salehi S, DeMara RF (2015) Energy and area analysis of a floating point unit in 15 nm CMOS process technology. In: Proceedings of IEEE Southeast conference 2015 (SECon-2015), Fort Lauderdale, FL, 9–12 April 2015
10. Clark LT, Hoffman EJ, Miller J, Biyani M, Liao Y, Strazdus S, Morrow M, Velarde KE, Yarch MA (2001) An embedded 32-b Microprocessor Core for low-power and high performance applications. IEEE J Solid-State Circ 36(11):1599–1608
11. Deleganes DJ, Barany M, Geannopoulos G, Kreitzer K, Morrise M, Milliron D, Singh AN, Wijeratne S (2005) Low-voltage swing logic circuits for a Pentium ® 4 Processor Integer Core. IEEE J Solid-State Circ 40(1):36–43
12. Pilmeier MR (2002) Barrel shifter design, optimization and analysis. Lehigh University, Jan 2002
13. Priyanka Mandal, Siddhant Malani, Palsodkar PM. VLSI implementation of barrel shifter. In: Proceedings of SPIT-IEEE colloquium and international conference, Mumbai, India
14. Pappachan R, Vijaykumar V, Ravi T, Kannan V (2013) Design and analysis of 4-bit low-power barrel shifter in 20 nm FINFET technology. IJES 2(3):17–25
15. Bhattacharyya P, Kundu B, Ghosh S, Kumar V, Dandapat A (2015) Performance analysis of a low-power high-speed hybrid 1-bit full adder circuit. IEEE Trans Very Large Scale Integr (VLSI) Syst 23(10):2001–2008
16. Srivastava A, Srinivasan C (2002) ALU design using reconfigurable CMOS logic.Preview the documentView in a new window. Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on. (vol 2), IEEE, 2002

Buffer Overflow Attack and Prevention for an FPGA-Based Soft-Processor System

Chamandeep Singh, Sripadam Satish, Jubin Mitra and Sandeep Shukla

Abstract Embedded systems have a plethora of role in all spheres of life ranging from medical application to nuclear application to industrial applications. For rapid prototyping of designs, reconfigurable hardware like FPGA is the most preferred solution. Such designs often rely on an embedded processor in the FPGA—sometimes implemented as a soft configuration in FPGA fabric or hard-wired as the integrated circuit while still connected to the FPGA fabric. Typically, embedded systems are exposed to various types of attacks, of those major vulnerability is the buffer overflow attack on the processor used. The buffer overflow attack corrupts the return address of a function or process and subsequently changes the execution order. Our goal is to study the effect of buffer overflow attack in an embedded processor, demonstrate the attack on a full chain of embedded system and providing cost-effective mitigation solution to prevent this type of attack.

Keywords Intel FPGA · Soft-processor · Buffer overflow attack · Embedded systems security · Vulnerabilities in C

1 Introduction

An FPGA is a device that contains programmable arrays of logic elements. Basic elements of the FPGA are CLBs (configurational logic blocks), programmable interconnects and I/O blocks. These logic blocks are further made up of transistor pairs, logic gates, lookup tables and multiplexers. The reason FPGA-based systems found so wide application because of shorter development time, easy maintenance and remote upgrade. The major area of applications ranges from data processing and

C. Singh (✉) · S. Satish
NIT, Tiruchirappalli, Tamil Nadu, India
e-mail: divergentchaman@gmail.com

J. Mitra · S. Shukla
IIT Kanpur, Kanpur, Uttar Pradesh, India
e-mail: jubinmitra@cse.iitk.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_43

storage to instrumentation, telecommunications and digital signal processing. A typical FPGA solution consists of multiple interface links, dedicated logic cores, and a processor to supervise, control or monitor the system. It is easier for the design user to go for a soft-processor-based solution within the FPGA itself than to venture for SOC (System-On-Chip) Hard processor Systems (HPS) or heterogeneous computing solution.

Embedded systems are exposed to various types of attacks. These attacks are broadly classified into three categories: logical, physical and side channel. Logical attacks can further be classified into software and cryptographic attacks. Cryptographic attacks take advantage of the weakness of cryptographic protocols for stealing information, whereas software attacks are more common due to the ease of implementation. These attacks try to find the security loopholes in the operating systems and application software. Buffer overflow attacks, format string attacks and integer overflows are some of the examples of software attacks.

In our study, we have focused on buffer overflow attacks. The attack is emulated in latest FPGA hardware from Intel with soft-processor. The attack manipulates the function call procedure and function pointers. By corrupting the return address in the stack, the attacker can cause the program to jump to exploit code when the victim function returns. This type of buffer overflow is called "*stack smashing attack*". Once the attacker overwrites the return address, they can choose what instructions to be executed next, for example, spawning a new shell. The instructions (exploit code) can be placed in the buffer that is overflowing and then overwrites the return address, such that it points back into the buffer. This form of attack is called "*Injected Code*" attack. The attacker can pass characters representing binary instructions in hexadecimal notation [1]. The code (instructions) to be injected can be generated manually, provided one has the target processor instruction set architecture knowledge. It can also be first written in the C language and then compiled and assembled to get the equivalent binary instructions.

2 Literature Review

In the late 1980s, 10% of the Internet nearly came to a halt, when Robert T. Morris designed a worm using a buffer overflow vulnerability in UNIX's fingerd program. It disrupted many computers on what was then a fledgling internet. He was arrested and prosecuted for the same, later on. Nearly three decades later, in 2014, this attack surfaced again exposing this vulnerability even in modern computing systems [2, 3]. A buffer overflow vulnerability in the cryptography library of OpenSSL was disclosed to the public. This flaw came to be known as "Heartbleed." It exposed millions of users of online services and software platforms to a vulnerable version of the OpenSSL software. One more example of this attack is the Apache HTTP server's htpasswd.c program, which manipulates password file for Apache HTTP server was found to be vulnerable to stack-based buffer overflow attacks when the binary has root permissions or if the script is accessible through a CGI (Common

Gateway Interface) [4]. The vulnerable part of the code takes user supplied 'user' and copies it to a fixed size local buffer using strcpy.

In another instance, Benjamin Kunz-Mejri from security firm Vulnerability Lab discovered that Skype software of Microsoft is also vulnerable to the buffer overflow attack. Vulnerability Lab reported the flaw to Microsoft on 16 May 2017, Microsoft released a patch on 8 June 2017 in Skype version 7.37.178. The fault present in Skype software was exposed to attacks by remote attackers without the authenticate user's knowledge and was possible even with the basic Skype account. Attackers could potentially crash the software with one request to overwrite to overwrite the EIP register of the active software process [5]. Thus, allowing local or remote attackers to execute own codes on the affected and connected computer systems via the Skype software.

Even Google's Android software development kit (SDK) was found to be vulnerable to this attack by the droidsec group. In this particular case, an attacker could initiate an Android debug bridge server that interfaces with Android devices on a multi-user system and then waits for ADB clients, started by developers wanting to debug apps or send commands to devices for connection. Due to exploit occurring early in the protocol negotiations, any command that communicates with the ADB server will lead to successful exploitation. Droidsec team also discovered that ADB Binary failed to have a non-executable stack [6]. Having a non-executable stack is one of the ways to mitigate the buffer overflow vulnerability.

3 Design Methodology

Intel® Cyclone IV FPGA development board from Terasic is used for the emulation of the Buffer overflow attack. Platform Designer, a design integration tool available with Intel Quartus Prime software suite is used for the design development. The soft-processor used for the experiment is NIOS II [7], a RISC-(Reduced Instruction Set Computer) based processor. The application code on the soft-processor is written in Micrium MicroC/OS-II RTOS kernel [8], a variant of ANSI C. MicroC/OS-II RTOS is a priority-based real-time multitasking operating system kernel primarily used for embedded system designs. SDRAM-based memory space is used to store the software codes.

System Integration: Platform Designer simplifies the task of defining and integrating customized IP Components (IP Cores) into the designs. The interconnection of individual design components are shown in Fig. 1. The purpose behind the use of individual components are listed below.

Utility of individual design components:

- *Source clock (clk_0)*: On board crystal oscillator to provide the reference clock of 50 MHz.

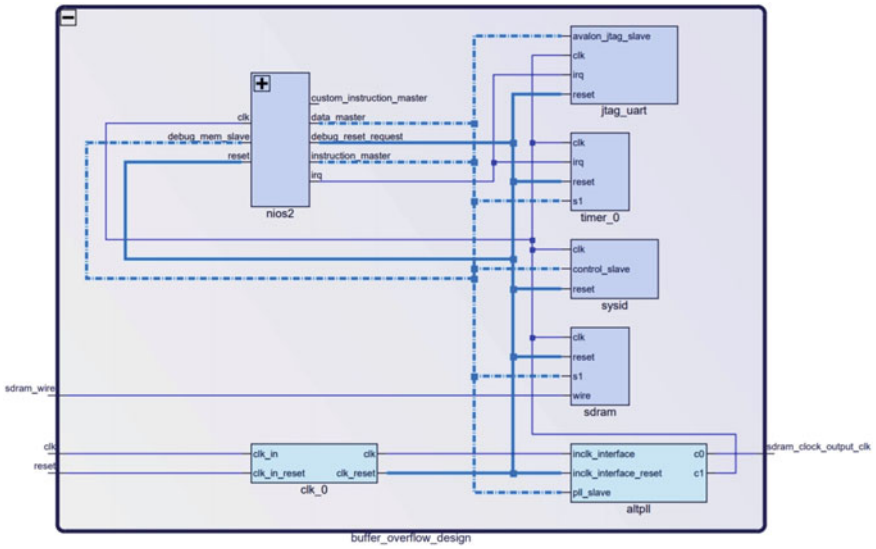


Fig. 1 System design using Nios II processor with SDRAM

- *Altera PLL (altpll)*: Is used for correcting the clock skew before forwarding it to the SDRAM to compensate for PCB routing delay and the other output is to clean the reference clock before using it in the design.
- *JTAG-UART (jtag_uart)*: This logic core enables the communication between the host PC and the SOPC Builder system of an FPGA using the serial character streams.
- *SDRAM (sdram)*: Used for storing and processing of the data and instructions.
- *NIOS II Processor (nios2)*: A RISC-based soft-processor to control the system. It is composed of two master lines and all peripheral logic components are connected to it.
- *System ID (sysid)*: Provides Platform Designer systems with a 32-bit unique identifier associated with the NIOS processor. Any mismatch in the system ID results in the improper execution by the software.
- *Timer (timer_0)*: A 32-bit counter is needed since we are dealing with thread-based real-time operating system.

3.1 Performing Buffer Overflow Attack

Once the firmware connection for the processor is completed, the application software codes are written in Micro C/OS II kernel that operates over the Hardware Abstraction Layer (HAL) framework provided by the Intel. C language has various string handling library functions like *strcpy*, *strcat* in *string.h* Header file. The *strcpy* function copies

a null-terminated string from one memory location (called source array) to another (destination array). These functions do not check for the length of the input argument in their implementation. Therefore, when the input argument size is more than the destination variable, data overflow spreads into adjacent memory regions. Moreover, a buffer overflow enables us to change the return address of a function. In this manner, we can change the flow of execution of the program [4, 9]. Other C library functions like *gets*, *scanf*, *sprintf* are also buffer overflow prone. A vulnerable C-code is given in Listing 1.1, over which we will perform the buffer overflow attack.

```
#include <stdio.h>
#include "includes.h"
#include "string.h"

/*Definition of Task Stacks*/
#define TSK_STKSZ 2048
OS_STK task1_stk[TSK_STKSZ];
OS_STK task2_stk[TSK_STKSZ];

/* TASK PRIORITY */
#define TASK1_PRIORITY      1
#define TASK2_PRIORITY      2

/* data passed is copied into local
     buffer */
void user_input(char *data) {
 char name[4];
 strcpy(name, data) ;
 return;

}

/* TASK1 */
void task1(void* pdata)
{
user_input("\x3a\xf0\x4a\x29\xe8\
    x02\x00\x02");
  while (1)
  {
    printf("Credential Mismatched\n
    ");

    OSTimeDlyHMSM(0, 0, 3, 0);
  }
}

/* TASK2 Never Called */
void task2(void* pdata)
{
  while (1)
  {
    printf("Credential Matched\n");

    OSTimeDlyHMSM(0, 0, 3, 0);
  }
}

/* The main function creates
   two task and starts
   multi-tasking */
int main(void)
{

printf("MicroC/OS-II based Buffer
    Overflow\n");

  OSTaskCreateExt(
  task1,
  NULL,
  (void *)&task1_stk[TSK_STKSZ-1],
  TASK1_PRIORITY,
  TASK1_PRIORITY,
  task1_stk,
  TASK_STACKSIZE,
  NULL,
  0);

  OSStart();
  return 0;
}
```

Listing 1.1. An example of a standard application code

NOP Sled Technique to do buffer overflow: The attacker first passes (an arbitrary) string to the function to see if the program malfunctions. If a string of this length does not have any effect on the program and output is still the same. The attacker continues this process by adding NOPs until the program malfunctions, and the flow of execution is altered. The attacker now knows the initial stack address and the effective length of the buffer. At this stage, the attacker uses NOP sled technique with the initial stack address to point the new return address (ra) into the injected code function. Nios II implements NOP as 'add r0, r0, r0', as r0 is a constant register with a value of zero. But 'add r0, r0, r0' when encoded will have null bytes in it. Hence, in order to avoid this problem, the NOP is implemented as 'xor r5, r5, r5'. This instruction simply flushes r5. To supply (input) a string in hexadecimal format to C

programs, we need to use an '\x' character for each byte [10]. Since the Nios II architecture is little endian, the equivalent input string for the 'xor r5, r5, r5' instruction is '`\x3a\xf0\x4a\x29`'. With NOP sled technique, the program malfunctions for a string of '`\x3a\xf0\x4a\x29\x3a\xf0\x4a\x29\x3a\xf0\x4a\x29`' of length 12 bytes or 3 NOPs. Hence, the buffer overflow attack can be made by appending the injected code's address to this string. Nevertheless, when the processor pops out the return address 'ra' it sees a different return address than initially stored in the stack. In this way, the flow of execution of the program can be altered by the manipulated the string.

4 Mitigation of Buffer Overflow Vulnerability

Several mitigation methods are available to counter the discussed attack. In our particular case of embedded soft-processor, we implemented the use of safer library functions and replaced the vulnerable functions with them. C is vulnerable to this attack due to direct access to memory and lack of strong object typing. The string handling functions like strcpy and strcat copy a string into a buffer and then append the contents of one buffer upon another, respectively. These two are unsafe because they do not check any bounds on the target buffer and would write past the buffer's limits if given enough bytes to do so. So, it is suggested to use their associated **strn-** versions. These versions only write to the maximum size of the target buffer.

5 Conclusion

We discussed the menace of software-based buffer overflow attack on FPGA-based embedded systems. Here, we demonstrated the attack using Intel's Cyclone IV FPGA board with the Nios II soft-processor and Micrium MicroC/OS-II RTOS kernel. NOP Sled technique is used to implement the attack and discussed a mitigation method to prevent the attack. As a future research direction, hardware-based robust mitigation solutions are planned.

Acknowledgements The research work is funded by the DST for C3I center at Department of Computer Science, IIT Kanpur.

References

1. Cowan C, Wagle F, et al (2000) Buffer overflows: attacks and defenses for the vulnerability of the decade. In DISCEX'00 Proceedings, IEEE, 2000
2. Synopsis Editorial Team. Prevention of buffer overflow attack. Synopsis Security Blog Post, 2017

3. Synopsys Editorial Team. How to detect, prevent, and mitigate buffer overflow attacks. Synopsis Security Blog Post
4. James CF, Vitaly O, Nish B (2005) Buffer overflow attacks: detect, exploit, prevent. Syngress. ISBN: 9780080488424
5. Paganini P (2017) Experts found a critical remote buffer overflow vulnerability in Skype. Security Affairs Web-Post
6. Duckett C (2014) Android SDK suffers from buffer overflow and lack of hardening. ZDNet newsletter
7. Intel Corporation. Processor architecture. White Paper, 2015
8. Using MicroC/OS-II RTOS with the Nios II Processor Tutorial, 2011
9. FreeBSD Developers' Handbook. Buffer overflows, 2018
10. Nios II processor reference handbook. Intel Corporation, December 2010

High-Voltage Gain CMOS Charge Pump at Subthreshold Operation Regime for Low Power Applications

C. Arul Murugan, B. Banuselvasaraswathy and K. Gayathree

Abstract This paper presents an enhanced complimentary metal oxide semiconductor (CMOS) charge pump (CP) circuit with improved gain and efficiency by dynamically controlling substrate and gate terminals of PMOS transistor. The proposed novel charge pump provides good performance even at low voltage. But, the efficiency of the charge pump circuit is being influenced by body effect and threshold voltage as the number of stage increases. Hence, numerous charge pump strategies are used to minimize the effect of threshold voltage and body effect. In this paper, the above-mentioned issues are overcome by implementing charge pump circuit in 180-nm standard in Cadence Virtuoso operated at low-voltage CMOS technology. From the result obtained, it is analyzed that the proposed CMOS charge pump achieves an improved voltage conversion ratio, efficiency and low operating voltage as compared to existing conventional CMOS charge pumps.

Keywords Charge pump · Voltage gain · Threshold voltage · Body effect · Clock signal generator

1 Introduction

In the modern era, the utilization of portable devices is progressing due to reduced power consumption. In the nanometre-scale technology, still there is continuous demand for minimizing the power consumption and supply voltage of integrated circuits (IC) [1, 2]. But, there exists several applications such as EEPROM programmers, MEMS devices, power switches and line drivers [3] that need high voltage to perform its operation. This made the necessity for the utilization of charge pumps (CPs) at light-load applications and requirement of step-up dc–dc converters in heavy-load

C. Arul Murugan
Department of ETE, Karpagam College of Engineering, Coimbatore, India

B. Banuselvasaraswathy (✉) · K. Gayathree
Department of ECE, Sri Krishna College of Technology, Coimbatore, India
e-mail: banu.saraswathy74@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_44

applications. An on-chip charge pump (CP) circuit is highly preferred in non-volatile memories and analog circuits to deliver maximum output voltage and voltage gain. In some cases, high-voltage generators are embedded on the IC to provide solutions for small form-factors [4–7].

Charge pumps are widely used than inductor-based converters since it provides an optimum result in terms of electromagnetic interference and integration. In conventional methods, high voltage are generated through charge pump continued by LDMOS output switch to result in maximum output voltage with additional costs for mask. In nanometre technology [8], the LDMOS output switch is eliminated and stacking devices operating at low-voltage devices are introduced [9, 10]. These devices require predriver circuit to increase the reliability of the circuit. But the complexity arises when the number of stacked devices increases. At increased state, it cannot adapt in high-voltage drivers. This leads to a limitation of using stacked devices at more number of stages in nanometre technology. It is overcome by cascading the identical charge pump to n stages based on the requirement. The number of stages is limited by voltage handling. The pumping capacitors should be charged and discharged during each switching period. Therefore, the charge pump-based drivers are suitable only for slow switching applications. At low-voltage devices, the charge pump-based drivers result in a high voltage, low capacitance value. Hence, it is widely used than LDMOS output switches. The efficiency of the architecture is increased by minimizing the rise time of charge pump. Besides all these factors, the utilization of charge pump is progressing due to its small size and reduced amount of power dissipation [11, 12].

2 Related Works

Several researchers discussed the design features of charge pump they are as follows. The fundamental conventional four-stage Dickson CP circuit consists of NMOS and diodes. The NMOS transistors are used as pass devices. The entire circuit helps to pump up the input voltage [13]. But the major drawback of these circuits is that the threshold voltage drop across the diode-connected transistors degrades the voltage gain of the Dickson CP circuit. In addition, the body effect of NMOS transistor raises the threshold voltage during each pumping stage at varying supply voltage. This, in turn, minimizes the voltage gain of CP circuit. In [14], a dynamic control circuitry was introduced to decrease the voltage drop. In this technique reduced number of diodes are used to connect the NMOS transistors. Even though the number of diodes is reduced, the voltage gain of the circuit remains worse by minimizing the use of diode-connected NMOS pass devices. In [15], the diodes are connected across PMOS transistors and two auxiliary PMOS transistors are placed to bias the substrate terminal to eliminate the body effect. But still, there is a limitation in voltage gain of the circuit due to the constant threshold voltage drop that occurs due to the diode-connected PMOS device. In [16], the Ron (on-resistance) of PMOS transistors is reduced in this technique. But the body terminal of the PMOS transistors remains in

floating condition during substrate control approach reducing the power efficiency and voltage gain. In [3], employed a bootstrap circuit to help in obtaining high voltage which is required to turn ON the charge transfer switches in the triode region. This, in turn, eliminates the threshold voltage drop. The series–parallel topology [17] was introduced for discrete IC implementation. The major drawback of this system is that the high-voltage swing on parasitic capacitance values results in maximum power loss. In [18] and [19], a charge-pumping architecture with two branches is introduced with a six-phase clocking scheme. In this approach, both forward and backward control schemes are used [20]. In [21], clock generator operating at high voltage is demonstrated to get a linear output voltage irrespective of extended of cascaded stages. In [22], a high-voltage CMOS process to obtain high voltage and density of charge pump circuit. In [23], double charge-pumping circuit with a four-phase clocking scheme is implemented. In [24] and [25], the number of stages is optimized to reduce the rise time. In [26], instead of fixed frequency, different pumping clock frequency is utilized to decrease the rise time with maximum power efficiency. In [27], charge pump optimized capacitor value expression is estimated to reduce the output impedance of the circuit.

In this paper, several optimization features and advanced control schemes are introduced into the CP circuit to enhance the voltage gain and the power efficiency and reduce the threshold voltage, body effect and power consumption. The paper is organized as follows: Sect. 3 discusses about the design and operation of proposed charge pump circuit, Sect. 4 deals with different results obtained as a result of simulation and Sect. 5 describes the conclusion.

3 Proposed Charge Pump Circuit

To overcome the above-discussed concerns, a new approach with parallel substrate and dynamic gate control is introduced with three stage CP circuit. The proposed design provides biasing to all the body terminals of PMOS transistors with various clock phases. The design is free of the diode from input to output to connect transistors. In addition the CP circuit small on-resistance on each pass transistors. All these features result in reduced voltage drop across each transistor to maximize the voltage gain and minimize the threshold voltage and body effect of the CP circuit. Moreover, the performance and effect of pumping capacitors and its sizing on the rise time are also taken into account.

This circuit comprises of five PMOS pass transistors T1–T5 and four on-chip pumping capacitors C1–C4 for providing the output voltage of 5 V under different clock phases [28]. The transistor T1 drain terminal present in the first stage is connected to the clock signal Θ_1 to guarantee the improved operation of the CP circuit. All transistors (M1–M5) from the input to output stage operate in the triode region. This is to ensure that once the transistors are turned ON, the circuit will not experience threshold voltage drop across the connected transistors. Hence, the circuit's voltage gain remains unaffected. Similarly, the shoot current that occurs during transistor

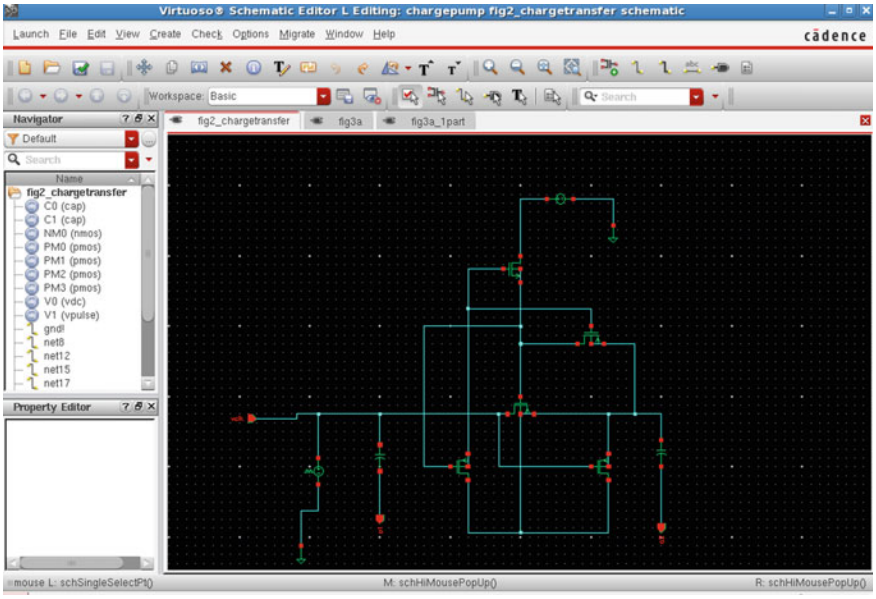


Fig. 1 Single stage charge pump circuit

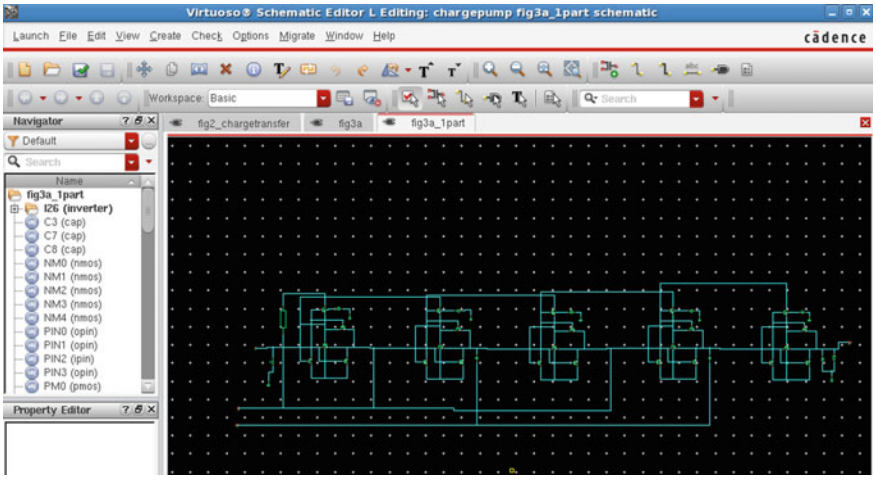


Fig. 2 Proposed charge pump circuit

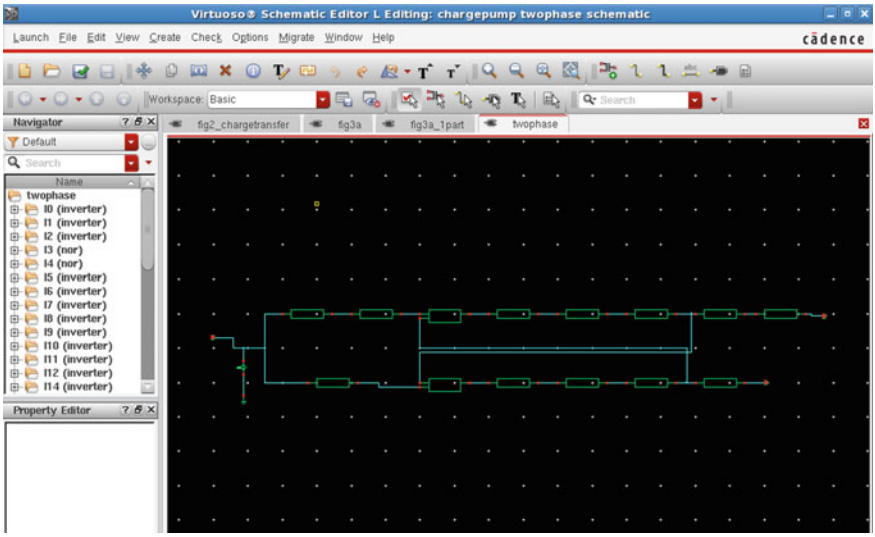


Fig. 3 Clock signal generator with different clock inputs

switching between on-state and off-state of the pass transistors is reduced to increase the power efficiency. This is carried out by using non-overlapping clock signal generator inputs Θ_1 and Θ_2 into the charge pump circuit [28]. In the CP circuit, rise in voltage at the built-to-source of NMOS transistor at increased stages will increase its corresponding threshold voltage. As long as the transistor T5 threshold voltage remains lesser than the voltage of gate to source of transistors the NMOS transistor performs as a switch without affecting the charge pump circuit (Figs. 1, 2 and 3).

The rise time value at different capacitor loads (CL) are given by the following equation [29]

$$R_{\mathrm{T}} = \mathrm{CUE} - C_{\mathrm{E}}/C_{\mathrm{E}}$$

where C_{UE} determines unequal pumping capacitors and C_{E} indicates equal pumping capacitors and R_{T} is the rise time.

4 Results and Discussion

Figure 4 shows the output obtained from the clock signal generator obtained with two different clock signals.

Figures 5, 6, 7, and 8 shows the comparison of capacitor values C1, C2, C3, and C4 at different stages $N = 2$, $N = 3$, and $N = 4$ (Table 1; Fig. 9).

Fig. 4 Output of clock signal generator

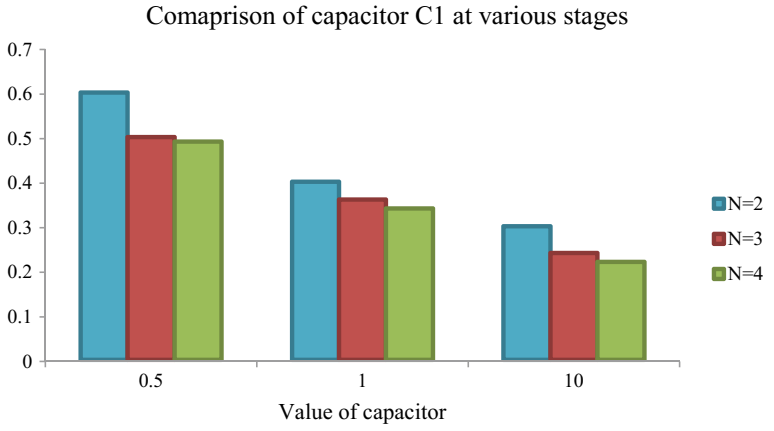


Fig. 5 Comparison of capacitor C1 at different stages

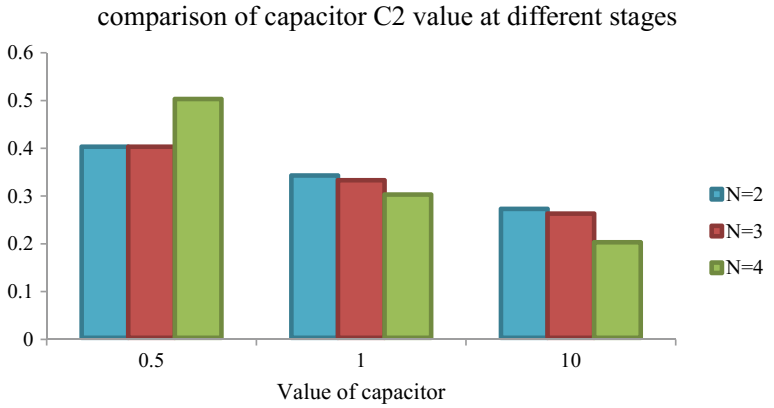


Fig. 6 Comparison of capacitor C2 at different stages

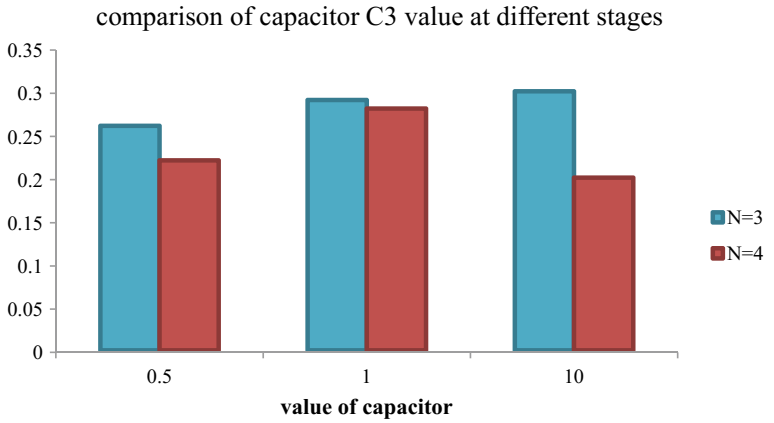


Fig. 7 Comparison of capacitor C3 at different stages

Table 1 Comparison of different conventional charge pump circuit with the proposed technique

Characteristics	Dickson	Ref. [17]	Pelliconi Ref. [4]	Proposed technique
Output voltage (V)	5.73	6.19	6.39	7.41
Voltage gain	4.78	5.16	5.33	5.95
Start up response (μs)	0.55	1.11	1.20	1.47
Load current (mA)	5	5	5	5

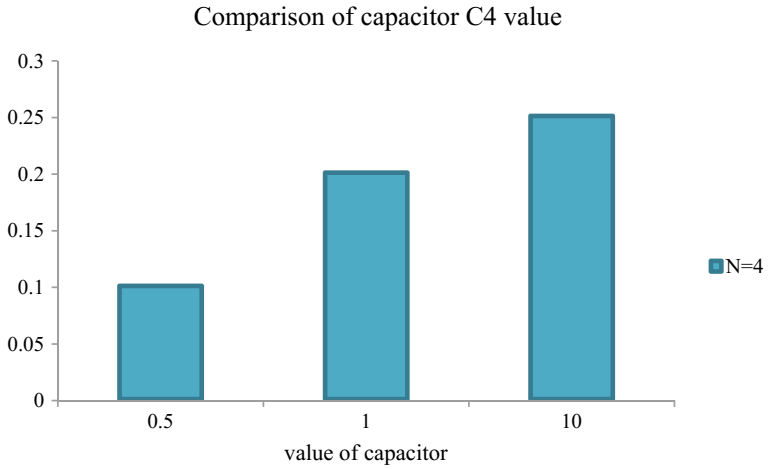


Fig. 8 Comparison of capacitor C4 at different stages

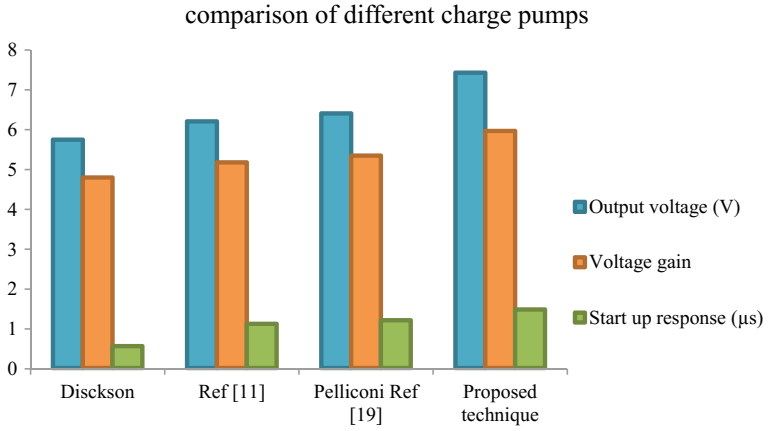


Fig. 9 Comparison of different charge pump circuits

5 Conclusion

A CMOS charge pump which exhibits high-voltage gain was designed. In this design, the diode switches are completely removed to maintain the threshold voltage drop across the input to the output stages. Moreover, the body terminal is biased with different clock pulses and on-resistance on every transistor is minimized. All these in-built features result in reduced threshold voltage drop across the connected transistors and body effect. The entire design is simulated using 180 nm-Cadence Virtuoso operated at low-voltage CMOS technology. From the result obtained, it is analyzed

that the proposed CP circuit provides better voltage gain and output voltage when compared to the other existing conventional charge pump circuits.

References

1. AbdElFattah M, Mohieldin A, Emira A, Sanchez-Sinencio E (2011) A low-voltage charge pump for micro scale thermal energy harvesting. In: Proceedings of IEEE ISIE, June 2011, pp 76–80
2. Abdelaziz S, Emira A, Radwan A, Mohieldin A, Soliman A (2011) A low start up voltage charge pump for thermoelectric energy scavenging. In: Proceedings of IEEE ISIE, June 2011, pp 71–75
3. Umezawa A, Atsumi S, Kuriyama M, Banba H, Imamiya K, Naruke K, Yamada S, Obi E, Oshikiri M, Suzuki T, Tanaka S (1992) A 5-V-only operation 0.6-μm flash EEPROM with row decoder scheme in triple-well structure. IEEE J Solid-State Circ 27(11):1540–1546
4. Ker M-D, Chen S-L, Tsai C-S (2006) Design of charge pump circuit with consideration of gate-oxide reliability in low-voltage CMOS processes. IEEE J Solid-State Circ 41(5):1100–1107
5. Arul Murugan C, Banuselvasaraswathy B (2015) High gain enhanced CMOS charge pump with reduced leakage and threshold voltage. Int Res J Eng Technol (IRJET) 02(03)
6. Gayathree K, Arul Murugan C, Banuselvasaraswathy B, Ishwarya Niranjana M, Agnes Shiny Rachel N (2018) A robust single ended 10 T Schmitt Trigger based SRAM cell with enhanced read/write assist techniques. Int J Pure Appl Math 118(20):411–416
7. Arul Murugan C, Banuselvasaraswathy B, Gayathree K, Ishwarya Niranjana M (2018) Efficient high throughput decoding architecture for non-binary LDPC codes. Int J Eng Technol 7(2.8):195–200
8. Cha HK, Zhao D, Cheong JH, Guo B, Yu H, Je M (2013) A CMOS high-voltage transmitter IC for ultrasound medical imaging applications. IEEE Trans Circ Syst II Express Briefs 60(6):316–320
9. Annema AJ, Geelen GJGM, de Jong PC (2001) 5.5-V I/O in a 2.5 V 0.25-μm CMOS technology. IEEE J Solid-State Circ 36(3):528–538
10. Ismail Y, Yang CKK (2014) A 12-V charge pump-based square wave driver in 65-nm CMOS technology. In: Solid-state circuits conference (A-SSCC), 2014 IEEE Asian, Nov 2014, pp 237–240
11. Saeed A, Ibrahim S, Ragai HF (2016) A sizing methodology for rise-time minimization of Dickson charge pumps with capacitive loads. IEEE Trans Circ Syst—II
12. Banuselvasaraswathy B (2013) A new enhanced trellis based decoding architecture for punctured codes using modified max product algorithm. Int J Adv Inf Sci Technol (IJAIST) 18(18):36–43
13. Dickson JF (1976) On-chip high-voltage generation in MNOS integrated circuits using an improved multiplier technique. IEEE J Solid-State Circ 11(3):374–378
14. Wu J-T, Chang K-L (1998) MOS charge pumps for low-voltage operation. IEEE J Solid-State Circ 33(4):592–597
15. Shin J, Chung I-Y, Park YJ, Min HS (2000) A new charge pump without degradation in threshold voltage due to body effect. IEEE J Solid-State Circ 35(8):1227–1230
16. Mensi L, Colalongo L, Richelli A, Kovács-Vajna ZM (2005) A new integrated charge pump architecture using dynamic biasing of pass transistors. In: Proceedings of European solid-state circuits conference, Sept 2005, pp 85–88
17. Cheng K-H, Chang C-Y, Wei C-H (2003) A CMOS charge pump for sub-2.0 V operation. In: Proceedings of IEEE international symposium on circuits systems, May 2003, pp V-89–V-92
18. Khouri O, Gregori S, Cabrini A, Micheloni R, Torelli G (2002) Improved charge pump for flash memory applications in triple well CMOS technology. Proc IEEE ISIE 4:1322–1326

19. Pan J, Yoshihara T (2007) A charge pump circuit without overstress in low-voltage CMOS standard process. In: Proceedings of IEEE conference on EDSSC, 2007, pp 501–504
20. Wang X, Wu D, Qiao F, Zhu P, Li L, Pan L, Zhou R (2009) A high efficiency CMOS charge pump for low voltage operation. In: Proceedings of IEEE 8th international conference on ASIC ASICON, 2009, pp 320–323
21. Huang WC, Cheng JC, Liou PC (2008) A charge pump circuit—cascading high-voltage clock generator. In: Proceedings of 4th IEEE international symposium on DELTA, 2008, pp 332–337
22. Richard J-F, Savaria Y (2004) High voltage charge pump using standard CMOS technology. In: Proceedings of 2nd annual IEEE NEWCAS, 2004, pp 317–320
23. Nozaki M, Tangsrirar W, Suzuki Y, Yoshid M, Saitoh S, Teramoto M, Yamaguchi A (1998) New double charge-pumping circuit for high-voltage generation. In: Proceedings of IEEE APCCAS, Nov 1998, pp 719–722
24. Tanzawa T, Tanaka T (1997) A dynamic analysis of the Dickson charge pump circuit. IEEE J Solid-State Circ 32(8):1231–1240
25. Emira A, AbdelGhany M, Elsayed M, Elshurafa AM, Sedky S, Salama KN (2013) All-pMOS 50-V charge pumps using low-voltage capacitors. IEEE Trans Ind Electron 60(10)
26. Arul Murugan C, Karthigai Kumar P (2018) Survey on image encryption schemes, bio cryptography and efficient encryption algorithms. Mobile Netw Appl. https://doi.org/10.1007/s11036-018-1058-3
27. Kim KY, Kim Y, Lee D, Kim SW, Park J (2011) An energy efficient VPP generator with fast ramp-up time for mobile DRAM. IEEE J Solid-State Circ 46(6):1488–1494
28. Seeman MD, Sanders SR (2008) Analysis and optimization of switched-capacitor DC-DC converters. IEEE Trans Power Electron 23(2):841–851
29. Banuselvasaraswathy B (2014) Trellis based decoding architecture for non-binary LDPC codes using modified Fano algorithm to achieve high throughput. Int J Adv Inf Sci Technol (IJAIST) 23(23):383–389

Part V
Miscelleneous

Robot Path Planning Using Memory

Gadhamsetty Ravi Theja and Srinath R. Naidu

Abstract Proposed work implements the robot path planning using memory. Nowadays, robots are mostly used in agriculture, industrial, and military purposes where the robots are moving autonomously. So, the robot requires planning a path for moving from one spot to other location, which is called as path planning for a robot. Robots are moving in different kinds of environment, mainly obstacle and obstacle-free environment. In obstacle-free environment, path planning must ensure the most direct path between start and end locations. In an environment surrounding with obstacles, path planning must ensure a collision-free shortest path between start and end locations by avoiding obstacles. Most recent works use the A-Star algorithm to catch the most direct route between start and end locations. In the proposed work, once the robot arrives at the end location then the path is stored in memory. The idea behind this is that when the robot reaches the same location in future, there is no requirement to forecast the path again and it can be retrieved from memory which will save the computational time and power requirements of the robot.

Keywords A-Star algorithm · Shortest path · Memory · Retrieval

1 Introduction

In today's generation, the usage of robots in agriculture is incremented and becoming an exciting high-technology industry for new professionals, investors, and companies. The technology is developing very rapidly in order to advance the production capabilities of the farmers using robotic and automation technologies. Agriculture

G. R. Theja · S. R. Naidu
Department of Computer Science and Engineering,
Amrita School of Engineering, Bengaluru, India
e-mail: teja2656@gmail.com

G. R. Theja · S. R. Naidu (✉)
Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: r_srinath@blr.amrita.edu

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_45

robots mechanize continual, passive, and dull tasks of the farmers so that farmers will focus on bettering complete production yields. Agriculture robots are mainly used in the areas of harvesting and picking, weed control, autonomous mowing, pruning, seeding, spraying, thinning, and utility platforms. Nowadays, robots are used in industries in order to automate the applications throughout the factory system to free time and energy. Applications of robots in industries comprise painting, welding, assembly, pick and place, and packaging and testing.

In space applications, robots play a decisive part in analyzing and exploring the outer space environment which is hostile in nature. The advancement in technologies making it easier to send robots to space for gathering information because they will not get tired and can work in air-less environment. Space robots are mainly robot arms, rovers, landers, and orbiters.

In case of military applications, robots perhaps used to carry weapons from one locale to other locale. On a war field, robots can be sent before humans to clear out landmines and thus potentially save human lives.

1.1 In Industries

In case of industrial applications, the robots are mainly used to carry things from one section to other locations. There could be many paths available to find the target location but the optimal collision-free path can be selected and the path is stored in memory and the path is retrieved whenever required.

1.2 In Agriculture

Today, in agriculture domain the conventional tasks like tilling, sowing, and harvesting can be done by robots. In case of fruit picking, robots have to maneuver from one spot to another location to perform tasks and the shortened direction between these points by avoiding obstacles can be found using A-Star algorithm. Once the goal point is reached the path is stored in memory so that the robot becomes familiar with the environment in which it is moving.

1.3 In Space

In case of space applications, robots are used to investigate things in space. In order to get this, the robots have to be competent to shift to different locations. Given start and end locations, a robot should be adequate to figure out the optimal collision-free path. Having once computed this path, the robot can store it in memory and retrieve it whenever required.

2 System Architecture

2.1 Existing System

Traditionally, people have done path planning using Dijkstra's algorithm but this algorithm constitutes to take too much time.

Some researchers [1, 2] investigated methods to boost the accomplishment of Dijkstra's algorithm for different applications.

As an alternative to Dijkstra's algorithm, robot path planning can be done using the A-Star algorithm. Here, the robot is having innate capacity to find the optimal collision-free path under the constraint that it can move only along X- and Y- directions which will increase the time required to reach the target point [3]. A modified A-Star algorithm has been proposed in [4], where the robot is able to move along the X, Y, and diagonal directions which will decrease the computational time and cost required to reach the goal position.

Using the methods described above the robot is smart to find the optimal collision-free path but given the same start and end locations, the robot has to calculate the path every time which will make the system inefficient in certain conditions. In our work, we will use simulation to test tactual the working of a robot in an environment filled with obstacles. In previous work [5], experiments have been brought out using webots simulator [5] to test the working of the robot using A-Star path planning. In [6], experiments are done to get the shortest path using the Hough transform in image processing. Some research is done in [7] to analyze the multipath robot environment.

2.2 Proposed System

This will implement a system in two phases. In first aspect, the precise route between begin and end spots can be found using modified A-Star algorithm. In second step, once the goal is reached the path is stored in sequential memory so that in future when the same path occurs in the robot path then there is no requirement to figure out the path again. It just retrieves the path from memory which will save the computational cost, time, and power requirements of the robot.

2.3 A* (A-Star Algorithm)

A-Star algorithm is an algorithm which is used to compute the short-lived way among given spots. A-Star method is really smart algorithm than conventional algorithms. Many web-based maps and games used this method to figure out the shortest path efficiently between two points. A-Star algorithm is called as the best first search that

it finds the shorter route between two spots by finding all the possible paths for the one incurred to be the smallest one.

A-Star algorithm selects the smallest path that satisfies the below equation:

$$a(m) = b(m) + c(m) \tag{1}$$

where

m is the final point on the route,
$b(m)$ is the computation cost of the route from starting point to the final spot on the route,
$c(m)$ is the heuristic which approximates the cost of the shortened way from spot m to the final spot, and
$a(m)$ is the total cost.

A-Star uses priority queue to select minimum cost points to expand. The priority queues are known as open list and closed list. At each step, the points with the low $a(m)$ are deleted from open list and added to the closed list, the a, b, and c values of neighbors are updated and these neighbors are included to open list. Algorithm will search until a goal point with low f cost is found. Finally, shortest path can be drawn with closed list points.

2.4 Pseudocode

OSX=Old Starting Point Of X-Coordinate
OSY=Old Starting Point Of Y-Coordinate
OEX=Old Ending Point Of X-Coordinate
OEY=Old Ending Point Of Y-Coordinate
NSX=New Starting Point Of X-Coordinate
NSY=New Starting Point Of Y-Coordinate
NEX=New Ending Point Of X-Coordinate
NEY=New Ending Point Of Y-Coordinate
NPX=Neighbor Point X-Coordinate
NPY=Neighbor Point Y-Coordinate
OL=OPEN LIST=Priority queue
CL=CLOSED LIST=Priority queue
POL=Point in OPEN LIST
PCL=Point in CLOSED LIST
CP=Current Point
PL=Parent LIST
G=Cost of path from starting point to final point on route
H=Heuristic which approximates cost of the shortened way from m to the end spot
F=Total Cost (G+H)

1. Define (OSX, OSY) and (OEX, OEY)
2. Generate map.
3. Add (OSX, OSY) to OL
4. Add (OSX, OSY) to CL
5. Add (NPX, NPY) to OL

   -If (OL !=NULL);
   -If ((NPX, NPY) !=POL));
   -If ((NPX, NPY) !=PCL));
   -Compute G, H and F values. Record CP to PL
   -If (CP==POL);
   Compute G value of CP.
   -If (G (CP)<G (POL))
   Update PL with CP

6. If (OL==NULL)

   -No Possible Path

7. If ((OEX, OEY)==POL)

   -Define Path using PL

8. Find the lowest cost (NPX, NPY). Add it to CL
9. Go to Step 5.
10. Store the path in memory.
11. Define new start and goal points.
12. Load the map area.
13. Compare new start and end points with the old start and end points.
14. If matches retrieve the path otherwise go back to point 3.
15. If (NSX<OSX && NSY<OSY) && (NEX==OEX && NEY==OEY)

    -Calculate the path from (NSX, NSY) to (OSX, OSY)
    -Retrieve path from (OSX, OSY) to (OEX, OEY)

16. If (NSX==OSX && NSY==OSY) && (NEX>OEX && NEY>OEY)

    -Calculate the path from (OEX, OEY) to (NEX, NEY)
    -Retrieve path from (OSX, OSY) to (OEX, OEY)

17. If (NSX<OSX && NSY<OSY) && (NEX>OEX && NEY>OEY)

    -Calculate the path from (NSX, NSY) to (OSX, OSY)
    -Retrieve path from (OSX, OSY) to (OEX, OEY)
    -Calculate the path from (OEX, OEY) to (NEX, NEY)

18. If (NSX>OSX && NSY>OSY) && (NEX==OEX && NEY==OEY)

    -Retrieve path from (NSX, NSY) to (OEX, OEY)

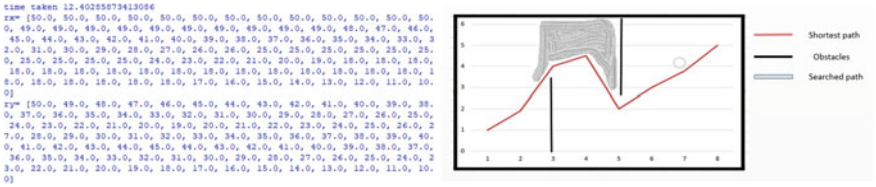


Fig. 1 *X*- and *Y*- points of the traveled path and graph representing the robot given start and end points

19. If (NSX==OSX && NSY==OSY) && (NEX<OEX && NEY<OEY)

    -Retrieve path from (OSX, OSY) to (NEX, NEY)

20. If (NSX>OSX && NSY>OSY) && (NEX<OEX && NEY<OEY)

    -Retrieve path from (NSX, NSY) to (NEX, OEY)

2.5 Time Complexity of Different Operations Used

1. Search—O (n)
2. Append—O (1)
3. Insert—O (n)
4. Delete—O (n).

3 Experiments and Results

The proposed idea of robot path planning using memory has been tested in the following scenarios which consist of obstacles in robot path and the performance is compared to the case when there will be no obstacles in the robot path. We use the simulation environment and determine the time taken to find the final location with and without our algorithm.

3.1 Test with Given Start (10,10) and End Points (50,50)

Here, the robot will find the optimal collision-free path using A-Star algorithm between begin spot (10,10) and end spot as (50,50) and the time taken to find the end spot is 12.40 s which is shown in Fig. 1.

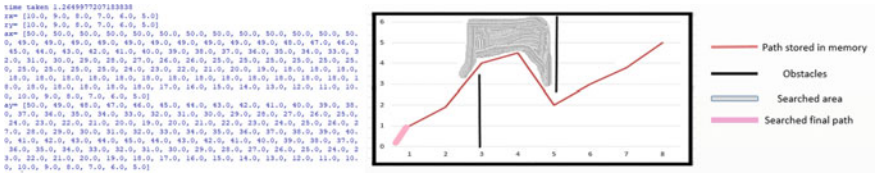


Fig. 2 *X*- and *Y*-spots of traveled path and graph representing the robot given start and end points

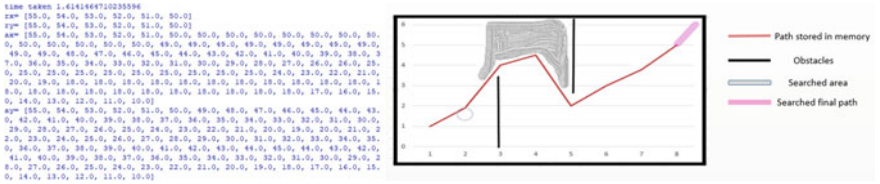


Fig. 3 *X*- and *Y*-locations of the traveled path and *X*- and *Y*-spots of the traveled path stored in memory

3.2 Test with Given Start Point (5,5) and End Point (50,50)

Here, the robot will find the optimal collision-free path using A-Star algorithm between the begin point coordinates as (5,5) and final spot (10,10) and retrieve the path from (10,10) to (50,50) from memory by comparing begin and final spots with the old start and end locations of route stored in memory and the time required to reach the final spot is 1.26 s and is shown in Fig. 2.

3.3 Test with Given Start Point (10,10) and End Point (55,55)

Here, robot will retrieve the path from (10,10) to (50,50) from memory by comparing begin and final locations with the old start and end spots of route stored in memory and will find the optimal collision-free path using A-Star algorithm between begin spot coordinates as (50,50) and end location (55,55) and the time taken to find end point is 1.61 s which is shown in Fig. 3.

3.4 Test with Given Start Point (5,5) and End Point (55,55)

Here, robot will find the optimal collision-free path using A-Star algorithm between start location coordinates as (5,5) and final location (10,10) and retrieve the path from (10,10) to (50,50) from memory by comparing start and end spots with the old begin and final spots of the route stored in memory and will find the optimal collision-free

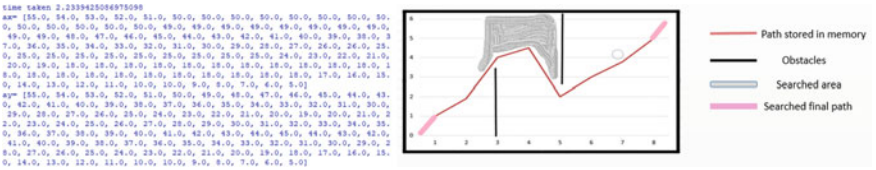


Fig. 4 X- and Y-points of the traveled path and X- and Y-spots of the traveled path stored in memory

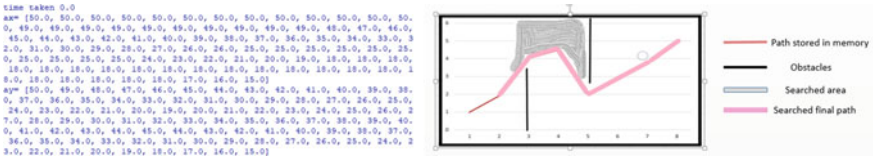


Fig. 5 X- and Y-points of the traveled path and X- and Y-locations of the traveled path stored in memory

Fig. 6 X- and Y-coordinates of the traveled path and X- and Y-coordinates of the traveled path stored in memory

path using A-Star algorithm between start location coordinates as (50,50) and end spot as (55,55) and the time taken for the goal location is 2.23 s which is shown in Fig. 4.

3.5 Test with Given Start Point (15,15) and End Point (50,50)

Here, robot will retrieve the path from (15,15) to (50,50) from memory by comparing start and end spots with the old begin and end locations of the path stored in memory and the time taken in order to figure out the goal is 0.00 s which is shown in Fig. 5.

3.6 Test with Given Start Point (10,10) and End Point (45,45)

Here, robot will retrieve the path from (10,10) to (45,45) from memory by comparing start and end points with the old begin and final spots of the route stored in memory and the time taken to find the goal spot is 0.00 s which is shown in Fig. 6.

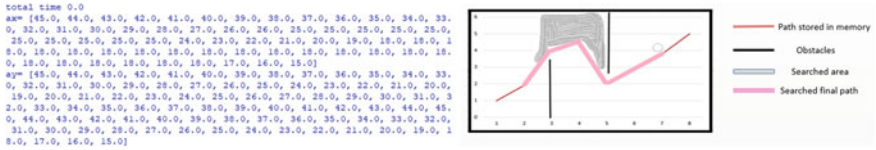


Fig. 7 *X*- and *Y*-locations of the traveled path and *X*- and *Y*-spots of the traveled path stored in memory

```
time taken= 1.280432939529419
rx= [50.0, 49.0, 48.0, 47.0, 46.0, 45.0, 44.0, 43.0, 42.0, 41.0, 40.0, 39.0, 38.
0, 37.0, 36.0, 35.0, 34.0, 33.0, 32.0, 31.0, 30.0, 29.0, 28.0, 27.0, 26.0, 25.0,
 24.0, 23.0, 22.0, 21.0, 20.0, 19.0, 18.0, 17.0, 16.0, 15.0, 14.0, 13.0, 12.0, 1
1.0, 10.0]
ry= [50.0, 49.0, 48.0, 47.0, 46.0, 45.0, 44.0, 43.0, 42.0, 41.0, 40.0, 39.0, 38.
0, 37.0, 36.0, 35.0, 34.0, 33.0, 32.0, 31.0, 30.0, 29.0, 28.0, 27.0, 26.0, 25.0,
 24.0, 23.0, 22.0, 21.0, 20.0, 19.0, 18.0, 17.0, 16.0, 15.0, 14.0, 13.0, 12.0, 1
1.0, 10.0]
```

Fig. 8 *X*- and *Y*-coordinates of the traveled path

3.7 Test with Given Start Point (15,15) and End Point (45,45)

Here, robot will retrieve the path from (15,15) to (45,45) from memory by comparing begin and end locations with the old start and end spots of the route stored in memory and the time taken to find the goal is 0.00 s which is shown in Fig. 7.

3.8 Test with Given Start Point (10,10) and End Point (50,50) with No Obstacles in Robot Path

Here, robot will find the optimal collision-free path using A-Star algorithm between begin spot coordinates as (10,10) and end location as (50,50) and the time taken to reach the goal point is 1.28 s which is shown in Fig. 8.

3.9 Test with Given Start Point (10,10) and End Point (55,55)

Here, robot will retrieve the path from (10,10) to (50,50) from memory by comparing start and end spots with the old begin and final locations of the route stored in memory and will find the optimal collision-free path using A-Star algorithm between the start point coordinates as (50,50) and end point coordinates as (55,55) and the time taken to reach the goal is 0.95 s which is shown in Fig. 9.

```
time taken= 0.9501223564147949
rx= [55.0, 54.0, 53.0, 52.0, 51.0, 50.0]
ry= [55.0, 54.0, 53.0, 52.0, 51.0, 50.0]
ax= [55.0, 54.0, 53.0, 52.0, 51.0, 50.0, 49.0, 48.0, 47.0, 46.0, 45.0, 44.0, 43.
0, 42.0, 41.0, 40.0, 39.0, 38.0, 37.0, 36.0, 35.0, 34.0, 33.0, 32.0, 31.0, 30.0,
 29.0, 28.0, 27.0, 26.0, 25.0, 24.0, 23.0, 22.0, 21.0, 20.0, 19.0, 18.0, 17.0, 1
6.0, 15.0, 14.0, 13.0, 12.0, 11.0, 10.0]
ay= [55.0, 54.0, 53.0, 52.0, 51.0, 50.0, 49.0, 48.0, 47.0, 46.0, 45.0, 44.0, 43.
0, 42.0, 41.0, 40.0, 39.0, 38.0, 37.0, 36.0, 35.0, 34.0, 33.0, 32.0, 31.0, 30.0,
 29.0, 28.0, 27.0, 26.0, 25.0, 24.0, 23.0, 22.0, 21.0, 20.0, 19.0, 18.0, 17.0, 1
6.0, 15.0, 14.0, 13.0, 12.0, 11.0, 10.0]
```

Fig. 9 *X*- and *Y*-spots of the traveled path

4 Conclusion

From the above experiments, it is understood that the present proposed idea is producing better results in case of obstacle environment when compared with the obstacle-free environment.

5 Future Work

The algorithm can be enhanced to produce better results in case of obstacle-free environment and works in case of a dynamic environment where obstacles are in motion. Retrieving data from memory can be made faster using cache memory and instead of storing full path, we can store only the path fragments from the grid cells which are used very often to improve the performance.

References

1. Liu YQ, Huang QL, Zhang YX, Luan GF, Xu MH (2007) An improved Dijkstra's shortest path algorithm for sparse network. Appl Math Comput 185(1):247–254
2. Fan DK, Shi P (2010) Improvement of Dijkstra's algorithm and its application in route planning. In: Proceeding of 7th international conference on fuzzy systems and knowledge discovery (FSKD), Yantai, Shandong, 2010, pp 1901–1904
3. Jeddisaravi K, Alitappeh RJ, Guimari'ies FG (2016) Multi-objective mobile robot path planning based on A* search. In: 6th International conference on computer and knowledge engineering (ICCKE 2016), 20–21 October 2016, Ferdowsi University of Mashhad
4. Duchon F, Babineca A, Kajana M, Beoa P, Floreka M, Ficoa T, Jurišicaa L (2014) Path planning with modified A star algorithm for a mobile robot. In: Modelling of mechanical and mechatronic systems MMaMS, 2014
5. Vargas GA, Rubiano OG, Castillo RA, Avilés OF, Mauledoux MF (2016) Simulation of e-puck path planning in Webots. Int J Appl Eng Res 11:9772–9775. ISSN 0973-4562
6. Li C, Li Y, Zhou F, Fan C (2007) Multi-objective path planning for the mobile robot. In: Proceedings of the IEEE international conference on automation and logistics, 18–21 August 2007, Jinan, China

7. Radhakrishnan G, Gupta D, Sindhuula S, Khokhawat S, Sudarshan TSB (2015) Experimentation and analysis of time series data from multi-path robotic environment. In: 2015 IEEE International conference on electronics, computing and communication technologies (CONECCT), 2015, pp 1–6

A Comparative Analysis of Breadth First Search Approach in Mining Frequent Itemsets

M. Sinthuja, N. Puviarasan and P. Aruna

Abstract The ensuing paper contrasts the algorithms using breadth first search technique of mining frequent itemset. The Apriori algorithm is a well-known algorithm which adopts breadth first search approach to mine frequent itemsets. The process of finding the frequent itemsets is simple for smaller database and complex for larger databases. Candidate itemset generation is a major disadvantage for Apriori algorithm. In this research paper, the proposed BFS-LP-Growth algorithm is compared with another breadth first search algorithm Apriori. The proposed BFS-LP-Growth algorithm overcomes the limitation of Apriori algorithm by processing the database twice and also avoids the exploration of candidate itemset. The structure of LP-Growth algorithm makes the proposed BFS-LP-Growth algorithm more advantageous. The breadth first search enhances the performance of the proposed BFS-LP-Growth algorithm by creating subheader table which avoids the necessity of conditional pattern base and conditional LP-tree. Standard databases used for comparison are Chess and Mushroom. Based on the results, we find that the proposed BFS-LP-Growth algorithm is more worthwhile than the Apriori algorithms as execution time is low and the need for candidate patterns is diminished.

Keywords Association rule mining · Depth first search approach · Data mining · Frequent itemset mining · Linear tree · Minimum support · Pruning · Breadth first search approach

M. Sinthuja (✉) · P. Aruna
Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India
e-mail: sinthujamuthu@gmail.com

P. Aruna
e-mail: arunapuvi@yahoo.co.in

N. Puviarasan
Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India
e-mail: npuvi2410@yahoo.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_46

1 Introduction

Data mining is the phenomenon of consolidating data into information. Association rule mining is a vital section in the domain of data mining [1, 2]. Numerous techniques have been experimented for mining association rules in the research studies [3, 4]. Apriori algorithm is the most favored algorithm among association rule mining techniques, which explores candidate patterns [3, 5, 6]. FP-Growth algorithm is remarkable among frequent pattern mining. It constructs a frequent pattern tree called FP-tree [7–9].

The LP-Growth algorithm which adopts linear structure was introduced [3, 10, 11]. The LP-Growth algorithm derives advantage from its structure. The algorithm limitation in mining frequent itemsets—usage of bottom-up approach which needs to construct conditional pattern base at every stage and conditional LP-tree for each itemset.

2 Comparison of Various Frequent Itemset Mining Techniques

The following theory mainly concentrates on mining frequent patterns. In this section, the breadth first search algorithms of Apriori is compared with the proposed breadth first search LP-Growth algorithm (BFS-LP-Growth). On comparison of these algorithms, it is derived that the proposed BFS-LP-Growth is more efficient than Apriori algorithm.

2.1 Apriori Algorithm

Apriori algorithm [5, 12, 13] uses a simple technique to explore all frequent itemsets from the database. Apriori is an iterative technique known as level-wise search [10]. This algorithm uses breadth first search (BFS). The iterative method exposes frequent patterns. Pruning is the underlying concept of the above method that outstages the less frequent itemsets.

Definition 1 It states that any subset of frequent items must be most occurring.

In Table 1, Minimum support count of 2 is considered for this example. The working pattern of the algorithm is as follows as portrayed in Fig. 1:

Step 1: Initially, size of the transaction is inserted into the transaction database.
Step 2: The algorithm scans the database to count the frequency of each item. In the primary iteration, individual item is a component of candidate 1-itemset, C1.

Table 1 Transaction database

TID	Items	SOT
100	f, c, a, d	4
200	f, c, a, g, e	5
300	f, c, b, a, h	5
400	f, b, g	3
500	c, b	2

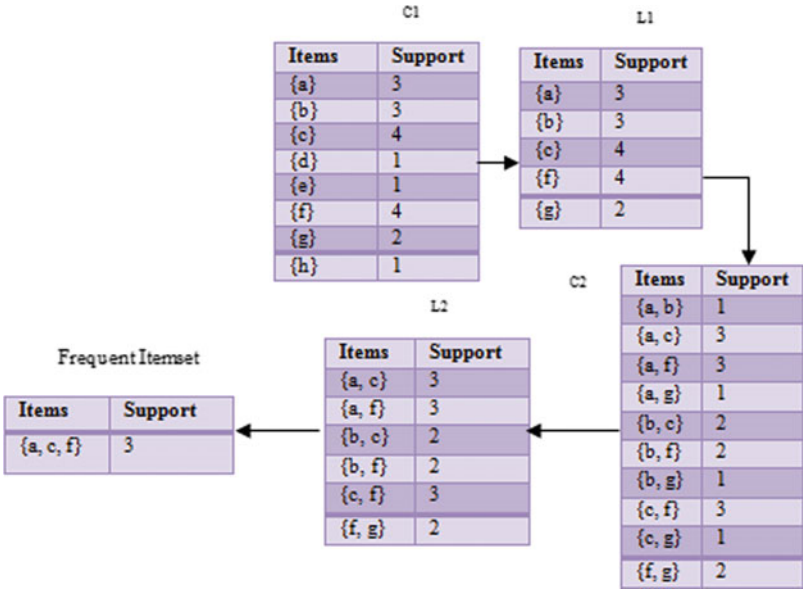


Fig. 1 Steps to derive frequent itemsets using Apriori

Step 3: This algorithm will create numerous frequent items. It is called the Size Of Transaction (SOT).

Step 4: As minimum support is 2, items with count lower than 2 is deleted in C1. Thus set of frequent 1-itemset, L1 is found. It contains the candidate 1-itemset, satisfying the minimum support threshold.

Step 5: Join step is performed in C2 combining each item with other item and finding the support of individual itemset which is a candidate 2-itemset.

Step 6: As the minimum support is 2, items with count lower than 2 is deleted. Now C2 is pruned. Thus set of frequent 2-itemset, L2 can be determined.

Step 7: The algorithm will terminate and provide all the frequent itemsets as there is no iteration.

Apriori algorithm has to undergo multiple times of database scan. Apriori algorithm has limitation of overconsumption of time and memory and its rate of accuracy diminishes with early pruning of less occurring items.

2.2 *Proposed LP-Growth Algorithm Based on Breadth First Search Approach (BFS-LP-Growth)*

The proposed LP-Growth algorithm is based on breadth first search (BFS-LP-Growth) which overcomes the limitation of Apriori algorithm. In the proposed BFS-LP-Growth algorithm, two scans of the database are recommended. By using breadth firsts search approach, conditional pattern base and conditional LP-tree is avoided. Thus, the consumption of runtime is reduced. As this algorithm is based on tree, generation of candidate itemset is avoided. It leads to lower usage of memory and runtime. The proposed BFS-LP-Growth algorithm takes advantage of linear prefix tree structure. The structure of LP-tree is in array forms to avoid pointers between nodes. Thus, for the generation of the frequent itemsets, LP-tree is highly effective compared to the conventional methods. The main advantage of LP-tree is that separate arrays are used to store each transaction. They explore the simplicity of array while inserting transactions and traversing through LP-tree. The algorithm proceeds as follows as portrayed in Fig. 2.

Transaction database is shown in Table 2, which contains the information for five transactions. Initially, the database is scanned and the support of each item is found. The items prioritized are depicted in Table 3, which is arranged in descending order. In this example, the minimum support is denoted as greater than or equal to two. The items lower than minimum support threshold value is pruned. Remaining items

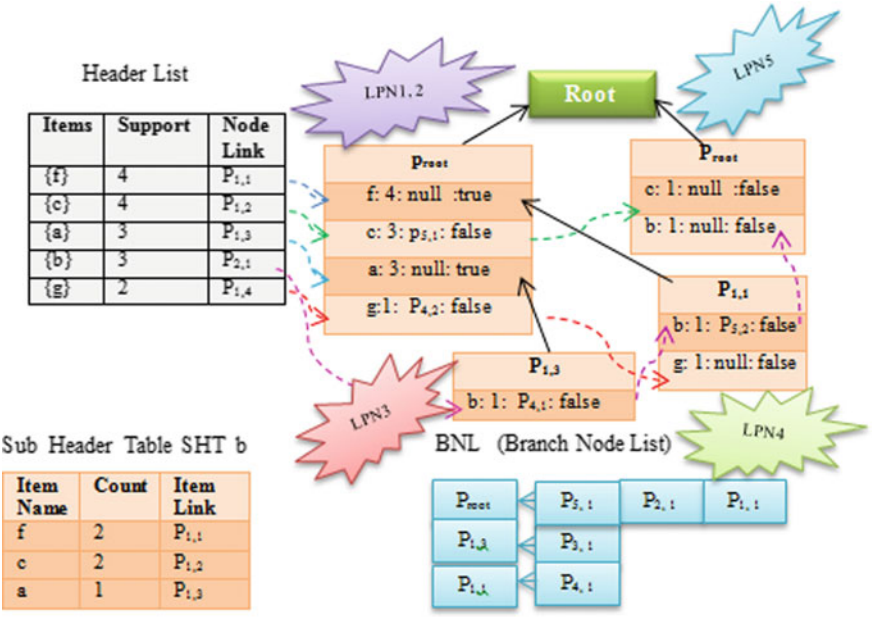


Fig. 2 Construction of BFS-LP-Growth

Table 2 Transaction database

TID	Item bought	Ordered item
100	d, c, a, f	f, c, a
200	g, c, a, f, e	f, c, a, g
300	b, a, c, f, h	f, c, a, b
400	g, b, f	f, b, g
500	c, b	c, b

Table 3 Prioritize the items

Items	Support
{f}	4
{c}	4
{a}	3
{b}	3
{g}	2
{d}	1
{e}	1
{h}	1

are ordered in support descending order {f: 4, c: 4, a: 3, b: 3, g: 2}. Based on this prioritization, the items are ordered in Table 2. The tree always begins with null node. The topmost transaction which contains the itemset {f, c, a} is started with, into the LP-tree based on the length of the transaction. The length of first transaction is three so four internal nodes are created which includes the header node. As this transaction is linked from root node, the header node is updated to P_{root}. The item "f" is infused into LP-tree with count 1, as there is no link to the corresponding node, it is stated as "null" and the flag is stated as "false" as it does not have any branch information. Likewise, remaining items are inserted into LP-tree. Simultaneously, the branch node is updated in the list as P_{root} linked to $P_{1,1}$. Likewise, remaining transactions are portrayed in Fig. 2.

The process of breadth first search approach is described as follows: breadth first search approach begins with the initial node of the header list as depicted in Fig. 2. "Item_f" is the initial item that presents in the header list. By walking through the path of "Item_f" in the header list denoted as HL it is found that there is no path in the LP-tree. As an "Item_f" appears at the leading of the tree. The item "f" is declared as frequent item with count 4. By walking through the path of "Item_c" two paths are found. At the time of walk up, link up all scanned nodes of the same item by an item link and aggregate the count for such nodes. The first path is "Root-f-c". As path is found subheader table has to be created for item "c" named as SHT_c. The item "f" is added into SHT_c with count 3 as this path occurs thrice in the database. The node link is denoted as $P_{1,1}$ in SHT_c. In the second path, there is no prefix path for "Item_c". At the end, item {f, c: 3} is the frequent itemset as it satisfies the condition of minimum support value. Likewise, the remaining items are inserted using breadth

first search approach. The frequent itemsets are {f, c: 3} {f, c, a: 3} {f, b: 2} {c, b: 2} and {f, g: 2}.

3 Performance Evaluation

In this paper, experimental analysis were performed to explore frequent itemsets. Comparative studies were conducted on two breadth first search algorithms of Apriori and proposed BFS-LP-Growth algorithm. In order to prove that the experiments are reasonable, evaluation was based on two criteria: execution time and memory usage.

3.1 Test Environment

The experiments are conducted on Intel® corei3™ CPU with 2.13 GHz, and 2 GB of RAM computer. The algorithms have been implementation in Java.

3.2 Dataset Description

In this paper, two standard databases were used from different applications. These datasets have been acquired from Tunedit Machine Learning Repository. Each database contains different instances and attributes. The analysis is performed on two databases of Chess and Mushroom which are widely used for evaluating frequent pattern mining algorithms.

Execution time

In this subsection, the proposed BFS-LP-Growth algorithm is compared with Apriori algorithm. From the Figs. 3 and 4, the latitudinal axis shows various minimum support threshold values and longitudinal axis shows runtime in milliseconds. In Fig. 3, the database Chess is tested with Apriori and the proposed BFS-LP-Growth algorithm with different minimum support values. Scanning of the transaction database is done multiple times to mine frequent itemset in Apriori algorithm. It adopts join and prune property to mine frequent itemset which results in generating candidate itemset. Thus, consumption of time is more for the algorithm Apriori in generating frequent itemset. Comparatively, the proposed BFS-LP-Growth algorithm performs better than the Apriori algorithm as the former avoids candidate itemset generation in mining frequent itemsets.

With reference to Fig. 4, the dataset Mushroom is tested with the algorithms Apriori and the proposed BFS-LP-Growth algorithm with distinct support threshold values. The processing time of Apriori algorithm is more in exploring candidate

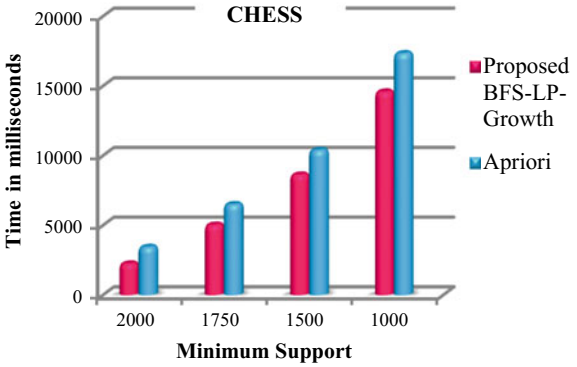


Fig. 3 Execution time of FIs using Chess dataset

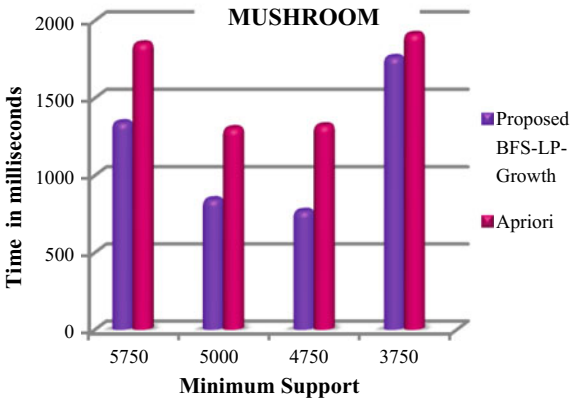


Fig. 4 Execution time of FIs using Mushroom dataset

itemset for each iteration as described above. The algorithm has to generate lot of candidate itemset for each iteration which is time-consuming in mining frequent itemsets for lower minimum support. Comparably, the proposed BFS-LP-Growth algorithm outperforms the existing Apriori algorithm in terms of execution time for both higher and lower minimum support because of the quick traversing of the tree and its linear structure. Thus, it is clearly shown that the runtime performance of BFS-LP-Growth is better than Apriori algorithm.

Memory usage

The experimental results of consumption of memory are portrayed in Figs. 5 and 6. The memory consumption of proposed BFS-LP-Growth and Apriori algorithms for mining frequent itemsets are plotted for varying minimum support threshold values of 1000, 1500, 1750, and 2000 for the database Chess in Fig. 5. It shows that Apriori does a considerable amount of work to deal with a large number of candidate frequent itemsets. Thus memory utilization to store candidate frequent itemset is also more. Meanwhile, the proposed BFS-LP-Growth algorithm requires lower consumption of

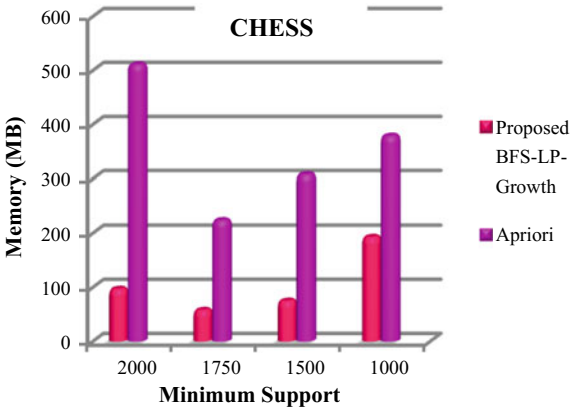


Fig. 5 Memory usage of FIs using Chess dataset

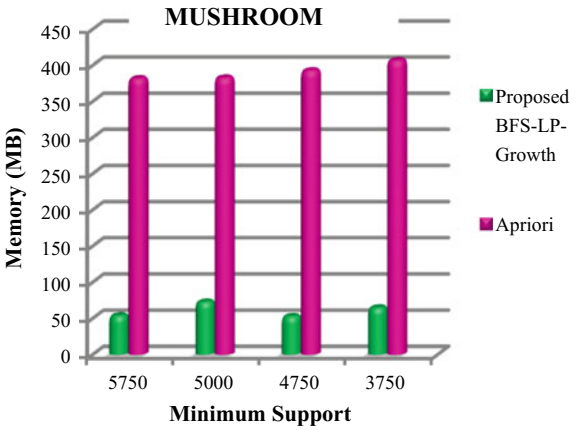


Fig. 6 Memory usage of FIs using Mushroom dataset

memory due to minimal usage of pointer between nodes. The main advantage in using breadth first search approach is that it does not generate conditional pattern base and conditional LP-tree.

The comparison ratio of proposed BFS-LP-Growth and Apriori algorithm for the database Mushroom is reported in Fig. 6. Consumption of memory for Apriori algorithm increases rapidly for both higher and lower minimum support. Meanwhile, the proposed BFS-LP-Growth algorithm consumes lower memory space when compared to Apriori algorithm. For the minimum support threshold value of 5000 the proposed BFS-LP-Growth consumes 73 MB and Apriori consumes 383 MB. The proposed BFS-LP-Growth algorithm generates only subheader table for mining frequent itemset. Specifically, when the minimum support is high the usage of memory for the proposed BFS-LP-Growth algorithms is low. Because the generation of frequent itemset is reduced. In case of Apriori algorithm, consumption of memory is

more when compared to proposed BFS-LP-Growth algorithm. Thus, the proposed algorithm is best for all cases when compared to the Apriori algorithm.

4 Conclusion

In this research paper, breadth first search algorithms of Apriori and the proposed BFS-LP-Growth algorithm were compared. In the process of mining, proposed BFS-LP-Growth is more efficient than Apriori algorithm with respect to long itemsets. The Apriori algorithm needs multiple scans of transactional database as it explores candidate itemsets. It results in huge amount of execution time and memory. The frequency of database scan for the proposed BFS-LP-Growth algorithm is only twice and also avoids the candidate itemset generation which results in lower consumption of time and memory. Here, breadth first search algorithm plays a significant role in LP-Growth algorithm which avoids the creation of conditional pattern base and conditional LP-tree. Thus, it gives guarantee of both faster runtime and lower memory usage when compared to Apriori algorithm.

Acknowledgements We gratefully acknowledge the funding agency, the University Grant Commission (UGC) of the Government of India, for providing financial support for doing this research work.

References

1. Agarwal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. ACM-SIGMOD, pp 207–216
2. Agarwal R, Srikanth R (1995) Mining sequential patterns. ICDE
3. Pyun G, Yun U, Ryu KH (2014) Efficient frequent pattern mining based on Linear Prefix tree. Knowl-Based Syst 55:125–139
4. Tanbeer SK, Ahmed CF, Jeong BS, Lee Y (2008) Efficient single pass frequent pattern mining using a prefix-tree. Inf Sci 179(5):559–583
5. Sinthuja M, Puviarasan N, Aruna P (2016) Experimental evaluation of Apriori and equivalence class clustering and bottom up lattice traversal (ECLAT) algorithms. Pak J Biotechnol 13(2):77–82
6. Wang K, He Y, Han J (2000) Mining frequent pattern using support constraints. VLDB, pp 43–52
7. Lin K-C, Liao I-E, Chen Z-S (2011) An improved frequent pattern growth method for mining association rules. Expert Syst Appl 38:5154–5161
8. Sinthuja M, Puviarasan N, Aruna P (2017) Comparative analysis of association rule mining algorithms in mining frequent patterns. Int J Adv Res Comput Sci 8(5)
9. Sinthuja M, Puviarasan N, Aruna P (2018) Research of improved FP-growth (IFP) algorithm in association rules mining. Int J Eng Sci Invent 24–31
10. Mannila H, Toivonen H (1997) Level wise search and borders of theories in knowledge discovery. Data Min Knowl Disc 1(3):241–258

11. Sinthuja M, Puviarasan N, Aruna P (2018) Mining frequent Itemsets Using Top Down Approach Based on Linear Prefix tree, Springer. Lecture Notes on Data Eng Commun Technol 15:23–32
12. Agrawal R, Srikant R (1994) Fast algorithm for mining association rules in large databases. In: Proceedings of the 1994 international conference VLDB, Santiago, Chile, pp 487–499
13. Sinthuja M, Puviarasan N, Aruna P (2017) Evaluating the performance of association rule mining algorithms. World Appl Sci J 35(1):43–53. ISSN 1818-4952

Email Authorship Attribution

Suman Patil, Sandeep Varma Nadimpalli and Pavithra K. Yadav

Abstract Email correspondence is regularly manhandled for directing social designing assaults including spamming, phishing, data fraud, and circulating malware. This is to a great extent credited to the issue of obscurity intrinsic. Finding the authorship of email which can be stated as attribution problem is contemplated as content classification issue where the styles of writing of people are displayed in view of their already composed documents. In this paper, Multiple Association Rules for Authorship Attribution (CMARAA) is proposed to solve the problem of authorship attribution. The proposed method makes the remarkable composition style of features of a person. The experimental evaluation shows that CMARAA classifies with 92% of accuracy being the most accurate algorithm when compared to other classification algorithms for different combinations of authors.

Keywords Authorship attribution · Email classification · Association mining · Stylometric features

1 Introduction

Data security is increasing great consideration from specialists in the group particularly after the developing infiltration of the e-based frameworks and e-data that scale everywhere around the world [1]. The vast majority of the data is presently gathered, handled, and put away on electronic PCs and transmitted crosswise over systems to different PCs. Sadly, alongside the favorable circumstances, the arbor of the web is regularly abused in different routes for illicit purposes. Unmistakably, without any

S. Patil (✉) · S. V. Nadimpalli · P. K. Yadav
Department of Information Science and Engineering, BMS College of Engineering, Bengaluru, Karnataka, India
e-mail: sumanpatil2012@gmail.com

S. V. Nadimpalli
e-mail: sandeepvarma.ise@bmsce.ac.in

P. K. Yadav
e-mail: pavithrakyadav123@gmail.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_47

authentication, the most widely recognized approach to transmit data over the web is sending emails. Emails that are spam, for example, are not anymore simply unsolicited emails. Digital offenders utilize spam emails to spread the virus over the web to allure individuals to phishing destinations that take crucial individual data. Besides assaults might be the type of email mishandle, for example, inappropriate behavior, phishing, transmitting worms and viruses, forgery, tricks, youngster erotica, phony, email bombing, and email infection.

1.1 Research Problem

The project intends to naturally and proficiently demonstrate the writeprint or the patterns of individual speculate. At that point, it is possible to present such patterns as proof distinguishing the creator of the pernicious email.

Given S is a chance to be the gathering of authors that are suspected with an unknown email E. Given E_a is a chance to be a moderately extensive accumulation of messages written by presuming $S_a \in S$. Given V is a chance to be an arrangement of words that are distinct in $\bigcup E_a$. The issue of creation identification means identifying the creator S_1 from the speculates S who is the most conceivable, whose email gathering E_1 mostly coordinated with the patterns of Stylometric including things in the malicious email E. Naturally, an email accumulation E_a matches E if E_a and E share comparative patterns of composing style including a strongly illustrative blends.

As far as associate classification is concerned, what precisely is a writeprint? In particular, there is a need to extract rules got from patterns that unequivocally and interestingly speak of the composition style of each presume suspect S_a yet do not exemplify the written style of composition of some other suspect S_b, where $a \neq b$.

2 Related Work

Attribution of authorship is contemplated as a content arrangement and characterization issue in the writing [2]. By and large, an arrangement shown is manufactured utilizing the beforehand composed archives of the presumed creators. The names of creators are utilized for naming the classes during the testing and training procedures for developing the model. Not at all like and two-class [3] and one-class [4] characterization issue, which is considered to verify authorship, present-day attribution of origin, and by perusing Stamatato's review [5] can better comprehend this; can be drawn closer than a multi-class grouping issue.

The single standard predefined set of highlights is not there, which separates the individual scholar's written work style, however, a few examinations [6] have distinguished the most illustrative highlights as far as precisely characterizing mysterious or debated writings. Accentuation and n-gram highlights have turned out to be exceptionally illustrative all alone. One exhaustive investigation on complex highlights displayed by [7] talks about highlights with adequate detail. Initiation

examination has been very effective in settling creation attribution arguments about different kinds of compositions [8]. Nonetheless, email origin attribution postures uncommon difficulties because of its qualities of size, vocabulary, and synthesis when contrasted with artistic works [9, 10].

Artistic records have a clear semantic and syntactic structure and are normally extensive in the estimate, involving no less than a few passages. Interestingly, messages are short and for the most part, do not take after very much characterized sentence structure or syntax rules. In this way, displaying the composition examples of their creator is very hard [11]. Attribution of initiation may not be exceptionally exact for writings involving fewer than 500 words, making the requirement of a model, which is more better [12] ready to deal with the qualities inborn in messages.

The attribution of email initiation issue may be tackled by outlining calculations that arrange with the particular difficulties identified with email creation investigation which is demonstrated by [13]. The Support Vector Machine (SVM), which is prominent arrangement strategy was connected [2, 14] above a basic also expressive highlights arrangement for email origin. It is found by [15, 9, 10] additionally that when the quantity of words utilized expanded from 122 to 320, the execution of SVM was reduced, negating precept SVM underpins high dimensionality and prompting the conclusion, which says expanding the quantity of highlights do not enhance precision [12]. As of late, [16] proposed a deliberate way to deal with picture writeprints by coordinating syntactic and *n*-gram words highlights between the unknown email and test email.

3 Methodology

Authorship examination incorporates authorship attribution, confirmation, profiling, and additionally similitude recognition. In the proposed approach, distinctive sorts of features are extracted and a few machine learning calculations are utilized. The past examinations demonstrate that the predefined uncommon list of the set of features is not there which can be utilized to decide the composition style. The style of composition contains: Lexical highlights, Syntactic highlights, Content-particular highlights, Structural highlights, and Idiosyncratic highlights. The agent catches the style of composition of each speculate by obtaining the features that are frequent of his works. The extricated highlights can be used to produce a model of classification. At last, the conceivable features contained in the noxious email are extricated and encourage the model of the classifier to distinguish the work style of writing of presume whose writing style intently coordinates the email under scrutiny.

3.1 Dataset Preparation

The informational index utilized here, is the Enron email accumulation, which is part into a preparation set and a test set containing 40 sends and 20 sends individually,

separated similarly among inbox and sent things messages. You will effortlessly perceive the sends as it contains inbox and sent things in its filename. In any content mining issue, content cleaning is the initial step where those words from the record which may not add to the data are expelled. Messages may contain a great deal of bothersome characters like accentuation marks, stop words, digits, and so on which may not be useful in identifying the spam email. The messages can be preprocessed in the accompanying ways: (i) Removal of stop words and (ii) Lemmatization.

3.2 Classification of Emails—CMARAA

The algorithm exhibits the presented origin recognizable proof procedure for deciding the initiation of a pernicious email E from the predetermined speculates. An arrangement of preparing messages for all the conceivable presumes S_1, S_2, and S_k is essentially for the training procedure.

Algorithm: Multiple Association Rules Authorship Attribution (CMARAA)

Input: A large collection of preparing emails $\{E_{1a}, E_{1b}, \ldots\ldots, E_n\}$, $\{E_{2a}, E_{2b}, E_{2c}\}$, ..., $\{E_{ka}, E_{kb}, E_{kn}\}$ beforehand composed by potential presumes S_1, S_2 and S_k, respectively.
A noxious email E written by suspect.
Output: Recognize the author of the noxious email E
/* Preparing*/
for each speculate *Si* **do**
  for each email $\{E_1, .., E_n\}$ **do**
    Obtain patterns that are frequent of powerful features
  end for
  Give the obtained highlights as input to a few supervised machine learning calculations.
end for
A few models of classification are made
/* Find the creator of the noxious email *E* from S_1, S_2 and S_k */
Obtain a few powerful highlights from E.
Give the obtained highlights as input to the classification model which is created
if recognizable proof score ≥ limit defined in advance
  Recognize the speculate having noteworthy score.
Else
  Classification is not there 'The email E is not coordinated with any of the speculates style of composition'.
end if

4 Results and Discussion

CMARAA is assessed against some outstanding classifiers like Ensemble of Nested Dichotomies (END), Decision Trees, Bayesian Networks (BayesNet), and Naïve Bayes. It demonstrates that CMARAA coordinates or outflanks alternate strategies incorporating CBA and AM regarding average precision. The experiment was performed on a 2.2 GHz Intel(R) Core(TM) 2 Duo CPU with 2.00 GB, running on Windows. The email accumulations utilized as a part of this examination are all from the Enron email informational collection which is the most appropriate gathering of freely accessible emails.

Figure 1 and Table 1 indicate normal order correctness's over arrangements of accumulations of messages for 2–10 creators for every calculation tried in this investigation. Exactness's range from 31 to 91% with CMAR being the most precise calculations for Authorship Attribution (CMARAA), Classification by Association (CBA), Classification by Multiple Association Rule (CMAR), and AuthorMiner (AM), individually. As the quantity of creators increments correctness diminish, with CMAR showing the greatest diminish when there are 9 creators.

For instance, if the help limit is 10% and there are 10 creators with 100 messages each, at that point an element exceptional to one creator would need to show up in

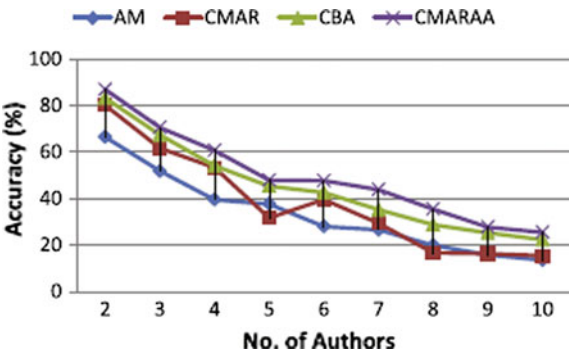


Fig. 1 Accuracy of algorithms and number of creators

Table 1 Authors values and accuracy of algorithms

Authors	AM	CMAR	CBA	CMARA
2	66.489	80.106	83.229	86.923
3	51.83	61.63	67.37	70.54
4	39.52	53.33	54.12	60.73
5	37.602	31.636	45.34	47.9
6	28	39.535	42.555	47.69
7	26.5275	29.37	35.2075	43.86
8	19.8425	16.66	28.705	35.44
10	15.78	16.32	25.216	27.69

each and every email of that creator so as to be viewed as incessant. One of the primary commitments of this examination, executed in CMARAA, diminishes the help limit to consider just a single creator's preparation set, permitting highlight things that are a visit, regardless of whether one of a kind to one creator, to be considered.

5 Conclusions

Cybercrime examinations need a writeprint demonstrating calculation which will give dependable proof to help attribution of initiation. Utilizing a promising information mining procedure named classification based on multiple association rules for authorship attribution on the email authorship attribution issue is addressed through this research work. Our proposed algorithm proved the accuracy of classification toward authorship attribution.

Acknowledgements The authors would like to acknowledge and thank technical education quality improvement program TEQIP phase III, BMS College of Engineering for funding this research work.

References

1. Casey E (2010) Handbook of digital forensics and investigation. Elsevier
2. de Vel O (2000) Mining e-mail authorship. KDD workshop on text mining
3. Iqbal F, Khan LA, Fung BCM, Debbabi M (2010) E-mail authorship verification for forensic investigation. In: Proceedings of the 25th ACM SIGAPP symposium on applied computing (SAC), ACM Press
4. Koppel M, Schler J, Argamon S, Messeri E (2006) Authorship attribution with thousands of candidate authors. In: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval
5. Stamatatos E (2009) A survey of modern authorship attribution methods. J Am Soc Inf Sci Technol (JASIST)
6. Grieve J (2007) Quantitative authorship attribution: an evaluation of techniques. Lit Linguist Comput
7. Abbasi A, Chen H (2008) Writeprints: a stylometric approach to identity-level identification and similarity detection in cyberspace. ACM Trans Inf Syst 26(2):7
8. Mendenhall TC (1887) The characteristic curves of composition. JSTOR
9. de Vel O, Anderson A, Corney M, Mohay G (2001) Multi-topic e-mail authorship attribution forensics. In: Proceedings of ACM conference on computer security e workshop on data mining for security
10. de Vel O, Anderson A, Corney M, Mohay G (2001) Mining e-mail content for author identification forensics. SIGMOD
11. Ledger GR, Merriam TVN (1994) Shakespeare, fletcher, and the two noble kinsmen. Lit Linguist Comput
12. Iqbal F, Binsalleeh H, Fung BCM, Debbabi M (2010) Mining writeprints from anonymous e-mails for forensic investigation. Digit Investing
13. Iqbal F, Binsalleeh H, Fung BCM, Debbabi M (2013) A unified data mining solution for authorship analysis in anonymous textual communications. Inf Sci Special Issue Data Min Inf Secure

14. Teng G, Lai M, Ma J, Li Y (2004) E-mail authorship mining based on SVM for computer forensic. In: Proceedings of the 3rd international conference on machine learning and cyhemetics
15. Corney M, de Vel O, Anderson A, Mohay G (2002) Gender-preferential text mining of e-mail discourse. In: Proceedings of the 18th annual computer security applications conference (ACSAC)
16. Ding SHH, Fung BCM, Debbabi M (2015) A visualizable evidence-driven approach for authorship attribution. ACM Trans Inf Syst Secur (TISSEC)

A Security-Aware Trust-Based SLA Framework Using SDN in Cloud Environment

H. M. Anitha and P. Jayarekha

Abstract Software-defined network is an emerging technology in the IT industry. Cloud computing provides the services to the users based on the requirements. Service-level agreement is signed between user and cloud provider. SLA compliance is an important feature in cloud environment. Trust-based SLA is considered for the system. Monitoring of activities is carried out using SDN controller. SDN controller gathers the information completely about cloud provider and cloud user. The main idea of this paper is to monitor and notify any unusual activities such as attacks and hence, trust-based SLA framework is proposed. The framework strongly monitors the malicious users by SDN controller through OpenFlow switches. If the user is found to be unreliable, resources allocated will be released and reallocated to another user. Hence, security monitoring is carried out using SDN controller.

Keywords Service-level agreement · Trust · Software-defined network · SDN controller · OpenFlow switch

1 Introduction

Cloud computing has been used by a wide range of applications since many years. Cloud computing basically provides three different types of services to users such as software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). SaaS provides complete software services to the users. PaaS offers platforms for computations using underlying hardware provided by the IaaS. Infrastructure can be used by the users for computations through virtual machines. IaaS is used by SaaS and PaaS. Virtual machines are deployed to the users based on their

H. M. Anitha (✉) · P. Jayarekha
Department of Information Science and Engineering, BMS College of Engineering, Bangalore 560019, India
e-mail: anithahm.ise@bmsce.ac.in

P. Jayarekha
e-mail: jayarekha.ise@bmsce.ac.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_48

demand-matching capability requirements. Capability requirements are computations, storage, and one more parameter is also considered here, i.e., security. All the requirements are mentioned in the service-level agreement (SLA). Conformance of the cloud provider toward the agreement is a very important feature. Hence, there is a requirement for the intelligent framework, which will take care of security feature and other QoS features assured by the cloud provider. The security issues, which are to be addressed are attacks from the malicious VMs, unauthorized access of users data, and DOS attacks. Software-defined network can be used to monitor all activities in the virtualized network.

SDN is emerging as one of the promising approaches for latest applications. SDN is different from the traditional networks. Here, there are two planes namely control plane and data plane [1]. Control plane is the main unit, which controls all the actions and data plane follows all the rules enforced by the control plane. Control plane is present in the SDN controller. Controller is equipped with all intelligence to monitor activities in the virtualized network. The interaction between the control plane and data plane takes place with OpenFlow. OpenFlow is the protocol used for communication between control plane and data plane. Northbound APIs are used for communication between controller and applications, southbound APIs are used for communication between the controller and switches. Cloud provider receives the requests from the users for resources such as virtual machines. Requests are passed from the cloud provider to the controller, which takes decision to allot the VMs to the users. Controller maintains the forwarding table of all the networking information. If the enough resources are available, the users are granted with required resources. User's authentication information is checked in the controller. Controller maintains the forwarding table of all the details of users and cloud providers. Unauthorized users are not granted any resources.

1.1 Traditional Versus Software-Defined Network

Traditional Network: In the traditional network, both control plane and data plane are coupled together which is shown in Fig. 1. Traditional networks [2] specify their requirements by negotiations and manual allocation of resources is carried out. They do not have methods to present a full range of user requirements such as reliability, jitter, throughput, and availability of resources. Normally traditional networks will assume user requirements and offer the resources along with cost specification.

Software-defined network: SDN controls everything with controller and collects information from the switches as shown in Fig. 2. Switches are dumb and pass on whatever information is given by the SDN controller. As all the activities are managed by the controller, decisions are taken by the controller. SDN architecture consists of three planes, namely, Application plane, Control plane, and Data plane. Northbound APIs are used for communication between application plane, control plane, and southbound APIs are used for communication between control plane and data plane.

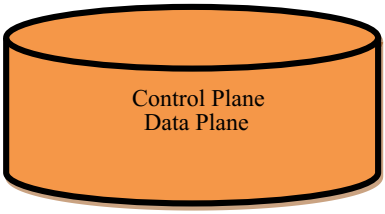


Fig. 1 Traditional network

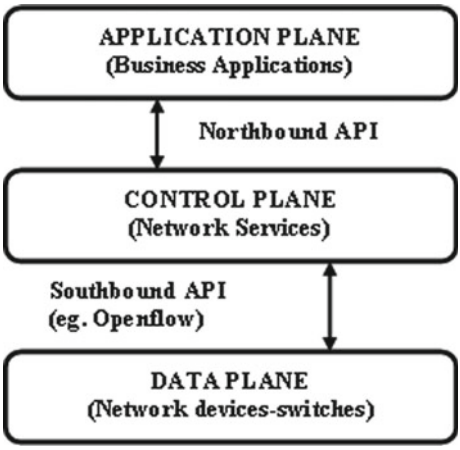


Fig. 2 Software-defined network architecture

SDN in cloud: The properties of SDN pave way for advantages in dynamically changing requirements of users for resources. Cloud computing supports multitenancy. SDN controller operates centrally so that it can dynamically update any changes with respect to the requests. Control decisions [2] are made timely and centrally regarding the resource availability, scheduling, and communication toward lower level data plane.

The main idea of the paper is to take the benefits of SDN in cloud monitoring and provides better security for the virtualized environment in turn enhancing QoS parameters. The remainder of the paper is organized as follows: Sect. 2 provides the related work, Sect. 3 gives the proposed framework, Sect. 4 gives results and discussions and Sect. 5 provides the conclusion.

2 Related Work

Zheng Yan et al. [3], has proposed a security framework for trust evaluation and management of security in computing environments to achieve security in cloud computing. It specifies the trust and security parameters in the virtualized environments. Security framework satisfies all the features of the security mentioned in requirements. They have suggested two types of services: Trust and Security.

Zhaowen Lin et al. [4] proposed a security service architecture, which has three planes: infrastructure, control, and application. Infrastructure consists of hardware devices. Application plane controls certain applications. Control plane includes security services to provide security to software-defined network.

da Silva et al. [5] proposed an approach to automation of SLA contracts between the users and cloud providers. It specifies an approach to monitor the security feature in the cloud environment. The authors have used the security metrics of the user to measure the security value. Monitoring agents are deployed in all the infrastructure components. Information collected is used in evaluation of security-based SLA.

Gaolei li et al. [6], has proposed scheme for multitenant software -defined network to provide the finely grained SLA-aware QoS parameters. SDN supports the provisioning of resources. End-to-end negotiation of SLA is carried out here. Based on the service requested by the users, SLA parameters are fixed. SLA parameters is mentioned by users. Any modification in SLA, will be completed by the users. The authors have provided the dynamic bandwidth allocation.

de Jesus et al. [7] presented that SDN can be used in cloud computing to take advantage of central control over the network by the use of SDN controller. The policies can be defined which can change dynamically as per changes in resource utilization in cloud environment such as deployment of the VMs.

Kaaniche et al. [8] proposed rSLA to monitor the security of the network. They have specified some of the mechanisms for enforcement of security and monitor whether the system is SLA compliant. Users can view all the security specifications of the SLA.

3 Proposed Framework

3.1 Trust-Based SLA in Cloud

Service-level agreement is the formal agreement of the cloud user with cloud provider for the QoS parameters required for the application to be deployed. Many QoS parameters are specified by the user. Trust is the one among QoS parameters. Trust is evaluated in the cloud with the parameters such as security feature, availability of resources, reliability of the cloud provider, and throughput. Based on the trust score, cloud provider is chosen by the cloud user. If the cloud provider fails to provide the SLA compliance, penalties are incurred. Hence, conformance of the SLA is needed with user's perspective to achieve the good performance. It is required that somebody has to monitor the SLA conformance. SDN is the emerging technology, which can be used for monitoring the SLA compliance and improve the security at infrastructure level. Any kind of unusual behaviors and threats are to be monitored by the SDN controller.

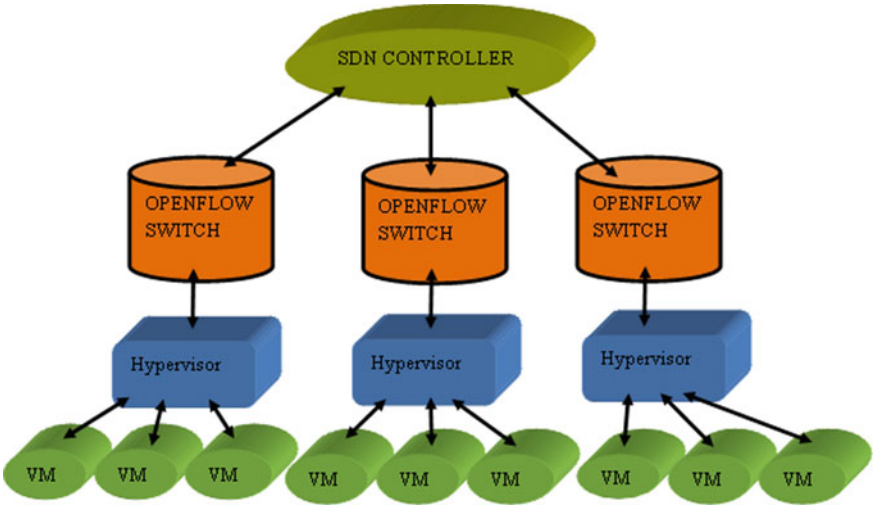


Fig. 3 SLA-based SDN security framework

3.2 Contribution of SDN in Cloud Computing

In this framework, SLA compliance is monitored. Normally certain parameters are mentioned in the SLA. In this SLA, along with QoS parameters, which are used regularly trust and security are given higher priority. Hence, timely monitoring of the activities in the virtual network is carried by the SDN controller. SDN is a more promising technology to deploy and programme. It is further divided into two planes, namely, control plane which makes decisions [9, 10], and data plane which follows the instructions and acts as forwarding elements. Control plane acts as a central control system for the entire network [10]. Unlike normal intrusion detections systems detecting unusual behaviors of users and sending notification, SDN controller sets policies to block the malicious user's traffic in the network.

Security-based framework as shown in Fig. 3 to monitor the SLA compliance and address security has been provided. The proposed framework consists of

- Virtual Machines
- Hypervisor
- SDN controller
- OpenFlow switches

Virtual Machines: Virtual machine runs operating systems, software required by the users. They are not the physical computer but abstract view of computer. Many virtual machines can be present in the same physical computer.

Hypervisor: Hypervisor acts as the middle layer and is well informed about the VMs. Hypervisor is the software which manages the VM lifecycle activities such as

VM deployment, VM migration, and VM shutdown. Basically it takes care of entire lifecycle of VM. The information is forwarded to the SDN controller.

SDN Controller: SDN-based controller is placed on the top of hypervisor. It has all the information about the cloud provider as well as users. Controller is the brain to control all the operations in the system. It sets the policies to block malicious user's traffic toward the network [11] so that data integrity, confidentiality, and reliability are increased.

OpenFlow Switches: They act as the forwarding elements in the network. They are not intelligent but dumb and obey the instructions of the SDN controller. Communication between switch and controller takes place using OpenFlow protocol. Policies are set to monitor malicious users. The information is sent to the controller. Using SDN, security can be enhanced [12].

The phases involved in monitoring the trust-based SLA are the following:

1. Negotiation phase: SLA parameters [13, 14] are negotiated between the user and provider.SLA parameters such as reliability, throughput, data integrity, and availability of resources, which contribute for trust score are included in SLA. A cloud provider whose values are satisfactory for the user is chosen.
2. SLA compliance is monitored by SDN controller checking the parameters of QoS and security.
3. Monitoring of security attacks is carried out by the SDN controller.
4. If any security attack is noticed, notification is sent by the switch to the SDN controller to further take action.
5. If repeatedly attacks are seen, then the resources are released from malicious user and allocated to other users.

4 Results and Discussions

Monitoring process is carried by OpenFlow switches and notification is sent to SDN controller. Measured values are received by the SDN controller. Only true warning messages are sent to the cloud provider from controller after deciding about attack packets. Two cases are considered wherein first case is without using SDN controller and second one is using SDN controller. QoS features are very important as cloud providers have to offer the services to the users. Ensuring the availability of the resources and monitoring the SLA compliance in the network by the SDN controller will result in better performance. The proposed framework will enhance the security features and comparison is very well shown in Table 1. The QoS features obtained in both traditional and SDN-based network are listed as shown in Table 1.

Simulation is conducted by considering the attacks like SYN flood and DOS attacks. Simulation time considered is 60 min. Around 150 attack packets are sent toward the network among 350 packets. In traditional network, notification is sent to the cloud provider and action is not taken immediately. But, whereas in SDN-based

Table 1 QoS features in traditional and SDN-based network

QoS features obtained	Traditional network	SDN-based network
Availability	Yes	Yes
Reliability	Yes	Yes
Data integrity	No	Yes
Confidentiality	No	Yes
Throughput	Yes	Yes

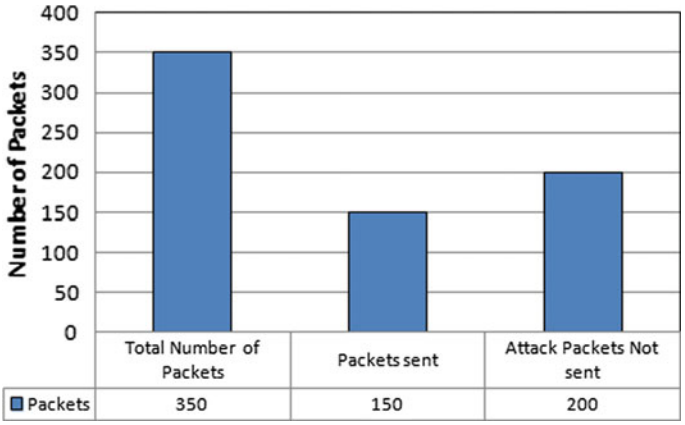


Fig. 4 Attack detection and action taken

network attack packets are blocked. Figure 4 lists the attack detection and action taken.

5 Conclusion

Cloud computing is a technology which offers the services to the users anytime, anywhere as per the need. Security at IaaS level has to be given higher priority as the upper levels makes use of infrastructure. Hence, the hardware has to be trusted so that users can complete their tasks. Security-aware trust-based SLA framework using SDN is proposed in which conformance of SLA is monitored using SDN controller. If any attack is found in the network, it is detected by the OpenFlow switch and action policy is set to block the traffic and this information is sent to the SDN controller. Controller is the brain of the network, which takes decision whether the resource allocated to malicious user has to be released or not. Initially, warning notification is sent and continued attacks from the same user might release user's resources. Protection of the users data is very important. Many virtual machines will be deployed on the same physical machine which is a security threat. So in order to

take care of the system security, monitoring is carried out. It is observed that attacks are noticed by the controller.

Acknowledgements The authors would like to thank TEQIP Phase III and BMSCE management for their continuous support for carrying out this research and completion of the work.

References

1. Desai A, Nagegowda KS, Ninikrishna T (2015) Secure and QoS aware architecture for cloud using software defined networks and Hadoop. In: 2015 international conference on computing and network communications (CoCoNet), IEEE
2. SDN Architecture Overview, Open Network foundation (2012)
3. Yan Zheng, Zhang Peng, Vasilakos Athanasios V (2016) A security and trust framework for virtualized networks and software-defined networking. Secur Commun Netw 9(16):3059–3069
4. Lin Zhaowen, Tao Dan, Wang Zhenji (2017) Dynamic construction scheme for virtualization security service in software-defined networks. Sensors 17(4):920
5. da Silva CA, de Geus PL (2014) An approach to security-SLA in cloud computing environment. In: 2014 IEEE Latin-America conference on communications (LATINCOM), IEEE
6. Li G, Wu J, Li J, Zhou Z, Guo L (2017) SLA-aware fine-grained QoS provisioning for multi-tenant software-defined networks. IEEE access
7. de Jesus WP, da Silva DA, Júnior RT, da Frota FVL (2014) Analysis of SDN contributions for cloud computing security. In: Proceedings of the 2014 IEEE/ACM 7th international conference on utility and cloud computing. IEEE Computer Society
8. Kaaniche N, Mohamed M, Laurent M, Ludwig H (2017) Security SLA based monitoring in clouds. 2017 IEEE International Conference on Edge Computing (EDGE), IEEE
9. Iqbal A, Javed U, Saleh S, Kim J, Alowibdi JS, Ilyas MU (2017) Analytical modeling of end-to-end delay in openflow based networks. IEEE Access 5:6859–6871
10. Farahmandian S, Hoang DB (2016) Security for software-defined (cloud, SDN and NFV) infrastructures–issues and challenges. In: Eight international conference on network and communications security
11. Tsugawa M, Matsunaga A, Fortes JAB (2014) Cloud computing security: What changes with software-defined networking? Secure Cloud Computing. Springer, New York, NY, pp 77–93
12. Seeber S, Rodosek GD (2014) Improving network security through SDN in cloud scenarios. In: 2014 10th international conference on network and service management (CNSM), IEEE
13. Kumawat S, Tomar D (2014) Sla-aware trust model for cloud service deployment. Int J Comput Appl 90(10)
14. Buyya R, Garg SK, Calheiros RN (2011) SLA-oriented resource provisioning for cloud computing: challenges, architecture, and solutions. In: 2011 international conference on cloud and service computing (CSC), IEEE

Big Data Computing with Distributed Computing Frameworks

Gurjit Singh Bhathal and Amardeep Singh

Abstract Big Data volume, velocity, and veracity characteristics are both advantageous and disadvantageous during handling large amount of data. It is really difficult to process, store, and analyze data using traditional approaches as such. To process data in very small span of time, we require a modified or new technology which can extract those values from the data which are obsolete with time. The distributed computing frameworks come into the picture when it is not possible to analyze huge volume of data in short timeframe by a single system. Distributed Computing is the technology which can handle such type of situations because this technology is foundational technology for cluster computing and cloud computing. Distributed Computing compute large datasets dividing into the small pieces across nodes. Numbers of nodes are connected through communication network and work as a single computing environment and compute parallel, to solve a specific problem. Hadoop is an open-source framework that takes advantage of Distributed Computing. It is implemented by MapReduce programming model for distributed processing and Hadoop Distributed File System (HDFS) for distributed storage. With time, there has been an evolution of other fast processing programming models such as Spark, Strom, and Flink for stream and real-time processing also used Distributed Computing concepts.

Keywords Distributed computing · Big data · Hadoop · Stream processing · Batch processing

1 Introduction

In the era of the Internet, which provides services from banking to healthcare, it has created the demand for new-generation data management and considered as a platform for everything. The growth of technology in last decade like virtualization

G. S. Bhathal (✉) · A. Singh
Department of Computer Science and Engineering, Punjabi University, Patiala, Punjab, India
e-mail: gurjit.bhathal@gmail.com

A. Singh
e-mail: amardeep_dhiman@yahoo.com

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_49

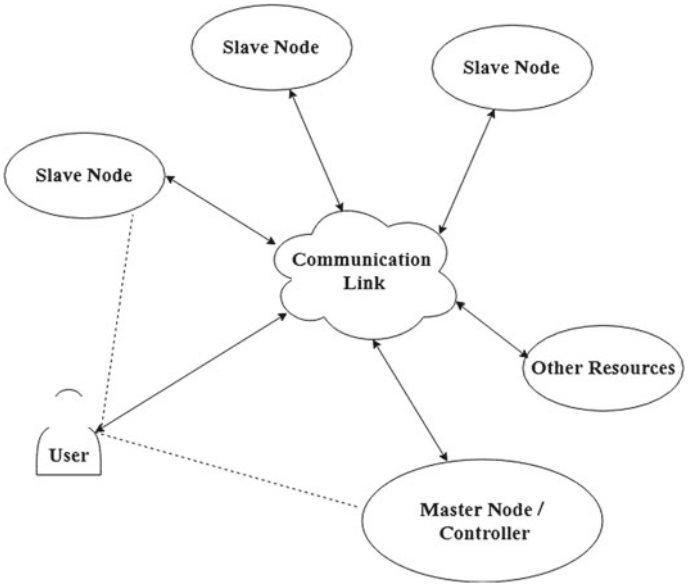


Fig. 1 This figure shows a Distributed Computing Architecture of Hadoop consisting of master node and slave nodes are connected with other resources via a network communication link

and commodity computers that can be clustered in a rack and networked as alternate of high-end supercomputing. Economics of computing changed dramatically with decreasing the cost of purchasing storage and computing resources. Distributed Computing has a major contribution and played an important role for this change. The concepts of scalable and affordable computing of Distributed Computing are adopted by Big Data. The cluster of commodity computers with low power and low cost are used for Big Data processing and clusters are working on scale-out concept as per requirement for processing, data is replicated on different nodes for high reliability and fault tolerance. However, not all problems require Distributed Computing; in some individual computing entities simply pass messages to each other. In other situations, the concept of multithreading and multiprocessing of Distributed Computing is adopted to enhance the parallelization of processing. In real-time analytics like velocity characteristics of Big Data, Spark is using multithreading concept to tackle problem.

Distributed Computing is "A system where the hardware and software components have been installed in geographically dispersed computers that coordinate and collaborate their actions by passing messages between them". At present, Big Data is creating a huge impact on industries. Consequently, 50% of world's data has already been moved to Big Data technologies (Hadoop), and 75% data will be moved to this technology by end of this year, predicted as per the demand of market (Fig. 1).

Apache Hadoop [1, 2], a distributed system, also takes advantage of parallel computing technology for processing of data simultaneously on multiple machines

using scale-out clustering. MapReduce is a computing model designed for highly scalable processing of huge volume of data. In MapReduce [3], jobs are split into small independent tasks and process in parallel on multiple nodes of cluster, which refers to the collection of commodity computers connected with each other.

Recently, the growth of data has exploded, and the inception of data from social media and transaction logs has been streamed. Thus, new tools and improved frameworks have been added according to the market need. New popular frameworks such as Spark and Flink take a lead to MapReduce. Spark is a multithreaded Distributed Computing programming model that splits a large dataset into small chunks and processes and stores data on a distributed file system. Nowadays, the completely stream-based frameworks, such as Strom and Samza, are also being used in the industry. Section 2 presents various computing techniques that follow the same principle as that of Distributed Computing. Section 3 discusses various batch-based and stream-based frameworks used for fast data processing using basic of Distributed Computing. In addition, the features of all these frameworks are compared.

2 Distributed Computing Techniques

Big Data inherits advantages of both Distributed Computing and parallel computing. In Distributed Computing, the resources are connected by a fast communication network and computing tasks are distributed on these resources. This method increases the speed and efficiency of the system. Distributed Computing is about 50-year-old technology, used with new programming concepts in Big Data, which is much faster and more efficient than the traditional methods of computing and also suitable for Big Data to process a massive amount of data in limited time [4]. There are many Distributed Computing techniques in operation today, where following are the important computing techniques:

- Grid Computing
- Utility Computing
- Cluster Computing
- Cloud Computing

2.1 Grid Computing

Grid Computing [5] is a mixture of heterogeneous computing resources with different administrative control domains for a common task and resources can be distributed on several locations and interconnected with each other using communication network of high latency and low bandwidth, mostly by the Internet. In grid environments, large jobs are divided into smaller jobs and run on many nodes parallel. The grid might be owned by various companies or private. The grid nodes are independent

and loosely coupled; the resources in a grid distributed on different physical location and are operated independently.

2.2 Utility Computing

Utility Computing [6] technologies provide services such as storage, applications, computing power and computing resources at a very low cost to the customers. Utility Computing model is on-demand computing like software as a service, applications, etc. Utility computing provides more computing power and storage space than a single time-sharing computer. Utility computing requires an infrastructure like cloud to provide different services. Various advance servers are used at the back end support to make it possible.

2.3 Cluster Computing

Cluster Computing [7] is a collection of homogenous tightly coupled computers connected with each other using high-speed LAN and work together as a single supercomputer. The clusters are widely used to improve computing performance with centralized scheduling and management. In this, computers are connected with a very fast speed communication link at the same physical location. The resource manager of the cluster is managed resources using master node.

The cluster is single-handed collection of nodes owned by an organization. The cluster is high fault-tolerant and a fast speed network with low latency and high bandwidth in contrast to Grid computing.

2.4 Cloud Computing

Cloud Computing [8–10] is a combination of all above computing techniques like Grid, Utility, Cluster, and Distributed, in which all the devices are controlled remotely on a network. This setup decreases the user demands for software and super hardware. Cloud computing is a model where computing is done at remote computers away from local computers. Cloud provides different services, such as e-mail to multipart data analysis programs, using the Internet. Cloud reduces the cost of processing and storage for all small, medium and large companies. The cloud is a dynamic scale-in and scale-out computing infrastructure which is highly scalable, fault-tolerant, and load balancing. The cloud provides its services to the users on demand, and its inter-connection network is a high-speed network with low latency and high bandwidth. Every node and application in the cloud is an independent entity.

Table 1 Comparison of different computing techniques considering different function

Grid Computing	Utility Computing	Cluster Computing	Cloud Computing
Loosely coupled	On-demand pricing	Tightly coupled systems	On-demand self-service
Diversity and dynamism	Uniform utility computing services	Single system image	Broad network access
Distributed job management and scheduling	Share the resources in the shared pool of machines	Centralized job management and scheduling system	Resources pooling and rapid elasticity
High-end computers (servers, clusters)	High-end computers (servers)	Commodity computers	Commodity computers, high-speed network and high-end servers and NAS

3 Big Data Distributed Computing Processing Frameworks

Distributed Computing has a great role in the success of Big Data. Big Data requires very low costing storage space and infrastructure, which is provided by cloud computing. Cloud Computing is a branch of Distributed Computing [11]. In order to process a huge quantity of data at a very high speed, we required the power of cluster computing, which is also a branch of Distributed Computing. Thus, to enhance the processing speed of Big Data, there are two features: batch processing and stream processing [12]. By default, Hadoop MapReduce is a batch processing system, but with the invent of social media inception of data in real- time, there is a need for a stream processing framework for Big Data [13]. MapReduce processing framework is designed to handle large data sets and split datasets into small batches. In contrast other frameworks process data in an uninterrupted stream, as it flows into the system. MapReduce workings as its default processing engine but, for instance, Apache Spark and other framework are hooked into the Hadoop Ecosystem to handle real-time streaming data (Table 1).

Batch processing frameworks split the large data jobs into small chunks and distribute these chunks on a large number of nodes according to the size of computer cluster to process Big Data. The execution time of a batch processing is determined with the number of active nodes in a cluster and the size of the job. The batch processing model is inappropriate to satisfy real-time constraints due to having high latency to process Big Data.

Stream processing is a model for handling real-time stream synchronized with the data flow and returns the results in a low-latency fashion. Stream processing also has some features of batch processing such as fault tolerance, high availability, and resource utilization. Real-time stream processing systems give guarantee to be up and available all the time for real-time data. Stream processing achieves incremental scalability automatically by distributing processing power as well as storage capacity

across multiple computers without any human interaction. The following frameworks are hooked in Hadoop environment:

- MapReduce (Batch Processing frameworks)
- Spark (Stream and Batch Processing frameworks)
- Flink (Stream and Batch Processing frameworks)
- Storm (Stream processing frameworks)
- Samza (Stream Processing frameworks).

3.1 MapReduce

Apache Hadoop MapReduce [14] has been built for batch processing, and takes a large dataset as the input, all at once, processes it, and produces the results. Batch processing is very capable but slow in processing where high-volume data involved. An output gets delayed due to the data size and the computational power of the system. Hadoop's three main components are the distributed file system HDFS similar to the Google File System, MapReduce, and YARN [15] as scheduler. In Hadoop, the data are processed in a chain of stages, at each stage, the user progresses using the output of the earlier stage and produces input for the next stage. MapReduce wires batch processing only and does not support processing of streamed data. MapReduce is highly fault-tolerant in case of any crash and there is no need to restart the application. Hadoop framework runs on commodity computers, which are less-expensive hardware. It supports mainly Java and other languages such as C, Ruby, C++, Groovy, Perl, and Python.

3.2 Strom

Storm [16] is an open-source distributed real-time processing system. Strom processes unbounded data streams for real-time processing. It is simple to use with machine learning and real-time analytics using any programming language like distributed RPC, ETL, etc. Storm records the new benchmark for fast processing per node to process millions tuples per second. Storm is simple and easy to operate and guarantees that the data will be processed. It is scalable and fault-tolerant. Storm cluster has two types of nodes: masters and workers. The master node called "Nimbus", which is responsible for code distributing in cluster, whereas "Supervisor" program runs on each worker node, listens for work assigned to its machine and also responsible to start and stop workers' processes according to the requirement. Storm is integrated with the queuing and database technologies already in use.

3.3 Apache Samza

Apache Samza [17] is a distributed framework used for stream processing removes the shortcomings of MapReduce and avoids the large turn-around times involved in batch processing. Samza was developed by LinkedIn. It takes advantages of Kafka's such as state storage, fault tolerance and buffering. Samza is a lightweight framework, to process continuous data stream and Jobs are written in JVM support languages like Java and Scala. Every Job is split into smaller execution units called tasks, just like a partition of a stream for parallel execution to achieve scalability. Each task processes messages sequentially from each of its input partitions, and a task only first consumes data delivered by single of the partitions. Each task is operated independently if the partitions have no defined ordering. Samza relies on the rich features of YARN, and YARN used as a resource negotiation.

3.4 Apache Spark

Apache Spark [18] is developed for stream processing hybrid processing framework in UCB of Berkeley, next-generation processing framework for Big Data analytics. Spark is hybrid processing engine and much faster than MapReduce, which is 100 times faster in-memory model and 10 times faster on disk model than MapReduce. It is fast for in-memory processing that processes batch, streaming, and interactive analytics. Apache Spark uses the resilient distributed dataset (RDD) data structure, which is fault-tolerant, wherein data items are distributed over nodes of a cluster, which are read-only multiset. Spark Streaming works on a micro-batching mode, in which the batch size is much smaller than that in the conventional batch processing. Spark processes Hive structured data and supports multiple programming languages such as Java, Python, and Scala, and also streams data from a variety of sources including HDFS, Flume, Kafka, Amazon S3, etc. Spark is used for processing social media data for performing sentiment analysis in real-time and when used with Kafka it provides excellent processing results (Table 2).

3.5 Apache Flink

Apache Flink [17] is a Hybrid processing platform for batch, stream and interactive processing. Flink is the only platform which is also used for running applications like machine learning and graph processing. However, its implementation is totally different than Spark. Flink provides a powerful streaming programming model with fault management, high throughput, automated memory management, and flexible windowing schemes. Flink has a unique capability of fine-grained event-level processing architecture; owing to this, stream processing operates in real-time. Apache

Table 2 Comparison of different Computing frameworks considering different function [17]

Functions	MapReduce	Spark	Flink	Storm	Samza
Data processing engine	Batch	Hybrid	Hybrid	Stream	Stream
Developed in	Java	Scala	Java and Scala	Clojure	Java and Scala
Supported language	Java, C++, Ruby, Groovy, Perl	Java, Scala, Python and R	Java as well as Scala	Ruby, Python, JavaScript and Perl	Java, Scala JVM language only
Latency	Higher than both spark and Flink	High as compared to Apache Flink	Process the data in sub-second range without any delay	Extremely low latency as compared to others	Low latency
Security	Authentication using Kerberos	Authentication via shared secret	User authentication via Ranger	Authentication with Kerberos Thrift	Authentication with Kerberos

Flink uses native closed-loop iteration operators, for fast processing of application like machine learning and graph than MapReduce and Spark. Flink supports programs written in Java or Scala, which are compiled and optimized automatically. Flink input source data comes from variety of sources like distributed storage system and Kafka because Flink does not have its own data storage system.

4 Experiment Setup and Results

An experiment is conducted for this paper on single node Hadoop installation to check the performance of Hadoop MapReduce and Apache Spark using word count program on different sized dataset. To perform the experiment a desktop computer with i5 processor quad core with 8 GB RAM, 1 TB hard disk, Ubuntu 16.4 (64-bit) Linux operating system and Hadoop 2.7 release was used. Hadoop single node installation components are NameNode, jobtracker, DataNode, and tasktracker on same machine. As mentioned earlier four different sizes of datasets are used to test the performance of MapReduce and Apache Spark using word count programs and results are as shown in Table 3. During the experiment Linux word count default command also considered with Spark and MapReduce, result shows file size and time taken by each system to process these datasets.

According to the experiment result table data, the traditional system and methods are incompetent to process the large data sets like the word count facility of a Linux operating system, which is unreachable means taking much more time to process very

Table 3 Comparison data-on-disk different size datasets and time is taken by frameworks

Datasets file size in MB (Approx)	Processing time (s)		
	Spark	MapReduce	Linux
2 MB	7.32	30.74	150.42
4 MB	7.91	31.52	170.33
40 MB	12.93	32.88	unreached
60 MB	14.59	39.12	unreached

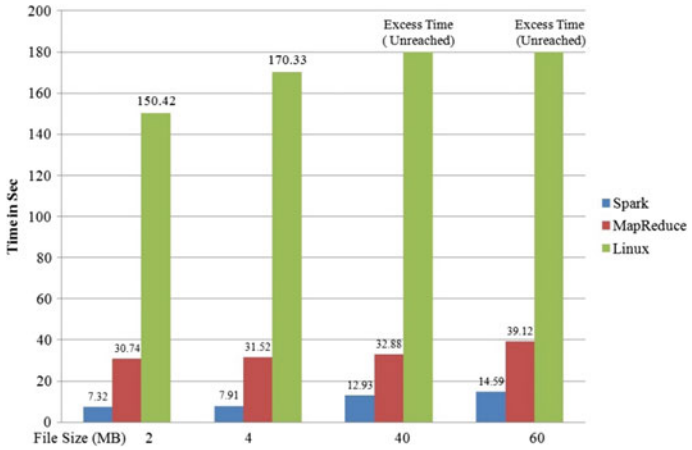


Fig. 2 The graph figure shows performance comparison of Spark, MapReduce and Linux WC command for word count programme using different size of files. The result of Spark and MapReduce processing in seconds and Linux in minutes

small files of 40 MB. Spark taking very less time as compared to the other framework like MapReduce and LINUX operating system. If we extend this experiment at large scale that is cluster of hundreds of commodity computers and dataset files 10 GB and above these frameworks process those files infraction of a second but we required Distributed Computing concept and parallel computing in cluster of hundreds of nodes. Hadoop was introduced to process large data sets and successful to achieve the target only with the help of distributed systems (Fig. 2).

5 Conclusion

The above analysis indicates that the role of Distributed Computing and Parallel Computing both in Hadoop framework is highly useful and will stay in future. MapReduce cannot solve every problem of real-time data. It is still a good choice for research, experimentation, and everyday data manipulation workloads that are not time-sensitive are processed with batch-only. Hadoop is a less expensive and good

choice when it is expensive to implement hardware required for some other successful frameworks. For only stream processing data, Storm is very low-latency processing and consists of a beamy communication assist and cannot guarantee requests in its failure configuration. For mixed processing workloads stream and batch, Spark provides speedy batch processing and streaming processing breaks in micro-batches. Flink provides real-time stream processing with support of batch processing. It can run tasks written for other platforms and provides low-latency processing and heavily optimized. Spark exhibits good performance on dedicated large main memory, where the entire data of stream or micro batch can fit in the memory on clusters, in contrast, Hadoop perform well even less main memory along with other services, even when only the part of data fit in memory. Storm is a good choice when application requirements sub-second latency without data loss, whereas Spark can be used in state full computations to ensure that the event is processed only once.

The best framework depends on situation and the condition of the data to be processed, how time constraint is, as per the requirements, and in what category of results a user is curious. There are always a trade-offs between implementing as one package and working with closely focused specific projects, and considerations with similar characteristics when evaluating brand new and progressive successes over well-tested equivalents.

References

1. Apache Software foundation. http://hadoop.apache.org/ [Online] (2017, Dec)
2. David T. https://wiki.apache.org/hadoop/PoweredBy [Online] (2017, Dec)
3. Ghemawat S, Dean J (2004) MapReduce: simplified data processing. In: 6th symposium on operating system design and implementation (OSDI 2004), San Francisco, California, USA, pp 137–150
4. Hortronworks. https://hortonworks.com/ [Online] (2018, Jan)
5. Grid Computing. http://en.wikipedia.org/wiki/Grid_computing [Online] (2017, Dec)
6. Wiki Pedia. http://en.wikipedia.org/wiki/Utility_computing [Online] (2017, Dec)
7. Cluster Computing. http://en.wikipedia.org/wiki/Computer_cluster [Online] (2018, Jan)
8. Cloud Computing. http://en.wikipedia.org/wiki/Cloud_computing [Online] (2018, Jan)
9. Botta A, de Donato W, Persico V, Pescapé A (2016) Integration of Cloud computing and Internet of Things: A survey. Future Gener Comput Sys 56:684–700
10. Purcell BM (2013) Big data using cloud computing
11. Tanenbaum AS, van Steen M (2007) Distributed Systems: principles and paradigms. Upper Saddle River, NJ, USA: Pearson Higher Education
12. de Assunção MD, Buyya R, Nadiminti K (2006) Distributed systems and recent innovations: challenges and benefits. InfoNet Mag 16(3)
13. Steve L. https://wiki.apache.org/hadoop/Distributions%20and%20Commercial%20Support [Online] (2017, Dec)
14. Corporation D (2012) IDC releases first worldwide hadoop-mapreduce ecosystem software forecast, strong growth will continue to accelerate as talent and tools develop
15. Thusoo A, Sarma JS, Jain N, Shao Z, Chakka P, Anthony S, Liu H, Wyckoff P, Murthy R (2009) Hive. Proceedings of the VLDB Endowment 2(2):1626–1629
16. Apache Strom (2018). http://storm.apache.org/releases/1.1.1/index.html [Online] (2018)
17. https://fxdata.cloud/tutorials/hadoop-storm-samza-spark-along-with-flink-big-data-frameworks-compared [Online] (2018, Jan)

18. Justin E. https://www.digitalocean.com/community/tutorials/hadoop-storm-samza-spark-and-flink-big-data-frameworks-compared [Online] (2017, Oct)
19. Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH, M. G. Institute J. Manyika (2011) Big data: the next frontier for innovation, competition, and productivity, San Francisco
20. Ed Lazowska (2008) Viewpoint Envisioning the future of computing research. Communications of the ACM 51(8):28
21. Dollimore J, Kindberg T, Coulouris G (2015) Distributed systems concepts and design, 4th ed. England, Addison-Wesley, London
22. Hadoop Tutorial (Sep, 2017). https://data-flair.training/blogs/hadoop-tutorial-for-\beginners/

A Novel Approach for Weather Prediction Using Forecasting Analysis and Data Mining Techniques

Md Rashid Mahmood, Raj Kumar Patra, Rohit Raja and G. R. Sinha

Abstract The Weather forecast is used to predict the atmospheric condition such as rain, wind, heat, pressure, humidity etc. It is very useful especially in agriculture, air traffic and military purposes. Weather forecast is done by analysis on bulk data, received from the satellite for specific purpose. It is a very tedious and time taking for analysis on such a huge volume of data. Now a day, Data miming is very popular to handle such type of data. Some algorithm like Back Propagation, Naïve Bayesian and Decision Tree Induction is being used to predict the weather condition on diverse parameters. Here, we are proposing Cumulative Distribution Function (CDF) for modeling and analysis of complex data for weather prediction. This method gives us the better accuracy in case of climate change in near future.

Keywords Data mining · Weather forecasting techniques · Forecast types · Forecast methods

1 Introduction

A historical data is perceived and analyses by the data scientist, but it is almost impossible to get accurate prediction of rainfall and any other characteristic of climate by simply analyzing the historical data. So, it critical need to see how the atmosphere

M. R. Mahmood (✉)
Department of Electronics and Communications, Guru Nanak Institutions Technical Campus, Hyderabad, India
e-mail: er.mrashid@gmail.com

R. K. Patra
CMR Technical Campus, Hyderabad, India
e-mail: patra.rajkumar@gmail.com

R. Raja
Sreyas Institute of Engineering and Technology, Hyderabad, India
e-mail: drrohitraja1982@gmail.com

G. R. Sinha
Myanmar Institute of Information Technology, Mandalay, Myanmar
e-mail: gr_sinha@miit.edu.mm

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_50

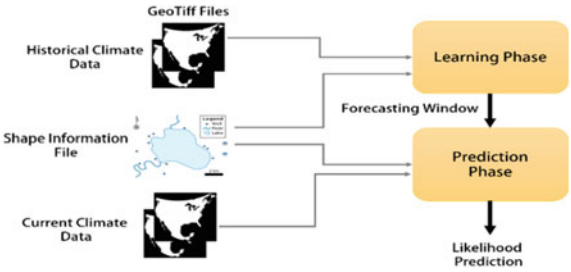


Fig. 1 Extreme climate prediction, learning and prediction phase [12]

inconstancy impacts the area. A weather and prediction model of Jordanian region is proposed by Omary et al. This model is built by gathering data from local and web resources. A tool is built for weather prediction by using websites that store weather related information. Data Mining (DM) and Artificial Intelligence (AI) algorithm are used to predict the weather forecast and expected climate change [1]. Data mining is very popular technique for weather forecast. This is carried out using Neural Networks (NN), Decision Tree Algorithm (DTA) and weather-related data collected in particular duration of time. This algorithm predicts the optimum results as compared to standard performance metrics [2–4]. K-means clustering is used to group similar type of data sets with J48 classification method using Linear Regression Analysis (LRA) is also being used in weather forecast [5–8]. The prediction of rainfall is major challenging issues in weather forecast. Artificial Neural Networks, Clustering and Regression Analysis are widely used techniques for weather forecasting. Multiple linear regression (MLR) techniques are also used for the early prediction of rainfall [9, 10]. The average rainfall in Udupi, Karnataka, India has been predicted by ANN models by training and testing of the data sets and calculating the number of hidden neurons in the layers [11]. There is various prediction model available for weather forecast. Here, we are proposing extreme climate change for a particular are by using CDF methods. This is segmented into two steps, learning and prediction as shown in Fig. 1. In learning, a particular window sized model is given input for training and output of this phase is used to predict the future climate condition in prediction phase.

In learning, we concern about peak atmospheric condition such as extreme low/high temperature or extreme high precipitation in the area of interest. Here, Empirical Cumulative Distribution Function (ECDF) based frequency examination for extreme occasion events is used to explain information to benchmark the authentic atmosphere information. The examination gives the ideal window-measure models (ideal displaying and estimating windows) for anticipating future events of similar occasions over the region of interest. Learning is further classified into feature extraction and cumulative distribution function as shown in Fig. 2. This model having two data sources, Shape record file and GeoTiff file. The Shape record file consist of geospatial vector related with chosen land region of interest. On the other hand, GeoTiff records hold the geo-referenced transient atmosphere occasion data examined at a particular interval of time over a specified time window. In this research,

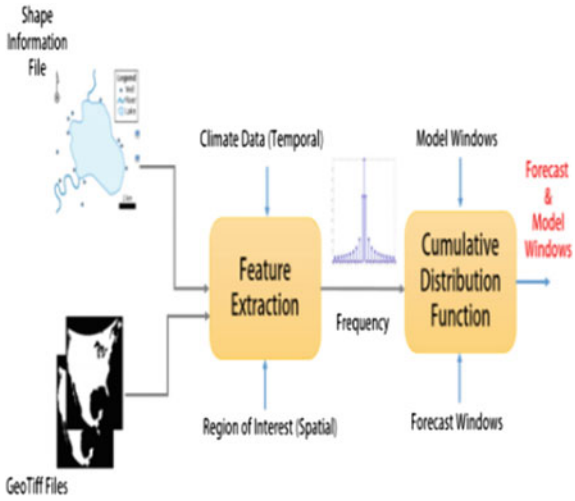


Fig. 2 Framework of learning phases

we utilize the GeoTiff documents for temperature and precipitation occasions examined at month to month determination. The steps of features extraction are shown in Fig. 3 temporal atmosphere information comprises of two-dimensional matrices sorted based on longitudes and latitudes of the land focuses in the GeoTiff documents. We utilize the Shape record to veil those geological focuses that are not included in the area of interest. At last, the area of interest ought to comprise just of those focuses in the 2D framework related with topographical territories with particular qualities that are region of interest. On the other side, in prediction stage, similar region is to be considered and it utilize the window-measured model. Here, present metrological information is contributed to an expectation module. In this module, we utilize ideal demonstrating window to create the empirical cumulative distribution function model. By and large, the probability data is utilized to assess with specific certainty (extreme occasion happen inside the ideal window measure). An interesting and profoundly instructive portrayal of empirical cumulative distribution function model is the survival display, which gives the probability to an occasion surpassing a specific esteem, in particular, $P(X > x)$, frequently known as likelihood of exceedance and is expressed as $1 - \text{ECDF}$. The prediction stage utilizes the ideal displaying and gauging windows from the learning stage to play out the ECDF-construct frequency examination with respect to the latest atmosphere information. This stage is worked by an expectation module, which is basically like that of the learning stage. The prediction module evaluates the probability of certain atmosphere occasion events in a given topographical locale. The contributions to the forecast module are the latest atmosphere information and a client inquiry. The component extraction arranges changes over the time-area atmosphere occasions to recurrence space atmosphere occasions. The CDF generator gathers the frequency-domain tests in the ideal demonstrating window to create the ECDF.

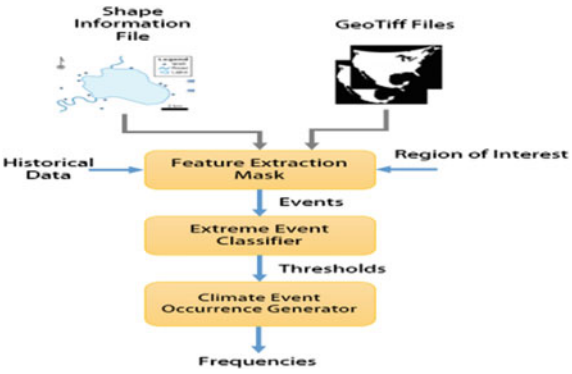


Fig. 3 Feature extraction components [12]

2 Experimental Results and Discussions

In this section various experiments are conducted and its results are presented in the form of bar graphs and charts. In the presents work results performed using CFD with various feature extraction techniques of climate data. To evaluate our metrics, we have used data mining tools and MATLAB. We have taken climate data of Chattisgarh, India to utilize our CDF framework for analysis on climate data, we are presenting three cases, through which the framework can be evaluated and predication can be made abnormal high-temperature, abnormal-low temperature, abnormal high-precipitation.

2.1 Abnormal High Temperature

This prediction part deals with the abnormal high temperature events. For this case, we need to anticipate the normal extent of the region watching abnormally high temperatures amid the four seasons and estimates the odds that future years may watch more prominent extents of the locale (more than the normal) encountering curiously high temperatures various circumstances (surpassing high-temperature events). Figures 4, 5, 6 and 7 demonstrates the three dimensional disseminate space (hit rate versus demonstrating window as opposed to estimating window) gotten in the learning stage for the preparation temperature information. To get the normal extent of the area that may encounter strangely high temperature, the forecast stage utilizes the events tests gathered in the displaying window. Review that an event test in this setting is the extent of the locale encountering strangely high temperature. Since all the occasion tests in the demonstrating window are similarly likely, the forecast stage gets the normal extent as

$$E[X] = \frac{1}{N} * \sum_{i=1}^{N} x_i$$

where, $x_1, x_2, x_3 \ldots x_{N\rightarrow}$ sample of event in window (four in a year).

The prediction obtains result as Winter (January), Spring (April), Summer (July) and October (Fall). Here, we determine the likelihood of multiple events that occurs frequently with high temperature and the recurrence period.

Figures 4, 5, 6 and 7, presents output prediction for each season window. The recurrence equation in which event will happen again is given by the equation

$$P = \left(\frac{y}{i}\right) * p^i * (1-p)^{y-i}$$

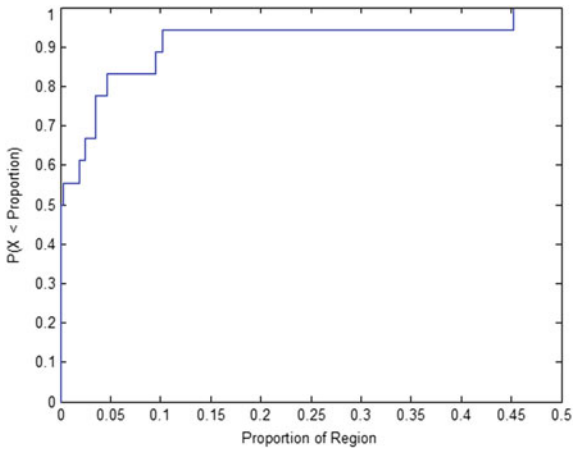


Fig. 4 ECDF model (Jan)

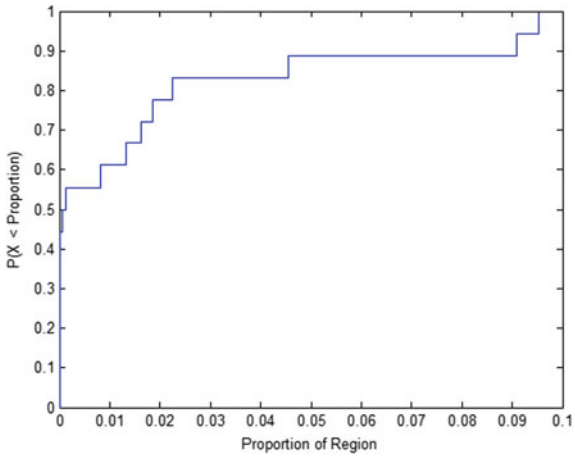


Fig. 5 ECDF model (Apr)

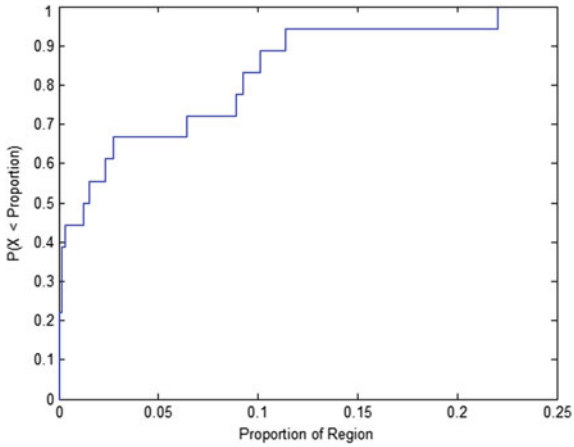


Fig. 6 ECDF model (Jul)

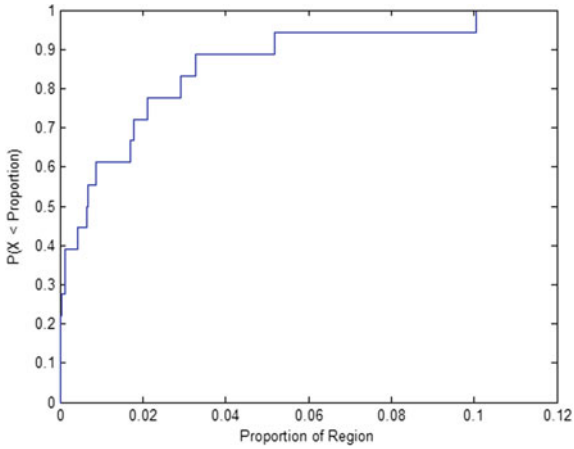


Fig. 7 ECDF model (Oct)

For this particular case study, the forecasting windows is ($y = 5$ years). Therefore, the event re-happening can occur 0–5 times in next 5 years (2019–2024). ECDF model for four different seasons (winter, spring, summer and fall) is shown in Figs. 4, 5, 6 and 7. By using above equation, likelihood of re-happing of events with high temperatures are shown in Figs. 10, 11, 12 and 13. Recurrence period for various seasons are obtained for low temperature are shown in Figs. 8, 9, 10 and 11 with recurrence period of 1.3, 1.5, 1.4, 1.5 years for month January, April, July, October respectively.

As shown in Fig. 8, the January of year 2020 will be having extreme high temperature with recurrence period of 1.3 years. The area affected by this low temperature is 7% of total area. Figure 9 showing forecasting window for the month of April. Here, April of year 2019 and 2020 will be having extreme high temperature with recurrence period of 1.5 years. The area affected by this low temperature is 8.5% of

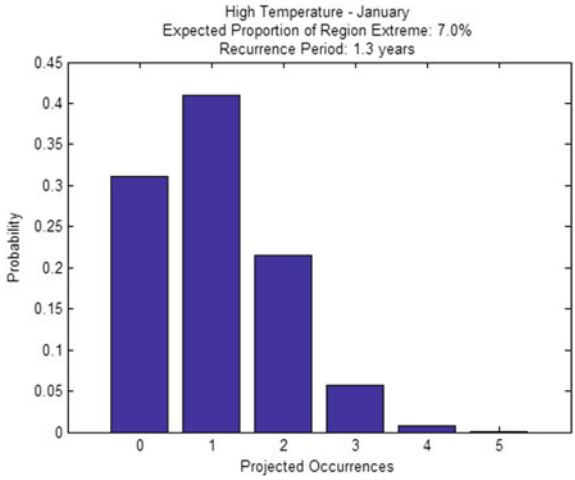


Fig. 8 Forecast window (Jan)

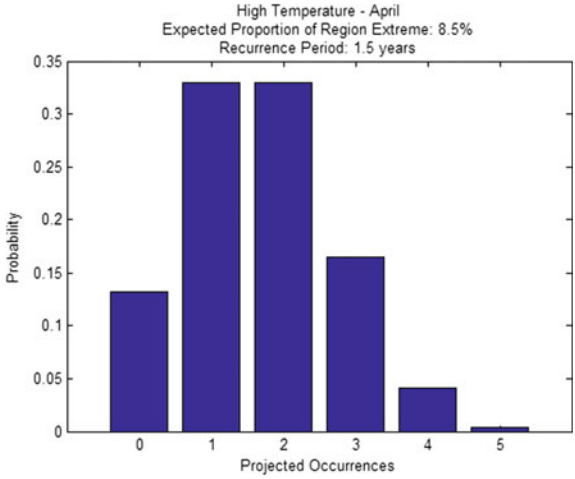


Fig. 9 Forecast window (April)

total area. Figure 10 is showing the forecast window for the month of July. As per this forecast model, July month of year 2020 will be having extreme high temperature with recurrence period of 1.4 years. The area affected by this low temperature is 10.5% of total area. In other case, for the month of October of year 2020 and 2021 will be having extreme high temperature with recurrence period of 1.5 years. The area affected by this low temperature is 8.8% of total area as shown in Fig. 11.

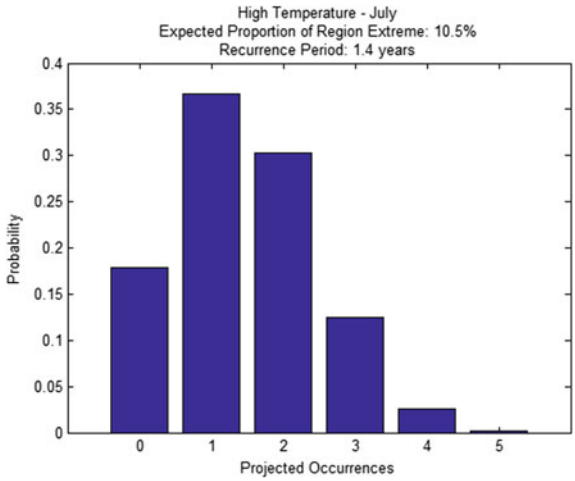


Fig. 10 Forecast window (July)

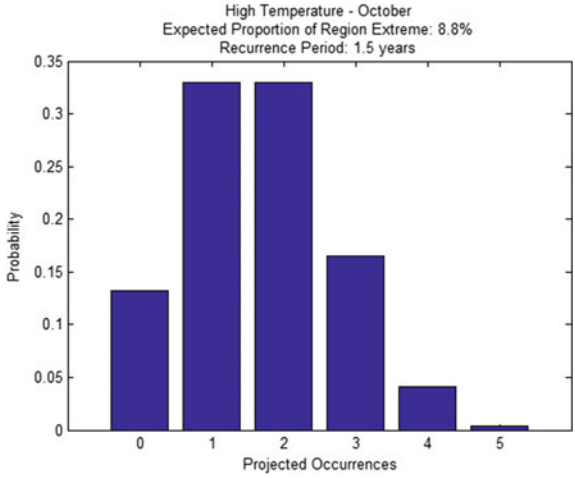


Fig. 11 Forecast window (Oct)

2.2 Abnormal Low Temperature

Here, we are dealing with the abnormal low temperature area and events. The area of focus for our analysis is considering the recurrence of event and percentage of area affected by this climate change. Figures 12, 13, 14 and 15 shows the externally low temperature areas and its re-occurrences. It is indicated that January of year 2019 and 2020 will be having extreme low temperature with recurrence period of 1.2 years. The area affected by this low temperature is 5.2% of total area which is clearly shown in Fig. 12. In Fig. 13, it is clear that April of year 2020 and 2021 will be having extreme low temperature with recurrence period of 1.4 years. The area affected by this low temperature is 1.3% of total area. For the month of July, year

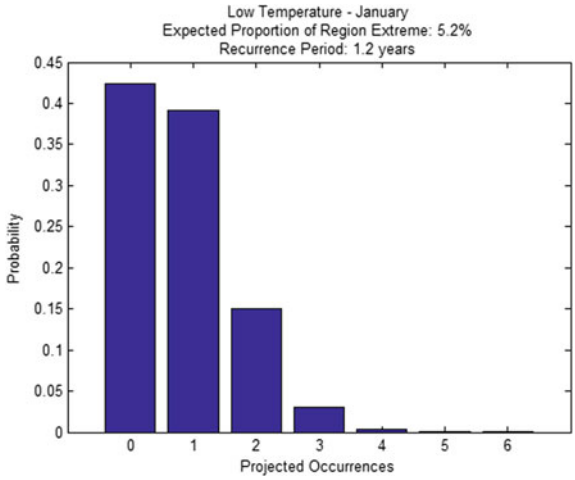


Fig. 12 Forecast window (Jan)

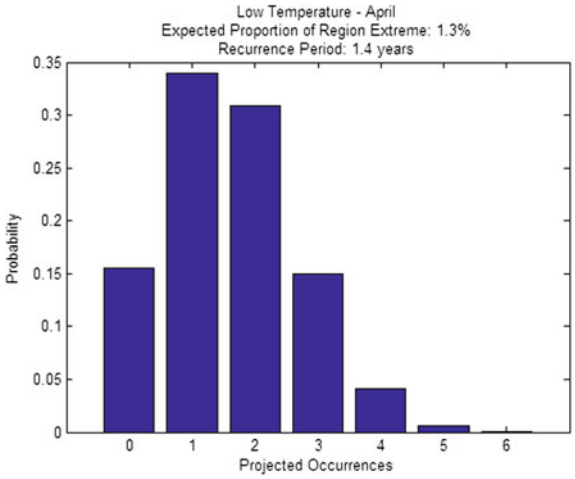


Fig. 13 Forecast window (Apr)

2020 will be having extreme low temperature with recurrence period of 1.3 years. The area affected by this low temperature is 3.6% of total area which is shown in Fig. 14. Figure 15 is forecasting for the month of October, in this month, year 2021 will be having extreme low temperature with recurrence period of 1.5 years. The area affected by this low temperature is 1.3% of total area.

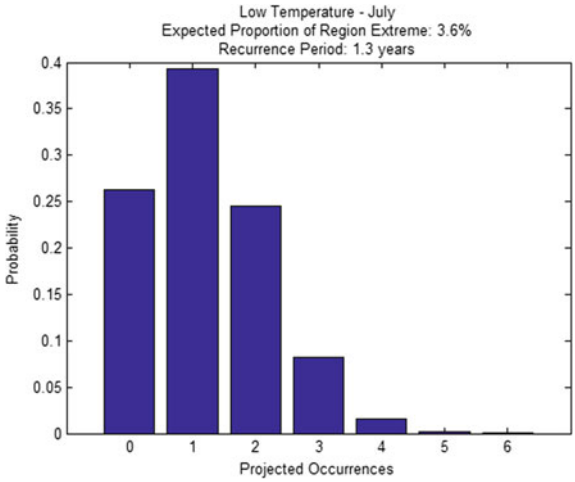


Fig. 14 Forecast window (July)

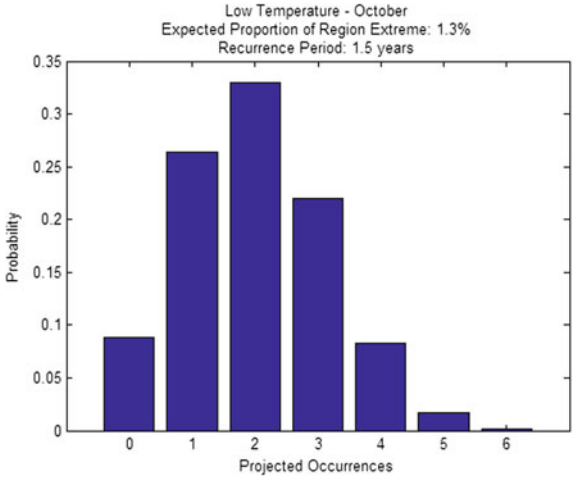


Fig. 15 Forecast window (Oct)

3 Conclusion

We have analyzed and studied two different cases of climate change of Chattisgarh, India region. The reason behind taking Chattisgarh region is because that particular climate goes sudden change in its environment frequently. So, it is interesting to apply some rule for prediction of climate activities to forecast some of the interesting occurrences patterns. We investigated two cases on various seasons with various factors and regions. This research presents re-occurrences period and region affected by that sudden climate changed to alert the civilians. Similarly, more number of cases can be included and prediction can be made. In most of the case, up-to coming three

years, it predicted in well, but for the remaining two years, it is not showing much concern. This weather forecast is very much useful to predict for the next three years.

References

1. Omary A, Wedyan A, Zghoul A, Banihani A, Alsmadi I (2012) An interactive predictive system for weather forecasting. 2012 international conference on computer, information and telecommunication systems (CITS), Amman
2. Mandale MA, Jadhawar BA (2015) Weather forecast prediction: a data mining application. Int J Eng Res Gen Sci 3(2). ISSN 2091-2730
3. Rajini Kanth TV, Balaram VVSSS, Rajasekhar N (2014) Analysis of Indian weather data sets using data mining techniques, pp. 89–94, © Cs & It-Cscp
4. Sethi N, Garg K (2014) Exploiting data mining technique for rainfall prediction. (IJCSIT) Int J Comput Sci Inf Technol 5(3):3982–3984
5. Sawaitul SD, Wagh KP, Chatur PN (2012) Classification and prediction of future weather by using back propagation algorithm-an approach. Int J Emerg Technol Adv Eng, 2(1) January. ISSN 2250-2459
6. Georgakakos KP, Hudlow MD (1984) Quantitative precipitation forecast techniques for use in hydrologic forecasting. Bull Am Meteorol Soc 65(11)
7. Newsome DH, Collier CG (1991) COST 73: possible hydrological applications of weather radar in Western Europe. Hydrol Appl Weather Radar. I.D. Cluckie and C.G. Collier (Editors), Ellis Horwood, pp 623–635
8. Hardaker PJ, Collier EG (1995) Radar and storm model-based estimation of probable maximum precipitation in tropics. In 3rd International Symposium on Hydrology Applications of Weather Radar, Sao Paolo, Brazil
9. KaveriKar Rohit Raja (2017) A review on weather prediction using data mining techniques. Int J Comput Math Sci IJCMS, ISSN 6(5):2347–8527
10. Kar K, Raja R, Chopra J (2017) Extreme weather event change prediction using CDF. Int J Innov Adv Comput Sci, IJIACS 6, ISSN 2347–8616
11. Raja R, Sinha TS, Patra RK, Tiwari S (2018) Physiological trait based biometrical authentication of human-face using LGXP and ANN techniques. Int J Inf Comput Secur 10(2/3):303–320 (Special Issue on: Multimedia Information Security Solutions on Social Networks)
12. McKinney R, Pallipuram VK, Vargas R, Taufer M (2015) From HPC performance to climate modeling: transforming methods for HPC predictions into models of extreme climate conditions. 2015 IEEE 11th international conference on e-science, Munich, pp 108–117. https://doi.org/10.1109/escience.2015.33

A Cloud-Based Inventory Management System Using a Smart Trolley for Automated Billing and Theft Detection

B. Karunakara Rai, J. P. Harshitha, Radhika S. Kalagudi, B. S. Priyanka Chowdary, Palak Hora and B. Sahana

Abstract Currently, self-checkout counters contribute about 90% of the supermarkets worldwide. However, there is no system to provide product details while shopping to customers and monitor theft which has led to loss in supermarkets and customer dissatisfaction, respectively. Nowadays, billing is a tedious task as customers have to stand in long queues to get their products barcode scanned. This paper proposes an architecture of a smart trolley which provides an automated billing, anti-theft system, and facilitates inventory management through a web application. This is achieved on a cloud-based platform using RFID and Wi-Fi technology. The goal of this proposed smart trolley system is to provide easily scalable, economical, and technology-oriented shopping system thereby reducing queuing time, anti-theft, and labor cost. Results show that there is a 26% reduction in time spent during shopping as compared to conventional shopping methods.

Keywords Anti-theft system · Automated billing · RFID · Smart trolley · Wi-Fi

B. Karunakara Rai (✉)
Department of ECE, NMIT, Bengaluru 560064, India
e-mail: karunakara.rai@nmit.ac.in

J. P. Harshitha · R. S. Kalagudi · B. S. Priyanka Chowdary · P. Hora · B. Sahana
Department of ECE, RVCE, Bengaluru, India
e-mail: jp3march1996@gmail.com

R. S. Kalagudi
e-mail: radhikakalagudi04@gmail.com

B. S. Priyanka Chowdary
e-mail: priyankachowdarybs@gmail.com

P. Hora
e-mail: horapalak05@gmail.com

B. Sahana
e-mail: sahanab@rvce.edu.in

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_51

1 Introduction

Supermarkets are a place where people visit to purchase required products according to their propensity ranging from food, clothing, and electrical appliances to improve their quality of life [1]. Currently, supermarkets provide baskets and trolleys for customers to aid their shopping, by helping them move around and pick the products of their choice. However, with the looming technology aiming at providing convenience, comfort, and efficacy to our diurnal lives, there is a deficit in this technology in the billing process and inventory management of supermarkets. Customers have to stand in long queues to get their products billed. This tedious process which involves the scanning of barcodes employed by a large number of supermarkets increases time and results in queuing delay. There is a need to revolutionize the billing process by adopting self-billing processes. Current developments in chip manufacturing technology have proposed practical approaches for new applications [2]. Barcode-based scanning results in waste of time and this can be overcome by using RFID (Radio Frequency Identification) tags to each product and by using RFID scanners which are compact and reliable [3].

This paper emphasizes on a smart trolley which provides automated billing and an anti-theft system. The trolley has RFID reader, LCD (Liquid Crystal Display), and ARM7 processor as the primary components. The customers have the advantage of self-scanning each RFID tag present on the products using the RFID scanner and view its information such as the price and its name. The bill gets augmented with each scanning and decreases with the removal of scanned items thereby providing flexibility to the customers. This system allows customers to be aware of their total bill while shopping and helps in budget management. The price for each item is obtained from a centralized database system which also apprises the stock management team when there is a need to replace racks with new stock based on a minimum threshold value. The customers only have to pay the bill amount generated at the billing court thereby avoiding waiting in long queues.

A centralized automated billing system is proposed in this work by integrating RFID technology and Wi-Fi (Wireless Fidelity) communication. Radio-frequency identification is a rapidly growing field which can have a great economic impact on the industries, particularly, in item level scanning [4]. These tags are tending to become the most pervasive tags in history.

The 16 bit ARM7 (Advanced Risk Machine) microcontroller is the heart of the smart trolley which communicates with the centralized billing system and the RFID readers. It is one of the most widely used microcontrollers which provides a fast and efficient embedded system platform. The anti-theft system is incorporated in the smart trolley using IR (InfraRed) sensors which keep count of the number of scanned items. The goal of this proposed smart trolley system is to provide easily scalable, economical, and technology-oriented shopping system thereby reducing queuing time and theft.

The organization of this paper is as follows: Sect. 2 presents the literature survey, Sect. 3 provides the implementation details, Sect. 4 discusses the result and feasibility issues, and the conclusion is provided in the last section.

2 Literature Survey

An IoT-based intelligent trolley using RFID tag and with each cart having RFID reader, microprocessor, 16 * 2 LCD, and an ESP (Espressif) module are discussed in paper [1]. An owner of multiple shops or malls will have the advantage of monitoring the shopping activities using this system. The ESP provides real-time information about the activities going on in all shops from any location. The advantage listed in the proposed model is that the system does not require any special training, customers get information throughout their shopping and queuing delay is reduced. The efficiency is improved because of the use of RFID and emancipates staff from repetitive scanning during checkout.

In paper [2], the population that goes for budget shopping is discussed. A study shows that the influence of real-time shopping feedback for budget shoppers helps them to spend 22% more than the average purchase. The proposed system uses RFID technology and an automated shopping cart to generate the bill for the products in the cart making it a self-serviced store. The key idea of the proposed model was to implement a technology-oriented, low-cost, easily scalable, and rugged system for aiding shopping.

The authors in the paper [3] discuss the components necessary for building automated smart trolley using RFID tag, RFID reader, EEPROM, Zigbee transceiver, LCD, and AT89S52 microcontroller. The specifications in terms of data rate, frequency, communication protocol, range, voltage values, number of inputs, and outputs are elaborated. Details of I2C (Inter-Integrated Circuit) protocol between the EEPROM and microcontroller is given. The challenges faced in this setup are memory size and the need for additional communication protocols.

Paper [4] gives an account of the speeding growth in the shopping mall. The paper talks about the growth in the number of shopping malls between 2000 and 2005, which had increased to about 31% in Mumbai and 21% in Delhi. The author anticipates the growth to exponentially increase every year. The project is implemented using ARM-7 family processors. A detailed description of the methodology is presented to reduce the queuing time spent at the shopping malls. The drawback of the proposed model includes the lack of security in its implementation. No robust algorithms are used to encrypt the data.

Paper [5] proposes a smart shopping cart along with face recognition feature to the user. The customer also is assisted with the recommendation of products based on the history of purchase. Therefore, the purchasing efficiency can further be enhanced and the shopping time can be reduced. However, the algorithms used for navigation of the cart and for face recognition are complex.

In paper [6], the model proposes a concept of Intelligent Shopping Assistant System (ISAS) with steady recognition and navigation of the cart. This model has an additional app feature where the customers can view the shopping list and direct the shopping cart to pick up goods. The app also has product recommendation, information on the discount details. The advantages of the model include a hands-free shopping experience with minimum time spent at the shopping mall.

3 Proposed System

In this paper, the proposed architecture is to provide an automatic billing system using RFID and Wi-Fi technology on a cloud platform. An anti-theft system is incorporated using infrared sensors. A frontend and backend inventory system are created to provide necessary information and generate an E-bill.

Each trolley will have RFID scanner, Infrared sensor, ARM7 microcontroller, and Wi-Fi module attached to it. When a product is selected, the customer has to scan the RFID tag before putting it into the trolley. The product details and bill is extracted from the centralized database system through cloud and the information is displayed on the LCD as well as web page. The bill gets updated each time the item is added or removed accordingly.

The infrared sensor senses each item added to the cart and keeps a count of it. If the total number of products sensed is not equal to the number of items billed, theft is signaled through an alarm. The block diagram of hardware connection is illustrated in Fig. 1. IR sensor is used to count the number of items added in the cart. A passive IR sensor is used in this project. IR sensor constantly emits IR rays which are invisible. Each time there is an obstruction in the path of the rays, which is reflected back to the receiver, the count is increased. This phenomenon is particularly used for anti-theft. If the items added in the cart are not equal to the number of items billed, there is a theft, and hence the buzzer rings to alert the customer and the management.

The overall pin schematic of the hardware connection is given in Fig. 2. The hardware is mounted on the cart and the overall setup is as shown in Fig. 3.

Once the cart is powered, the setup is initialized, i.e., Pin initialization, UART initialization, and Wi-Fi initialization. The RFID reader searches for any tags around its peripheral radius of 35 cm. Once the tag is read, RFID reader sends the information to ARM board to check its existence in its database. The relevant information is displayed on the LCD screen. This information is also sent to the cloud database for billing purpose and for Inventory management. On pressing the interrupt switch, shopping ends and an E-bill is generated. The flowchart for overall process is depicted in Fig. 4 [7–11].

A database named ShoppingCart has been created in MYSQL. Three tables are created in this database, namely, Stock table, Shopping cart application table, and product table. Stock table which consists of stock details is used to notify the management to refill items when the stock falls below a certain threshold value. From the Shopping cart application table, the information of the products that a user is

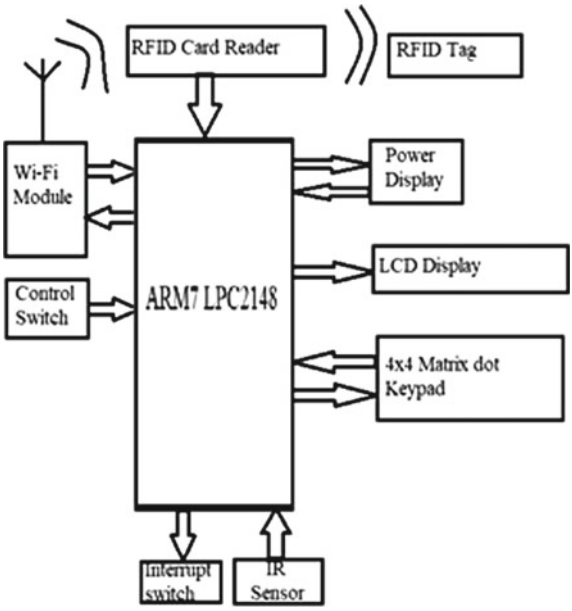


Fig. 1 Block diagram of hardware connections

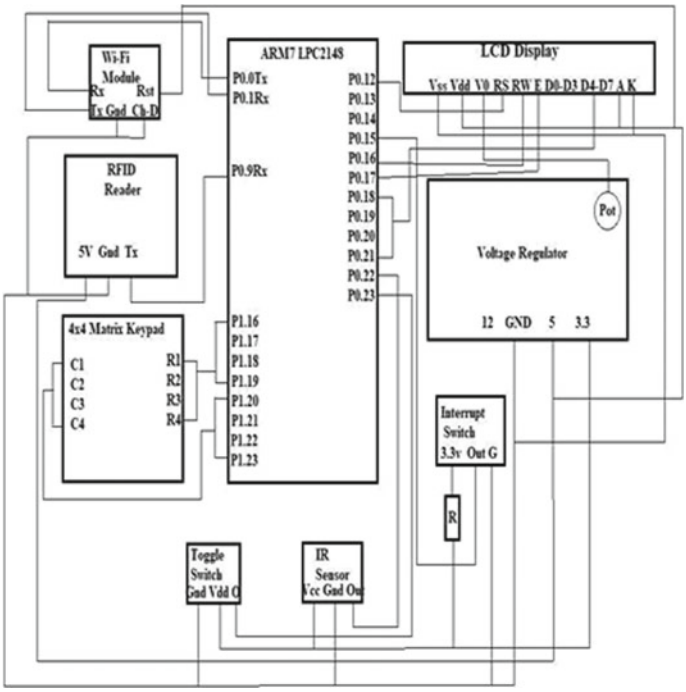


Fig. 2 Pin schematic of the overall setup

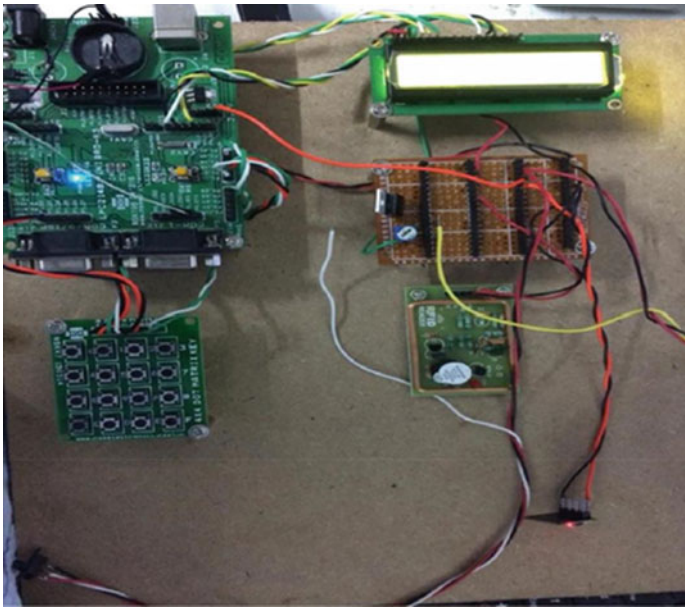


Fig. 3 Overall setup of the hardware

buying is displayed along with the total price. The stock details and shopping cart details from these tables are displayed as shown in Figs. 5 and 6, respectively. The third one, product table is used to map the product IDs with the product names using hashmap [12, 13].

In the message column of the Stock table, a notification about the declining stock is displayed as shown in Fig. 7.

The flow of URL mapping is shown in Fig. 8. The entered URL is sent to the Rest controller where mapping is done which is followed by the MYSQL Access through JDBC to establish a connection with the database.

C and Java languages are used for programming and Digital Ocean is used as a cloud platform to access the uploaded data.

The final E-bill generated is sent to the billing counter via Wi-Fi module and can also be accessed by the customer at any point of time using a URL.

4 Results

The hardware is tested with several RFID tags. Once the cart is powered up and shopping is started, "SHOPPING STARTED" is displayed on the LCD unit as shown in Fig. 9. The Wi-Fi module is initialized before shopping begins and the items can be added or removed using the toggle switch placed on the trolley.

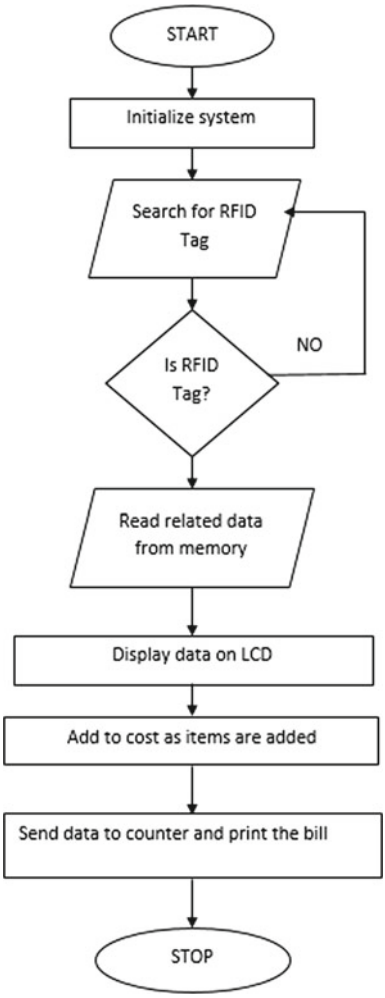


Fig. 4 Process of intelligent trolley

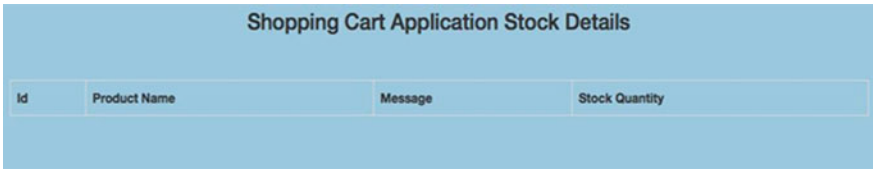


Fig. 5 Stock table

Shopping Cart Application

Id	Date	Shopping Status	Product Name	Product Quantity	Price

Fig. 6 Shopping cart application table

Shopping Cart Application Stock Details

Id	Product Name	Message	Stock Quantity
1	pen		24
2	pencil		35
3	book		30
4	shampoo	shampoo is running out of stock which has only 2 quantity left in stock	2

Fig. 7 Inventory management

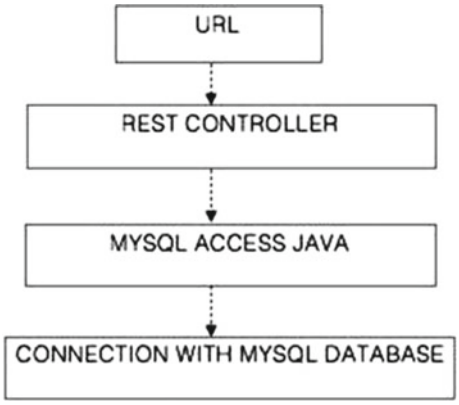


Fig. 8 Flow between database and URL

Fig. 9 Start of shopping

Fig. 10 End of shopping

Shopping Cart Application

Id	Date	Shopping Status	Product Name	Product Quantity	Price
1	02/05/2018 10:47:05	Shopping Started		0	0
2	02/05/2018 10:47:20		pen	1	5
3	02/05/2018 10:48:14		shampoo	2	200
4	02/05/2018 10:48:42	Shopping Stopped	Total	3	205

Fig. 11 Final E-bill

Table 1 Comparison between existing system and proposed system

Number of items purchased	Existing traditional technology (shopping + billing time)	Proposed technology (approx.) (shopping + billing time)	Percentage reduction (approx.)
10	(20 + 10) min	(20 + 2) min	27
20	(35 + 10) min	(35 + 2) min	24

The shopping ends once the interrupt switch present on the trolley is pressed. The final total number of items purchased along with the total cost is displayed on the LCD unit as shown in Fig. 10. An E-bill is generated through URL mapping and is uploaded on the cloud dynamically while shopping. It consists of the number of products purchased, quantity, price of each item, total cost along with the time, and date of purchase as shown in Fig. 11.

Table 1 describes the approximate comparison between the actual shopping time between the conventional method and the proposed system. It is noted that there is a huge reduction in billing time. The average percentage reduction in shopping time results to 26%.

5 Conclusion

In this work, it is observed that customers have advantages in terms of huge reduction in the time spent at the self-checkout counter. The cloud system enables the user and the management to look into the bill through different URLs. The inventory system reduces the burden of manual check on the stock. Further, there is a reduction of 26% in shopping time compared to the conventional shopping methods.

Smart Trolley hardware can be easily attached to any shopping cart chosen by developers. However, if the Smart Trolley is ideally developed and commercialized, the cost can be significantly reduced. The work can be further enhanced by incorporating additional features such as discounts and product recommendation. Also, the shopping time can be significantly reduced by exploiting an inbuilt navigation technology.

References

1. Dhavale Shraddha D, Dhokane Trupti J, Shinde Priyanka S (2016) IOT based intelligent trolley for shopping mall. Int. J. Eng. Dev. Res. 4(2):1283–1285
2. Ali Z, Sonkusare R (2014) RFID based smart shopping: an overview. In: Proceedings of international conference on advances in communication and computing technologies, Mumbai
3. Yewatkar A, Inamdar F, Singh R, Bandal A (2014) Smart cart with automatic billing, product information, product recommendation using RFID and Zigbee with anti-theft. In: Proceedings of 7th international conference on communication, computing and virtualization, Chennai, India
4. Jalkote V, Patel A, Gawande V, Bharadia M, Shinde GR, Deshmukh AA (2013) Futuristic trolley for intelligent billing with amalgamation of RFID and ZIGBEE. Int J Comput Appl 5:0975–8887
5. Chiang H et al (2016) Development of smart shopping carts with customer-oriented service. In: Proceedings of international conference on system science and engineering, Puli, Taiwan
6. Wu B et al (2016) Intelligent shopping assistant system. In: Proceedings of international automatic control conference, Taichung, Taiwan
7. Chandrasekar P, Sangeetha T (2014) Smart shopping cart with automatic billing system through RFID and ZigBee. In: Proceedings of international conference on information communication and embedded systems, Chennai, India
8. Hicks D, Mannix K, Bowels HM, Gao BJ (2013) Smart mart: IoT-based in-store mapping for mobile devices. In: Proceedings of 9th IEEE international conference on collaborative computing: networking, applications and worksharing, 20–23 October 2013
9. Yadav S, Aggarwal S, Yadav M, Gupta N, Karkra S (2016) Ingenious shopping cart: RFID enabled for automated billing. Int J Comput Sci Mob Comput 5(5):209–214
10. Li R, Song T, Capurso N, Yu J, Couture J, Cheng X (2017) IoT applications on secure smart shopping system. IEEE Internet Things J 4(6):1945–1954
11. Manikandan T, Mohammed Aejaz MA, Nithin Krishna NM, Mohan Kumar AP, Manigandan R (2017) RFID based advanced shopping trolley for super market. J Chem Pharm Sci
12. Gangwal U, Roy S, Bapa J (2016) Smart shopping cart for automated billing purpose using wireless sensor networks. In: Proceedings of the seventh international conference on sensor technologies and applications, pp 168–172
13. Berdaliyev Y, James A (2016) RFID-cloud smart cart system. In: Proceedings of international conference on advances in computing, communications and informatics, Jaipur, India

Long Short-Term Memory-Convolution Neural Network Based Hybrid Deep Learning Approach for Power Quality Events Classification

Rahul, Kapoor Rajiv and M. M. Tripathi

Abstract This paper presents novel method of deep learning for automatic selection of features and classification of power quality events. The power quality events analysis based on hybrid approach of long short-term memory-convolution neural network is proposed here to ensure reliability, security of supplied power. It is necessary to recognize and classify the power quality events. Power quality disturbance classification methods mostly based on unique feature extraction such as statistical behavior, spatial-temporal information, nonlinear, and nonstationary characteristics of power quality disturbances. The performance of the presented long short-term memory-convolution neural network is evaluated on a set of synthetic power quality disturbances. The developed model is shown to be novel for feature training and classification of various disturbances. Totally different from conventional methods such as Support vector machine (SVM), fuzzy logic, probabilistic neural networks (PNN), and artificial neural networks (ANN). The achieved results show that deep learning based classifier is more efficient than the earlier state of art. In addition, the proposed method can be efficiently used to classify the complex power quality events.

Keywords Long Short-Term Memory (LSTM) · Convolution neural network (CNN) deep learning (DL) · Power quality (PQ) disturbances

1 Introduction

Deep learning is a part of machine learning techniques that is based on human brain's ability to learn and react. However, proficiency of training data and computational power makes deep learning efficient in complex pattern recognition and power quality disturbances analysis. Power quality disturbances are defined by the international

Rahul (✉) · K. Rajiv
EC Department, Delhi Technological University, Delhi, India
e-mail: rahulverma@dtu.ac.in

M. M. Tripathi
EE Department, Delhi Technological University, Delhi, India

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9_52

standards, such as IEC 61000-4-30 based on time, frequency, phase and magnitude of voltage and current of power quality events. It defines the disturbances as the deviation of voltage or current out of a predefined tolerance limit. The main reason which makes power quality events a serious issue is that it can be the possible reason for the malfunctioning of the electrical instruments in the power system. Increasing demand of solid state switching, microprocessor-based power electronic devices, heavy nonlinear loads and unbalanced power systems are the major cause of power quality events. The power quality disturbances such as sag, swell, transient, interruption, flickers and spikes and their combinations degrade the quality of supplied power. These disturbances cause disruption of end-user equipment and huge economic loss. Detection of power quality events is the first step to develop compensatory equipment for improving the quality of supplied power.

The power quality disturbances can be divided into two classes such as stationary and nonstationary signals. In most of the cases, power quality events are nonstationary in nature because of heavy load changing or switching continuously in power grid. In the last decade, numerous time and frequency domain methods have been applied for monitoring of power signal disturbances. These methods include Fourier transforms (FT) [1], discrete Fourier transforms (DFT), short time Fourier transforms (STFT), wavelet transforms (WT) and S-transform for analyzing of power quality events. The FT has been the suitable choice for stationary disturbances in power signal. STFT is the enhancement of DFT which is inefficient for the monitoring of the nonstationary signal due to fixed window size. The drawbacks of STFT have been corrected using wavelet transform. The various class of FT is not feasible choice to analyze the nonstationary signals and DFT has some limitations of resolution and leakage in spectrum. The major limitation of STFT is that it cannot use large window size at small frequencies and short window size at high frequencies DWT is the discrete form of WT and which extended in the form of discrete wavelet packet transform [2]. In WT power quality signal can be breakdown into distinct frequency bands by the concept of multiresolution analysis and used to calculate various parameters such the RMS values, total harmonic distortion and DC values [3]. Moreover, selection of suitable mother wavelet and the sampling frequency are the major issues of WT [4, 5]. S-transform is mainly used for feature extraction purpose but not appropriate for real-time application due to heavy complex computations [6–8]. HHT is mainly explored for detection point of view but having major drawback of signal end point contamination [9]. A Volterra series technique is also implemented for power quality events detection with the features of power spectral entropy, standard deviation, and kurtosis but having large computational overheads [10, 11].

Numerous classification techniques are implemented in earlier state of the art work such as artificial neural network (ANN) [12] and probabilistic neural network (PNN) [13], support vector machine [14], decision tree (DT) [15], multilayer perceptron (MLP) [5] and hidden Markov model (HLM) [3]. All the extracted features are fed as input to these classification methods. These classification techniques are having drawbacks of heavy complexity, this limiting the implementation in real-time conditions. To overcome these restrictions a proficient, systematic, and time-efficient method is required to address the issue of power quality.

In this article, deep learning, which is a robust classifier in scenario of massive data to train, is applied here for power quality disturbances data analysis. In comparison with earlier state of the art, proposed methodology takes the advantage of image file based classification in contrast with sampled database analysis techniques, sampled data of the power quality disturbances are not exercised here, but image files of the disturbances are used by utilizing the superiority of deep learning approach on image file based classification. As in PQ disturbances analysis, we are dealing with patterns such as voltage sag, swell, harmonics and flickers, etc., DL is competent of analyzing these patterns in separate level of information. Each and every layer utilize previous layer data such as input and output for processing result in a large gain in the form of system's capability to generalize power quality image data. Hence, the uniqueness of the proposed methodology is that images of power quality disturbances are classified with high accuracy rather than classifying the sampled data strings of power quality events.

2 Fundamentals of Deep Learning

Deep learning is a robust and time-efficient technique in machine learning, artificial intelligence, and data processing. Most of the machine learning algorithm works on extracted features, which needs tedious effort to extract adequate features. Deep learning approaches are data-dependent while feature extraction aspect is function specific. In artificial intelligence methods, data are continuously generating and a stage will come where automatic extraction of feature is impossible but this is not the case in deep learning techniques. DL process combines training which includes interactive learning of data samples and inference step extracts the output from trained model. In training network, huge data is decomposed into training and testing sets, and finally a validation module as well. A machine learning algorithm fed the given training set data to review some representation that can be in terms of function approximation of the given feature extraction or in the form of decision sets based on the addition of each feature to others.

The role of validation comes into picture during training, in terms of method to validate the robustness of the training action, obtained results fed to tune learning module of the algorithm and revamp the concluding efficiency. The final test set then gives unobserved data set to evaluate the last accuracy of the training module and is broadly the source of the obtained accuracy scores and other potentiality metrics. The role of inference mainly is to feeding a data set into a trained and executed machine learning module and results in inferred output. The initial small data fed into the deep architecture are adequate to train features and to classify. A simple deep learning network has multiple deep layers to extract unique features of power quality disturbance's data. The yield from previous layer is passed as input to the next layer and trains multiple levels of sparse data [16]. For instance, the first layer can extract information about voltage amplitude, and the second layer can collect the information how much voltage dip occurred and the third gives the information

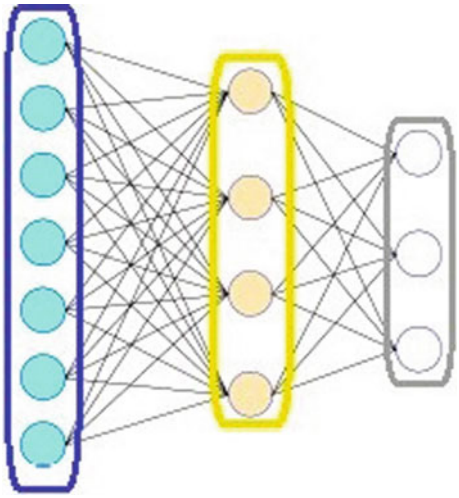


Fig. 1 Basic deep learning network

which portion have voltage sag in power signal and the final layer combines all the information from previous layers and takes decision regarding classification of power quality disturbance. A general deep learning model has an input layer, several hidden layers and finally output layers as shown in Fig. 1.

In Fig. 1, first blue layer is the input layer that takes the input data sets and then forwards the information to multiple middle, yellow hidden layers, and finally output grey layers classify the signals on the basis of training data sets and inference system build up in middle hidden layers.

3 Proposed Hybrid Network of Long Short-Term Memory (LSTM)-Deep Convolution Neural Network

LSTM is a specific type of recurrent neural network that solves the problem of vanishing gradient [17]. The block diagram of an LSTM module is shown in Fig. 2. The main concept behind LSTMs is a memory cell network M_t which is capable to maintain state information for a long time that can make gradients to flow over entire sequences.

The data set information going in and out of the memory cell network M_t is controlled by three gates: an input gate p_t, a forget gate q_t and an output gate r_t. For every clock time instant, the LSTM cell collects input data of the present time y_t and the hidden information z_{t-1} from the previous state. The multiplication of input y_t and hidden information z_{t-1} is analyzed by passing over a tanh function:

$$d_t = \tanh\left(W_g y_t + U_g z_{t-1} + b_g\right) \tag{1}$$

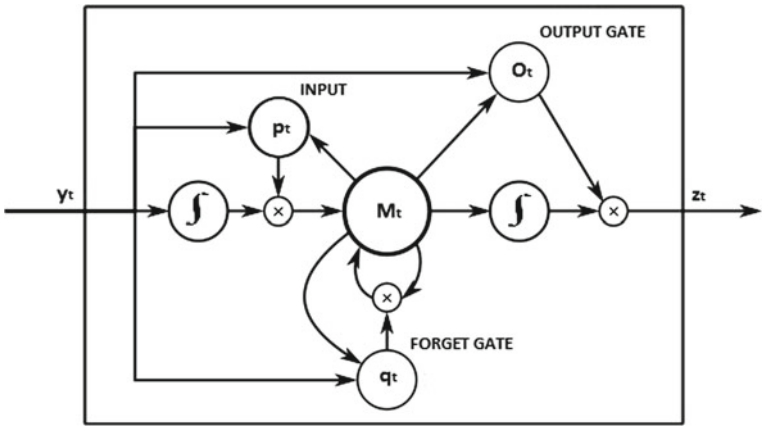


Fig. 2 Block diagram of LSTM

Here, W_g, U_g and b_g represents input weights, recurrent weights and bias, respectively. The role of forget gate q_t is to decide which data should be erased from the memory M_t over an element-wise sigmoid function:

$$q_t = \sigma\left(W_q y_t + U_q z_{t-1} + b_q\right) \tag{2}$$

However, the input gate takes a decision regarding which data will be saved in the memory cell M_t over an element-wise sigmoid function:

$$p_t = \sigma\left(W_p y_t + U_p z_{t-1} + b_p\right) \tag{3}$$

Further, the information in the memory is refreshed through fractional erasing of data saved in the last memory cell M_{t-1} through the following step:

$$M_t = q_t * M_{t-1} + p_t * d_t \tag{4}$$

Here, $*$ stands for element-wise product operation of two constituent. Now the role of forgetting gate q_t at this stage is to decide the extent to which previous data saved in M_{t-1} will be clean up. The set q_t having value between 0 and 1. If $q_t \to 0$, the previous state data will be completely clean up and if $q_t \to 1$, the previous data will be stored as it is in the memory. Finally, the output hidden state h_t will be refreshed on the basis of evaluated cell state M_t as shown:

$$z_t = o_t * \tanh(M_t) \tag{5}$$

Here, network input weights $W_{\{d,p,q,o\}}$, bias $b_{\{d,p,q,o\}}$, and recurrent weights $U_{\{d,p,q,o\}}$ are computed through the training and inference process.

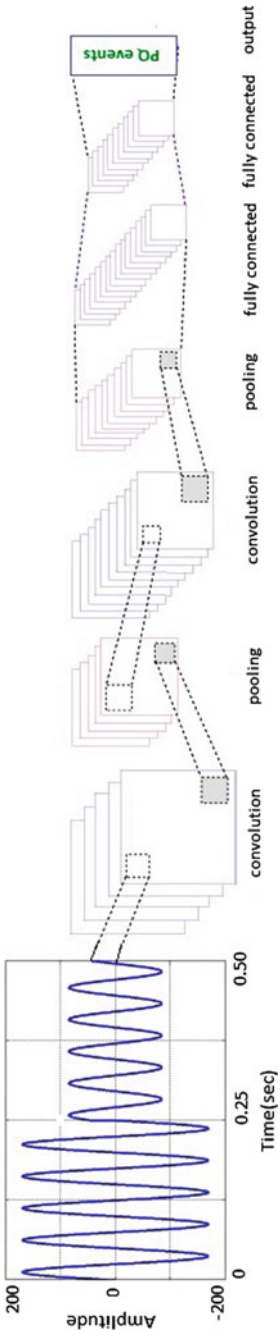


Fig. 3 CNN architecture for power quality disturbance

The CNN architectures made up of three main layers, first layer is convolution layer, pooling layer is second layer, and the fully connected layer is the last layer. The role of fully-connected layer is same as in the general neural networks and convolution layer do the job of convolution on the previous layer many times [18]. The task of downsampling is assigned to pooling layer, maximum of each stage of the previous layer. The stacked network of three layers makes complete CNN architecture as shown in Fig. 3.

A convolution neural network (CNN) is a generalized form of recursive neural network (RNN). In case of RNNs, iteration process is executed to recursively restore the weights of the neurons based on the loss function obtained after each learning stage. In the architecture of CNN, alternating convolution and pooling layers are combined before fully connected layer. The task of convolution layers is to filter out huge multidimensional matrices [19]. The function of pooling layer is to spatially shorten the size of the feature vectors into a smaller and more feasible matrix. In abstract, the convolution layers overcome the complexity of the power quality disturbance image by some filtering operation which does the job of sharpening and edge detection and the pooling layers perform crucial task of compression of filtered data. Role of multiple filters is generally to extract parallel and comprehensive features.

4 Proposed Methodology for Classification of Power Quality Events

In this article, to enhance the efficiency of power quality disturbance classification by learning some supplementary features, hybrid architecture of LSTM-CNN is implemented here as shown in Fig. 4, many steps involved in the novel hybrid architecture. The proposed architecture comprises of an initial layer of CNN then next stage is max pooling which performs the task of finding low-level dependencies in layers. To develop a concise feature map, the features extracted in first stage are applied as input to subsequent layer of CNN and max pooling layers. The first CNN layer contains 128 filters with a filter length of four for convolving input information. The next CNN layer has 256 filters with length of four. The max pooling layer has normal size of 3, which mixes the features from last layer. The final combined feature matrix developed from max pooling layer are then applied to an LSTM layer having memory cells. The LSTM layer finds large temporal dependencies and develops a more concise feature vector. The LSTM-CNN novel architecture having 100 memory cells containing many gates and activation controllers. The final output of LSTM layer has passed to a last stage, i.e., fully connected layer. Commonly, a feed-forward neural network layer performs the job of fully connected layer, which involves units conforms to the number of power quality events classification. The fully connected layer having nonlinear activation module termed as softmax task is to evaluate the probabilities of accurate classification of power quality events.

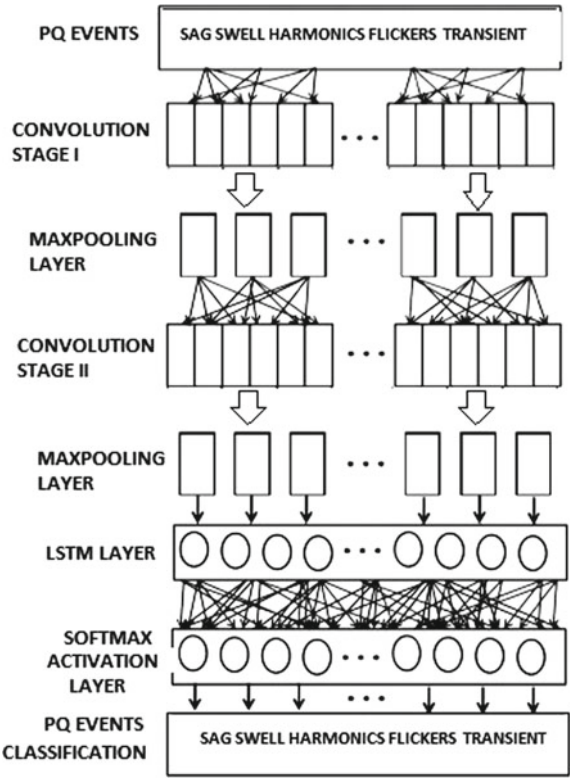


Fig. 4 Hybrid architecture of LTSM-CNN for PQ events classification

For LSTM-CNN, the performance is totally relying on the number of filters used in the network. In the article two form of experiments performed with novel LSTM-CNN architecture and max pooling layer by changing the number of filters as 64, 128, 256 and filter length as 4, 5, and 6. The CNN model with number of filters as 128, each filter having a length of 4 has achieved the highest accuracy and hence these values are appropriate for detailed study of power quality disturbances. After setting the network parameters, experiments are performed to select the most suitable network model for all deep learning process.

5 Results and Discussion

In the proposed novel architecture of LSTM-CNN, power quality events are generated synthetically by using parametric equation in MATLAB for voltage sag, swell, harmonics, flickers, and transient events. These events then applied as a raw data to the proposed architecture; first, the various features are extracted by LSTM network. Features extracted in each layer are flowing through to next subsequent layer and last

Table 1 The accuracy of deep learning algorithms for power quality disturbances

PQ events	LSTM		CNN		LSTM-CNN	
	TPR	FPR	TPR	FPR	TPR	FPR
Sag	0.91	0.098	0.91	0.098	0.097	0.007
Swell	0.97	0.003	1	0	1	0
Harmonics	1	0	0.99	0.098	1	0
Flicker	0.97	0.039	1	0	1	0
Transient	0.99	0.05	0.94	0.099	0.98	0.006
Accuracy	0.956		0.97		**0.989**	

LSTM Long Short-Term Memory, *CNN* Convolution neural network, *TPR* True positive rate (Classifier parameter), *FPR* False positive rate (Classifier parameter)

activation layer is capable to separate the data into different classes of power quality events and then extracted features fed as input to the CNN network which performs the function of classification for various power quality events. The results obtained for different PQ events are mentioned in Table 1.

The LSTM model has attained an accuracy of 95.6% for five different power quality disturbances occurring in power system. The second model of CNN achieve the accuracy of 97% little better performance than LSTM model, but proposed novel architecture of LSTM-CNN model attained the highest accuracy of 98.90%. In comparison of conventional classifiers such as artificial neural network (ANN) [12] and probabilistic neural network (PNN) [13], support vector machine [14], decision tree (DT) [15] proposed novel methodology of LSTM-CNN achieved highest accuracy of 98.90%. The conventional PQ disturbance classification techniques are not time efficient mainly because of extraction of features. While the deep learning architecture has eliminated the stage of the feature extraction and that's why the processing time is less than other methods.

6 Conclusions

In this article, a deep learning based novel model is developed for five different power quality events such as voltage sag, swell, harmonics, flicker and transient. In contrast with the previously proposed techniques of PQ disturbances sampled data set of the PQ events are not utilized for feature extraction, but image files of the PQ disturbances are implemented by taking the advantage of the deep learning architecture on image-based classification. This justifies the novelty of the proposed method to classify PQ disturbances with 98.90% accuracy, rather than classifying the sampled data set of PQ events with low accuracy. This reveals that the proposed novel approach generates a more inventive system with classification performance better

than that of each and every approach discussed in Table 1. This confirms that the proposed method is applicable for real-time analysis of power signal disturbances.

References

1. Heydt G, Fjeld P, Liu C, Pierce D, Tu L, Hensley G (1999) Applications of the windowed FFT to electric power quality assessment. IEEE Trans Power Deliv 14(4):1411–1416
2. Vatansever F, Ozdemir A (2009) Power parameters calculations based on wavelet packet transform. Int J Electr Power Energy Syst 31(10):596–603
3. Dehghani H, Vahidi B, Naghizadeh R, Hosseinian SH (2013) Power quality disturbance classification using a statistical and wavelet-based hidden markov model with dempster–shafer algorithm. Int J Electr Power Energy Syst 47:368–377
4. Eris H, Demir Y (2012) Automatic classification of power quality events and disturbances using wavelet transform and support vector machines. IET Gener Transm Distrib 6(10):968–976
5. Uyar M, Yildirim S, Gencoglu MT (2008) An effective wavelet-based feature extraction method for classification of power quality disturbance signals. Electr Power Syst Res 78(10):1747–1755
6. Uyar M, Yildirim S, Gencoglu MT (2009) An expert system based on S-transform and neural network for automatic classification of power quality disturbances. Expert Syst Appl 36(3):5962–5975
7. Erişti H, Yıldırım Ö, Erişti B, Demir Y (2014) Automatic recognition system of underlying causes of power quality disturbances based on S-transform and extreme learning machine. Int J Electr Power Energy Syst 61:553–562
8. Salem ME, Mohamed A, Samad SA (2010) Rule based system for power quality disturbance classification incorporating S-transform features. Expert Syst Appl 37(4):3229–3235
9. Shukla S, Mishra S, Singh B (2009) Empirical-mode decomposition with hilbert transform for power-quality assessment. IEEE Trans Power Del 24(4):2159–2165
10. Rahul, Kapoor R, Tripathi MM (2017) Prot Control Mod Power Syst 2:9. https://doi.org/10.1186/s41601-017-0039-z
11. Kapoor R, Kumar R, Tripathi MM (2017) Volterra bound interval type-2 fuzzy logic-based approach for multiple power quality events analysis. IET Electr Syst Transp. https://doi.org/10.1049/iet-est.2017.0054
12. Biswal B, Biswal M, Mishra S, Jalaja R (2014) Automatic classification of power quality events using balanced neural tree. IEEE Trans Ind Electr 61(1):521
13. Bhende S, Mishra B (2008) Panigrahi, Detection and classification of power quality disturbances using S-transform and modular neural network. Electr Power Syst Res 78(1):122–128
14. De Yong D, Bhowmik S, Magnago F (2015) An effective power quality classifier using wavelet transform and support vector machines. Expert Syst Appl 42(15–16):6075
15. Ray PK, Mohanty SR, Kishor N, Catalao JP (2014) Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems. IEEE Trans Sustain Energ 5(1):200–208
16. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
17. Hochreiter S, Schmidhuber J (1997) Long term short memory. Neural Comput 9(8):1735–1780
18. Wu F, Hu P, Kong D (2015) Flip-rotate-pooling convolution and split dropout on convolution neural networks for image classification, arXiv preprint arXiv: 1507.08754
19. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 1097–1105

Author Index

H. S. Saini et al. (eds.), *Innovations in Electronics and Communication Engineering*, Lecture Notes in Networks and Systems 65,
https://doi.org/10.1007/978-981-13-3765-9