Lecture Notes in Electrical Engineering 887

Vijay Nath
Jyotsna Kumar Mandal *Editors*

Microelectronics, Communication Systems, Machine Learning and Internet of Things

Select Proceedings of MCMI 2020

Springer

Lecture Notes in Electrical Engineering

Volume 887

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at https://link.springer.com/bookseries/7818

Vijay Nath · Jyotsna Kumar Mandal
Editors

Microelectronics, Communication Systems, Machine Learning and Internet of Things

Select Proceedings of MCMI 2020

Springer

Editors
Vijay Nath
Department of Electronics
and Communication Engineering
Birla Institute of Technology
Ranchi, Jharkhand, India

Jyotsna Kumar Mandal
Department of Computer Science
and Engineering
University of Kalyani
Kalyani, West Bengal, India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-19-1905-3 ISBN 978-981-19-1906-0 (eBook)
https://doi.org/10.1007/978-981-19-1906-0

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

Preface

This volume contains the papers that were presented at the First International Conference on **Microelectronics, Communication Systems, Machine Learning and Internet of Things(MCMI-2020)** organized by the Indian Society for VLSI Education (ISVE) and IETE Ranchi Centre with support of BSNL at ARTTC BSNL, Ranchi, during 12–13 September 2020. It provided a great platform for researchers from across the world to report, deliberate, and review the latest progress in the cutting-edge research pertaining to smart computing and its applications to various engineering fields. The response to MCMI-2020 was overwhelming with a good number of submissions from different areas relating to artificial intelligence, machine learning, Internet of Things, VLSI design, micro- and nanoelectronics circuits and devices, green energy, communication systems, signal processing, smart computing, computational intelligence, agriculture cyberphysical system, and its applications in main tracks. After a rigorous peer review process with the help of programme committee members and external reviewers, only quality papers were accepted for publication in this volume of LNEE series of Springer. Several special sessions were offered by eminent professors in many cutting-edge technologies. Several eminent researchers and academicians delivered talks addressing the participants in their respective field of proficiency. Our thanks are due to Prof. P. S. Neelakanta, Florida Atlantic University, USA; Prof. Gopal Pathak, Vice-Chancellor, Jharkhand Technical University, Ranchi; Prof. Nand Kumar Yadav 'Indu', Vice-Chancellor, Central University of Jharkhand; Sh. Anurag Dubey, Xilinx, USA; Prof. Abhijit Biswas, University of Kolkata; Prof. A. A. Khan, Former VC, Ranchi University; Sh. K. K. Thakur, CGMT, BSNL, Ranchi; Dr. Raj Kumar Singh, Ranchi University; Sh. Ajay Kumar, Governing Council Member, IETE New Delhi; Sh. Viney Kakkar, Treasurer, IETE New Delhi; Prof. D. K. Yadav, NIT Jamshedpur; Dr. Shylashree N., RVCE, Bangalore; Dr. Sujata Sanjay Kotabagi, KLE Technological University, Hubballi; Prof. Rama Komaragiri, Bennett University, Greater Noida; Dr. Ambrish Mishra, NIT Patna; Dr. Dharmveer Yadav, Katihar College of Engineering, Katihar; Dr. Ritu Sharma, MNIT Jaipur, and Prof. M. Bhaskar, NIT Trichy, for their valuable talks/session chairs for the benefits of the participants. We would like to express our appreciation to the members of the programme committee for their support and cooperation in this publication. We are also thankful to the team from Springer for providing a meticulous service for the timely production of this volume.

Our heartfelt thanks to our loving founder Dr. J. W. Bakal, President, IETE New Delhi, and Sh. K. K. Thakur, Chairman, IETE Ranchi Centre; Prof. P. R. Thakura, President, ISVE Ranchi, and Executive Committee of IETE and ISVE Ranchi for extending excellent support to host this in ARTTC BSNL Campus. Professor P. S. Neelakanta, Florida Atlantic University, USA, and Professor Ramjee Prasad, Arhus University, Denmark, deserve a big round of applause from all of us for their continuous guidance and support from the beginning of the symposium. Without their support, we could never have executed such a mega event.

Special thanks to all guests who have honoured us in their presence in the inaugural day of the conference MCMI-2020. Our thanks are due to all special session chairs, track managers, and reviewers for their excellent support. Last but not least, our special thanks go to all the authors who submitted papers and all the attendees for their contributions and fruitful discussions that made this conference a great success.

Ranchi, JH, India Vijay Nath

Acknowledgements

We extend our thanks to all the authors for contributing to this book/proceeding by sharing their valuable research findings. We specially thank all reviewers for promptly reviewing the papers submitted to the conference **MCMI-2020**. We are grateful to the volunteers, invited speakers, session chairs, sponsors, sub-committee members, members of International advisory committee, and members of national advisory committee, members of technical programme committee, members of joint secretary, and members of scientific advisory committee for successful conduct of conference. The editor expresses their heartfelt gratitude towards **Dynamic and Visionary Chief Minister of Jharkhand Shri Hemant Soren** and **Honourable Governor of Jharkhand Smt. Droupadi Murmu**. The editor also express their heartfelt gratitude towards Prof. Ramjee Prasad, Professor of Aarhus University, Denmark, and Founder President of CTIF Capsule; Dr. J. W. Bakal, President of IETE New Delhi; Dr. K. T. V. Reddy, Immediate Past President of IETE New Delhi; Dr. A. K. S. Chandele, Past President of IETE New Delhi; Smt. Srimati Dagur, Former President of IETE New Delhi; Dr. Ajit Khosla, Professor of Yamagata University, Japan; Prof. S. Pal, Former VC, IAT Pune, and Former Scientist of SAC ISRO; Dr. D. K. Tiwari, Chief Secretary of Government of Jharkhand; Sri Vinay Kumar Choubey, Principal Secretary IT, Government of Jharkhand; Sh. Rakesh Sharma, Principal Secretary of Higher Education Government of Jharkhand; Sh. Sudhir Tripathy, Former Chief Secretary of Government of Jharkhand; Sh. Sanjay Kumar Jha, Former Chairman of IETE Ranchi, and Executive Engineer Government of Jharkhand; Prof. Bernd Michel, Micro Materials Centre (MMC), Berlin, Germany; Prof. Bharath Bhushan, Ohio Eminent Scholar, and The Howard D. Winbigler, Professor and Director of NBLL, The Ohio State University, Columbus, Ohio, USA; Prof. P. S. Neelakanta, C. Engg. Fellow of IEE Florida Atlantic University (FAU), USA.; Sh. K. K. Thakur, Chairman of IETE Ranchi, and CGMT (Rtd.), BSNL, Jharkhand Circle, Ranchi; Sh. Prasad Vijay Bhushan Pandey, DTO Term Cell1 BSNL, Ranchi, and Chairman of ISVE Ranchi; Prof. A. A. Khan, Former VC of Ranchi University; Prof. M. K. Mishra, Former VC of BIT Mesra; Prof. Gopal Pathak, VC of Jharkhand Technical University, Ranchi; Prof. Nand Kumar Yadav, 'Indu', VC of Central University of Jharkhand, Ranchi; Prof. R. P. Yadav, Professor (HAG) of MNIT Jaipur, and Former

VC of University of Jharkhand; Prof. R. K. Pandey, VC of Ranchi University, Ranchi; Prof. S. Konar, Dean (Faculty Affairs and Sponsored Research), and Former VC, BIT Mesra; Prof. I. Manna, VC of BIT Mesra, Ranchi, and Former Director, IIT Kanpur; Sh. R. Mishra, Former CMD, HEC, Ranchi; Dr. M. Chakraborty, Professor, IIT Kharagpur; Dr. Ramgopal Rao, Professor, IIT Bombay, and Director, IIT Delhi; Dr. P. Chakraborty, Professor, IIT BHU, and VC, BESU, Kolkata; Dr. S. Jit, Professor, IIT BHU; Dr. J. K. Mandal, Professor, Kalyani University; Dr. Sujata Sanjay Kotabagi, Professor, KLE Technological University, Hubballi; Dr. Rama Komaragiri, Professor, Bennett University, Greater Noida (UP); Dr. Ambrish Mishra, Assistant Professor, NIT Patna; Dr. Abhijit Biswas, Professor, Kolkata University; Dr. Subir Kumar Sarkar, Professor, Jadavpur University; Dr. Gaurav Trivedi, Associate Professor, IIT Guwahati; Dr. Y. S. Chauhan, Professor, IIT Kanpur; Dr. B. K. Kaushik, Professor, IIT Roorkee; Dr. Shree Prakash Tiwari, Assistant Professor, IIT Jodhpur; Dr. P. Kumar, Associate Professor, IIT Patna; Dr. M. Bhaskar, Professor, NIT Trichy; Dr. Adesh Kumar, Associate Professor, UPES University, Dehradun; Dr. Manish Kumar, Associate Professor, MMMUT, Gorakhpur; Dr. Manish Mishra, Professor, DDU University, Gorakhpur; Dr. Umesh Yadav, Professor, DDU University, Gorakhpur; Dr. Arpan Deyasi, Assistant Professor, RCC Institute of Technology, Kolkata; Dr. N. Gupta, Professor, BIT Mesra, Ranchi; Dr. S. Pal, Professor, BIT Mesra; Dr. Vibha Rani Gupta, Professor, BIT Mesra; Dr. B. K. Mishra, Principal, Jumeritelaya Government of Jharkhand; Dr. V. K. Jha, Associate Professor, BIT Mesra; Sh. Ajay Kumar, AGM (admin) ARTTC BSNL, Ranchi, and Imm Past Chairman, IETE Ranchi, and GC Member, IETE New Delhi; Dr. M. Chandra, Former Executive Member of IETE Ranchi, and Former Professor of BIT Mesra Ranchi; Dr. S. K. Ghorai, Executive Member of IETE Ranchi and Professor of BIT Mesra, Ranchi; Dr. Dharmveer Yadav, Assistant Professor of Katihar College of Engineering, Katihar; Dr. B. Chakraborty, Executive Member of IETE Ranchi, and Executive Engineer, Mecon, Ranchi; Dr. S. Chakraborty, Executive Member of IETE Ranchi, and Professor of BIT Mesra, Ranchi; Dr. Vandana Bhattacharya, Professor of BIT Mesra, Ranchi; Dr. D. Mohanta, Professor of BIT Mesra; Dr. S. S. Solanki, Professor of BIT Mesra Ranchi; Dr. S. Pal, Professor of BIT Mesra, Ranchi; Dr. B. Acharya, Assistant Professor of NIT Raipur; Dr. Chandrashekhar, Assistant Professor of NIT Jamshedpur; Dr. D. K. Yadav, Professor of NIT Jamshedpur; Dr. Subhash Yadav, Professor of Central University of Jharkhand Ranchi; Dr. S. Kumar, Executive Member of IETE Ranchi, and Associate Professor of BIT Mesra Ranchi; Dr. B. K. Bhattacharya, Professor of NIT Agartala; Dr. N. Shylashree, Associate Professor, RVCE, Bengaluru; Dr. V. Anandi, Associate Professor of Ramaih Institute of Technology, Bengaluru; Dr. Ritu Sharma, Associate Professor of MNIT, Jaipur; Dr. Anand Kumar Thakur, Treasurer of IETE Ranchi, and Assistant Professor and Director of FM Ranchi University; Dr. Raj Kumar Singh, Executive Member of IETE Ranchi, and Assistant Professor and Coordinator of UGC Refresher Course, Ranchi University, Dr. R. K. Lal, Associate Professor of BIT Mesra, Ranchi, Dr. Sudip Sahana, Dr. P. Pal, Dr. Amritanjali, Dr. Rishi Sharma, Dr. K. K. Senapati, Dr. S. S. Sahu, Dr. Aminul Islam, Dr. Sudip Kundu, Dr. S. Pushkar, Dr. M. K. Mukul, Dr. K. Bose of BIT Mesra; Dr. Parul Patel Head of RESPOND, SAC ISRO, Ahmadabad; Dr. H. S. Gupta, Sh. Manish Mehta;

Sh. V. Vedam, Senior Scientist, and Smt. Ankita Patel, SAC ISRO, Ahmadabad; Smt. Saroj Treasurer of ISVE Ranchi; Prof. Jyoti Singh, Joint Secretary of ISVE Ranchi; Dr. A. K. Pandey, Secretary of ISVE Ranchi; Sh. Suraj Kumar Saw, IIIT Surat, Sh. Subro Chakraborty, Executive Engineer of IBM Bangalore; Sh. Dipayan Ghosh, Executive Engineer of Aircell; Sh. Ramkrishna Kundu, Executive Member of ISVE Ranchi, and Design Engineer, IBM; Ashish Kumar, QUALCOMM; Sh. Deepak Prasad, SRF CSIR BIT Mesra; Mohd Javed Khan, Ms. Namrata Yadav, Sh. Om Prakash, MentorGraphics, Noida; Sh. Rohan Kumar, MentorGraphics; Sh. Ankit Kumar, Synopsis, Bangalore; Sh. Sumit Singh, IBM Ahmedabad; Sh. Hiranmoy Kar, Sh. Rajanish Yadav, Sh. Aditya Nath, Sh. Shrill Kumar for their endless support, encouragement, motivation to organize such prestigious event that paved the way for this book on Proceeding of First International Conference on **Microelectronics, Communication Systems, Machine Learning and Internet of Things (MCMI-2020)**. At last, we express our sincere gratitude towards the staff members of Springer (India) who helped in publishing this book.

Contents

About the Editors

Vijay Nath received his Bachelor's degree in physics and Master's degree in electronics from DDU Gorakhpur University, in 1998 and 2001, respectively. He received PGDCN (GM) from MMMUT Gorakhpur in 1999 and a Ph.D. degree in VLSI Design and Technology from Dr. RML Avadh University Ayodhya in association with CEERI-Pilani in 2008. He is presently an Associate Professor in the Department of ECE, BIT Mesra Ranchi. His research interests include micro and nano-electronics, analog and digital VLSI design, ASICs, embedded systems designs, Internet of Things, and machine learning. He has to his credit around 195 publications in reputed journals and conferences and three scholars were awarded Ph.D. degrees under his supervision.

Jyotsna Kumar Mandal received his M.Sc. in Physics from Jadavpur University in 1986 and M.Tech. in Computer Science from the University of Calcutta and was awarded a Ph.D. in Computer Science and Engineering by Jadavpur University in 2000. Currently, he is a Professor of Computer Science and Engineering and was Dean of the Faculty of Engineering, Technology, and Management at Kalyani University, West Bengal for two consecutive terms. He has 30 years of teaching and research experience. His research interests include coding theory, data, and network security; remote sensing and GIS-based applications, data compression, error correction, visual cryptography, steganography, security in MANET, wireless networks, and unified computing. He has over 30 proceedings volumes, six books, 160 journal publications to his credit. He has supervised 23 scholars for their Ph.D. degrees.

Design, Implementation, and Study of an IoT-Based Battery Life Cycle Tester and State of Charge Indicator

Shinjini Mukherjee, Sudipta Sarkar, Ashoke Mondal, and Shilpi Bhattacharya

Abstract Batteries have varied uses in our day-to-day lives starting from using them in several electronic instruments like mobile phones, UPS, car batteries to even submarines, and non-conventional power plants. Even with the increase in electric vehicles, there is also greater demand of batteries. So, to know the quality of the battery, a benchmarking is very necessary to find out the battery with highest reliability and life. In this paper, we have illustrated an IoT-enabled battery life cycle tester, which comprises intelligent charging and discharging units, measurement units, data logging, remote monitoring, and communication sections for data acquisition (Mutagekar et al. in Designing a high performance battery life cycle tester. 2016 first international conference on sustainable green buildings and communities (SGBC), 2016 [1]). For this device, input parameters could be manually set depending upon the specification of the battery (e.g., rated battery voltage, battery capacity, full charged voltage, low battery voltage, and charging/discharging rate), thus making it compatible with all types of batteries. This device charges and discharges the given battery continuously, monitors the parameters, and logs the data. The gathered data are further used to decide upon the state of charge/depth of discharge and also predict the life cycle of the battery without spending the entire time of complete life cycle of the battery. This device can also precisely measure the actual capacity of a battery. State of charge indicator is like a fuel gage especially in the application areas of electric vehicle. Suitable measures have been taken to improve the life of the battery by preventing substantial damage to the battery by prevention of overcharging or deep discharging at specified temperatures.

S. Mukherjee (✉) · S. Sarkar · A. Mondal · S. Bhattacharya
RCC Institute of Information Technology, Kolkata 700015, India
e-mail: tiai.mukherjee@gmail.com

S. Sarkar
e-mail: sudiptasarkar400@gmail.com

A. Mondal
e-mail: ashoke.mondal@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_1

Keywords Intelligent charging and discharging · State of charge (SoC) · Battery life cycle · Depth of discharge (DoD)

1 Introduction

Battery is one of the mostly used energy storage units either in portable or stationary form. With recently predicted increase in electric vehicle and reduction of conventional fuel-based vehicle, there will be huge demand of suitable batteries to fulfill the demand. On the other hand, renewable sources of energy like solar, wind, tidal wave, and others are of variable nature. Therefore, it is very important that the produced energy must be stored in some sort of accessible device that will provide us with power according to our requirement. Since use of such forms of energy has escalated, so is the use of batteries. Additionally, batteries are also used in several instruments like health instruments, mobile phones, laptops, automobiles, wireless electronic items, wireless sensor networks, and many stand-alone applications. Life of a rechargeable battery is limited and needs replacement after certain time period parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables, are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow. There comes the demand to check the life expectancy of a particular type of battery or battery of same capacity from different manufacturers both from the quality, reliability, and economic point of view. Life cycle test is a measure of life expectancy before capacity falls to 80% of the rated capacity. Though a guideline has been formed for such test [2], not much work has been done till date in this area in India; in the year 1993, the method suggested by Mondal and Saha [3] was a stand-alone approach and also a time-consuming process; afterward, model-based approach has also been suggested [4]. Another method of estimation through capacity fading and increase in internal resistance had been reported by Wang et al. for traction batteries [5].

Batteries are a collection of cells, where the chemical reaction creates a flow of electrons in the circuit [6]. In the present work, two types of rechargeable batteries, namely sealed lead–acid battery and lithium-ion battery, have been selected for their wide range of uses and popularity. Sealed lead–acid battery uses sponge lead and lead peroxide to chemical energy into electrical energy [7]. It can provide higher cell voltage at a lower cost thus used in vehicles, power stations, etc. The active elements are lead peroxide, sponge lead, and dilute sulfuric acid which acts as an electrolyte [8, 9]. Pb–acid battery has a specific energy of 40 Wh/kg. On the other hand, for lithium-ion (Li-ion) battery, the key component is lithium ion. During the discharging condition, lithium atoms in the anode are ionized and separated from their electrons [10]. The lithium ions move from anode to cathode through the electrolyte, thus recombining with the electrons from the cathode and getting neutralized. Lithium-ion batteries can provide comparatively higher voltage 3.7 V compared to 2 V in Pb–acid and high charge storage per unit mass and unit volume owing to lithium's smaller size 150 Wh/kg and 250 Wh/l, respectively [11, 12]. The state of charge of a battery is the level of charge of a battery relative to its original capacity usually expressed in percentage [13]. The depth of discharge of a battery is the percentage of battery capacity that has been discharged relative to the overall capacity of the battery. A battery life cycle tester is a device which is used to find the life cycle of a given battery. Here it has been tried out to find the state of charge or depth of discharge of the battery at different operational points and the actual capacity of the battery. It also aids in benchmarking of different manufacturer's battery which are either of the same rated capacity or batteries of different types. As it is not practical in some cases to wait and count all the charge and discharge cycles, the battery completes before the capacity falls to 80% at a rated current of C20 or even C10 (20 or 10 h of discharging time to get 100% rated output). Life estimation of a system is a time-consuming process [2], but if statistical model or machine learning tool is used [14], that time could be greatly reduced; on the other hand in some cases, accelerated test or statistical tools like linear regression technique could be used. In this work, linear regression technique has been used to predict the life cycle. Though lower depth of discharge (DoD) can increase cycle life of battery, but in this work, 80% DoD has been selected for both the types [11, 15]. However, the designed system is independent of battery type, and all the parameters could be set prior to experiment for a particular type of battery. Prediction of life for VRLA batteries through early current/voltage data has been proposed by Sexton et al. [16]. Life cycles test on a lithium battery system had been reported by Vellucci et al. along with the performance degradation [17].

2 Proposed System

Figure 1 illustrates the system block diagram of the battery life cycle tester. Here we have used an AVR controller [18] which reads battery voltage and current either charging or discharging from a voltage divider and a current sensor through a 10-bit inbuilt analog-to-digital converter.

First, the input data are then scaled to get the actual values of voltage and current. Initially, when a battery is connected to the battery life cycle tester, the device asks for specifications of the battery connected (e.g., battery rated voltage, battery capacity, full charge voltage, low battery voltage, and current rating for charge/discharge). The inputs are entered through two SPST push buttons. These data get logged into the microSD memory card [19]. The data are then processed, and the battery is allowed to be discharged to 20% of its full capacity. When the value of battery voltage falls below the specified low voltage, then charging is started and that marks the beginning of the first charging cycle. Thus, the cycle number is incremented by 1. When the battery is fully charged, then discharging starts until the depth of discharge becomes 80%. This entire process of charging and discharging constitutes one cycle of the battery. For both charging and discharging states, the battery voltage, current, and time are measured, logged in a microSD card, and sent to the cloud. Furthermore,

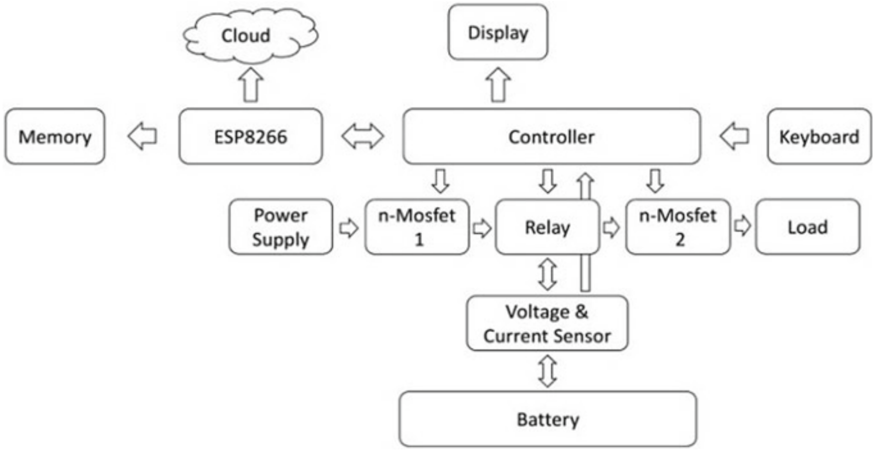


Fig. 1 System block diagram of the proposed tester

using these three parameters, the device calculates ampere-hour, watt-hour, and state of charge/depth of discharge [20], which is also logged in the microSD card and consequently sent to the cloud. This whole process of data logging and sending the data to the cloud is executed by another controller. The communication mode between the two controllers is serial communication. This device also constitutes 16×2 LCD displays which shows the different parameters, i.e., state of charge/depth of discharge, charging/discharging state, number of cycles passed, battery voltage, and current. To switch between the charging and discharging modes, a DPDT relay is used. Furthermore, the relay is controlled through a signal that originates from the AVR controller. From two PWM pins of the AVR controller, two PWM signals are generated which are then directed to the two n-MOSFETs are used as switch. The n-MOSFETs are used for current control. Two opto-isolators are used for the protection of the AVR controller AT Mega328P by optically isolating the control circuit and the power circuit. The power circuit also consists of external power supply, load, and battery terminals. Suitable load and power supply need to be connected according to the battery requirements. During the charging cycle, power streams from the power supply to the battery through the pathway containing n-MOSFET 1, relay, and voltage and current sensor. In the course of the discharging cycle, power flows from the battery to load through the route comprising voltage and current sensor, relay, and n-MOSFET 2.

3 System Algorithm

In this proposed system, there are two controllers connected by UART (i.e., Controller 1 and Controller 2) to control the total task [18, 19, 21]. From the very beginning, these two controllers start together and with proper synchronization accomplish the complete task. The entire system flowchart has been described in Figs. 2 and 3. At first, the battery life cycle tester takes the following inputs to get initialized in similar order: battery rated voltage, low battery voltage, full charge voltage, battery capacity, and rate of charging/discharging current. After the inputs are taken, all these data are then stored in the memory card. The data are also presented in the LCD display so that it may be verified whether the data entered are accurate or not and change accordingly if required. The charging/discharging rate so entered is used to determine the discharging current. The battery is then made to discharge at the determined value of discharge current. While discharging, the voltage of the battery is measured

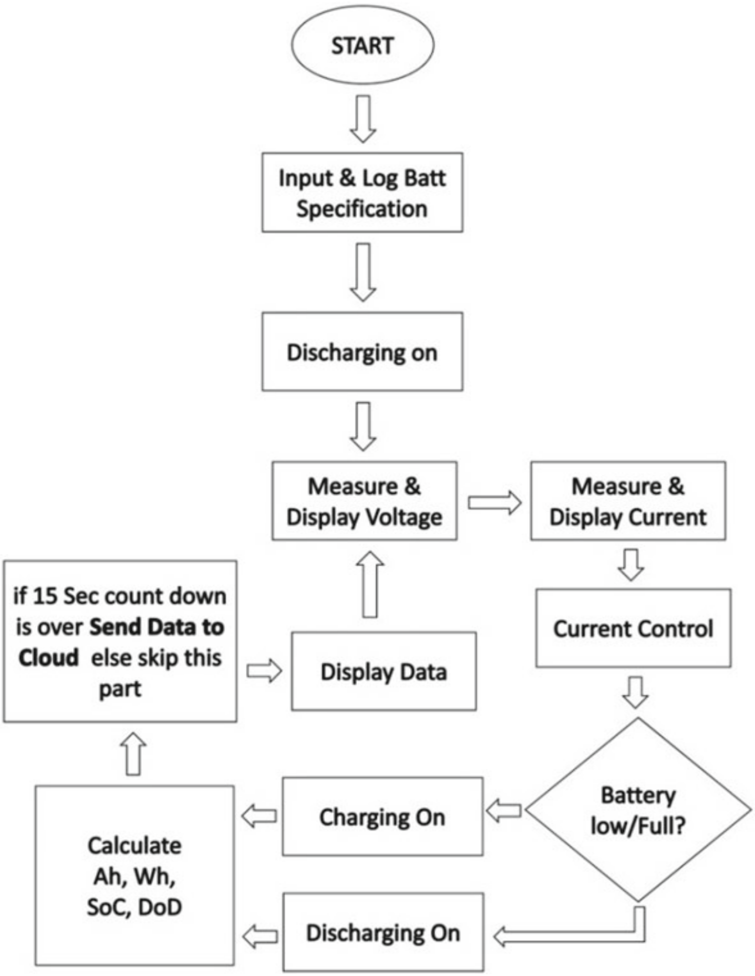
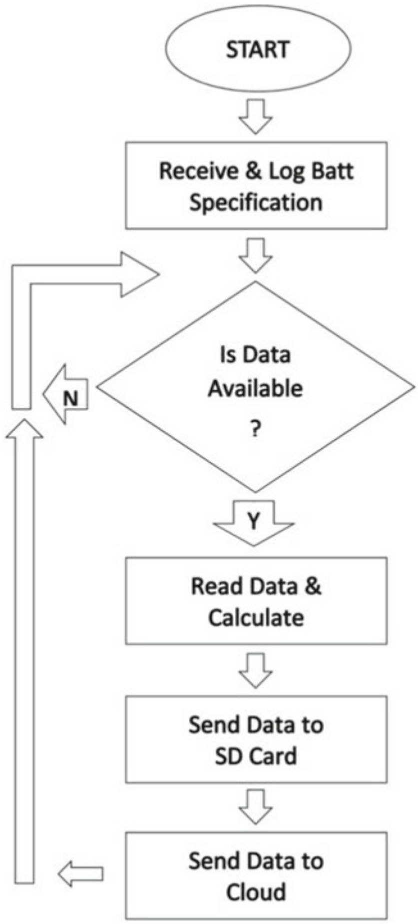


Fig. 2 System software flowchart I

Fig. 3 System software flowchart II



and displayed. Next, the battery discharging current is also measured and displayed. Here hall current sensor ACS 712 has been used to measure charging and discharging currents [22]. The subsequent part is current control. During current control, if the current increases above the predetermined discharging current, then by changing the on time of PWM signal the discharging current is decreased, and similarly if the current decreases below the preset discharging current, then by altering the PWM signal the discharging current is increased. Thus, a predefined set current level is maintained through a closed loop controlled PWM technique. The same method is followed during the charging cycle. The charging current is determined using the entered charging/discharging rate. In the next part, the battery voltage is checked, and if the voltage falls lower than the predefined minimum battery voltage, then the battery stops discharging as n-MOSFETS 2 gets cut off and again starts charging through n-MOSFET 1 being on. Otherwise, if the battery voltage increases above the maximum full charge voltage, then the battery is considered to be fully charged

Table 1 Data format for CSV file in cloud

Created_at	Entry_id	Field1	Field2	Field3	Field4	Field5	Field6	Field7	Field8
Date and time	S. No.	Voltage	Current	Ch/disch	Time	SoC/DoD	Cycle	Ah	Wh

Table 2 Data format for txt file in SD card

Battery voltage: 12.00	Battery capacity: 1.00	Battery full: 14.00 V	Battery low: 10.80 V	Cx: 20.00

Battery voltage	Battery current	Battery chg	Battery disk	SoC	Cycle	Chg Ah	Disch Ah

and the discharging cycle starts by closing the witch n-MOSFET 2 and opening n-MOSFET 1. During the discharging cycle, the discharging ampere-hour, watt-hour, and depth of discharge are calculated, and after a hiatus of every 15 s, these data are sent to the Controller 2. During the next charging cycle, the charging ampere-hour, watt-hour, and state of charge are calculated and stored, and after an interval of every 15 s in a similar manner, the data are sent to the Controller 2 for data logging and IoT-based remote monitoring to a cloud server (ThingSpeak). Next, the state of charge/depth of discharge charging/discharging state, the cycle number is displayed on the LCD screen.

As soon as the Controller 2 gets the battery specifications data from Controller 1 through the method of serial communication, Controller 2 logs the data in the SD card using serial peripheral interface. Controller 2 then checks whether there are any more serial data available, if not, then it checks again and if yes, then it at first, reads the data, calculates several parameters (e.g., ampere-hour, watt-hour, etc.), stores the data in the SD card, and also sends them to the cloud server.

After this whole process is completed, it waits for more data to arrive. In Table 1, format for CSV data in the cloud is given. In Table 2, data format for data log in the SD card is presented.

4 Results and Discussion

The results acquired are illustrated in Table 3 which demonstrates the CSV data from the cloud. This sample data were recorded on March 19, 2020.

Figure 4 demonstrates battery voltage versus time curve. It may be observed that the battery voltage increases with time, and then after a period of time, it becomes stable. Constant current charging occurs during the period when battery voltage rapidly increases. Constant voltage charging occurs during the next phase when the battery voltage becomes stable. During the discharging period, battery discharges at

Table 3 Sample data in cloud

Created_at	Entry_id	Field1	Field2	Field3	Field4	Field5	Field6	Field7	Field8
Date and time	S. No.	Voltage	Current	Ch/disch	Time	SoC/DoD	Cycle	Ah	Wh
2020-03-19 11:20:51 IST	392	11.0303	−0.25367	0	14	0	0	0	0
2020-03-19 11:21:21 IST	393	13.82503	0.03464	1	15	20.03278	1	0.00028	0.00393
2020-03-19 12:02:08 IST	528	15.88151	0.0559	1	56	35.89408	1	0.1201	1.90729
2020-03-19 12:02:26 IST	529	14.62733	−0.30413	0	56	99.9415	1	−0.00048	−0.00702
2020-03-19 12:53:19 IST	705	10.61493	−0.26571	0	107	82.28656	1	−0.20025	−2.12561
2020-03-19 12:53:35 IST	706	13.88141	0.00911	1	107	20.03105	2	0.00027	0.00373
2020-03-19 13:25:18 IST	825	15.65286	0.07257	1	139	35.67326	2	0.12016	1.88079
2020-03-19 13:25:34 IST	826	15.05388	−0.04092	0	139	99.9903	2	−0.00008	−0.00116
2020-03-19 14:06:37 IST	980	10.68106	−0.25686	0	180	82.60177	2	−0.19547	−2.08779
2020-03-19 14:06:53 IST	981	12.96155	−0.08446	1	180	20.0143	3	0.00013	0.00172
2020-03-19 14:39:40 IST	1104	15.5916	0.06701	1	213	35.42363	3	0.11871	1.85084
2020-03-19 14:39:56 IST	1105	14.96851	−0.04445	0	213	99.99279	3	−0.00006	−0.00086
2020-03-19 15:21:30 IST	1260	10.61908	−0.2563	0	255	82.8508	3	−0.19379	−2.0579
2020-03-19 15:21:46 IST	1261	13.60956	−0.01392	1	255	20.02644	4	0.00023	0.00317
2020-03-19 15:58:33 IST	1390	15.53202	0.06655	1	292	35.28617	4	0.1181	1.83434
2020-03-19 15:58:49 IST	1391	14.67525	−0.05548	0	292	99.9916	4	−0.00007	−0.00101
2020-03-19 16:49:12 IST	1560	10.69463	−0.24975	0	343	82.90201	4	−0.19185	−2.05176
2020-03-19 16:49:29 IST	1561	13.02568	−0.03634	1	343	21.9407	5	0.01788	0.23288

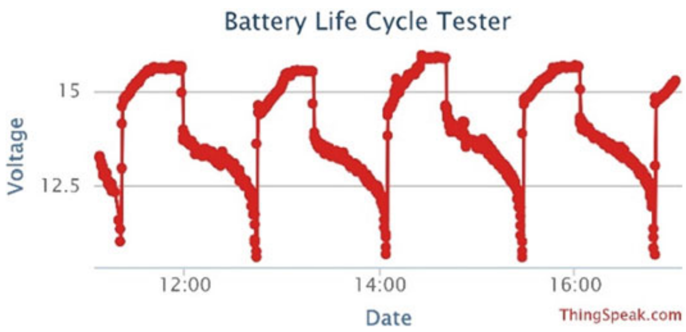


Fig. 4 Variation of voltage in SLA battery with time

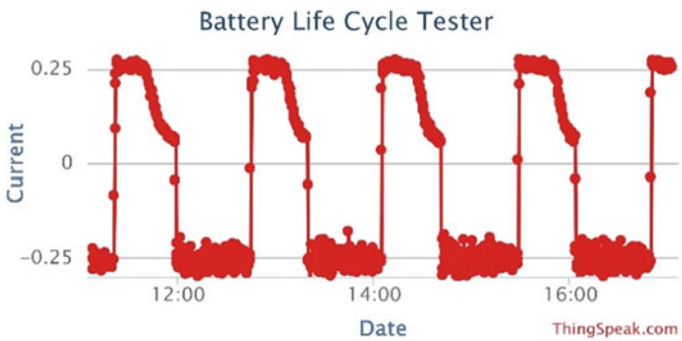


Fig. 5 Variation of current in SLA battery with time

a constant current of 250mAmps as previously defined by us as charging/discharging rate which is C25. Figure 6 exhibits the charging/discharging current discussed.

Figure 5 is the graphical representation of charging/discharging current with time. Here we can comprehend that the positive *Y*-axis depicts the charging current and the negative *Y*-axis depicts the discharging current. Previously, we have mentioned that constant current charging occurs while the voltage increases which can be seen in Fig. 5. In Fig. 5 at the beginning of charging, charging current stays constant and after that decreases rapidly. The state of charge (during charging) and depth of discharge (during discharging) both are demonstrated in Fig. 6. In Fig. 7, the charging and discharging Ah of the battery are represented. In Fig. 8, the Wh during the charging and discharging period is demonstrated (Fig. 9; Table 4).

Linear Regression Model

The data collected so far from the cloud used to predict the entire life cycle of the battery through linear regression model as applicable to a time series data, and it has been observed that the cycle life of the Pb–acid battery comes around 500 cycles and that for lithium-ion battery is around 1200. The data in the form of CSV.xlsx file

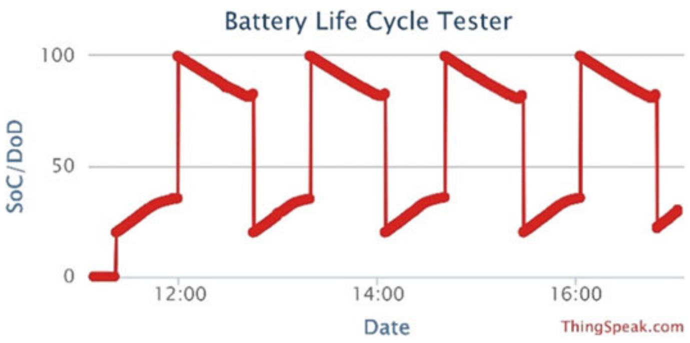


Fig. 6 SoC/DoD of different testing cycle of SLA battery

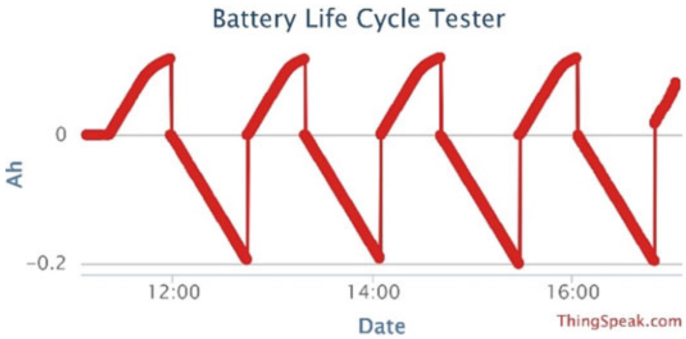


Fig. 7 Ah in different testing cycle of SLA battery

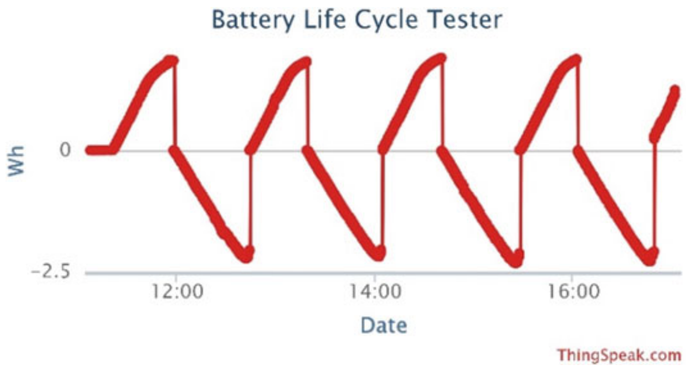


Fig. 8 Wh in different testing cycle of SLA battery

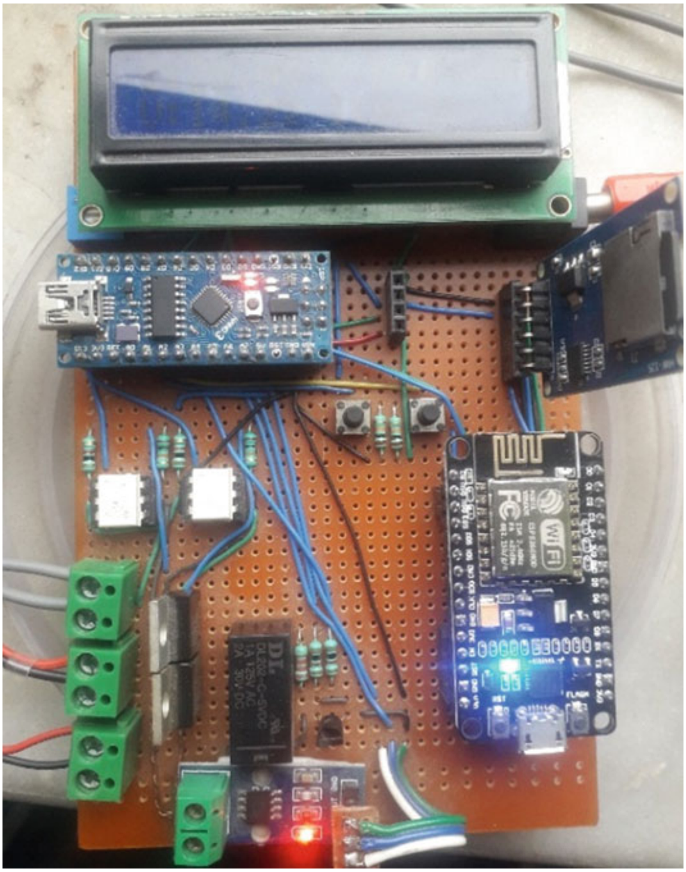


Fig. 9 Hardware setup

Table 4 System components

Parameters	Specifications
LCD (display)	16 × 2 character
Controller	ATMEGA328P
Wi-Fi module	ESP8266
Current sensor	ACS712
MOSFET	IRF540
Cloud	ThinkSpeak

uploaded in Jupiter notebook and a data mining by Python programming are applied. The data then shaped into two different arrays [x, y]. The data size is [361, 361]. We are taking field 6 time cycle and field 8 capacity (Wh) as a X and Y-dimensional array. After splitting the data, one set is used for training the model and then the other is used for testing the battery (Figs. 10 and 11).

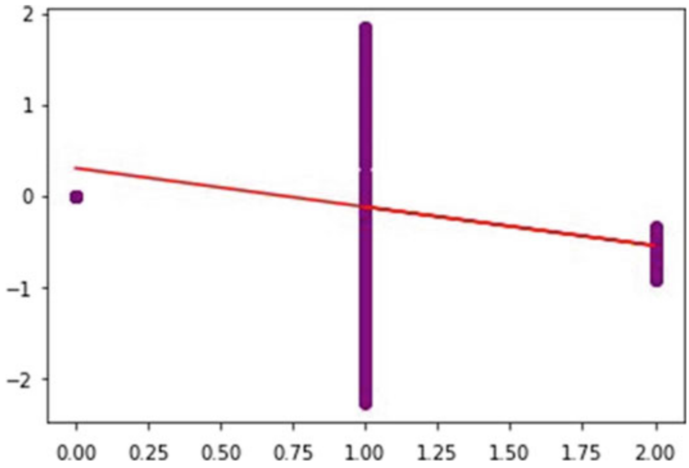


Fig. 10 Training model of linear regression

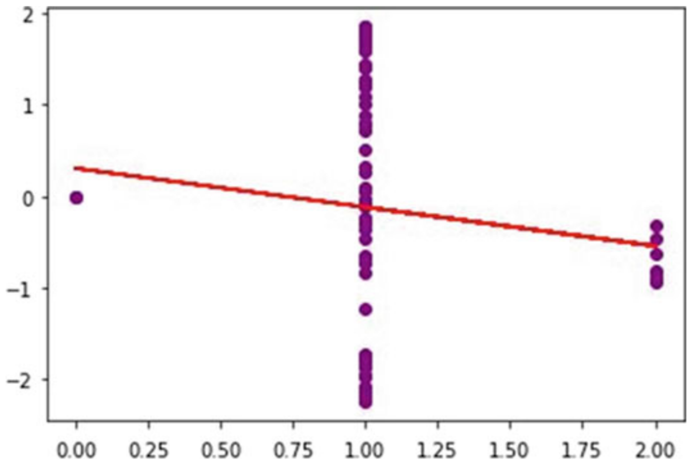


Fig. 11 Testing model of linear regression

5 Conclusion

The system could be integrated into a device to detect life cycle, SoC, DoD, and life expectancy of a battery on board for various applications. If database of battery characteristics of famous manufacturers and their model could be loaded previously in the system, then manual setting of battery parameters could be avoided, and only selection of model will do the rest of the job. Moreover from the existing database of characteristics, it would be much easier to find out and predict the desired parameters.

References

1. Mutagekar S, Kurian JP, Jhunjhunwala A, Kaur P, Gunaki S (2016) Designing a high performance battery life cycle tester. In: 2016 first international conference on sustainable green buildings and communities (SGBC)
2. https://law.resource.org/pub/in/bis/S05/is.1652.1991.pdf [BIS standards for testing batteries]
3. Mondal A, Saha H (1994) Low cost digital battery life cycle tester. IETE Tech Rev 2(5 and 6):347–351
4. Agarwal V, Uthaichan K, Decarlo RA, Toukalas LH. Development and validation of a battery model useful for discharging and charging power control and lifetime estimation. IEEE Trans Energy Convers 25:821–835. https://doi.org/10.1109/TEC.2010.2043106
5. Wang F, Liu S, Xiao C, Meng X (2014) Cycle-life testing & evaluation research of traction battery. In: 2014 IEEE conference and expo transportation electrification Asia-Pacific (ITEC Asia-Pacific), Beijing, pp 1–3
6. McKeon BB, Furukawa J, Fenstermacher S. Advanced lead–acid batteries and the development of grid-scale energy storage systems
7. May GJ, Davidson A, Monahov B (2018) Lead batteries for utility energy storage: a review. J Energy Storage 15:145–157
8. https://docs.exideindustries.com/pdf/industrial-export-Batteries/products/ups-batteries/2v-agm-vrla-catalogue.pdf
9. Schmidt-Rohr K (2018) How batteries store and release energy: explaining basic electrochemistry. J Chem Educ 95(10):1801–1810
10. https://www.capacitorsite.com/lithium.html
11. https://www.exideindustries.com/products/know-your-battery/battery-defination.aspx
12. https://www.researchgate.net/figure/Schematic-of-the-lithium-ion-battery
13. How DNT, Hannan MA, Hossain Lipu MS, Ker PJ. State of charge estimation for lithium-ion batteries using model-based and data-driven methods: a review
14. Huang P, Wang Q, Li K, Ping P, Sun J. The combustion behavior of large scale lithium titanate battery
15. Carullo A, Vallan A, Hirano M, Oka K, Tagawa Y (2012) Outdoor experimental laboratory for long-term estimation of photovoltaic-plant peformance. IEEE Trans Instrum Meas 61(5):1307–1314
16. Sexton ED, Nelson RF, Olson JB (2001) Life prediction for valve regulated lead acid batteries from early current/voltage data. In: Sixteenth annual battery conference on applications and advances, proceedings of the conference, Long Beach, CA, p 193
17. Vellucci F, Sglavo V, Pede G, Pasca E, Malvaldi V, Scalari S (2014) Life cycles test on a lithium battery system. In: IECON 2014—40th annual conference of the IEEE industrial electronics society, Dallas, TX, pp 3129–3134
18. https://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-7810-Automotive-Microcontrollers-ATmega328P_Datasheet.pdf

19. Mahzan NN, Omar AM, Mohammad Noor SZ, Mohd Rodzi MZ. Design of data logger with multiple SD cards. In: 2013 IEEE conference on clean energy and technology (CEAT)
20. Yang F, Song X, Xu F, Tsui K-L. State-of-charge estimation of lithium-ion batteries via long short-term memory network
21. https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_en.pdf
22. https://www.sparkfun.com/datasheets/BreakoutBoards/0712.pdf

Websites

23. https://www.scribd.com/
24. uk.mathworks.com

Design and Integration of Collaborative Features Along with an Efficient GUI to the Industrial Robot

Pottem Samhith Kumar, K. R. Usha Rani, K. S. Geetha, Shanmukha Nagaraj, Prathamesh Zingade, and Akshata Shet

Abstract With the current advancements in technology, industries are adopting and employing the robots to perform complex and repetitive tasks. Performing these operations sometimes requires the human involvement to complete and finish those complex tasks. But, safety is one of the key aspects involved in this. This collaboration of robots with humans paves the way for development of the modern and future industries where both humans and robots work together thus improving the precision and accuracy of the operations. Usually, industrial robots operate with high speeds to finish of the operations rapidly and can perform the maximum operations daily. Since safety is a critical parameter in this human–robot paradigm, the collaborative robots must operate with less force compared to the industrial ones by augmenting human abilities as well as providing much more flexibility. Thus, the integration of such automated systems and operators costs very high compared to an ordinary industrial robot. In this work, the team considered to integrate the collaborative features on a 4 Degrees of Freedom (DoF) industrial robot with a payload of 3 kg and reach of 550 m with a repeatability of ±0.1 mm. In order to provide effective human–machine interface which includes automation and operators, a proximity sensor with range up to 20 cm and a tactile sensor with range up to 4 kPa are interfaced to the robot. Real-time values are taken from sensors and are sent to controller section, and stop function is initiated in order to stop the robot operation.

Keywords Collaborative robots (CoBots) · Four axis motion controller (FAMC) · Graphical user interface (GUI) · Sensors · Robot · Degrees of Freedom (DoF) · Torque

1 Introduction

The collaborative robots or CoBots are a special class of robots that are mainly intended to work together with humans making the work environment flexible and

P. S. Kumar (✉) · K. R. Usha Rani · K. S. Geetha · S. Nagaraj · P. Zingade · A. Shet
R.V. College of Engineering, Bangalore 560059, India
e-mail: pottemsk.ec16@rvce.edu.in; samhith99@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_2

safe [1]. Safety factor in CoBots not only lies in reducing the torque and speed of motor, but also on using the lightweight materials and rounded edges. Additional safety factor relies on the desired sensors and software where the sensors are continuously monitored and the feedback is given to the controller to take appropriate decisions and drive motor accordingly.

Collaborative robots are evolved from traditional industrial robots which consist of drivers and motors for performing complex and repetitive tasks that work together with humans. It differs from humanoids [2] where they mimic the human action not only in functioning of the system but also resemble humans structurally. Collaborative industrial robots can be used to automate repetitive, ergonomic tasks such as fetching and carrying heavier parts, machine feeding and final assembly, and a typical picture of CoBot can be seen in Fig. 1.

As there are many incidents that robot harms the products and the human in the industries, it is very much essential to replace them with the collaborative robots (CoBots) to provide the safety. Thus, incorporating the collaborative features into an industrial robot where the speed of operation could be compromised upon and giving utmost priority to safety opens a new field of research in industrial robots. Normally, such features in robots are preferred in the situation where robots work along with human intervention.

Cost effectiveness, i.e. the CoBot features are added with less cost which gives an upper hand for study-purpose robots when compared to industrial and conventional robots as in [4]. Also, the GUI can easily be customized and that paves the way in the research area for these kind of robots as compared to the conventional robots [5].

Fig. 1 Typical structure of a Pick and Place CoBot [3]

1.1 Literature Review

S. No.	Paper title	Journal/conference	Major observations
1.	"The concept and implementation of a passive trajectory enhancing robot"	W. Book, R. Charles, H. Davis, and M. Gomes, Proc. ASME Dyn. Sys. Cont. Div., vol. DSC-58, p. 633, 2004	Designing the robots for passive trajectory enhancement
2.	"Methods for safe human–robot interaction"	Foundation and trends in robotics 5(4), 261–349, 2017	Human–robot-interaction (HRI) safety through control, motion planning and prediction
3.	"Perspective for indirect human–robot interaction"	Tsui KM, Desai M and Yanco HA. In: Proceedings of the fifth ACM/IEEE international 2010, pp. 129–130	The interactive task involved a humanoid robot's use of pointing gestures
4.	"Automatic categorization of haptic interactions—what are the typical haptic interactions between a human and a robot?"	T. Tajika, T. Miyashita, and H. Ishiguro, in Proceedings of the 2006 IEEE-RAS International Conference on Humanoid Robots, Genova, Italy, December 2006	For categorizing haptic interactions between a human and a robot by using tactile sensors embedded in the soft skin covering the robot's entire body

2 System Development

The overall system view where in collaborative features are added to the industrial robot to make it interactive with operator has been described in Fig. 2.

2.1 CoBot Design

The design of robot is as shown in Fig. 3 that provides 4 Degrees of Freedom, provides 550 mm of reach and can carry 3 kg of payload. The torque is calculated for each axis of robot, and motors are selected with respect to the gear ratio. Axes 1, 3 and 4 have gear ratio 1:50 as compared to gear ratio 1:160 of Axis 2. Axis 2 carries more weight compared to other three axes as it holds weight of both Axes 3, 4. Motors in each axis have its own incremental encoder and power off break, and incremental encoder is used to provide feedback loop to motion controller. Degrees of Freedom provided by each axis is given as follow.

Joint angle ranges:

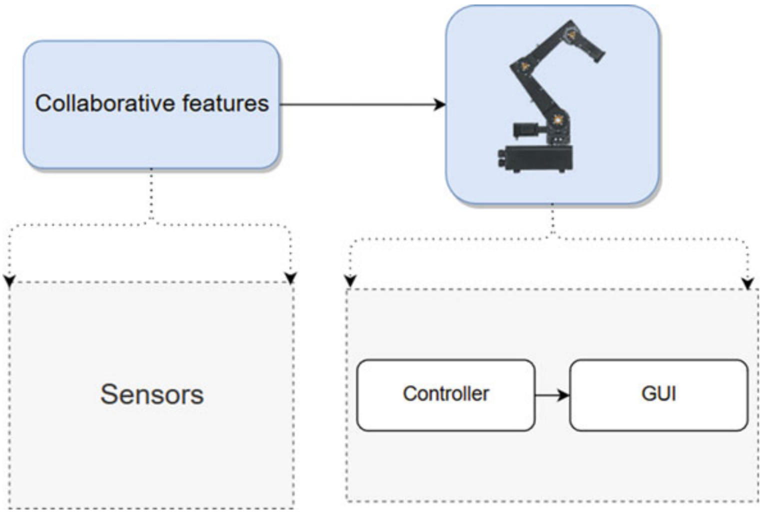


Fig. 2 System design

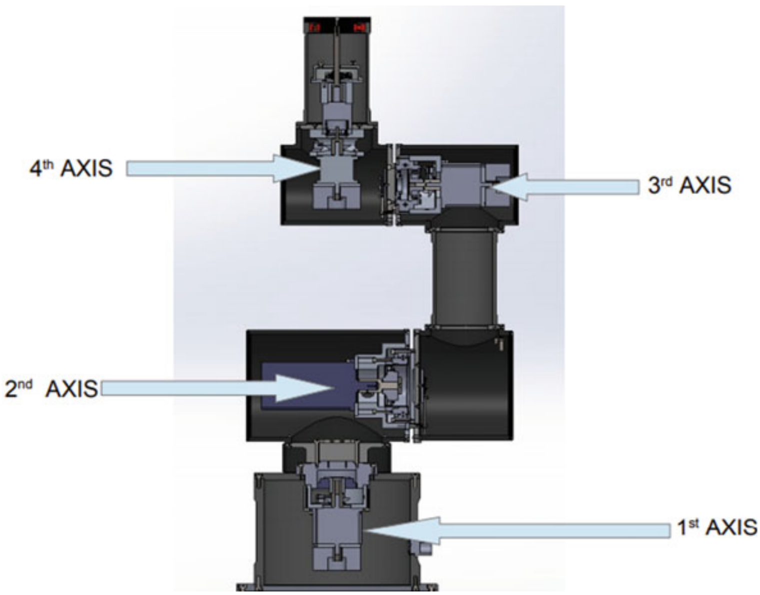


Fig. 3 Arm unit of a CoBot

1. Joint 1: ±170°
2. Joint 2: ±90°
3. Joint 3: ±135°
4. Joint 4: ±360°.

CoBot system is divided into two subgroups

1. CoBot's arm unit
2. CoBot's controller unit.

CoBot's Arm Unit

It consists of four axes as mentioned in image (Fig. 3).

CoBot's Controller Unit

CoBot's controller unit as shown in Fig. 4 is the major part to interface all the motor drivers and the controllers which provide communication. Control unit majorly has eight hardware units which are interfaced to the CoBot. Each has its own specifications.

1. AC line filter
   - Two-phase AC line filter is used to filter the noise in the AC power line
   - System input supply: 230 V AC, 50 Hz
2. Miniature circuit breaker (MCB)
   - MCB is used for main circuit breaking and also as a power on switch
3. AC to DC converter (24 V DC)
   - 230 V AC is converted into 24 and 5 V
   - 24 V is used for drivers and brakes

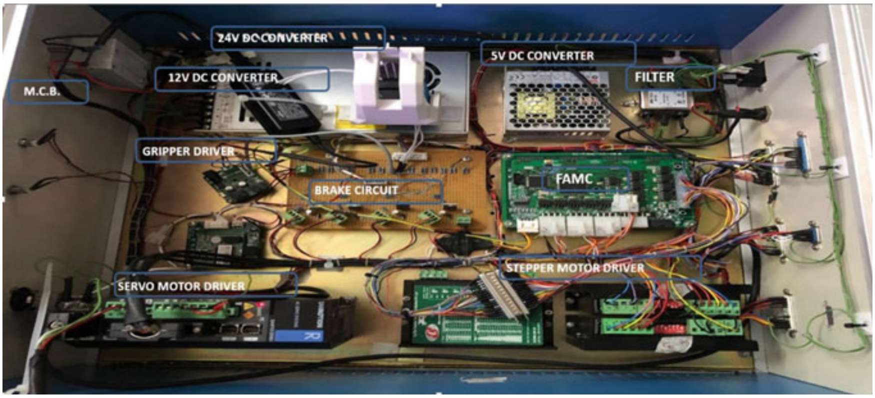


Fig. 4 Control unit of a CoBot

4. AC to DC converter (5 V DC)
    - 5 V is used for FAMC, encoders and sensors
5. Four axis motion controller (FAMC) PCB
    - FAMC card contains microcontroller and motion controller IC
    - From GUI, different commands are given to the microcontroller which commands the motion controller through on-board parallel communication
    - Motion controller will generate the pulse and direction based on the command received by the microcontroller
    - Individual encoder signals are fed to the FAMC card for position read
    - From FAMC, signals are generated to control the brakes while arm operation
6. 24 V brake switching circuitry
    - 24 V electromechanical brakes are used in each axis of CoBot arm which is used in energizing and de-energizing brakes according to the operation
7. Stepper drivers (base, elbow, wrist)
    - There are three two-phase stepper drivers used in the system to control the base, elbow and wrist which operate in 24 V
    - From FAMC card, pulse and direction signals are given to individual stepper drivers, respectively
8. Brushless direct current (BLDC) driver (shoulder)
    - It is used in the system to control the shoulder which operates in 230 V AC
    - From FAMC card, pulse and direction signals are given to BLDC driver.

Figure 5 shows the overall connection between controller and sensors. Main supply is stepped down to 5 V which is required for controller and sensors, and common ground is provided to both controller and sensors.

As the sensor has three pins, the 5 V VCC is given with the supply and GND is connected to common ground. The output pin is given to the input of FAMC. Two pins are still unused and can be used for future applications. The distance covered by proximity sensor is 2–30 cm, and tactile sensor has a sensitivity of 10.56 Ω/kPa at the range of 0–4 kPa. The block diagram, shown in Fig. 5, depicts the major components and flow of input output within the system.

2.2 Methodology Adopted

To design the sensor fusion circuit, existing method is used to check the sensor characteristics. Both the tactile and proximity sensor were interfaced to the Raspberry Pi. Functions related to sensing, sensitivity of the sensors, fixing of threshold, range of sensors were studied. Suitable sensors were selected, and the circuitry was developed.

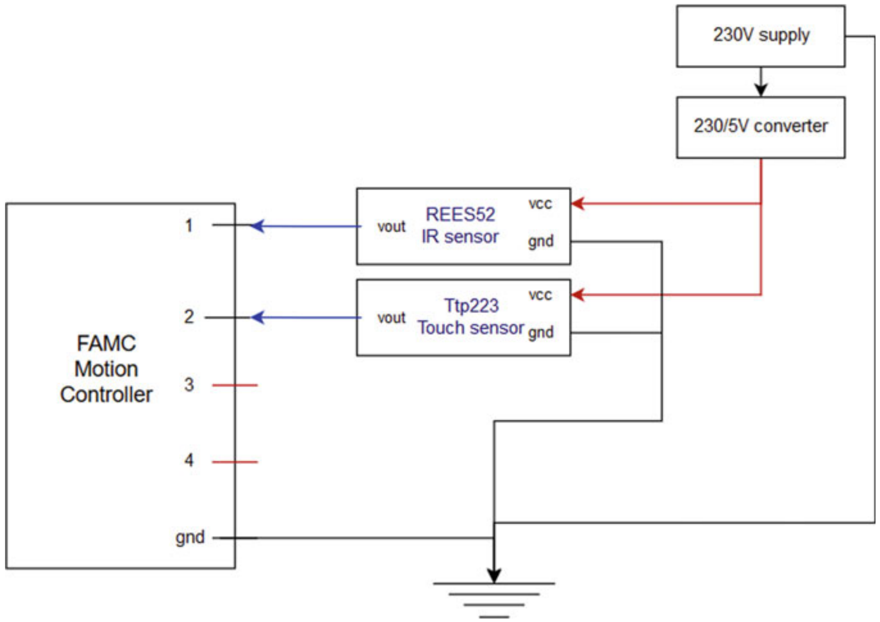


Fig. 5 Block diagram of interfacing sensors to the motion controller

An efficient graphical user interface (GUI) has been designed to connect the GUI to controller and operate in different desired modes. There are different operating modes that has been implemented and integrated to the CoBot. They are:

- Position mode: In this mode, user can enter the values in degree both negative and positive for particular axis. Once the command is received by motion controller card from GUI, axis arm will move in respective target position.

  i. Trapezoidal
  ii. Sinusoidal.

- Jog mode: The continuous movement of an arm in a direction along a selected axis. It is a velocity profile, and user can enter the values in °/s with clockwise or counterclockwise direction for particular axis.

  i. Trapezoidal (CW/CCW)
  ii. Sinusoidal (CW/CCW).

- Linear interpolation: Linear interpolation moves tool point from start point to the target point along a straight line and can be implemented both in 2D plane and 3D space.
- Ten-point teach mode: The user has to move the axes by using position profile and allowed the GUI to record that position. Once all the ten sequences are recorded, the robotic arm will move continuously in a loop.

- Maintenance mode: Maintenance mode is used to make each axis to reach its home position. Home position is nothing, but all axes are in zeroth position.
- Four programmable general purpose input/output (GPIO) 5 V logic level.

The circuit developed as mentioned before was interfaced to motion controller's available GPIOs; in GUI, ON/OFF icon was created to start the sensing operation. A thread was created in GUI to sense the input and take the action. A typical GUI screen that has been designed which has various control features is shown in Fig. 6.

Also, a flowchart describing the sequence of operations in brief is shown in Fig. 7.

- When supply turns ON, the ON/OFF future in the GUI is checked continuously. This feature is added so that controller can start sensing the inputs only when this is activated.
- If this feature is not activated, then robot will be performing its normal functions. But when it is activated, then it activates collaborative features.
- Once it is ON, the timer will be activated so that sensing is restricted to the certain period.
- Then, controller will start monitoring Pins 1 and 2 to which the sensors are connected.
- Once the sensor detects the touch or the object passing by it, then it will generate HIGH/LOW signals.
- If the signals are high, that indicates there is some detection at the input. Hence, the user should call the stop function so that robot stops its function.
- If nothing is detected, i.e. if the signals are low, it must wait and sense the input to a certain period; hence, the timer will be decremented and checked if it reaches to zero.

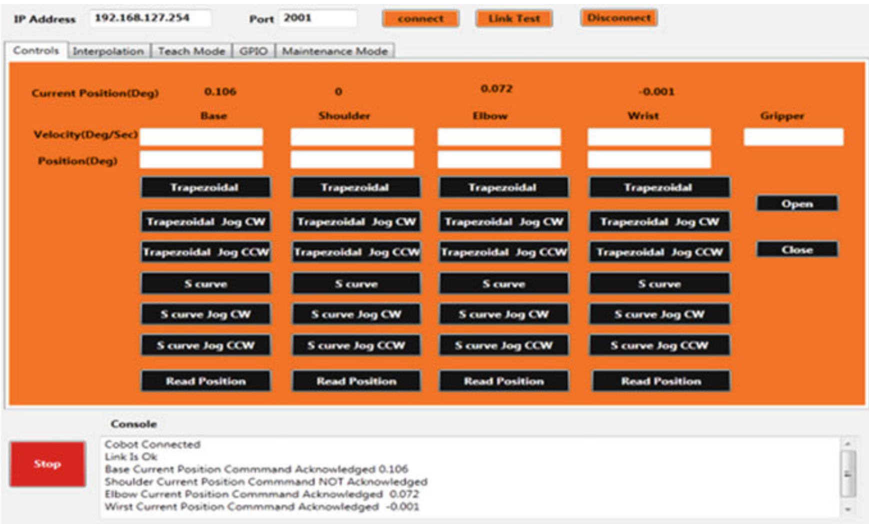


Fig. 6 GUI designed having various control features

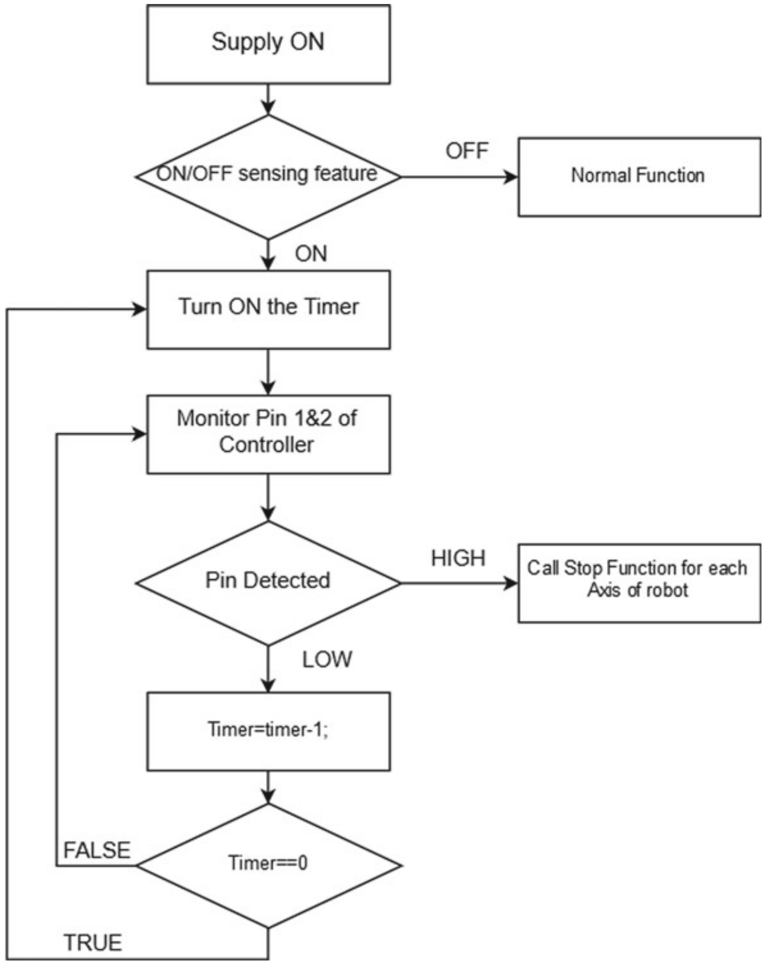


Fig. 7 Process flowchart

- If it is zero, that indicates sensing period is over and deactivates the option so that timer can be restarted.
- If timer is not yet zero, then it will continuously monitor the pins until timer is zero.

2.3 Experimental Details

Heart of the CoBot design is control system unit which consists of AC to DC converter, FAMC, MCB, gripper driver, brake circuit, stepper driver and servo motor driver. The controlling of motor rotation, griper movement, feedback control, etc. is

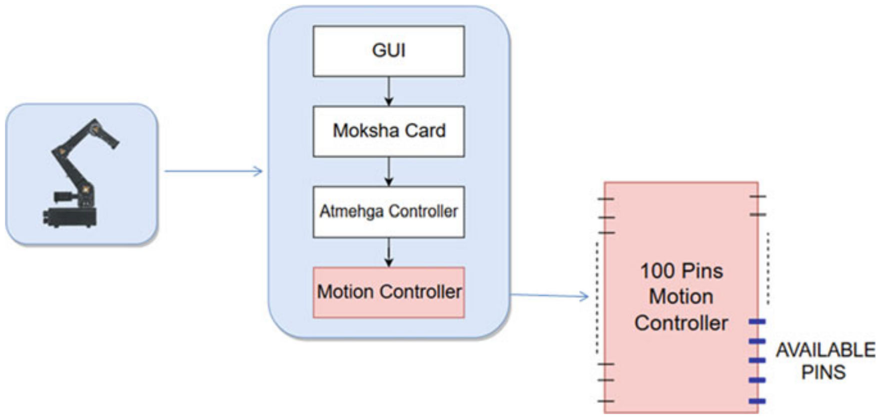


Fig. 8 Collaborative robot components

performed together through the control unit. The output from the sensors is continuously monitored by the GPIO pins, and thread is created when it detects. This flow of operation is explained and described in Fig. 7.

As explained above, the main component is FAMC to which sensors need to be interfaced. Robot that has been designed will have four major components which provide this application as shown in Fig. 8. Using this, the user must check the available pins in FAMC. As there are five pins available one for ground and four for GPIO, it can be used for sensor interfacing.

3 Results and Discussions

If the object is found near robot with the range up to 30 cm, the light emitting diode (LED) glows indicating that robot should stop its functioning. This output can be seen in Fig. 10.

Similarly, when sensor detects touch-on-touch sensor with range and sensitivity up to 4 kPa, 10.56 Ω/kPa, respectively, the immediate decision would be taken by Raspberry Pi to stop the functioning of CoBot by giving the signals to the FAMC.

The desired sensors to make the robot collaborative are placed on various positions of the robot especially on the arms, end manipulator or gripper and on all other places where there might be chances of human intervention.

The hardware results are shown in Figs. 9 and 10, and GUI response to the hardware is recorded and shown in Fig. 11. It is observed that GPIO status window will show status of the sensor parallel to the hardware.

Figure 12 shows the status of sensors connected to the controller. This is the outcome of GUI interfaced to the FAMC of robot that indicates different status of the sensors. Among four pins, only two are used, and they are configured as input.

Fig. 9 Without object sensing (normal function)

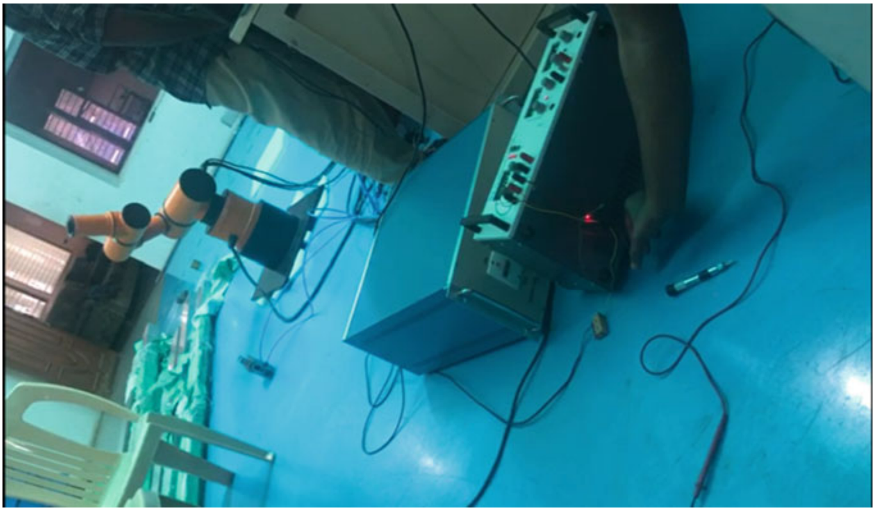


Fig. 10 With object sensing (stop function)

4 Conclusions

Human–machine interaction would make the system more complicated in design; however, it offers much flexibility for the automating the applications. The industrial robot incorporated with collaborative features, such that the operation is stalled when there is human intervention, is the first step towards the CoBot. In this work, the

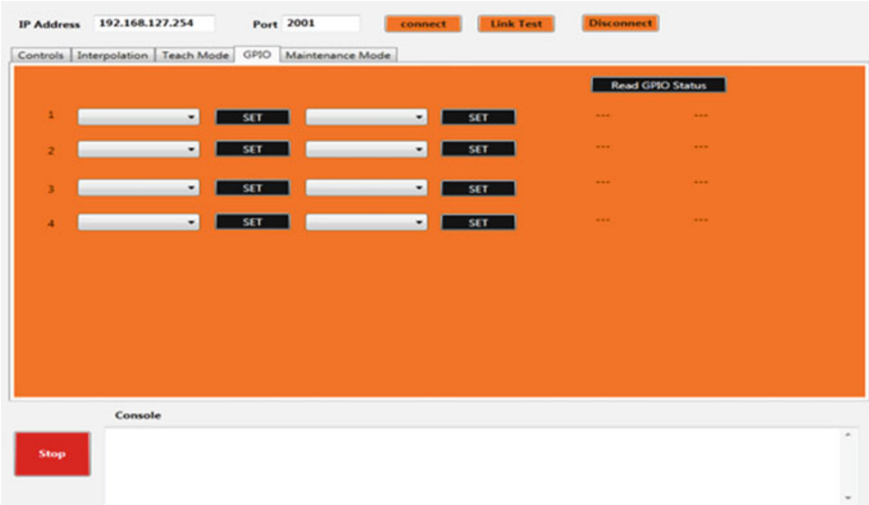


Fig. 11 GUI status screen

Fig. 12 GPIO status of different sensors

team proved that the ranging and characteristics of infrared (IR) and touch sensor by connecting it to the Raspberry Pi have successfully incorporated collaborative features. Touch sensor detects the touch and provides high output, and IR sensor is found to be detecting the objects passing by it in the range of 20 cm. It is concluded that these sensors can be used to interface with the robot.

The tested sensors are interfaced to the motion controller. When the touch sensor senses the touch, it generates high output, and when IR sensor senses the object, it generates low output to the controller which will invoke the thread. This thread has the function related to stop each axis. Then, robot will stop its functioning once it detects the object; hence, the objectives of the project are satisfied by extracting collaborative features of the robot. Thus, the designed CoBot can be used as a "Pick and Place" robot for industrial applications as well as a service robot for research-oriented applications.

4.1 Future Scope

- Light curtain sensor can be implemented which is mounted surrounding the CoBot to detect the motion.
- Sensor readings and conditions can be directly monitored through IoT application using suitable platforms like AWS, ThinkSpeak, etc.
- As there are two more pins available in the controller, robot-related sensors like force sensors can also be interfaced.

Acknowledgements The authors thank TE Connectivity for funding the project. They also extend gratitude to the company for their continuous guidance and support, because of which the project is able to finish in a stipulated time.

References

1. Gillespie RB, Colgate JE, Peshkin MA (2001) A general framework for cobot control. IEEE Trans Robot Automat 17:391–401
2. Haddadin S (2014) Towards safe robots: approaching Asimovs 1st law. Springer, Berlin, Heidelberg
3. Guizzo E (2015) Universal robots UR3 arm is small and nimble, helps to build copies of itself. IEEE Spectrum, 17 Mar 2015. https://spectrum.ieee.org/universal-robots-ur3-robotic-arm
4. Phan T, Chao PC, Cai J, Wang Y, Wang S, Wong K (2018) A novel 6-DOF force/torque sensor for COBOTs and its calibration method. In: IEEE international conference on applied system invention (ICASI), Chiba, pp 1228–1231. https://doi.org/10.1109/ICASI.2018.8394511
5. Nakamura T, Yuguchi A, Maël A, Garcia Ricardez GA, Takamatsu J, Ogasawara T (2019) Ontology generation using GUI and simulation for service robots to operate home appliances. In: Third IEEE international conference on robotic computing (IRC), Naples, pp 315–320. https://doi.org/10.1109/IRC.2019.00058
6. Vogel C, Walter C, Elk Mann N (2013) A projection-based sensor system for safe physical human–robot collaboration. In: Proceedings of IEEE/RSJ international conference on intelligent robots and systems, Tokyo, 3–7 Nov 2013. https://doi.org/10.1109/IROS.2013.6697132
7. Haddadin S, Albu-Schffer A, Fromm Berger M et al (2009) The 'dlr crash report': towards a standard crash-testing protocol for robot safety—part I: results. In: Proceedings of IEEE international conference on robotics and automation (ICRA), Kobe, 12–17 May 2009, pp 272–279. https://doi.org/10.1109/ROBOT.2009.5152602

8. Rybski P, Anderson-Sprecher P, Huber D et al (2012) Sensor fusion for human safety in industrial workcells. In: IEEE international conference on intelligent robots and systems, Vilamoura, 7–12 Oct 2012. https://doi.org/10.1109/IROS.2012.6386034
9. Jaward M, Mihaylova L, Canagarajah N et al (2006) Multiple object tracking using particle filters. In: Proceedings of the IEEE aerospace conference, Big Sky, MT, 4–11 Mar 2006, pp 8–16. https://doi.org/10.1109/AERO.2006.1655926
10. Tsui KM, Desai M, Yanco HA (2010) Considering the bystanders perspective for indirect human-robot interaction. In: Proceedings of the fifth ACM/IEEE international conference on human robot interaction, pp 129–130. https://doi.org/10.1109/HRI.2010.5453230
11. Lumelsky VJ, Cheung E (1993) Real-time collision avoidance in teleoperated whole-sensitive robot arm manipulators. IEEE Trans Syst Man Cybern 23(1):194–203. https://doi.org/10.1109/21.214777
12. Bajcsy R, Canny J, Fukuda T, Giralt G, Harrigan R, Goldberg K, Moraff H (1996) The grand challenges for robotics and automation. In: Proceedings of IEEE international conference on robotics and automation, Minneapolis, MN, Apr 1996. https://doi.org/10.1126/scirobotics.aar7650

Developed Optimized Routing Based on Modified LEACH and Cuttlefish Optimization Approach for Energy-Efficient Wireless Sensor Networks

Pallavi Joshi, Shashank Gavel, and A. S. Raghuvanshi

Abstract The main concern for efficient data transmission in a wireless network is energy dissipation. With the enhancement in the size of the network, there is more demand to aggregate the information which leads to depletion of energy in the nodes. To solve such issues, an efficient data transmission model has to be designed. This paper implements the modified LEACH (MOD-LEACH) concept for node clustering. Along with the efficient cluster head selection, the proposed model is backed with an optimized approach called cuttlefish optimization (CO) which uses two objective functions to optimize the parameters of the network. The proposed novel work is compared with two existing approaches, and the experimental and theoretical results reveal that the proposed model surmounts the two preexisting similar methods in terms of average energy consumption, number of alive nodes, and total packets delivered to the base station. The proposed approach records approx. 75% reduction in the energy consumption of the network.

Keywords Wireless sensor network · Data aggregation · Modified LEACH · Cuttlefish algorithm · Clustering

1 Introduction

A wireless sensor network (WSN) is a collection of several autonomous nodes which can monitor, sense, and process the parameters of a region where it is deployed. WSN is a technology specially designed to serve many applications like military surveillance, monitoring of habitat, smart cities, and many more. Along with its

P. Joshi (✉) · S. Gavel · A. S. Raghuvanshi
Department of Electronics and Communications Engineering, NIT Raipur, Raipur, India
e-mail: pjoshi.phd2017.etc@nitrr.ac.in

S. Gavel
e-mail: sgavel.phd2016.etc@nitrr.ac.in

A. S. Raghuvanshi
e-mail: asraghuvanshi.etc@nitrr.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_3

advantages, there are some challenges and problems faced by these networks. These challenges and problems are listed in Table 1.

The routing [1] also plays a major role in transmitting the data from one node to another. It provides a short and energy-efficient path to deliver the packets without being lost. Many research works manifest the use of optimization approaches while

Table 1 Description of some of the problems encountered by wireless sensor networks

Problems	Description
Environment	WSNs are deployed in areas where human access is very difficult, so they must have the ability to serve in harsh surroundings [2]
Anomalies and faults	Some nodes in a sensor network may work like anomalies and create faults in the network. To avoid this, fault-tolerant approaches and anomaly detection techniques must be employed [3, 4]
Scalability	The quantity of sensor nodes may increase as per the need of the application, so the network should be able to efficiently handle a large number of nodes too
Coverage and node connectivity	The coverage of the network should be such that it should cover almost full area of interest and due to its self-organizing nature, it should connect all nodes to enhance the communication abilities even in harsh surroundings like landslides or volcanoes, etc.
Clustering	Clustering is the way to organize a huge network into some decentralized units called clusters. Each cluster possesses its leader which is accountable to collect the data from the cluster members. An effective clustering of nodes can prevent network latency, enhance the life of the network, and make it more energy efficient [5, 6]
Energy	Since the nodes are powered by batteries, they have limited energy storage with them. The use of the channel bandwidth should be done efficiently to transmit the information from sensors to the base station at low energy expenditure [7]
Data redundancy	Repeated and redundant data are sometimes present in the network and cause energy drain problems in the nodes. It should be avoided by implementing mathematical techniques like MIN, SUM, COUNT, MAX, etc., and some data aggregation approach [8, 9]
Quality of service	QoS parameters are the measures that determine the network reliability and its efficiency. QoS parameters like throughput, delay, packet delivery ratio, and network lifetime need to be taken care to get desired outcomes from the network
Security	The information contained by the sensor network can be leaked or some intrusion may occur due to a lack of security measures. Some cryptographic approaches and intrusion detection techniques can prevent the network from malicious attacks [10]

finding the routes in the network [11]. There is one crucial challenge in the path to routing which is energy depletion. The nodes have very limited power, and they need to deliver their information using the available energy. The management of energy should be done in a way to achieve more transmission in very less utilization of energy. The applications like target tracking, event monitoring, surveillance, etc., need continuous monitoring and observations so there is no chance to replace the batteries or the nodes until the task is done. There are some techniques that can tackle this issue. One of them is data aggregation which enables the network to divide into some clusters or the nodes to send their data to an intermediate node in their vicinity, thereby reducing the amount of energy consumed in the network. Also, the nodes carry redundant and repeated data which is also a cause of energy drain. Data aggregation can solve the redundancy problem by using mathematical or graphical techniques, thereby enabling the nodes to send highly correlated data [12].

Data aggregation aims to retrieve only specific non-redundant information and forward it to the destination [13]. This results in the considerable minimization of the packets delivered and energy usage, and also the bandwidth can be maximized with the help of the data aggregation technique. Figure 1 gives a picturization of an aggregation task done using aggregator and monitoring nodes along with the sensor nodes. These aggregator nodes collect the information, aggregate it, and forward it to the forwarding node where it is forwarded to the base station. The aggregation task is challenging, though it can be solved by efficient clustering. But another major problem is to get desired QoS parameters which can be done by applying various

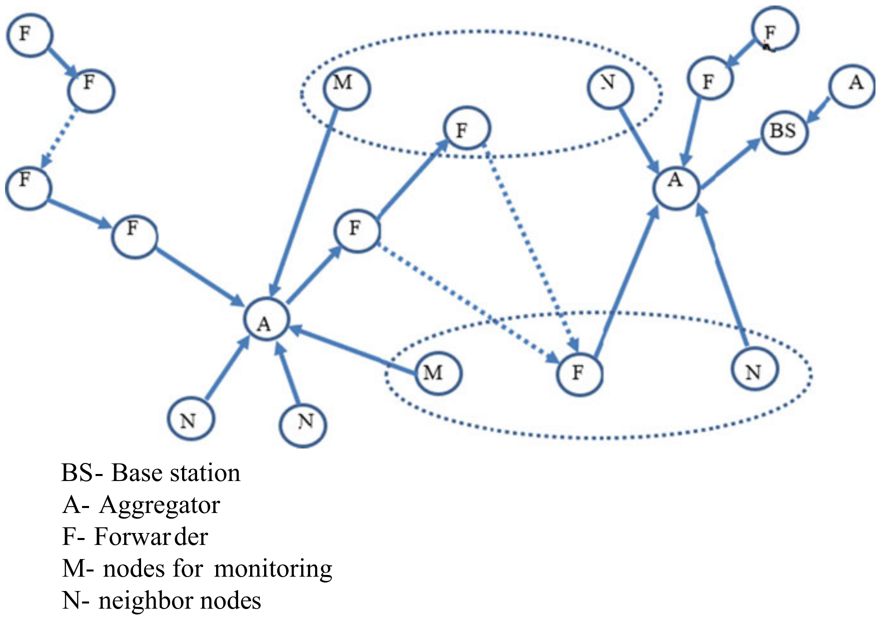


Fig. 1 Aggregation of data using aggregators in a WSN

optimization approaches. This serves as the main motivation behind this paper. The paper combines the notion of both clustering and optimization to aggregate the data efficiently in a sensor network.

The contributions made in this paper to avoid high energy usage are listed below:

(1) The modified version of LEACH performs clustering using the energy and distance parameters that account for the efficient cluster formation.
(2) The data collection technique reduced the packet loss and also prevented redundant data to be transmitted to the BS.
(3) The number of hops is minimized using optimization in combination with the clustering approach by creating efficient routes from source to sink and placing the aggregators in the neighborhood of the source nodes.
(4) Maximizing the coverage and node connectivity by minimizing the number of nodes in a cluster.
(5) The optimization also provides a way to minimize the residual energy of the nodes, thereby maximizing the network life.

The paper is structured in various sections which are described as follows: Sect. 2 gives a compilation of various works relevant to the one discussed in this paper. Section 3 describes the proposed work and explains the methodology using a block diagram. Section 4 investigates the outcome from the proposed work and shows the comparative analysis with other preexisting works. In Sect. 5, the whole paper is concluded and provides a future direction.

2 Related Works

Ezhilarasi et al. [14] investigate various routing approaches with their advantages and disadvantages and their way to solve different issues such as energy conservation and a lifetime of the network. Table 2 shows the classification of various routing protocols along with their examples.

Table 2 Classification of various routing protocols

Routing protocol classification	Routing protocols
Location-based	SPAN, GAF, GEAR, MECN
Data-centric	Gradient-based routing, COUGAR
Hierarchical network protocol	APTEEN, TEEN, HEED, DEEC, LEACH
Node mobility-based	SEAD, data MULES, dynamic proxy tree-based data dissemination
Multipath routing approach	*N*-to-1 multipath discovery, braided multipath
Heterogeneity-based	CADR, CHR
QoS-based	SAR, energy-aware routing

Liu et al. [15] suggest two objectives to be achieved in a WSN that is minimizing the energy dissipation by maximizing accuracy in sensing by using an evolutionary algorithm called fast differential evolution algorithm for optimizing these objectives. He et al. [16] solve the multi-objective problem in optimization by implementing a stochastic algorithm to enhance the network lifetime. The task is achieved by employing a quasigradient algorithm and a Lagrange dual decomposition technique to decompose the multi-objective dual primal problem. In [17], the authors attempt to prevent the circle routes by implementing ladder diffusion and cat swarm optimization algorithms during the routing phase in WSN. An optimized LEACH using PSO and ant colony optimization has been implemented in [18], and the result is compared with the LEACH protocol. The authors in [19] investigate the routing in a clustered sensor network using particle swarm optimization (PSO) to acquire energy efficiency in the network. Tate et al. [20] employ protocol tuning based on multi-objective optimization to gain efficiency in energy consumption. The Pareto evolutionary approach is employed to design and model the system to extirpate the useless factors and to get near-optimal regions.

Above all methods investigate the use of only optimization or only the routing protocols in the wireless sensor networks; very few of the methods amalgamate the optimization with the routing techniques. But somewhere these methods are not much successful to reduce the trade-off problems. The proposed model in this paper implements the cuttlefish optimization which is a very efficient metaheuristic approach, along with the modified LEACH protocol to enhance the network life by minimizing the average residual energy of nodes.

3 Methodology of the Proposed Approach

The MOD-LEACH protocol [21] is employed to obtain efficient clustering of the randomly deployed sensor nodes. The probability function of the LEACH is modified concerning the energy and distance to enable efficient cluster head selection. Next, to minimize the average residual energy of nodes in clusters and maximize the coverage area of the sensor node to get good connectivity, a two-objective cuttlefish algorithm in conjunction with the MOD-LEACH protocol is used.

The Cuttlefish algorithm was first implemented and proposed by Adnan Mohsin and Adel Sabry Eesa in the year 2013. It works like a population-based metaheuristic algorithm where a certain population follows its leader to forage food, prey, etc. The best fitness values given by the optimization algorithm are obtained to optimize the network parameters like energy consumption, throughput, delay, etc.

Figure 2 shows the block diagram of the proposed model. Initially, the nodes in a wireless network are randomly deployed in an area of interest. The LEACH protocol is implemented with highly efficient cluster head selection using distance and energy parameters. The residual energy and the number of sensor nodes are the constraints to the objective function used in the cuttlefish optimization. The result of this fitness function is to minimize the residual energy and maximize the coverage and

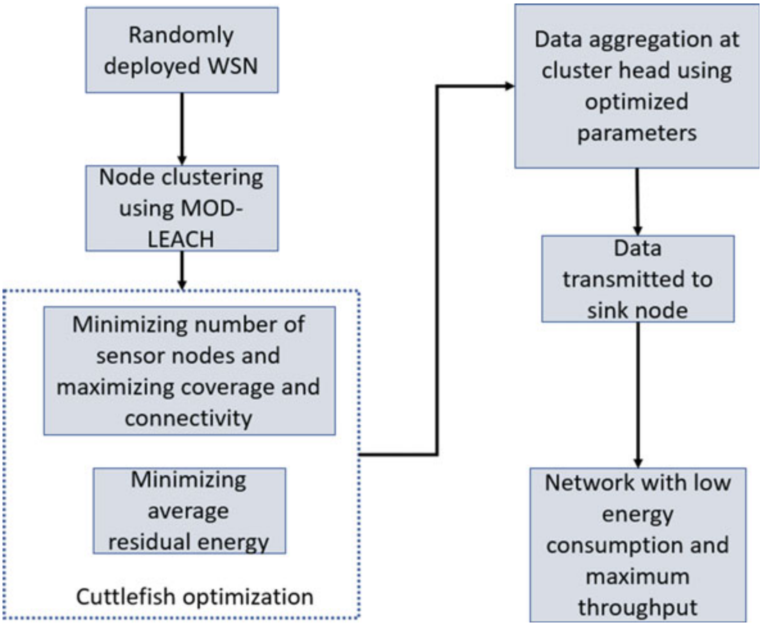


Fig. 2 Flow schema of the proposed MOD-LEACH-CO protocol

connectivity in the network by lowering the number of sensor nodes to be clustered. Next, the data is aggregated at the cluster heads of the respective clusters, and from there, it is transmitted to the sink node.

Steps to perform cuttlefish optimization (CO)

Step 1: Initialize the parameters used in CO.
Step 2: Define the fitness function $f(c)$, $c = (c_1, c_2, \ldots, c_n)^k$.
Step 3: Calculate the best position and current fitness values.
Step 4: Determine the fitness value and local best position.
Step 5: Generate more fitness and local best solutions.
Step 6: Compute the average of all best solutions.
Step 7: Update fitness value and local best positions.
Step 8: Obtain the final best position and best fitness value for the objective function.

The above steps are followed to perform the CO algorithm. Initially, the population is randomly created [22]. Then, the fitness function is defined and the best positions are obtained using that function. The average of the best local values gives the final best-optimized solution. Table 3 gives the description of two objective functions [23] considered in the proposed model. The cuttlefish optimization approach is used to model these two objective functions, and the network is clustered using modified LEACH protocol.

Table 3 Implemented objective functions

Fitness function	Parameter optimized	Technique used
Min $A = \sum_{m=1}^{p\times q}\left(1-\mathrm{e}^{(\mathrm{sd}-\mathrm{cd})}\right)$	'sd' represents the range of sensing for each node, 'cd' indicates the communication range of a sensor. 'A' specifies the minimum quantity of nodes with maximum connectivity and coverage	The cuttlefish optimization approach is modified to achieve energy efficiency by keeping the sensing range constant and varying only the communication range of the cluster member nodes
Average residual energy = $\frac{\sum_{i=1}^{S} E(s_i)}{S}$	S = Number of nodes $E(s_i)$ = energy assigned to each node	The LEACH protocol's modified version is used in which the fitness function for minimizing the average residual energy is used and the clustering is done by prioritizing the distance and energy parameters

4 Results and Comparative Analysis

This paper uses MATLAB R2016a software for simulating both modified LEACH and Cuttlefish algorithms. Table 4 shows a comparison between the fitness values of cat swarm optimization (CSO), particle swarm optimization (PSO), and cuttlefish optimization (CO) with a single objective and two objectives considered in the proposed work, i.e., minimizing the number of nodes with maximum coverage and minimizing the average residual energy in the network. It can be noted that CO with two objectives has a very low variance between the best fitness values and hence the standard deviation for CO is zero as compared to other algorithms. It indicates the stability of the algorithm. The time taken by the CO algorithm is also very less in contrast to other optimization approaches and noted to be 12.7 s when both the 'sd' and 'cd' are kept constant. The network implemented in this paper consists of 200 nodes, and the parameters for simulation are given in Table 5. The modified LEACH with cuttlefish optimization algorithm (MOD-LEACH-CO) is compared with energy-aware routing protocol using cat swarm optimization (EARP-CSO) [17] and LEACH-PSO [18].

Table 4 Comparison of the fitness values for various optimization algorithms

Optimization algorithm	Best fitness value	Average fitness value	Time in s
CSO	39.0327	36.338	25
Single objective CO with varying parameters	34.5841	33.145	23.1
PSO	37.4362	37.4012	16.3
CO with two objectives in proposed model	35.1849	35.1849	12.7

Table 5 Parameters used while simulating the proposed algorithm

Name	Value
Size of network	200 m × 200 m
Quantity of sensors	200
Initial energy of nodes (E_0)	0.5 J
Energy consumption in circuit (E_{elec})	50 nJ/bit
Free space channel specification ($\varepsilon_{\mathrm{fs}}$)	10 pJ/bit/m^2
Multipath channel limit ($\varepsilon_{\mathrm{mp}}$)	0.0013 pJ/bit/m^4
Length of a data packet (l)	1000 bits
Threshold for distance (d_0)	$\sqrt{\frac{\varepsilon_{\mathrm{fs}}}{\varepsilon_{\mathrm{mp}}}}$

Figure 3 conveys the average values of energy consumed by every node for every data transmission. Many rounds of simulations are performed and averaged to obtain the values of average energy consumption in Joule. The proposed work, i.e., MOD-LEACH-CO, is compared with EARP-CSO and LEACH-PSO. It can be inspected that the MOD-LEACH-CO offers the lowest amount of energy consumption in contrast to the other two algorithms because of the objective functions used in CO which optimize the network parameters. Also, in MOD-LEACH-CO, the parameters considered here focus on constructing efficient cluster and their leaders. Figure 4 depicts the alive number of nodes decreasing with the increase in the number of rounds. After 50 rounds, all three protocols show a significant decrease in the alive nodes. For the proposed method, after the 150th round, the number of nodes decreases to 185 approx., whereas the EARP-CSO shows a much steep graph where the alive nodes are only limited to 10 in number after 300 rounds. The LEACH-PSO performs better than EARP-CSO and shows the reduction of nodes to 165 after 100 rounds.

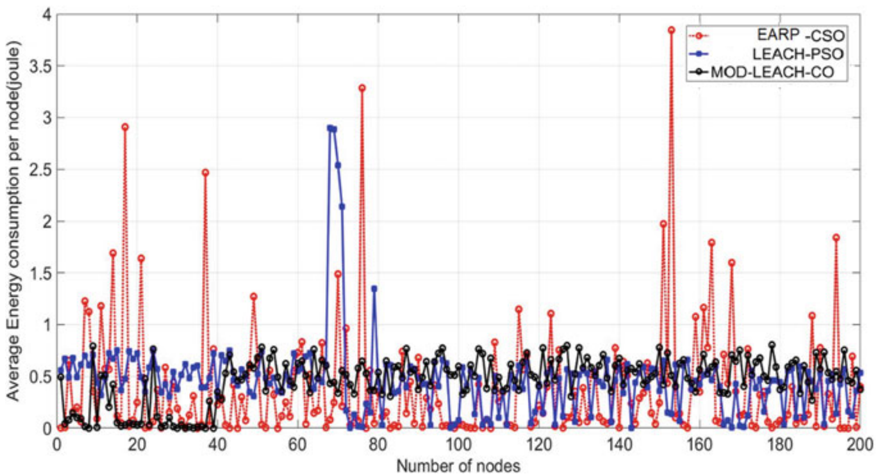


Fig. 3 Average energy consumption by the sensor nodes per transmission

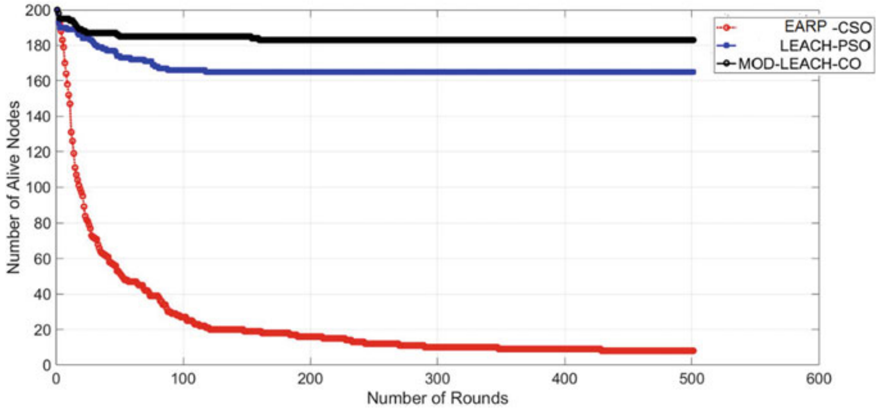


Fig. 4 Number of alive nodes versus number of rounds

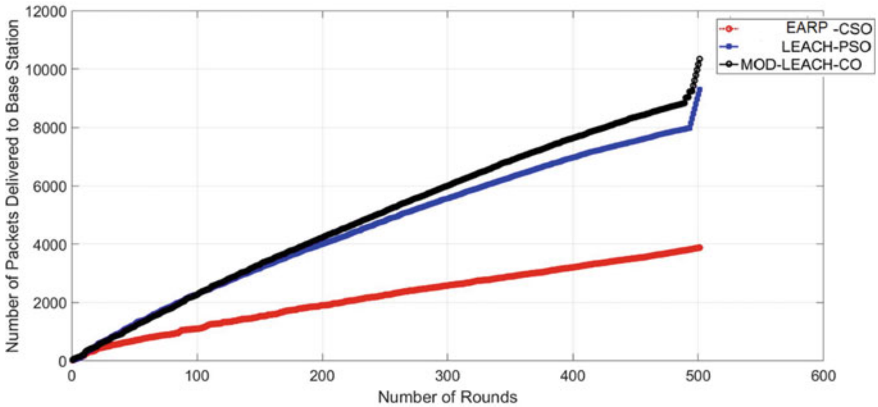


Fig. 5 Number of packets delivered to BS per round

Figure 5 specifies the total packets delivered at the base station. The maximum packets delivered by the EARP-CSO protocol reach 4000 at 500 rounds. A sudden surge can be recorded for both MOD-LEACH-CO and LEACH-PSO approaches after the 490th round indicating the increase in packets delivered. The MOD-LEACH-CO surpasses LEACH-PSO by the difference of 800 packets as seen in the figure.

5 Conclusion

In this paper, we have proposed an optimized clustering model which uses modified LEACH protocol along with the cuttlefish optimization. The model aims to optimize the parameters such as average residual energy of nodes, minimizing the number of nodes and maximizing their coverage and connectivity, thereby optimizing the clustering task and enhancing the lifetime as well as the throughput of the system. The work has been compared with the other two previously published similar works, and it is observed that the proposed work gives superior results to the other two methods. The work can be extended by using cryptographic methods or intrusion detection methods to remove redundant and anomalous nodes, thereby ensuring the safety of the system. Also, multi-objective optimization has proved itself as an efficient solution to obtain reliable systems with lower complexity.

References

1. Singh SK, Singh MP, Singh DK (2010) Routing protocols in wireless sensor networks—a survey. Int J Comput Sci Eng Surv (IJCSES) 1(2):63–83
2. Hossain A, Biswas PK, Chakrabarti S (2008) Sensing models and its impact on network coverage in wireless sensor network. In: 2008 IEEE region 10 and the third international conference on industrial and information systems, Dec 2008, pp 1–5
3. Su L, Wan H, Qin Y, Zhao X, Gao Y, Song X, Lu C, Gu M (2019) Synthesizing fault-tolerant schedule for time-triggered network without hot backup. IEEE Trans Ind Electron 66(2):1345–1355
4. Gavel S, Raghuvanshi AS, Tiwari S (2020) Comparative study of anomaly detection in wireless sensor networks using different kernel functions. In: Advances in VLSI, communication, and signal processing. Springer, Singapore, pp 81–89
5. Raghuvanshi AS, Tiwari S, Tripathi R, Kishor N (2012) Optimal number of clusters in wireless sensor networks: a FCM approach. Int J Sens Netw 12(1):16–24
6. Dasgupta K, Kalpakis K, Namjoshi P (2003) An efficient clustering-based heuristic for data gathering and aggregation in sensor networks. In: 2003 IEEE wireless communications and networking, 2003. WCNC 2003, vol 3, Mar 2003, pp 948–1953
7. Ramar C, Rubasoundar K (2015) A survey on data aggregation techniques in wireless sensor networks. Int J Mob Netw Des Innov 6(2):81–91
8. Hasson ST, Abd Al-Kadhum H (2017) Developed clustering approaches to enhance the data transmissions in WSNs. In: 2017 international conference on current research in computer science and information technology (ICCIT), Apr 2017, pp 99–106
9. Gherbi C, Aliouat Z, Benmohammed M (2017) A survey on clustering routing protocols in wireless sensor networks. Sens Rev
10. Roy S, Conti M, Setia S, Jajodia S (2014) Secure data aggregation in wireless sensor networks: filtering out the attacker's impact. IEEE Trans Inf Forensics Secur 9(4):681–694
11. Yen HH (2009) Optimization-based channel constrained data aggregation routing algorithms in multi-radio wireless sensor networks. Sensors 9(6):4766–4788
12. Harb H, Makhoul A, Tawbi S, Couturier R (2017) Comparison of different data aggregation techniques in distributed sensor networks. IEEE Access 5:4250–4263
13. El Fissaoui M, Beni-Hssane A, Saadi M (2019) Energy efficient and fault tolerant distributed algorithm for data aggregation in wireless sensor networks. J Ambient Intell Humaniz Comput 10(2):569–578
14. Ezhilarasi M, Krishnaveni V (2018) A survey on wireless sensor network: energy and lifetime perspective. Taga J Graph Technol 14:280
15. Liu W, Qin G, Li S, He J, Zhang X (2015) A multiobjective evolutionary algorithm for energy-efficient cooperative spectrum sensing in cognitive radio sensor network. Int J Distrib Sens Netw 11(5):581589

16. He S, Chen J, Xu W, Sun Y, Thulasiraman P, Shen X (2010) A stochastic multiobjective optimization framework for wireless sensor networks. EURASIP J Wireless Commun Netw 1
17. Kong L, Chen CM, Shih HC, Lin CW, He BZ, Pan JS (2014) An energy-aware routing protocol using cat swarm optimization for wireless sensor networks. In: Advanced technologies, embedded and multimedia for human-centric computing. Springer, Dordrecht, pp 311–318
18. Jiang S (2018) LEACH protocol analysis and optimization of wireless sensor networks based on PSO and AC. In: 2018 10th international conference on intelligent human-machine systems and cybernetics (IHMSC), Aug 2018, vol 2, pp 246–250
19. Kuila P, Jana PK (2014) Energy efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach. Eng Appl Artif Intell 33:127–140
20. Tate J, Woolford-Lim B, Bate I, Yao X (2011) Evolutionary and principled search strategies for sensornet protocol optimization. IEEE Trans Syst Man Cybern Part B (Cybern) 42(1):163–180
21. Amirthalingam K, Anuratha (2016) Improved LEACH: a modified LEACH for wireless sensor network. In: 2016 IEEE international conference on advances in computer applications (ICACA), Oct 2016, pp 255–258
22. Verma VK, Singh S, Pathak NP (2014) Comprehensive event based estimation of sensor node distribution strategies using classical flooding routing protocol in wireless sensor network. Wireless Netw 20(8):2349–2357
23. Iqbal M, Naeem M, Anpalagan A, Ahmed A, Azam M (2015) Wireless sensor network optimization: multi-objective paradigm. Sensors 15(7):17572–17620

Evaluation of Performance on 81-Level Switch-Ladder Inverter Using Minimum Components

L. Vijayaraja, R. Dhanasekar, M. Balaji, and S. Ganesh Kumar

Abstract An inverter is modeled to generate 81-level output voltage using a two-stage switch-ladder circuit. To design the source value of DC in the inverter, two different methods are identified and compared in terms of yielding more output levels with less number of circuit components. Detailed working of the inverter is studied with the switching sequence to obtain 81-level output. Comparative analysis is made with the classical inverter and some recently developed inverters. MATLAB simulation is carried out for the design, and the performance of 81-level inverter with RL load is analyzed in terms of THD. Simulation results of 81-level inverter are compared for both the algorithms, and presence of harmonic in the output voltage is found to be within the limits of IEEE standards.

Keywords Harmonic reduction · Switch ladder · Voltage sources

1 Introduction

The advancement and the importance of the inverter producing multilevel (MLI) are studied here. While comparing the conventional two-level inverter with MLI, the total harmonic distortion (THD) is very low. Also, parameters like electromagnetic interference (EMI) and change in voltages across the devices with respect to time are low [1–3].

In this paper, without varying the input voltages, output voltage waveform is generated using an ordinary PWM method for the pulse generation. According to the pulse generation algorithm, constant voltage levels are compared with the reference sinusoidal voltage for generating the switching sequences [4–7].

L. Vijayaraja (✉) · R. Dhanasekar
Department of EEE, Sri Sairam Institute of Technology, Chennai, Tamil Nadu, India
e-mail: vijayaraja.eee@sairamit.edu.in

M. Balaji · S. Ganesh Kumar
Department of EEE, Anna University, CEG Campus, Chennai, Tamil Nadu, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_4

In the new design of MLIs, the harmonic distortion is reduced, and it is explained in [8, 9]. In the proposed new topology, the quantity of devices is minimized when comparing with classical cascaded H-bridge multilevel inverter. By the selection of switching angles properly, reduction in harmonics is obtained.

The analysis, simulation, and harmonic distortion of the seven-level inverter is presented in [10]. The performance evaluation of seven-level inverter is verified with MATLAB [10].

In [11], a new cascade inverter configuration to yield multilevel output is shown. A fundamental multilevel inverter is compared with new cascaded topology. The proposed cascade topology has its advantages in generating more output levels by utilizing less switches and voltage sources. From the results obtained, it is concluded that a new multilevel inverter topology requires fewer numbers of components.

For fixing the amplitude of input sources in the asymmetric MLI, a mathematical method is obtained. Using the minimum number of input voltage sources, output voltage level and amplitude of voltage is increased. Here, a charge balance control design is used [12].

A cascaded MLI is framed by a sequence connection of several new basic units, and it is capable of generating positive levels. To yield a sine wave, H-bridge is added in the developed cascaded MLI. Comparing with the other presented cascaded design, the proposed cascaded topology [13] utilizes few IGBT's with drivers, source voltages, and diodes.

The performance of five-level inverter with the modeling of single-phase induction motor is discussed with the simulation results for speed of the induction motor and the current flow in the stator windings [14, 15].

Design of different multilevel inverter configurations is carried out to reduce the harmonics, DC voltage sources, and circuit elements [16–19].

This paper is incorporated with three parts and is given as follows: In Sect. 2, the design of 81-level multilevel inverter with switch-ladder inverter is demonstrated; Sect. 3 provides the performance of 81-level ladder inverter via simulation study for impedance load, and conclusion is provided in Sect. 4.

2 81-Level Multilevel Inverter

In the proposed work [20], cascaded SLMLI is designed to generate staircase load voltage levels based on circuit model. Generalized structure of switch-ladder multilevel inverter for which the quantity of output levels will be increased by increasing the size of DC sources or by cascading number of stages by fixing DC sources in each stage (Fig. 1a).

In the proposed work, cascaded SLMLI gives rise to the number of output voltage levels which depends on the circuit configuration. In the inverter structure as shown in Fig. 1a, the number of output levels can be amplified by either increasing the magnitude of DC sources as shown in Fig. 1b or by cascading a number of stages by fixing some DC sources in each stage.

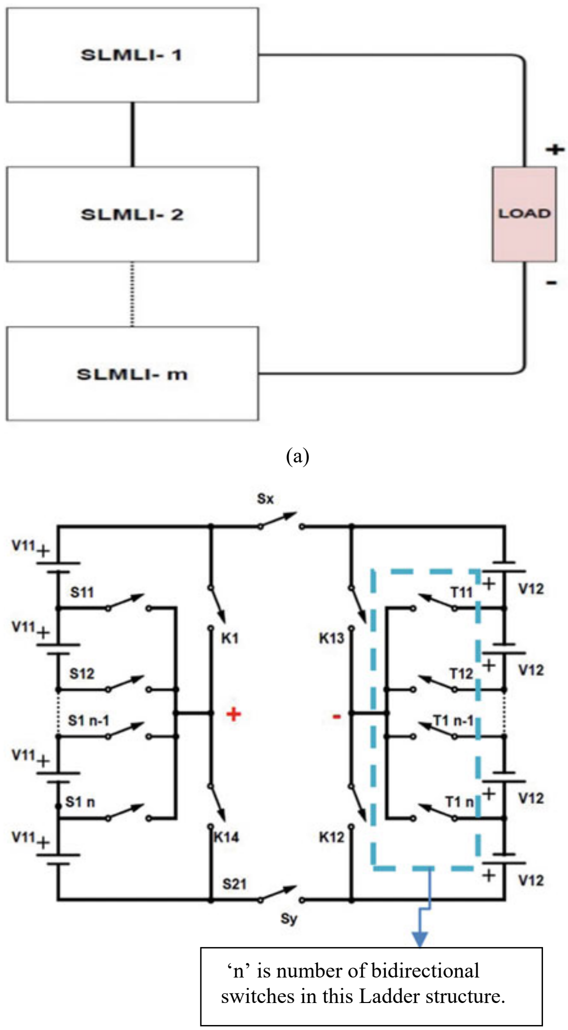


Fig. 1 Proposed SLMLI [20]

In the proposed SLMI, with higher voltage levels, makes the system huge and high cost. Moreover, the voltage values of the switches *Sx* and *Sy* are equal to the sum of all the values of the individual sources. Cascade topology with series connection of ‘*m*’ SLMLIs is prescribed to reduce the components and voltages on *Sx* and *Sy*, and it is shown in Fig. 1b. It is named as cascade switch-ladder multilevel inverter (CSLMLI). The desired output voltage of CSLMLI (V_{out}) is equal to the sum of the individual voltages of the all SLMIs, and it is given by in (1).

$$V_{\text{out}} = V_{\text{o1}} + V_{\text{o2}} + \cdots + V_{\text{o}m} \quad (1)$$

where
V_{o1} is the output voltage produced by the stage-1 (SLMLI-1)
V_{o2} is the output voltage produced by the stage-2 (SLMLI-2)
V_{om} is the output voltage produced by the stage m (SLMLI-M).

With the persistent number of power electronic devices, the maximum number of levels is obtained by considering the number of dual switches in each SLMI is equal. Alternatively,

$$2a_1 = 2a_2 = \ldots = 2a_m = 2_a. \tag{2}$$

Proposed structure shown in Fig. 1b consists of unidirectional and bidirectional switches.

2.1 *Framework for 81-Level Inverter*

A two-stage MLI is designed by two methods, and the framework is provided as below:

Method 1:

The values taken for stage-1

$$V_1 = V_3 = V_{\mathrm{dc}} \tag{3}$$

$$V_2 = V_4 = (b + 2) * V_{\mathrm{dc}} \tag{4}$$

The maximum output voltage using the abovementioned method is obtained as follows:

$$V_{\mathrm{cascaded,max}} = 2 * [(b + 1) * (V_1 + V_4)] \tag{5}$$

where b is the number of bidirectional switches in ladder structure.

The number of voltage level, switches, and DC voltage sources are given in (6), (7), and (8)

$$N_{\mathrm{level}} = 2 * \left[\left(V_{\mathrm{cascaded,max}}\right)/V_1\right] + 1 \tag{6}$$

$$N_{\mathrm{switch}} = 2 * (2b + 3) \tag{7}$$

$$N_{\mathrm{source}} = 4 * (b + 1) \tag{8}$$

Method 2: V_3 and V_4 values are taken as follows:

$$V_1 = V_{dc}, \tag{9}$$

$$V_2 = (b + 2) * V_{dc} \tag{10}$$

$$V_3 = (4b) * V_1 \tag{11}$$

$$V_4 = (b + 2) * V_3 \tag{12}$$

The maximum output voltage, number of voltage levels, switches, and DC voltage sources are obtained as follows:

$$V_{\text{cascaded,max}} = (2) * (V_1 + V_2 + V_3 + V_4) \tag{13}$$

$$N_{\text{level}} = 2 * \left[\left(V_{\text{cascaded,max}}\right)/V_1\right] + 1 \tag{14}$$

$$N_{\text{switch}} = 2 * (4b + 6) \tag{15}$$

$$N_{\text{source}} = 4 * (b + 1) \tag{16}$$

2.2 Cascaded SLMLI to Generate 81-Levels (2-Stages)

Inverter shown in Fig. 1b is able to yield ‘*n*’ levels, but rating of IGBT increases with the magnitude of voltage increase. This problem is solved by cascading number of switch-ladder multilevel inverters to generate ‘*n*’ levels of output. Figure 2 gives the circuit diagram of cascaded SLMLI to generate 81-level output. The requirement of switch and sources is same for stage-1 and stage-2.

From Fig. 2 [20], for stage-1, unidirectional switches used are represented as T_{11}, T_{12}, T_{13}, T_{14}, S_{1X}, S_{1Y}, and the bidirectional switches are represented as S_{11}, S_{12}. For stage-2, unidirectional switches are represented as T_{21}, T_{22}, T_{23}, T_{24}, S_{2X}, S_{2Y}, and the bidirectional switches are represented as S_{21}, S_{22}. By cascading two stages, stress because of voltages between the switches is reduced. Thus, the overall system cost will be reduced.

To generate multilevel voltages at the output, turn on the corresponding set of switches for the particular instances. For stage-1, three switches should be turn on at a time, and the rest of the switches should be in ‘OFF’ state. In cascaded SLMLI, numbers of stages are cascaded, and at any instant of time, three switches are on.

Modes for obtaining various voltage levels are provided in Table 1. Zero voltage

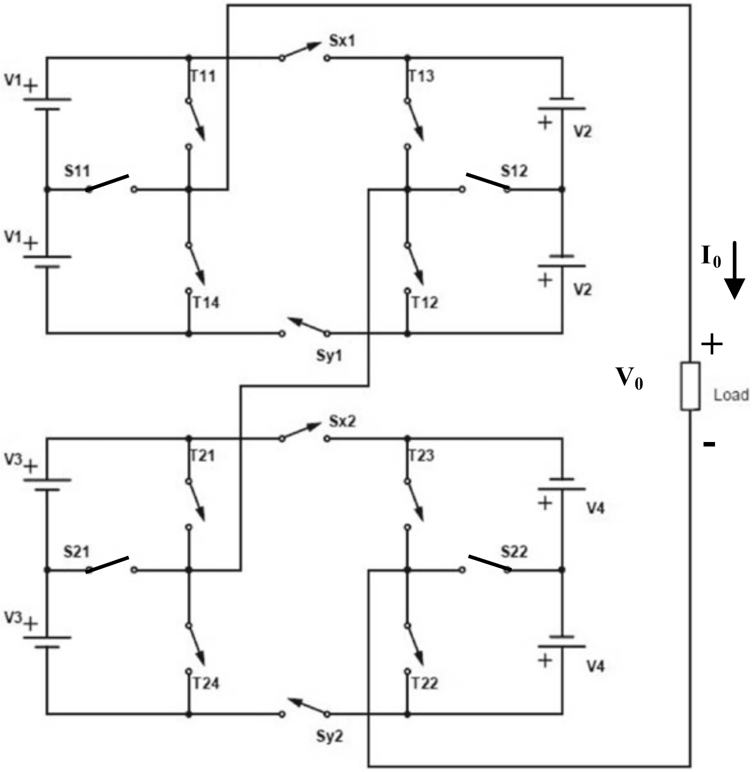


Fig. 2 Cascaded two stages of SLMLI [20]

is achieved by turning on T_{11}, T_{13}, S_{1X} in stage-1 and T_{21}, T_{23}, S_{2X} in stage-2. V_1 voltage is achieved by turning on T_{12}, S_{11}, S_{y1} in stage-1 and T_{21}, T_{23}, S_{2X} for stage-2.

Performance of two-stage SLMLI is compared by considering two sets of DC source values used in the construction of inverter. DC source magnitudes and the performance of inverter using both the methods are in Table 2.

From Table 2, it can be concluded that for the method 2, one can realize higher voltage levels with few devices. Furthermore, peak value obtained using method 2 is high in comparison with that obtained using method 1.

From Eqs. 6, 7, and 8, the method 1 yields a voltage with 33-level output at load. Also, from Eqs. 14, 15, and 16, the method 2 gives a voltage with 81-levele at load. From method 1 and method 2, using same quantity of switches and sources, the method 2 produces higher voltage steps compared to method 1.

This type of inverter with method 2 can be used in many electric power applications like FACTS devices, electrical vehicle, and clean energy power systems.

Table 3 indicates the comparison of existing MLIs with the proposed 81-level MLI from Table 3; it can be concluded that SLMLI seems to be better than DC, FC, and CHB multilevel inverters.

Table 1 Switching status of inverter to generate 81-level utilizing cascaded two-stage circuit

No. of level	Voltage levels	T_{11}	T_{12}	T_{13}	T_{14}	S_{11}	S_{12}	S_{1X}	S_{1Y}	T_{21}	T_{22}	T_{23}	T_{24}	S_{21}	S_{22}	S_{2X}	S_{2Y}
1	0	1	0	1	0	0	0	1	0	1	0	1	0	1	1	1	0
2	V_1	0	1	0	0	1	0	0	1	1	0	1	0	0	0	1	0
3	$V_1 + V_1$	1	1	0	0	0	0	0	1	1	0	1	0	0	0	1	1
4	V_3	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	1
5	V_4	0	0	0	0	1	1	0	1	1	0	1	0	0	0	1	1
6	$V_1 + V_2$	1	0	0	1	0	1	0	1	1	0	1	0	0	0	1	1
7	$V_1 + V_1 + V_2$	0	0	1	0	1	0	0	1	1	0	1	0	0	0	1	1
8	$V_1 + V_2 + V_2$	0	0	1	0	0	0	0	1	1	0	1	0	0	0	1	1
9	$V_1 + V_1 + V_2 + V_2$	1	0	1	0	0	0	0	1	1	0	1	0	0	0	1	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
14	$V_1 + V_4$	0	1	0	0	1	0	0	1	0	0	0	1	0	1	0	1
15	$V_1 + V_1 + V_4$	1	1	0	0	0	0	0	1	0	0	0	1	0	1	0	1
16	$V_2 + V_4$	0	0	0	1	0	1	0	1	0	0	0	1	0	1	0	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
34	$V_1 + V_3 + V_3 + V_4 + V_4$	0	1	0	0	1	0	0	1	1	0	1	0	0	0	0	1
35	$V_1 + V_1 + V_3 + V_3 + V_4 + V_4$	1	1	0	0	0	0	0	1	1	0	1	0	0	0	0	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
39	$V_2 + V_2 + V_3 + V_3 + V_4 + V_4$	0	0	1	0	1	0	0	1	1	0	1	0	0	0	0	1
40	$V_1 + V_2 + V_2 + V_3 + V_3 + V_4 + V_4$	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1
41	$V_1 + V_1 + V_2 + V_2 + V_3 + V_3 + V_4 + V_4$	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1

Table 2 Comparison of method 1 and 2 for stage-two circuit

Details	Method 1	Method 2
	$V_1 = V_3 = 5$ V and $V_2 = V_4 = 15$ V	$V_1 = 5$ V, $V_2 = 15$ V, $V_3 = 20$ V and $V_4 = 60$ V
Maximum output voltage (V)	80	200
Number of output levels	33 levels	81 levels
Number of switches	20 switches	20 switches
Number of DC voltage sources	8	8

Table 3 Comparison of basic MLI with SLMLI

S. No.	Topology	81-levels	
		Number of switches	Number of DC sources
1	Diode-clamped MLI	160	1
2	Flying capacitor MLI	160	1
3	Cascaded h-bridge MLI	160	40
4	**SLMLI**	**20**	**8**

SLMLI is to make comparision with conventional inverters

Table 4 Comparison of SLMLI with other MLI

S. No.	Topology	Number of switches	Number of DC sources	Number of levels
[21]	AMLI	10	5	15
[22]	CCNPI	20	4	17
[23]	AMCI	12	4	49
[24]	DMCI	12	3	27
Proposed	SLMLI	20	8	81

Table 4 explains the comparison of existing MLIs with the proposed 81-level MLI. From Table 4, it can be shown that SLMLI seems to be better than the MLI such as AMLI [21], CCNPI [22], AMCI [23], and DMCI [24].

3 Performance of Ladder Inverter from Simulation Study for RL Load

For generating 81-levels in the output using cascaded two stages of switch-ladder multilevel inverter, a pulse generator is used for creating pulses for the switches in two stages.

Table 5 Specification of two-stage SLMLI

S. No.	Parameters	Values
Input voltages of SLMLI-1		
1	V_1	5 V
2	V_2	15 V
Input voltages of SLMLI-2		
3	V_3	20 V
4	V_4	60 V
Output frequency		
5	f	50 Hz
Load		
6	R	2 Ω
7	L	2 mH

The simulation design is carried with MATLAB, and the design values are as mentioned in Table 5.

FFT analysis of 1.27 and 0.44% as THD for load voltage and current is shown in Fig. 3b. From Fig. 3b, current THD is reduced due to inductance load added in the circuit which smoothen the current waveform, and small phase swift occurs due to inductance.

Figure 4 represents the load voltage and current waveform connected to a load of $R = 2\ \Omega$, $L = 2$ mH with cascaded two stages of switch-ladder multilevel inverter.

4 Conclusion

Analysis, design, and the software implementation of the ladder inverter are completed in this paper. The comprehensive literature survey was presented for various multilevel inverter topologies, control schemes, and THD improvement methods. The simulation work is completed in MATLAB, and the performance of the proposed inverter is analyzed in terms of minimizing the circuit components with the algorithms proposed.

Out of the two algorithms proposed, for the design of voltage sources with the same number of circuit components, second algorithm is able to generate 146% more output voltage levels in comparison to the first algorithm. Specifications used for the second method are shown in Table 6.

The result shows the performance of 81-level inverter is better while using second algorithm than that of the other algorithm presented and satisfactory results are obtained. The harmonic presence in the output voltage and current is obtained as 1.27% and 0.44%, respectively, and it is found to be less than the IEEE standard.

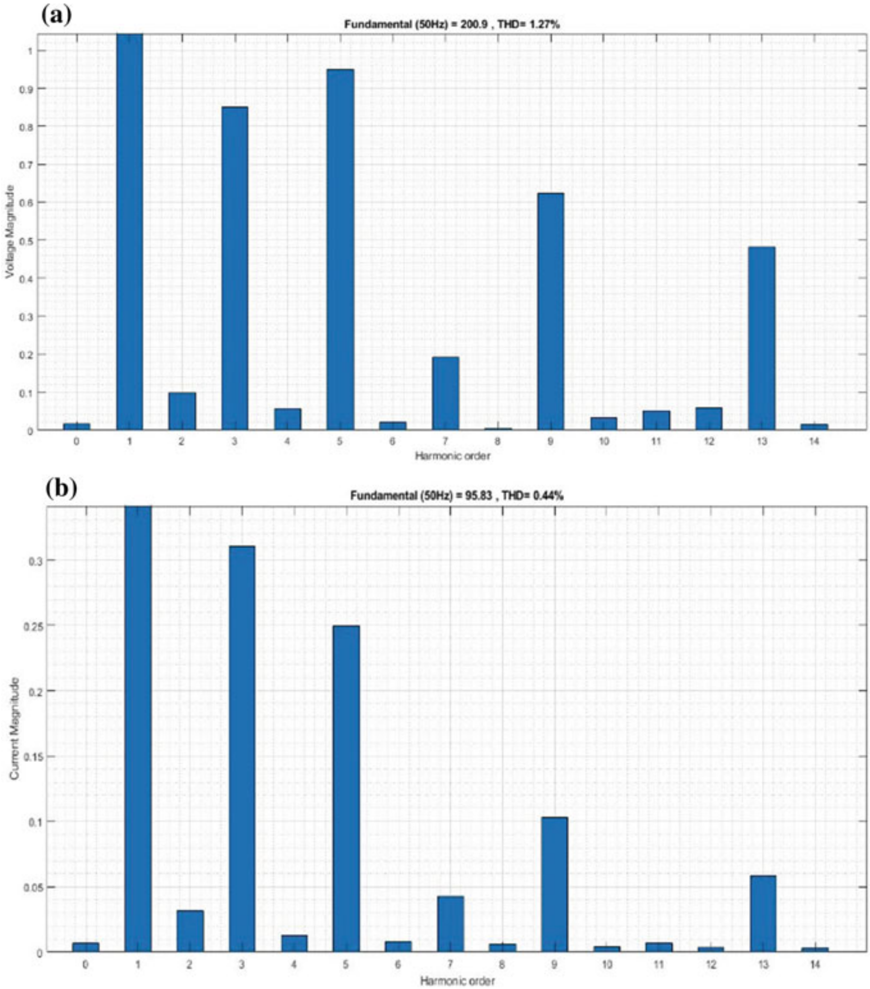


Fig. 3 FFT analysis for voltage (**a**) and current (**b**) for RL load

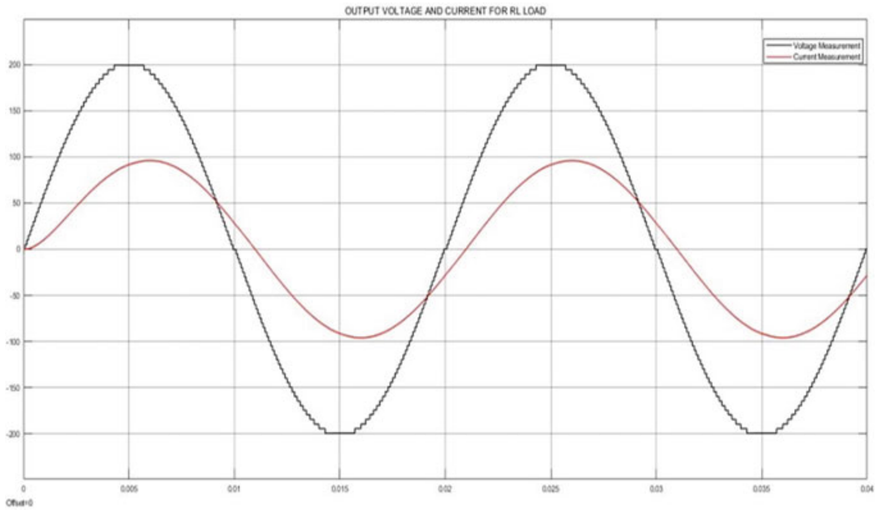


Fig. 4 Output voltage and current waveforms for RL load

Table 6 Specifications for second algorithm

S. No.	Parameters	Details
1	Source value	$V_1 = 5$ V, $V_2 = 15$ V, $V_3 = 20$ V and $V_4 = 60$ V
2	Number of voltage levels	81
3	Peak voltage	200 V
4	Number of devices and sources used	20 and 8

References

1. Vijayaraja L, Kumar SG, Rivera M (2016) A review on multilevel inverter with reduced switch count. In: 2016 IEEE international conference on automatica (ICA-ACCA), Curico, pp 1–5
2. Koshti AK, Rao MN (2017) A brief review on multilevel inverter topologies. In: International conference on data management, analytics and innovation (ICDMAI)
3. Singh J, Dahiya R, Saini LM (2018) Recent research on transformer based single DC source multilevel inverter: a review. Renew Sustain Energy Rev 82(Part 3):3207–3224
4. Hajizadeh M, Fathi SH (2016) Fundamental frequency switching strategy for grid-connected cascaded H-bridge multilevel inverter to mitigate voltage harmonics at the point of common coupling. IET Power Electron 9(12):2387–2393
5. Sanjay PS, Tanaji PR, Patil SK (2018) Symmetrical multilevel cascaded H-bridge inverter using multicarrier SPWM technique. In: 2018 3rd international conference for convergence in technology (I2CT), Pune, pp 1–4
6. Pilli NK, Meena RK, Singh SK (2017) A simplified gate pulse generation technique for modified multilevel DC-link inverter. In: 2017 IEEE industry applications society annual meeting, Cincinnati, OH, pp 1–6
7. Omer P, Kumar J, Surjan BS (2020) A review on reduced switch count multilevel inverter topologies. IEEE Access 8:22281–22302

8. Hosseinzadeh MA, Sarebanzadeh M, Rivera M, Babaei E, Wheeler P (2021) A reduced single-phase switched-diode cascaded multilevel inverter. IEEE J Emerg Sel Top Power Electron. https://doi.org/10.1109/JESTPE.2020.3010793
9. Ganesh B et al (2019) Implementation of twenty seven level and fifty one level inverter using constant voltage sources. In: 2019 IEEE CHILEAN conference on electrical, electronics engineering, information and communication technologies (CHILECON), Valparaiso, pp 1–4
10. Bais YS, Deshpande SB, Muley SP (2017) Implementation of a new H-bridge 7-level inverter fed induction motor with lesser harmonic values. In: 2017 international conference on energy, communication, data analytics and soft computing (ICECDS), Chennai, pp 257–262
11. Babaei E, Laali S, Bayat Z (2015) A single-phase cascaded multilevel inverter based on a new basic unit with reduced number of power switches. IEEE Trans Ind Electron 62(2):922–929
12. Vijayaraja L, Kumar SG, Rivera M (2018) A new topology of multilevel inverter with reduced part count. In: 2018 IEEE international conference on automation/XXIII congress of the Chilean association of automatic control (ICA-ACCA), Concepcion, pp 1–5
13. Raman G, Imthiyas A, Raja MD, Vijayaraja L, Kumar SG (2019) Design of 31-level asymmetric inverter with optimal number of switches. In: 2019 IEEE international conference on intelligent techniques in control, optimization and signal processing (INCOS), Tamil Nadu, pp 1–3
14. Ganesh P, Shanmugavadivu N, Santha K (2018) Single-phase 63-level modular multilevel inverter fed induction motor drive for solar PV applications. In: 2018 4th international conference on electrical energy systems (ICEES), Chennai, pp 596–604
15. Sivakumar K, Das A, Ramchand R, Patel C, Gopakumar K (2010) A five-level inverter scheme for a four-pole induction motor drive by feeding the identical voltage-profile windings from both sides. IEEE Trans Ind Electron 57(8):2776–2784
16. Rathore V, Yadav KB, Dhamudia S (2021) An implementation of nine-level hybrid cascade multi-level inverter (HCMLI) using IPD-topology for harmonic reduction. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. Lecture notes in electrical engineering, vol 692. Springer, Singapore
17. Paul S, Mahato B, Majumdar S, Jana KC (2021) Diode switch-based 17-level inverter with lesser power electronic elements. In: Nath V, Mandal JK (eds) Proceedings of the fourth international conference on microelectronics, computing and communication systems. Lecture notes in electrical engineering, vol 673. Springer, Singapore
18. Mahato B, Majumdar S, Vatsyayan S, Jana KC (2021) Implementation of a multi-level inverter (MLI) with a new structure with different values of DC sources. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. Lecture notes in electrical engineering, vol 692. Springer, Singapore
19. Ponkumar S, Rivera SM, Kamroon F, Kumar SG (2017) Realization of cascaded multilevel inverter. In: 2017 CHILEAN conference on electrical, electronics engineering, information and communication technologies (CHILECON), pp 1–7
20. Alishah RS, Hosseini SH, Babaei E, Sabahi M (2017) Optimal design of new cascaded switch-ladder multilevel inverter structure. IEEE Trans Ind Electron 64(3):2072–2080
21. Mahato B, Majumdar S, Vatsyayan S, Jana KC (2020) A new and generalized structure of MLI topology with half-bridge cell with minimum number of power electronic devices. In: IETE technical review. Taylor & Francis, p 0256-4602
22. Odeh C, Agu V (2017) Cascaded common nodal point multi-level inverter topology. Electr Power Compon Syst 45:1–14
23. Salari E, Darvish Falehi A (2019) A novel 49-level asymmetrical modular multilevel inverter: analysis, comparison and validation. Analog Integr Circ Sig Process 101:611–622
24. Francis R, Meganathan D (2019) A dual-mode cascaded H-bridge multilevel inverter for improving THD. Electr Eng 101:225–237

Design of Error-Tolerant and Low-Power Approximate Full Adder

Ketki C. Pathak, Anand D. Darji, Jignesh N. Sarvaiya, Zinal Bhatt, Anjali Gangadwala, Shreya Diwan, and Azba Patel

Abstract Many engineering problems, which do not require exact computations, lead to a loss of dynamic power and device utilization. “Approximate Computing” term indicates that it is solving the problems approximately instead of an exact computation, by which we can save power consumption, area, device utilization, and critical path timings with acceptable loss in quality. Approximate computing helps us to improve the energy efficiency and reducing area which will be beneficial in mobile devices as well as online groundwork. Human beings can gather useful information from apparent (approximate) outputs in most of the multimedia applications. In this paper, we are analyzing majority logic-based approximate adder (Zhang et al. in Design of majority logic (ML) based approximate full adders. 2018 IEEE international symposium on circuits and systems (ISCAS). IEEE, 2018, pp 1–5 [1]) and propose the design of an approximate adder named as almost full adder. The designs are analyzed, and improvement is achieved in comparison with ML-based approximate adder. The proposed designs are found more prior in terms of error distance (ED), mean error distance (MED), normalized mean error distance (NMED), and power and device utilization. The proposed design consumes 6% less power than an accurate full adder of one bit. More than that, it also has 14% less errors compared to ML-based full adder and 14% less error distance compared to carry-based approximate full adder.

Keywords Approximate arithmetic circuits · Accurate full adder · ML-based approximate adder · Almost full adder

K. C. Pathak (✉) · Z. Bhatt · A. Gangadwala · S. Diwan · A. Patel
Electronics and Communication Department, Sarvajanik College of Engineering and Technology, Surat, India
e-mail: ketki.joshi@scet.ac.in

A. D. Darji · J. N. Sarvaiya
Electronics Engineering Department, S.V.N.I.T., Surat, India
e-mail: add@eced.svnit.ac.in

J. N. Sarvaiya
e-mail: jns@eced.svnit.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_5

1 Introduction

Many scientific and engineering problems do not require exact computations. Digital circuits are vital in most of the emerging technology. In nanoscale technology, we are trying to reduce the area of silicon devices, for such cases if we approach the concept of approximation rather than exact, so we will be able to reduce power, delay, and area effectively. To perform arithmetic operation for any digital circuit through exact computations, it is slightly difficult compared to approximation concept. Hence, approximation has been needed for the hour!

Multimedia is easy to interrupt by humans. Human perception is not that sensitive to high frequency changes; hence, approximate or less than optimal results have also satisfied the need. Approximate arithmetic circuits have been considered as a considerable option for error-tolerant applications to trade-off some accuracy to maintain other circuit-based metrics, such as area, delay, and power.

Addition is the basic operations in circuit design. Addition is a key component in a logic-based circuit. The delay and power of arithmetic circuits improve the performance. Basically, any arithmetic circuit can be implemented using an adder. As we consider the approximation theory, we are neglecting or modifying circuit design of adder so that we can efficiently reduce the power, area, and delay of the digital circuit. There are many types of approximate adders available as shown in Fig. 1.

In many emerging and nanoscale technologies, increasing the value of power is a bit challenging for designing digital circuits. In this technical literature, we are

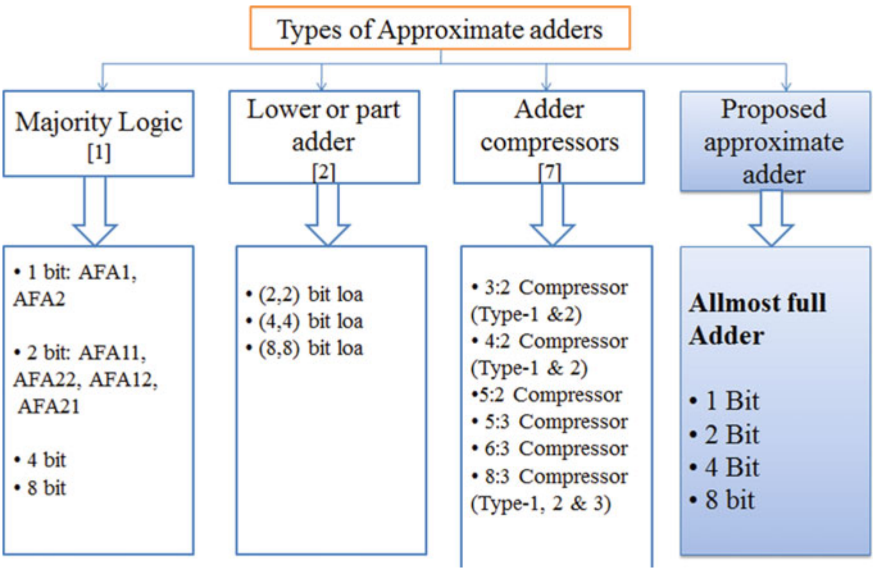


Fig. 1 Types of approximate full adder

analyzing ML-based approximate adder, also comparing the proposed design of approximate full adder in terms of error distance (ED), mean error distance (MED), and normalized mean error distance (NMED) to evaluate the efficiency of proposed design.

Proposed design in this technical literature is modified in terms of carry of the accurate full adder. In many image processing applications such as transform-based image compression, where we can neglect carry, however, the visual quality will not be degraded.

Another portion of this technical literature is as follows. Section 2 consists of reviewing different technical literatures. The proposed design of approximate full adder is in Sect. 3. Section 4 represents the comparison with ML-based full adder, accurate full adder, and carry-based approximate full adder. Section 5 summarizes the technical literature.

2 Literature Review

In this section, approximate adders are briefly introduced which are widely used nowadays. References [1–3] are for detailed information of ML-based approximate full adder, lower or part approximate full adder, and adder compressors, respectively.

2.1 Lower or Part Approximate Adder

This paper detailed several sequential adders and also discussed approximate adder, lower or part adder (LOA), and probabilistic full adder (PFA). From the term MED, this literature review suggests that probabilistic full adder (PFA) shows little advantage over the LOA, in terms of more gate error rates, but when it comes to reliability-area trade-off, in that case, LOA is more suitable, due to use of approximate logic.

2.2 ML-Based Approximate Full Adder

Review of ML-based full adder is shown in this section. This literature review has designed ML-based approximate full adders for 1 bit, 2 bit, 4 bit, and 8 bit. They had evaluated efficiency of their proposed solution with various parameters such as ED, MED, and NMED. According to this literature, these designs have less area, delay, and less number of logic gates. Ensuring that their proposed design results in an improvement of at least up to 50% in delay and 67% in area for the 4-bit design, compared to accurate full adder [1].

2.3 Adder Compressors

Compressions between 8-2 adder compressors are shown in this literature review, which are simulated at 100 MHz frequency. Smaller hierarchical versions cannot be represented by 8-2 monolithic adder compressors, which can be done by their proposed solution. Shorter critical path leads to significantly higher maximum operational frequencies for their proposed circuits, but the power consumption is almost similar to existing design [3].

Adder compressors may give us high error tolerance, but it consumes similar power as an accurate one [4–6]. Similarly, in LOA design, power consumption, area utilization, and delay are less, but it is not that much efficient compared to other approximate full adders. Hence, for emerging technology and for betterment, representing a new design of approximate adder, namely almost full adder, has similar power consumption and utilization with high error tolerance compared to ML-based full adder [7–12].

3 Proposed Design for Approximate Adder

This portion incorporates the proposed design of approximate full adder named as almost full, which is implemented with the help of VIVADO 2015.2 software.

3.1 Almost Full Adder

We represent a new approximate full adder, namely almost full adder. Consider Fig. 2 which shows the design of almost full adder, where we had done modification in the equation of carry.

$$\mathrm{Sum} = A \,\mathrm{XOR}\, B \,\mathrm{XOR}\, C$$

$$\mathrm{Carry} = BC \tag{1}$$

Therefore, in Eq. (1), we are considering only B and C, neglecting the rest of the part of the traditional equation of carry. Consider Table 1 for the truth table of almost full adder, which we are also comparing with accurate almost full adder.

Carry of almost full adder is nearly the same as accurate full adder. As per Table 1, there are only three terms in carry which contains error, but that will not affect significantly in most of the applications. To identify the efficiency of approximate adders, ED is one of the most important parameters. These terms are defined as follows.

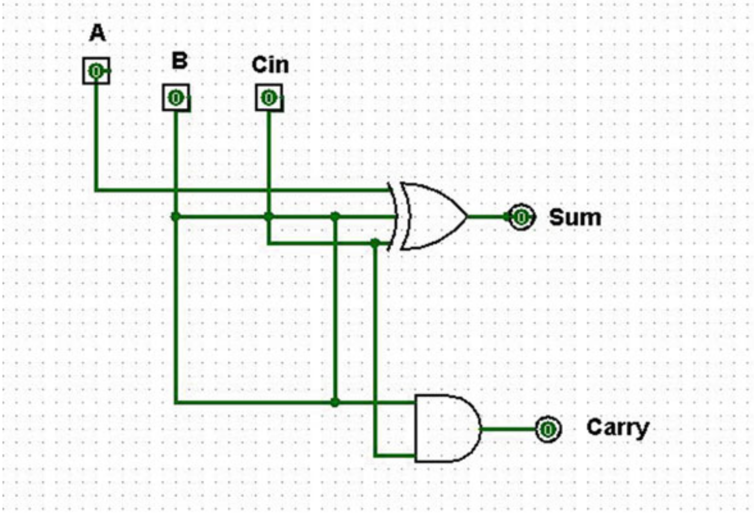


Fig. 2 Design of one-bit almost full adder

Table 1 Truth table for accurate and almost full adder

Inputs			Results of accurate full adder		Results of almost full adder	
A	*B*	*C*	Sum	Carry	Sum	Carry
0	0	0	0	0	0	0
0	0	1	1	0	1	*1*
0	1	0	1	0	1	0
0	1	1	0	1	0	1
1	0	0	1	0	1	0
1	0	1	0	1	0	*0*
1	1	0	0	1	0	*0*
1	1	1	1	1	1	1

$$\text{Error Distance (ED)} = |\text{Accurate Result} - \text{Approximate Result}| \tag{2}$$

$$\text{Mean Error Distance (MED)} = \frac{\sum \text{ED}}{N} \tag{3}$$

$$\text{Normalized Mean Error Distance (NMED)} = \frac{\text{MED}}{M} \tag{4}$$

where N and M are the number of all possible inputs and maximum value of the result, respectively.

Evaluate efficiency of one-bit almost full adder, mean error distance, and normalized mean error distance are given by:

$$\mathrm{MED} = \frac{1}{8}(0+0+0+0+0+0+0+0) = 0$$

$$\mathrm{NMED} = \frac{\mathrm{MED}}{3} = 0 \tag{5}$$

Considering one-bit almost full adder, we can extend for 2-bit, 4-bit, and 8-bit almost full adder which will be useful in many applications. Figure 3 shows the design configuration of extended almost full adder.

We are implementing all the designs by using VIVADO 2015.2 software, and the results of which are shown in Fig. 4. Table 2 comprises other measuring parameters such as device utilization, input–output (I/O), and power consumption.

4 Result Analysis

4.1 Comparison of ML-Based Full Adder and Almost Full Adder

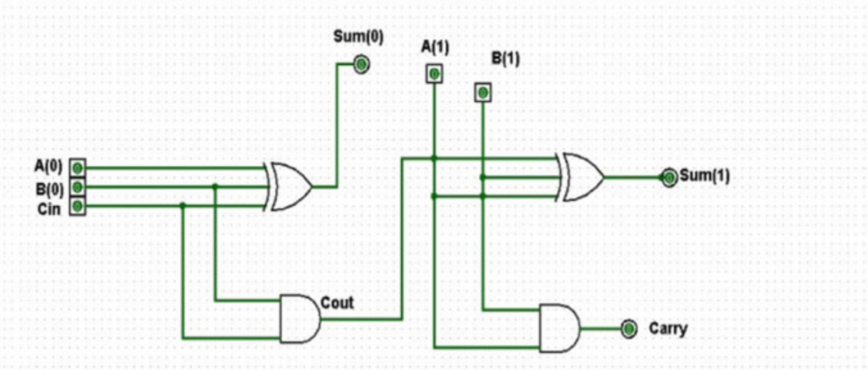
In this section, comparison in terms of power, device utilization, and error distance between the ML-based full adder and almost full adder is shown in Table 4. Before that, simulation results of ML-based full adder are shown in Table 3. Proposed approximate one-bit full adder comprises 0 units of error distance, whereas the ML-based one-bit full adder comprises two units of error distance shown in Table 5. Error distance for proposed design and for ML-based full adder can be calculated as shown in Eqs. (6) and (7), respectively.

$$\text{Error Distance (ED)} = |\text{Accurate Results} - \text{Approximate Results}| = |8-8| = 0 \tag{6}$$

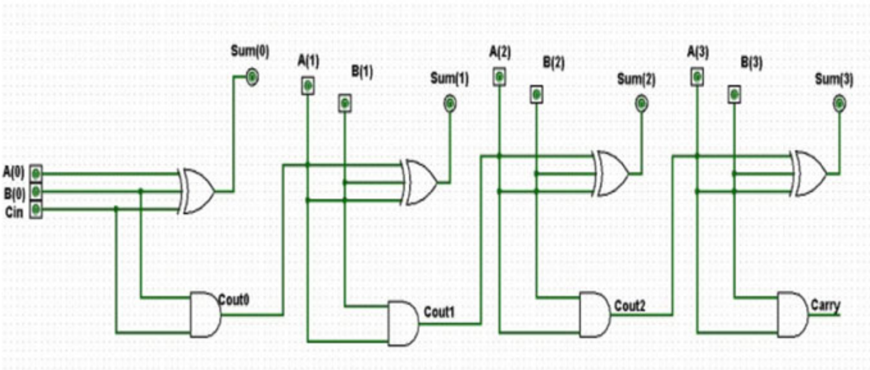
$$\text{Error Distance (ED)} = |\text{Accurate Results} - \text{Approximate Results}| = |8-6| = 2 \tag{7}$$

The proposed design of approximate adder comprises less device utilization and also consumes the equivalent power with less amount of error distance, and that entire phenomena are shown in Fig. 5a, b, respectively.

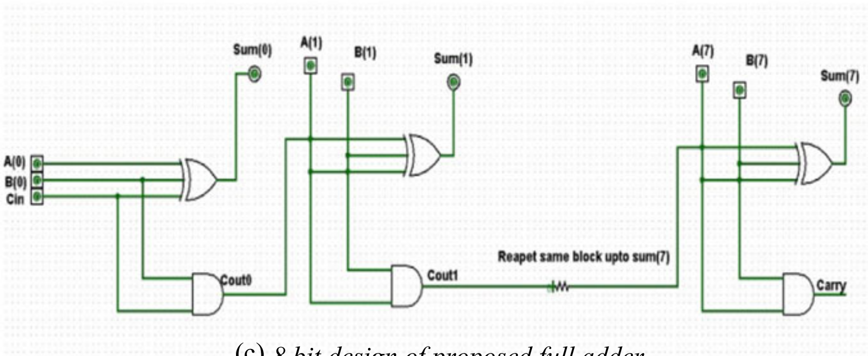
More than that, from the comparison of the truth table of two bit, also we can ensure that our proposed algorithm is way better than ML-based full adder. Table 6 comprises that error which is highlighted as yellow in color. In this table, ML-based full adder is more comparable to proposed approximate full adder. Same as the one bit, we can evaluate MED and NMED for extended bits as per Table 7.



(a) *2 bit design of Proposed full adder*

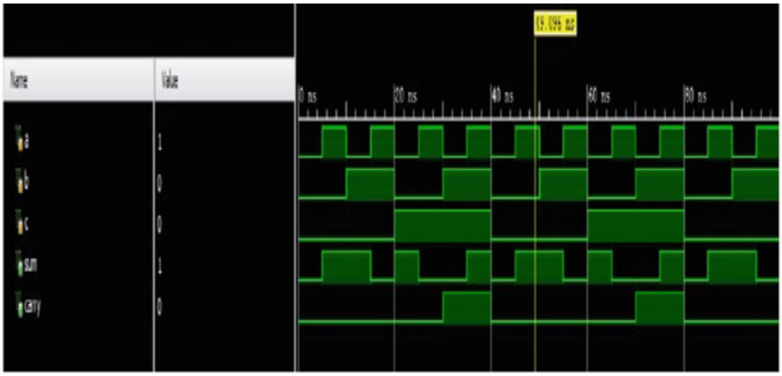


(b) *4 bit design of Proposed full adder*

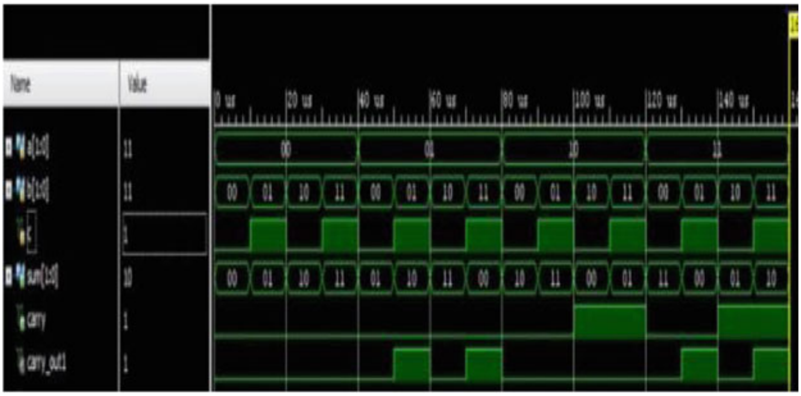


(c) *8 bit design of proposed full adder*

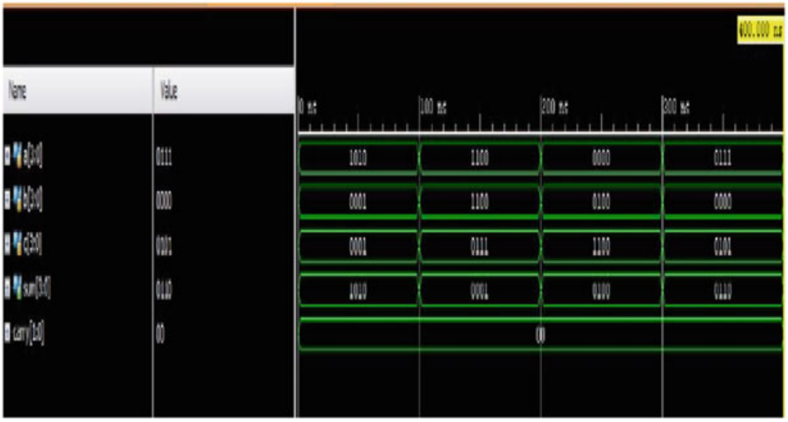
Fig. 3 Design of proposed full adder



(a) *Result of 1 bit Proposed full adder*



Result of 2 bit Proposed full adder



Result of 4 bit Proposed full adder

Fig. 4 Configuration of almost full adder

Results of 8 bit Proposed full adder

Fig. 4 (continued)

Table 2 Results of almost full adder

No. of bits	Dynamic power	LUT		LUT (%)	IO		IO (%)
		Utilization	Available		Utilization	Available	
1	0.864	1	32,600	0.0	5	150	3.33
2	1.365	2	32,600	0.01	8	150	5.33
4	1.319	2	32,600	0.01	7	150	4.67
8	4.879	8	134,600	0.01	25	400	6.25

Table 3 Results of ML-based full adder

Circuit type	Circuit name	Dynamic power	LUT		I/O	
			Utilization	Available	Utilization	Available
1 bit	AFA1	0.72	2	32,600	5	150
	AFA2	0.544	2	32,600	3	150
2 bit	AFA11	1.613	2	32,600	8	150
	AFA12	1.406	2	32,600	8	150
	AFA21	1.297	2	32,600	9	150
	AFA22	1.217	2	32,600	9	150
4 bit	AFA12 21	2.202	5	32,600	15	150
	AFA12 12	2.193	5	32,600	15	150
	AFA21 12	2.157	4	32,600	15	150
	AFA21 21	2.202	4	32,600	15	150
8 bit	AFA12 121212	4.499	10	32,600	29	150
	AFA12 122121	4.468	9	32,600	29	150
	AFA21 211212	4.444	9	32,600	29	150
	AFA21 212121	4.376	8	32,600	29	150

Table 4 Comparison of ML-based full adder [1] and proposed approximate full adder

Approximate adders		Power (W)	Utilization (%)	IO (%)
ML-based approximate full adder [1]	1 bit (AFA1)	0.72	0.01	3.22
	2 bit (AFA22)	1.217	0.01	6
	4 bit (AFA2121)	2.202	0.01	10
	8 bit (AFA12121212)	4.499	0.03	19.33
Almost full adder	1 bit	0.864	0.00	3.33
	2 bit	1.365	0.01	5.33
	4 bit	1.319	0.01	4.67
	8 bit	4.897	0.01	6.25

Table 5 Truth table of one-bit full adder

Input			Accurate one-bit full adder		Approximate one-bit full adder			
					Almost full adder		ML-based AFA 2 [1]	
A	*B*	*C*	Cout	*S*	Cout	*S*	Cout	*S*
0	0	0	0	0	0	0	0	0
0	0	1	0	1	*1*	1	*1*	*0*
0	1	0	0	1	0	1	0	1
0	1	1	1	0	1	0	1	0
1	0	0	0	1	0	1	0	1
1	0	1	1	0	*0*	0	1	0
1	1	0	1	0	*0*	0	*0*	*1*
1	1	1	1	1	1	1	1	1

4.2 Comparison with Accurate Full Adder

This portion comprises the comparison between accurate full adder and almost full adder in terms of dynamic power and device utilization. Consider graph (2) shown in Fig. 6, which shows that one-bit accurate full adder consumes 0.984 W dynamic power, whereas one-bit almost full adder consumes 0.864 W dynamic power which is 6% less than accurate one. As we extend the bits, the difference percentage will also increase, which means that we will be capable of storing maximum power with minimal error compared to accurate full adder.

4.3 Comparison with Carry-Based Approximate Full Adder

This section comprises comparison in terms of error distance between carry-based approximate full adder [13] and proposed almost full adder. Error distance is 14%

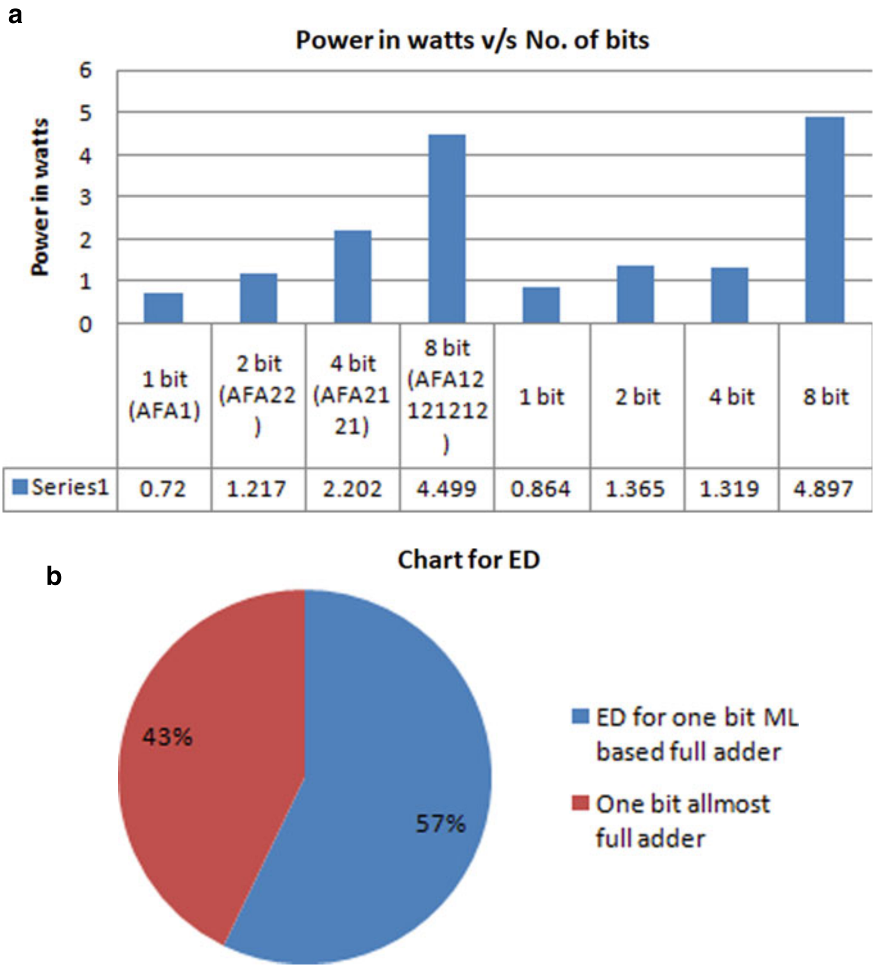


Fig. 5 **a** Comparison power utilization of proposed full adder. **b** Graph of comparison for error distance measurement

less as per the graph shown in Fig. 7. Along with that, we can have more details from the comparison Table 8 of carry-based approximate full adder and proposed approximate full adder.

Table 6 Truth table of 2-bit approximate full adder

Input				Cin	Accurate 2-bit full adder		ML-based 2-bit full adder [1]		Almost 2-bit full adder	
A		*B*								
*a*0	*a*1	*b*0	*b*1		Sum	Carry	Sum	Carry	Sum	Carry
0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	1	*1*	1	0
0	0	1	0	0	1	0	*0*	0	1	0
0	0	1	1	1	11	0	11	*1*	11	*1*
0	1	0	0	0	1	0	*0*	0	1	0
0	1	0	1	1	10	0	*1*	*1*	10	*1*
0	1	1	0	0	11	0	*0*	0	11	0
0	1	1	1	1	0	1	*11*	*0*	0	1
1	0	0	0	0	1	0	*0*	0	*10*	0
1	0	0	1	1	11	0	11	*1*	11	*1*
1	0	1	0	0	0	1	*10*	*0*	0	*0*
1	0	1	1	1	1	1	11	1	1	1
1	1	0	0	0	11	0	*0*	0	11	0
1	1	0	1	1	0	1	*11*	1	0	1
1	1	1	0	0	1	1	10	*0*	1	*0*
1	1	1	1	1	10	1	*11*	1	10	1

Table 7 Results of MED and NMED

	ML-based full adder [1]			Almost full adder		
	2 bit	4 bit	8 bit	2 bit	4 bit	8 bit
MED	0.75	2.87	47.02	0.125	0.125	0.143
NMED	0.107	0.092	0.092	0.042	0.042	0.045

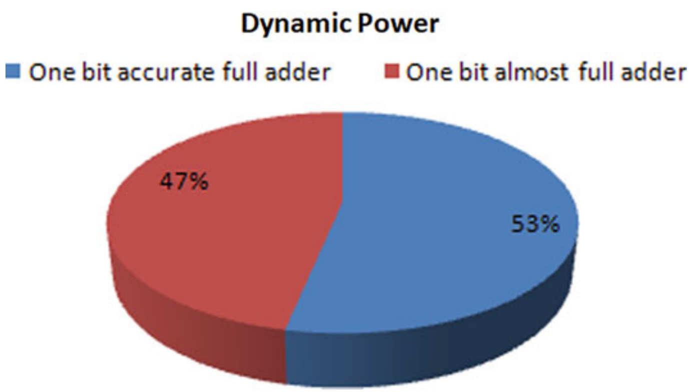


Fig. 6 Graph of comparison dynamic power

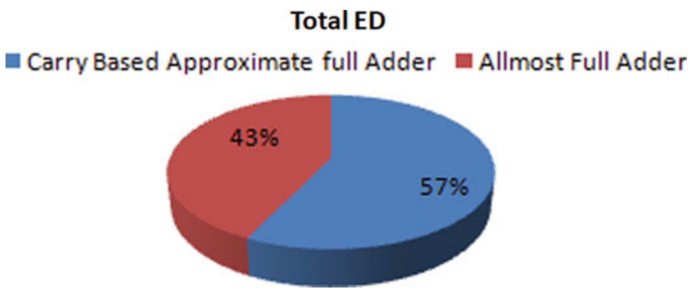


Fig. 7 Graph of comparison error distance

Table 8 Comparison truth table for carry-based approximate full adder [13] and proposed almost approximate full adder

Inputs			Outputs											
A	*B*	*C*	CBA1 [13]			CBA2 [13]			CBA3 [13]			Almost full adder		
			Sum	Carry	ED	Sum	Carry	ED	Sum	Carry	ED	Sum	Carry	ED
0	0	0	*1*	0	1	*1*	0	1	*1*	0	1	0	0	0
0	0	1	*0*	*1*	2	1	0	0	*0*	*1*	2	1	*1*	1
0	1	0	1	0	0	1	0	0	1	0	0	1	0	0
0	1	1	0	1	0	0	1	0	0	1	0	0	1	0
1	0	0	1	0	0	*0*	*1*	2	1	0	0	1	0	0
1	0	1	0	1	0	0	1	0	0	1	0	0	*0*	1
1	1	0	0	1	0	0	1	0	0	1	0	0	*0*	1
1	1	1	*0*	1	1	*0*	1	1	1	*0*	1	1	1	0
Total error distance			$\sum$ ED = 4			$\sum$ ED = 4			$\sum$ ED = 4			$\sum$ ED = 3		

5 Conclusion

The technical literature presents approximate adder that overcomes the previously proposed ML-based full adder circuit in terms of error performance, device utilization, and circuit complexity. The proposed approximate adder, almost full adder, can be useful for modifying the existing configuration of Dadda multiplier which will consume less device utilization. More than that, with the help of proposed algorithms, we can develop certain image processing applications, such as transform-based image compression, image contrasting, image smoothening, and many more. Hence, we are concluding through this paper that, depending upon the application, we can set the approximation level to maintain error rate as well as power consumption for energy efficient runtime quality configuration.

After all the comparisons, the proposed design consumes 6% less power than accurate full adder of one bit. More than that, it also has 14% less errors compared to ML-based full adder and carry-based approximate full adder. In an era of quality and quantity, we can have such designs like almost full adder which has less power and device utilization compared to accurate adder and less error compared to other approximate adders, such as ML-based full adder.

References

1. Zhang T, Liu W, McLarnon E, O'Neill M, Lombardi F (2018) Design of majority logic (ML) based approximate full adders. In: 2018 IEEE international symposium on circuits and systems (ISCAS). IEEE, pp 1–5
2. Liang J, Han J, Lombardi F (2011) On the reliable performance of sequential adders for soft computing. In: 2011 IEEE international symposium on defect and fault tolerance in VLSI and nanotechnology systems. IEEE, pp 3–10
3. Fontanari T, Paim G, Rocha LMG, Santana GM, da Costa EAC, Bampi S (2019) A fast monolithic 8-2 adder compressor circuit. J Integr Circuits Syst 14(1):1–7
4. Alam I, Lau KT (2017) Approximate adder for low-power computations. Int J Electron Lett 5(2):158–165
5. Jiang H, Liu C, Maheshwari N, Lombardi F, Han J. A comparative study of approximate adders and multipliers. A tutorial
6. Esposito D, Strollo AGM, Napoli E, De Caro D, Petra N (2018) Approximate multipliers based on new approximate compressors. IEEE Trans Circuits Syst I Reg Pap 65(12):4169–4182
7. Jiang H, Liu C, Lombardi F, Han J (2018) Low-power approximate unsigned multipliers with configurable error recovery. IEEE Trans Circuits Syst I Reg Pap 66(1):189–202
8. Porto R, Agostini L, Zatt B, Roma N, Porto M (2019) Power-efficient approximate SAD architecture with LOA imprecise adders. In: 2019 IEEE 10th Latin American symposium on circuits & systems (LASCAS). IEEE, pp 65–68
9. Jiang H, Santiago FJH, Ansari MS, Liu L, Cockburn BF, Lombardi F, Han J (2019) Characterizing approximate adders and multipliers optimized under different design constraints. In: Proceedings of the 2019 on Great Lakes symposium on VLSI, pp 393–398
10. Hasini PNVK, Krishna Murthy T (2015) A novel high-speed transistorized 8×8 multiplier using 4-2 compressors. Int J Eng Res Gen Sci 3(2) Part 2:359–365
11. Potluri US, Madanayake A, Cintra RJ, Bayer FM, Kulasekera S, Edirisuriya A (2014) Improved 8-point approximate DCT for image and video compression requiring only 14 additions. IEEE Trans Circuits Syst I Reg Pap 61(6):1727–1740
12. Almurib HAF, Kumar TN, Lombardi F (2017) Approximate DCT image compression using inexact computing. IEEE Trans Comput 67(2):149–159
13. Ramasamy M, Narmadha G, Deivasigamani S (2019) Carry based approximate full adder for low power approximate computing. In: 2019 7th international conference on smart computing & communications (ICSCC). IEEE, pp 1–4
14. Rezaalipour M, Tajasob S, Dehyadegari M, Bojnordi MN (2018) DrAx: an automatic approach to designing more precise and energy-efficient approximate adders. In: 2018 real-time and embedded systems and technologies (RTEST). IEEE, pp 8–15

Degradation of Optical Communication System Performance Due to SPM

Dhruv Aditya Mittal, Priyamvada, and Kuldeep Singh

Abstract In this paper, an optical transmission link simulation result has been presented for different optical source powers. The impact of source power on SPM has been investigated. Optical power has been obtained for an optical link for 2 spans of 100 km. The power variation is from 5 to 17.5 mW. It has been observed that BER increases considerably if the power is increased beyond 16 mW. However, no appreciable effect has been observed for optical powers below 15 mW. Further, optical power spectrum at input and output has been obtained with various types of filters. Eye diagrams at different values of optical powers have also been observed which signify that the optical power at the source plays an important role in the performance of optical transmission links.

Keywords Bit error rate (BER) · Cross-phase modulation (XPM) · Dispersion · Dense wavelength-division multiplexed (DWDM) · Fiber nonlinearities · Self-phase modulation (SPM)

1 Introduction

The current high time, in the global aspects, makes WDM a more attractive solution for speedy data communication, miniaturization in system size, network compatibility, and high efficiency for spectrum coverage [1–4]. Commercially available systems developed and installed at the data rate of 10 Gb/s are very common nowadays. However, system beyond this up to the range of Tb/s has also been modeled and designed. The research and development down the line last 4–5 decades in the field of semiconductor laser improves the coupling power at the fiber link, so that today's big data rate optical system can be designed can transmit enough power in

D. A. Mittal (✉) · Priyamvada
BML Munjal University, Gurugram, Haryana, India
e-mail: dhruvadityamittal@gmail.com

K. Singh
Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_6

signal to detect it error free at the goals [5–10]. The signal generated by the semiconductor laser is transmitted through optical fiber. The refractive index of the fiber is a dependent function of signal intensity in the transmission link [11–15]. Because of this dependence, the signal transmission is limited by fiber nonlinearities self-phase modulation (SPM) and Cross-phase modulation (XPM). SPM is caused by the intensity modulation of target channel, and XPM is caused by the intensity modulation of parallel propagating channel. After crossing the reference point of input power and 10 Gb/s of data speed, SPM starts dominating the transmission because of nonlinear function of refractive index [16–19].

Here, in this paper an optical communication link is modeled and simulated to see the degradation in the performance of transmission system by varying the impact of power on self-phase modulation using simulation package OptSim.

2 System Modeling and Description

A constructive model designed and simulated to illustrate the characteristics for optical system for SPM effects with variable power in an optical transmission link is shown in Fig. 1. The number of span is considered two, and each span is of 100 km in length. The system model described with an external intensity modulated source of 10 Gb/s having 25 samples per bit, 0.25 dB/km fiber loss. A random data source of $2^9 - 1$ length has been pushed in for the NRZ rectangular pulse driver which has a voltage swing between −2.5 and +2.5 V references. Dispersion of fiber is taken as 0.4 ps/nm/km, and power of optical source varies from 5 to 17.5 mW. The NRZ signal is modulated by Sin^2MZ modulator of maximum transmittivity offset value of 2.5 V, and an extinction ratio keeps ideal position with chirp factor calibrated at zero. The dispersion co-relation length is taken as 10 km; nonlinear refractive index of fiber is 3×10^{-20} m^2/W. The fiber effective core area is 67.57×10^{-12} m^2; fiber nonlinearity coefficient is 1.8 (W km)$^{-1}$. Fiber average beat length is 5 m at standard deviation of 0.5 m. Fiber polarization mode dispersion (PMD) is 0.1 ps/km$^{0.5}$. Rayleigh scattering and fiber reflectivity are considered as ideal. To

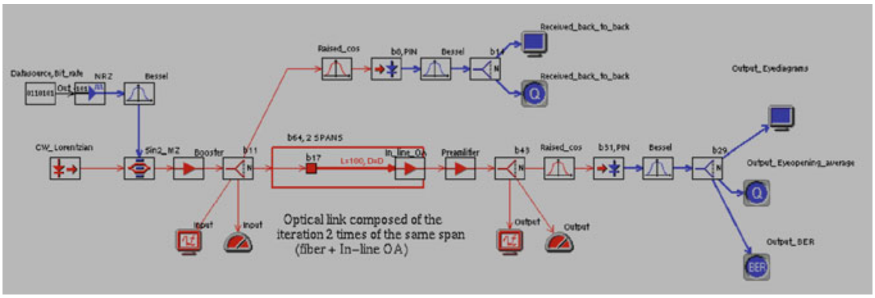


Fig. 1 Circuit arrangements for simulation of self-phase modulation

maintain the smooth transition for Fourier transform between frequency and time domains, the shape of data is considered super Gaussian I nature. An input source with continuous wave (CW) laser having adjustable power with modeling parameters of laser phase noise, line width has been considered in random and ideal nature, respectively. The calibration of a low-pass Bessel filter in transmitter block improves the stability and quality of system with numbers of poles 5 and a frequency spectrum of 10 GHz. To capture the signal at the receiver block, at input stage a band-pass filter with super-Gaussian distribution of order 1 having frequency spectrum range of 60 GHz has been installed. PIN diode to recover original signal characterized with quantum efficiency and dark current is used. To obtain and display the results for bit error rate values, average eye opening, *Q*-factor, a BER estimator and *Q*-meter are used at the end.

3 Results and Discussion

Bit error rate is plotted for various laser power sources varying from 5 to 17.5 mW for fiber lengths of 100 and 160 km in Fig. 2. It reveals that BER increases with increase in optical power at the source. In Fig. 3, optical power at the output is plotted against the power of the source and it has been found that the optical power received at the output depletes due to self-phase modulation. In Fig. 4, optical power spectrum at various optical powers has been obtained. A comparison is made between the optical spectrum for input power 17.5 mW at the input and output of the fiber by using Ideal, Lorentzian, and Rectangular filters in Fig. 5. It has been observed that the pulse gets broaden due to the impact of SPM. Moreover, it has been observed from the simulation that its power also gets decreased at the output as the pulse traverses through the fiber. Figure 6 gives the eye diagrams at the output for fiber length of

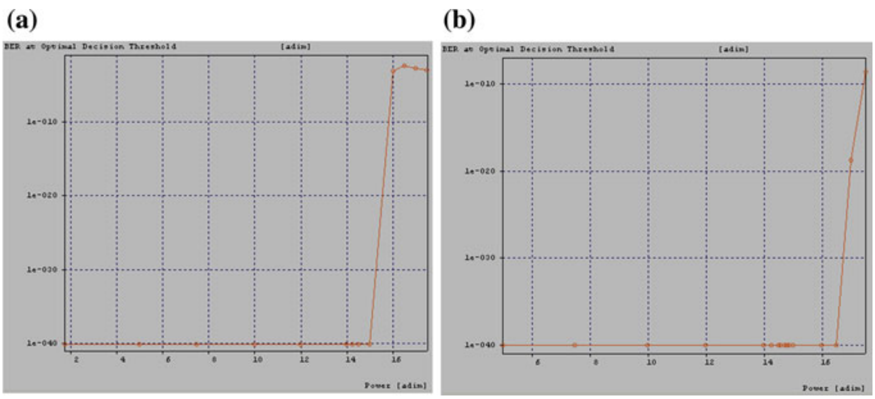
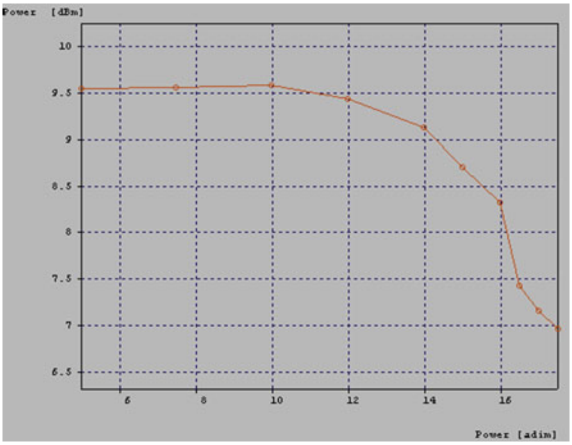


Fig. 2 **a** BER at optimal threshold at output for fiber length 100 km. **b** BER at optimal threshold at output for fiber length 160 km

Fig. 3 Optical power at output for fiber length 100 km



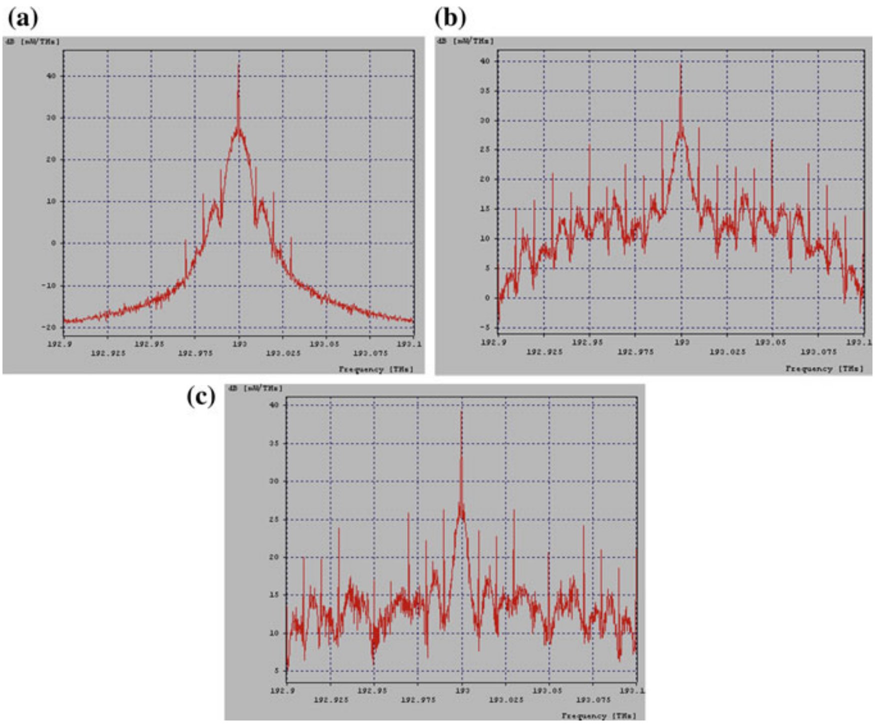


Fig. 4 **a** Optical power spectrum at 5 mW. **b** Optical power spectrum at 16 mW. **c** Optical power spectrum at 17.5 mW

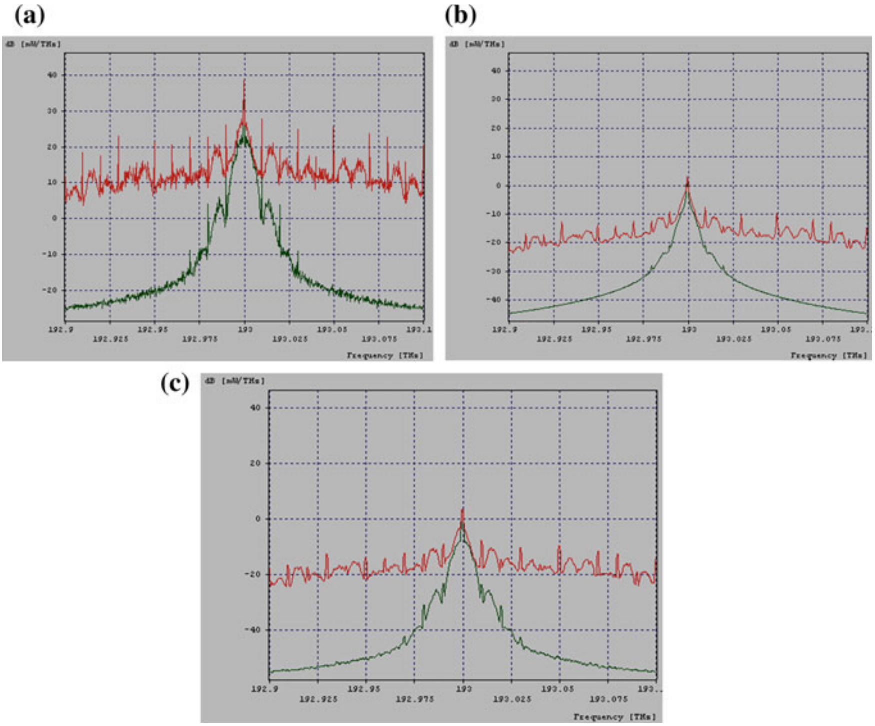


Fig. 5 **a** Optical power spectrum at 17.5 mW at output and input with ideal filter. **b** Optical power spectrum at 17.5 mW at output and input with Lorentzian filter of resolution 0.001. **c** Optical power spectrum at 17.5 mW at output and input with rectangular filter of resolution 0.001

100 km for various values of input optical power. From the figures, it depicts that the eye gets closed with the increase in optical power of the source.

4 Conclusion

BER at various values of optical power has been obtained for an optical link of length 200 km. The power variation is from 5 to 17.5 mW, and it has been observed that BER increases considerably if the power is increased beyond 16 mW. However, no appreciable effect has been observed for optical powers below 15 mW. Moreover, optical power spectrum at input and output has been obtained with various types of filters. Eye diagrams at various values of optical powers were observed which signify that with the increase in optical power the SPM plays a dominant role in degrading the overall performance of optical transmission system.

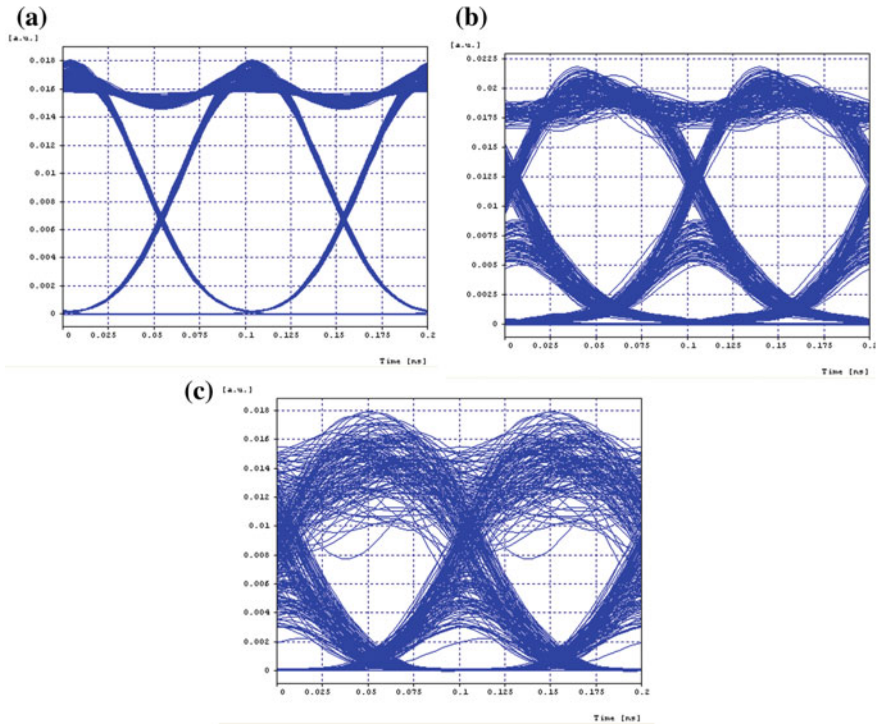


Fig. 6 **a** Eye diagram at 5 mW power for fiber length 100 km. **b** Eye diagram at 16 mW power for fiber. **c** Eye diagram at 17.5 mW power for fiber length 100 km

References

1. Olsson B, Olsson DJ (2002) Pulse restoration by filtering of self-phase modulation broadened optical spectrum. JLT 20(7)
2. Amemiya M (2002) Pulse broadening due to higher order dispersion and its transmission limit. JLT 20(4)
3. Kaler RS, Sharma AK, Kamal TS (2002) Simulation results for DWDM systems with ultra high capacity. Fiber Integr Opt 21(5):361–369
4. Singh K, Arya SK (2017) A novel method to compensate fiber nonlinearities by using NAFO model. J Opt Commun. https://doi.org/10.1515/joc-2017-0016
5. Agrawal GP (1995) Nonlinear fiber optics. Academic Press, New York
6. Sharma D, Prajapati YK (2016) Performance analysis of DWDM system for different modulation schemes using variations in channel spacing. J Opt Commun. https://doi.org/10.1515/joc-2016-0011
7. Konyshev VA, Leonov AV, Nanii OE, Treshchikov VN, Ubaydullev RR (2015) New method to obtain optimum performance for 100 Gb/s multi-span fiber optic lines. Opt Commun 355(2015):279–284. https://doi.org/10.1016/j.optcom.2015.06.048
8. Niu J-N, Sun Y-M, Cai C, Ji Y-F (2018) Optimized channel allocation scheme for jointly reducing four-wave mixing and Raman scattering in the DWDM-QKD system. Appl Opt 57(27). 1559-128X/18/277987-10. https://doi.org/10.1364/AO.57.007987

9. Singh K, Arya SK (2018) Estimation of signal-to-cross talk ratio of stimulated-Raman-scattering-induced cross talk in wavelength-division-multiplexing-based radio-over-fiber links. J Opt Commun. https://doi.org/10.1515/joc-2018-0039
10. Hicdurmaz B, Temurtas H, Karlık SE, Gunes Y (2013) A novel method degrading the combined effect of FWM and ASE noise in WDM systems containing in-line optical amplifiers. Optik 124:4064–4071. https://doi.org/10.1016/j.ijleo.2012.12.071
11. Jiao B (2015) Leveraging ultrascale architecture transceivers for high-speed serial I/O connectivity. White paper: ultrascale GTH/GTY transceivers. WP458 (v2.0), 29 Oct 2015
12. Hu H, Jopson RM, Gnauck AH, Randel S, Chandrasekhar S (2017) Fiber nonlinearity mitigation of WDM-PDM QPSK/16-QAM signals using fiber-optic parametric amplifiers based multiple optical phase conjugations. Opt Express 25(3):1619. https://doi.org/10.1364/OE.25.0016
13. Agrell E et al (2016) Roadmap of optical communications. J Opt 18:063002 (40 pp). https://doi.org/10.1088/2040-8978/18/6/063002
14. Kaur G, Singh ML, Patterh MS (2010) Impact of fibre nonlinearities in optical DWDM transmission systems at different data rates. Optik 121:2166–2171. https://doi.org/10.1016/j.ijleo.2009.11.001
15. Morsy E, Fayed HA, Aziz AAE, Aly MH (2018) SPM and XPM crosstalk in WDM systems with DRA: channel spacing and attenuation effects. Opt Commun 417:79–82. https://doi.org/10.1016/j.optcom.2018.02.028
16. Lawan SH, Mohammad AB (2018) Reduction of four wave mixing efficiency in DWDM systems using optimal PMD. Opt Quant Electron 50:91. https://doi.org/10.1007/s11082-018-1370-y
17. Manzoor HU, Manzoor T, Hussain A, Aly MH, Manzoor S (2019) FWM reduction using different modulation techniques and optical filters in DWDM optical communication systems: a comparative study. Iran J Sci Technol Trans Electr Eng. https://doi.org/10.1007/s40998-019-00189-4
18. Sharan L, Shanbhag AG, Chaubey VK (2016) Design and simulation of modified duobinary modulated 40 Gbps 32 channel DWDM optical link for improved non-linear performance. Cogent Eng 3:1256562. https://doi.org/10.1080/23311916.2016.1256562
19. Schidlack E (2017) Automatic insertion of debug logic for transceivers in synthesis DCP. Application note: Vivado HSSIO debug. XAPP1295 (v1.0), 19 Sept 2017

Real-Time Measurement and Analysis of Power Quality Issues in Distribution System

Sai Mounika Ramireddi, Narasimha Rao Kolipaka, Santosh Kumar Nagalla, and C. Venkata Krishna Bhanu

Abstract There is a growing concern for power quality with the increase in use of power electronic-based equipment in all sectors of electricity consumption. Besides causing maloperation/degradation of equipment, power quality problems are causing huge economic losses. A lot of work has been carried out on the impact of poor power quality on the equipment and the consequent financial implications. This paper presents analysis on extensive field data collected from various locations of LT distribution system in Visakhapatnam, Andhra Pradesh. Monitoring of power quality has been performed at PCC or secondary side of the transformer by using YOKOGAWA CW240 Clamp-on Power Meter. No power quality data is available in the literature on Indian LT distribution network. A comparative study on the harmonic characteristics of residential, industrial, agricultural and commercial feeders is also conducted. So, here an attempt is made to measure and analyze the power quality issues under different categories.

Keywords Power quality · Harmonics

1 Introduction

Due to increase in use of nonlinear loads, the effect of harmonics increased. The nonlinear equipment such as fluorescent lamps, computers, air conditioners, welding machines and mobile chargers used in residential, industrial and commercial loads causes distortions in voltage and current. These harmonics cause reverse effects in the various power quality surveys have been carried out to analyze all these issues [1]. It provides how the harmonic voltages and currents are presented at sending end of feeders, i.e., distribution substations [2]. It analyzes the comparison of harmonic

S. M. Ramireddi (✉) · N. R. Kolipaka · S. K. Nagalla · C. Venkata Krishna Bhanu
Department of Electrical and Electronics Engineering, GVP College of Engineering (A), Visakhapatnam, India
e-mail: saimounikareddy996@gmail.com

C. Venkata Krishna Bhanu
e-mail: bhanucvk@gvpce.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_7

characteristics of residential, industrial and commercial loads by conducting the extensive field survey [3]. They have monitored harmonic distortions present in the residential area LV network [4]. They provide existing level of harmonic characteristics of residential and commercial loads [5]. It presents the influence of harmonics on MV distribution system.

An attempt was made in this work to measure and identify power quality issue under different categories. Here extensive field survey has been carried out from residential, industrial, commercial and agricultural loads which are presented in LT distribution system to know the existing level of harmonic pollution presented in the distribution system. This paper provides continued research in the field of power system harmonics.

This paper is organized as follows. Section 2 describes field measurement activities. Section 3 provides comparative analysis of power quality parameters for multiple loads at different locations of a power system. Section 4 gives conclusions.

2 Field Measurements

2.1 Features of Power Quality Measurement Locations and Duration

Power quality survey was carried out in Visakhapatnam, Andhra Pradesh. Here totally 5 residential, 4 industrial, 2 agricultural and 1 commercial loads and individual equipment were collected. The measurement locations are at service panels and secondary side of the service transformers. The measurements were taken from each load for the duration of 24 h but 8 h for agricultural loads because the supply has given to these loads for 8 h.

2.2 Instrument and Software Used to Analyze the Data

The voltage and current measurements were taken at LV side of the distribution systems. The data was collected by using YOKOGAWA CW240 power quality analyzer. Measuring the power quality parameters, Clamp on Power Meter—YOKOGAWA CW240 is used. The instrument is capable of measuring three-phase voltages and currents, frequency and power factor and also can carry out power quality analysis. The power quality parameters like voltage sag/swell, interruptions and harmonics (up to 50th order) can be measured using meter. It has a measurement voltage range of 150–1000 V and a current range of 200 uA–3000 A with seven different current probes.

Current Setting

The current ranges selected differ depending on the clamp-on current probe. For all the residential and agricultural loads, 50 A current probe is used. For industrial and commercial loads, 500 A current probe is used.

Voltage Setting

Standard voltage (V_{Std})	Voltage sag (setting w.r.t. V_{Std})	Voltage swell (setting w.r.t. V_{Std})	Voltage interruption (setting w.r.t. V_{Std})
240 V (RMS)	90%	110%	10%

2.3 Standards for Power Quality Parameters

In this paper, for all loads total harmonic distortion in voltage is compared with IEEE 519 standards and total harmonic distortion current is compared with IEC standards.

The unbalancing voltage is compared with IEEE 519 standards. Here power is also calculated whether it is good or bad. Different voltage quality events like voltage sag, voltage swell, over voltage, under voltage and interruptions are observed.

3 Analysis of Power Quality Parameters

3.1 Residential Loads

The power quality survey was conducted for 5 residential loads. Among all, 3 loads are completely residential and other 2 are hostel-based loads. Load details are lighting, air conditioners, computers, lift, motors, fans, etc. Out of 5 residential loads, one load consists of more distortions in voltage and current as shown in Figs. 1 and 2.

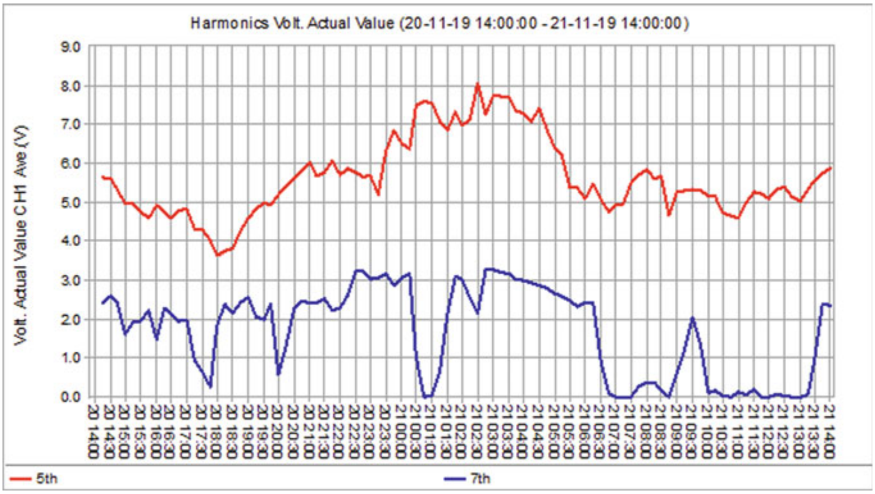


Fig. 1 Voltage harmonic distortions versus time

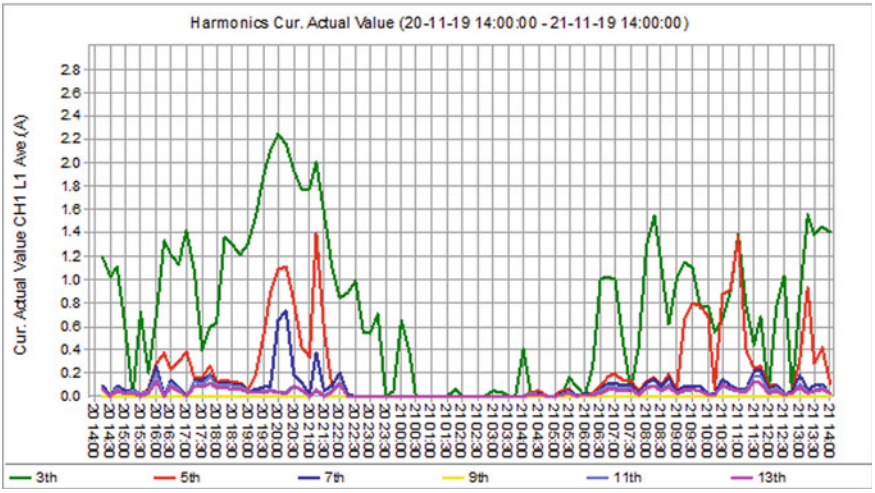


Fig. 2 Current harmonic distortions versus time

In voltage harmonics, 5th and 7th order harmonics presented and 5th order harmonics are dominant one. In current harmonics, up to 13th order harmonics presented in the system and 3rd order harmonics are dominant one. The voltage harmonics are within the limits, and current harmonics slightly violates its limits. Duration of current harmonics which violates its limits is 12 h: 50 min, and these harmonics are more during daytime and less during nighttime due to more usage of power electronic devices during daytime.

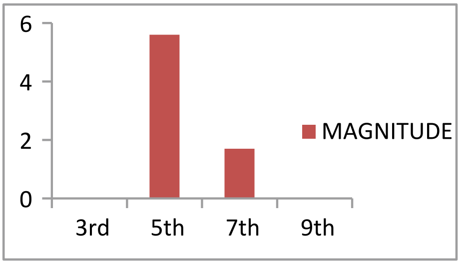


Fig. 3 Voltage harmonic spectrum

The voltage and current harmonic spectrums are shown in Figs. 3 and 4.

The unbalancing ratio and power factor in the system are shown in Figs. 5 and 6.

The power factor is unity for 75% duration of time, and unbalancing ratio in voltage is within the limits.

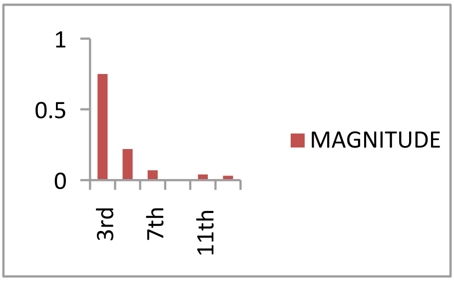


Fig. 4 Current harmonic spectrum

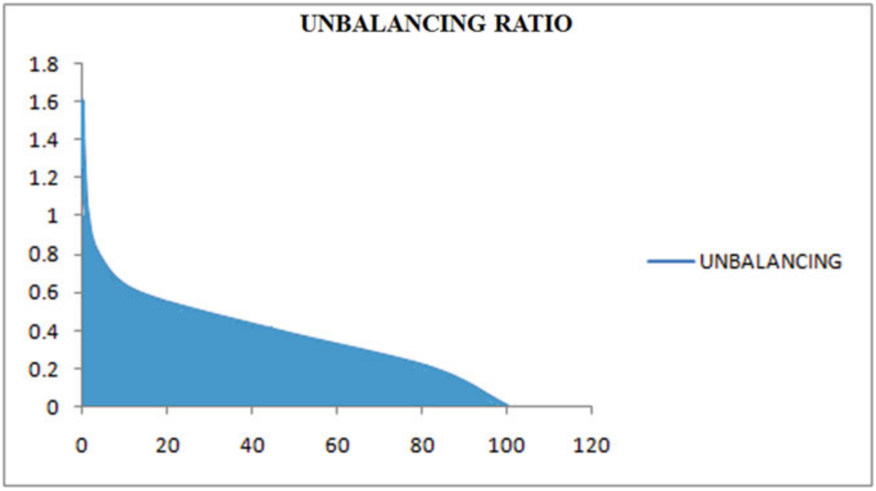


Fig. 5 Unbalancing ratio

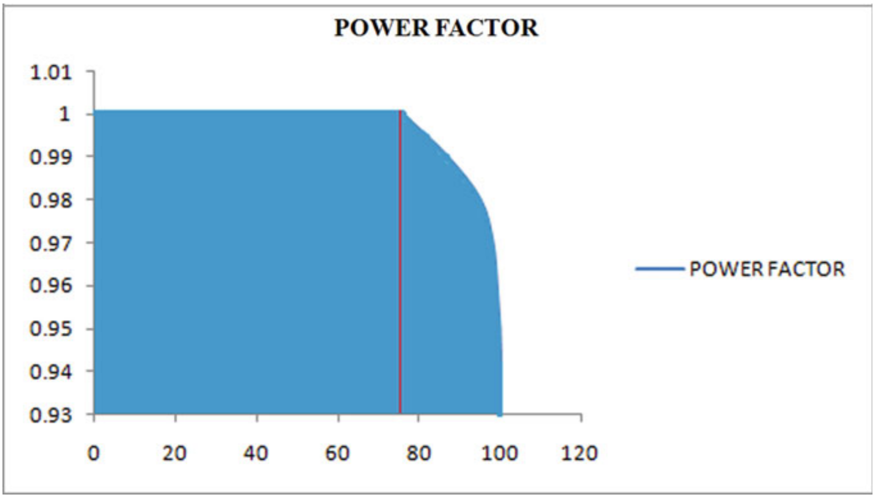


Fig. 6 Power factor versus duration of time

3.2 Industrial Loads

There are 4 industrial loads, named as sea food industry, ice plant, crab factory and steel industries. These all are located in Visakhapatnam. The two industries sea foods and crab factory are food processing units. Ice plant is manufacturing ice. Steel industry is one of the large manufacturing of steel in Visakhapatnam. Out of all these industries, ice plant is having more distortions in voltage and current (Figs. 7 and 8).

In voltage harmonics, 3rd and 7th order harmonics are dominant in voltage and 5th order harmonics are dominant in current. The duration of harmonics violates its limits. These harmonics are continuously presented in the system throughout the day. These harmonics are mainly due to motors.

The voltage and current harmonic spectrums are shown in Figs. 9 and 10.

The unbalancing ratio and power factor in the system are shown in Figs. 11 and 12.

The power factor is unity in the system, and unbalancing ratio in voltage is more and beyond the limits.

3.3 Agricultural Loads

The survey was conducted for 2 agricultural loads. One load is having 4 connections, and other load is having 9 connections. All the loads are motor loads. The voltage and current harmonics in voltage and current are shown in Figs. 13 and 14.

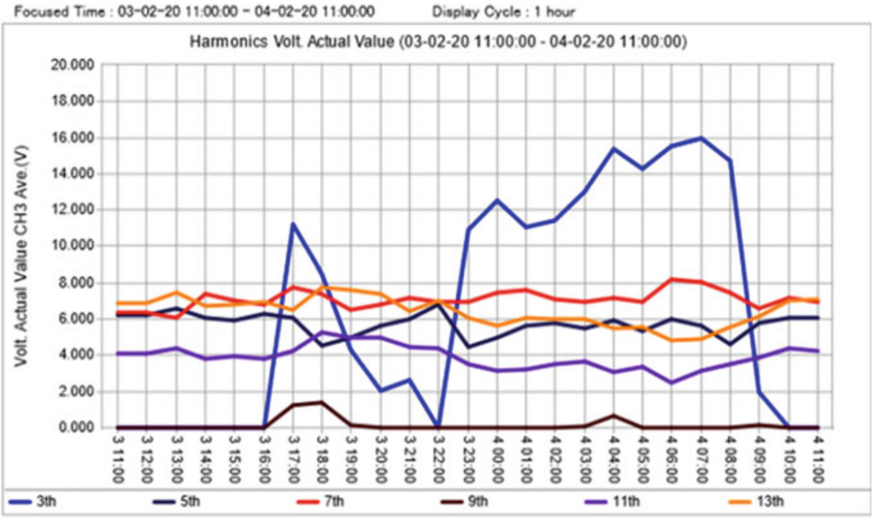


Fig. 7 Voltage harmonic distortions versus time

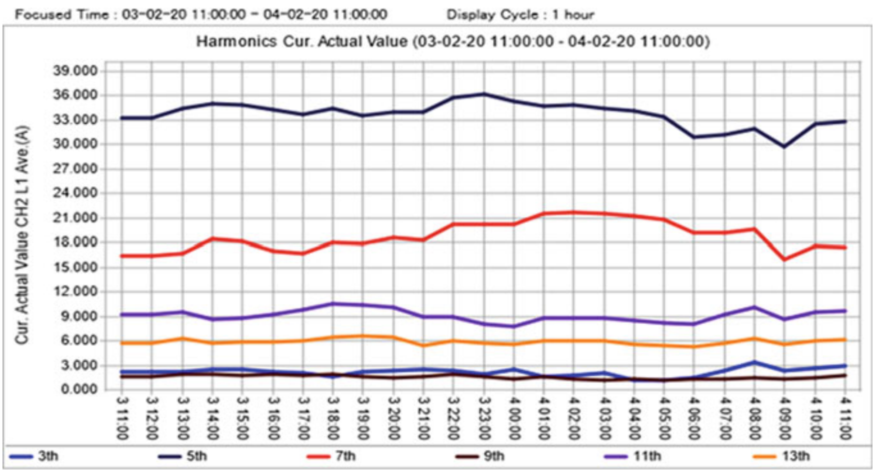


Fig. 8 Current harmonic distortions versus time

In both voltage and current, 5th order harmonics are more. These harmonics are more due motors.

The voltage and current harmonic spectrums are shown in Figs. 15 and 16.

The unbalancing ratio and power factor in the system are shown in Fig. 17.

The power factor is poor in the system, and the unbalancing ratio in the system is within the limits (Fig. 18).

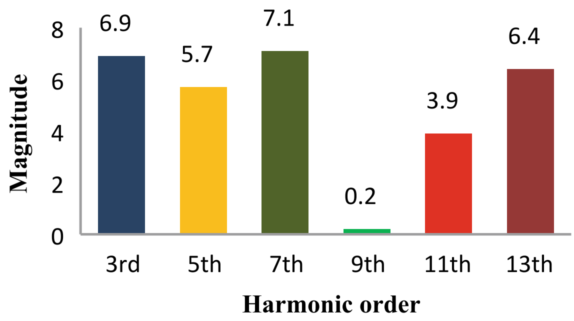


Fig. 9 Voltage harmonic spectrum

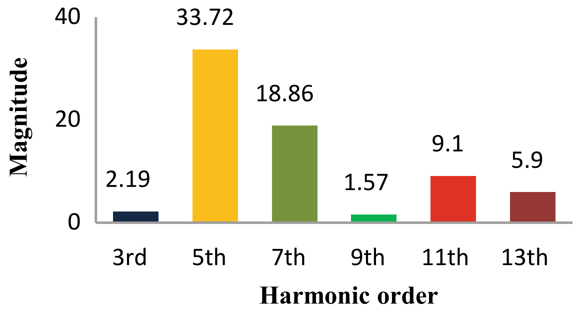


Fig. 10 Current harmonic spectrum

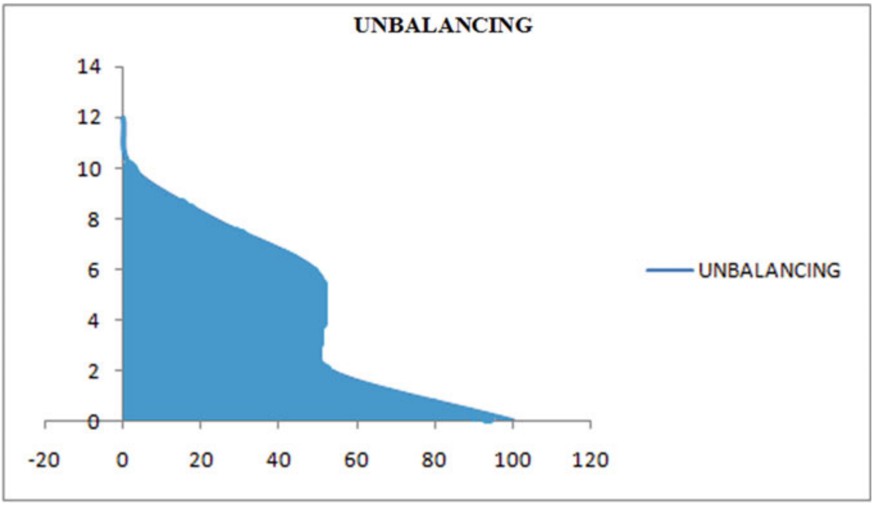


Fig. 11 Unbalancing ratio

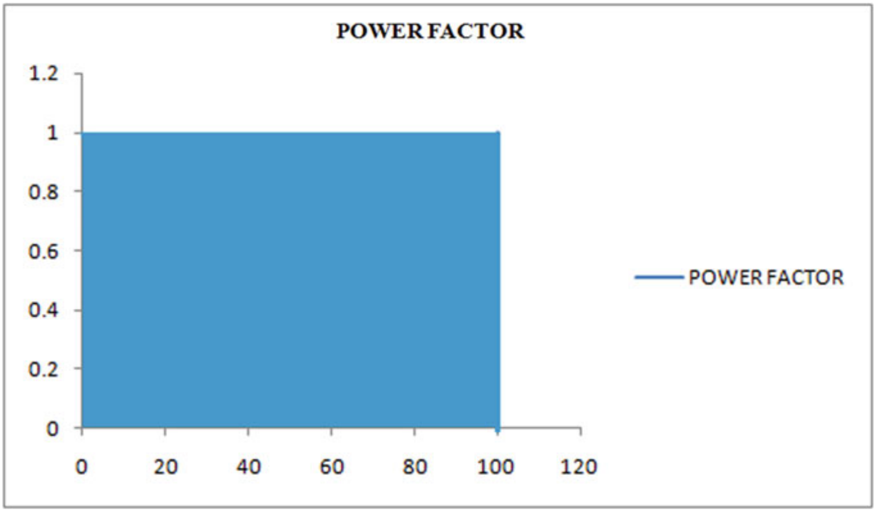


Fig. 12 Power factor versus duration of time

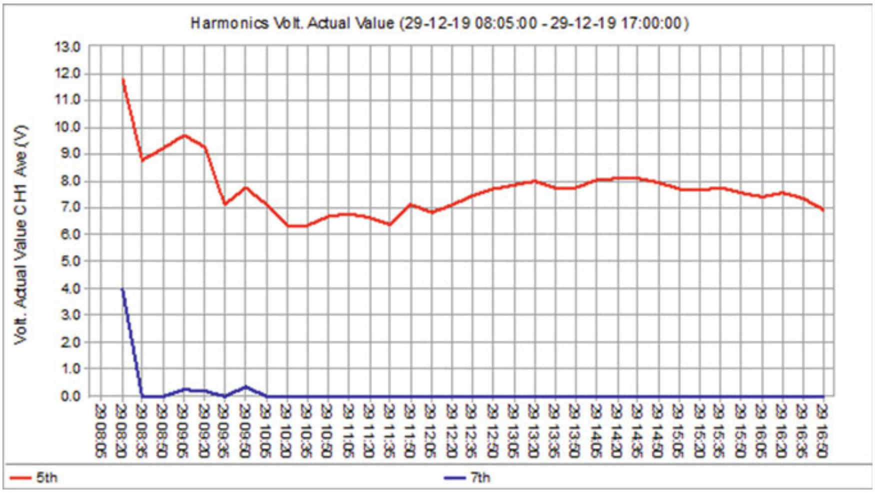


Fig. 13 Voltage harmonic distortions versus time

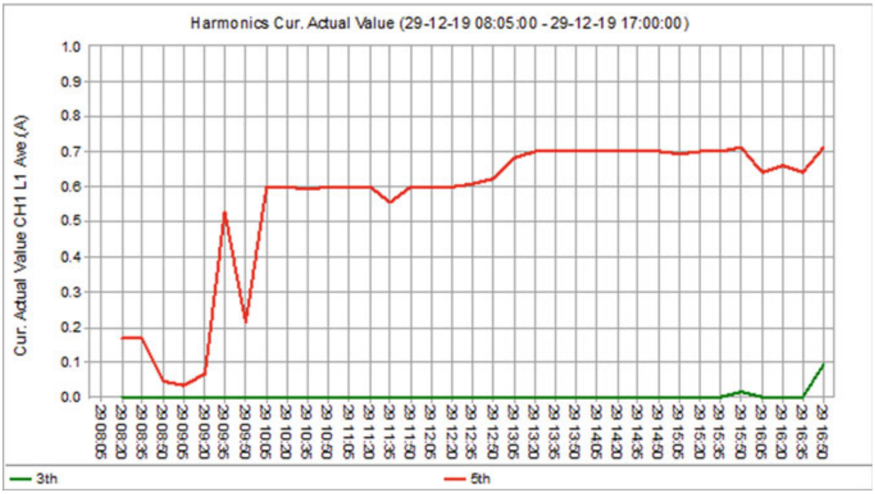


Fig. 14 Current harmonic distortions versus time

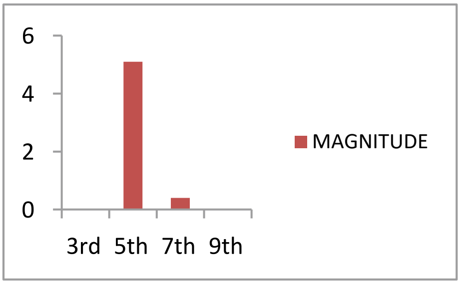


Fig. 15 Voltage harmonic spectrum

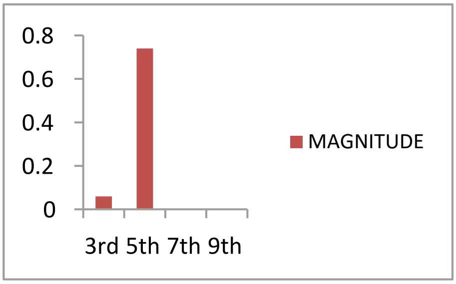


Fig. 16 Current harmonic spectrum

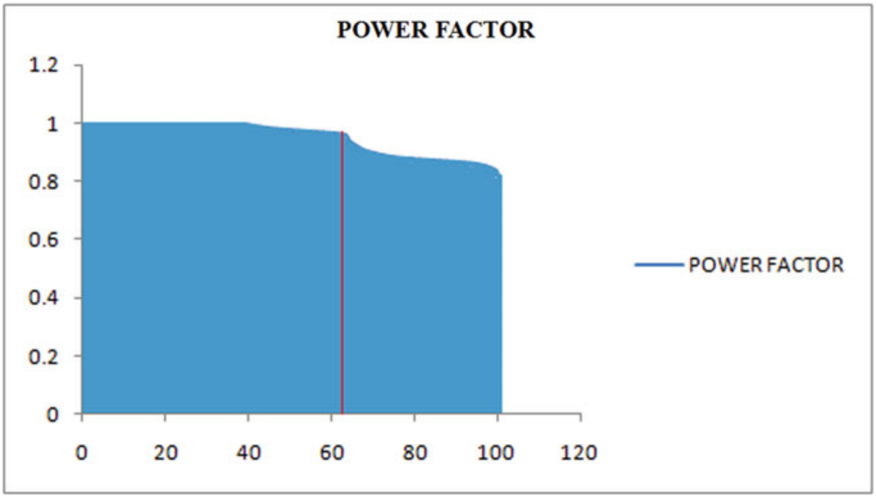


Fig. 17 Power factor versus duration of time

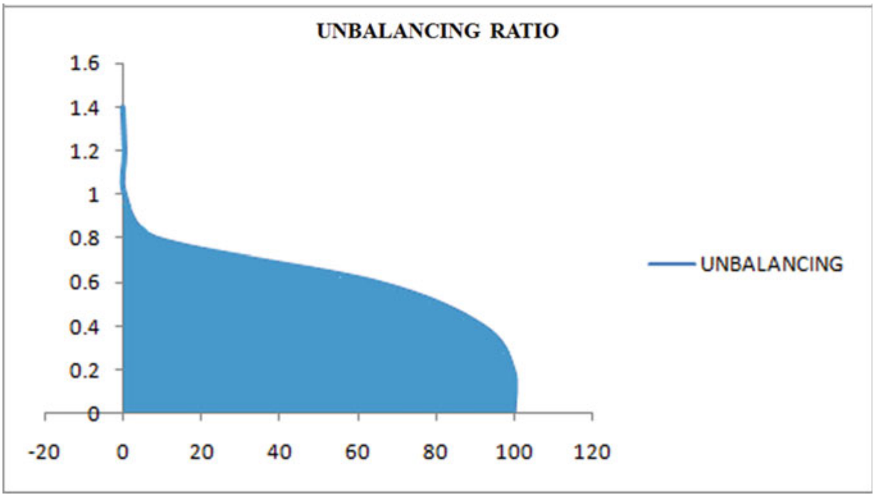


Fig. 18 Unbalancing ratio

3.4 Commercial Loads

The name of the commercial complex which is located in Visakhapatnam. This load consists of lighting load, air conditioners and fans.

The voltage and current harmonics in this load are shown in Figs. 19 and 20.

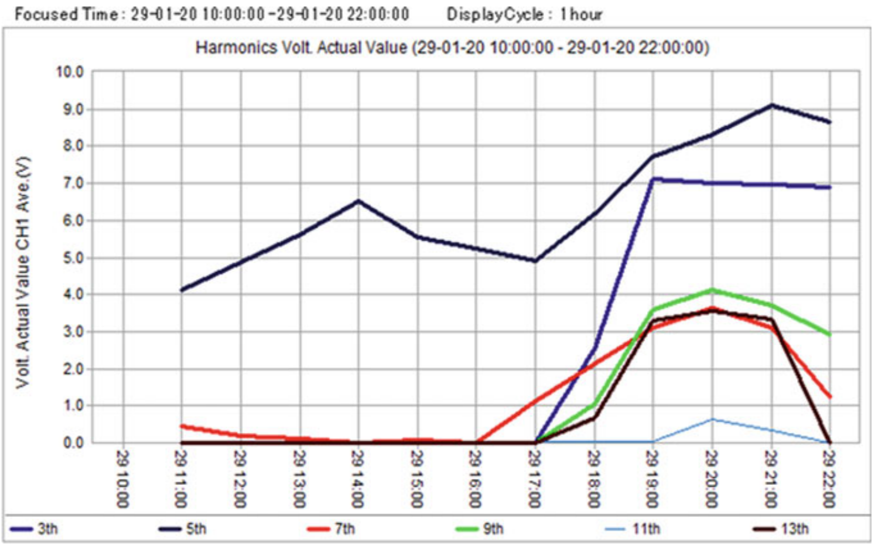


Fig. 19 Voltage harmonic distortions versus time

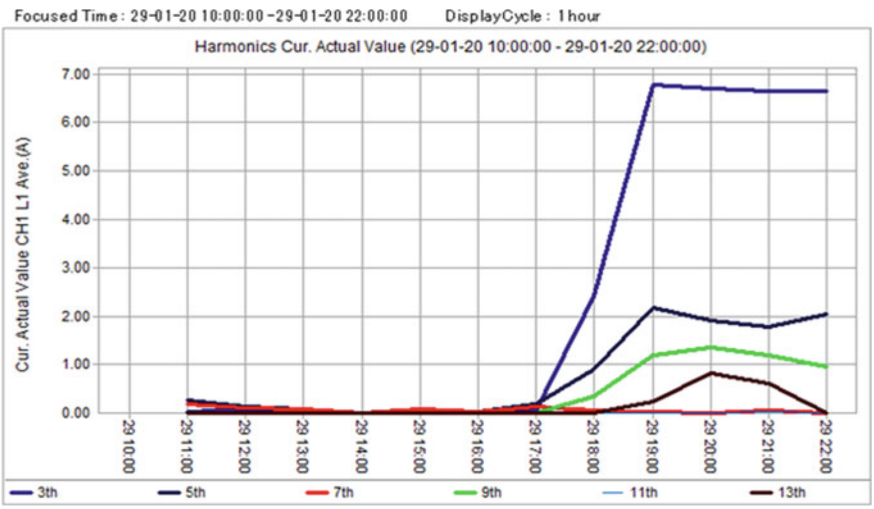


Fig. 20 Current harmonic distortions versus time

From this, 5th order harmonics are more in voltage and 3rd order harmonics are more in current. These harmonics are more during nighttime due to lighting. Here the voltage harmonics are within the limits but current harmonics are highly violating its limits.

The voltage and current harmonic spectrums are shown in Figs. 21 and 22.

From the harmonic analysis data in current harmonics, 3rd order harmonic is dominant one. In commercial load, there are significant current distortions.

The comparison of various power quality parameters are given in Table 1.

The total harmonic distortion in voltage and current for different loads is shown in Figs. 23 and 24.

3.5 *Some of the Individual Equipment*

(a) *One bulb (fluorescent)*

See Figs. 25 and 26.

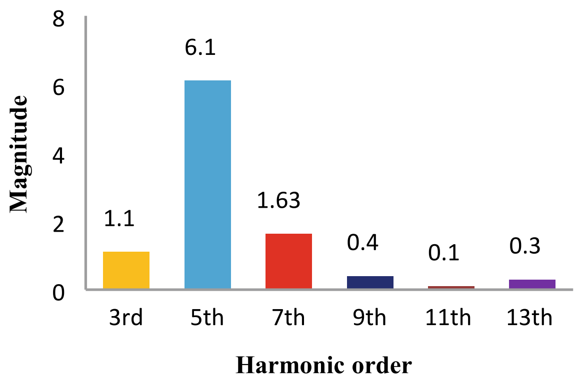


Fig. 21 Voltage harmonic spectrum

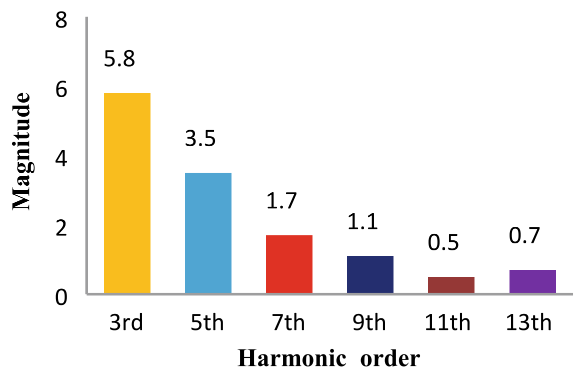


Fig. 22 Current harmonic spectrum

Table 1 Comparison of various power quality parameters

	Interruption	Sag	Swell	Over voltage	Under voltage	% of unbalancing	THD-V	THD-I	PF
AGL-1	0	8	10	0	0	8.35	5.14	6.41	0.77
AGL-2	4	4	9	0	0	1.36	1.58	2.25	0.94
RL-1	0	1	23	20	0	3.75	1.73	6.88	0.99
RL-2	1	1	18	0	0	1.56	1.88	0.51	1
RL-3	2	2	6	10	0	1.31	1.89	9.35	0.98
RL-4	0	1	0	14	0	3.47	2.28	7.06	0.99
RL-5	0	0	5	43	0	1.49	2.5	11.47	1
IND-1	0	0	7	21	0	0.6	5	5.4	0.7
IND-2	3	47	0	0	0	12.29	6.1	47.3	1
IND-3	0	1	16	29	0	0.37	3.7	13	0.9
IND-4	0	0	19	22	0	0.54	0.79	2	0.84
COM-1	0	0	21	11	0	1.78	2.4	60.2	0.72

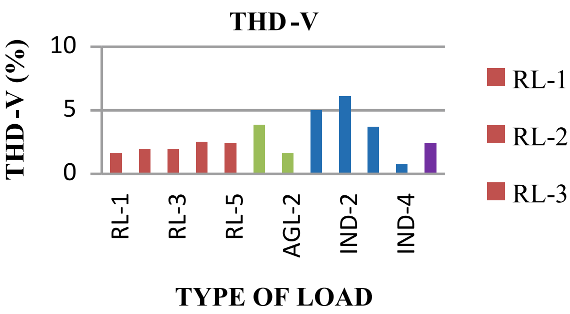


Fig. 23 Total harmonic distortion in voltage

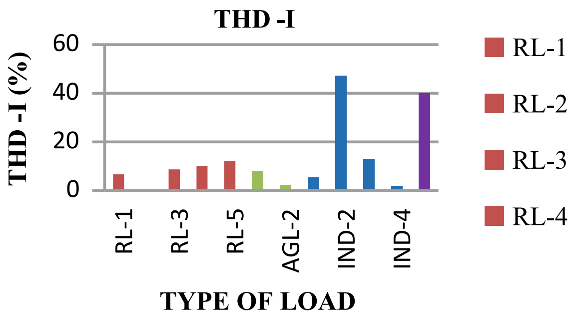


Fig. 24 Total harmonic distortion in current

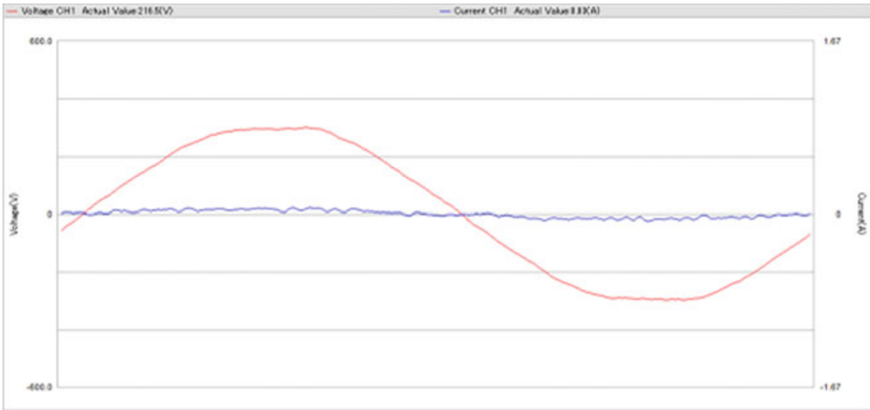


Fig. 25 Voltage and current waveforms

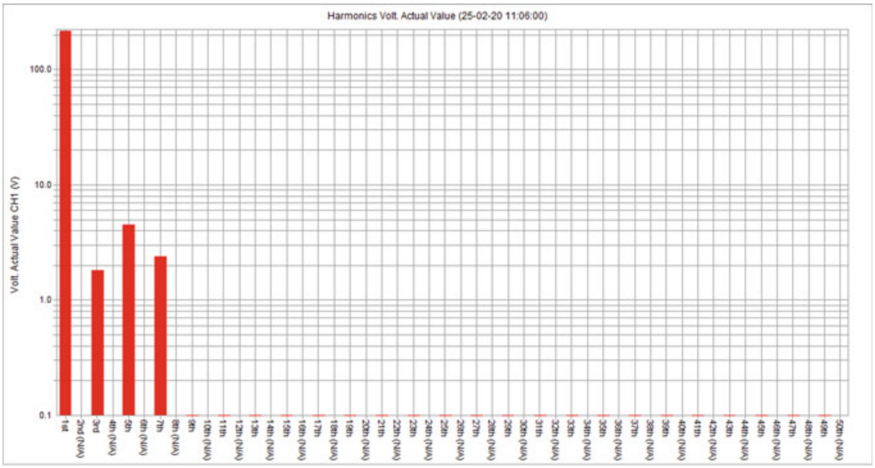


Fig. 26 Voltage harmonic spectrum for 1 bulb

Total harmonic distortion in voltage (%)—2.5
Total harmonic distortion in current (%)—0.

(b) *Two bulbs (fluorescent and LED)*

See Figs. 27 and 28.
Total harmonic distortion in voltage (%)—2.4
Total harmonic distortion in current (%)—0.

(c) *Three bulbs (fluorescent and LED)*

See Figs. 29 and 30.
Total harmonic distortion in voltage (%)—2.5

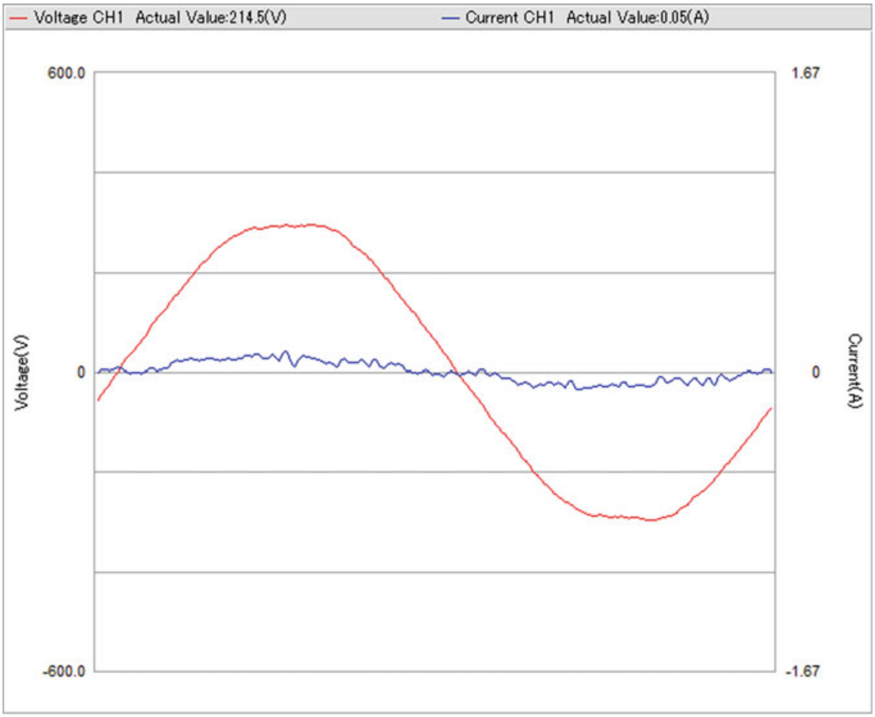


Fig. 27 Voltage and current waveforms

Total harmonic distortion in current (%)—0.

(d) *Exhaust fan*

See Figs. 31 and 32.
Total harmonic distortion in voltage (%)—2.6
Total harmonic distortion in current (%)—0.

(e) *Laptop*

See Figs. 33 and 34.
Total harmonic distortion in voltage (%)—2.4
Total harmonic distortion in current (%)—188.

(f) *Mobile charger*

See Figs. 35 and 36.
Total harmonic distortion in voltage (%)—2.3
Total harmonic distortion in current (%)—132.2 (Table 2).

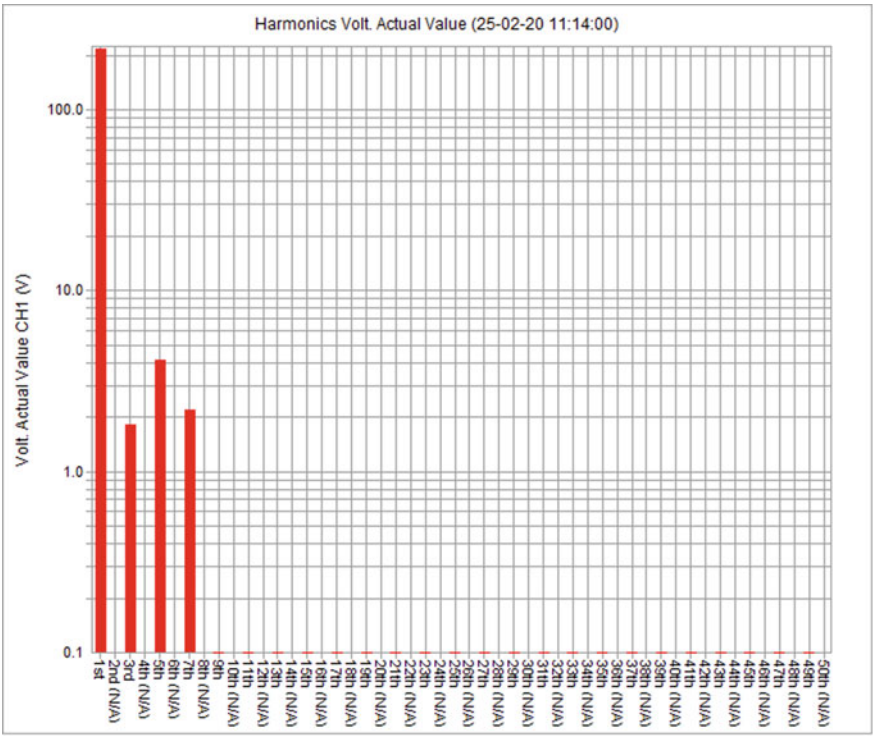


Fig. 28 Voltage harmonic spectrum for 2 bulbs

From all loads, laptop is having more THD in current comparing with other. Welding machine and mobile charger are also violating the limits. Here 3rd order harmonics are more in current.

4 Conclusions

Real-time measurement of power quality parameters in an LT distribution system under different load categories is presented in this work. From the analysis of the data presented, the number of voltage quality events is high for residential loads. Also in residential loads, the THD-I levels are more during daytime than the nighttime. In the case of industrial and commercial loads, the power factor is less than 0.9 which is of serious concern. In the case of agriculture loads, a special case where availability of supply is only for 8 h of duration per day, the harmonics are high. One of the major limitations of this work is the duration of the measurement which is only 24 h and measurement for one week might give a better picture. In the future, it is proposed

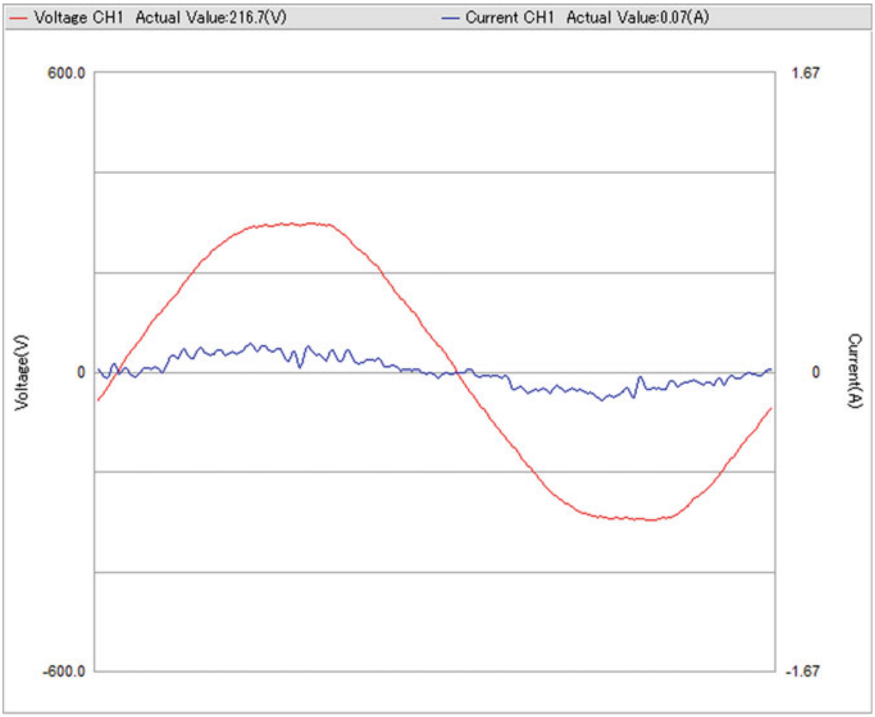


Fig. 29 Voltage and current waveforms

to measure the power quality parameters for at least one week continuously in each category and with more number of the sample under each category.

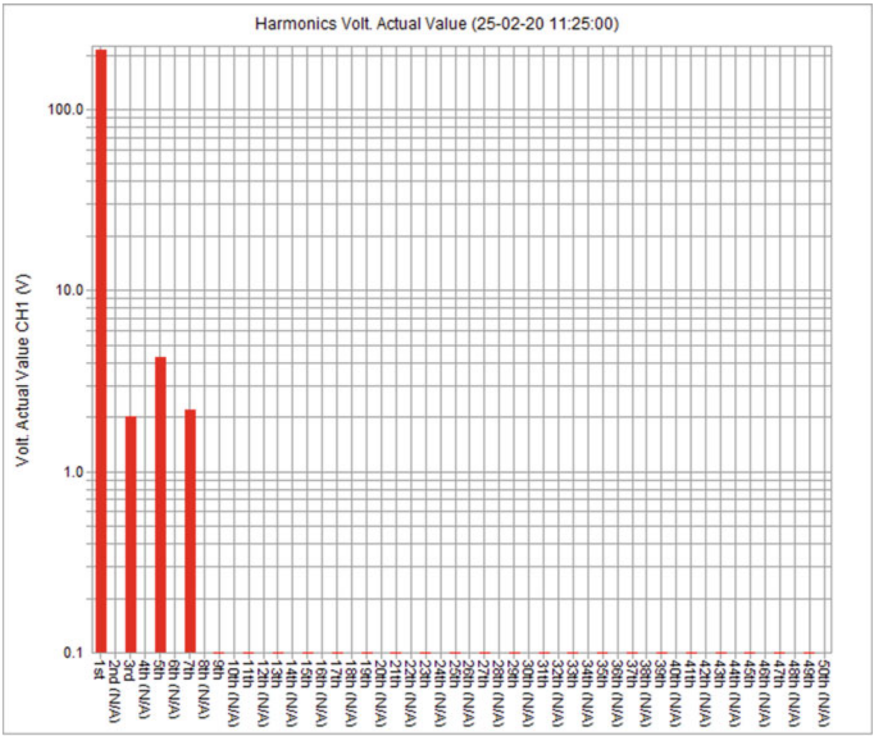


Fig. 30 Voltage harmonic spectrum for 3 bulbs

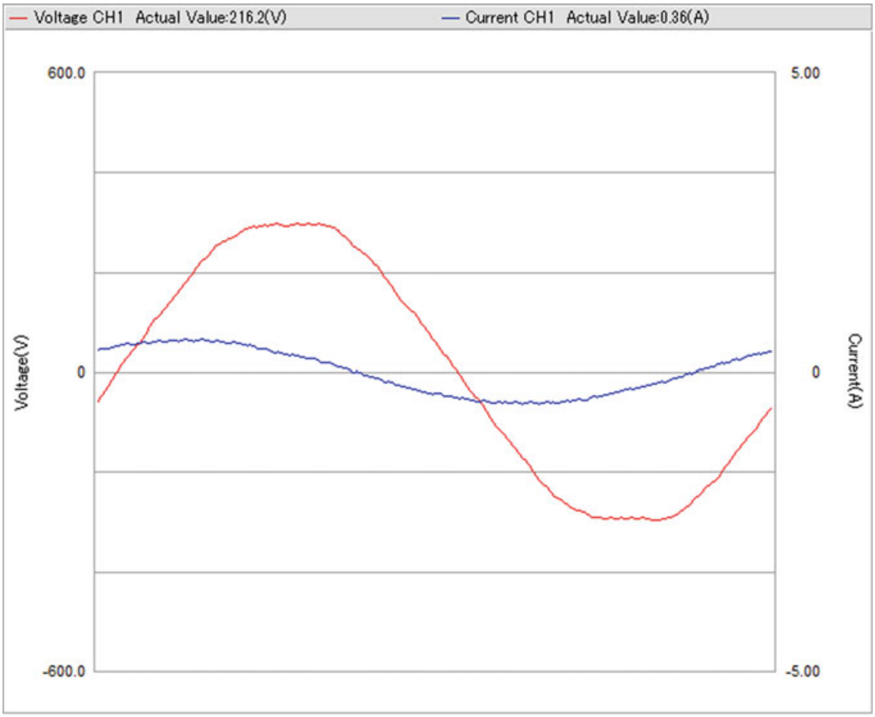


Fig. 31 Voltage and current waveforms

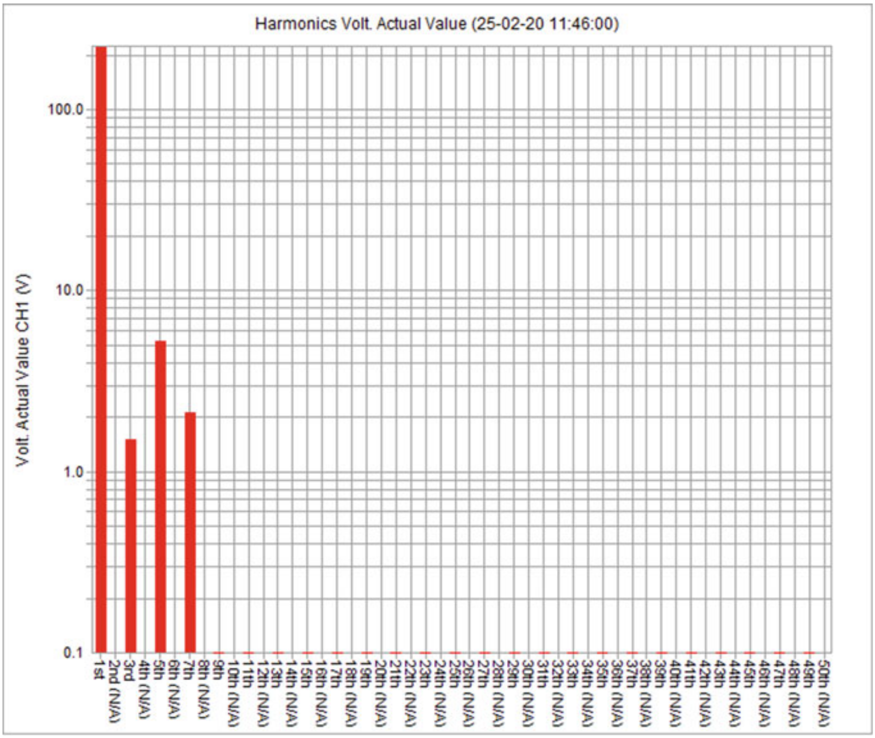


Fig. 32 Voltage harmonic spectrum for exhaust fan

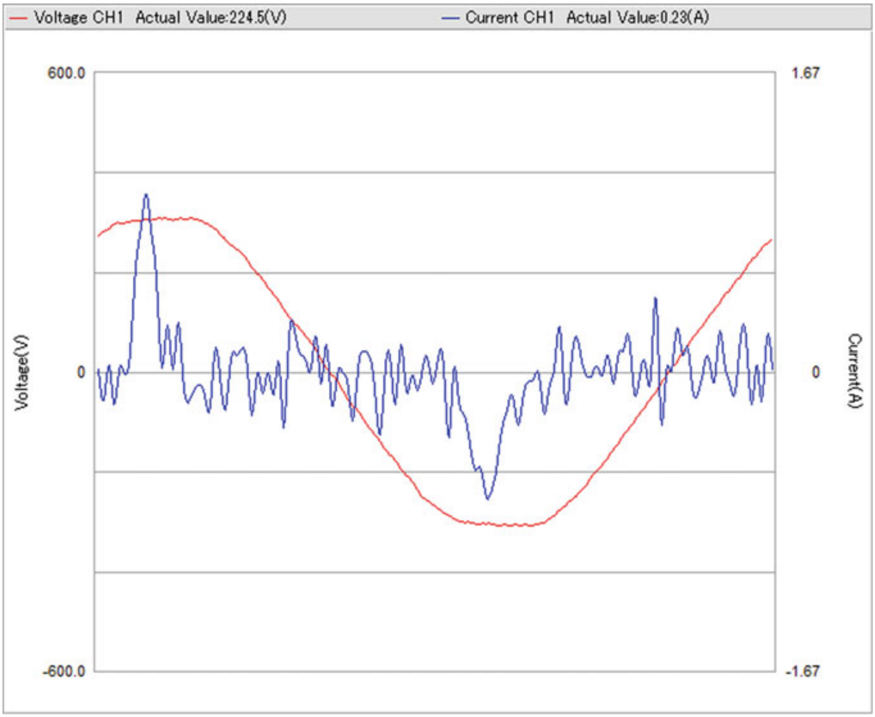


Fig. 33 Voltage and current waveforms

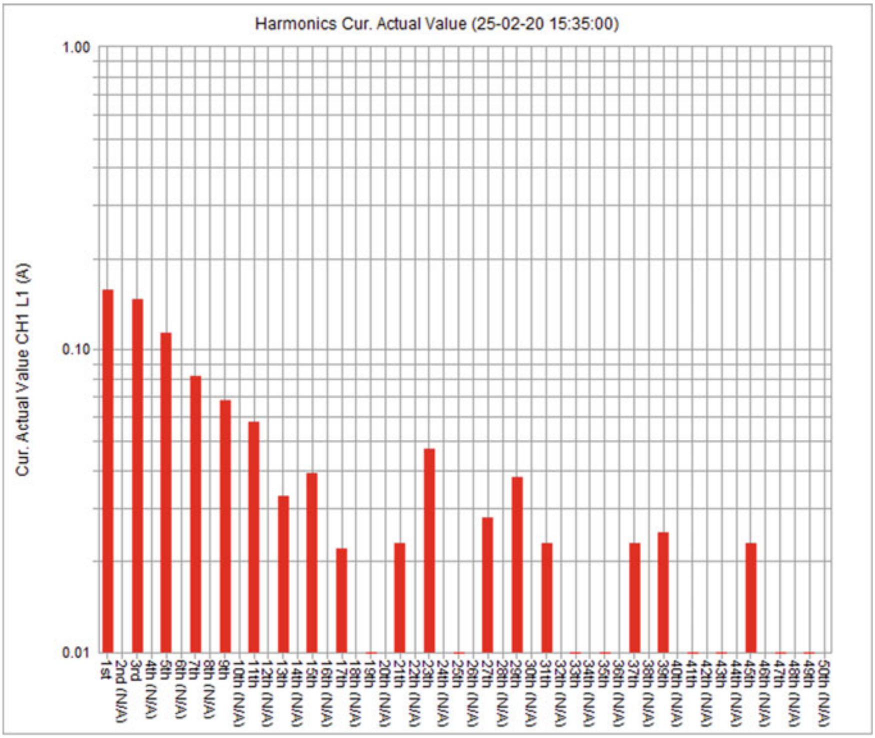


Fig. 34 Voltage harmonic spectrum for laptop

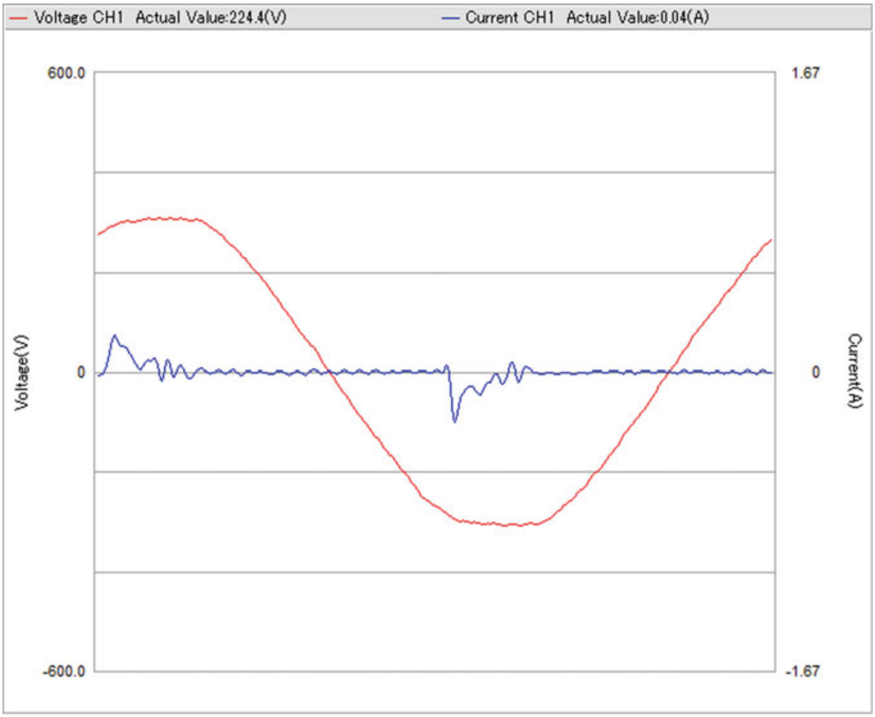


Fig. 35 Voltage and current waveforms

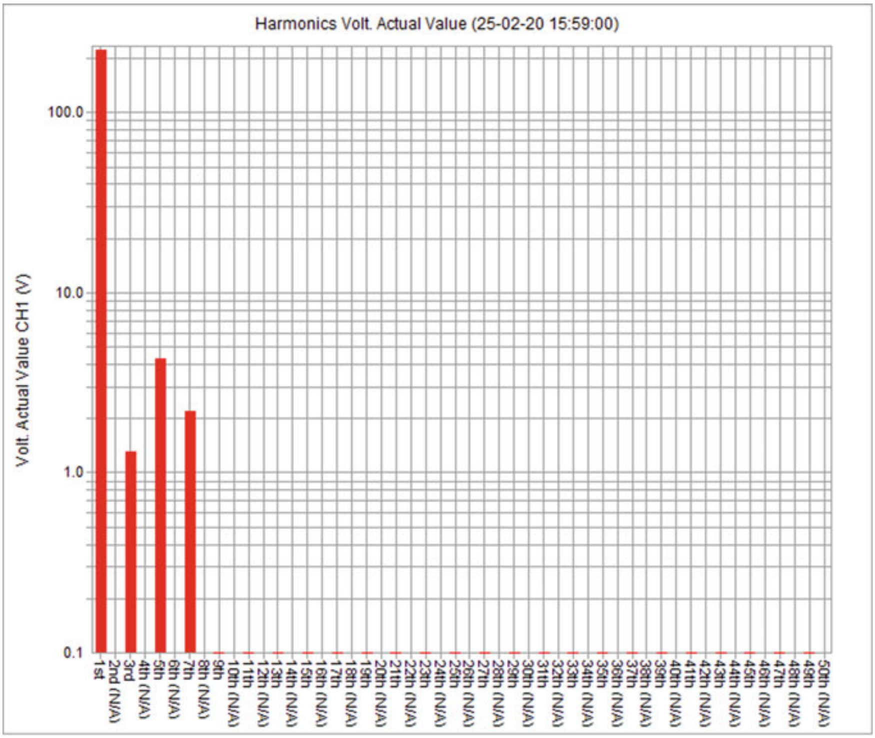


Fig. 36 Voltage harmonic spectrum for mobile charger

Table 2 Power quality parameters for some of the individual equipment

Device	Current THD (%)	Most dominant harmonic order	Other harmonics	PF
Fluorescent lamp	–	–	–	1
Laptop	188	3	5, 7, 9, 11	0.4 (lead)
Mobile charger	132.2	–	–	0.5 (lead)
Welding machine	40.1	3	3, 7, 9, 13	0.41
Woodland (small commercial load)	60.2	3	5, 7, 9, 11, 13	0.72

References

1. Barbulescu C, Cornoiu M, Kilyeni S, Stoian C, Stoian P (2013) Electric power quality issues: harmonic analysis for real network. In: IEEE EUROCON, July 2013
2. Emanuel AE, Orr JA, Cyganski D, Gulachenski EM (1993) A survey of harmonic voltages and currents at the customer bus. IEEE Trans Power Delivery
3. Jain S, Agarwal P, Gupta HO (2004) A survey of harmonics: Indian scenario. In: IEEE India annual conference 2004, Indicon 2004, Dec 2004, pp 84–89

4. Wang Y, Yong J, Sun Y, Xu W, Wong D (2017) Characteristics of harmonic distortions in residential distribution systems. IEEE Trans Power Delivery 32(3):1495–1504
5. Radmehr M, Farhangi S, Nasiri A (2006) Effect of power quality distortion on electrical drives and transformer life in paper industries: simulation and real time measurements. In: IEEE pulp and paper industry technical conference, 18–23 June 2006, pp 1–9
6. Puchalapalli S, Pindoriya NM (2016) Harmonics assessment for modern domestic and commercial loads: a survey. In: 2016 international conference on emerging trends in electrical electronics & sustainable energy systems (ICETEESES), Sultanpur, pp 120–122
7. Vahidnia A, Dastfan A, Banejad M (2009) Determination of harmonic load characteristics in distribution networks of cities. In: 2009 international conference on power engineering, energy and electrical drives, Lisbon, pp 442–446
8. Lalli MS, Paul IPS (1998) Field measurement of power quality in steel rolling mills. In: Power quality '98, Hyderabad, pp 279–282
9. Marzband M, Shaikholeslami A (2006) Power quality monitoring at the industrial, commercial and educational centers of Mazandaran province and presenting the related solution. In: 2006 international conference on power electronic, drives and energy systems, New Delhi, pp 1–6
10. Milanović JV et al (2014) International industry practice on power-quality monitoring. IEEE Trans Power Delivery 29(2):934–941
11. Aziz T, Nandi SK, Rahman MS, Riadh RR (2015) Study of power quality with changing customer loads in an urban distribution network. In: 2015 3rd international conference on green energy and technology (ICGET), Dhaka, pp 1–6
12. Ahmad BA, ElSheikh HH, Fadoun A (2015) Review of power quality monitoring systems. In: 2015 international conference on industrial engineering and operations management (IEOM), Dubai, pp 1–8
13. Püvi V, Tukia T, Lehtonen M, Kütt L (2019) Survey of a power quality measurement campaign in low-voltage grids. In: 2019 electric power quality and supply reliability conference (PQ) & 2019 symposium on electrical engineering and mechatronics (SEEM), Kärdla, pp 1–7
14. Gasch E et al (2016) Assessment of power quality performance in distribution networks part I—measurement campaign and initial analysis. In: 2016 17th international conference on harmonics and quality of power (ICHQP), pp 164–169
15. Bokhari A et al (2014) Experimental determination of the ZIP coefficients for modern residential, commercial, and industrial loads. IEEE Trans Power Delivery 29(3):1372–1381
16. Mubiana A, Mutingi M, Gope G, Mbohwa C (2016) Analysis of power quality factors—a case study of Windhoek-Namibia
17. Shao S, Pipattanasomporn M, Rahman S (2013) Development of physical-based demand response-enabled residential load models. IEEE Trans Power Syst 28(2):607–614
18. Bollen MHJ (2000) Understanding power quality problems. IEEE Press, Piscataway, NJ
19. Swachh power—a glimpse of power quality in India

Design of Low-Power 10T Full Adder Circuit using DG—MOSFET at 180 nm Technology

Aakrati Agrawal, Pramod Kumar Jain, and D. S. Ajnar

Abstract Keeping the current scenario of digital era in outlook where devices with high portability and performance are in demand, the paper proposes a technique, i.e., double-gate MOSFET (DGFET) which holds well. This technique replaces single-gate transistor logic with double gate results in reducing problems like short channel effects in FETs. The paper includes design of 10T full adder digital logic circuit, parametric analysis such as power consumption and resistance offered. The implementation of design and simulation results has done using Cadence Virtuoso tool at 180 nm technology.

Keywords Double-gate FET · Full adder · Single gate · Power consumption · MOSFETs

1 Introduction

The most considerate predictions done in the field of microelectronics and VLSI are by Gordon Moore who described a law named Moore's law which states that "the number of transistors in a dense integrated circuit (IC) doubles about every two years." To keep following the law, the dimensions of transistor get reduced by half in a period of every three years. Since transistor is the integral component in microelectronics and VLSI that every feature and parameter is essential to look after. The prime objective of switching technologies from using BJT (bipolar transistor) to FETs (field effect transistor) and MOSFETs is improving performance parameters such as power consumption, increased speed, smaller chip area, cost reduction.

A MOSFET is a semiconductor device which is widely used in electronic devices due to various features it possess like amplification and high switching speed when

A. Agrawal (✉) · P. K. Jain · D. S. Ajnar
Department of Electronics and Instrumentation Engineering, SGSITS, Indore, M.P., India
e-mail: aakrati1506@gmail.com

D. S. Ajnar
e-mail: dsajnar@sgsits.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_8

dealing with electrical signals. It is considered as an ideal switching device which has bidirectional capability and offers low static power dissipation, high scalability, and high packing density in integrated circuits [1–8].

There are two modes of operating MOSFETs:

(1) Depletion mode in FET
(2) Enhancement mode in FET.

There are several types of MOSFETs but here described are two types, i.e., single-gate MOSFET and double-gate MOSFET.

1.1 SG—MOSFET

A single-gate MOSFET is the basic component which has four terminals, i.e., source, gate, drain, and body. In this type of device, all four terminals operate independently and offering different functionality. Single-gate FET operates in both modes of operations, i.e., enhancement mode and depletion mode [1, 9–11].

1.2 DG—MOSFET

The double-gate MOSFET (DGFET) is designed using two single-gate MOSFETs that are placed in such a way the identical terminals of both the components get connected. The symbol now has two gates, i.e., front gate and back gate. These two gates can operate simultaneously also which is termed as symmetrical driven (SDDG). The gates can be operated independently by applying voltage bias to one and switching the other which is known as independent driven (IDDG) [1–5, 8, 11, 12].

Figure 1 shows circuit symbol of p-type and n-type DG-MOSFET.

Transistors and Fig. 2 show modes of operations of double-gate transistor.

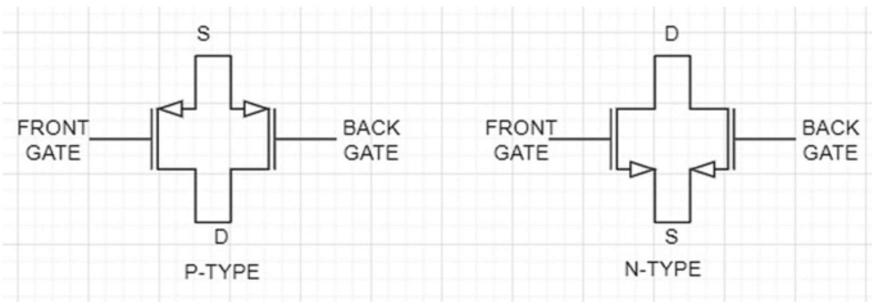


Fig. 1 Circuit symbols

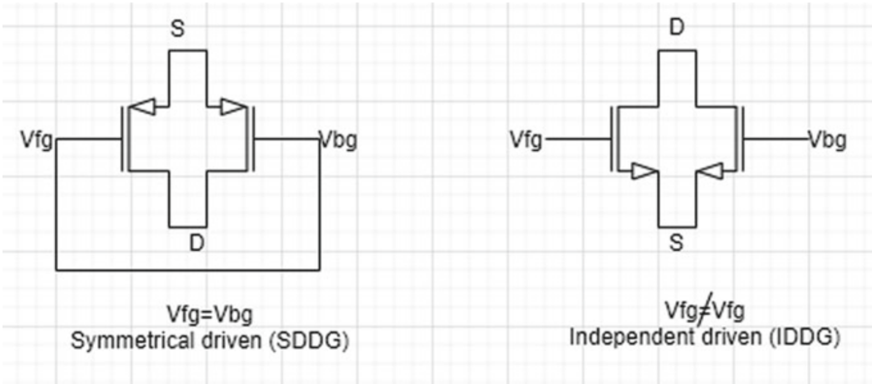


Fig. 2 Operating modes

2 Full Adder

The efficiency of any digital device depends entirely on the internal devices and components used in designing the particular unit. As the market growing and dependency on digital devices is increasing rapidly, there is always a need to improve the performance based on some parameters such as smaller size, cost per unit, durability, and power consumption, these parameters helps a device to be more reliable and portable. Full adder is one integral unit in any digital device, i.e., microcontrollers and microprocessors in its arithmetic logic unit (ALU).

The basic block diagram using logic gates of one bit full adder is shown in Fig. 3.

Figure 3 shows a combinational circuit of full adder that is used to perform arithmetic operations in ALUs. It takes three bits as inputs (A, B, Cin), performs the required mathematical operation using the gates and gives two outputs (S, Cout).

The Boolean expression for this logic is described as:

$$S = A \oplus B \oplus C\text{in}$$

$$C\text{out} = AB + AC\text{in} + BC\text{in}$$

For designing of full adder circuit, an additional degenerate module is required, i.e., a 5T XOR-XNOR circuit. This module helps in reducing the number of transistors thus reduces the integration size and also helps in making leakage current less with less static power dissipation.

This module uses combination of p-MOS and n-MOS transistors which takes two inputs (A, B) and give outputs (C, D).

The circuit is biased with the input pulses and has a ground pin (gnd, 0 V).

The schematic design of 5T XOR-XNOR module is shown in Fig. 4.

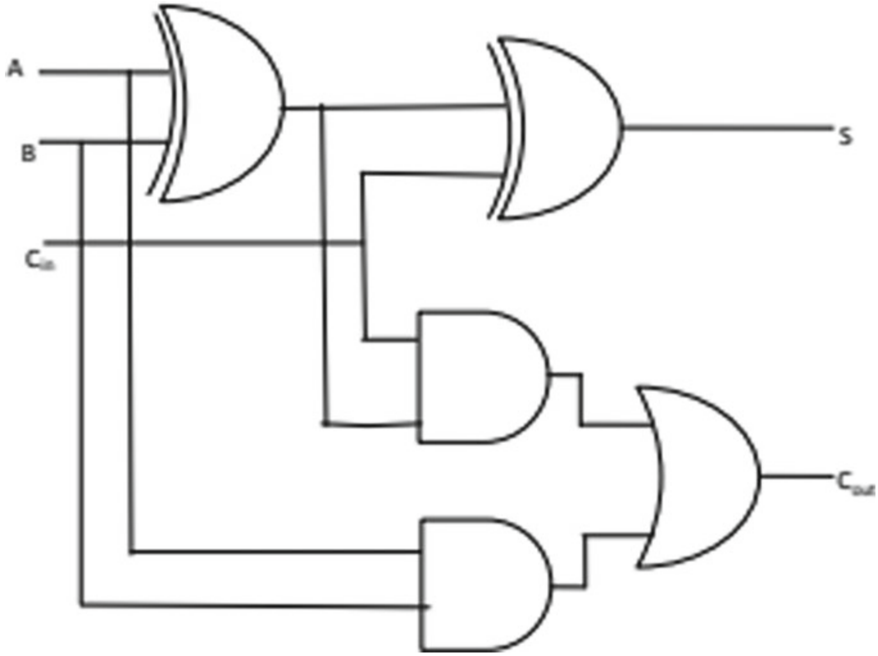


Fig. 3 One-bit full adder

The module is best suitable even where there is requirement of low-power operations since it excludes the use of voltage.

The circuit performs two operations, i.e., XOR and XNOR on the input pulses taken as *A* and *B*. This operation can be well described through its truth table that is described in Table 1.

The two signals (1−) and (0+) are degraded signals due to threshold voltage loss. In the case where both the inputs are "0," the XNOR output becomes floating which is indicated by the symbol (X). Hence, module is considered logically degenerated.

Combinations of all the inputs are not satisfied this module, but it is sufficient enough to produce the output signals (*C*, *D*) needed for the appropriate functioning of the full adder.

The transient response for the circuit module is recorded at 100 ns shown in Fig. 5.

3 Full Adder Design Using Single-Gate Transistor Logic

The designing of full adder using single-gate transistor logic is implemented by using 5T XOR-XNOR module discussed above. The output signals (*C*, *D*) taken from previous module are used as input to the other half module of single-gate full adder design circuit.

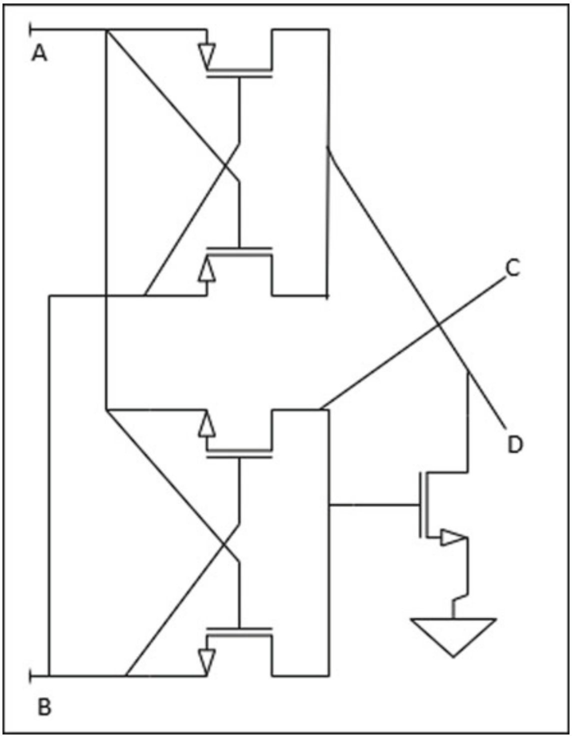


Fig. 4 5T XOR-XNOR module

Table 1 XOR and XNOR operations on the inputs *A* and *B*

A	*B*	$\overline{D}$ (XNOR)	*D* (XOR)
0	0	X	0+
0	1	0	1
1	0	0	1
1	1	1−	0

The schematic design of full adder using single-gate transistor logic is shown in Fig. 6.

The Boolean expression for the 5T XOR-XNOR module is described below [13]:

$$\boxed{\begin{array}{l} S = (\overline{A \oplus B})C + (A \oplus B)\overline{C} \\ Cout = (A \oplus B)C + (\overline{A \oplus B})A \end{array}}$$

The Boolean expression for SG-MOSFET-based double adder using the 5T XOR-XNOR module is described below:

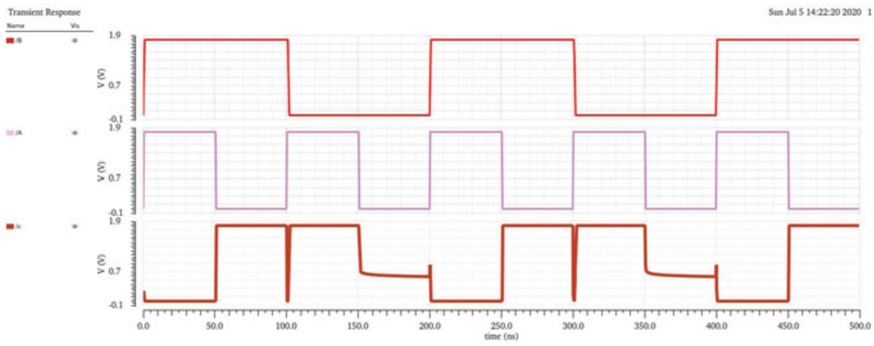


Fig. 5 Transient response at 100 ns

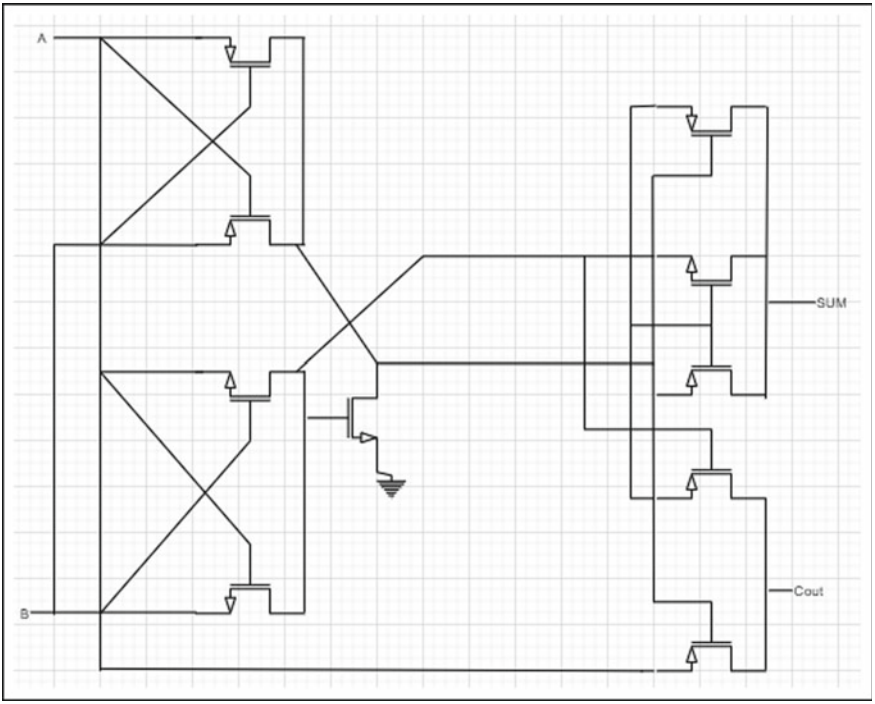


Fig. 6 Schematic design of full adder using SG transistor

$$\boxed{\begin{aligned} &S = D \oplus C \\ &Cout = D \oplus C + \overline{D} \oplus A \end{aligned}}$$

The layout design of full adder circuit using single-gate transistor logic is carried out by using Cadence Virtuoso tool at 180 nm technology, which is shown as Fig. 7.

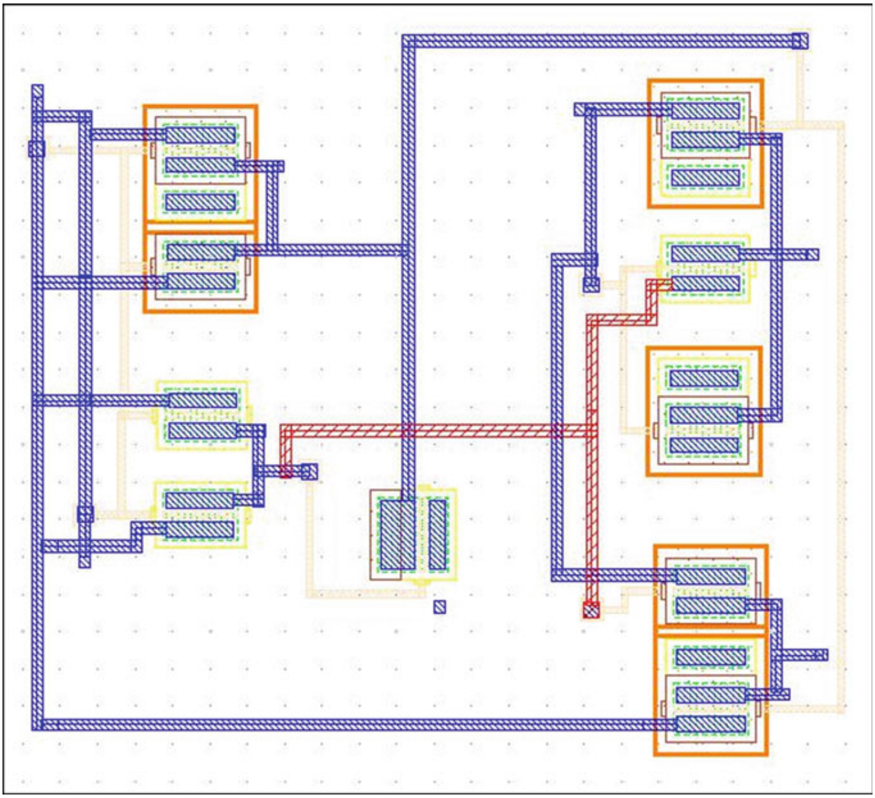


Fig. 7 Layout design of full adder using SG transistor

All paths in all layers of the layout design are dimensioned in λ units, and subsequently, λ can be allocated an appropriate value compatible with the feature size of the fabrications process. This layout design follows layout design rules, i.e., sub-micron rules such as:

Min poly width (L) = 2λ
Min spacing b/w poly = 2λ
Width of contacts = 2λ
Min metal to metal spacing = 3λ
Min metal width = 3λ or 4λ
(w) Min width of diffusion = 6λ.

To check the correctness of design rules, design rule check (DRC) is done using Assura; it ensures that the layout passes through the rules designed for faultless fabrication followed by LVS (layout vs. schematic) that determines whether a particular circuit layout corresponds to the original schematic of the design.

4 Full Adder Design Using Double-Gate Transistor Logic

The designing of full adder using double-gate transistor logic (DG-MOSFET) technique is carried out by using the single-gate transistor logic design in an equivalent way. Each transistor in the circuit is configured and connected in a manner that their source terminal and drain terminal must be connected to each other. The two (front and back) gates of DG-MOSFET logic are electrically coupled which operates in symmetrical mode of operation. This mode of operation establishes connection between the terminals in a way that channel and gates get near to each other which in turn provides good control over conductance of the channel and reduces leakage.

Also, DG-MOSFET technique drastically reduces the short channel effects that occur in basic MOSFET devices which reduces the performance. This technique is preferable more since it possess better scalability even in nanometer range in digital circuits.

The schematic design of double-gate MOSFET-based full adder is shown in Fig. 8.

The transient response of proposed circuit is carried out using an EDA tool, i.e., Cadence Virtuoso tool at 180 nm technology which is shown in Fig. 9.

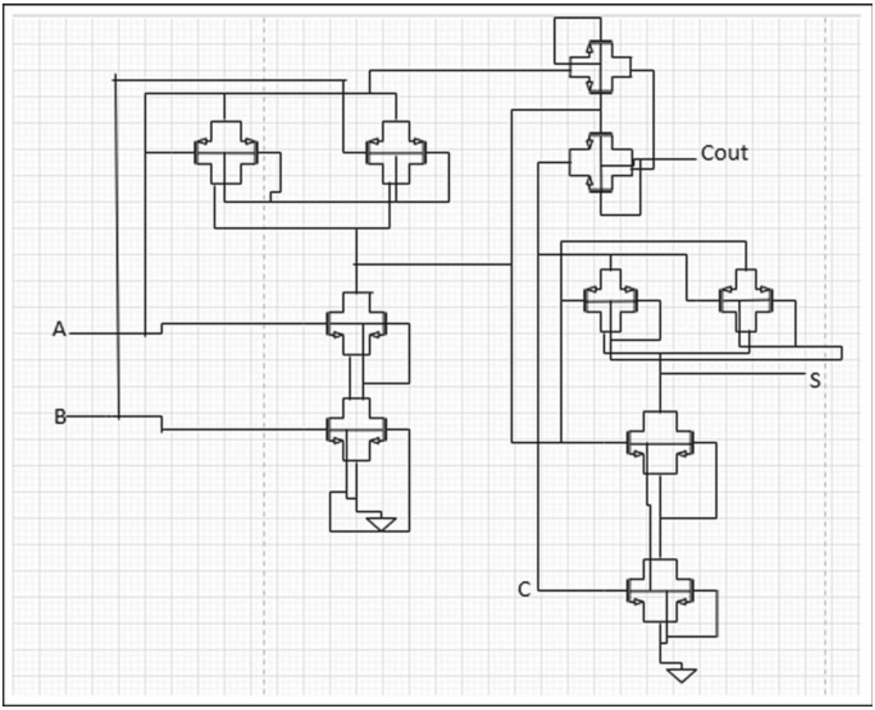


Fig. 8 Schematic design of full adder using DG-MOSFET

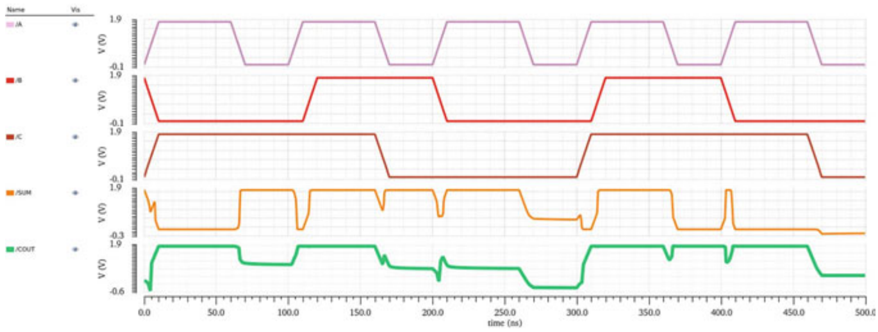


Fig. 9 Transient response of full adder using DG-MOSFET at 180 nm

5 Results and Observations

5.1 Average Power Consumption

The average power consumption for a CMOS circuit is described by the equation [9]:

$$\begin{aligned} P(\text{avg.}) &= P(\text{dynamic}) + P(\text{leak}) + P(\text{short-circuit}) \\ &= \text{CLVddVfclk} + \text{IleakVdd} + \text{IscVdd} \end{aligned}$$

From the above equation, it is clearly visible that the average dissipated depends upon different parameters such as current components (leakage + short-circuit), induced capacitance, and supply voltage. So, to lower the average power consumption, reducing the applied supply voltage will help. This concept will enhance the durability and overall performance of the MOSFET-based full adder circuit.

Figure 10 shows the graphical representation of power consumption against time (ns) taken at different levels of voltage, i.e., at 290, 310, 330, 350, 370 mV for double-gate MOSFET-based full adder circuit.

Figure 11 shows the numerical value of average power consumption in DG-based full adder circuit. This calculation is done using parametric analysis in EDA tool Cadence Virtuoso at 180 nm technology.

The average power consumption of double-gate MOSFET-based full adder calculated above is as follows:

$$P\text{avg (DG)} = 43.38\text{E} - 9$$

And the average power consumption of single-gate MOSFET-based full adder calculated using the same process of parametric analysis is as follows:

$$P\text{avg (SG)} = 1.769\text{E} - 3$$

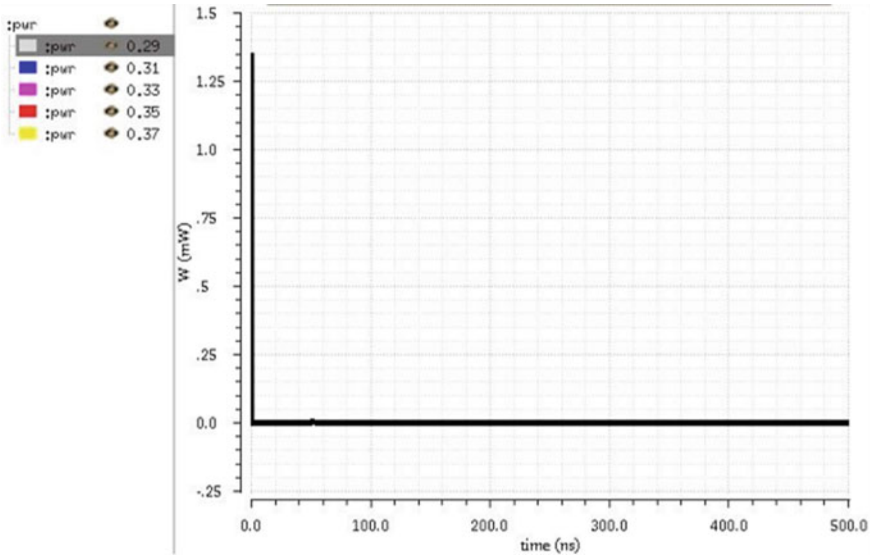


Fig. 10 Power consumption versus time in ns

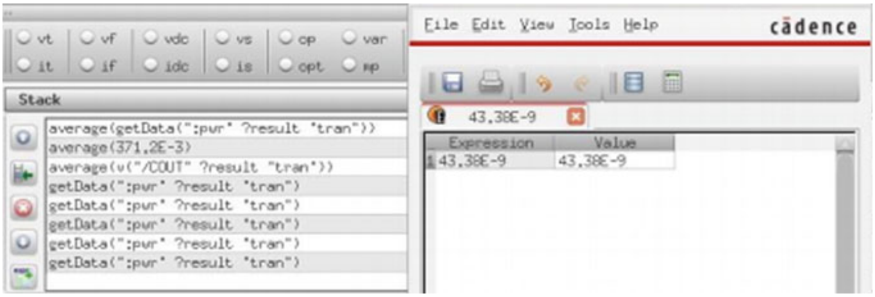


Fig. 11 Average power consumption

5.2 Variation of Power with Voltage

Figure 12 shows a graph of variation of power consumption with respect to input voltage of both single-gate and double-gate MOSFET-based full adder circuit.

From the graph, the percentage change in power consumption of SG and DG-MOSFET-based full adder circuit, it has been observed that the power consumption gets reduced by 22% in double gate logic circuit.

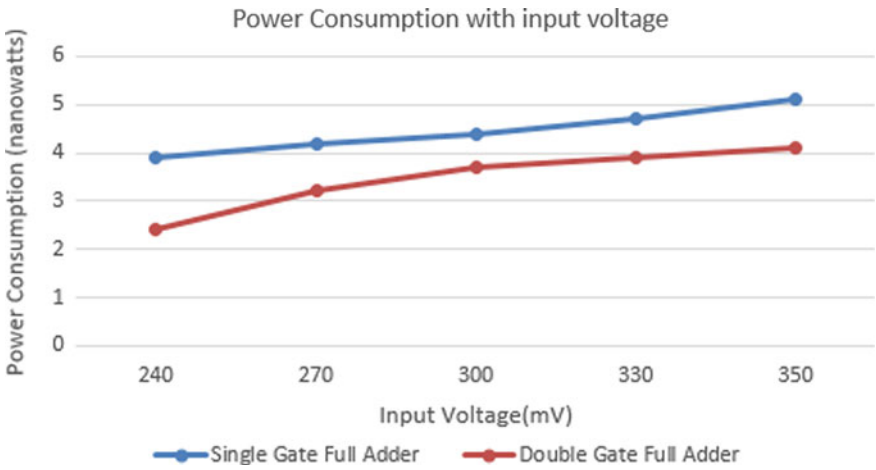


Fig. 12 Power consumption versus input voltage

6 Comparison Results of Parameters

Parameters	Conventional design	Proposed design
Power (avg.)	71.4 μW (14T)	43.38 nW (10T)
Tools and technology	Tanner EDA 13.0	Cadence Virtuoso (180 nm)
Resistance	More	Less
Current (Ids) (μA)	0.018	0.015

7 Conclusion

From the above observations and simulation results, it can be easily concluded that double-gate MOSFET-based full adder is way better suitable for the designing of digital devices than single-gate MOSFET based since DG based provides low power consumption, high speed, good scaling at low voltage range. These are most reliable devices that can be used for any microprocessors and other portable digital devices.

References

1. Gupta SK (2011) Double gate MOSFET and its application for efficient digital circuits. In: 2011 IEEE
2. Monisha A (2014) Performance and characteristic analysis of double gate MOSFET over single gate MOSFET. In: 2014 international conference on electronics and communication systems (ICECS—2014), Coimbatore
3. Vaddi R. Design and analysis of double gate MOSFETs for ultra low power radio frequency identification (RFID): device and circuit co-design. J Low Power Electron Appl
4. Moaiyeri MH, Mirzaee RF, Navi K (2009) Two new low-power and high performance full adders. J Comput 4
5. Roy SM (2003) Leakage current mechanisms and leakage reduction techniques in deep-submicron CMOS circuits. In: Proceedings of the IEEE 2003
6. Mishra S, Tomar SS, Akashe S (2013) Design low power l0T full adder using process and circuit techniques. In: Proceedings 7th international conference on intelligent systems and control
7. Lin J-F, Hwang Y-T, Sheu M-H (2012) Low power 10-transistor full adder design based on degenerate pass transistor logic. IEEE
8. Amara A, Parikh CD, Nagchoudhuri D (2011) A 0.7-V rail-to-rail buffer amplifier with double-gate MOSFETs. In: Faible tension faible consommation (FTFC). IEEE
9. Srivastava AK, Akashe S (2013) Design high performance and low power 10T full adder cell using double gate MOSFET at 45 nm technology. In: 2013 ICCCM
10. Srivastava AK (2013) Comparative analysis of low power 10T and 14T full adder using double gate MOSFET at 45 nm technology. Int J Comput Appl 75(3). 0975-8887
11. Sharma S, Kumar P (2009) Non overlapped single and double gate SOI/GOI MOSFET for enhanced short channel immunity. J Semicond Technol Sci 9
12. Gupta SK, Pathak GG, Das D, Sharma C (2011) Double gate MOSFET and its application for efficient digital circuits. In: 3rd international conference on electronics computer technology (JCECT). IEEE
13. Sahani JK (2015) Design of full adder circuit using double gate MOSFET. In: 2015 fifth international conference on advance computing and communication technologies
14. Weste NHE, Harris D, Banerjee A (2011) CMOS VLSI design: a circuit and system perspective, 3rd edn. Pearson Education
15. Shyam Kishore G. A novel full adder with high speed low area. In: Proceedings of 2nd national conference on information and communication technology
16. Chiang TK (2005) A scaling theory for fully-depleted, SOI double-gate MOSFETs: including effective conducting path effect. Solid State Electron 49
17. Abbasalizadeh S, Forouzandeh B (2012) Full adder design with GDI cell and independent double gate transistor. In: 20th Iranian conference (lCEE), May 2012
18. Sharma S, Kumar P (2008) Optimizing effective channel length to minimize short channel effects in sub 50 nm single/DG SOI MOSFETs. J Semicond Technol Sci
19. Srinivasa R, Scheinberg N. Short-circuit power dissipation estimation for CMOS logic gates. IEEE Trans Circuits Syst I Fundam Theory Appl 41
20. Colinge J-P (ed) (2008) FinFETs and other multi-gate transistors. Springer
21. Power-delay characteristics of CMOS adders. IEEE Trans VLSI Syst

Study the Effect of Pt–Ga Sensitization on Sol–Gel-Grown Nano-ZnO Thin-Film H_2 Sensor

Aniket Das, Gour Gopal Jana, Jyotirmoy Nandy, Bijoy Kantha, Subhashis Roy, and Subir Kumar Sarkar

Abstract The highly sensitive hydrogen gas sensor has been fabricated by investigating the catalytic effects of platinum (Pt) and galladium (Ga) on nanoparticle ZnO. The key aspects of the fabricated sensor like operating temperature, sensitivity, and selectivity are thoroughly studied. The most effective fabrication process sol–gel method has been adopted, and Si substrate is chosen as base material for sensor. The spin-coated sensor is tested in presence of hydrogen gas in the 500–3000 PPM range. The significantly gained sensitivity of the sensor has been reported which is from 58.7 to 85.8% by introducing the foreign materials Pt and Ga into ZnO lattice. The characterization of the sensor is carried through state-of-art techniques such as scanning electron microscopy (SEM) and X-ray diffraction (XRD) which present the surface morphology and phase of the sensor. The achieved ZnO nanoparticle average size is 70–90 nm.

Keywords Hydrogen · ZnO · Pt · Ga · Sol–gel · Nanoparticle · Sensitivity · Selectivity

A. Das · B. Kantha
Electronics and Communication Engineering, Netaji Subhash Engineering College, Kolkata, India

G. G. Jana
Electronics and Communication Engineering, GKCEM, Kolkata, India

J. Nandy
Home Department, Government of West Bengal, Kolkata, India

S. Roy (✉)
Electronics and Communication Engineering, Techno India University, Kolkata, India
e-mail: subhashis.r@technoindiaeducation.com

S. K. Sarkar
Electronics and Telecommunication Engineering, Jadavpur University, Kolkata, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_9

1 Introduction

In recent times, there are plenty of application of ZnO-based thin film sensor in commercial markets. These are like electronic, optoelectronic, and MEMS-based gas sensors. Due to high band gap energy (3.37 eV), good binding energy (60–65 meV), n-type conductivity, high structural stability, the ZnO becomes most popular sensing material. Hence, in literature, there are many research works available on ZnO-based sensor fabricated on different substrates like GaAs, sapphire, ITO-coated glass, n-Si, and p-Si. Further, variety of fabrication processes also adopted which are compatible with ZnO material like pulsed laser deposition, molecular beam epitaxy, chemical vapor deposition, metal organic CVD, RF sputtering, sol–gel [1–6]. This makes ZnO the most effective material as sensor. Hence, in this research work, it is taken, and further, study has been conducted.

The surface-to-volume ratio of nanoparticle plays very important role in sensing applications. The grain size of nanoparticle sensor can be modulated by adding different foreign materials like manganese (Mn), magnesium (Mg), aluminum (Al), antimony (Sb), tin (Sn), ytterbium (Yb), silver (Ag), and gallium (Ga) [4, 6–15]. By doping or adding specific metal ions into the lattice of ZnO, the characteristics like electrical and optical can be modified in the nanoscale regime [4–7]. Interestingly, the doping of the trivalent group elements such as Al, Ga into the ZnO lattice structure causes enhancement of sensing performance [3, 4].

Hence, Ga-doped ZnO and Pt-doped ZnO thin-film sensors on p-silicon substrate have been fabricated, and their sensing characteristics are investigated in presence of H_2 gas. Scanning electron microscopy (SEM) and X-ray diffraction (XRD) methods are used to investigate the surface morphology and phase of the sensor. However, to obtain a better understanding of the fabricated sensor, the study of selectivity also carried out.

2 Experimental Work

There are few reports available on ZnO-based sensors with Pd and Ga modifications [2, 14]. But, a ZnO-based sensor with Pt modification research paper has not been published so much for sensing H_2 gas. It is reported that the sensitivity and the operating temperature of thin-film sensor is effected by the addition of noble metals like Pt, Pd, and Au [3, 4]. In the present work, PZ1 (Pd–Ag/Si–(Pt–ZnO)/Pd–Ag), PZ2 (Pd–Ag/Si–ZnO/Pd–Ag), and PZ3 (Pd–Ag/Si–(Ga–ZnO)/Pd–Ag) sensors have been prepared for sensing H_2 gas. Sol–gel process is used for the preparation of sensing element. The preparation route of ZnO sol (milky white color) has been taken from our previous work [8]. The spin coating method is utilized for the preparation of PZ2 sensor on Si wafer, and the heat treatment is applied at the temperature of 500 °C for 2 h for better morphological stability. Then, PZ1 is fabricated following the routes as described in our work [13] using H_2PtCl_6 solution and ethylene glycol. Then, the obtained black color Pt nanoparticles are mixed with deionized water (10 ml) and ZnO sol. Finally, using this solution, the Pt-modified ZnO thin-film sensor was fabricated by calcining at 600 °C in air for 2 h. Finally, the PZ3 fabrication is done using sol of ZnO and Ga powder mixed by using spin coating method and evaporated at the temperature of 150 °C for 30 min. The E-beam evaporation method is used to make palladium silver [Pd–Ag] (70–30%) contacts on both side of each sensor.

3 Results and Discussion

3.1 The Characterization of Thin Film

The X-ray diffraction (XRD) method is used to investigate the structure and phase of thin-film sensors. The average crystallite size is estimated at about 70–90 nm. The Scherrer's formula is used to measure the crystallite size of thin-film sensor [9].

$$D = \frac{k\lambda}{\beta \cos\theta} \quad (1)$$

where k is the space factor (0.9) and λ, θ, β are the wavelength of the X-ray radiation, the Bragg angle, and full width at half maximum of the diffraction line, respectively. Figure 1 shows the XRD image of ZnO sensor.

The topographical features of thin film sensors have been investigated by using scanning electron microscopy (SEM) method (Fig. 2a–c).

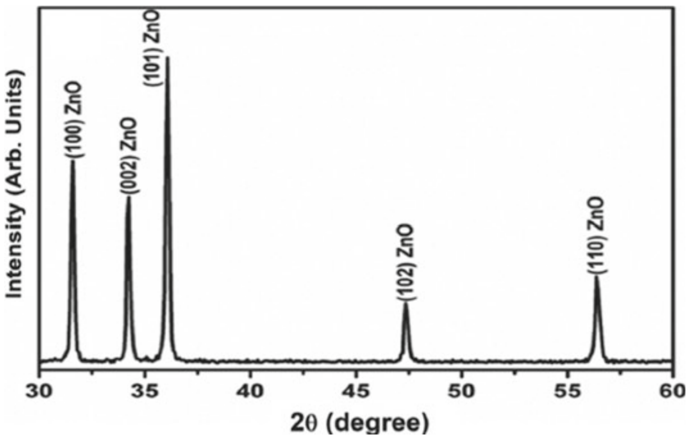


Fig. 1 XRD image of ZnO sensor

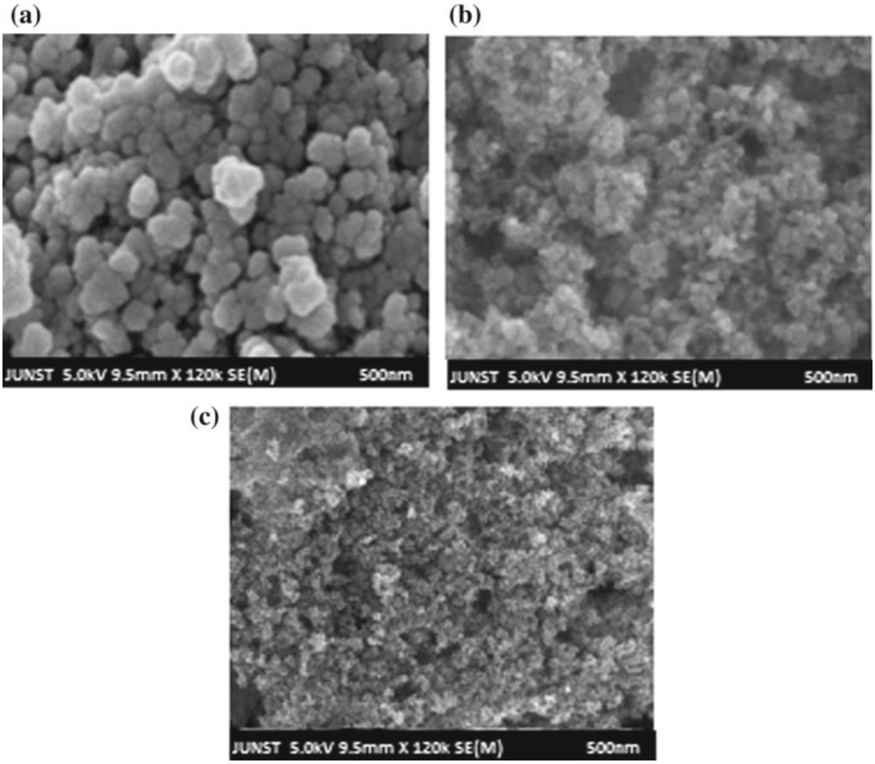


Fig. 2 **a** SEM of ZnO. **b** SEM of Pt–ZnO. **c** SEM of Ga–ZnO

3.2 The Characterization of Sensor

The vital parameters of a sensor are response magnitude, operating temperature, selectivity, response, and recovery times. All these parameters have played significant role to sense a target gas (H_2).

$$S\% = \frac{R_A - R_G}{R_A} \times 100 \quad (2)$$

where R_A represents as the resistance of sensor in presence of N_2 gas, and R_G represents as the resistance of sensor in presence of H_2 gas.

The Pt-modified ZnO-based sensor presents the lowest operating temperature of 160 °C, whereas the bare ZnO-based sensor is 250 °C (operating temperature). This phenomenon indicates that the modified sensor is more suitable compare to unmodified sensor for sensing field. Actually, the modified sensor has more interaction space (between sensing surface and gaseous molecules) at optimum operating temperature. Table 1 shows that the response and recovery times of Pt-modified sensor have better

Table 1 Response time and recovery time chart of fabricated sensors

H_2 (PPM)	Sample PZ1 (Pt–ZnO)		Sample PZ2 (ZnO)		Sample PZ3 (Ga–ZnO)	
	T_{RES}	T_{REC}	T_{RES}	T_{REC}	T_{RES}	T_{REC}
500	47	52	71	88	58	63
1000	40	64	62	102	51	71
1500	31	74	48	117	37	88
2000	26	86	35	129	27	101
3000	20	97	29	141	21	114

values compare to other sensors. These parameters have shown vital role to sense a target gas (H_2). The sensing characteristics will depend on both the response and recovery times.

Further, Fig. 3 shows the experimental setup for the sensing of H_2 gas. N_2 has been used as carrier gas, and temperature controller is used to increase the temperature to the operating temperature. Figure 4 presents the selectivity results of all the sensors in presence of different hazardous gas such as methane, methanol, ethanol, and hydrogen at their respective operating temperatures. As shown in Fig. 5, Pt-modified sensor reveals better sensitivity than unmodified sensor which can also explain the performance improvement of Pt-modified sensor. The highest sensitivity is found for the sensor with Pt modification as it provides very less amount of lattice discontinuation, higher stability, and highly reactive with target gas H_2. The response time and recovery time of each fabricated sensor are characterized and shown in Table 1.

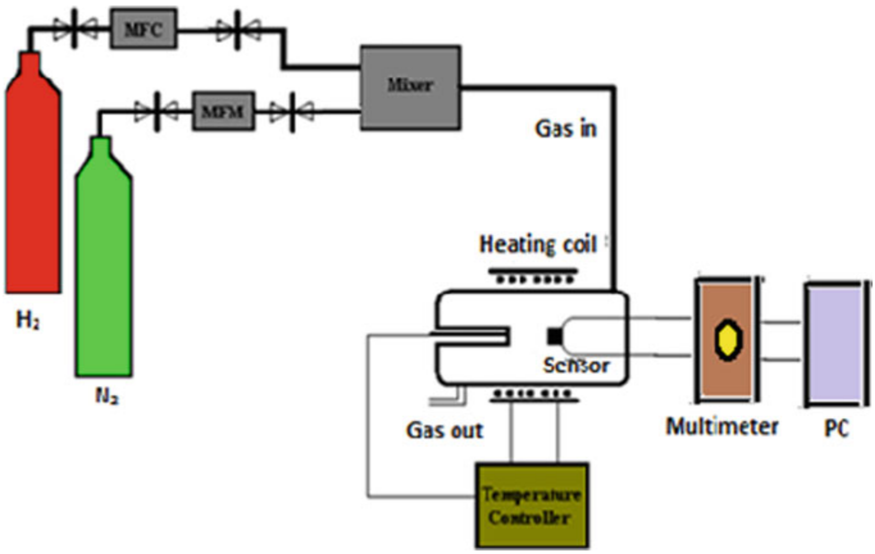


Fig. 3 Experimental setup

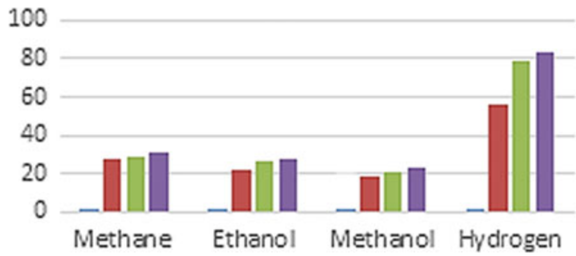


Fig. 4 Selectivity of the fabricated sensors

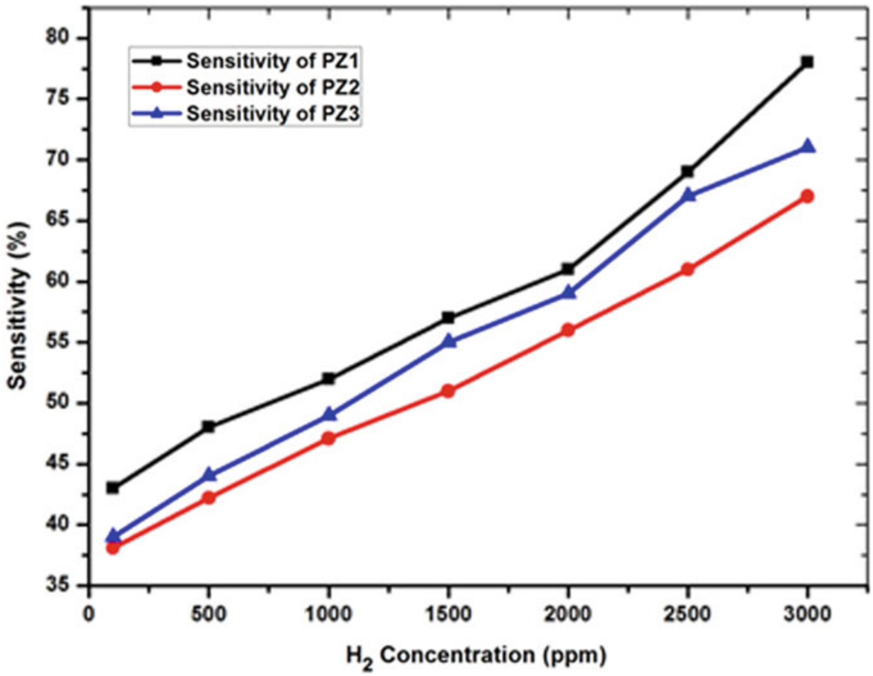


Fig. 5 Sensitivity versus H_2 concentration variation

As the concentration of the H_2 increases, the T_{RES}, T_{REC} also vary differently. The response time of sensor decreases with increase of H_2 concentration as surface reaction enhanced with increase of gas concentration. However, the recovery time acts differently, i.e., the recovery time increases with increase of gas concentration. The Pt-modified ZnO sensor acts efficiently with respect to other two sensors. Further, a comparative study of current work with previous reported H_2 sensor is shown by Table 2. The similar research studies have shown that the efficiency of the present fabricated sensor is performing better.

Table 2 Comparative study of current work with previous reported H_2 sensor

The structure of sensor	The fabrication process of sensor	The operating temperature (°C) of sensor	The response magnitude (%) of sensor	The range of detection (ppm)	References
ZnO–CuO	Sol–gel dip coating	275	71.85@10,000 ppm	1000–10,000	[8]
Pd-modified WO_3-ZnO	Sol–gel by spin coating	175	71@1000 ppm	200–1000	[14]
Ga-modified ZnO	PLD	150–300	50@5000 ppm	5000	[2]
Fe_2O_3@ZnO	Hydrothermal approach	220–350	78.5@5000 ppm	100–5000	[9]
Present work Pt-modified ZnO	Sol–gel spin coating	160	85.8@3000 ppm	500–3000	

4 Conclusion

In the present work, a bare ZnO, Ga–ZnO, and Pt–ZnO have been prepared using sol–gel technique on p-Si wafer. The effect of Pt metal on the gas sensing characteristics of ZnO sensor is investigated toward H_2 gas. SEM and XRD processes are used to examine the material properties of sensor. Pt–ZnO sensor shows maximum sensitivity (85.8%) in presence of H_2 gas at the operating temperature of 160 °C compare to other ZnO sensors. It has been also observed that Pt–ZnO sensor presents excellent result than other ZnO sensors in terms of selectivity, response time, and recovery time.

Acknowledgements The authors wish to thank the IC Center of Jadavpur University, Kolkata-700032, India for experimental support.

References

1. Hadeethi Y, Badran RI, Umar A, Heni SH, Raffah BM, Zhrani S (2020) Electrical properties of Ga-doped ZnO nanowires/Si heterojunction diode. Mater Express 10(6):794–801
2. Annathurai S, Chidambaram S, Rathinam M, Venkatesan GP (2019) Ga doping improved electrical properties in p-Si/n-ZnO heterojunction diodes. J Mater Sci Mater Electron 30(6):5923–5928
3. Dey A, Kantha B, Sarkar SK (2017) Sol–gel grown Pd modified WO_3 thin film based methanol sensor and the effect of annealing temperatures. Microsyst Technol 23(9):4195–4201
4. Kantha B, Roy S, Sarkar SK (2015) Implementation of Pd modified WO_3 thin film gas sensing system with bulk-CMOS and SOI-CMOS for monitoring leakage of hydrogen gas. J Nanoelectron Optoelectron 10:74–81

5. Kantha B, Roy S, Ghosh A, Khan A, Bag J, Sarkar SK (2012) Designing a micro-heater with genetic algorithm based optimized parameters and study its performance. In: International conference on communications, devices and intelligent systems (CODIS-2012), Jadavpur University, Kolkata, 28–29 Dec 2012
6. Liu Y, Hang T, Xie Y, Bao Z, Song J, Zhang H, Xie E (2011) Effect of Mg doping on the hydrogen-sensing characteristics of ZnO thin films. Sens Actuators B Chem 160:266–270
7. Kantha B, Sarkar SK (2014) The design of a low power MEMS based micro-hotplate device using a novel nickel alloy for gas sensing applications. J Nano- Electron Phys 6:1–5
8. Wang JX, Sun XW, Yang Y, Kyaw KKA, Huang XY, Yin JZ, Wei J, Demir HV (2011) Free-standing ZnO–CuO composite nanowire array films and their gas sensing properties. Nanotechnology 22(32)
9. Zhang J, Liu X, Wang L, Yang T, Guo X, Wu S, Wang S, Zhang S (2011) Synthesis and gas sensing properties of α-Fe_2O_3@ZnO core–shell nanospindles. Nanotechnology 22(18):1–7
10. Kantha B, Sarkar SK (2015) Comparative study of particle swarm optimization and genetic algorithm for the optimization of system parameters of MEMS based micro-heater. J Comput Theoret Nanosci 12:1641–1646
11. Siddiqui A, Mahboob R, Islam T (2017) A passive wireless tag with digital readout unit for wide range humidity measurement. IEEE Trans Instrum Meas 66(5):1013–1020
12. Wang C, Yin L, Zhang L, Xiang D, Gao R (2010) Metal oxide gas sensors: sensitivity and influencing factors. Sensors 10:2088–2106
13. Dey A, Roy S, Sarkar SK (2018) Synthesis, fabrication and characterization of ZnO-based thin films prepared by sol-gel process and H_2 gas sensing performance. J Mater Eng Perform (JMEP) 27(6):2701–2707
14. Acharyya D, Ghosal S, Roychaudhuri R, Bhattacharyya P (2018) Hierarchical MnO_2 nanoflowers based efficient room temperature alcohol sensor. In: 2018 IEEE SENSORS, New Delhi, pp 1–4
15. Bhowmik B, Dutta K, Bhattacharyya P (2019) An efficient room temperature ethanol sensor device based on p-n homojunction of TiO_2 nanostructures. IEEE Trans Electron Dev 66(2)

Effect of Electromagnetic Radiations on Children

Rahul Bramhane, Santosh Narayankhedkar, Sandeep Dabhade, and Vishakha Mote

Abstract In this research work the Electromagnetic radiation especially radiation from Mobile Towers is studied for multiple carriers and multiple operators in the Aurangabad City, state of Maharashtra of India. Out of 550 towers, samples of 50 towers are taken for consideration. Field strength in terms of micro Tesla is measured by the radiation meter and its corresponding Power density in terms of Watts per meter square is obtained. The effect of EM radiation on adult is calculated theoretically by computation and the same criteria made applicable for children of average age of 7 years and half of the height of adult under consideration for single carrier and single operator. In this paper we have covered practical observations for multiple carrier and multiple operators are compared with the standards of ICNIRP, FCC and WHO. It is found that the additive interference of radiations due to towers adjacent to each others have tremendously increased the power density and crossed the limits applicable for radiations. Since the children have small size head, delicate bones, thin skin and if exposed to radiations beyond the standard limits may cause harm in long run.

Keywords ICNIRP · Field strength · Power density

1 Introduction

In India, currently more than 80% of population is using Cell phones [1]. And to cater the services it has enabled the system to install no. of mobile towers to meet the communication demand. The number of Cell towers and cumulative radiation power are increasing without giving due respect to its disadvantages. Worldwide it is widely accepted that high radiation causing many health problems. And accordingly the guidelines of ICNIRP are revised with changing scenario of advancement in technology. The current paper is discusses about the analytical and experimental

R. Bramhane (✉) · S. Narayankhedkar · S. Dabhade · V. Mote
Dr. Babasaheb Ambedakar Marathwada University, Aurangabad, India
e-mail: rsbramhane26@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_10

comparison of radiation caused by mobile phone handset and cell tower and its probable implications on children.

2 Parameters Associated with RF EMF Radiation

According to the guidelines of ICNIRP the parameters associated with radiation are Absorbed energy density, Incident energy density, Plane-wave equivalent incident energy density, Absorbed power density, Incident power density, Plane-wave equivalent incident power density, Induced electric field strength, Incident electric field strength, Incident electric field strength, Incident magnetic field strength, and Specific energy absorption rate SAR. For any system the safe radiation limits are advised to follow as per the values of these parameters for specific frequency and the duration of exposure. The Incident electric Field Einc, Induced Electric Field Eind (Penetrated Electric Field) and Specific energy absorption rate are calculated for the purpose of study in this paper.

3 Determination of Penetrated Electric Field

There are three conditions used under which Eind is calculated,

1. Considering Mobile phone handset as a point source radiating EM wave at frequencies 900, 1800 and 2450 MHz, respectively, at a distance from 1 to 15 cm.
2. Considering Cell tower as a point source radiating EM wave at frequencies 900, 1800 and 2450 MHz, respectively, at a distance from 1 to 50 m.
3. It is assumed that the radiated electric field penetrates into the human body and causes Penetrated Electric field in V/m at the conditions of 0.1–0.5 mm Depth. In this paper Human body tissue skin is used for the computation purpose.

The equation derived for Penetrated EF is [2],

$$\mathrm{Eind} = \mathrm{Einc}\, \mathrm{e}^{\left(-\frac{z}{\delta}\right)} \tag{1}$$

Einc, is incident Electric Field in V/m, derived for 2 W and 20 W, respectively, for Mobile phone handset and Cell tower. Eind is the field inside the depth z and δ is skin depth.

Eind is the field inside the depth z and δ is skin depth.

The result obtained are as in Table 1.

Table 1 shows the penetrated electric field inside the skin [3, 4] due to the electromagnetic wave at frequency 900 MHz of mobile phone handset at depth, 'z' varying from 0.1 to 0.5 mm inside the body under consideration from 'r' which varies from 1 to 15 cm from the mobile phone is shown in Fig. 1.

Table 1 Penetrated electric field inside the skin

‘*r*’ in cm	Penetrated electric field in V/m inside the skin at depth in mm				
	0.1	0.2	0.3	0.4	0.5
1	1068.260	1042.502	1017.403	992.905	969.003
2	534.120	521.205	508.701	496.407	484.502
3	356.070	347.500	339.103	330.907	323.000
4	267.070	260.602	254.304	248.230	242.250
5	213.650	208.500	203.408	197.930	193.800
3	178.030	173.705	169.506	165.408	161.500
7	152.600	148.903	145.304	141.804	138.404
8	133.530	130.31	127.170	124.101	121.102
9	118.690	115.830	113.040	110.302	107.606
10	106.820	104.205	101.750	99.290	96.900
11	97.105	94.760	92.408	90.250	88.090
12	89.002	86.880	84.707	82.720	80.704
13	82.107	80.109	78.260	76.309	74.504
14	76.208	74.406	72.606	70.902	69.202
15	71.201	69.480	67.820	66.100	64.508

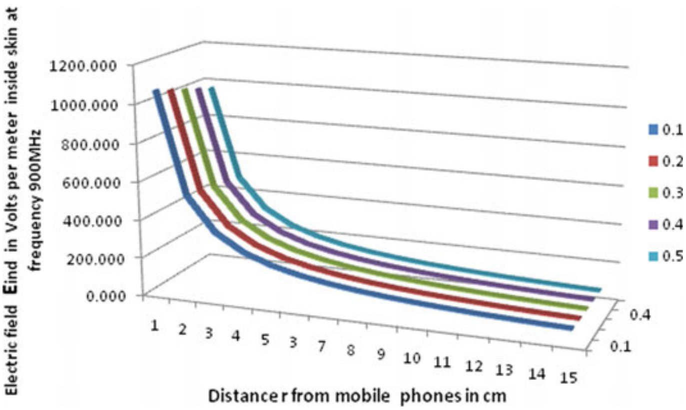


Fig. 1 Graph showing incident electric field versus distance from mobile phone

4 Determination of Specific Energy Absorption Rate (SAR)

For the above parameters, the Specific Energy Absorption Rate can be determined and presented by as [5],

Table 2 SAR inside the skin

Distance r cm	Specific energy absorption rate (SAR in W/kg) inside the skin with different depths in mm				
	0.1	0.2	0.3	0.4	0.5
1	900.8216	857.9496	817.1506	778.2856	741.2547
2	225.2012	214.4874	204.2836	194.5714	185.3098
3	100.0838	95.32408	90.78737	86.47096	82.35653
4	56.3003	53.62185	51.0689	48.64089	46.32555
5	36.03287	34.31667	32.6841	30.92234	29.64835
6	25.01954	23.83102	22.6955	21.61643	20.58913
7	18.38239	17.50884	16.6749	15.88146	15.12702
8	14.07297	13.40443	12.76622	12.15924	11.58043
9	11.12042	10.59095	10.08689	9.607304	9.149593
10	9.007373	8.579167	8.171024	7.782229	7.412088
11	7.443483	7.088317	6.751319	6.429651	6.124175
12	6.254183	5.957069	5.672537	5.402802	5.146008
13	5.329906	5.076138	4.834735	4.605241	4.386032
14	4.594394	4.376622	4.167579	3.969246	3.781209
15	4.002902	3.811866	3.630852	3.45842	3.293241

$$\mathrm{SAR} = \frac{\sigma Ei^2}{\rho} \tag{2}$$

where Ei is the field inside that material, σ is the conductivity of the material, ρ is density [6]. The specific energy absorption rate (SAR) inside the skin due to the electromagnetic wave at frequency 900 MHz of mobile phone handset at depth, 'z' varying from 0.1 to 0.5 mm inside the body under consideration from 'r' which varies from 1 to 15 cm from the mobile phone is shown in Table 2.

According to the guidelines of ICNIRP 1998, For Frequency ranging from 400 to 2000 MHz Induced EF is $1.375f^{½}$ and for $f = 900$ MHz Eind is 41,250 V/m is safe [7]. And as per the revised guidelines of ICNIRP March 2020 [8], the safe limit for exposure for Induced Electric field for 400–2000 MHz > 400–2000 MHz $4.72fM^{0.43}$. Eind is 33,440 V/m.

For the frequency of Electromagnetic Magnetic Wave ranging from 10 MHz to 10 GHz its safe limit for SAR is 0.4 W/kg as per ICNIRP.

From the results obtained from Tables 1 and 2, it is observed that the mobile phone handset at 2 W of radiated power with single carrier and single operator penetrates an electric field within safe limits of ICNIRP standards. Whereas, the computed values of SAR increases if the mobile handset is moved from 15 to 1 cm. And do not fall in the safe limits as prescribed by ICNIRP guidelines.

5 Measurement of Incident Electric Field

In this paper, the radiation due to Mobile phone handset is measured for several handset of which 5 samples are taken in to consideration and presented in Table 3. These values are taken for comparison with the analytical method. Comparison based on radiation due to handsets is shown in Fig. 2.

In second part of the experiment the radiation due to Cell towers is measured in Aurangabad city, the state of Maharashtra of India. The city has more than 550 cell towers and 10% of the samples are taken into account.

In Table 1, the penetrated electric field is derived from the incident electric field at 900 MHz, whereas the Incident electric field is derived from the field strength in micro Tesla using RF calculator [9] obtained practically for five different mobile phone handset is shown in Table 3. The increased incident Electric field is due to multiple carriers and multiple operators using different radiation powers. And both the results shows similarity as EF increase as the object is moved toward the mobile phone handset. Here electromagnetic field measurement method is adopted [10].

In further analysis it is found that the Penetrated electric field has more effect on certain tissues like Skin, Blood, Skeletal Muscles and Bone. The Dielectric properties like permittivity and electrical conductivity of human body cells are taken into account for computation purpose [3, 4]. And we got varied response due to change in conductivity which leads to in temperature of the tissue. If we consider the change in temperature in the body cells of Children, then it has noticeable effect.

Table 3 Incident electric field for different handset

Distance *r* from mobile phone handset in cm	Incident electric field, Einc in V/m		
	Handset 1	Handset 2	Handset 3
1	63,290.10143	90,755.61715	86,874.62036
2	60,006.18107	59,110.56643	60,901.79572
3	59,110.56643	57,319.33715	34,928.97107
4	27,764.05393	36,123.12393	38,809.96786
5	20,897.675	30,152.35964	29,256.745
6	18,210.83107	24,181.59536	23,285.98072
7	17,912.29286	20,599.13679	26,271.36286
8	17,315.21643	18,210.83107	19,404.98393
9	15,225.44893	17,613.75464	17,912.29286
10	14,926.91072	16,419.60179	17,016.67822
11	14,628.3725	15,225.44893	15,822.52536
12	13,135.68143	14,329.83429	14,628.3725
13	13,434.21964	13,135.68143	13,135.68143
14	13,135.68143	12,837.14321	12,837.14321
15	13,135.68143	12,538.605	12,538.605

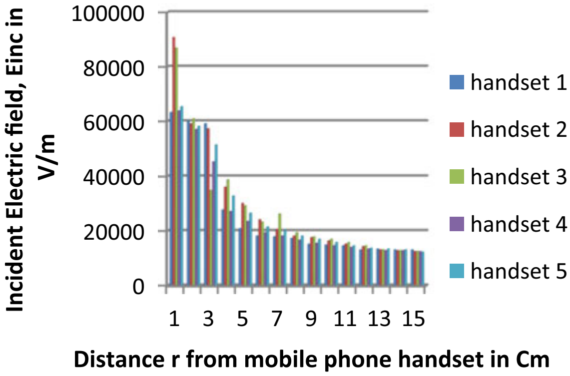


Fig. 2 Graph of incident electric field versus distance from mobile handset

This can be better explained with the example associated with the report submitted by Kumar [11], if we assume an average children with age of 7 years and half of the above height and weight, the area will be 0.718 m^2, Power received by children body will be Power density *X* area, i.e., 3.38 W. In one hour, microwave energy absorbed will be 12.14 kW s. In one day microwave energy absorbed will be 12.15 $\times$ 24 = 291.56 kW s. And with this concept Children body can safely be kept in oven for 291.56 kW s/500 W = 583 s = 10 min approx. per day. This example denote the single carrier frequency power radiated at GSM900 with safe exposure limits as per ICNIRP 1998 guidelines and its associated power density. Practically there are multiple carrier and multiple operators.

From the Experimental method used above, the Power density derived from practically obtained values of Incident Electric field is much more beyond limits of exposure. Practically we are exposed to heavy RF EMF exposure for 24 $\times$ 7 and throughout the years. It impacting slowly as the symptoms are hidden. The tissues of Children during growth are very sensitive and vulnerable to the changes associated due to radiation.

There are several litigations pending and several orders are issued by the court of law of India regarding cell towers and radiation limits [12, 13]. The DOT has issued guidelines to install towers at nearer distance with much less radiation than the ICNIRP and WHO guidelines. Even due to court order common public can get test reports for the tower in their area by paying Rs. 4000 at TERM cell of DOT [14]. The area Aurangabad city has urban population as per provisional reports of Census India, in 2011 is population is 1,193,167 of which 618,845 are males and 574,322 are females. The number of children between 1 and 12 age groups are assumed to be 15% of the total population [15] and as per our observations the cell towers in the city are such that at several places the distance among their installation is almost 200–300 m. This will certainly harmful for the growing children. In our experimental results for cell tower radiations it is found that the Measured Incident Electric Field is much more than the safe limits and uniformly spread over the city area due to additive interference of radiated power by multiple frequency operators.

6 Conclusion

From Analytical method, if we consider the safe limits advised by ICNIRP, and then probable effect is none. But with the same method if used for Specific Energy Absorption rate (SAR) then it has considerable impact on body cells. Specially tissues of growing children.

From the experimental results based on the observations, if compared with ICNIRP guidelines, IEEE [16], NCRP [17], ARPANSA [18], DOT Guidelines [19–22], certainly the values obtained are alarming. The increased number of towers and increased number of mobile phone handsets in families has increased the risk of health with hidden symptoms, particularly on children who has to live life more than present adults.

References

1. https://www.statista.com/statistics
2. Polk C, Postow E (1996) Handbook of biological effects of electromagnetic fields, pp 1–23
3. Gabriel C (1996) Compilation of the dielectric properties of body tissues at RF and microwave frequencies. Report N.AL/OE-TR-1996-0037. Occupational and Environmental Health Directorate, Radiofrequency Radiation Division, Brooks Air Force Base, TX
4. https://itis.swiss/virtual-population/tissue-properties/database/dielectric-properties/
5. Adair ER, Peterson RC (2002) Biological effects of radio frequency radiation. IEEE Trans Microw Theory Tech 50:953
6. Stuchly MA, Stuchly SS (1981) Biological effects of electromagnetic fields. Phys Rev 61:435
7. ICNIRP GUIDELINES (1998) Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields. Health Phys 74(4):509
8. ICNIRP Guidelines (2020) Health Phys 118(5):496
9. https://www.compeng.com.au
10. Test procedure for measurement of electromagnetic fields from base station antenna (for telecommunication sector) document no. TEC/TP/EMF/001/01, Sept 2009
11. Kumar G (2010) Report on cell tower radiation. Submitted to Secretary, DOT, Delhi by Electrical Engineering Department, IIT Bombay, Powai, Mumbai, Dec 2010
12. High Court of Judicature at Allahabad Order in PUBLIC INTEREST LITIGATION (PIL) No. 48084 of 2015 in Smt. Asha Mishra vs. State of U.P. and 7 others
13. Delhi High Court order in W.P. (C) No. 5550/2015 in the matter Kapil Choudhary and anr Vs Union of India and others
14. www.dot.gov.in
15. www.infobase.co.in
16. Safety issues associated with base stations used for personal wireless communications (2001) IEEE, New York, pp 110–114
17. National Council on Radiation Protection and Measurement (NCRP) (1986) Biological effects and exposure criteria for radio frequency electromagnetic fields. Report No. 86. NCRP, Bethesda, MD
18. Australian Radiation Protection and Nuclear Safety Agency (ARPANSA) (2002) Maximum exposure level to radio frequency field (3 kHz to 300 GHz). Radiation protection series publication no. 3. ARPANSA, Victoria
19. Letter issued by Government of India (2013) Ministry of Communications and Information Technology, Department of Telecommunications (As-II Cell) dated $th Nov 2008, No. 842-998-2008-AS-IV/14

20. Letter issued by Government of India, Department of Telecommunications (Access Cell) dated 10th January 2013, File No. 800-15/2010 VAS
21. Letter issued by Government of India, Department of Telecommunications (Access Services Cell) dated 26th June 2013, File 800-115/2010—VAS Incidence
22. Letter issued by Government of India, Department of Telecommunications (Access Services Cell) dated 20th Nov 2013, File 800-15/2010—VAS

Experimental Investigation of Solar PV Panel Using Phase Change Material Cooling by Capric Acid ($C_{10}H_{20}O_2$)

Nimesh Patel, Pavit Shah, and Alkesh Mavani

Abstract The conversion efficiency of PV panel is affected by its surface temperature. Therefore, cooling techniques is required to reduce temperature. This paper presents the results of an experimental study for cooling of solar PV panel using phase change material. Capric acid is used as phase change material. It has been filled in vacuum bags and mounted on back surface of PV panel to absorb excess heat. The result shows that the back surface temperature is reduced by average 1.60% and maximum 9.19%. The average % increase in efficiency is achieved up to 5.80%. It can be concluded from this work that using capric acid as phase change material is found to be an efficient method for temperature regulation of the PV panel.

Keywords Capric acid · PCM bag · Phase change material · PV panel

1 Introduction

The demand of global energy for 21st century has turned into one of the most important concern for the researchers working on the energy conservation and management. For year 2008, the total energy consumption was 474×10^{18} J out of which 80–90% has been allocated to fossil fuels [1]. Many environmental concerns such as global warming, greenhouse effect caused by conventional resources have lead toward more use of non-conventional resources, and solar energy is one of the most important natural and renewable resource. The total energy received from sun is about 1.8×10^{11} MW [2]. This energy is more than the present world energy consumption. Solar photovoltaic cell converts this solar energy directly into electricity without any thermo-mechanical conversion step.

N. Patel · P. Shah (✉) · A. Mavani (✉)
L.D.R.P.—ITR, Kadi Sarva Vishwavidyalaya, Gandhinagar, Gujarat, India
e-mail: pavitshah@gmail.com

A. Mavani
e-mail: alkeshmavani@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_11

The conversion efficiency of PV cell is affected by many factors, but operating temperature of cell is major factor which affects the performance of PV panel. For the crystalline silicon cells, the higher operating temperature causes a drop in production of electricity between 0.4 and 0.65% K^{-1} [3, 4]. Therefore, many cooling techniques have been developed in order to reduce cell temperature. The phase change materials are widely used for passive temperature control. Various types of material and their capsulation have been employed for better performance. Abdelrahman et al. [5] carried out experimental investigation of PV/PCM system. The configuration involved PCM RT35HC, cylindrical fins-heat sink, and nanoparticles Al_2O_3. Three configurations were compared: PV system with heat sink, PV system with heat sink and PCM, PV system with heat sink, PCM and nanoparticles Al_2O_3. In the conclusion, RT35HC PCM with nanoparticles (0.77%) Al_2O_3 and heat sink was considered as optimum solution where 36.9–52.3% temperature drop was obtained on front surface of PV panel. Radziemska et al. [6] performed experiment study on PV/T system. Three phase change materials: paraffin wax42-44, RT22 [7], and ceresin were used along with water cooling. The results showed that it was more convenient and cheaper to apply modifications to the model with PCM without water cooling, which lowers temperature of module by 7 K.

Kant et al. [8] conducted heat transfer study on PV panel integrated with PCM RT35 [7]. The numerical model was solved by using Multiphysics 5.0 with finite element analysis. It was found that the cell temperature was reduced by maximum 6 °C for PV/PCM with convection mode of heat transfer, and 3 °C for PV/PCM system with only conductive heat transfer. Klemm et al. [9] developed a 3D model and solved in COSMOL Multiphysics. PCM fiber component was prepared from metallic fiber structure, in which the PCM was filled in voids. Three PCMs RT44HC, RT50, and RT54HC [7] were investigated. The optimum thickness for RT25HC with 1% porosity was 50 mm. Further, addition in thickness had no significance improvement.

Numerical and experimental study on hybrid BIPV/PCM system was performed by Aelenei et al. [10]. The melting range of PCM was 18–23 °C. The maximum thermal and electrical efficiency were achieved up to 12% and 10%, respectively. Mousavi et al. [11] used copper foam as porous medium to investigate the thermal performance of PV/T/PCM system. The CFD model was proposed to investigate five PCM (Paraffin C15, C18, C22, sodium phosphate salt, palmitic acid) as well as the effects of copper foam as the porous medium. In result, it was found that sodium phosphate salt had highest electrical efficiency 13.8%, while paraffin wax C22 had maximum thermal efficiency 83.5%. From literature survey, it is concluded that organic PCMs are widely used for PV/PCM system because they are easily available in market with large melting temperature range. Very less work has been carried out for pure fatty acid—capric acid with PCM bags. Thus, in this work, capric acid was used as phase change material, and it was encapsuled in vacuum bag. In Sect. 2, experiment set up and procedure is explained followed by Sect. 3 in which result of an experiment is discussed. At last, the conclusion is discussed.

2 Methodology

For experimental study, two regular PV panels, manufactured in Topsun Energy Ltd. were used. The specification of panel is described in Table 1. The front surface temperature and back surface temperature were measured by using non-contact type infrared thermometer, and the solar radiation was measured by using solar power meter. Current and voltage were measured by using digital multimeter.

2.1 Phase Change Material Selection

Phase change material can absorb heat energy as latent heat and release it when temperature drops below its melting temperature. The past work by researchers shows that organic phase change materials are generally employed due to their availability with wide range of melting temperature range and large volumetric heat. For selection of PCM, several properties are considered. The most important property is that it should be non-hazardous to handle with high melting enthalpy. Another important property is that the material should have a low expansion factor which makes simpler the calculations for the storage in PCM bag. The material's lifetime and maximum working temperature were also analyzed. The lifetime is most essential parameter because the PCM should not discontinue work before the end of the PV panel's lifetime. The maximum sustainable temperature is a curious factor to observe because it will stop working beyond this limit until the temperature drops again. Some important properties of capric acid are shown in Table 2.

2.2 Preparation of PCM Bags

To determine quantity of PCM, total heat energy which should be taken out from PV panel is calculated. This energy is assumed to be collected by capric acid in two ways: sensible heat and latent heat. If m is mass of PCM in Kg, L is latent heat in KJ/Kg, c_p is the specific heat in KJ/ Kg °C, and ΔT is temperature difference; then, following equation was used to calculate mass of PCM.

Table 1 Specification of PV panel

Model	TEL24P320
Rated maximum power (P_m)	320 W
Rated operating voltage (V_m)	37.20 V
Rated operating current (I_m)	8.61 A
Open circuit voltage (V_{OC})	45.28 V
Short circuit current (I_{SC})	9.20 A

Table 2 Thermo-physical properties of PCM

PCM	Capric acid
Type	Organic/fatty acid
Chemical formula	$C_{10}H_{20}O_2$
Density Kg/m^3	893
Temperature °C	32
Latent heat KJ/Kg	169
Specific heat KJ/Kg °C	2.761
Thermal conductivity W/mK	0.372

$$
\begin{aligned}
Q &= \text{Specific Heat} + \text{Latent Heat} \\
&= \mathrm{mc_p}\Delta T + mL \qquad (1)
\end{aligned}
$$

Capric acid was delivered in powder form. Hot water bath was used to melt capric acid. 200 gm (225 ml) of capric acid was filled in each PCM bags. Packing sealing machine was used to seal PCM bags, and total numbers of 68 bags were prepared. Extra bags were also prepared in case of damage to the original one. PCM bags were allowed to cool down for a period of night on the flat surface so that perfect and smooth surface could be obtained. The PCM bags with flat and smooth surface could be easily attached to PV panel surface.

2.3 *Experiment Setup and Procedure*

To get maximum solar radiation, solar PV panels were kept at latitude angle of Gandhinagar (23.5°) in the south direction. Total numbers of 68 PCM bags of capric acid were attached on the back side of one solar PV panel as shown in Fig. 1. Enough

Fig. 1 Experiment setup

space was left in between two PCM bags for temperature measurement. Thermal grease has higher thermal conductivity, so it was used to enhance heat transfer rate to PCM and to reduce air gap. Adhesive tape was used to attach PCM bags on PV panel back surface. Total eight temperature readings were taken for both front surface and back surface of panel to get more accurate temperature. Out of two panels, one was used as reference panel without any cooling method and other as panel with modifications. The readings were taken at every interval of 15 min from 9 am to 4 pm in month of February, 2020. For the calculation of efficiency of PV panel, fill factor was also included which was 0.78. The collected data were compared with reference panel.

3 Result and Discussion

In this section, detail discussion of experiment result is given. The experiment was performed on the terrace of LDRP-ITR, Gandhinagar. The altitude and longitude of Gandhinagar are 23.5° N and 72.6° E. The experiment was performed on February 11th, 2020. Figure 2 shows the solar radiation data during the day. As the time is rising, solar radiation is increasing and becomes maximum around 1:00 pm, and then again, solar radiation decreases.

Figures 3 and 4 show the temperature of front surface and back surface of the PV panel, respectively. For initial first hour, the front surface temperature and back surface temperature of PV panel were almost equal to the reference panel. But, as the solar radiation was raising, the front surface and back surface temperature was also increased in case of reference panel. The melting temperature of capric acid is 32 °C. As soon as the temperature of PV panel approached to melting temperature of capric

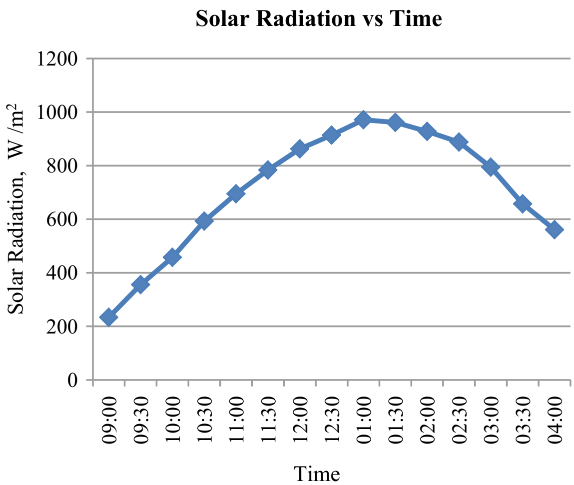


Fig. 2 Solar radiation versus time

Fig. 3 Front surface temperature versus time

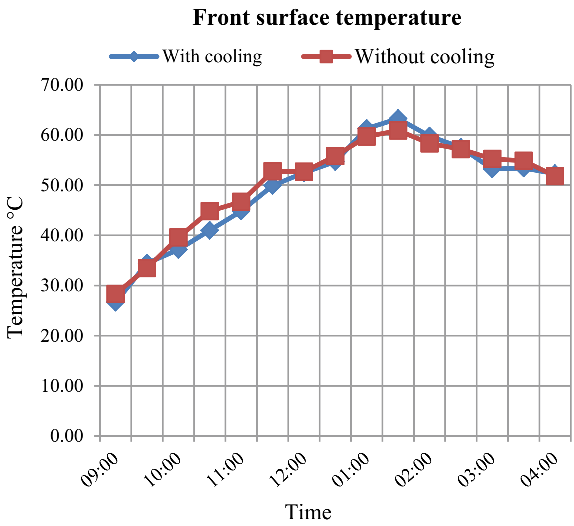
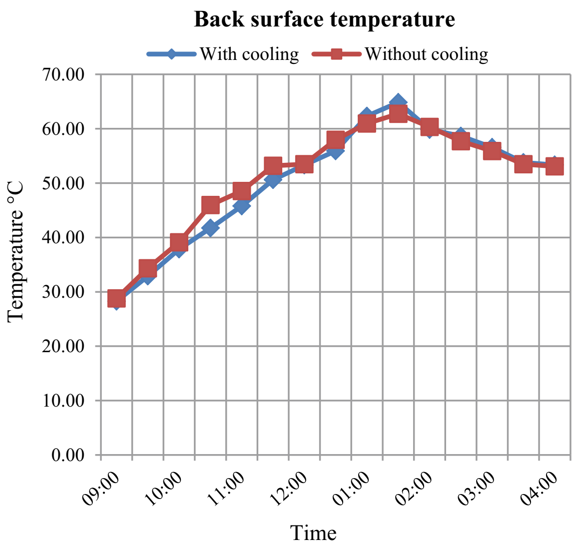


Fig. 4 Back surface temperature versus time



acid, it began to melt and reduce the surface temperature of PV panel. It is seen that for initial two to three hours, panel surface temperature is lesser than reference panel due to cooling effect provided by melting process of capric acid. Once the capric acid was completely melted, then it again raised surface temperature for last two to three hours.

Figures 5 and 6 show the data of power output and efficiency, respectively. As the solar radiation was increased with time, the power output for both panels was

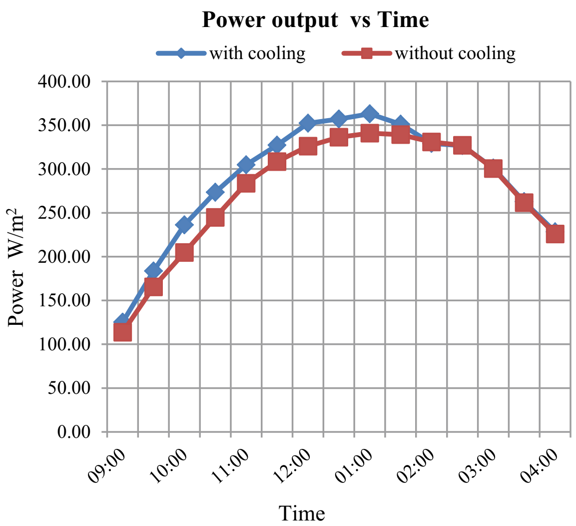


Fig. 5 Power output versus time

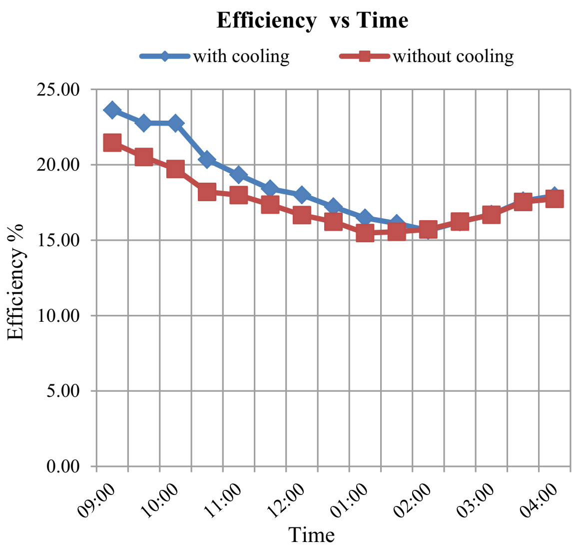


Fig. 6 Efficiency versus time

increased. For the panel with capric acid, power output was more as compared to reference panel due to lower surface temperature. After complete melting of capric acid, the power output was almost equal for both panels. The average power output of PV panel with cooling and reference panel was 288.12 W and 273.94 W.

As the solar radiation was increased with time, the efficiency of both panels was decreased due to increase in operating temperature. For the panel with capric acid, efficiency was higher as compared to reference panel due to lower surface

Table 3 Result table

Parameter	Capric acid	
	Average %	Maximum %
% decrease in back surface temperature	1.60	9.19
% decrease in front surface temperature	1.67	8.59
% increase in efficiency	5.80	15.39

temperature. After complete melting of capric acid, the efficiency was almost equal for both panels. The average efficiency of PV panel with cooling and reference panel was 18.61% and 17.54%.

Table 3 shows the maximum and average change in efficiency and temperature of the PV panel.

4 Conclusion

This experimental work contributes to the research area of cooling of solar PV panel by using capric acid as phase change material. Organic PCM is widely used for passive temperature control because it is available with wide melting temperature range. Due to proper distribution of PCM bags on the back surface of the panel, nearly, uniform cooling effect is obtained on entire surface area.

The performance of solar PV panel is improved due to PCM cooling by capric acid. The average and maximum % reduction in temperature of back surface is obtained by 1.60% and 9.19%, respectively. The average and maximum % increase in efficiency is obtained by 5.80% and 15.31%, respectively.

With PCM cooling by capric acid, it is observed that for the few hours the operating temperature of PV panel decreases due to cooling effect produced by melting process of capric acid. Once capric acid melts completely, it leads to increase the solar PV panel temperature as a result of reduction in heat extraction rate of PCM. Figure 7 shows comparison of electrical efficiency with different research work. For current work, highest average and maximum electrical efficiency has been achieved.

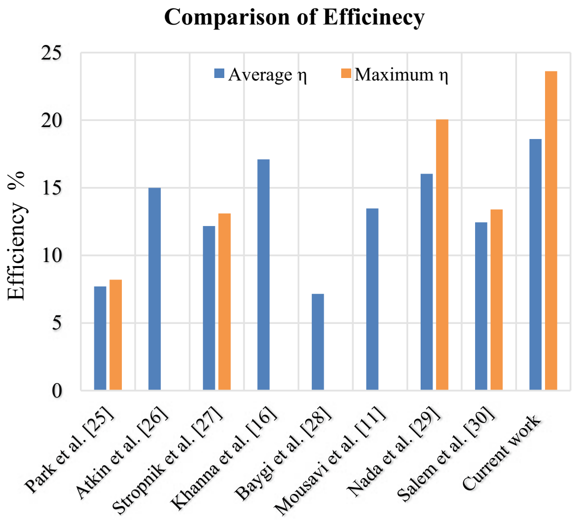


Fig. 7 Comparison of efficiency

Acknowledgements We would like to highly acknowledge the Government of Gujarat for partly funding the project under Laghu Sahay Yojana of Gujarat Energy Development Agency (GEDA), Climate Change Department. We would like to express our gratitude to Kadi Sarva Vishwavidhyalaya for providing us the remaining financial assistance under "Kadi Sarva Vishwavidyalaya—SVKM Student Project Grant no-005."

References

1. Zhang X, Zhaoa X, Smith S, Xu J, Yu X (2011) Review of R & D progress and practical application of the solar photovoltaic/thermal (PV/T) technologies. Renew Sustain Energy Rev 16:599–617
2. Parida B, Iniyan S, Goic R (2011) A review of solar photovoltaic technologies. Renew Sustain Energy Rev 15:1625–1636
3. Biwole P, Eclache P, Kuznik F (2011) Improving the performance of solar panels by the use of phase-change materials. In: World renewable energy congress. Sweden
4. Raziemska E (2003) The effect of temperature on the power drop in crystalline silicon solar cells. Renew Energy 28:1–12
5. Abdelrahman HE, Wahba MH, Refaey HA, Moawad M, Berbish NS (2019) Performance enhancement of photovoltaic cells by changing configuration and using PCM (RT35HC) with nanoparticles Al_2O_3. Sol Energy 177:665–671
6. Klugmann-Radziemska E, Wcisło-Kucharek P (2017) Photovoltaic module temperature stabilization with the use of phase change materials. Sol Energy 150:538–545
7. Rubitherm Phase Change Material. Retrieved from https://www.rubitherm.eu/en/index.php/ January 2020
8. Kant K, Shukla A, Sharma A, Biwole PH (2016) Heat transfer studies of photovoltaic panel coupled with phase change material. Sol Energy 140:151–161

9. Klemm T, Hassabou A, Abdallah A, Andersen O (2017) Thermal energy storage with phase change materials to increase the efficiency of solar photovoltaic modules. Energy Proc 135:193–202
10. Aelenei L, Pereira R, Gonçalves H, Athienitis A (2014) Thermal performance of a hybrid BIPV-PCM: modeling, design and experimental investigation. Energy Proc 48:474–483
11. Mousavi S, Kasaeian A, Behshad Shafii M, Jahangir MH (2018) Numerical investigation of the effects of a copper foam filled with phase change materials in a water-cooled photovoltaic/thermal system. Energy Convers Manage 163:187–195

Performance Enhancement of PV Panel by Cooling Front Surface of PV Panel with the Use of Water as a Cooling Medium

Pranav Patel, Pavit Shah, and Alkesh Mavani

Abstract This paper introduces a cooling system in a commercial photovoltaic panel using water to examine the improved output through a reduced operating temperature. It presents an alternative cooling technique for photovoltaic (PV) panels that include a water flow over panel surfaces. Solar radiation and operating temperature are two main parameters that affect the effectiveness of a photovoltaic panel. For these conditions, the electrical efficiency of the solar panel will be degraded as the operating temperature of the solar panel rises. Water flowed over the panel at a constant flow rate. The average temperature fall of the front and back surfaces is 3.54 °C and 2.79 °C, respectively, mainly the front water flow over the solar panel. Front cooling provides a 9.64% enhancement inefficiency on average. The average temperature fall of the front and back surfaces is 3.54 °C and 2.79 °C, respectively, mainly to front water flow over the solar panel. Front cooling provides a 9.64% enhancement in efficiency on average.

Keywords Efficiency · PV module · Temperature drop · Water cooling

1 Introduction

While the globe struggles with energy scarcity, global warming, and the degradation of energy and the environment, there is a need for green technologies to substitute fossil fuels, water, and other natural resources. PV (photovoltaic solar energy) is one of the alternative energy sources. The photovoltaic solar cell, as its name implies, is a semiconductor device that creates energy when exposed to light. A photovoltaic cell changes only a small portion of the irradiance into electrical energy (less than 20%). PV cells' electrical efficiency diminishes as the temperature goes up [1–25]. The

P. Patel (✉)
LDRP-ITR, Kadi Sarva Vishwavidyalaya, Gandhinagar, Gujarat, India
e-mail: pranavpatel1119@gmail.com

P. Shah · A. Mavani
Mechanical Engineering Department, LDRP-ITR, Kadi Sarva Vishwavidyalaya, Gandhinagar, Gujarat, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_12

power produced of conventional flat-panel PV modules can be enhanced by cooling because it stops the PV cells from reaching temperatures that induce significant harm. The effectiveness and power output of a PV module is found to be inversely proportional to temperature [13–20]. Solar energy is just one of many alternative energy sources accessible. Solar energy is an almost endless source of energy. Solar power has the ability to supply all of the world's present and future energy needs on a consistent basis. As a result, it is one of the most viable energy sources. The sun's power absorbed by the earth is around 1.81×10^{11} MW, which is greater than the current consumption rate of all commercial energy sources on the planet [10]. As a result, solar energy may provide all of the world's current and future energy needs on an ongoing basis. Cooling can enable typical flat-panel PV modules to create more electricity because it prevents the solar cells from reaching the temperature where lasting damage occurs.

Effect of Operating Temperature on Solar Panel:

Because of the absorption and reflection caused by the glass cover and PV module in conventional photovoltaic (PV) setups, over 80% of solar radiation is transformed into heat, resulting in a massive gain in PV cell operating temperature [2]. Like all other semiconductor devices, solar cells are sensitive to temperature. Increasing temperature diminishes the bandgap of a semiconductor, altering many of the properties of the material. The decrease in a semiconductor's bandgap with rising temperatures can be seen as increasing the energy of the material's electrons. Therefore, it takes less energy to sever the bond. The reduction of bond energy in the bond model of a semiconductor also decreases the bandgap. Raising the temperature, therefore, reduces the difference in the band. Within a solar cell, the open-circuit voltage is the parameter most influenced by an increase in temperature. The P–V characteristic is the interaction between the electrical power consumption of a solar PV module, P, and the output voltage, V, while solar radiation, E, and cell temperature, Tm, stay unchanged. If any of those two elements, Tm or E, alters, then all features change as well. The effect of temperature increase is shown in Fig. 1.

2 Objectives

The major goal of this project is to investigate passive water cooling systems in order to improve efficiency. To conduct experimental research on passive water cooling settings in order to improve panel efficiency and performance.

3 Research Gap

According to the literature review, water cooling is the most effective approach for panels [1–25]. Only a few researchers have experimented with passive water cooling.

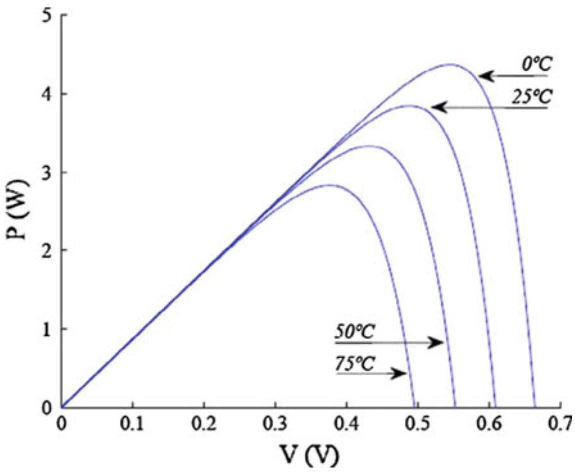


Fig. 1 Temperature variation P–V characteristics across 0° and 75 °C [3]

We can reduce the operating temperature of PV panels and increase their efficiency at a lower cost by using an effective passive water cooling technology. Flowing water on the front surface can be a good technique to keep cool.

4 Methodology

In this section, experimental setup and procedure are discussed. A non-contact type infrared thermometer was used to measure the front and back surface temperatures, and a solar power meter was used to monitor the sun's radiation. For the experimental study, two regular PV panels, manufactured in Topsun Energy Ltd. were used. The panel's specifications are listed in Table 1. Current and voltages were measured by using a digital multimeter.

The solar panel should be at an angle that is equal to the latitude of the location where it is installed. For Gandhinagar, the latitude is 23°, so the panel was kept at 23° of inclination angle, and then, the experiment was performed.

Table 1 PV panel specifications

Model	TEL24P320
Rated maximum power (P_m)	320 W
Rated operating voltage (V_m)	37.20 V
Rated operating current (I_m)	8.61 A
Open-circuit voltage (V_{OC})	45.28 V
Short-circuit current (I_{SC})	9.20 A

4.1 Experiment Setup and Procedure

For getting maximum solar radiation, solar PV panels were kept at the latitude angle of Gandhinagar (23°) in the south direction. We have used a PVC pipe of 2-m length and 1-inch diameter and fixed it above the panel. We have made 40 holes of 1.5 mm diameter on the surface of that pipe in a straight line. These numbers of holes were required for water uniformity on the front surface of the panel. So for every cell column on the panel, there are three holes provided. The 500 L water tank was used as storage tank. The valve is used for water flow control, and a 2 m long pipe is used for the water supply on the panel. The readings are taken at every half-hour interval. The water was supplied over the surface 5 min before reading and flowed for only 10 s every time. The water flow rate was constant at a rate of 8 ml/s. The big size PVC pipe pieces are used as collectors under the panel to make sure there is no water wastage. But there is a slight loss of water due to evaporation because of high surface temperature. To acquire a more precise temperature, eight temperature readings were taken for both the front and back surfaces of the panel. One of the two panels was utilized as a reference panel with no cooling mechanism and the other as a modified panel. The fill factor, which was 0.78 for the solar panel, was also taken into account when calculating the PV panel's efficiency (Fig. 2).

5 Results and Discussion

In this section, detailed data and observations of this experimental research are given. The experiment was performed on the terrace of LDRP-ITR, Gandhinagar. The reading was taken from 9:00 A.M to 4:00 P.M. The observation table was prepared which is given below. A detailed discussion is given below.

As can be seen in Figs. 3 and 4, the PV panel's front and back surface temps were nearly identical to the reference panel in the early hours. In the case of the reference panel, however, as the solar energy increased, the front and back surface temperatures increased as well. Front cooling via water flow improves the performance of the solar PV panel. 3.54 °C is the average difference in temperature of the front surface. Due to the cooling effect, the panel surface temperature is lower than the reference panel for the first two to three hours. After that, due to high front surface temperature, the difference of temperatures between with and without cooling is lesser at 12 P.M and 2 P.M.

The solar radiation increased with time and was maximum between 1:00 and 2:00 PM. The average difference in back surface temperature is obtained by 2.79 °C. But from Fig. 4, we can see that the PV panel with front cooling gives a lower back surface temperature than without the cooling panel.

The power output of the PV panel grows in tandem with the rise in solar radiation over time. The results for power output and efficiency are shown in Fig. 5 and 6. The power output of both panels rose as solar radiation increased over time. Because the

Fig. 2 Experiment setup

front-cooled panel had a lower surface temperature, it produced more power than the reference panel. 308.67 W and 274.54 W are the average power outputs of a PV panel with cooling and a reference panel, respectively.

The efficiency of solar panels with cooling is always better than that of solar panels without cooling, as shown in Fig. 6.

The panel efficiency is highest during the initial hours, as shown in the graph. However, when the working temperature rises over time, so does the efficiency. However, the reference panel's average efficiency is 17.43%, whereas the panel with front cooling has a 19.11% efficiency.

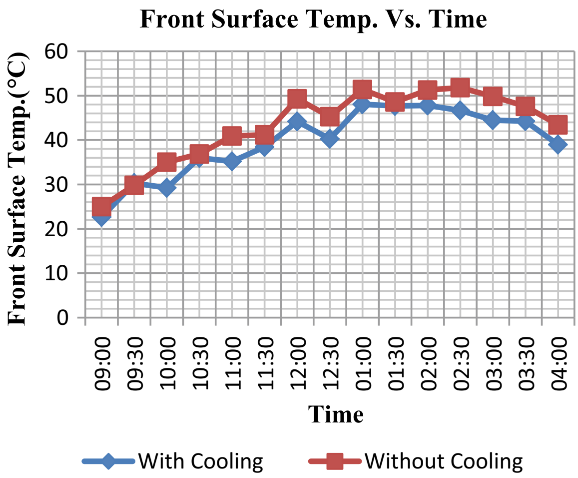


Fig. 3 Front surface temperature versus time

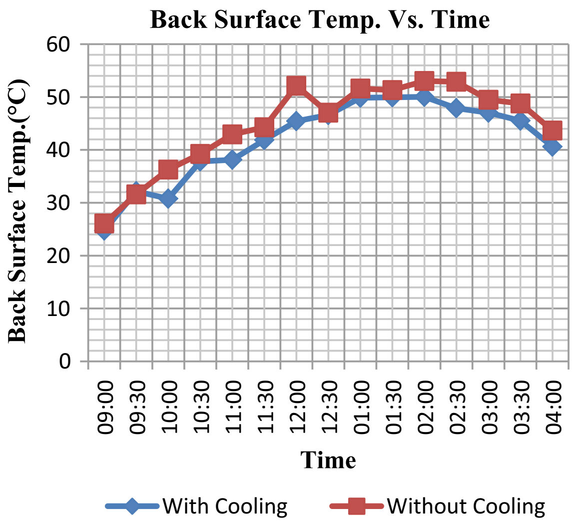


Fig. 4 Back surface temperature versus time

6 Conclusion

Finally, it is feasible to assert that the suggested water cooling approach has a positive impact on the performance of PV panels, which is advantageous in PV applications. Due to the front water flow, a nearly uniform cooling effect is obtained. The available solar radiation at the site for the reference panel varies from 245 to 1001 W/m^2, while for cooling panel varies from 315 to 998 W/m^2. The results showed a significant decrease in the temperature of the PV panel.

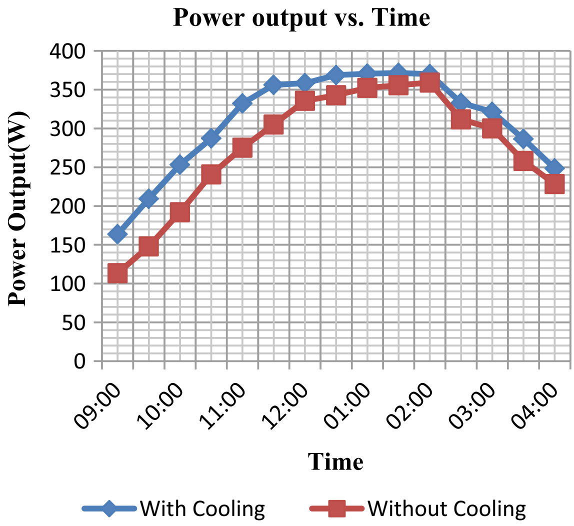


Fig. 5 Power output versus time

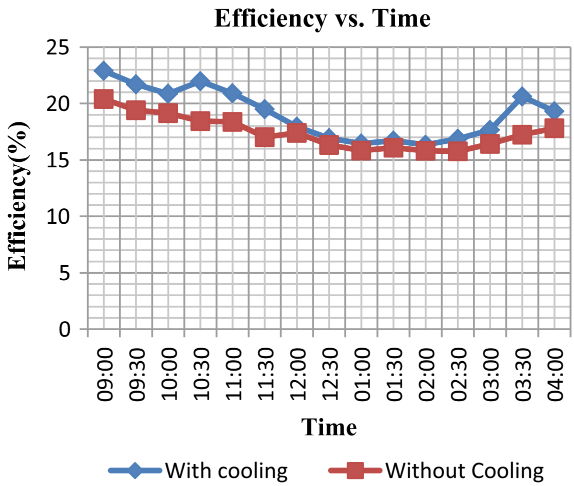


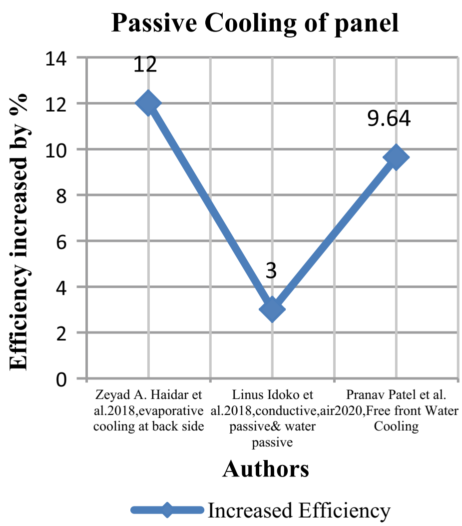
Fig. 6 Efficiency versus time

Front cooling via water flow enhances the performance of the solar PV panel. The front surface seems to have a temperature reduction of 3.54 °C, whereas the back surface has a temperature reduction of 2.79 °C.

The average percentage increase in efficiency and power achieved by front cooling is 9.64% and 12.44%, respectively. No water is wasted due to the use of water collectors (Fig. 7).

When enhancing electricity efficiency is the primary goal, water cooling has been proved to be an excellent option. As a result, the future research goal should be to achieve effective water cooling of PV panels.

Fig. 7 Comparison of passive cooling results with other findings obtained using different methods used by various authors



Acknowledgements We highly acknowledge the Government of Gujarat for partly funding the project under Laghu Sahay Yojana of Gujarat Energy Development Agency (GEDA), Climate Change Department. We are also grateful to Kadi Sarva Vishwavidyalaya for providing us the remaining financial assistance under "Kadi Sarva Vishwavidyalaya—SVKM Student Project Grant no-006."

References

1. Benato A, Stoppato A (2019) An experimental investigation of a novel low-cost photovoltaic panel active cooling system. Energies 12(8):1448
2. Dubey S, Sarvaiya J, Seshadri B (2013) Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world—a review. Energy Proc 33:311–321. https://doi.org/10.1016/j.egypro.2013.05.072
3. Bijjargi Y, Kale SS, Shaikh KA (2016) Cooling techniques for photovoltaic module for improving its conversion efficiency: a review 7(4):22–28. Article ID: IJMET_07_04_003
4. Moharram K, Abd-Elhady M, Kandil H, El-Sherif H (2013) Enhancing the performance of photovoltaic panels by water cooling. Ain Shams Eng J 4(4):869–877
5. Matias C, Santos L, Alves A, Calixto W (2017) Increasing photovoltaic panel power through water cooling technique. Trans Environ Electr Eng 2(1)
6. Nižetić S, Čoko D, Yadav A, Grubišić-Čabo F (2016) Water spray cooling technique applied on a photovoltaic panel: the performance response. Energy Convers Manage 108:287–296
7. Schiro F, Benato A, Stoppato A, Destro N (2017) Improving photovoltaics efficiency by water cooling: modelling and experimental approach. Energy 137:798–810
8. Krauter S (2004) Increased electrical yield via water flow over the front of photovoltaic panels. Sol Energy Mater Sol Cells 82(1–2):131–137. https://doi.org/10.1016/j.solmat.2004.01.011
9. Irwan Y et al (2015) Indoor test performance of PV panel through water cooling method. Energy Proc 79:604–611. https://doi.org/10.1016/j.egypro.2015.11.540
10. Milind N, Antony M, Francis F, Francis J, Varghese J, Sajith UK (2017) Enhancing the efficiency of solar panel using cooling systems. Int J Eng Res Appl 07(03):05–07

11. Patil M, Sidramappa A, Angadi R (2018) Experimental investigation of enhancing the energy conversion efficiency of solar PV cell by water cooling mechanism. IOP Conf Ser Mater Sci Eng 376:012014. https://doi.org/10.1088/1757-899x/376/1/012014
12. Haidar Z, Orfi J, Kaneesamkandi Z (2018) Experimental investigation of evaporative cooling for enhancing photovoltaic panels efficiency. Results Phys 11:690–697
13. Zubeer S, Mohammed H, Ilkan M (2017) A review of photovoltaic cells cooling techniques. E3S Web Conf 22:00205. https://doi.org/10.1051/e3sconf/20172200205
14. Rahman M, Hasanuzzaman M, Rahim N (2015) Effects of various parameters on PV-module power and efficiency. Energy Convers Manage 103:348–358. https://doi.org/10.1016/j.enconman.2015.06.067
15. Siecker J, Kusakana K, Numbi B (2017) A review of solar photovoltaic systems cooling technologies. Renew Sustain Energy Rev 79:192–203. https://doi.org/10.1016/j.rser.2017.05.053
16. Sargunanathan S, Elango A, Mohideen S (2016) Performance enhancement of solar photovoltaic cells using effective cooling methods: a review. Renew Sustain Energy Rev 64:382–393. https://doi.org/10.1016/j.rser.2016.06.024
17. Haidar Z (2016) Cooling of solar PV panels using evaporative cooling. J Therm Eng 2(5). https://doi.org/10.18186/jte.72554
18. Abdulgafar SA, Omar OS, Yousif KM (2014) Improving the efficiency of polycrystalline solar panel via water immersion method. Int J Innov Res Sci Eng Technol 3(1). ISSN: 2319-8753
19. Özkul F, Kayabasi E, Çelik E, Kurt H, Arcaklioğlu E (2018) Comparison of different cooling options for photovoltaic applications. 978-1-5386-7538-0/18/$31.00 ©2018 IEEE
20. Grubišić-Čabo F, Nižetić S, Giuseppe Marco T (2016) Photovoltaic panels: a review of the cooling techniques. ISSN 1333-1124, eISSN 1849-1391
21. Shukla A, Kant K, Sharma A, Biwole P (2017) Cooling methodologies of photovoltaic module for enhancing electrical efficiency: a review. Sol Energy Mater Sol Cells 160:275–286. https://doi.org/10.1016/j.solmat.2016.10.047
22. Sainthiya H, Beniwal NS (2017) Different types of cooling systems used in photovoltaic module solar system: a review. In: IEEE WiSPNET 2017 conference, 978-1-5090-4442-9/17/$31.00_c 2017 IEEE
23. Balamuralikrishnan B, Deepika B, Nagajothi K, Shubaa shree S, Subasini PT (2014) Efficiency enhancement of photovoltaic cell. IJAREEIE 3(4)
24. Hachicha AA, Ghenai C, Hamid AK (2015) Enhancing the performance of a photovoltaic module using different cooling methods. Int J Energy Power Eng 9(9)
25. https://firstgreenconsulting.wordpress.com/2013/05/04/effect-of-temperature-on-module-performance

Drinking Water Quality of Different Industrial Sources

(Environmental Science and Engineering, Urban Planning)

Pallavi Praveen, Tejaswini Singh, and Harshavardhana Singh

Abstract The quality of different drinking water sources at Chas (Bokaro) which is prone to different waterborne diseases was assessed in the present study. It was observed that several important constituents have crossed the permissible limit suggested by the World Health Organization and the Indian Council of Medical Research.

Keywords Drinking water sources · Bokaro · Waterborne diseases · WHO · Pathogenic microorganisms

1 Introduction

Drinking water is defined as water free from undesirable odours, flavours and pathogenic microorganisms and chemicals that are harmful to human health, i.e. water which is pure, wholesome and state for human consumption [1]. The World Summit on Sustainable Development Aid in Johannesburg set a target of having by the year 2015 the proportion of the people who access safe drinking water in India. The major sources of drinking water for human beings are river streams, well, hand pump municipal water supply, etc. Due to this, a large number of people get affected by enteric diseases every year. WHO estimates that about 80% of all diseases are caused due to lack of safe drinking water and sanitation in the recent past. There have been sensutbreak of such diseases in different parts of Chas Bokaro Drinking water contains ron cabum, magnesium, manganese silica fluoride, nitrate potassium,

P. Praveen (✉)
Bokaro Steel City College, Bokaro, India
e-mail: pallavipraveen8@gmail.com

T. Singh · H. Singh
Birla Institute of Technology, Ranchi, India
e-mail: be10049.18@bitmesra.ac.in

H. Singh
e-mail: be10056.19@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_13

etc [2, 3]. When the quantity of any of these components increases either by abiotic acutis then such water affects health waters consumption of the body systems and pose great problems for human animals. Keeping in consideration an effort has been made to evaluate the drinking water quality of three Parc 09 des common drinking water sources at the Chas Town. They include Hand 45 feet deep which are situated boring water 300 feet deep and well water Crash Main Road Ram Nagar Colony and Jodhadih More areas of Chas Township, respectively.

2 Materials and Methods

Water samples from hand pipe, boring and well water were collected in polythene bottles for the estimation of different chemical constituents in the laboratory. Water temperature was recorded by mercury centigrade thermometer pH was determined by pH metre free CO, and alkalinity was determined by the method suggested by Welch [4] Dissolved (1985) were followed for the estimate of different chemical constituents. Biological oxygen demand was measured by incubating the sample at 20 °C for five days.

3 Result and Discussions

The data of three different drinking water sources have been depicted Table 1. The water temperature was 25 °C, 28 °C and 25.5 °C in the hand pump, boring and well water, respectively. Turbidity value was 20 NTU different sources, respectively, which are below the 10 NTU and 20 NTU in these three upper limits of WHO (25 NTU) [2]. The dissolved solids of boring and well water were 300 ppm and 490 ppm which are below the permissible limit of WHO [4, 5] standard (i.e. 500–1500 ppm), whereas hand pump value was found to be 510 ppm, pH was 72 in the tube well water and 71 in both the bong and well water [1, 6]. The concentration of dissolved O_2 was 5 ppm. 4.8 ppm and 6.8 ppm, respectively. Higher concentration of dissolved O_2, observed in well water might be due to low temperature which enhances the oxygen dissolving capacity of water. Percentage O_2 saturation was maximum in well water (83%) whereas it was 62% in hand pump water and 61% in boring water. The value of free CO_2 was 16.0 ppm in the hand pump water 70 ppm in boring water and 110 ppm in well water. Lower value (70 ppm) of boring water might be due to lower respiratory activities with respect to other two sources (Figs. 1 and 2).

Table 1 Water composition index

Parameters/sources	Hand pump water (100-feet deep) Chas road	Boring water (300-feet deep) Ram Nagar colony	Boring water (300-feet deep) Chas block
Water temp (C)	25	28	25.5
Turbidity (N.T.U.)	20	10	20
T.D.S. (ppm)	510	300	490
pH	7.2	7.1	7.1
DO_2	5.0	4.8	6.8
% O_2 saturation	62	61	83
Free CO_2 (ppm)	16.0	7.0	11.0
Total alkalinity (ppm)	294	92	152
Total hardness (ppm)	370	48	100
Ca-hardness (ppm)	336	30	76
Ca (ppm)	134.4	12.0	30.4
Mg (ppm)	8.39	4.48	5.93
Chloride (ppm)	10.25	0.75	6.10
Silicate (ppm)	62.5	55.5	55.6
NO_3N (ppm)	2.68	1.17	1.78
PO_4 (ppm)	Nil	Nil	0.01
Na (ppm)	63.2	36.0	54.0
K (ppm)	6.0	1.5	1.5
B.D.O (ppm)	3.2	2.1	4.3

Hydrological properties of three different drinking water sources in Chas

4 Conclusion

Total hardness value was 370 ppm, 48 ppm and 100 ppm, whereas calcium hardness was 336 ppm, 30 ppm and 76 ppm in the hand pump, boring and well water, respectively. Higher value of hardness in hand pump water is due to higher concentration of calcium and magnesium salts with respect to boring and well water and the present total hardness value of hand pump water (370 ppm) has crossed the range of maximum permissible standard of ICMR. The concentration of calcium was 134.4 ppm, 120 ppm and 30.4 ppm, whereas magnesium value was 8.39 ppm, 4.48 ppm and 5.93 ppm in hand pump, bong and well water, respectively. Chloride served as a basis for detecting pollution of groundwater by sewage Chloride value was 10.25 ppm 0.75 ppm and 6.10 ppm and silicate showed 62.5 ppm, and 55.6 ppm concentration in the hand pump, boring and well water, respectively (Fig. 3).

Higher value of nitrate nitrogen (268 ppm) is boring and well water, respectively. The value is below the maximum permissible limit of ICMR (i.e. 20 ppm). Phosphate–phosphorus was absent in the hand pump water in comparison with boring and

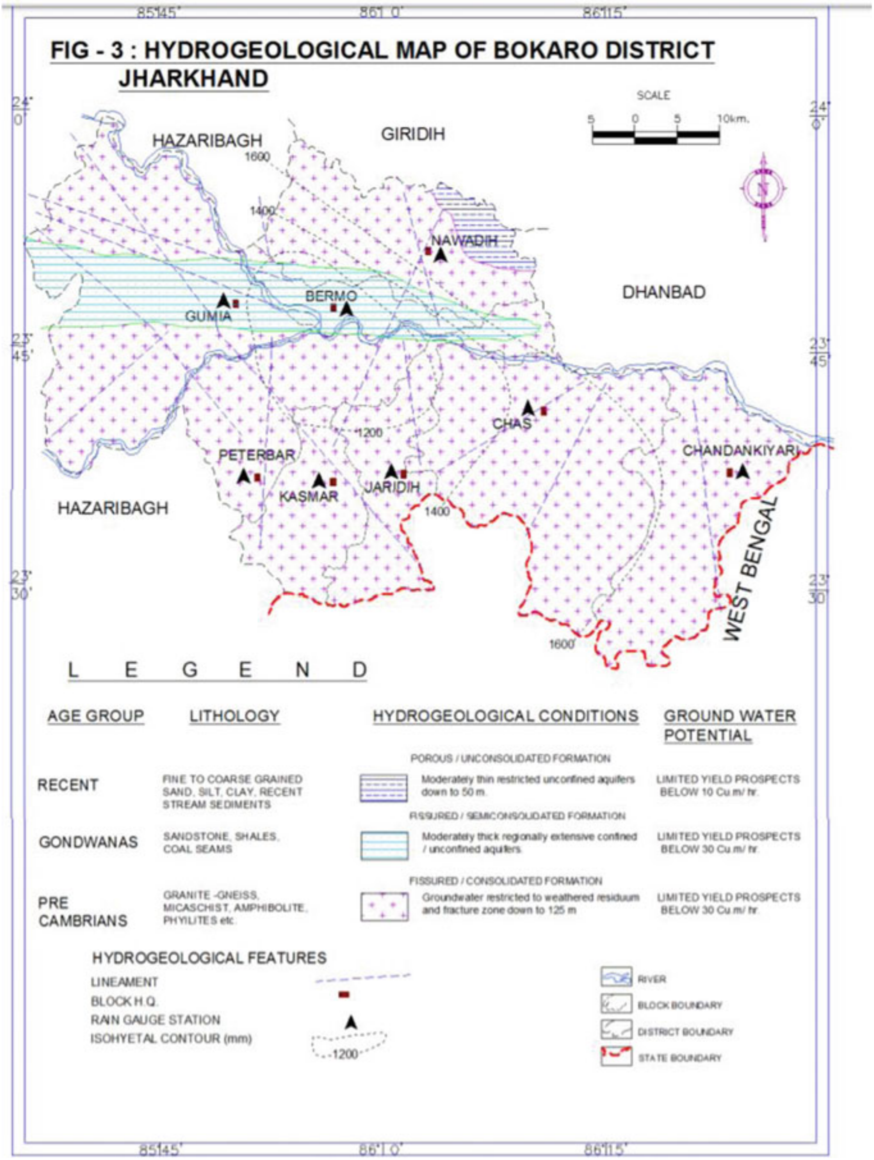


Fig. 1 Hydrological map. *Source* Hydrological information booklet, Bokaro District, Jharkhand

well water. The biological oxygen demand was found to be 3–2 ppm and 4–3 ppm in hand pump, boring and well water, respectively. Lower value of BOD in boring water may possibly be due to lower bacterial population with respect to other two sources.

Fig. 2 Bokaro District water supply. Source of drinking water for Bokaro District

Fig. 3 Factory waste disposal. High volume of chemical and sewage waste disposal in Damodar River

Acknowledgements The authors are thankful to Chas Municipality, Chas; and for laboratory facilities provided by Bokaro Steel Plant, Bokaro Steel City.

References

1. APHA (1985) Standard methods for the examination of water and wastewater
2. Groundwater information booklet, Bokaro District, Jharkhand
3. Chas-Bokaro Water Project, District Profile, Govt. of Jharkhand
4. Welch PS (1952) Limnology. McGraw Hill, New York
5. BIS (1991) Drinking water—specification IS 10 500. International Standards for drinking water (1984) Issued by World Health Organisation Geneva
6. Manual of Standards of Quality of Drinking Water Supplies (1971)—issued by Indian council of Medical Research (ICMR). New Delhi

Design of Low-Power High-Efficient Single-Tail Comparator Using 180 nm CMOS Technology

Rahul Nagore, Pramod Kumar Jain, R. S. Gamad, and Rahul Priyadarshi

Abstract Promoting dynamic single-tail comparators to optimize input speed and productivity is the clandestine to low-cost, area-efficient, high-speed analog to digital converters. An overview of the functional comparators' low power will be discussed in this article, and theoretical expressions obtained modern dynamic comparator is introduced based on the provided research, in which circuit of a standard single-tail comparator is adjusted for little power and debauched action level in limited voltages of supply. Deprived of making simple design complicated and adding some transistors, we present functional analysis of a single tail comparator. Speed and power consumption are both important parameters. The full circuit supply required 1.8 V. Use 0.18 μm technologies, Cadence modeling methods are used to conduct simulation analyzes at the comparator.

Keywords Single-tail comparator · Cadence · Spice · ADCs

1 Introduction

Comparator is one of the major design squares of commonly easy-to-automate (ADC) converters. Some high-speed ADCs instances incorporate low-power, fast comparators with a confined chip region High-speed comparators in really significant sub-micrometer (UDSM). CMOS advancements experience the evil impacts of defenseless stock voltages, explicitly, considering the way that the contraption edge voltages were not scaled at a comparative rate as the standard CMOS process supply voltages [1]. Therefore, quick plan comparators become truly testing when the yield voltage is more humble. In various terms, to show up at a quick in a given development, greater semiconductors are critical to address the reducing in supply

R. Nagore (✉) · P. K. Jain · R. S. Gamad
Department of Electronics and Instrumentation Engineering, SGSITS Indore, Indore, Madhya Pradesh, India
e-mail: rahulnagore38@gmail.com

R. Priyadarshi
Department of Electronics and Communication Engineering, NIT Patna, Patna, Bihar, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_14

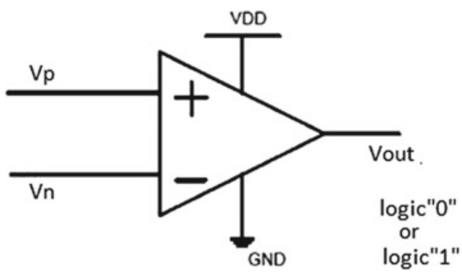
voltage, which consistently surmises more kick the pail locale, and power is required. Moreover, low-voltage development achieves a little data extent of ordinary mode, which is essential for some high-speed ADC structures, for instance, streak ADC. A few strategies have been created to meet the necessities of low-voltage supply, for example, supply helping techniques [2, 3], body-driven semiconductor procedures [4, 5], current-mode plan [6], and double oxide processes that can permit more prominent stockpile voltages. Supporting and bootstrapping are two methods centered on further developing the voltage source, reference, or clock to determine input reach and exchanging issues. Such are single reason in any case, especially in UDSM CMOS advances, they cause unwavering quality issues. The body-driven strategy presented by Block [4] diminishes the edge voltage measure so that body-driven MOSFET goes about as an instrument of exhaustion structure. A 1-bit quantize for sub-1 V modulators is proposed in [5] dependent on this technique. Given the advantages, the body-driven semiconductor experiences less Tran's conductance (equivalent to the semiconductor gmb) comparative with its entryway-driven identical in light of the fact that diverse assembling techniques, for example, profound n-well, permit both nMOS and pMOS semiconductors to work in body-driven arrangements. Thinking about this present methodology's viability, the effect of gadget crisscross in extra hardware on comparator yield ought to be considered [7].

2 Comparator Operation

*C*omparator can take two analog efforts and has one digital yield. Presume, two analog inputs are Vp and Vn and the optical output is Vo. VDD is the voltage of the DC voltage (Fig. 1).

Assuming Vp's ability is more prominent than Vn's, the yield Vo has reasoning 1. At the point when Vp is not exactly Vn's ability, then at that point, yield Vo has reasoning 0. The comparator utilized here just works on the distinction of info voltages and does not work if the information voltages are equivalent. A heartbeat voltage is applied at Vp for this differentiation, and a similar DC voltage is applied in Vn. Vout got a rationale 1 worth if a non-feedback comparator, which is considered an Op-Amp A regenerative comparer measures the frequency of two inputs using

Fig. 1 Comparator



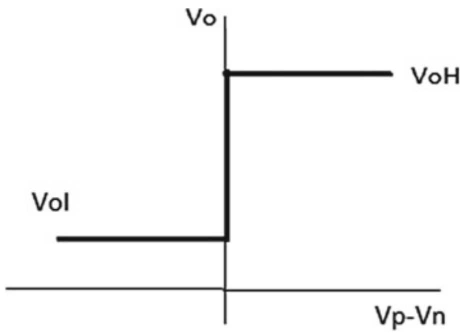


Fig. 2 Characteristics of the ideal comparator

constructive feedback; single-tail comparator has a transistor tail friendly to the circuit clock Op-amps. Regenerative comparator and single-tail comparator input are the most commonly used comparators (Fig. 2).

3 Existing Comparators

Obvious sturdy positive feedback in recreating latch, clocked recreating comparators have seen large requests in countless high-speed ADCs because they may take reckless verdicts. Existing comparators are single-tail comparators and traditional singular tail dynamic comparators.

4 Single-Tail Comparator

The single-tail comparator circuit chart is seen in Fig. 3. It is essentially utilized in A\D converters, with high information impedance, no dispersal of static control, and swing rail-to-rail yield. Utilizing two phases, the comparator activity can be explained. This circuit works in the reset venture by utilizing CLK = 0. During this progression, the Mtail semiconductor is taken off, and semiconductors (M7 and M8) are reset to all yield hubs Out and Out to VDD to show a beginning state and to have a legitimate intelligent level during this stage. At the point when VINP < VINN, the circuits work the other way around the combined deferral for the single-tail comparator as per the following: tdelay = t0 + tlatch t0 mirrors the hooking postponement of the release blunder T-lock delay.

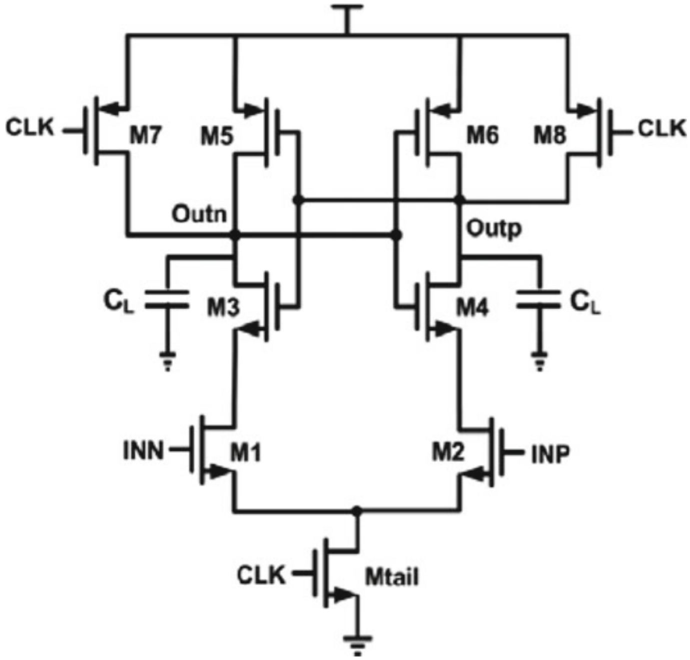


Fig. 3 Single-tail comparator

5 Proposed Circuit

The key principle of the theoretical high-speed comparator is to upsurge speed of the latch replication to improve speed of the bolt. For this purpose, parallel to M3 or M4 transistors, two control transistors (Mc1 and Mc2) were attached to primary phase. Many technical problems have to be addressed when constructing the planned comparator. Time it receipts for one of control transistors to turn on necessity be lesser than t0 when influential size of the tail transistors (Mtail1 and Mtail2) (Starting bid generation) (Fig. 4).

$$t_{\mathrm{on,Mc1(2)}} < t_0 \rightarrow \frac{|V_{\mathrm{Thp}}| \cdot C_{L,\mathrm{fn}(p)}}{I_{n1,2}} < \frac{V_{\mathrm{Thn}} C_{L\mathrm{out}}}{I_{B1}}$$
$$\rightarrow \frac{|V_{\mathrm{Thp}}| \cdot C_{L,\mathrm{fn}(p)}}{\frac{I_{\mathrm{Tail1}}}{2}} < \frac{V_{\mathrm{Thn}} \cdot C_{L\mathrm{out}}}{\frac{I_{\mathrm{Tail2}}}{2}}.$$

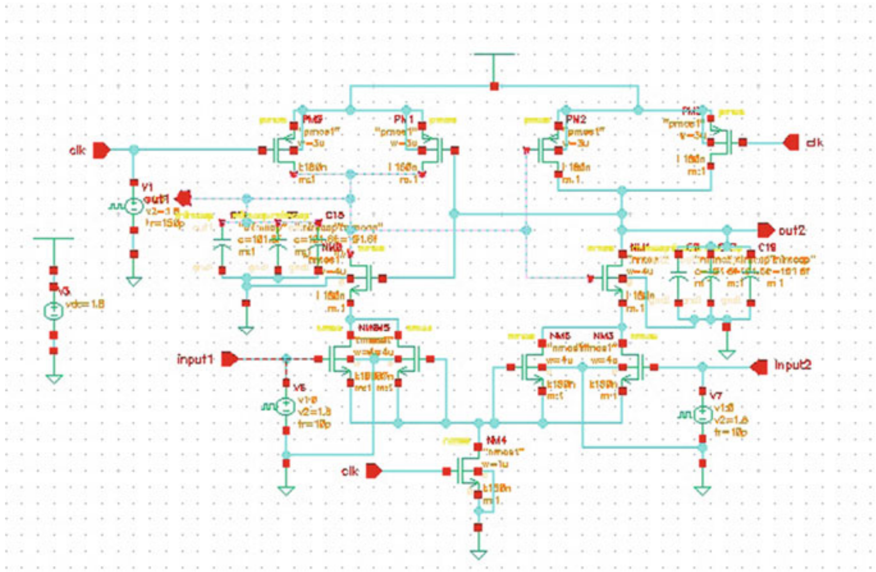


Fig. 4 Schematic design of single tail comparator

6 Symbol Creation

Schematic is turned into a figure representing circuit diagram. This symbol is used in construction of test benches to produce necessary [7] (Fig. 5).

7 Test Form Plan

Symbol is encompassed in test bench field, and appropriate pins are given as input and output. A voltage of 1 V DC is applied to VDD, other end ashore. The inputs are given clock pulses, and circuit is simulated to obtain yield waveforms [7] (Fig. 6).

8 Results and Simulations

It is possible to build the 180 nm high-speed comparator in T-spice applications. Various parameters can be evaluated using this program (Table 1).

Identify applicable funding agency here. If none, delete this. The transient analysis of the outputs from 180 nm as shown in Fig. 7 and then Fig. 8. Transitory analysis describes operational principles (Figs. 9 and 10).

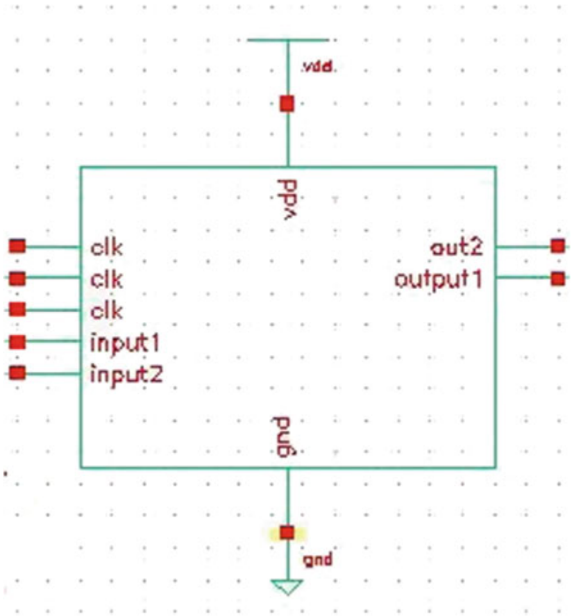


Fig. 5 Symbol of a single tail

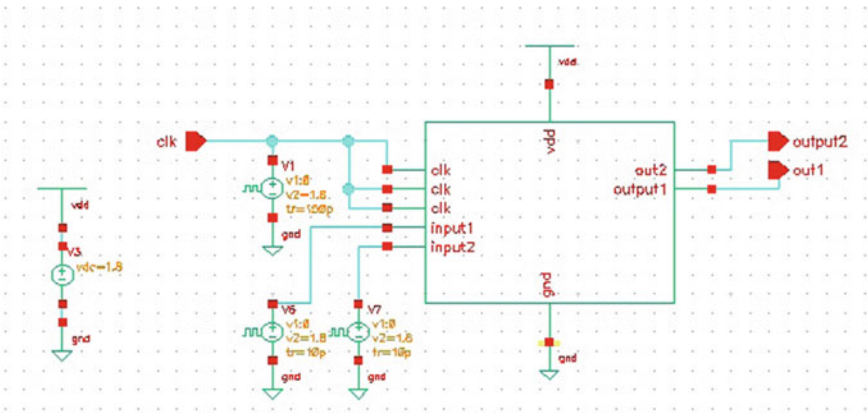


Fig. 6 Test bench

9 Conclusion

The 180 nm high-speed comparator was constructed by means of imitation software. In elevation, speed here may be reached by reducing the delay. This dissertation provides the study of the delays for clocked dynamic comparators. Single-tail comparator systems are analyzed. A modern high-speed comparator with little power

Table 1 Results for contrast

Comparator structure	Conventional dynamic comparator (2)	Conventional dynamic comparator (1)	Proposed dynamic comparator
Technology (nm)	180	180	180
Power supply voltage (V)	1.8	1.2	1.8
Power dissipation (μW)	244	329	209
Delay/log (*V*in) (p)	959.55	940	922.4

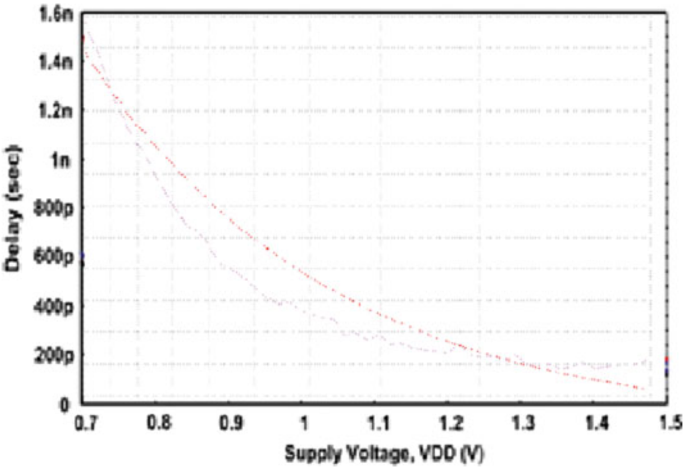


Fig. 7 Delay of the proposed comparator supply voltage (VDD)

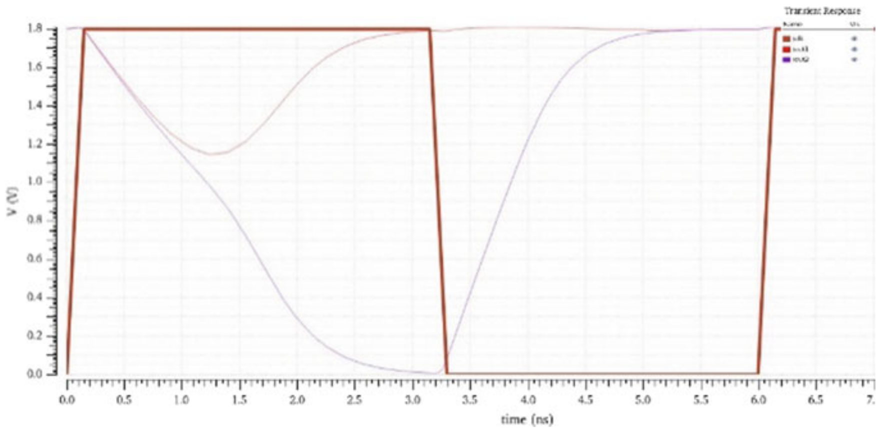


Fig. 8 Transient simulations

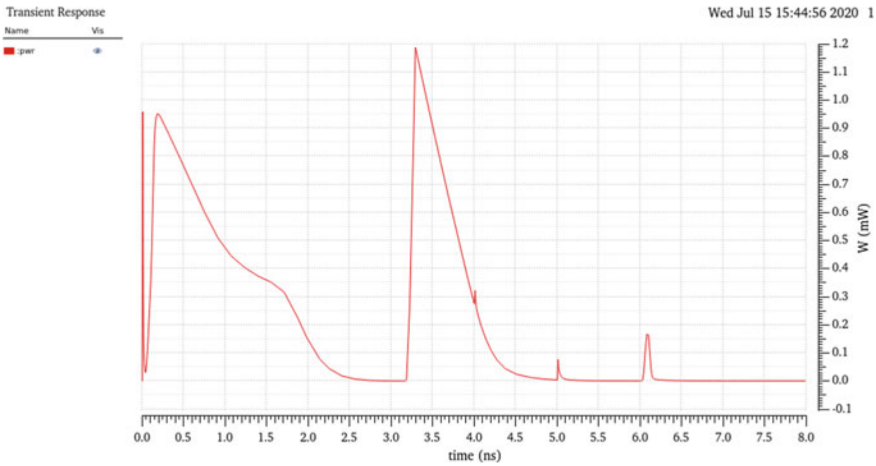


Fig. 9 Average power overall dynamic comparator

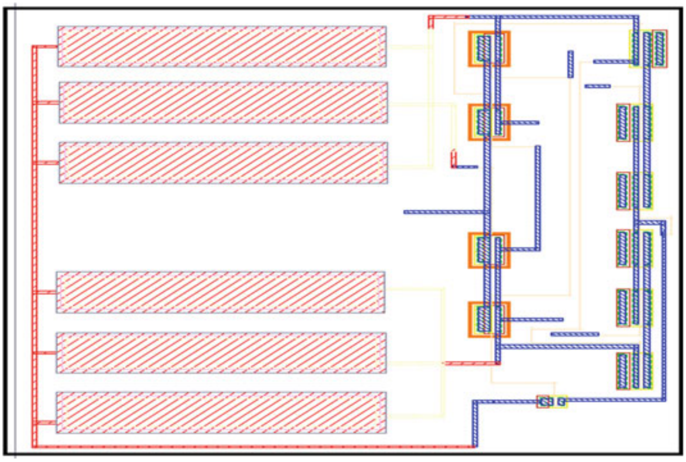


Fig. 10 Layout schematic diagram

capabilities has been developed to boost the comparator's efficiency and reduce the delay post-layout simulation findings in 180 nm. CMOS equipment verified that new comparator's delay and energy per adaptation are significantly decreased as opposed to the traditional dynamic comparator.

References

1. Rameshkumar R, Bharathiraja S (2014) A comparative analysis of high-speed dynamic comparator in 180 nm and 90 nm using H-spice 1(12)
2. Ay SU (2011) A sub-1 Volt 10-bit supply boosted SAR ADC design in standard CMOS. Int J Analog Integr Circ Sig Process 66(2):213–221
3. Mesgarani A, Alam MN, Nelson FZ, Ay SU (2010) Supply boosting technique for designing very low-voltage mixed-signal circuits in standard CMOS. In: Proceedings—IEEE international Midwest symposium on circuits and systems, pp 893–896
4. Blalock BJ (2000) Body-driving as a low-voltage Analog design technique for CMOS technology. In: Proceedings—IEEE southwest symposium mixed-signal design, Feb 2000, pp 113–118
5. Maymandi-Nejad M, Sachdev M (2003) 1-bit quantiser with rail to rail input range for sub-1 V modulators. IEEE Electron Lett 39(12):894–895
6. Goll B, Zimmermann H (2007) A 0.12 µm CMOS comparator requiring 0.5 V at 600 MHz and 1.5 V at 6 GHz. In: Proceedings—IEEE international solid-state circuits conference. Digest of technical papers, Feb 2007, pp 316–317
7. Bhagavatula VSR (2015) Design of a single tail comparator on a 90 nm technology. (IJERT) ISSN: 2278-0181 IJERTV4IS100269 www.ijert.org (This work is licensed under a Creative Commons Attribution 4.0 International License.) 4(10)
8. Babayan-Mashhadi S (2013) Analysis and design of a low-voltage low-power double-tail comparator. 1063-8210/$31.00 © 2013 IEEE
9. Okaniwa Y, Tamura H, Kibune M, Yamazaki D, Cheung T-S, Ogawa J, Tzartzanis N, Walker WW, Kuroda T (2005) A 40 Gb/s CMOS clocked comparator with bandwidth modulation technique. IEEE J Solid-State Circuits 40(8):1680–1687
10. Babayan-Mashhadi S, Lotfi R (2014) Analysis and design of a low voltage low power double tail comparator. IEEE Trans VLSI 22(2)
11. Babayan-Mashhadi S, Lotfi R (2012) An offset cancellation technique for comparators using body voltage trimming. Int J Analog Integr Circ Sig Process 73(3):673–682
12. He J, Zhan S, Chen D, Geiger RJ (2009) Analyses of static and dynamic random offset voltages in dynamic comparators. IEEE Trans Circuits Syst I Reg Papers 56(5):911–919
13. Shinkel D, Mensink E, Klumperink E, van Tuijl E, Nauta B (2007) A double-tail latch-type voltage sense amplifier with 18 ps setup + hold time. In: Proceedings—IEEE international solid-state circuits conference. Digest of technical papers, Feb 2007, pp 314–315
14. Thakur R, Nema M (2012) Design of low offset voltage dynamic latched comparator. ISRO J Eng 2(4):585–590
15. Babayan-Mashhadi S (2014) Analysis and design of a low voltage low-power double-tail comparator. IEEE Trans Very Large Scale Integr Syst 22(2)
16. Razaawi (2000) Analog integrated circuit design, 3rd edn. PHI
17. Lokesh Krishna K, Al-Naamani YMA (2019) Analysis of different comparator architectures. Taylor and Francis: J Inf Optim Sci 40(2):367–376
18. Pancholi S, Meena RS (2015) A new design of a CMOS comparator using 45 nm technology. IJISET Int J Innov Sci Eng Technol 2(9)
19. Shinde AD, Sharma M (2017) Low power, area efficient dynamic comparator with reduced activity factor. In: IEEE international conference on electrical, computer and electronics (UPCON). IEEE

Design of Protection Circuit for Biomedical Signals Using 180 nm Technology

Prashant Yadav, P. P. Bansod, D. K. Mishra, and Rupali Jarwal

Abstract This paper presents designing of protection circuit for biomedical signals that is electrocardiograph (ECG) and electroencephalogram (EEG). Protection circuit is an essential part of any biomedical device, and it is designed using Cadence Virtuoso environment in 180 nm SCL technology with 1.8 V power supply. During signal acquisition of ECG and EEG, events like voltage surges and current spikes can occur because of defibrillator or electrostatic discharge. The altered voltage input may transfer high current to the patient and can be even fatal. To minimize these effects, protection circuit is used to protect patient and electronic circuit against serious damage. Proposed protection circuit can successfully limit high-voltage surges up to 1.5 to 2.7 V and generate almost no current. Low-pass filter (LPF) provides noise cancelation to electrosurgical RF units and eliminates phase distortion in EEG and ECG signals with 3.49 kHz cutoff frequency.

Keywords Electrocardiograph (ECG) · Electroencephalogram (EEG) · Electrostatic discharge (ESD) · Low-pass filter (LPF) · SCL technology · Cadence Virtuoso

1 Introduction

Human body produces electrical signs having low recurrence and extremely low abundancy, to manage these low recurrence and plentifulness, high CMRR is needed to recognize powerless biomedical signs as normal mode sign may modify the first sign. To satisfy these prerequisites, instrumentation enhancer is utilized as pre-speaker as it has high differential addition and high CMRR [1–4].

P. Yadav (✉) · P. P. Bansod · D. K. Mishra · R. Jarwal
Electronics and Instrumentation Engineering, SGSITS Indore, Indore, India
e-mail: prashantyadav1135@gmail.com

P. P. Bansod
e-mail: pbansod@sgsits.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_15

Electrocardiograph is a high accuracy electronic gadget for clinical consideration. Utilizing ECG gadget, the ideal powerless voltage–current pointers are intensified, so they might put something aside for finding and clinical examination. ECG [5] alerts range from few millivolts with electricity band aspect of 0.05 to 100 Hz [6]. The alerts are much liable to defibrillation, diverse undesirable electric susceptibilities, and other noises. Excessive voltage for 5 to 20 ms is created because of these factors, which is a very long duration for any electronic device to pull through [7]. Therefore, this excessive voltage required to be clamped using safety device and noises to be filtered out.

Electroencephalogram is an electrophysiological signal that maps brain activity, which is extremely important to diagnose and analyze neurological disorders. EEG [8–11] signal amplitude ranges from 10 to 100 µV with frequency band of 0.5 to 100 Hz which is very susceptible to electrosurgical RF noises and phase distortion [12].

The biomedical signs [13, 14] are taken from human body, so tape of these signs is an extraordinary test in light of the fact that biomedical signs take little recurrence and extremely low plentifulness, thusly high common mode rejection ratio (CMRR) is needed biomedical signal front-end design to detect the exceptionally frail biomedical signs in light of the fact that the normal mode sign may influence the biomedical signs. Along these lines, the instrumentation intensifier is utilized as the pre-enhancer. It has high differential increase just as high normal mode dismissal proportion (CMRR). The common technique for measuring the biomedical signals is electroencephalography and electrocardiography [9, 15].

Table 1 demonstrates dissimilar biomedical signals with their amplitude and frequency range [16].

Figure 1 shows block diagram of biomedical signals acquisition structure. Fortification circuit is provided with the signals from leads that are connected to the patient, it secures patient and device from great voltage–current events. Further, the signals are passed to instrumentation amplifier (IA) for signal amplification so it can be detected and analyzed easily. In order to reduce the interference noise due to high resistance used in IA, band-pass filter is used.

Table 1 Biomedical signals

Biomedical signals	Frequency range (Hz)	Amplitude range
ECG	0.05 to 150	1 to 5 mV
EEG	0.5 to 40	0.5 to 100 µV

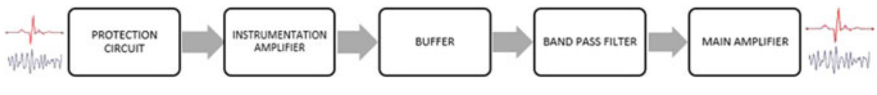


Fig. 1 Block diagram of biomedical signal acquisition unit

Table 2 Sources of input transient

Input transient sources	Magnitude (kV)
Defibrillation	~6
Electrostatic discharge (ESD)	≥2
Electrosurgical RF units	~0.3–2
Accidental contact with power supply	~0.11–0.22

2 Sources of Input Transient

Input transient degree offered by different sources of ECG and EEG, respectively, is as follows [17]:

Defibrillation: Defibrillators are gadgets that re-establish a typical heart beat by transfer an electric heartbeat or shock to the heart. They are exploited to foresee or discourse an arrhythmia and can likewise forestall unexpected demise among individuals who have high danger of a dangerous arrhythmia (Table 2).

ESD: Electrostatic discharge is the exchange of electrical charge between two bodies at various possibilities because of direct contact or an actuated electric field, and if this charge exceeds safety limit, it will damage the IC. ESD is a major concern in case of MOSFET as it can destroy gate oxide, metallization, and junction. ESD can harm patient as current may flow from circuit to the patient causing an electric shock that can be even fatal [18].

Electrosurgical RF units (ESU): Biomedical sign should be shielded from different natural or outer emanations. Close by clinical gears can likewise produce high recurrence. During electrosurgery, ECG and EEG transmissions can turn out to be significantly twisted because of the impedance of the radio recurrence signals which are ensured by electrosurgical unit. ESU recurrence goes from 100 kHz to 100 MHz with greatness up to a few of kilovolts [19].

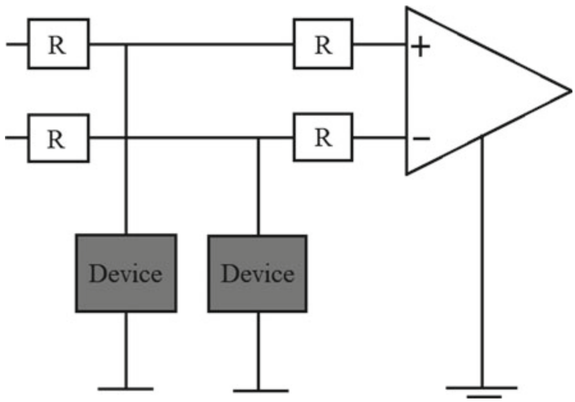
The insurance circuit liked in this paper includes voltage restricting gadget that can clasp voltage more modest than 40 V. Information separating unit turns down ESU clamor without mutilating ECG and EEG signal. Patient is gotten from over voltage–recent developments by thinking about these variables.

3 Filtering and Protection

Resistor R shown in Fig. 2 [17], in voltage limiting device that can limit voltage and eventually the current flow. Range of resistor can be varying from 10 to 20 kΩ [12, 20].

Instead of diode D1, MOSFET is used with drain and source shorted because it have characteristics equal as diode [21] (Fig. 3).

Fig. 2 Input protection circuit



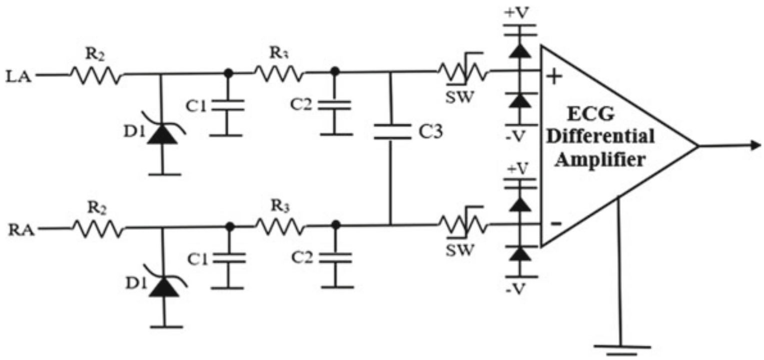


Fig. 3 ECG input filtering circuit

4 Implementation of Input Filtering Protection Circuit

Figure 4 shows schematic of protection circuit. Voltage limiting device clamps high voltage, and low-pass filter limits radio frequency/ESU interference. With the support of electrode of left and right arm, *R* and *C* values are obtained. Amplifier is supplied with the clamped and filtered outputs.

The filtering circuit successfully limits frequency of 3 kHz to 5 kHz approximately that is sufficient to diminish ESU noise however can also be prevented phase distortion in biomedical signals.

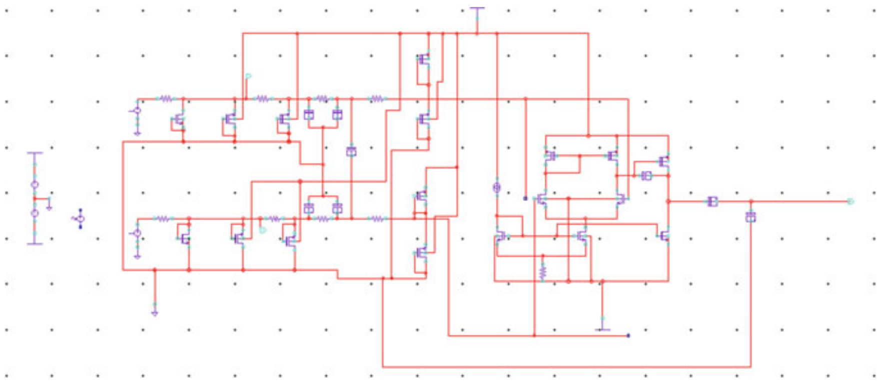


Fig. 4 Schematic of protection circuit

5 Analysis of Circuit

5.1 *Voltage Clamping Analysis*

Clipping device comprises MOSFETs before which resistor is used that limits the current and even protects the device from damage. One MOSFET clipper is operated in reverse saturation as shown in Fig. 5.

5.2 *ESU Filtering Analysis*

Figure 6a shows low-pass filter (LPF) with 3.49 kHz cutoff frequency which can successfully eliminate ESU noise and can also control the phase distortion of biomedical signals. This enables error-free biomedical signal diagnosis and analysis.

Figure 6b shows LPF output gain plot with respect to frequency.

5.3 *Overall Analysis*

Pulse of high voltage about 2 to 7 kV is applied to circuit, but protection circuit avoids the defilement of electronic device. Considering safety of patient, leakage current should be negligible because it can be even fatal.

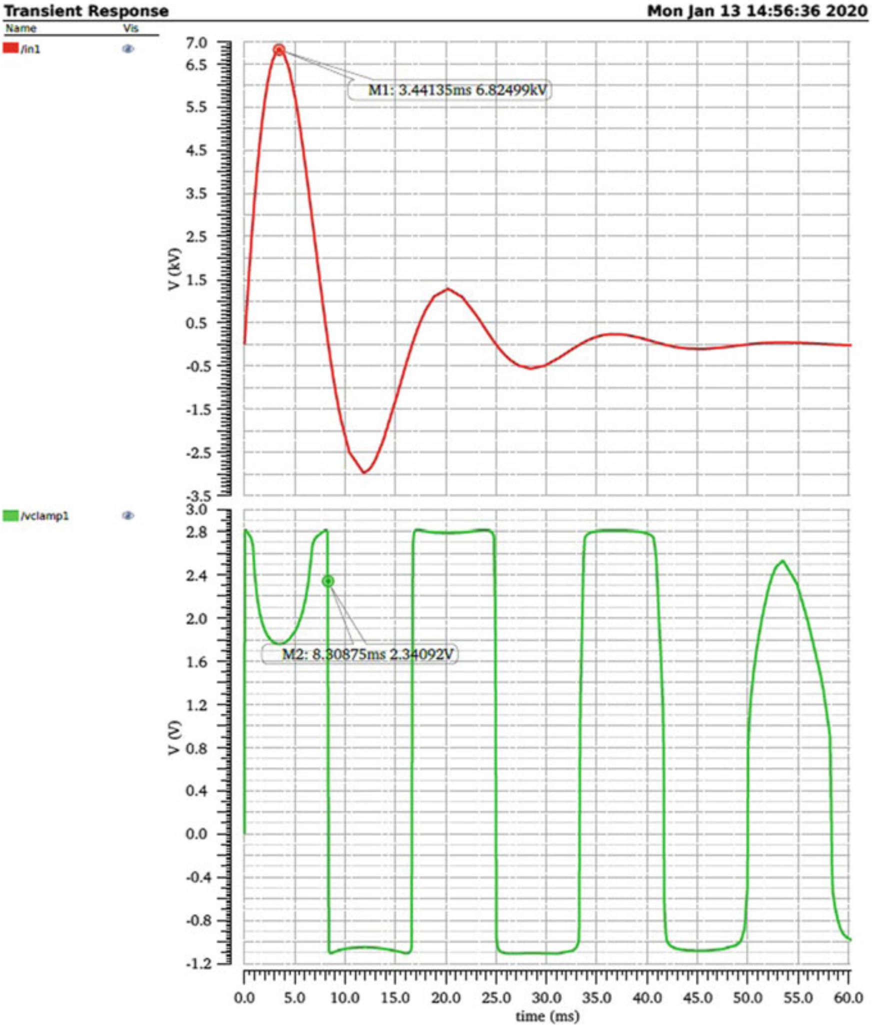


Fig. 5 Graphical representation of voltage clamping

Figure 7 shows defibrillator of 2.53 kV with high rise time, supplied to protection circuit which limits it to 2.7 V with negligible leakage current. The final output comes under few millivolts (Figs. 8 and 9).

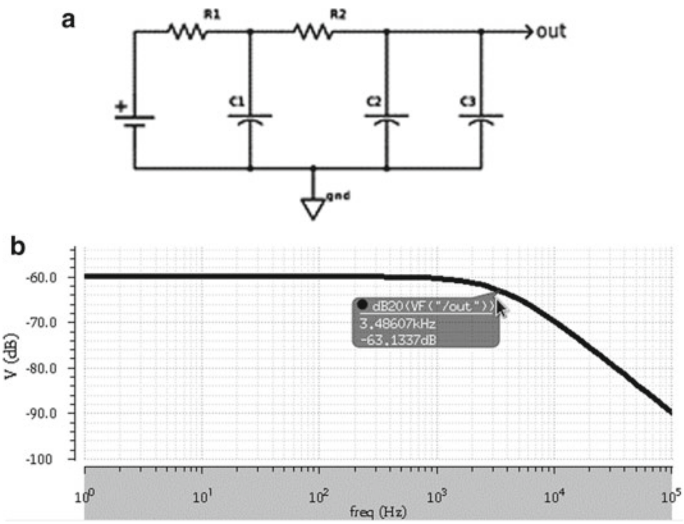


Fig. 6 **a** Diagram of ESU filter, **b** Graphical plot representation of output gain

6 Conclusion

Favored insurance circuit in this paper can effectively restrict high voltage–recent developments that give security to patient and stay away from pollution of electronic gadget. Voltage scope of kilovolts can be braced from 1.5 V to 2.7 V, and spillage current is practically unimportant. ESU clamor and stage mutilation are additionally disposed of utilizing low-pass channel (LPF). Indeed, even with the vulnerable recurrence range distinction among ECG and EEG flags, the favored insurance circuit can ensure the two signs for blunder-free conclusion and investigation, yet most huge job of the security circuit is the well-being of the patient. The circuits are carried out and recreated utilizing rhythm Virtuoso Climate in 180 nm SCL innovation (Table 3).

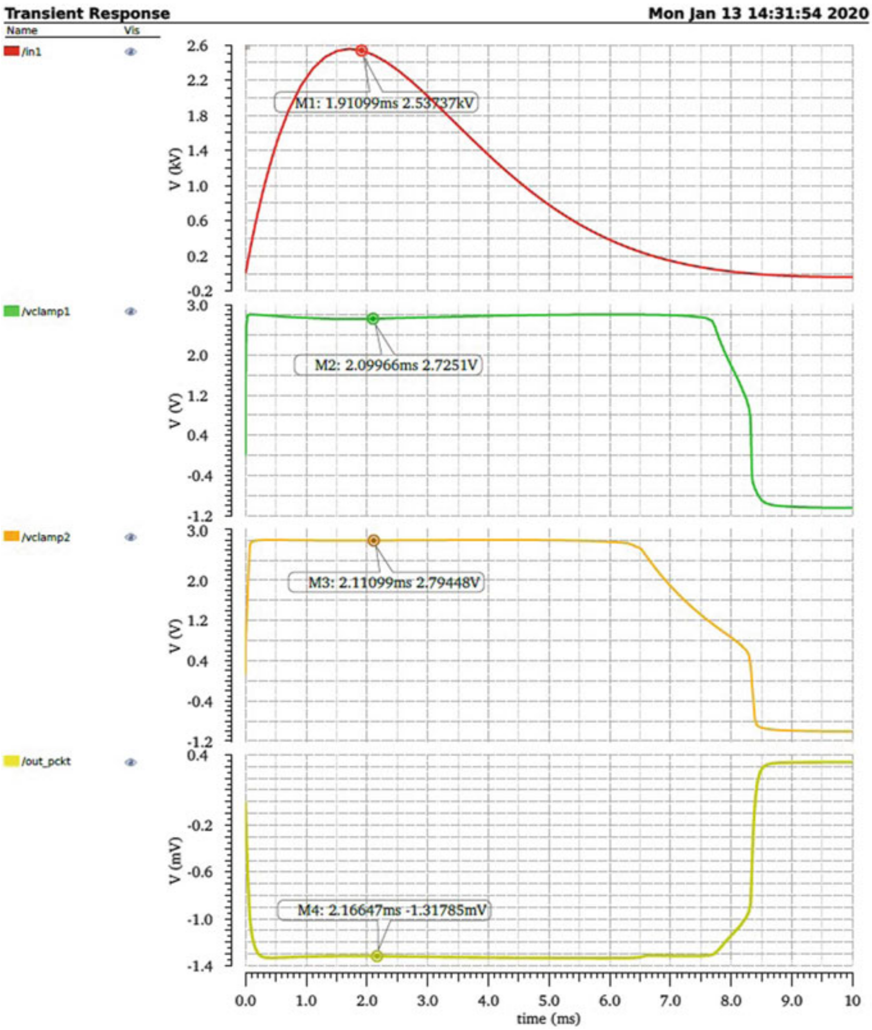


Fig. 7 Graphical plan symbol of clamped voltage

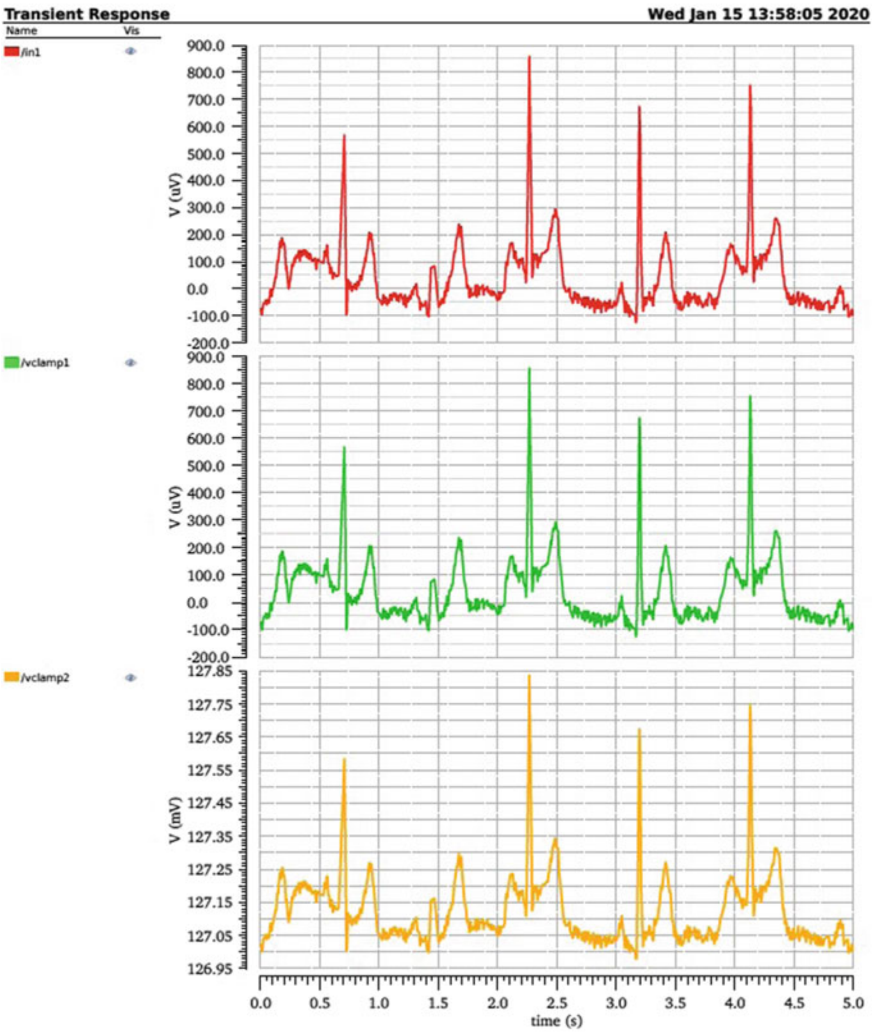


Fig. 8 Graphical plot representation of transient response of protection circuit when ECG signal is applied

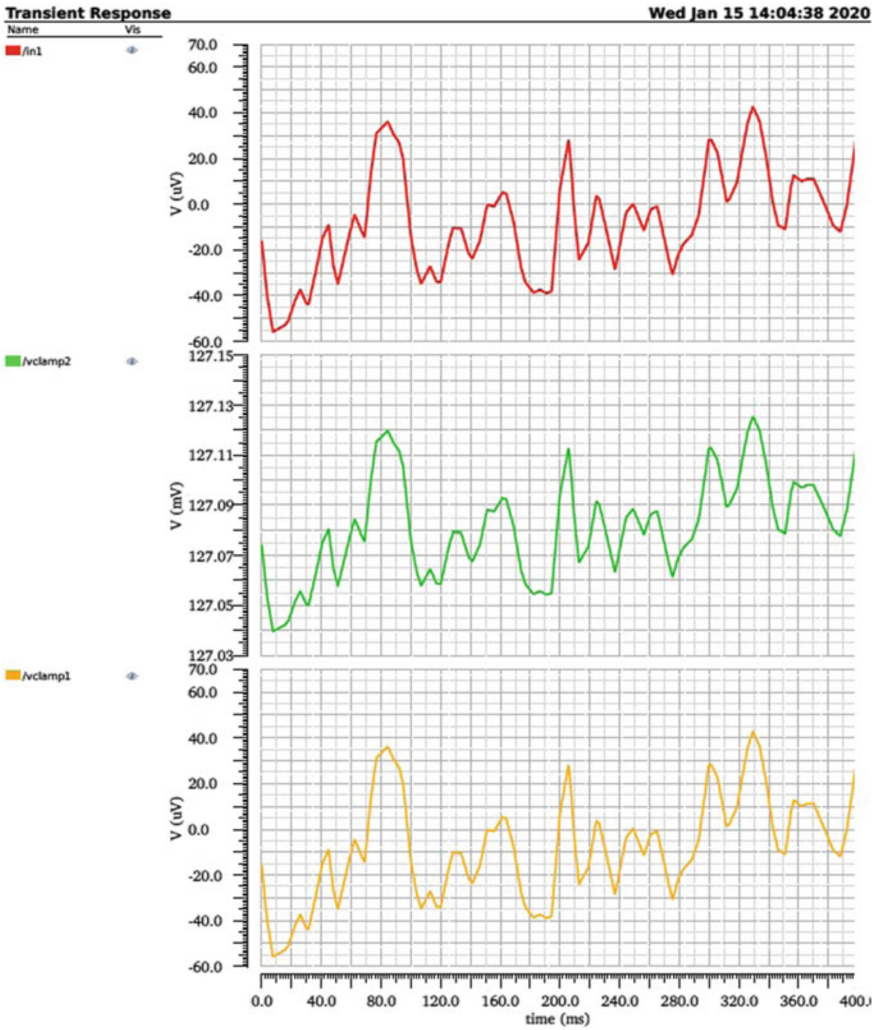


Fig. 9 Graphical plot representation of transient response of protection circuit when EEG signal is applied

Table 3 Simulated results of protection circuit

S. No.	Simulated results of protection circuit		
	Parameters	Target specifications	Achieved specifications
1	Technology (nm)	180	180
2	Supply voltage (V)	1.8	1.8
3	Defibrillator voltage (kV)	4–7	5
4	Defibrillator current (A)	40–60	50
5	Clamped voltage due to primary protection circuit (V)	≤ 5	2.7
6	Clamped voltage due to secondary protection circuit (V)	≤ 1.8	1.3
7	Cutoff frequency of RC low-pass filter (KHz)	3–5	3.49
8	ECG amplifier output voltage	≤ 0.5 V	0.83 mV
9	EEG amplifier output voltage	≤ 0.5 V	40.3 μV
10	Power dissipation (uW)	≤ 2	1.5

References

1. Application report: ECG implementation on the TMS320C5515 DSP medical development kit (MDK), by Texas Instruments, June 2010
2. Sedra AS, Smith KC (2015) Microelectronics circuits, 7th edn. Oxford University Press, pp 82–83
3. Razavi B (2000) Design of Analog CMOS integrated circuits. Tata McGraw Hill Education
4. Lin C, Chiu Y (2017) High-voltage driving circuit with on-chip ESD protection in CMOS technology. In: 2017 international conference on intelligent informatics and biomedical sciences (ICIIBMS). Okinawa, pp 223–224
5. Viheriäkoski T, Kokkonen M, Tamminen P, Kärjä E, Hillberg J, Smallwood J (2014) Electrostatic threats in hospital environment. In: Electrical overstress/electrostatic discharge symposium proceedings 2014. Tucson, AZ, pp 1–9
6. Yongjun L, Xiaorong X (2011) Analysis and implement of ECG front-end signal processing circuit. In: 2011 International conference of Information technology, computer engineering and management sciences, pp 309–312
7. Laumeister B (2014) Lightning bolts, defibrillators, and protection circuitry save lives. www.electronicdesign.com/power/lightning-bolts-defibrillators-and-protection-circuitry-save-lives
8. Lee SY, Cheng CJ (2009) Systematic design and modeling of a OTA-C filter for portable ECG detection. IEEE Trans Biomed Circ Syst 3(1):53–64
9. Harrison RR, Charles C (2003) A low-power low-noise CMOS amplifier for neural recording applications. IEEE J Solid-State Circ 38(6):958–965
10. Gardella PJ, Villa Fernandez E, Baez E, Biberidies N, Cesaretti JM (2019) Low noise front-end and ADC for real-time ECG system in CMOS process. In: 2019 IEEE 10th Latin American symposium on circuits and systems (LASCAS). Armenia, Colombia, pp 1–4
11. Barsaiyan M, Bansod PP (2017) Protection circuit design for electrocardiograph (ECG) with input filtering. In: 2017 8th International conference on computing, communication and networking technologies (ICCCNT). Delhi, pp 1–4
12. Wu C-Y, Huang C-W, Tsai T-L (2019) A 2.36 μW/Ch CMOS 8-channel EEG acquisition unit with input protection circuits for applications under transcranial direct current stimulation. In: 2019 IEEE biomedical circuits and systems conference (BioCAS)

13. Gardella PJ, Villa Fernandez E, Baez E, Cesaretti JM (2017) A chopped front-end system with common-mode feedback for real time ECG applications. In: 2017 IEEE 8th Latin American symposium on circuits and systems (LASCAS). Bariloche, pp 1–4
14. Ishida T, Nitta S, Xiao F, Kami Y, Fujiwara O (2015) An experimental study of electrostatic discharge immunity testing for wearable devices. In: 2015 IEEE international symposium on electromagnetic compatibility (EMC). Dresden, pp 839–842
15. Wang T-Y (2014) A fully reconfigurable low-noise biopotential sensing amplifier with 1.96 noise efficiency factor. IEEE Trans Biomed Circ Syst 8(3):411–422
16. Singh KS, Bansod PP, Gamad RS (2019) ECG monitoring system using 180 nm technology. In: 2019 3rd international conference on electronics, materials engineering and nanotechnology (IEMENTech)
17. Krishnakant Nammi (2003) Voltage-limiting devices for ECG amplifiers. BME 462 Medical Instrumentation
18. Kohani M, Bhandare A, Guan L, Pommerenke D, Pecht MG (2018) Evaluating characteristics of electrostatic discharge (ESD) events in wearable medical devices: comparison with the IEC 61000-4-2 standard. IEEE Trans Electromagn Compat 60(5):1304–1312
19. Liu J et al (2011) ESD protection and biomedical integrated circuit co-design techniques. In: 2011 IEEE biomedical circuits and systems conference (BioCAS). San Diego, CA, pp 405–408
20. Carr JJ, Brown JM (1981) Introduction to biomedical, equipment technology, 4th edn. pp 222–228
21. Zhou J et al (2017) An ESD demonstrator system for evaluating the ESD risks of wearable devices. In: 2017 39th electrical overstress/electrostatic discharge symposium (EOS/ESD). Tucson, AZ, pp 1–7

Analyzing Performance of IoT Applications—A Software Metric Approach

Rishabh Deo Pandey and Itu Snigdh

Abstract Software metrics are an integral part of any software. Currently, there are many IoT applications that are being created. Our article tries to apply metrics to the basic applications that an IoT application employs. We analyze the network performance using the applications in context to metrics like reliability, functionality, and efficiency. We also show through simulations the performance variations in the applications subject to similar network conditions.

Keywords Software metrics · Internet of Things · Energy efficiency · CoAP · CBR

1 Introduction

Internet of Things abbreviated as IoT is a network of physical devices or things embedded with electronic devices such as sensors, actuators, software, and an interconnecting network which enables it to collect and manage data without involving human interaction [1]. By the term 'things,' we refer to various smart devices such as smart watches, smart homes, heart rate monitoring chips and similar hardware for IoT based applications. It helps to integrate human society and physical devices. It combines various technologies into a single intelligent unit. Due to its tremendous growth in present market, IoT is being considered as the Fourth Industrial Revolution. Machine to machine (M2M) is considered as the forerunner of IoT [2]. However, due to its cloud architecture, IoT is more scalable as compared to M2M.

Though IoT is not a very old term in industrial market, its base was provided way back in 1800s [3]. A brief evolution of IoT technology is mentioned in Table 1.

R. D. Pandey (✉) · I. Snigdh
Department of CSE, BIT Mesra, Ranchi, India
e-mail: phdcs10056.19@bitmesra.ac.in

I. Snigdh
e-mail: itusnigdh@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_16

Table 1 Evolution of IoT

Year	Description
1832	Electromagnetic telegraph was designed through which two devices can directly communicate with each other with the help of electrical signals
1980	Coca Cola vending machine located at Carnegie Mellon University used microswitches and Internet for tracking the cooling of bottles and cans
1990	John Romkey designed a toaster that could be turned on and off via Internet with the help of TCP/IP protocol
1999	Kevin Ashton in 1999 while working for Proctor & Gamble. He gave the term 'Internet of Things' during a presentation on the use of RFID in supply chain management
2000	LG electronics developed an Internet refrigerator that keeps track of the items that were stored in it using RFID
2008	First global conference on IoT was held at Zurich in Switzerland
2011	IPV6 was launched which allowed 2^{128} different addresses to communicate

1.1 Introducing Metrics for IoT Applications

The different domains where IoT is being utilized can be enlisted as follows:

1. **Wearables**: includes Fitbit, smart watches.
2. **Smart Home**
3. **Health Care**: heart rate monitoring devices.
4. **Agriculture**: involves technologies such as smart greenhouse.
5. **Industrial Automation**: application of IoT in quality check, logistics, and supply chain management.
6. **Smart Cities**: optimized traffic system.

Common to all the mentioned applications, the IoT application comprises four major segments [4]. The sensor and actuators with the communication capabilities are the integral part and mostly hardware oriented. The software segment consists of the user interface modules, connectivity, and adaptability of the apps across different machine capabilities and interfaces. Since an important part of an IoT application is the software in addition to its hardware, evaluation of the performance of an IoT application becomes necessary [5]. It is known that software metrics are a measure of characteristics of the software that are quantifiable or countable. The main objective of any software metric is to analyze the product or process, to determine its quality and suggest improvements as well as predict when the software development process is over [6]. Software metrics, therefore, need to be developed for IoT applications, in order to monitor IoT performance for fulfilling the following requirements:

- Able to monitor devices that run on different processor architecture.
- Able to monitor IoT application written in different programming languages.
- Overhead in monitoring should be minimal.
- Able to receive data generated.

There are many software metrics that are generally used for IoT systems [7]. Some of them are as follows:

- Mean time between failures (MTBF).
- Mean time to repair (MTTR).
- Function-oriented metrics (computing function point).
- Defect removal efficiency (DER).
- Application crash rate (ACR).

2 Related Work

In [8], IoT framework, which is a computational middleware, was discussed briefly. It was designed on an open-source component. Likewise, in [9], the authors examine about IoTivity Cloud, which is one of the most prominent open-source platforms responsible for gathering, analyzing, and interpreting enormous amount of information accessible in the IoT climate. As suggested by [5], to give a profound knowledge into IoT Web application execution, two test applications were carried out. In the first application, latencies initiated by various communication protocols were with the help of message encodings techniques as well as graphic rendering performance. They additionally analyzed the working of various Web platforms by implementing them separately. In the subsequent application, Web execution of different IoT messaging protocols, for example, MQTT, AMQP, XMPP, and DDS, was assessed by estimating the delay in the sensor message delivery and the message delivery throughput rate. On the one hand, the data storage and processing capabilities were measured [9, 10], and on the other hand, literatures also provide instances to practically test the performance of routing protocol for low-power and lossy networks (RPL) in different scenarios in order to evaluate its efficiency in terms of packet delivery ratio (PDR), throughput, and latency [11]. While considering security as a performance metric, BenchIoT is a benchmark suite and assessment system to address squeezing difficulties and constraints for assessing IoT-μCs security. The assessment system empowers programmed assessment of 14 measurements covering security, execution, memory use, and energy utilization. The BenchIoT benchmarks give a curated set of five genuine IoT applications that cover both IoT-μCs with and without an OS [12].

3 Simulation and Analysis

- In order to determine performance of a software model, we have predefined ISO-9126 software quality characteristics which is used for evaluating the software model [13]. It mainly comprises six attributes for evaluating a model. For analysis, we have considered attributes illustrated below:

- **Functionality as a metric of performance** includes suitability, accurateness, security, interoperability, and compliance. It is observed by the throughput of the application.
- **Reliability as a metric for fault tolerance** comprises maturity, fault tolerance, and recoverability. It has been measured as the packet delivery ratio.
- **Efficiency** includes time behavior and resource behavior. Here, we have considered delay and energy efficiency.

Our model is assessed on the basis of these factors. To analyze the efficiency of an IoT-based application, a system model has been implemented via NetSim version 11.0.21 as shown in Fig. 1. The IoT model comprises a hypothetical outlay of 16 sensors deployed in an area of 500 * 250 (grid length 500 m and sensor grid length 250 m). These 16 nodes are connected through Zigbee protocol (IEEE 802.15.4) to a gateway which is further connected to a wireless server with the help of a router and an access point.

For the simulation, we have placed 16 nodes randomly and considered ideal condition. We assume that there is no path loss. The routing algorithm is AODV with four different applications, namely COAP, VIDEO, CBR, and VOICE. All the 16 sensors are connected to a server through these applications. The sensor and their respective application details are given in Table 2.

We have considered constrained application protocol (CoAP) as it is a specialized Web transfer protocol and is generally used with constrained nodes and constrained networks in IoT [14]. CoAP is created in such a way so as to enable simple, constrained devices to join the IoT network even through constrained networks with low bandwidth and low availability. Application CBR stands for constant bitrate and is an encoding technique which keeps the bitrate constant [15] and is employed in

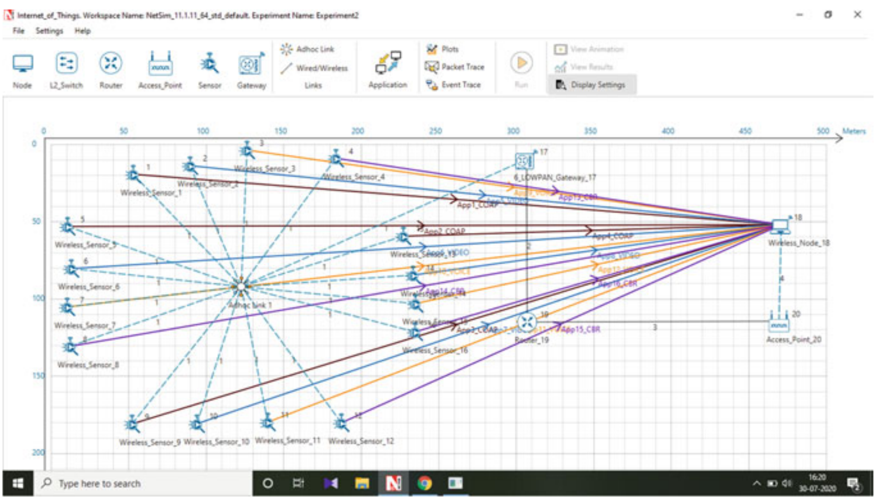


Fig. 1 NetSim connection

Table 2 Sensor application details

Application	Sensor ID
COAP	1, 5, 9, 13
VIDEO	2, 6, 10, 14
VOICE	3, 7, 11, 15
CBR	4, 8, 12, 16

most networks. The main advantage of using the CBR is that the multimedia data is processed faster as compared to other alternatives like variable bitrate (VBR). Voice and video applications attribute to the continuous data in IoT like camera and sound feeds. To analyze the performance of a network, four different parameters are considered which are illustrated as follows:

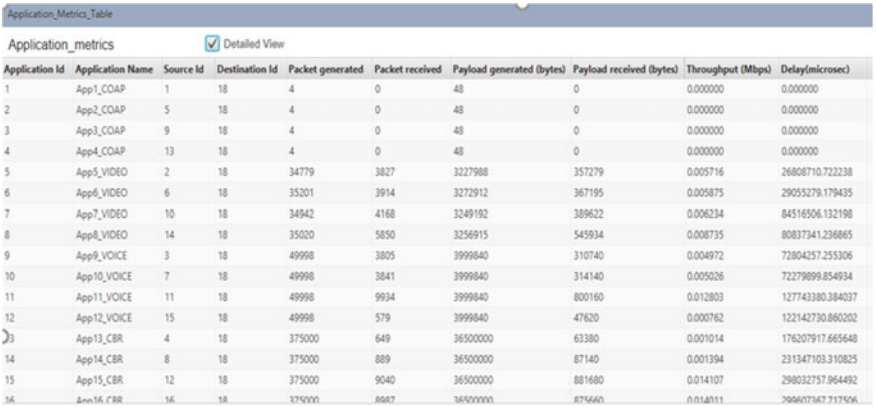
- **Throughput**: Rate of successful message delivery over a channel.
- **Packet delivery ratio (PDR)**: Ratio of the total number of packets received to the total number of packets sent in a network.
- **Energy/battery**: Amount of energy consumed.
- **Delay**: Amount of time a bit takes to travel from one endpoint to another endpoint in a communicating network.

A network's performance can be evaluated using abovementioned parameters as when they are mapped with the characteristics of ISO-9126 model [16, 17]. It is observed that throughput and delay would provide information regarding efficiency (time behavior characteristics), energy/battery gives information regarding efficiency (resource behavior), and throughput involves functionality (accurateness) behavior of a network.

4 Results

In this section, the detailed comparison of all four applications on the basis of four different parameters is done via comparison table. Three different scenarios of varying simulation times, 100, 200, and 500, have been used. The parameters were studied with the help of various metric tables that were generated in the end of NetSim simulation, and a sample has been represented in Fig. 2. Tables 3, 4, and 5 present the result obtained from the simulations.

From the results, we infer that, in order to determine the performance of different applications, the simulation result suggests that throughput for the multimedia applications such as VIDEO and VOICE is better as compared to CoAP. Also, when the simulation time is increased, throughput for CBR also increases but simultaneously delay also increases. When analyzing the performance of Zigbee network, CoAP and VIDEO applications are more effective than CBR and VOICE.



Application_Metrics_Table

Application_metrics ☑ Detailed View

Application Id	Application Name	Source Id	Destination Id	Packet generated	Packet received	Payload generated (bytes)	Payload received (bytes)	Throughput (Mbps)	Delay(microsec)
1	App1_COAP	1	18	4	0	48	0	0.000000	0.000000
2	App2_COAP	5	18	4	0	48	0	0.000000	0.000000
3	App3_COAP	9	18	4	0	48	0	0.000000	0.000000
4	App4_COAP	13	18	4	0	48	0	0.000000	0.000000
5	App5_VIDEO	2	18	34779	3827	3227988	357279	0.005716	26808710.722238
6	App6_VIDEO	6	18	35201	3914	3272912	367195	0.005875	29055279.179435
7	App7_VIDEO	10	18	34942	4168	3249192	389622	0.006234	84516506.132198
8	App8_VIDEO	14	18	35020	5850	3256915	545934	0.008735	80837341.236865
9	App9_VOICE	3	18	49998	3805	3999840	310740	0.004972	72804257.255306
10	App10_VOICE	7	18	49998	3841	3999840	314140	0.005026	72279899.854934
11	App11_VOICE	11	18	49998	9934	3999840	800160	0.012803	127743380.384037
12	App12_VOICE	15	18	49998	579	3999840	47620	0.000762	122142730.860202
13	App13_CBR	4	18	375000	649	36500000	63380	0.001014	176207917.665648
14	App14_CBR	8	18	375000	889	36500000	87140	0.001394	231347103.310825
15	App15_CBR	12	18	375000	9040	36500000	881680	0.014107	298032757.964492
16	App16_CBR	16	18	375000	8987	36500000	875660	0.014011	299607367.717506

Fig. 2 Application metrics

Table 3 Simulation results for 100 s

Appln	Avg. throughput (in Mbps)	Avg. energy consumed (in mJ)	Avg. performance of Zigbee (IEEE 802.15.4)	Average delay (in milliseconds)
COAP	4.025×10^{-5}	1187.45	0.27	17.23
VIDEO	8.075×10^{-3}	1592.54	0.24	13,965.56
VOICE	5.186×10^{-3}	1389.63	0.20	23,539.03
CBR	4.909×10^{-3}	1416.87	0.20	47,848.64

Table 4 Simulation results for 200 s

Appln	Average throughput (in Mbps)	Avg. energy consumed (in mJ)	Avg. performance of Zigbee (IEEE 802.15.4)	Average delay (in milliseconds)
COAP	3.217×10^{-3}	3574.87	0.25	2255.10
VIDEO	7.996×10^{-3}	4780.00	0.24	40,592.04
VOICE	5.247×10^{-3}	4180.18	0.21	69,251.72
CBR	4.746×10^{-3}	4253.11	0.20	137,435.87

Table 5 Simulation results for 500 s

Appln	Avg. throughput (in Mbps)	Avg. energy consumed	Avg. performance of Zigbee (IEEE 802.15.4)	Avg. delay (in milliseconds)
COAP	2.441×10^{-3}	5948.96	0.26	12,368.06
VIDEO	6.64×10^{-3}	6906.72	0.24	55,304.46
VOICE	5.89×10^{-3}	6450.46	0.22	98,742.57
CBR	7.63×10^{-3}	6980.01	0.23	251,298.79

5 Conclusion

Our paper presents the metrics that are important for identifying the performance of IoT applications. We are aware that many IoT applications are based on text, audio, and video, and their delay and throughput under the same network conditions vary, which we have portrayed in our article. We have also shown that though we may have dedicated video applications in the sensor devices, CoAP is more robust against packet delivery losses as well as better in conserving energy. Also the delay is least in CoAP and is worst for CBR that confirms we do not employ CBR for IoT applications.

References

1. Chen H, Jia X, Li H (2011) A brief introduction to IoT gateway. In: IET international conference on communication technology and application (ICCTA 2011). IET
2. Alam M, Nielsen RH, Prasad NR (2013) The evolution of M2M into IoT. In: 2013 First international black sea conference on communications and networking (BlackSeaCom). IEEE
3. Suresh P et al (2014) A state of the art review on the Internet of Things (IoT) history, technology and fields of deployment. In: 2014 International conference on science engineering and management research (ICSEMR). IEEE
4. Udoh IS, Kotonya G (2018) Developing IoT applications: challenges and frameworks. IET Cyber-Phys Syst Theory Appl 3(2):65–72
5. Babovic ZB, Protic J, Milutinovic V (2016) Web performance evaluation for internet of things applications. IEEE Access 4:6974–6992
6. Fenton NE, Neil M (2000) Software metrics: roadmap. In: Proceedings of the conference on the future of software engineering
7. Lee M-C (2014) Software quality factors and software quality metrics to enhance software quality assurance. Curr J Appl Sci Technol 3069–3095
8. Vandikas K, Tsiatsis V (2014) Performance evaluation of an IoT platform. In: 2014 Eighth international conference on next generation mobile apps, services and technologies. IEEE
9. Dang T-B et al (2017) Performance evaluation of an IoT platform. In: Proceedings of the Korea information processing society conference. Korea Information Processing Society
10. Asiminidis C, Kokkonis G, Kontogiannis S (2018) Database systems performance evaluation for IoT applications. Int J Database Manage Syst (IJDMS) 10
11. Kharrufa H et al (2019) A performance evaluation of RPL in mobile IoT applications: a practical approach. IFAC-PapersOnLine 52(24):312–317
12. Almakhdhub NS et al (2019) Benchiot: a security benchmark for the internet of things. In: 2019 49th annual IEEE/IFIP international conference on dependable systems and networks (DSN). IEEE
13. Miguel JP, Mauricio D, Rodríguez G (2014) A review of software quality models for the evaluation of software products. arXiv preprint arXiv:1412.2977
14. Bormann C, Castellani AP, Shelby Z (2012) Coap: an application protocol for billions of tiny internet nodes. IEEE Internet Comput 16(2):62–67
15. Kershaw S, Hughes-Jones R (2010) A study of constant bit-rate data transfer over TCP/IP. Futur Gener Comput Syst 26(1):128–134
16. Moridi MA et al (2018) Performance analysis of ZigBee network topologies for underground space monitoring and communication systems. Tunn Undergr Space Technol 71:201–209
17. Guceglioglu AS, Demirors O (2005) Using software quality characteristics to measure business process quality. In: International conference on business process management. Springer, Berlin, Heidelberg

Study and Comparison of Various Protection Configurations in Optical Networks

Sandeep Dabhade, Santosh Narayankhedkar, Rahul S. Bramhane, and Vishaka R. Mote

Abstract In optical networks, various protection mechanisms are used. Network survivability is critical in optical networks so that in any case, traffic will not be down. In protected scenarios, there are work path and backup path so that even if work path fiber is cut, then traffic will switch to protect path. In this paper, we have covered sub-network connection protection (SNCP), optical line protection (OLP), Y cable, line- and client-side protections, comparison between these protection schemes.

Keywords SNCP · Line protection · Client protection

1 Introduction

In optical networks, optical add drop multiplexers are connected through optical fiber cables. In actual field, users will create 1GBE, 10GBE, 100GBE, etc., optical connections passing through multiple nodes. These nodes will carry audio, video, data traffic, and the optical connections which will be either unprotected or protected. Network survivability plays critical role; since in any case, traffic should not be down. Hence, various protection schemes have evolved like sub-network connection protection (SNCP), optical line protection (OLP), and line-side and client-side protections. Depending upon the customer configuration and requirements, various protection mechanisms are used in field [1–3]. For example: Some customers will need path-level protection, while others will need port-level protection. We have created these setups on simulators [4–6].

2 Optical Protection: Client Protection

Client-level protection used in optical networks.

S. Dabhade (✉) · S. Narayankhedkar · R. S. Bramhane · V. R. Mote
Dr. Babasaheb Ambedakar Marathwada University, Aurangabad, India
e-mail: sandeep.dabhade@nokia.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_17

Fig. 1 Client protection with 100GBE packs service

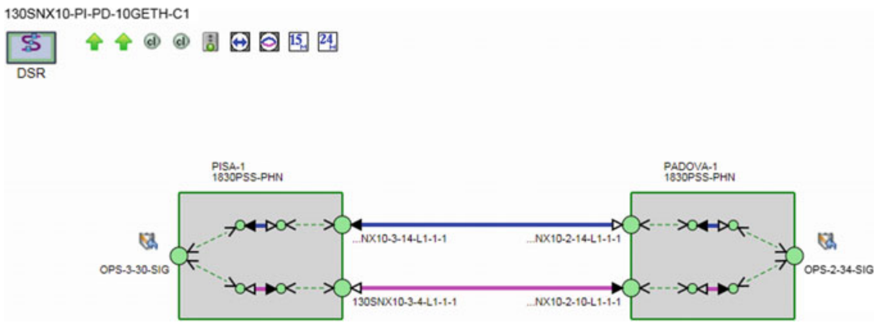


Fig. 2 Routing display of 100GBE connection

We have created three client-side protection setups using real nodes, one setup is with 100GBE transponder, second setup is using 500GBE transponder packs, and third setup is with 200GBE packs [7–10].

2.1 *Client Protection with 100GBE Packs*

See Figs. 1 and 2.

2.2 *Client Protection with 500GBE Packs*

See Figs. 3 and 4.

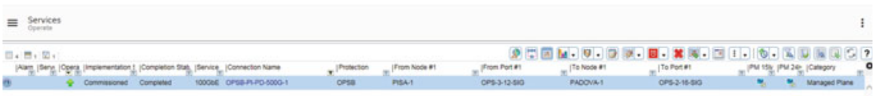


Fig. 3 Client protection with 500GBE pack service

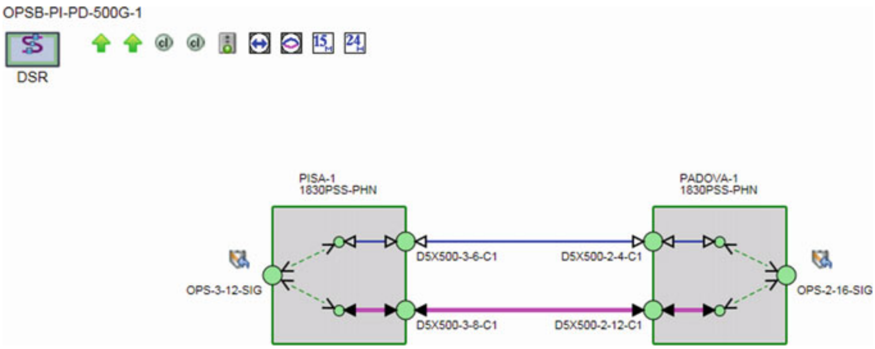


Fig. 4 Routing display of 500GBE client-protected connection

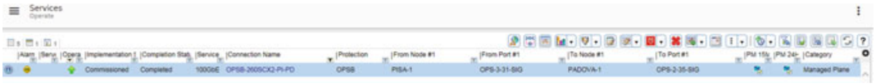


Fig. 5 Client-protected service

2.3 *Client Protection with 200GBE Packs*

See Figs. 5 and 6.

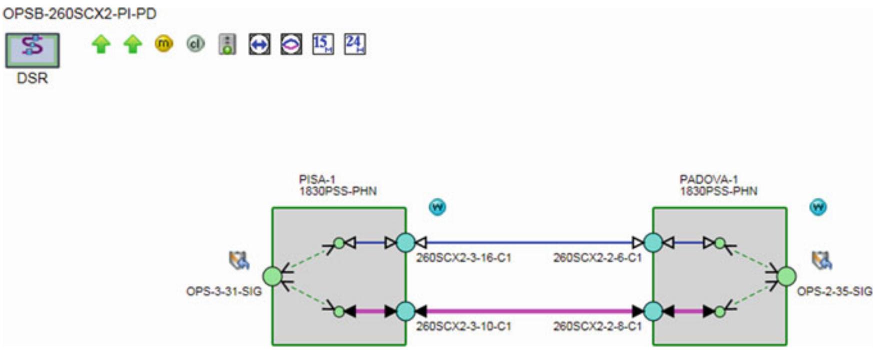


Fig. 6 Routing display of client protection with 200GBE packs

3 Sub-Network Connection Protection (SNCP)

SNCP provides path-level protection and is widely used in optical networks. In our setup, SNCP is created using MUXPONDER which has eight client ports and two line port [11].

L1 of node 1 is connected to L1 of node 2, and L2 port of node 1 is connected to L2 port of node 2. When traffic is switched from L1 port to L2 port, L1 port can be switched off by software-defined network. Also, if traffic is flowing through only one client port, then remaining unused client ports can be switched off using software defined network [12] (Figs. 7, 8 and 9).

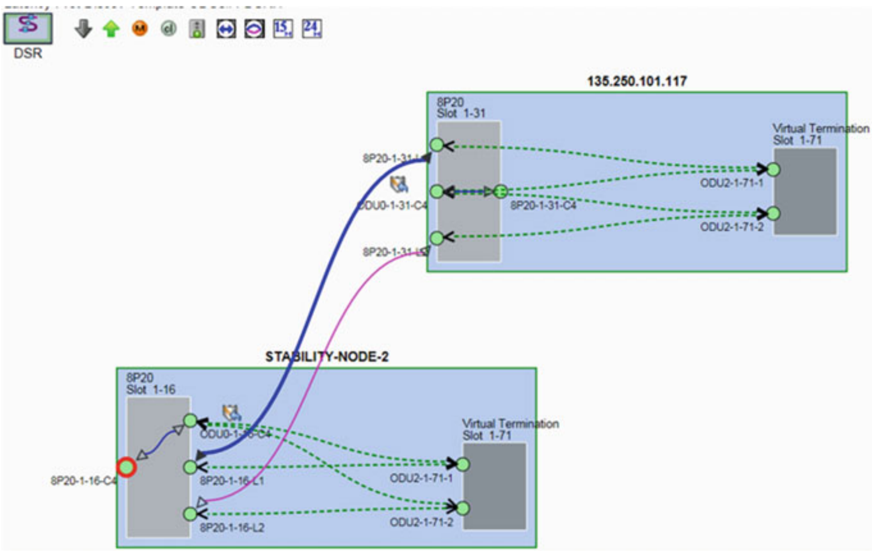


Fig. 7 End-to-end routing display of SNCP

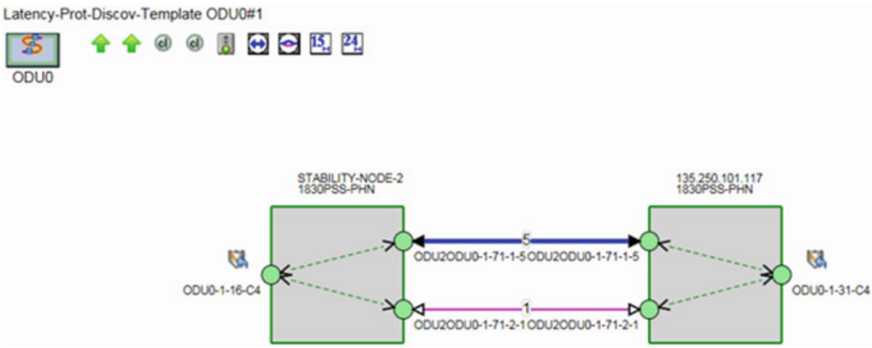


Fig. 8 1GBE SNCP

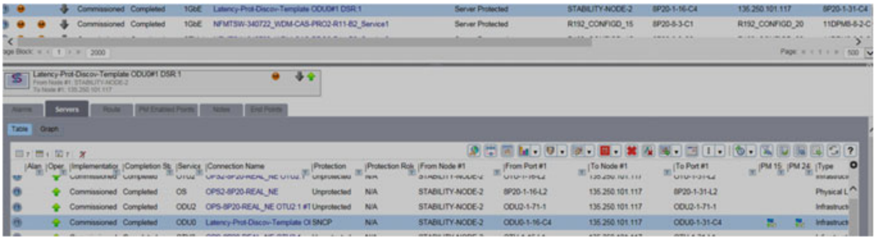


Fig. 9 1GBE SNCP service connection

4 Line Protection

Line-side protection is port-level protection. There are various methods in line-side protection (Figs. 10 and 11).

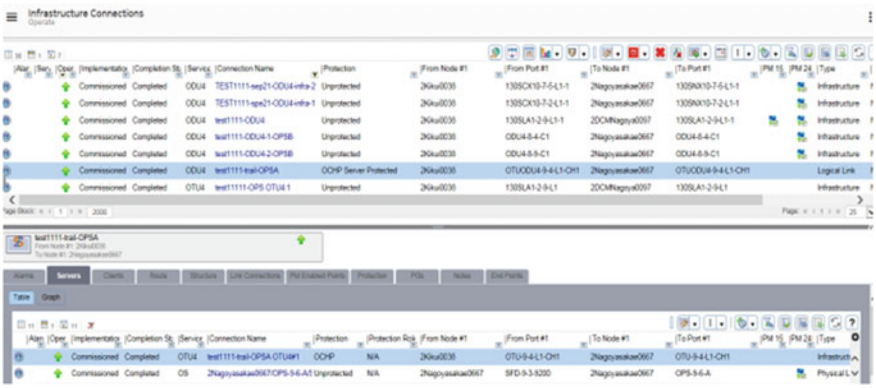


Fig. 10 Line-protected service

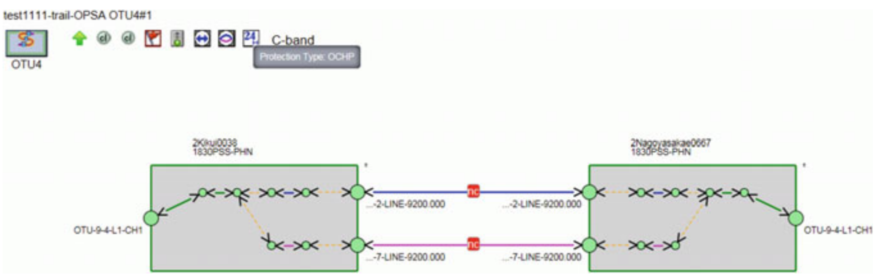


Fig. 11 Routing display of line-side protection

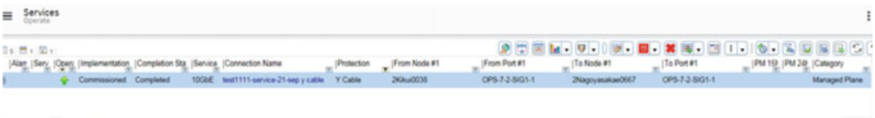


Fig. 12 Y cable-protected service connection

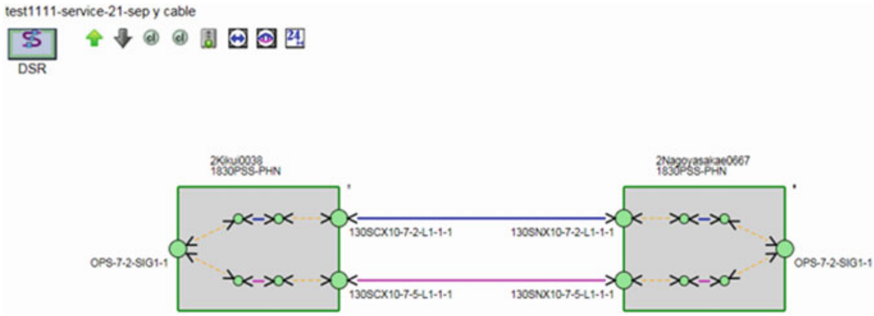


Fig. 13 Routing display of Y cable service

5 Y Cable Protection

Y cable protection provides both line- and client-side protections. Y cable is like optical splitter (Figs. 12 and 13).

6 Conclusions

In optical networks, audio, video, and traffic need to be protected in any situation like fiber cut. Whenever work path fiber gets cut, traffic switches to protect path, and once work path is restored, then traffic can be moved back to work path. Y cable protection can be used for both line and client side, sub-network connection protection is used for path protection, line protection is for port-level protection, while client protection scheme is used to protect traffic at client level.

References

1. Telco VoIP Scalability Test Results … for 10 Million Subscribers. http://www.communigate.com/Papers/CommuniGate-Superdome-VoIP-Benchmark.pdf
2. Black U Optical networks. Third Generation Transport Systems
3. Eiselt M (2011) Programmable modulation for high-capacity networks. In: ECOC. Geneva

4. Aboul-Magd O, Jamoussi B Automatic switched optical network (ASON) architecture and its related protocols. https://shiraz.levkowetz.com/pdf/draft-ietf-ipo-ason-02.pdf
5. Chen L-K, Cheung M-H, Chan C-K From optical performance monitoring to optical network management: research progress and challenges. http://www.lightwave.ie.cuhk.edu.hk/publication/document/Conference/ICOCN/2004/lkchen_icocn04.pdf
6. Samadi P, Wang H An optical programmable network architecture supporting iterative multicast for data-intensive applications. http://lightwave.ee.columbia.edu/files/Samadi2014a.pdf
7. Gifre L, Velasco L Experimental assessment of a high performance backend PCE for flex-grid optical network re-optimization. http://personals.ac.upc.edu/lvelasco/docs/research/2014-OFC-4.pdf
8. Okumura Y, Kato M, Fujii K (2014) Signal shaping to achieve OOK and PSK co-existence for improved optical access network performance. Am J Netw Commun 3(4):56–62
9. Chen L-K, Li M, Liew SC Optical physical-layer network coding—another dimension to increase network capacity? http://www.lightwave.ie.cuhk.edu.hk/publication/document/Conference/IPC/2014/2014_Chen_IPC.pdf
10. Xia M, Shirazipour M, Zhang Y, Green H, Takacs A SDN and optical flow steering network function virtualization. https://www.usenix.org/system/files/conference/ons2014/ons2014-paper-xia.pdf
11. Prabazhini M (2014) Performance improvement of optical network using dynamic bandwidth allocation. IJAICT 1(1)
12. Queiroz IM, Assis KDR (2018) A new strategy for energy saving in spectrum-sliced elastic optical networks. www.mdpi.com 11 May 2018

Evolutionary Tool for Denoising DNA Microarray Images Using CNN

R. Sunitha and H. B. Phani Raju

Abstract Deoxyribonucleic acid (DNA) microarray technology has promised rapid improvement in recent studies. On DNA microarray images, there are several spots. Spots on microarray images represent gene expression and show the status of normal and malignant cells. One of the approaches for enhancing and analysing an image is to use digital image processing. In this paper, a convolutional neural network (CNN)-based denoising technique is developed and deployed. It is based on feature extraction from images. It is no longer necessary to manually extract features. The simulation is carried out using MATLAB to evaluate the performance parameters like peak signal-to-noise ratio (PSNR), mean squared error (MSE) and structural similarity index (SSIM). The comparison between different filtering techniques and CNN is carried out checking for the efficiency. The traditional image noise reduction methods such as median filter method, linear filter method and Weiner filter methods are compared with CNN. The advantage is that the CNN model's parameters may be fine-tuned by network testing and training, whereas the parameters of traditional image denoising methods are fixed and cannot be changed throughout the filtering process and resulting in a lack of configurability.

Keywords Linear filter · Standard median filter · Weiner filter · Convolutional neural network · Mean square error · Peak signal-to-noise ratio · Structural similarity index · Deoxyribonucleic acid

1 Introduction

Microarray analysis is increasingly becoming a common research technique. However, the images generated by microarray studies are not perfect, as they contain

---

R. Sunitha (✉)
Associate Professor, Department of ECE, RajaRajeswari College of Engineering, Bangalore, Karnataka, India
e-mail: sunithar@rrce.org

H. B. Phani Raju
Professor, Department of EEE, RGIT, Bangalore, Karnataka, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_18

noise sources that taint them at various stages of their development. To provide trustworthy and accurate downstream analysis, these microarray pictures must be denoised. Distinguishing genuine gene expression values from experimental noise is a fundamental difficulty in DNA microarray analysis. Using an appropriate denoising tool, this paper proposes an efficient noise reduction strategy for microarray image Fig.1.

The ability to successfully filter away image noise has a direct effect on subsequent processing in image processing, such as object classification, edge recognition, extraction of features and so on. The degradation model, which is expressed as $N = x + \mathrm{Ng}$, where x is a clean image, N is a noisy image, and Ng is the additive Gaussian noise with standard deviation σ, is extensively employed in denoising problems to recover clear images. We normally apply two types of filtering on images. One is based on the spatial domain, and the other is on the frequency domain. Image filtering in the spatial domain uses the coefficients of several frequencies, whereas in the frequency domain, image filtering uses the pixel intensity and its neighbours. The respective denoising approaches for these two domains can be classified as spatial domain-based or frequency domain-based methods. The intensity of each pixel of an image is directly addressed by spatial domain-based approaches. After image decomposition, frequency domain-based algorithms handle picture denoising by modifying corresponding coefficients of numerous frequencies. These traditional image denoising methods are fixed and cannot be changed throughout the filtering process and resulting in a lack of configurability.

To achieve image denoising efficiently, we use convolutional neural network method. During network training, the CNN model has the advantage of continuously improving the weight of the convolution kernel. We also built the CNN model to see how it compares to traditional filters. We get two models that correlate to Gaussian noise after the network training. Denoising using the proposed CNN model has demonstrated improved performance when compared to other filtering techniques (Fig. 1).

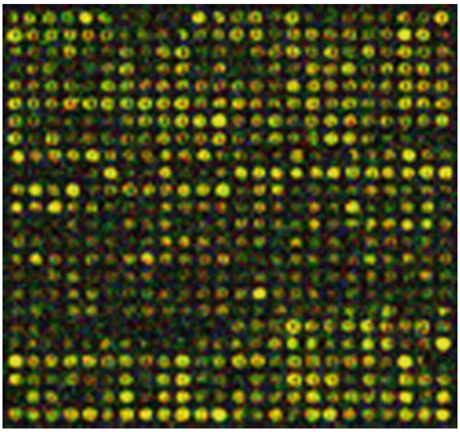


Fig. 1 Microarray image

2 Related Work

A significant amount of research has been done in the field of microarray image denoising, according to the findings of the literature review. It merges the conclusion of many different sources to explain the overall understanding of the topic. Microarray preparation, image acquisition and analysis, data pre-processing and normalisation and data in detail analysis have all been discussed by Leung et al. [1]. A vector processing-based framework for DNA microarray image denoising was proposed by Lukac and Plataniotis [2]. Tian et al. [3] introduced a new methodology for deniosing microarray pictures that combines Gaussian mixture methods (GMM) and vector flow analysis (VFA). Shao et al. [4] have suggested a novel, totally automated method for successfully detecting a deformed grid structure in a microarray image. The gridding process can be thought of as an optimisation issue that can be addressed using a genetic algorithm (GA). Nagaraja [5] has presented a strategy for utilising mathematical morphology for the enhancement of microarray pictures that is automated. On the basis of a vascular bio-marker on tissue microarrays, Ng and Ma [6] suggested a system for correctly assessing the presence and amount of stain. Liu et al. [7] developed and implemented a linear CNN-based denoising algorithm. Their results reveal that the suggested CNN method effectively removes Gaussian noise and greatly improves standard picture filtering methods. The denoising with CNN model has proven to be superior in terms of filtering operations. Zuo et al. [8] suggested the enhanced convolution neural denoising network as a new method. To overcome the problem of training challenges and expedite network convergence, they applied residual learning and batch normalisation approaches [9–16].

3 Methodology

The block diagram of Fig. 2 is divided into two parts. The first filter is the noisy image using traditional image filtering methods like: median filtering, Weiner filtering and linear filtering. The second method is CNN method.

The steps are followed for both traditional and CNN methods. Later, we are comparing both the results to know the efficiency of the image. In CNN method, we use a particular tool called as deep learning tool which helps in removing the noises present in the input images and thus helps in comparing the results with other traditional filtering methods with respect to the parametric efficiency. The proposed methodology is divided into five steps: (i) acquiring the microarray input image, (ii) adding the noise to the image, (iii) denoising the image using different traditional filtering techniques, (iv) denoising the image using CNN and (v) evaluating the performance parameters.

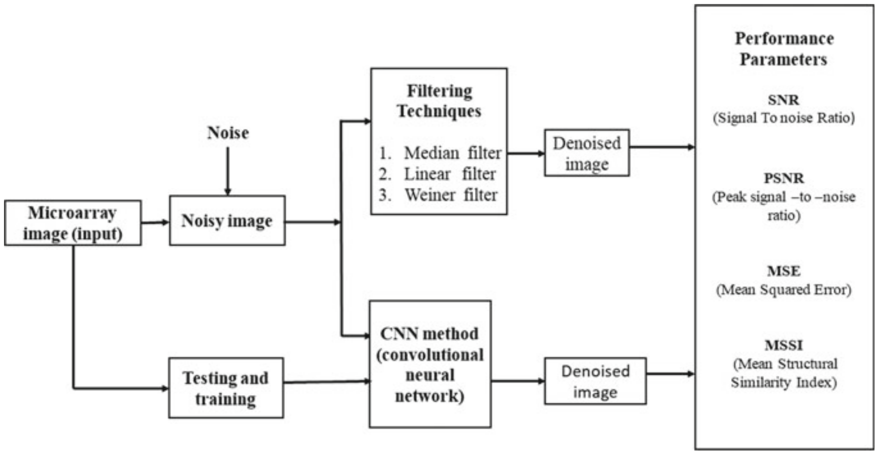


Fig. 2 Block diagram of the proposed methodology

4 Implementation

4.1 Traditional Filtering Techniques

Filtering is a technique for changing or enhancing a picture in digital image processing. We can apply filters to an image to highlight certain features while removing others.

The process followed is shown in Fig. 3. Here, we acquire a microarray image and induce a certain amount of noise such as Gaussian noise. The filtering techniques are applied on these noisy images to remove the noise and to obtain denoised image. Every filtering technique is different and has its particular algorithm. It executes the matrix's two-dimensional filtering operations. The median value in the 3×3 region around the corresponding pixels in the input image is stored in each output pixel.

When the goal is to reduce noise while maintaining edge, a median filter is more successful than convolution. It plays out the filtering activities of the grid in two measurements. Each output pixels contains the median an incentive in the 3×3 neighbourhood around the comparing pixels in the image. A median filtering [17] is more compelling than the convolution when the objective is to at the same time decrease the noise and safeguard edge.

A linear filter [18] is a filtering technique in which the output pixel's value is a linear combination of the pixel value in the input pixel's neighbourhood. Random noise is reduced, edges are sharpened, and unequal illuminations are corrected using the linear filtering technique. The technique is carried out by applying a correlation filter to the image. The weighted sum of nearby pixels is used to calculate the value of the output pixel.

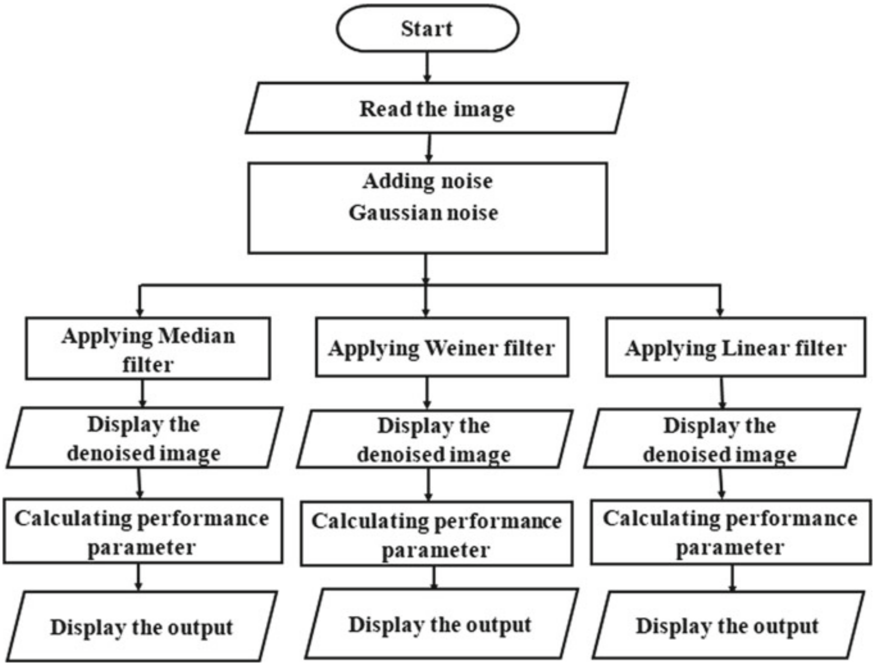


Fig. 3 Flowchart for traditional filtering technique

Weiner filtering is a filtering technique that uses linear time-invariant LTI filtering of an observed noisy process to obtain an estimate of a desired or target random process. The mean squared error between the estimated random process and the desired process is minimised using this filtering technique. It employs a pixel-by-pixel adaptive Weiner method based on statistics estimated from each pixel's local neighbourhood. Once the noisy microarray image is filtered to obtain denoised image, the performance parameters are calculated to find whether these techniques are effective for denoising.

4.2 Convolution Neural Network (CNN)

The convolutional neural network (CNN) system is as shown in Fig. 4. It is a type of improved artificial neural network that is used to classify images, group them according to similarity and identify objects. They interpret images as three-dimensional objects rather than flat canvases with only width and height measurements, and they assimilate them as three distinct layers of colour layered one on top of the other. To receive a normal colour image in the form of a rectangle box whose width and height are determined by the number of pixels along their lengths, the depth is three layers deep, one for each RGB colour. Deep learning toolbox R2020A gives

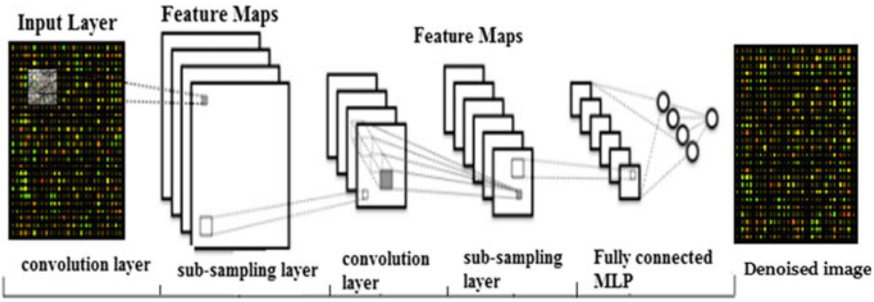


Fig.4 Convolutional neural network system

a system to structuring and executing profound neural systems with calculations, pre-prepared models and applications.

For image handling, analysis, perception and computation improvement, the image processing toolbox provides a comprehensive range of reference standard calculations and work process applications. It has access to typical image processing operations and can automate them.

5 Results

Tables 1, 2 and 3 show the comparison of performance parameters for Gaussian

Table 1 Comparison of PSNR (dB) obtained from filtering techniques and CNN

PSNR (dB)	Noise 10%	Noise 20%	Noise 50%	Noise 90%
Image	20.7268	16.9769	10.0596	6.2580
Median filter [4]	21.1892	17.2281	10.0113	0.0241
Weiner filter [5]	20.7164	16.8659	9.9707	6.2370
Linear filter [18]	20.8000	17.0165	10.0972	6.2266
Proposed CNN	30.9143	30.6562	29.9633	29.4354

Table 2 Comparison of MSE obtained from filtering techniques and CNN

MSE	Noise 10%	Noise 20%	Noise 50%	Noise 90%
Image	0.8914	0.7835	0.4516	0.0850
Median filter [4]	0.8780	0.7572	0.4059	0.0306
Weiner filter [5]	0.8432	0.7322	0.3813	0.0331
Linear filter [18]	0.8932	0.7847	0.4167	0.1000
Proposed CNN	0.8101	0.8597	1.0084	1.1397

Table 3 Comparison of SSIM obtained from filtering techniques and CNN

SSIM	Noise 10%	Noise 20%	Noise 50%	Noise 90%
Image	0.6684	0.7319	0.6532	0.4961
Median filter [4]	0.8475	0.8223	0.7245	0.5632
Weiner filter [5]	0.7905	0.7838	0.6954	0.5661
Linear filter [18]	0.884	0.8344	0.7154	0.5084
Proposed CNN	0.9579	0.9447	0.9145	0.9152

noise evaluated using filtering techniques and CNN. The accuracy and quality of denoising schemes influence the selection of these parameters. The following are the parameters:

I. **Mean square error (MSE)** is defined as the cumulative squared error between the input image and the denoised image, and it should be less than one. It is calculated as shown in Eq. (1)

$$\mathrm{MSE} = \frac{1}{P * Q} \sum_{i=1}^{P-1} \sum_{j=1}^{Q-1} [I(i, j) - O(i, j)]^2 \tag{1}$$

II. **Peak signal-to-noise ratio (PSNR)** is an estimated measurement of the quality of the reconstructed image. The PSNR is calculated as in Eq. (2)

$$\mathrm{PSNR} = 10 \log \frac{255^2}{\mathrm{MSE}} \tag{2}$$

III. **The structural similarity index (SSIM)** was proposed by Wang et al. [19].

$$\mathrm{SSIM} = \frac{\left(2\mu_x \mu_y + c_1\right)\left(2\sigma_{\mathrm{xy}} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{3}$$

Tables 1, 2 and 3 show comparison between median filter, linear filter, Weiner filter and proposed CNN method in terms of PSNR, MSE and SSIM values with different densities of noise. Results show that PSNR is high and SSIM is closely equal to one in CNN method as compared with traditional filter methods.

6 Conclusion

In this paper, the major concentration is upon the enhancement of DNA microarray image. Here, we use the basic traditional filtering techniques and CNN method on the microarray image for its comprehensive enhancement. The quantitative is calculated for each technique and is compared with CNN model. We can see from the findings that the CNN approach produces better residues than the traditional methods. The experimental results demonstrate that the proposed algorithm implements effectively in various noise densities. These results prove that our algorithm is more exclusively feasible and meaningful.

Future work could include extending the suggested approach to multiple metrics for characterising and measuring the qualities of a picture, as well as those of a filtering process, colour pictures and video framework that can help in improving video denoising.

References

1. Leung YF, Cavalieri D (2003) Fundamentals of cDNA microarray data analysis. Trends Genet 19
2. Lukac R, Plataniotis KN (2005) A data-adaptive approach to cDNA microarray image enhancement. LNCS 3515, pp 886–893
3. Tian C, Xu Y, Fei L, Wang J, Wen J, Luo N (2019) Enhanced CNN for image denoising. CAAI Trans Intell Technol
4. Shao G, Mi H, Zhou Q, Luo L (2009) Noise estimation and reduction in microarray images. In: WRI world congress on computer science and information engineering. Los Angeles, California, USA
5. Nagaraja J, Pradeep BS, Manjunath SS, Karthik SA (2017) An efficient technique for enhancement of microarray images. Elsiever
6. Ng P-E, Ma K-K (2006) A switching median filter with boundary discriminative noise detection for extremely corrupted images. IEEE Trans Image Process 15(6):1506–1516
7. Liu Z, Yan WQ, Yang ML (2018) Image denoising based on a CNN model. In: Fourth international conference on control, automation and robotics
8. Zuo W, Zhang K, Zhang L (2017) Convolutional neural networks for image denoising and restoration. Springer International Publishing
9. Zhao A (2016) Image denoising with deep convolutional neural network. Stanford University, Computer Science
10. Donald AA, Zhang Y, Parthe R (2006) On denoising and compression of DNA microarray images. Pattern Recogn 39(12):2478–2493
11. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition
12. Lim B, Son S, Kim H, Nah S, Lee KM (2017) Enhanced deep residual networks for single image super-resolution. In: IEEE conference on computer vision and pattern recognition workshops
13. Xie J, Xu L, Chen E (2015) Image denoising and inpainting with deep neural networks. CAAI Trans Intell Technol
14. Raza K (2015) Analysis of microarray data using artificial intelligence based techniques. arXiv: 1507.02870v1 10 July 2015
15. Sil D, Dutt A, Chandraell A (2019) Convolutional neural networks for noise classification and denoising of images. In: TENCON-2019. IEEE
16. Fu B, Zhao X-Y, Li Y, Wang X-H, Ren Y-G (2018) A convolutional neural networks denoising approach for salt and pepper noise. In: IEEE conference on computer vision and pattern recognition. Springer
17. Alagesan R, Manimekalai MAP (2013) An impressive method to remove high density salt-and pepper noise from microarray image. Int J Adv Res Electron Commun Eng 2(3). ISSN: 2278-909X

18. Smolka B, Lukac R, Chydzinskia A, Plataniotis KN, Wojciechowskic W (2003) Fast noise reduction in cDNA microarray images. Elsevier
19. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE

A Review on Parameters that Impact IoT Application—An Experimental Evaluation

Sakshi Gupta and Itu Snigdh

Abstract With the advent of the Internet of Things (IoT), diverse long-range, low-power, and low-bit-rate wireless network technologies have gained impetus. The success is, however, based on the capabilities of the underlying internetwork of "things." Our article focuses on a generic perspective of how scalability issues impact communications among IoT. We present a detailed analysis of the impacts of sensing range and environmental factors that affect the network performance. The simulations provide a realistic outlay to estimate the dependency of factors that can be controlled for expected IoT network performance and for deciding on configuring and reconfiguring IoT scenarios.

Keywords Internet of Things · Throughput · Packet delivery ratio · Controlled and uncontrolled factors

1 Introduction

Internet of Things is a concept based on ubiquitous computing. There are great literary works in [1–3] that focus on its definitions, significance, and requirements like security, interoperability, other characteristics, and the impact of being "smart." The background processes include sensor networks and intelligent devices communication. With an abundance of new applications and services, the future generation network involves a scalable smart system that is incessantly connected. These connections would be wireless, which would, in turn, be challenged by their inherent characteristics of battery constraints and environment dependence. The emerging WLAN (802.11 ah) has been developed to cater to IoT requirements. Since Zigbee networks (IEEE 802.15.4) are primarily used in wireless sensor networks, it most

S. Gupta (✉) · I. Snigdh
Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: phdcs10004.18@bitmesra.ac.in

I. Snigdh
e-mail: itusnigdh@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_19

appropriately finds its place in the enabling technologies for IoT applications. LoRa is one of the newest concepts and methodology with a good radio network compared to other long-range technologies like Bluetooth BLE [4], NB-IoT, LPWAN, and 5G.4G and 4G with LTE provide good quality of connectivity as required in IoT applications. With the current technological advancements, the challenges faced by 4G networks, namely more complicated communication, device computational capabilities, intelligence, would be overcome by 5G networks through wireless software-defined networking (WSDN). Therefore, technologies like 5G are claimed to be equipped with all basic requirements such as high throughput, low packet loss ratio, comprehensive coverage of devices with more battery life to cater to IoT environment constraints [5]. Every IoT application requires integrating live real-world data, including complex analytics, modeling, optimization, and visualization services, to make better operational decisions [1–3, 6].

This indicates that all the advantages of the communication channel and its flaws and limitations are applicable to any IoT application. Table 1 outlines the various standards and technologies from which IoT applications can customize advantages, but it solely depends on the type of application and environment where communication would take place. The choice of standards and protocols can surge more benefits owing to their pros and cons. Table 1 summarizes the communication technologies currently in use.

With IoT, requirements of diverse long range, low power, and low bit-rate have arisen in wireless networks. However, scalability comes as an issue when many end devices try to connect to the core network and is usually larger than the number of available channels, thus causing collisions and packet losses. Other issues are attributed to the prevalent environmental conditions. Figure 1 presents the overview of the factors that have been considered for analysis in the further sections. Related work has been outlined in Sect. 2. Section 3 explores the controlled and uncontrolled factors aforesaid which have been used for analysis and experimentation. Section 4 presents the analysis on what factors impact the network depending on the application type and scenario. Section 5 presents the analysis results followed by a conclusion in Sect. 6.

Table 1 State-of-the art technologies in IoT

S. No.	Types of communication	Technology used in communication
1	Short-range communication	Bluetooth BLE (802.15.1), RFID
2	Medium-range communication	Zigbee (802.15.4), Wi-Fi (802.11)
3	Long-range communication	(Licensed band) NB-IoT, LTE-M
		(Unlicensed band) LoRa, LoRaWAN
4	Cellular communication in IoT	2G, 3G, 4G, 4G with LTE, 5G

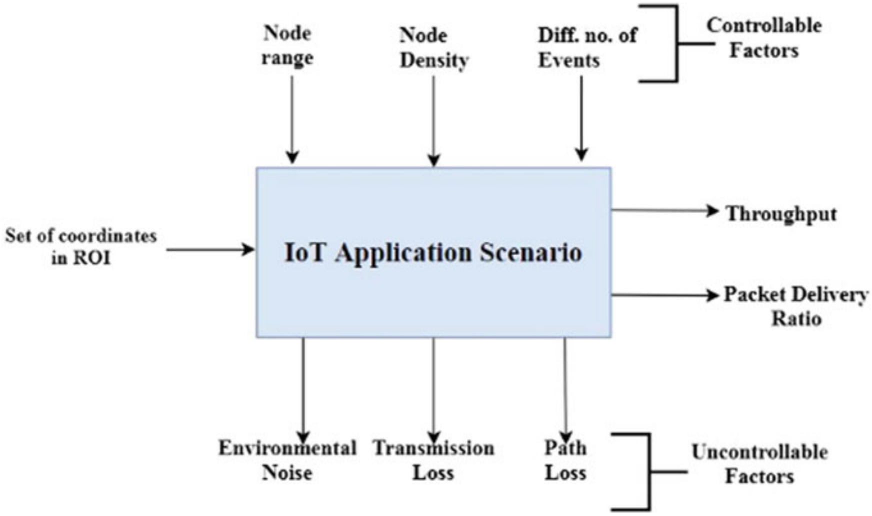


Fig. 1 Considered parameters for IoT application (updated diagram)

As the importance of these factors is significant in network communication, we have tried to identify the uncontrollable factors as well as models in predicting trends based on existing works of literature and as per our understanding. We have considered factors like path loss, shadowing, and fading models along with variable values of controllable factors for our experimental results.

Our article presents the requirements of applications in context to the network throughput. We also measure reliability in context to network failures experienced. The network failure is also assumed to be attributed to the number of failed transmissions, packet drops, and packets errored and presented as a packet delivery ratio (PDR).

2 Related Works

LoRa realizes the low energy and low resources concept [7] but only between end devices and gateway, limiting us for further analysis of LoRa in this article. Similarly, for 5G networks, the synchronization and harmony in terms of interoperability with existing technologies and applications to develop massive and critical IoT use cases are under analysis without any tangible results [8]. Further, Bluetooth technologies like BT5 beacons promise long battery and low maintenance with ease of installation but cannot satisfy the requirements of outdoor applications [9]. The medium-range 802.11 protocol, which is abundantly used, is not suitable for IoT applications, as reasoned in [10]. Comparison of 802.11 ah concerning 802.11 and 802.15.4 based on the throughput, delay, and association, time and coverage range have been mentioned in [11]. However, they do not address the factors that control the network behavior. We are also aware that multiple uses of heterogeneous devices are only possible if the devices communicate in an open and interoperable way [12]. The coexistence of Wi-Fi (802.11 b/g/n) and Zigbee (802.15.4) networks have also been presented in [13]. The performance analysis criteria are based on packet reception rate and throughput and are more oriented toward presenting the degradation and performance variation due to Wi-Fi (802.11x) and Zigbee (802.15.4) interference.

In our article, we employ the basic Wi-Fi and more frequently addressed Zigbee networks for our analysis. Since little has been discussed about the performance and impact of multiple devices connected on the communication interface, many factors affect wireless medium and the factors constraining wireless networks. For analysis, we enlist several factors that are categorized as controllable (chosen as per user intent or suitability of the application) and uncontrollable factors (parameters inherent as channel properties that carry on the necessary communication and cannot be varied as per the user's choice). Our article works on the different concepts to analyze the throughput, delay, and other network statistics when subjected to environmental and human-controlled factors.

3 Factors Affecting Communication in IoT Applications

3.1 Uncontrollable Factors

Environmental factors impact IoT-based applications and result in refraction, diffraction, saturation. We have categorized these factors that cannot be controlled by human intervention and assume them to be a constant value depending on the application while comparing variables controllable.

- *Environmental Factors*—attenuation and noise parameters are the main environmental factors affecting wireless communication systems and have been considered for our analysis. We also know that the higher the attenuation more is the distortion in communication. In wireless communication, attenuation depends on space, transmission distance, and physical surroundings like walls, windows, and water bodies. Thus, an IoT application that uses the wireless medium for data transfer is affected by attenuation depending on the situation (weather conditions) and position (Indoor and outdoor application) on the transmitter and receiver. Noises are universally present in communication systems due to natural causes and manmade noise (MMN), which are measured by signal-to-noise ratio (SNR) and can be resolved by repeaters. Radio noise is the aggregate of emissions from multiple sources that do not originate from the transmitter [14]. We consider natural noise that would also be present in any IoT application. This includes thermal, atmospheric, intermodulation, and impulse noise. In addition, refractions cause degradation in signal strength [14–17].

- Path Loss—path loss can be caused mainly by reflection, refraction, absorption, shadowing, scattering. Apart from these, path loss occurs due to the type and number of office equipment used, number of floors, and construction materials used in buildings. These also impact the received signal strength in communication [18], and hence, these conditions apply for IoT applications like smart cities. The signal strength estimates path loss based on models like free space propagation model, log-distance path loss model, and outdoor and indoor propagation model. Common models like Okumara model can model these, HATA Suburban, HATA Urban, COST 231 HATA Suburban, COST 231 HATA Urban, Indoor Offices, Indoor Factory, and Indoor Home [19].
- Transmission Loss—transmission loss is measured as the deviation in signal strength from an average value and affects how the devices connect and communicate with each other over a network. Loss in transmission power can be a result of fluctuation and variation in the received power signals due to various objects present in the path between transmitters and receivers or due to changes in the transmission medium and atmospheric conditions. These are referred to as shadowing and fading correspondingly in wireless communications.

3.2 Controllable Factors

For a generalized model of IoT applications aiding the realization of an IoT, we consider four important factors that can be controlled to reduce congestion, effective deployment, better reliability, and improved transmission reception. Therefore, we adopt variables like—

- Node range—it defines the distance a node can communicate. Generally, node range depends on transmission power, propagation losses, and receiver sensitivity. For our proposed work, we consider variable radio range accordingly set for simulations.
- Node density describes the number of nodes in the network, and in our simulation, we have considered node numbers varying as 4, 16, 25, 38, and 81.
- Number of events—This indicates the number of concurrent transmissions that define the number of communicating applications in one communication environment. Our simulation varies the number of applications from one application per simulation to four applications transmitting simultaneously.
- Routing algorithms—due to the ad hoc nature of communication, our simulations consider the routing algorithms applicable to MANETS and IoT specifically.

4 Simulation and Analysis

For studying the impact of the aforementioned controllable factors under the constraints of uncontrollable factors on the performance of IoT applications, we

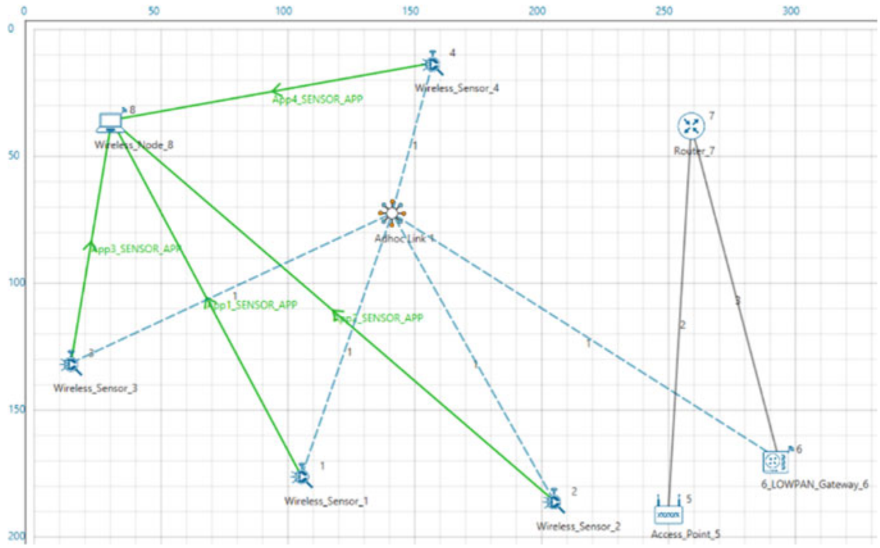


Fig. 2 Netsim scenario (static nodes)

Table 2 Variables for AODV experiments with no mobility (PDR metrics)

Fixed variables	Values	Model used
Application type	Sensor app	
Channel characteristics	Path loss model	Log distance, 4
Channel characteristics	Fading model	Rayleigh
Channel characteristics	Shadowing model	Log normal, 10
No. of nodes	4, 16, 25, 38, 81	
No. of applications	1, 2, 3, 4	

simulate our model in NetSim version 11.1. We consider a 50 × 50 region of interest to be fixed with varying the mentioned parameters iteratively. We assume randomly deployed nodes communicating with each other via UDP-based applications. Figure 2 depicts a sample scenario with nodes communicating to a sink node via a low-pan gateway. It is assumed that all nodes are static, and the network parameters assumed are mentioned in Table 2. Figure 3 represents a flowchart that shows how parameters will analyze in an IoT scenario.

4.1 Formula Used

- $\mathrm{PDR(data\ only)} = \frac{\text{Number of data packets successfully transmitted}}{\text{Number of data packets generated (Collided + Errored)}}$
- $\mathrm{PDR\ (data + control)} = \frac{\text{Number of data and control packets successfully transmitted}}{\text{Number of packets generated (data + control)}}$

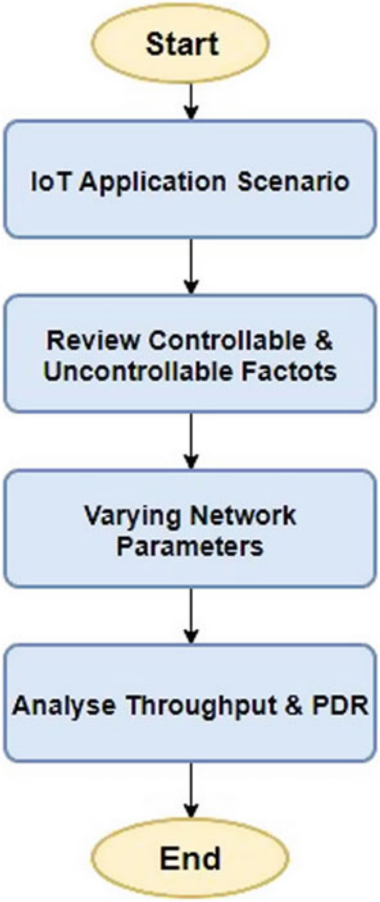


Fig. 3 Flowchart of parameter analysis

- $\text{Throughput} = \frac{\text{Total Payload Delivered to Destination in Bytes} * 8}{\text{Simulation Time (in microseconds)}}$

4.2 Simulation Parameters

The presented scenario is simulated with varying numbers of static nodes and applications in the AODV routing protocol to analyze the packet delivery ratio. We consider the log distance path loss model as it appropriately applies to Wi-Fi and Zigbee networks and is most commonly used for vehicle-to-vehicle scenarios and other related scenarios like the highway, urban intersection, and garage scenarios as it considers background path loss [20, 21]. The path loss model depends on the exponent value "*n*," which reflects the propagation environment. For propagation that nearly follows a free-space model or line of sight (LOS), "*n*" is set to 2, and propagation that is categorized as urban, suburban, or non-line of sight (NLOS), "*n*" is

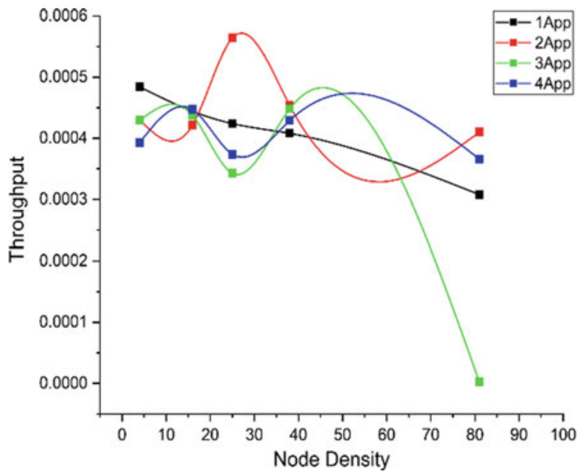


Fig. 4 Impact of node density on throughput

greater than 2. We assume the value of n as 4. We also consider log normal shadowing model with the shadowing factor with a standard deviation value 10 dB as generally defined values between 6 dB and 12 dB, but for large-scale shadowing in urban scenarios' owning to trees and other interferences, we assume a value between 8.2 dB and 10.6 dB [22]. For fading, the Rayleigh model is considered as only NLOS components are being considered, and we assume that no LOS path exists between transmitter and receiver in empirical urban studies as per references [23].

5 Results

5.1 Analysis of Node Density Variations on Throughput

Figure 4 depicts the different node densities and their impact on throughput. The graph is plotted for five node densities by placing different nodes in the same area 50×50 grid lengths. By increasing the number of applications, the communication interface is observed for the resulting throughput as shown in the graph and observed that as the node density increases, for example, in the case of 81 nodes, the throughput obtained is negligible as packets are dropped due to congestion for four simultaneous applications. However, for the other applications, throughput is comparatively high.

5.2 Analysis of Node Density Variations on PDR

Figure 5a and b are plotted to map different node densities' impact on the packet delivery ratio (PDR). Here, the reliability is represented as the packet delivery ratio

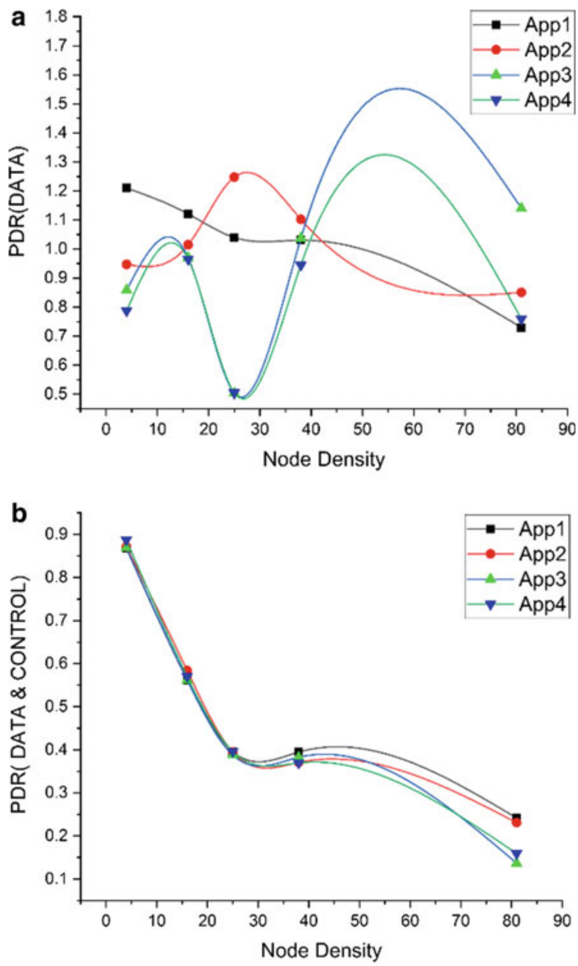


Fig. 5 **a** Impact of node density on PDR considering the only payload, **b** impact of node density on PDR considering data and control packets

considering only data packets and then all the packets, including both data and control. Here, we observe that when there is only one transmitter, the PDR for data packets reduces gradually with an increase in the number of nodes. However, when the number of simultaneous transmitters is increased, the PDR is non-uniform and erratic as compared to the PDR inclusive of both data and control packets. Hence, we cannot confirm that the PDR is solely dependent on the node density.

Table 3 shows the results data in tabular format.

Table 3 Numerical data values of results

Node density	Delay			
	App1	App2	App3	App4
4	90,040.95651	124,759.3416	114,385.5987	129,390.7264
16	283,507.175	308,567.9064	288,264.6634	292,993.4163
25	310,923.3468	263,195.8437	352,775.991	400,964.7909
38	412,260.4513	320,366.8748	419,570.8152	415,299.5598
81	341,318.1291	402,324.9155	91,343.0878	427,011.1418
Node density	PDR data			
	App1	App2	App3	App4
4	1.21	0.947136564	0.858666667	0.78757515
16	1.1212	1.014423077	0.970414201	0.965442765
25	1.04	1.247787611	0.503921569	0.504723347
38	1.0303	1.101941748	1.037037037	0.944933921
81	0.73	0.850622407	1.140828402	0.759447466
Node density	PDR data and control			
	App1	App2	App3	App4
4	0.8677	0.871381458	0.869037772	0.887037882
16	0.5614	0.584042925	0.561767956	0.570380463
25	0.395592498	0.39444909	0.387991038	0.396206079
38	0.394607589	0.370697129	0.382548914	0.369728203
81	0.241723788	0.231180647	0.136855796	0.159190675

6 Conclusion

Our article presents an elaborate analysis of the significant factors that affect communications in ad hoc network-based IoT applications. We have analyzed environmental and human or application-specific constraints like fading, path loss, number of nodes, and sensing range in the network parameters. We have also analyzed the impact of variance of these factors thoroughly. Therefore, in applications that need more simultaneous transmissions like WBANs and smart transportation, we should adopt scheduling, clustering, and data aggregation at edge devices.

Comparing the impact of node density and the number of transmissions on PDR shows that PDR solely depends on the number of applications communicating under fixed environmental conditions. However, network performance degrades when we consider larger values of fading and interference, even for a smaller number of communicating applications. Thus, we infer that for IoT applications; we need to consider environmental factors for computing the network performance.

References

1. Ahad MA, Paiva S, Tripathi G, Feroz N (2020) Enabling technologies and sustainable smart cities. Sustain Cities Soc 61:102301
2. Ismagilova E, Hughes L, Dwivedi YK, Raman KR (2019) Smart cities: advances in research—an information systems perspective. Int J Inf Manage 47:88–100
3. Buyya R, Dastjerdi AV (eds) (2016) Internet of things: principles and paradigms. Elsevier
4. Collotta M, Pau G, Talty T, Tonguz OK (2018) Bluetooth 5: a concrete step forward toward the IoT. IEEE Commun Mag 56(7):125–131
5. Li S, Da Xu L, Zhao S (2018) 5G internet of things: a survey. J Ind Inf Integr 10:1–9
6. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M (2014) Internet of things for smart cities. IEEE Internet Things J 1(1):22–32
7. Pötsch A, Haslhofer F (2017) Practical limitations for deployment of LoRa gateways. In: 2017 IEEE international workshop on measurement and networking (M&N). IEEE, pp 1–6
8. Goudos SK, Dallas PI, Chatziefthymiou S, Kyriazakos S (2017) A survey of IoT key enabling and future technologies: 5G, mobile IoT, sematic web and applications. Wireless Pers Commun 97(2):1645–1675
9. Boukhechba M, Bouzouane A, Gaboury S, Gouin-Vallerand C, Giroux S, Bouchard B (2017) A novel bluetooth low energy based system for spatial exploration in smart cities. Expert Syst Appl 77:71–82
10. Khorov E, Lyakhov A, Krotov A, Guschin A (2015) A survey on IEEE 802.11 ah: an enabling networking technology for smart cities. Comput Commun 58:53–69
11. Ahmed N, Rahman H, Hussain MI (2016) A comparison of 802.11 ah and 802.15. 4 for IoT. Ict Express 2(3):100–102
12. Waher P, Klauck R (2018) Internet of Things-discovery
13. Zhao Z, Wu X, Zhang X, Zhao J, Li XY (2014) ZigBee versus Wi-Fi: understanding issues and measuring performances of their coexistence. In: 2014 IEEE 33rd international performance computing and communications conference (IPCCC). IEEE, pp 1–8
14. ITU-R (2012) Radio communication sector of ITU, Methods for measurements of radio noise. https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.1753-2-201209-I!!PDF-E.pdf. Accessed 12 Jan 2020
15. Sean Victor Hum (2018) Atmospheric effects, radio and microwave wireless systems. http://www.waves.utoronto.ca/prof/svhum/ece422/notes/20a-atmospheric-refr.pdf. Accessed 12 Jan 2020
16. Siamack G, Jamil H, Tarlochan SS, Serguei P (2012) Effect of impulse noise on wireless relay channel. Wirel Sens Network
17. Stallings W (2009) Wireless communications and networks. Pearson Education India
18. Rath HK, Timmadasari S, Panigrahi B, Simha A (2017) Realistic indoor path loss modeling for regular Wi-Fi operations in India. In: 2017 twenty-third national conference on communications (NCC). IEEE, pp 1–6
19. Goldsmith A (2005) Wireless communications. Cambridge University Press
20. Ozpolat M, Bhargava K, Kampert E, Higgins MD (2021) Multi-lane urban mmWave V2V networks: a path loss behaviour dependent coverage analysis. Veh Commun 30:100348
21. Olasupo TO, Otero CE, Otero LD, Olasupo KO, Kostanic I (2017) Path loss models for low-power, low-data rate sensor nodes for smart car parking systems. IEEE Trans Intell Transp Syst 19(6):1774–1783
22. Abhayawardhana VS, Wassell IJ, Crosby D, Sellars MP, Brown MG (2005) Comparison of empirical propagation path loss models for fixed wireless access systems. In: 2005 IEEE 61st vehicular technology conference, vol 1. IEEE, pp 73–77
23. Chourabi H, Nam T, Walker S, Gil-Garcia JR, Mellouli S, Nahon K, Pardo TA, Scholl HJ (2012) Understanding smart cities: an integrative framework. In: 2012 45th Hawaii international conference on system sciences. IEEE, pp 2289–2297

A Novel Machine Learning Approach to Delay Efficient Offloading Strategy for Mobile Edge Computing

Piyush Bharti, Srishti Chaudhary, and Itu Snigdh

Abstract There have been studies on the problems experienced in offloading. But, how to leverage machine learning for resource-efficient computation offloading is not completely explored yet. We present this paper with the motivation to leverage machine learning and backward induction to suggest a method for the resource-efficient computation offloading in IoT devices. In particular, we use machine learning to predict the most efficient device to which a task should be offloaded through the use of a two-model machine learning system deployed in the edge cloud itself, since most high power IoT devices are fixed in their positions and display a certain pattern in their CPU utilization data. This is followed by using backward induction to determine the optimal execution location for each task component.

Keywords Mobile edge · Cloud · Resource · Offloading · Machine learning

1 Introduction

Cloud computing is defined as the on-demand availability of IT resources. A resource in the cloud is a computer system resource such as a computing resource, storage resource or memory [1]. As we move toward artificial intelligence and machine learning, applications are becoming more and more resource intensive. Intelligent IoT devices also make use of complex artificial intelligence and machine learning algorithms [2]. These devices need resources to complete their task. Instead of using far off and centralized data center resources, systems now resort to employing decentralized servers closer to the user devices for processing data. One way is to offload these tasks to the public cloud but it comes with the drawback of increased latency [3]. To overcome the shortcomings of offloading tasks to public cloud, a new paradigm

P. Bharti · I. Snigdh
Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, India
e-mail: itusnigdh@bitmesra.ac.in

S. Chaudhary (✉)
Software, Development, Microland Limited, Bengaluru, India
e-mail: chaudhary1srishti@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_20

known as mobile edge computing (MEC) [4, 5] is emerging. Mobile edge computing is implemented at cellular base stations and other edge nodes. In a mobile edge system, the cloud computing capabilities and an IT service environment are enabled at the edge of the network leveraging mobile base stations. Usually the resources are situated at the edge of the network and hence called edge devices. These computing devices bridge the gap between IoT end devices and the distant cloud servers. The devices are nevertheless, in comparison to the cloud resources, resource constrained in terms of computational capabilities and power source. They are usually heterogeneous, as they house processors with different architectures and also dynamic as their workloads change. Mobile edge computing alleviates the computation and power limitation of mobile devices. Additional latency is incurred when offloading tasks to a remote server. When users offload computing tasks, they have little information about the wireless networks to be accessed, including traffic load of the accessed network and the computation load of edge clouds [6]. Even though the capacity of edge cloud is more than that of a single IoT device, it is still limited because edge cloud has to handle a lot of devices. Therefore, for efficient task placement and utilization of computing resources, task offloading is suggested.

MEC technology can be utilized by IoT devices to offload their computationally intensive tasks to the edge cloud and operate in real time with low latency. Since, intelligent Internet of Things (IoT) devices and applications run complex algorithms and need fast processing of data, they require greater processing and computation power than other devices. Also, these intelligent IoT devices are incapable of achieving very high processing capabilities on their own because of the limitation on the physical size of computation resources. To meet the computational and processing requirements of intelligent IoT devices, task offloading strategy is used. In this strategy, the complex processing tasks are offloaded to the edge cloud. However, offloading a large number of tasks to the edge cloud results in latency due to the long network distance between IoT devices and clouds and higher payments for the cloud services by the user. Resource-efficient offloading is a way to minimize the number of tasks offloaded to the edge cloud while satisfying the QoS requirement by first trying to offload tasks to nearby helper devices with high computational capabilities [7–9]. Thus, reducing the occupancy on the edge cloud.

A. *Related work*

There are many constraints related to the use of MECs. Each device that aids in realizing the distributed computing resources architecture in the edge needs to provide support for multiple applications. Additionally, they also require to be energy efficient and sensitive to delay for supporting real time applications like for mobile healthcare [10], smart grid [11], Internet of vehicles, face recognition, interactive gaming and augmented reality to name a few [12]. A simple solution is to deploy numerous Fog notes (FNs) in the edge network. The end users could then offload their tasks to one of Fog servers or the cloud server thereby reducing the backbone traffic as well as latency for delay sensitive services. In applications like Internet of vehicles, it is needed to reduce the perception reaction time due to high speed mobility and inherent characteristics. For this, Collaborative task offloading and transmission mechanisms have been used. Similarly, latency aware workload offloading problem (LEAD) [13] have mathematically formulated the problem for minimizing the average response time for mobile users. Also, some literatures propose transmission power allocation of mobile users in conjunction with the computing resource allocation. In contrast to this, online optimization-based algorithm to solve resource allocation problem of mobile users at each time slot have been studied by [14]. Maximum allowed computing delays for different IoT application types and their resource requirements have been analyzed in [15]. The experimental evaluation in [7, 16] shows that the location of edge servers and the visualization technology used highly impacts the latency experienced.

B. *Related work*

Our primary goal is to propose a resource-efficient computation offloading strategy. We try to offload a minimum number of tasks to edge cloud while adhering to the QoS constraint. Instead, we first try to offload tasks to nearby helper devices which have desired computational capabilities. This approach ensures better utilization of resources in the mobile edge system and concurrent offloading of tasks to nearby IoT devices. Additionally, the lesser the number of tasks offloaded to the edge cloud, the lower the latency due to the long network distance between IoT devices and clouds. Smaller cloud occupancy also results in a lesser payment by the user for cloud services. Formally, we define the resource-efficient offloading problem as determining the percentage of tasks that should be offloaded to the edge cloud, subject to the constraint that the completion time of the task is within the deadline.

Previous works [17] on resource-efficient computation offloading do not take into consideration the communication strength of the target device with the helper devices. Moreover, they do not take into account the processing capacity of the helper devices. Our proposed method not only takes these factors into consideration but also the addition of any new parameters can be done easily since we rely on machine learning models. Thus, the scope of expanding our model to different scenarios is also increased.

2 Proposed Methodology

A. *Machine Learning-based Approach*

Our primary goal is to propose a resource-efficient computation offloading strategy. We try to offload a minimum number of tasks to edge cloud while adhering to the QoS constraint. Instead, we first try to offload tasks to nearby helper devices which have desired computational capabilities. This approach ensures better utilization of resources in the mobile edge system and concurrent offloading of tasks to nearby IoT devices. Additionally, the lesser the number of tasks offloaded to the edge cloud, the lower the latency due to the long network distance between IoT devices and clouds. Smaller cloud occupancy also results in a lesser payment by the user for cloud services.

Formally, we define the resource-efficient offloading problem as determining the percentage of tasks that should be offloaded to the edge cloud, subject to the constraint that the completion time of the task is within the deadline. We start by studying the percentage of CPU occupancy of the IoT devices at different times during the day. We train our first machine learning model to predict the percentage of CPU occupancy for different helper devices on any given day at any time. The first model takes time-series data of CPU utilization of IoT devices and forecasts the CPU utilization of any given device at any given date and time. For forecasting, we have used Facebook's Prophet Library [18] that gives us the ability to make better predictions with the parameters for seasonalities, change points and holidays that make it versatile and easy to adapt to any dataset. Since our dataset showed daily seasonality, we tuned it accordingly. For datasets exhibiting weekly, monthly, or yearly seasonality prophet's parameters can be tweaked to get accurate results. The process followed is elaborated in Fig. 1.

B. *Algorithm Used*

The algorithm used for our proposed work can be illustrated as:

Input: processing power of the helper device, the strength of communication between the target device and the helper device, and the percentage of CPU occupancy for the helper device (obtained from the first machine learning model).

Process:

Step I: Train model one to predict CPU utilization time and model two to predict total execution time including latency for any helper device.

Step II: Compute the time taken to execute a task component on different helper devices.

Step a: Determine the time each helper device will take to execute a task using models one and two.

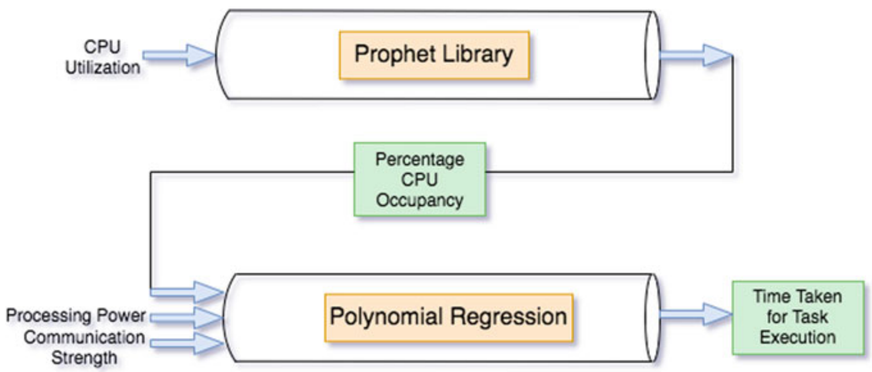


Fig. 1 Proposed method

Step b: Determine if a task component should be executed on the target device (n), helper device(h), or the edge cloud (c) based on the output from step (a).
Step c: iterate for all the task components.

Step III: Check if the total execution time obtained from the backward induction method is less than or equal to the maximum completion or deadline time for that task.
Step IV: If the deadline is met then we have the optimal execution locations else go to Step II and select more powerful edge cloud VM and recomputed the total time.

3 Results

As illustrated in Fig. 2, we evaluate the percentage of tasks offloaded to edge cloud as a function of the completion time requirement. To meet the constraint that the completion time of the task should be within the deadline, the percentage offloading varies. An increase in the completion time requirement implies that a small percentage of tasks will be offloaded to the edge cloud and vice versa. We can see from Fig. 2 that with our optimum virtual machine selection, we are able to reduce the completion time of the offloaded tasks.

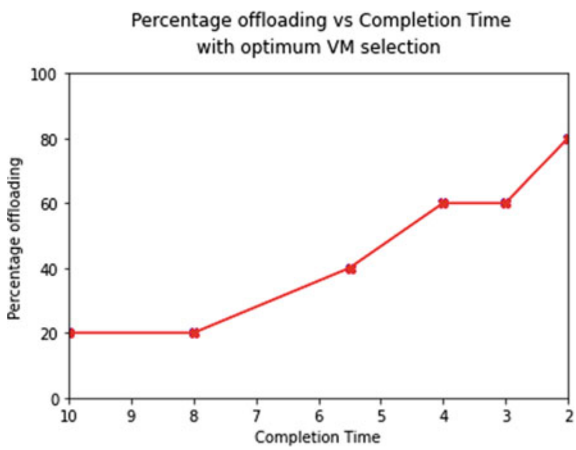


Fig. 2 Simulation results of percentage offloading for optimum virtual machines with their completion times

4 Conclusion

Our paper proposes the use of machine learning to select the best helper device to offload a given task. The main advantages of our approach are that we consider CPU utilization, processing power and communication strength to select the best device. In addition, we propose that the machine learning models be deployed in the cloud and therefore, it would require just one poll to the cloud to get the execution time and latency for the helper devices. Finally, we use backward induction to determine the optimal execution location for each task component. From our results we were able to achieve a smaller completion time with better utilization of the helper devices for offloaded tasks. Our strategy can be adopted for real time systems where the computations are data intensive as well as our methodology is well suited for aiding offloading decisions like, when, where, how and what part of the job should be uploaded.

References

1. Velte T, Velte A, Elsenpeter R (2009) Cloud computing, a practical approach. McGraw-Hill, Inc.
2. Goap A, Sharma D, Shukla AK, Krishna CR (2018) An IoT based smart irrigation management system using Machine learning and open source technologies. Comput Electron Agric 155:41–49
3. Parikh SM (2013) A survey on cloud computing resource allocation techniques. In: 2013 Nirma university international conference on engineering (NUiCONE). IEEE, pp 1–5
4. Alameddine HA, Sharafeddine S, Sebbah S, Ayoubi S, Assi C (2019) Dynamic task offloading and scheduling for low-latency IoT services in multi-access edge computing. IEEE J Sel Areas Commun 37(3):668–682
5. Tran TX, Pompili D (2018) Joint task offloading and resource allocation for multi-server mobile-edge computing networks. IEEE Trans Veh Technol 68(1):856–868
6. Liu CF, Bennis M, Poor HV (2017) Latency and reliability-aware task offloading and resource allocation for mobile edge computing. In: 2017 IEEE globecom workshops (GC Wkshps). IEEE, pp 1–7
7. Lan Y, Wang X, Wang D, Liu Z, Zhang Y (2019) Task caching, offloading, and resource allocation in D2D-aided fog computing networks. IEEE Access 7:104876–104891
8. Chen X, Liu Z, Chen Y, Li Z (2019) Mobile edge computing based task offloading and resource allocation in 5G ultra-dense networks. IEEE Access 7:184172–184182
9. Mach P, Becvar Z (2017) Mobile edge computing: a survey on architecture and computation offloading. IEEE Commun Surv Tutorials 19(3):1628–1656
10. Wu JH, Wang SC, Lin LM (2007) Mobile computing acceptance factors in the healthcare industry: a structural equation model. Int J Med Informatics 76(1):66–77
11. Yang C, Chen X, Liu Y, Zhong W, Xie S (2019) Efficient task offloading and resource allocation for edge computing-based smart grid networks. In: ICC 2019–2019 IEEE international conference on communications (ICC). IEEE, pp 1–6
12. Kan TY, Chiang Y, Wei HY (2018) Task offloading and resource allocation in mobile-edge computing system. In: 2018 27th wireless and optical communication conference (WOCC). IEEE, pp 1–4
13. Sun X, Ansari N (2017) Latency aware workload offloading in the cloudlet network. IEEE Commun Lett 21(7):1481–1484

14. Mseddi A, Jaafar W, Elbiaze H, Ajib W (2019) Intelligent resource allocation in dynamic fog computing environments. In: 2019 IEEE 8th international conference on cloud networking (CloudNet). IEEE, pp 1–7
15. Fan Q, Ansari N (2018) Application aware workload allocation for edge computing-based IoT. IEEE Internet Things J 5(3):2146–2153
16. Garcia Lopez P, Montresor A, Epema D, Datta A, Higashino T, Iamnitchi A, Barcellos M, Felber P, Riviere E (2015) Edge-centric computing: Vision and challenges
17. Chen X, Shi Q, Yang L, Xu J (2018) ThriftyEdge: resource-efficient edge computing for intelligent IoT applications. IEEE Netw 32(1):61–65
18. https://facebook.github.io/prophet

Design, Installation and Performance Analysis of an On-Grid Rooftop Solar PV Power Plant for Partial Fulfillment of Common Load

Sudipta Sarkar, Ajoyjit Ghosh, and Ashoke Mondal

Abstract With shortage of fossil fuels like coal, petroleum the energy generation is depicted toward renewable energy sources like solar, wind, biomass, etc. The renewable energy technologies present an emission free energy generation technique toward a sustainable tomorrow. Rooftop solar power plant (RTPV) is one of the good solar power generation technique. In this paper, a brief description on design, commissioning and techno economic analysis of a 50Kw$_p$ rooftop solar power plant design in Uluberia super specialty hospital Howrah, India have been described. The electricity generation in both input DC and output AC end of each inverter is recorded and monitored and analyzed in this paper. The input end DC power of the inverter is coming from solar panels. The open circuit voltage (V_{oc}), short circuit current (I_{sc}), maximum power (M_p) is noted as 48 V, 50 A, 49 KW, respectively, during 11am to 2 pm which is the bright sunny time of the day in West Bengal. The output power of the inverter is connected with the main transmission grid. The grid connectivity rule of solar power plant with main grid of West Bengal State Electricity Distribution Company Limited (WBSEDCL) is briefly mentioned. The generated current, voltage, power by solar panels of the RTPV power plant is taken out using a remote data logger to a dedicated cloud server and analyzed the output data in this paper. The total load of the hospital building is 450 kva, which was paid earlier to WBSEDCL. The plant is generating an average of 5000 units per month, saving Rs. 511,817 per year. The Capacitive Utilization Factor (CUF) of this solar power plant is 14.7%. An attempt has been made to calculate the payback period or break-even years is approximately 5 years, Reduction of CO_2 emission or carbon credit due to the plant has also been shown. A seasonal data analysis has been done to predict average output during any particular month of the year. Finally, it has been observed that this type of energy

S. Sarkar
Survey Engineer, Vikram Solar Ltd., Kolkata, India

A. Ghosh (✉)
Junior Engineer, Geetanjali Solar Enterprise, Kolkata, India
e-mail: ghoshajoyjit@gmail.com

A. Mondal
Department of Electronics and Communication Engineering, RCC Institute of Information Technology, IndiaKolkata

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_21

generation unit is suitable for all domestic as well as commercial applications for grid connected and off grid system with suitable battery backup in remote areas or sustainable and greener environment in the future.

Keywords Solar PV · Rooftop solar power plant · Islanded mode · Array junction box (AJB) · Performance ratio (PR) · Data logger · Renewable energy technology (RET)

1 Introduction

The main sources of energy generation in developing countries like India are non-renewable energy sources like coal, oil and gas [1]. These types of energy have limited resources and make a huge impact on environment. In India the main source of electricity generation is coal [2]. As of today, 198.5 GW of thermal power plant is installed in the country. Due to its limited resources, the future of electricity generation is delineated toward renewable energy resources like solar, wind, hydroelectric, etc [3]. These types of resources are environment friendly and have unlimited resources from nature. Among all other renewable energy sources, solar energy is the most popular energy generation to produce electricity in India [4]. Solar energy is obtained from sunlight which a thermal energy it converts to electrical energy via a photo diode [5]. Till now 37.6 GW of solar PV installed in the country out of which 1922 MW of solar PV are rooftop. In national renewable energy act 2015 tells that government institute should have a renewable energy generation system in their own premises. Here in this paper the remote data scheme is introduced [6]. The proposed power generation scheme is graphically shown below (Fig. 1):

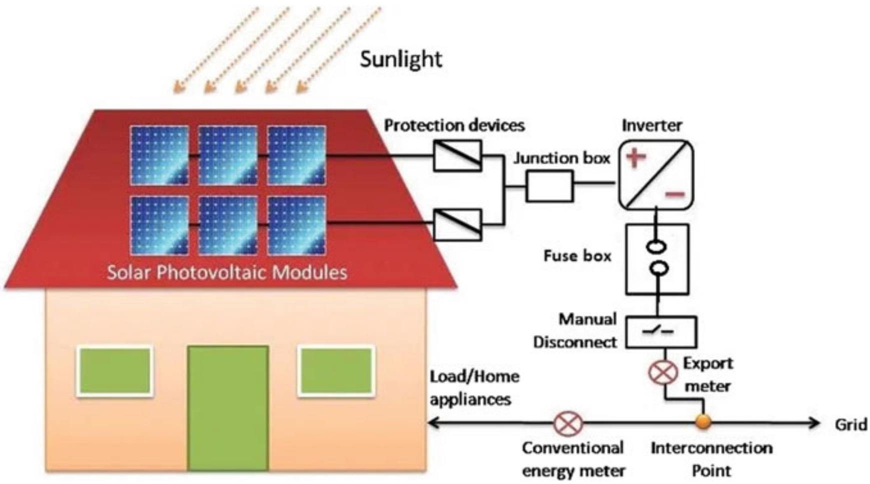


Fig. 1 Schematic diagram of the proposed power plant

2 Description of the Proposed Power Plant

The proposed PV system is an on-grid system which consists of solar panels, array junction box (AJB), inverter and a power conditioning meter which is further connected to grid. No storage device such as battery bank is used with this power plant.

A. *Site Description*

The geographical location of the power plant is Uluberia in the district of Howrah. The power plant is located in 22.469° N latitude and 88.095° E longitude at an altitude of 30 m from ground. The location is perfectly suitable for good solar energy generation. The average solar irradiance of the location is about 5.5 kWh/m^2/day with temperature between 13 and 37 °C throughout the year [7]. The survey is done on every buildings of the hospital before finally selected the super speciality hospital building which is the top most building and without any high rising beside the building. Thus, the power plant has not faced problems like shadow, smoke, etc., which makes lesser power plant maintenance cost. The total area of the rooftop is 1388.72 square meter.

B. *Plant Layout*

The solar PV modules are installed on the roof of new building. In this solar plant 310 square meter solar modules are placed, which is called active module area [8]. Whereas, the total plant area is 524.67 square meter. The modules are oriented at a fixed tilt angle of 23° and azimuth angle of 0° on the roof [9]. The PV modules with the power rating of 320 W_p are used in this power plant. Here 160 solar modules are used. The PV panels are connected in ten different orders. Thus, the module layout diagram is given below. This diagram is drawn on AutoCAD software (Fig. 2).

3 Design and Specification of Power Plant

The proposed rooftop solar PV power plant is consisting of solar PV modules, inverter, inverter, wires and protection fuses, etc [10]. The power plant is designed as it generates the maximum power. The total time taken for installation is about 4 months (From February 2019 to May 2019).

A. *Solar PV modules specifications*

In this power plant ware made 320 Wp modules are used. These modules are made using polycrystalline silicon technology [11]. These modules are cost effective and good for using in commercial rooftop solar PV plants. These modules are easily connected in series and parallel to adjust the voltage and current of the power plant (Table 1).

B. *Inverter Specifications*

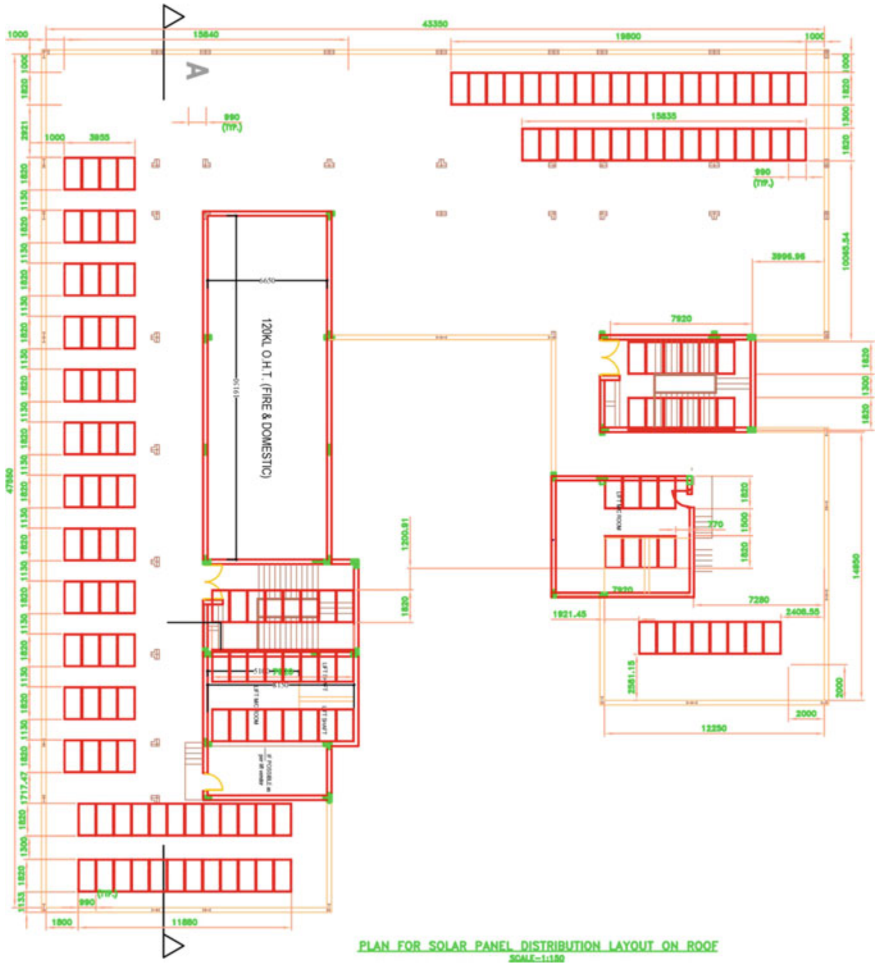


Fig. 2 Module layout diagram of the power plant

Here five inverters are used in this power plant to convert the DC power which is generated from the solar PV modules into AC power [12]. The inverters are Delta made and have the capacity of 10 KW. Each inverter is connected with each two strings of 5.12 KW, i.e., 10.24 KW solar PV. The inverter supports 20% of overloading of its full capacity. Anti-Islanding Protection is an important safety feature built into all grid connect inverters that can detect the loss of grid power in fractions of a second and switch off the inverter automatically [13]. It prevents sending power back into the inactive electricity supply lines when the grid is down due to fault or maintenance.

After installation of inverters, the time zone (GMT + 5.30), Country (India) and power factor (auto) is set in the control panel of the inverter. The technical specification is given (Table 2).

Table 1 Technical specification of the solar panel

Particulars	Specifications
Model number	WS-320
Mechanical characteristics	
Solar Cells per Module (Units)	72
Area (L*W*T) mm	1960*990*40
Weight (kg)	22.50
Thermal characteristics	
Operating temperature range (°C)	−40 to 85
Temperature coefficient of power (%/°C)	−0.4106
Electrical characteristics	
Open Circuit Voltage (Volt)	45.30
Short circuit current (Amps)	9.42
Maximum Power (Watt)	320
Maximum system voltage (V)	1000
Power tolerance	0/ + 5
Module efficiency (%)	12.50
Maximum series fuse rating (A)	15

Table 2 Technical specification of the inverter

Particulars	Specification
Model number	RPI-M10A
Dimension (l*b*w)	445*510*177
Weight (kg)	25
Operating temperature (°C)	−25 to 60
Noise level (dB)	< 50
INPUT (DC)	
Input DC power (KW)	12.5
Input voltage (V)	1000
Input current (I)	25
MPPT voltage range (V)	200–1000
No. of MPPT	2
OUTPUT (AC)	
Output power (kVA)	10.5
Output current (Amps)	16
AC voltage (V)	400, 3Phase
Output frequency (Hz)	50
Rated power factor	Unity
Efficiency (%)	98.30

4 Connection Methodology

In this power plant 16 modules are connected in series to make a string. Ten strings are made similarly followed by each two strings are connected in parallel in the array junction box (AJB). Depending upon the inverter MPPT input voltage range, the strings are made with 16 modules, i.e., 45.3 V $\times$ 16 = 724.8 V. Each AJB is further connected with an inverter. So, there are five inverters connected with five AJBs. The inverters then connected with an Inverter Interfacing Panel (IIP) to the feeding point. The feeding point is further connected to utility grid. There are several types of cables used to connect the different units of the power plant as follows:

- 4 and 6 mm^2 DC single core tinned copper flexible cables are used in the series connection between the PV modules.
- 6 and 10 mm^2 DC single core tinned copper flexible cables are used to connect the inverter and AJB.
- 4 mm^2 4-core AC copper flexible cable used between the inverter and IIP.
- 35 mm^2 4-core AC copper flexible cable used between IIP and feeding point.

The connection diagram of the rooftop solar power plant is given below (Figs. 3 and 4).

The picture of the installed solar rooftop power plant is shown below.

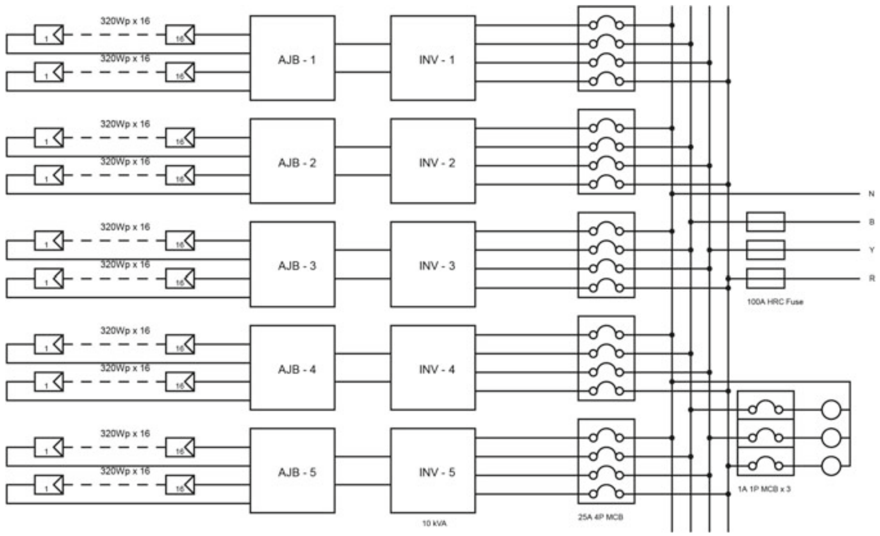


Fig. 3 Connection diagram of the power plant

Fig. 4 Picture of the installed rooftop solar power plant

5 Regulation of Distribution Grid Connectivity

This power plant is an on-grid type plant. So, it needs to connect with the local distribution grid. In West Bengal, West Bengal State Electricity Distribution Company Limited (WBSEDCL) is the local electricity distribution company. There is a power distributing unit (IIP) is installed followed by a three phase energy meter (kWh), which measures the total generated AC power from the solar power plant is supplied to the hospital's power distribution panel. The frequency, phase, voltage of the output AC power is matched with the power distribution panel bus. Where the frequency and phases are balanced with respected with the grid between 49.9 and 50.1 Hz and for three phases supply the voltage is kept nearly constant on 430 V.

6 Protection and Controls

A. DC side protection

- In this power plant all the module structures are grounded with chemical earthing. Chemical earthing is an earthing type, where the earthing electrode (Copper) is grounded surrounded by earthing compound consists of charcoal and salt (NaCl) to improve earth's pits resistance.
- The two terminals of each PV string are connected through 25 amps HRC fuses for over current protection to prevent excess current during faults in the power plant.

- DC Surge Protection Device (SPD) is placed in the AJB in parallel to strings to protect all the elements from lightning surges.

B. AC side protection

- 25 amps 4-pole Miniature Circuit Breaker (MCB) is used to isolate one inverter during fault or maintenance purpose.
- AC SPD is connected with AC bus to protect the power plant from any kind of surges.
- 3 nos. 100 amps HRC fuse is connected in between feeder and AC bus in series to protect any kind of faults form both sides.

Furthermore, three indicators (LED) are installed in each phase with the AC line to check the power supply of the power plant to the utility. Where each indicator is connected with 1amps 1-pole MCB.

7 Communication Interface

In this power plant Delta DelREMO—V2.0 data logger is used to server-based monitoring of the generation of power plant [6]. This data logger takes 230 V input which connects to the external AC supply of the hospital. Five inverters are connected with the data logger using cat5 cable in cascaded format. GSM enabled Data logger transmits the data to cloud of Delta Technologies. An Ethernet (RJ-45) port is available as an alternative of GSM. We get the data by logging in to their portal. The data logger gives data of 5 min interval throughout the day whereas monthly and yearly data of various parameters also generated. The following parameters are measured, displayed and recorded in the server.

- DC side voltage, current and power
- Frequency on AC side
- AC voltage, current and power
- Total energy
- Total yield of the power plant

An USB port is present in the data logger device to back up the data in a storage device (Preferably pen drive).

8 Output Generation

The output generation of the power plant from May 2019 to April 2020 is taken. The individual output generation through each inverter is recorded in every month. The output is varied in every month due to climate changes, which include temperature, humidity, rain, etc. The open circuit voltage (V_{oc}), short circuit current (I_{sc}), maximum power (M_p) of the power plant at running condition is noted as 48 V, 50 A, 49 KW, respectively, during 11 am to 2 pm which is the bright sunny time of the day (Table 3).

From the above table we can find out the total generated power of the power plant in every month from May 2019 to April 2020 (Table 4).

From the above table we can calculate that the total units of energy generation from May 2019 to April 2020 is 63,977.18 units. We can also observe that the generated energy variation throughout the year due to different season, different temperature and different solar irradiation. In winter, i.e., in the month of December and January, comparatively less energy generated from the plant. Whereas, in summer, i.e., March, April, May comparatively more amount of energy is generated.

The graphical representation of the generated energy data is given below (Fig. 5):

From the above table we can see that in the month of March and November, higher energy is generated compared to other months of the year. In the summer time, i.e., from April to June the plant generates a similar power of more than 5000 unit per month. In the winter months in India, i.e., December and January, the energy generation is comparatively lower than other time of the year. Whereas in July and September energy is generated 4640.31 and 4071.64 units, respectively, due to rainy season where cloudy sky makes hampered in the energy generation. Overall, we can see that the power plant can generate over 5000 units except rainy season and winter where climatic condition makes generation lesser.

Table 3 Monthly energy generation through each inverter

Month	Energy generation (kWh)				
	Inverter 1	Inverter 2	Inverter 3	Inverter 4	Inverter 5
May 19	1107.07	1116.00	1101.71	1029.65	1024.78
June 19	1134.25	1125.59	1106.31	1040.79	1034.09
July 19	967.75	960.38	939.85	887.82	884.51
Aug 19	1067.85	1064.88	1041.65	991.74	978.43
Sept 19	839.24	844.6	823.27	784.25	780.48
Oct 19	1140.67	1157.87	1118.43	1046.94	1040.32
Nov 19	1247.07	1269.0	1218.93	1170.91	1165.57
Dec 19	936.62	942.34	997.34	910.77	904.54
Jan 20	998.08	1023.38	1007.44	959.14	859.48
Feb 20	1158.32	1192.32	1183.87	1104.47	1087.2
Mar 20	1451.62	1488.43	1479.4	1364.81	1358.07
April 20	1085.32	1106.53	1093.79	1018.4	1013.13

Table 4 Monthly generation of the plant

Month	Energy generation (kWh)
May 2019	5379.21
June 2019	5441.03
July 2019	4640.31
August 2019	5144.55
September 2019	4071.84
October 2019	5504.23
November 2019	6071.48
December 2019	4691.33
January 2020	4847.52
February 2020	5726.18
March 2020	7142.33
April 2020	5317.17

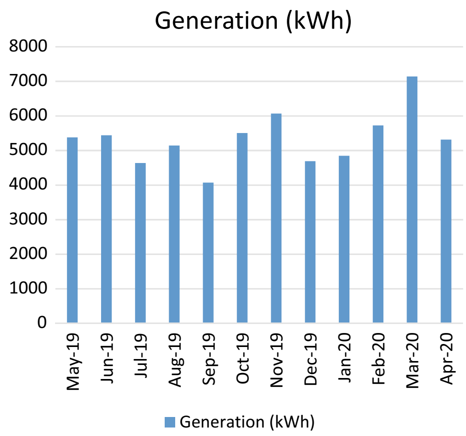


Fig. 5 Graphical representation of the generation data of the plant

9 Performance Analysis

The performance of the installed solar plant is analyzed in this section. Performance analysis is important to determine the durability and effectiveness of the system. There are several parameters include Final Yield (Y_F), Reference Yield (Y_R), Performance Ratio (PR), Capacitive Utilization Factor (CUF) and loss (L) [10, 14]. All these parameters are in Standard Test Condition (STC) and can be calculated in daily, monthly, yearly basis. In this paper yearly calculation values are shown [15].

Final Yield is the total generated power of the solar power plant in a particular span of time. We get Y_F from Table 2. Thus,

$$Y_F = E\_Grid = 63977\,\text{kWh} \tag{1}$$

Reference Yield is the actual value of generated power form a solar power plant. Here, the Y_R of one year of the solar power plant is-

$$\begin{aligned} Y_R &= (\text{Globlnc} * \text{PnomPV})\text{kWh} \\ &= (1461 * 50)\text{kWh} \\ &= 73050\,\text{kWh} \end{aligned} \tag{2}$$

Loss is referred as the power loss inside the power plant. It is the difference between the Reference Yield and the Practical Yield. Loss is measured in terms of kWh.

$$\begin{aligned} \text{Loss} &= [Y_R - Y_F]\text{kWh} \\ &= [(\text{Globlnc} * \text{PnomPV}) - \text{Y}_\text{P}]\text{kWh} \\ &= (1461 * 50) - 63977\,\text{kWh} \\ &= 9073\,\text{kWh} \end{aligned} \tag{3}$$

Performance Ratio (PR) is of a solar power plant is the ratio between the Final Yield and the Reference Yield. The PR of one year is shown below.

$$\begin{aligned} \text{PR} &= [E\_\text{Grid}/(\text{Globlnc} * \text{PnomPV})] * 100\% \\ &= [63977/(1461 * 50)] * 100\% \\ &= 87.57\% \end{aligned} \tag{4}$$

Capacitive Utilization Factor (CUF) represents the ratio of the actual output from a solar plant over the year to the maximum possible output from it for a year underr ideal conditions. CUF is also known as Plant Load Factor (PLF). Capacity utilization factor is usually expressed in percentage.

$$\begin{aligned} \text{CUF} &= [(E\_\text{Grid}/(365 * 12 * \text{PnomPV}))] * 100\% \\ (v) &= [(63977/365 * 12 * 50)] * 100\% \\ &= 14.72\,\% \end{aligned} \tag{5}$$

From Equation (1) to (5) where,

E_Grid = Total AC power generated from power plant in kWh.

Globlnc = Annual solar radiation in a particular area.

PnomPV = Total installed capacity of the power plant in KW.

Solar PV produce green and clean energy. Carbon-di-oxide emission is another vital part of solar PV power plant. It depends on energy generation of the power plant. Total CO_2 emission of the power plant in each month is shown in the Table 5.

On the other hand we can directly calculate the CO_2 emission of the power plant by using the below mentioned formula. Total CO_2 emission in the air due to this

Table 5 CO_2 emission in every month

Month	Energy generation (kWh)	CO_2 reduction in each month
May 2019	5379.21	4303.36
June 2019	5441.03	4352.82
July 2019	4640.31	3712.24
August 2019	5144.55	411,564
September 2019	4071.84	3257.47
October 2019	5504.23	4403.38
November 2019	6071.48	4857.18
December 2019	4691.33	3753.06
January 2020	4847.52	3878.02
February 2020	5726.18	4580.94
March 2020	7142.33	5713.86
April 2020	5317.17	4253.74

solar power plant in one year is-*b*.

$$
\begin{aligned}
&= (\text{Total generated power} * CO_2\text{emission per kWh})\text{kg} \\
&= (63977 * 0.8)\text{kg} = 51181.6\,\text{kg} \\
&= 51,\,181.6\,\text{kg}
\end{aligned}
$$

where 0.8 kg is the CO_2 emission of a rooftop power plant per unit energy generation [16]. Which is very less with compare to thermal or oil power plants. Thus, we conclude that solar energy is environment friendly.

10 Economic Benefit

Economic benefit is the main motivation behind installing this rooftop solar power plant. The solar power plant not only produce green, clean energy from sunlight but also reduces the overall electricity bill of the hospital building. The power plant is made in CAPEX mode [17]. Where the purchaser will provide the space and costing of the power plant such as materials cost, installation and freight of the materials. In return of benefit the purchaser will take over the power plant and uses the generated electricity as he wants to. To find out the economic benefit of the solar power plant we have gone through the previous monthly consumption bills of the hospital building. A three phase, one neutral net meter is provided by local electricity provider company to calculate the net power is taken from the local grid. As per WBSEDCL tariff system, one-unit (kWh) costs Rs. 8. Thus, we can say that monthly savings of the

Table 6 Energy consumption of the hospital building (without solar rtpv plant)

Month	Energy consumption (Unit)	Amount (Rupees)
May 2018	104,040.75	832,326.00
June 2018	108,624.22	868,993.76
July 2018	109,740.13	877,921.04
August 2018	103,680.30	829,442.40
September 2018	104,532.67	836,261.36
October 2018	107,280.10	858,240.80
November 2018	96,840.56	774,724.48
December 2018	97,092.41	776,739.28
January 2019	98,208.35	785,666.80
February 2019	101,136.54	809,092.32
March 2019	119,784.23	958,273.84
April 2019	112,344.56	898,756.48

entire hospital building is equivalent to monthly generation of the solar power plant * 8 rupees (as per WBSEDCL tariff system).

Monthly WBSEDCL consumption unit and amount form May 2018 to April 2019 (i.e., before installation of the solar plant) is given in the Table 6.

The graphical representation of total unit (kWh) consumption of the hospital building is shown below (Fig. 6).

The total amount of monthly generation bill (as per WBSEDCL tariff norms) from the solar power plant is shown in Tables 4 and 7.

From the above table we can observe that the plant generates Rs. 511,817 of energy in a year.

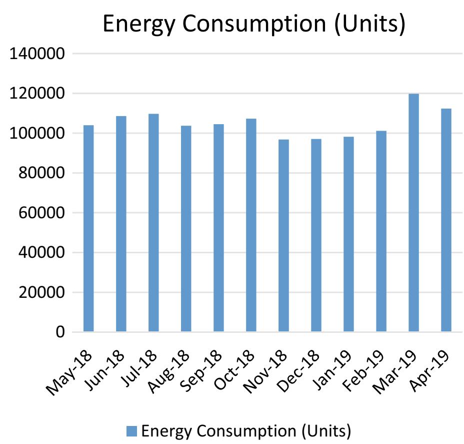


Fig. 6 Graphical representation of energy consumption of hospital building

Table 7 Total price of monthly energy generation of the plant

Month	Energy generation (kWh)	Amount (Rupees)
May 2019	5379.21	43,033.68
June 2019	5441.03	43,441.24
July 2019	4640.31	37,122.48
August 2019	5144.55	41,156.40
September 2019	4071.84	32,574.72
October 2019	5504.23	44,033.84
November 2019	6071.48	48,571.84
December 2019	4691.33	37,530.64
January 2020	4847.52	38,780.16
February 2020	5726.18	45,809.44
March 2020	7142.33	57,138.64
April 2020	5317.17	42,537.36

Payback period is another important factor of any solar power plant. It tells the total time taken to recover the economic investment of the power plant. Thus, the payback period of the power plant is-

$$\begin{aligned}
&\text{Payback period of the power plant}\\
&= \left[\text{Total plant cost}/(\text{Total power generation per year} * \text{per unit cost})\right]\text{Year}\\
&= \left[\text{Total plant cost}/(\text{Total units})\right]\text{Year}/\text{s}\\
&= 2743000/(511730)\,\text{Year}/\text{s}\\
&= 5.2\ \text{Year}/\text{s}
\end{aligned}$$

From the above calculation we can say that the entire payback will happened in the middle of 2024.

11 Conclusion

The rooftop solar power plant on the Uluberia super specialty hospital not only provides a pollution free and emission less energy generation but also reduces the overall electricity consumption. The cost of this power plant is Rs. 27,43,000 and it is takes about 4 months, i.e., from January 2019 to April 2019 to install this project. The payback period is estimated about nearly 5 years [18, 19]. This power plant is designed in such a way that total of five inverts are used. Thus, one inverter generates 20% of total AC power generation of the power plant at a time. In case of any accidental damage of any inverter, the faulty inverter trips and isolated from the rest of the plant without shutting the entire plant. Here also different size of the cables are used to reduce the cable resistance. As in DC side the current is high so the I^2R

loss is higher so reducing the cable resistance deduces loss [20]. On the other hand, increase in cable cross section area reduces the resistance but the price increases. Use of different size of cables solved the problem. An AC meter is installed between the AC distribution box and the grid feeder point to measure the final generation of the solar power plant. Solar energy is environment friendly. Total CO_2 avoided in the atmosphere is calculated about 51,181 kg (0.05 Tons) per year.

References

1. Chakravarthy M, Ramana Murthy KV, Devi BN (2015) Design, erection, testing and commissioning of 200Kwp rooftop grid tied solar photovoltaic system at vasavi college of engineering. In: 2015 IEEE IAS joint industrial and commercial power systems/petroleum and chemical industry conference (ICPSPCIC)
2. Nimje AA, Sawarkar PR (2017) Installation and commissioning of a 100 kW rooftop solar PV power plant. Ind J Sci Technol 10(27). https://doi.org/10.17485/ijst/2017/v10i27/111510, July 2017
3. Kathar SS, Thosar AG, Patil GC (2017) Design of rooftop solar PV. Int J Electr Eng Technol (IJEET) 8(2), pp 81–92, Article ID: IJEET_08_02_010, March- April 2017
4. Kumar NM, Subathra MSP, Moses JE, Small scale rooftop solar PV systems for rural electrification in India. In: 4th international conference on electrical energy systems (ICEES)
5. Meena RS, Rathore JS, Johri S, Grid connected roof top solar power generation: a review. Int J Environ Dev Res 3(1), ISSN: 2321–9939
6. Patil A, Deokar SA, Banderkar A (2018) GRID TIE solar power plant data acquisition system using internet of things. In: 2018 international conference on information, communication, engineering and technology (ICICET)Zeal College of Engineering and Research, Narhe, Pune, India. Aug 29–31 2018
7. https://globalsolaratlas.info/map?
8. Chakravarthy M, Ramana Murthy KV, Neelima Devi B (2015) Design, erection, testing and commissioning of 200Kwp rooftop grid tied solar photovoltaic system at Vasavi College of engineering. In: 2015 IEEE IAS Joint industrial and commercial power systems/Petroleum and chemical industry conference (ICPSPCIC)
9. Teja SC, Yemula PK (2016) Energy management of grid connected rooftop solar system with battery storage. In: 2016 IEEE innovative smart grid technologies—Asia (ISGT-Asia) Melbourne, Australia, Nov 28-Dec 1 2016
10. Jamil I, Zhao J, Zhang L, Jamil R, Rafique SF (2017) Evaluation of energy production and energy yield assessment based on feasibility, design, and execution of 3×50MW grid-connected solar PV pilot project in Nooriabad. Hindawi Int J Photoenergy 2017:18, Article ID 6429581, https://doi.org/10.1155/2017/6429581
11. Yadav SK, Bajpai U, Performance evaluation of a rooftop solar photovoltaic power plant in Northern India. Center of Excellence in Renewable Energy Education and Research, University of Lucknow (New Campus), Lucknow 226021, India, https://doi.org/10.1016/j.esd.2018.01.0060973-0826/© 2018 International Energy Initiative. Published by Elsevier Inc
12. Roy PK, Ghosh A, Mondal A, Bhattacharya S, Saha PK (2020) Design of a high gain DC-DC converter for renewable energy applications. In: 2020 international conference on emerging frontiers in electrical and electronic technologies (ICEFEET), 2020
13. Dondariya C, Porwal D, Awasthi A, Shukla AK, Sudhakar K, Murali Manohar SR, Bhimte A, Performance simulation of grid-connected rooftop solar PV system for small households: A case study of Ujjain, India. https://doi.org/10.1016/j.egyr.2018.08.002 2352–4847/© 2018 Published by Elsevier Ltd. This is an open access article under the CC BY NC-ND license

14. Meena RS, Rathore JS, Johri S (2014) Grid connected roof top solar power generation: a review. 2014 IJEDR | 3(1), ISSN: 2321–9939
15. Giraud F, Salameh ZM (2001) Steady-state performance of a grid-connected rooftop hybrid wind–photovoltaic power system with battery storage. IEEE Trans Energy Convers 16(1) (2001)
16. Zidane TEK, Adzman MR, Tajuddin MFN, Zali SM, Durusu A, Mekhilef S, Identifiability evaluation of crucial parameters for grid connected photovoltaic power plants design optimization. IEEE Access, https://doi.org/10.1109/ACCESS.2021.3102159
17. Ganga Prasanna M., Mahammed Sameer S., Hemavathi G., Financial analysis of solar photovoltaic power plant in India. IOSR J Econ Finance (IOSR-JEF) e-ISSN: 2321–5933, p-ISSN: 2321–5925 PP 09–15 www.iosrjournals.org
18. Okello D, van Dyk EE, Vorster FJ (2015) Analysis of measured and simulated performance data of a 3.2kWp grid-connected PV system in Port Elizabeth, South Africa. Energy Convers Manage 2015
19. Yoomak S, Patcharoen T, Ngaopitakkul A (2019) Performance and economic evaluation of solar rooftop systems in different regions of Thailand. Sustainability 11:6647. https://doi.org/10.3390/su11236647
20. Young M (1989) The technical writer's handbook. University Science, Mill Valley, CA

Image Authentication System Using Image Hashing with Color Filter Array

Ram Kumar Karsh

Abstract The availability of sophisticated multimedia editing tools makes image authentication a challenge in the modern age of multimedia communication. We suggested a hashing approach for image authentication in this paper. The proposed system extracts invariant vector distance features form a secondary image which is constructed through ring partitioning of the quaternion image. In parallel, the color filter array (CFA) interpolation coefficients are calculated from the CFA image, which has been used for identification of tampered regions. The concatenation of invariant vector distance and CFA coefficients construct the final hash. The results of experiments indicate that the approach proposed is robust in maintaining content against brightness/contrast adjustment, addition of noise, and arbitrary rotation etc., and identified the tampered regions. Compared to the existing image authentication methods, our approach provides better performance for forgery detection.

Keywords Image authentication · CFA interpolation coefficients · Ring partition · Invariant vector distance · Central orientation information · Robust image hashing

1 Introduction

The digital age has made the use of the Internet and multimedia an aspect of everyday life for image and video processing. The development of image editing tools also poses a number of challenges in academia and industry. The originality and integrity of the digital image must therefore be checked using sophisticated tools and techniques [1]. The techniques of image authentication can be divided into three groups via: (a) image watermark [2], (b) digital image forensic [3], and (c) perceptual image hash [4, 5]. In the watermark-based approaches, some extra information (watermark) are embedded into the digital image. It is assumed that the forgery on an image may alter the watermark, and a picture can be authenticated with a watermark match, in the receiver end. Here, the embedding process may degrade the image quality,

R. K. Karsh (✉)
Department of ECE, NIT Silchar, Silchar, India
e-mail: ram@ece.nits.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_22

which is unacceptable in many real-time applications. In the second scheme, the image features similarities/inconsistencies in the intrinsic features are used without the knowledge of watermark or hash from the transmitter side [6]. Being a blind approach, the non-existence of information from a transmitter to receiver side creates huge computation burden. The approaches based on perceptual image hashing are an enhancement of traditional cryptosystem [7]. In the traditional cryptosystem, the function of hash converts a long-length message to a short string [8]. The essential requirement for the cryptographic authentication is an avalanche effect. The excessive sensitivity of the cryptosystem hash function, however, reduces its application in the multimedia area. The following criteria should be applied in multimedia applications: first, robustness, in which an image hash of the original file, and after contents preservation operations should be approximately same [9–14]. The second one is discrimination, which means for visually different images, hash should be different. Security means an evildoer should not be able to trick the authentication system for malicious tampering. In order for this, the hash should differ for wrong key [9].

Schneider and Chang [15] implemented image hash for the first time. Thereafter, researchers paid large attention toward it, which have been divided into five major categories based on feature extraction approaches. Image hashing technique based on transform domain [10–12, 16–19]. The hashes are constructed from wavelet coefficients [10], coefficients of discrete Fourier transform (DFT) [11], Fourier–Mellin transform (FMT) [12], Radon transform (RT) [16], log-polar transform (LPT) [17], and low-frequency coefficients of discrete cosine transform (DCT) [18, 19]. These methods are invariant to some geometric distortions, but poor in discrimination. The hashing approaches via decomposition of matrix are [13, 20, 21]. The hashes are calculated based on non-negative matrix factorization (NMF) [13, 20] and tensor decomposition [21]. These approaches are geometrically sensitive, including arbitrary degree of rotation. Image hashing technique based on characteristics of local information [22–25]. Hashes are extracted from salient features based on the invariant transformation, i.e., SIFT and SURF [22, 23], adaptive local features [24], and local binary patterns (LBP) [25]. These methods are sensitive to noise addition and an arbitrary rotation. Moreover, the color forgery is not detected.

Image hashing technique based on moments [14, 26]. The hashes are constructed from Zernike moments [14, 26]. These methods are sensitive to large degree of rotation. Moreover, the discrimination needs to improve. Image hashing technique based on histograms [27, 28]. Hashes are generated from the relation of pixels in the different bins of histogram [27]. The content changes with similar histograms for various image information cannot be defined by these methods. In other works, Tang et al. [28] innovatively constructed image hash based on the histogram. These approaches are effective, in the case of rotation, but discrimination must be improved.

Some other approaches in image hashing have discussed in the past works. Karsh et al. [29] suggested ring-based hashing methods. The hashes are extracted from feature obtained from the secondary image based on ring partition. These methods are robust to rotation, but unexplored details during hash generation about 20%. Moreover, the discriminative capability needs to be improved further.

In this paper, we focused on robustness and small visual discrimination. We introduced content authentication via hashing approach using CFA interpolation coefficients along with existing invariant vector distance [30] and additional knowledge of central orientation information. The proposed method may detect and locate the forgery, even if the forged image has been rotated. Moreover, it can detect the small forgery. There are no publicly available database for global and local color change and also splicing and copy–move in the four outer boundary of an image. So, we have developed our own database for this study and also makes it publicly available [31]. The experimental results demonstrate robustness, discrimination, and identification of the counterfeit area improved with the proposed algorithm.

The remaining paper has the following structure. Section 2 describes the proposed algorithm for hashing. In Sect. 3, the hashing has been represented for authentication. Section 4 lists the experimental outcomes for the proposed model. The conclusion and possible scope of the study were discussed in Sect. 5.

2 Proposed Image Hashing Algorithm

Our hashing method is depicted in Fig. 1. In the subsequent sub-sections, the specifics of each step were analyzed.

A. *Preprocessing*

First, the source image of any size is framed to a fixed dimension of $M \times M$ using bilinear interpolation, which provide always fixed length of hash yields robust to scaling. Next, Gaussian filter is applied, which has been used to mitigate the effect of noises.

B. *Generation of first intermediate hash* (***h***)

There are four steps to this module: quaternion, a secondary image formation, extraction of statistics feature, and distance feature mapping to create a first intermediate hash.

In order to extract color information, the preprocessed image has been mapped in a quaternion domain. The color image consists of three parts at every pixel (x, y) that can be converted into a single component by the virtue of pure quaternion [32] as follows:

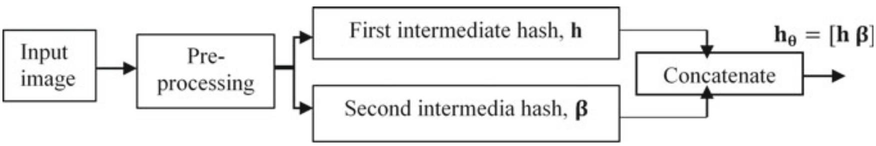


Fig. 1 Proposed image hashing

$$\mathbf{F}(x, y) = \mathbf{F}_R(x, y)\mathbf{i} + \mathbf{F}_G(x, y)\mathbf{j} + \mathbf{F}_B(x, y)\mathbf{k} \tag{1}$$

where $\mathbf{F}_R(x, y)$, $\mathbf{F}_G(x, y)$, and $\mathbf{F}_B(x, y)$ represents the RGB components of a color image.

Generally, the rotation operation keeps the center unaltered. Hence, the information of rotated image remains same in annular rings. The formation of a secondary image from a square image is done using ring partition. The ring-based method [29, 30] considers the image information around 80%. The corner information remains unused for a hash generation. There are significant amount of information in the corners. The ring-based method by including the corner information [33, 34], so as to identify malicious tampering of small size, in the corners of an image, has been carried out for first intermedia hash.

The pixels on the ring picture are arranged in the matrix columns. Four statistical elements, mean (μ), variance (σ^2) skewness (s), and kurtosis (k), are efficiently represented for the visual contents of each concentrated ring. This is because of these four parameters: the average image energy can be represented by mean, the variation of the pixel picture can be determined by variance, pixel asymmetry can be characterized by skewedness, and, finally, the distribution of pixel shape can be represented by kurtosis.

Now, the characteristics of statistics are used to generate a vector of features$\mathbf{v}_j$, reflecting the j-th column of a secondary image.

$$\mathbf{v}_j = \left[\mu_j, \sigma_j^2, s_j, k_j\right]^T \tag{2}$$

where $[.]^T$ represent the transpose.

Next, $\mathbf{V}$, i.e., $4 \times (n + 4)$ size generated from nth ring and four corners as

$$\mathbf{V} = \left[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+4}\right] \tag{3}$$

In a four-dimensional space, the vector that represents each ring and corner can be regarded as a point. In the feature space, the content preservation processes cause these points to be slightly disturbed. Thus, the shift in vector length is much less influenced by the digital operations. This vector distance property was used to describe the hash elements.

The vector obtained is initially normalized. If $\mathbf{f}_k = [f_k(1), f_k(2), \cdots, f_k(n), f_k(n+1), \cdots, f_k(n+4)]$ known for k-th row in $\mathbf{V}$ $(1 \leq k \leq 4)$. Next, $\mathbf{g}_k$ and $g_{\text{ref}}(j)(1 \leq j \leq 4)$ that are normalized vector and average values have been determined as

$$g_k(j) = \frac{f_k(j) - \overline{f_k(j)}}{\sigma_k^2};\ g_{\text{ref}}(j) = \frac{1}{(n+4)} \sum_{k=1}^{(\text{n+4})} g_k(j) \tag{4}$$

where σ_k^2 and $\overline{f_k(j)}$ represent variance and mean $\mathbf{f}_k$.

The vector distance $e(k)$ between $\mathbf{g}_k$ $(1 \leq k \leq (n+4))$ and $\mathbf{g}_{\text{ref}}$ is defined as

$$e(k) = \sqrt{\sum_{j=1}^{4} (g_k(j) - g_{\text{ref}}(j))^2} \tag{5}$$

The distance vector $e(k)$ is scrambled using secret key and generates first intermediate hash,

$$\mathbf{h} = [h(1), h(2), \ldots, h(n), h(n+1), \ldots, h(n+4)] \tag{6}$$

III. *Bayer array conversion and estimation of correlation coefficients of a CFA interpolated color image*

The preprocessed image is mapped to a Bayer array $[RGGB]$ [35] (i.e., single color at each pixel location) to get a CFA image. All other colors are interpolated using bilinear interpolation, which generates CFA interpolated color image $\mathbf{I}(x, y)$, where $1 \leq x \leq M$, and $1 \leq y \leq M$. In a CFA interpolated color image acquisition, a camera sensor captures a single color at every pixel; the remaining two colors are interpolated from the neighboring samples. As a consequence, within a color channel, a subset of pixel samples is linearly correlated to their neighboring samples (CFA detail reference [35]). However, even a small tampering may destroy this correlation. Hence, presence or absence of correlation may be used to authenticate an image. However, only this knowledge is not sufficient for an image hashing. Hence, it has been used as a complementary feature in the proposed model.

To estimate the correlation coefficients of a CFA interpolated color image, the expectation maximization (EM) algorithm [36] discussed in Appendix 1 has been applied on the red and green channels to get $\boldsymbol{\beta}_r$ and $\boldsymbol{\beta}_g$, respectively (where $\boldsymbol{\beta}_r$ and $\boldsymbol{\beta}_g$ are red, green channel correlation coefficients). These are concatenated to get final correlation coefficients $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_\mathbf{r}\boldsymbol{\beta}_\mathbf{g}\right]$, which is a second intermediate hash of length 16 digits.

3 Image Authentication Using Proposed Image Hashing

The proposed hashing method is applied in an image authentication application as shown in Fig. 2.

The receiver has the received image along with the transmitted hash $\mathbf{h}_{\boldsymbol{\theta}}$. The receiver model corrects the geometric correction (if any) using blind geometric correction approach as follows.

Let an image obtained be processed using combined geometric correction like RST. Next, rotation angle is determined as

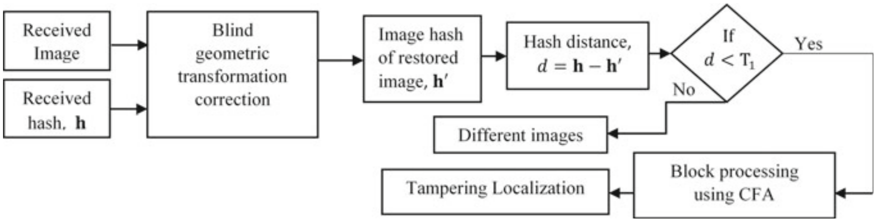


Fig. 2 Proposed image authentication

$$\theta = \arctan(\Delta Y/\Delta X); \theta' = \arctan\left(\Delta Y'/\Delta X'\right) \tag{7}$$

where, $\Delta Y = Y_b - Y_r$, $\Delta X = X_r - X_b$, $\Delta Y' = Y_l - Y_t$, and $\Delta X' = X_t - X_l$. Here, non-zero-pixel values indexes for top, bottom, left, and right are (X_t, Y_t), (X_b, Y_b), (X_l, Y_l), and (X_r, Y_r). If $\theta \cong \theta'$ the image has rotated with angle θ. In next operation the image needs to be anti-rotated with angle θ and region of interest is cropped yields a restored image.

Now, the intermediate hash ($\mathbf{h}'$) of restored image is generated. Next, the difference, d, between $\mathbf{h}$ and $\mathbf{h}'$ is found. If d exceeds the selected threshold (T_1), the image pair will be different (fake) or large tampered, otherwise perceptually same (authentic) or small tampered. Using the knowledge of $\boldsymbol{\beta}$, the perceptually similar pairs and small tampered pairs may be segregated using block processing as discussed below. Additionally, in the case of tampering, the tampered region will also be identified.

A. *Block processing for segregation and taper detection*

A block processing approach is used to differentiate between perceptually similar image and small content addition/deletion image. It is also applied to locate the tampered region. The perceptually same (authentic) or small tampered image of size $M \times M$ is received. Then, it is divided into $9M^2/4b^2$ number of blocks (block size = b) with 50% overlap between subsequent blocks. For each block, the EM algorithm is applied for the red and green channels to get interpolation coefficients, $\boldsymbol{\beta}_r'$ and $\boldsymbol{\beta}_g'$, respectively. These are concatenated to get $\boldsymbol{\beta}' = \left[\boldsymbol{\beta}_\mathbf{r}'\boldsymbol{\beta}_\mathbf{g}'\right]$.

A block is labeled as a tampered, if the correlation measure (C) between the received ($\boldsymbol{\beta}$) and calculated ($\boldsymbol{\beta}'$) of specified block is less than the threshold T_2; otherwise, it is considered authentic. Based on experiment, over 3922 perceptual similar image pairs and 500 forged image pairs empirically selected threshold $T_2 = 0.95$. If the blocks of an image is found to be tampered, then drawn a red rectangle in the received image in the corresponding block positions.

4 Results and Discussions

The proposed hashing technique has been validated with 37,855 image pairs (where 5830 perceptual similar, and 32,025 are forged or different pairs) generated from USC-SIPI database [37], CASIA Tampered database [38], from Nikon D3200 taken images, and extracted from Internet. Based on the rigorous experiments on large databases, the optimum parameter values used in the method proposed are as follows: $M \times M = 512 \times 512$, and $n = 32$, $N = 1$, $\sigma_0 = 0.0075$, $P_0 = \frac{1}{256}$, $m = 16$, $b = 32 \times 32$, $\delta = 10^{-5}$, $T_1 = 148$, and $T_2 = 0.95$.

A. *Analysis of robustness and discrimination*

For the analysis of trade-off between robustness as well as discrimination, a threshold (T_1) has been identified shown in Fig. 3. The 5830 content similar image pairs have been generated using 110 operations as shown in Table 1. 19900 different image pairs have been formed from 200 images. The hash distance probability plot for 25, 730 image pairs is shown in Fig. 3. Empirically from Fig. 3b, the optimal threshold hash distance (T_1) is selected 148 (which is the crossover point of two plots) to consider robustness as well as discrimination. It can be observed that most of the hash distances are situated left side of decided threshold, reverse for different image pairs. Hence, the proposed content authentication is better to maintain the balance for robustness and discrimination.

B. *Tampered area localization*

A set of 1400 images and their tampered versions (large area as well as small) have been taken from CASIA and TID database and created by own publicly available [31] for forgery detection capability of the proposed hashing technique. It is experimentally observed that for 700 tampered image pairs; the hash distances are larger than selected threshold (T_1). The remaining 700 small tampered image pairs hash

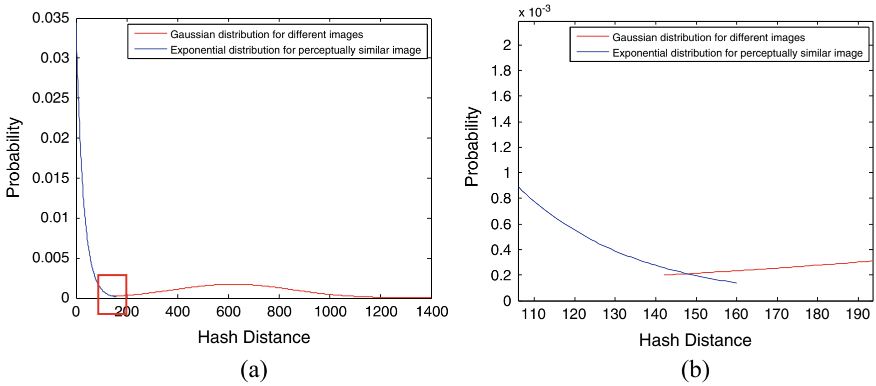


Fig. 3 Probability plot for hash distances (**a**) for content similar and different image pairs (**b**) Better view of cross over for red rectangle, shown in a

Table 1 Similar contents generation specifications

Tools	Operations	Parameters	Values of parameters
Photoshop	Brightness adjustment	Photoshop's scale	± 10, ± 20
Photoshop	Contrast adjustment	Photoshop's scale	± 10, ± 20
StirMark	JPEG contraction	Quality factor	30, 40, 50, 60, 70, 80, 90, 100
StirMark	Scaling	Ratio	0.5, 0.75, 0.9, 1.1, 1.5, 2.0
StirMark	Rotation	Rotation angle in degree	± 5, ± 15, ± 30, ± 45, ± 90
MATLAB	Gamma rectification	Γ	0.75, 0.9, 1.1, 1.5
MATLAB	3×3 Gaussian low-pass filtering	Standard deviation	0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
MATLAB	Salt and pepper noise	Density	0.001 to 0.01 (step size 0.001)
MATLAB	Speckle noise	Variance	0.001 0.01 (step size 0.001)
MATLAB	Translation	[x-pixels, y-pixels]	10 to 100 (Both axis)
MATLAB	Composite RST	Degree of rotation, Ratio, and [x-pixels, y-pixels]	# 1–(5°, 0.5, [10]), # 2–(10°, 0.75, [20]), # 3–(15°, 0.9, [30]), # 4–(30°, 1.1, [40, 40]), # 5–(45°, 1.5, [50, 50]), # 6–(87°, 2.0, [60, 60]. Made 36 combinations from above six parameters

distance are below the threshold (T_1), which has been correctly detected and tampered region(s) is/are highlighted by the red rectangle using block processing approach. For example, shown in Table 2.

III. *Performance comparisons with state-of-the-art methods*

To illustrate the effectiveness, the proposed image hashing technique is compared with some existing hashing techniques: such as RT-DFT [16], Zernike moments (ZM)-based hashing [14], ring-NMF-based hashing [29], and ring-invariant vector distance (IVD)-based hashing [30].

The comprehensive performance comparison is depicted in Fig. 4, which shows better TPR than other methods, during FPR 0 to 0.4. This indicates the robustness efficacy of the proposed model. Nevertheless, at the same time, low FPR shows the proposed method discriminative capability. However, the performance improvement is almost similar to that of the other methods above the FPR of value 0.4.

The performance comparison in the case of forgery detection and some other parameters are represented in Table 3. It can be find that proposed method can detect copy–move/splicing forgery in the corners of an image. The existing methods [14, 16] can also detect this forgery, but these methods do not possess the rotation invariant property for an arbitrary angle. However, it is experimentally find out that

Table 2 Forgery detection with content insertion and deletion of small size

Original images					
Forged images and Localization					
d	32.8953	18.6279	22.6936	9.7468	19.1311

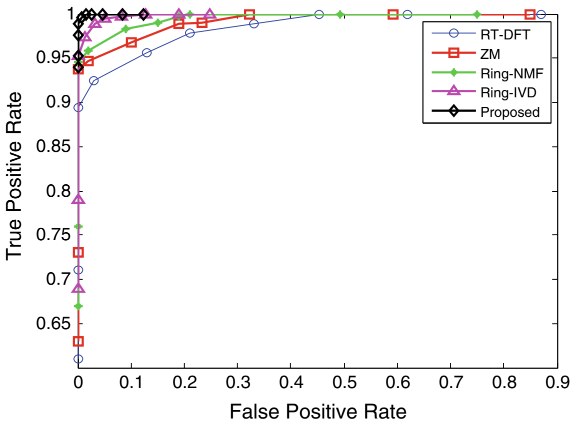


Fig. 4 Performance comparisons using ROC

Table 3 Performance comparison for discrimination with different existing techniques

Methods	RT-DFT based [16]	ZM based [14]	Ring-NMF [29]	Ring-IVD [30]	Proposed method
1. Detection of forgery such as splicing, copy-move on corners of an image	Yes	Yes	No	No	Yes
2. Color forgery detection	No	Yes	No	No	Yes
3. Robustness toward an arbitrary angle of rotation	No	No	Yes[1]	Yes[1]	Yes[2]
4. Hash length	15 digits	560 bits	74 digit	40 digit	52 digit
5. TPR when FPR = 0	0.8943	0.9378	0.9454	0.9523	0.9925
6. FPR when TPR = 1	0.1534	0.0732	0.0565	0.045	0.0011
7. Average time(sec)	3.4323	0.6235	2.1348	0.2943	3.2342
8. Identification of tampered areas gone through rotation	No	No	No	No	Yes

Yes[1]: Yes (loss of some information); Yes[2]: Yes (no loss of information) our approach is robust against an arbitrary rotation. Due to the inclusion of quaternion model, the proposed model can detect color forgery, but [16, 29, 30] are insensitive to color changes. The proposed image authentication model is also capable of detecting the tampered location, which is a limitation of [16, 29, 30]. The hash length of proposed method is a little bit large, but it is still below a 1 kb size which is usually considered to be a maximum size of a hash [39]. It is shown that TPR when FPR ≈ 0 of proposed method is 0.9925, which is comparatively better than the other methods. Similarly, when TPR ≈ 1, the FPR is observed to be lowest among all other method considered in this study. It is expected the method that provides high TPR and low FPR may be considered as better hashing technique. From this perspective, it may be inferred that the proposed method may be selected as an improved hashing approach than existing techniques.

The proposed hash generation approach has been implemented in desktop system with configuration 8 GB RAM, MATLAB 2015a, and Windows 8. The other methods that have been used for comparative analysis have also been implemented on the same platform to evaluate the computational demands. Table 3 represents that our algorithm is computationally demanding than some existing approaches, but fulfill the simultaneous requirement of good discrimination, robustness, and high sensitivity to content/color changes.

5 Conclusions and Future Scopes

Here, we design a hashing algorithm using CFA interpolation coefficients, vector distance, and a blind geometric correction, which has been applied for image authentication. The arbitrary rotation invariant and forgery detection along with arbitrary rotation are obtained in our approach due to a blind geometric correction. The CFA interpolation coefficients have enhanced the discrimination of the proposed method. The results of the experiments show that the proposed method is robust to digital operations, including arbitrary rotations. The proposed method forgery detection (like a small copy–move or splicing in the corners of an image, the global and local color changes, the small malicious tampering) performance is better to compared methods. The proposed model may also detect the malicious tampering regions. The ROC comparison shows that the proposed hashing algorithm has good discrimination and robustness, when applied to image authentication application.

Future work may be to explore machine learning for content authentication via image hashing.

References

1. Karsh RK, Saikia A, Laskar RH (2018) Image authentication based on robust image hashing with geometric correction. Multimedia Tools Appl 77(19):25409–25429
2. Cox I, Miller M, Bloom J, Fridrich J, Kalker T (2007) Digital watermarking and steganography. Morgan Kaufmann, Nov 2007
3. Birajdar GK, Mankar VH (2013) Digital image forgery detection using passive techniques: a survey. Digit Investig 10(3):226–245
4. Karsh RK, Laskar RH, Richhariya BB (2016) Robust and secure hashing using Gabor filter and Markov absorption probability. In: Proceeding of the IEEE conference on communication and signal processing (ICCSP), pp 1197–1202, Apr 2016
5. Saikia A, Karsh RK, Laskar RH (2017) Image authentication under geometric attacks via concentric square partition based image hashing. In: Proc.TENCON 2017–2017 IEEE Region 10 Conference, pp 2214–2219, Nov 2017
6. Farid H (2009) A survey of image forgery detection. IEEE Signal Process Mag 2:16–25
7. Monga V (2005) Perceptually based methods for robust image hashing. Ph.D. diss., 2005, pp 120
8. Stamp M (2011) Information security: principles and practice. vol 2., Wiley, Nov 2011
9. Karsh RK, Laskar RH, Aditi (2017) Robust image hashing through DWT-SVD and spectral residual method. EURASIP J Image Video Process 2017(1):31, Dec 2017
10. Ahmed F, Siyal MY, Abbas VU (2010) A secure and robust hash-based scheme for image authentication. Signal Process 90(5):1456–1470
11. Swaminathan A, Mao Y, Wu M (2006) Robust and secure image hashing. IEEE Trans Info Forensics Secur 1(2):215–230
12. Sun R, Zeng W (2014) Secure and robust image hashing via compressive sensing. Multimedia Tools Appl 70(3):1651–1665
13. Monga V, Mihcak MK (2007) Robust and secure image hashing via non-negative matrix factorizations. IEEE Trans Inf Forensics Secur 2(3):376–390
14. Sucharitha G, Senapati RK (2020) Biomedical image retrieval by using local directional edge binary patterns and Zernike moments. Multimedia Tools Appl 79(3):1847–1864
15. Schneider M, Chang SF (1996) A robust content based digital signature for image authentication. In: Proceeding of the IEEE international conference on image processing, pp 227–230, Sept 1996
16. Lei Y, Wang Y, Huang J (2011) Robust image hash in radon transform domain for authentication. Signal Process Image Comm 26(6):280–288
17. Ouyang J, Coatrieux G, Shu H (2015) "Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform. Digital Signal Process 41:98–109
18. Ching YL, Shih FC (2001) A robust image authentication method distinguishing JPEG compression from malicious manipulation. IEEE Trans Circ Syst Video Technol 11:153–168
19. Tang Z, Lao H, Zhang X, Liu K (2016) Robust image hashing via DCT and LLE. Comput Secur 62:133–148
20. Karsh RK, Laskar RH (2015) Perceptual robust and secure image hashing using ring partition-PGNMF. In: Proceeding of the TENCON 2015–2015 IEEE region 10 conference, pp 1–6, Nov 2015
21. Paul M, Karsh RK, Talukdar FA (2020) Image authentication using tensor decomposition and local features with geometric correction. In: Proceeding of the international conference on machine learning, image processing, network security and data sciences, pp 397–411 Jul 2020
22. Lv X, Wang ZJ (2012) Perceptual image hashing based on shape contexts and local feature points. IEEE Trans Inf Forensics Secur 7(3):1081–1093
23. Paul M, Karsh RK, Talukdar FA (2019) Image hashing based on shape context and speeded up robust features (SURF). In: Proceeding of the IEEE international conference on automation, computational and technology management (ICACTM), pp 464–468, Apr 2019
24. Patil V, Sarode T (2019) Adaptive CSLBP compressed image hashing. Int J Electric Comput Eng 9(4):2982

25. Qin C, Hu Y, Yao H, Duan X, Gao L (2019) Perceptual image hashing based on weber local binary pattern and color angle representation. IEEE Access 28(7):45460–45471
26. Zhao Y, Wang S, Zhang X, Yao H (2012) Robust hashing for image authentication using zernike moments and local features. IEEE Trans Inf Forensics Secur 8(1):55–63
27. Choi YS, Park JH (2012) Image hash generation method using hierarchical histogram. Multimedia Tools Appl 61(1):181–194
28. Tang Z, Huang L, Dai Y, Yang F (2012) Robust image hashing based on multiple histograms. Int J Digit Content Technol Appl 6(23):39–47
29. Karsh RK, Laskar RH, Richhariya BB (2016) Robust image hashing using ring partition-PGNMF and local features. Springerplus 5(1):1995
30. Tang Z, Zhang X, Li X, Zhang S (2015) Robust image hashing with ring partition and invariant vector distance. IEEE Trans Inf Forensics Secur 11(1):200–214
31. Color adjustment and copy-move/splicing in corner of an image database https://rishabhphukan.wixsite.com/rkkarsh
32. Kantor ILV, Solodovnikov AS, Shenitzer A (1989) Hyper complex numbers: An elementary introduction to algebras. Springer-Verlag, New York
33. Karsh RK, Mishra H, Vashist T, Laskar RH (2018) Image authentication using image hashing based on ring partition with corner inclusion. In: Proceeding of the. IEEE conference on communication and signal processing (ICCSP), pp 0557–05623, Apr 2018
34. Karsh RK (2020) Geometric invariant image authentication system using hashing. In: Proceeding of the IEEE conference on communication and signal processing (ICCSP), pp 1143–1149, Jul 2020
35. Bayer BE (1976) Color Imaging Array. US Patent, 3 971 065, Jul 1976
36. Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc Ser B (Methodol) 99(1):1–38
37. USC-SIPI image database (2007) http://sipi.usc.edu/database/. Accessed Feb 2007
38. CASIA tampered image detection evaluation database [Online] Available: http://forensics.idealtest.org/
39. Decker BD, Lapon J, Naessens V, Uhl A (2012) Communication and multimedia security. Springer ebook, Ghent, Belgium, Sep 2012

Smart Segregation and Defect Detection of Nuts for Fasteners Industries

Akhilesh Kumar, Gurleen Kaur, Bhavyae Mittal, Gaurav Sapra, Rajesh Kumar, and Garima Joshi

Abstract This paper presents the design and fabrication of economically viable nut segregation and defect detection machine for fasteners industries. This prototype is developed to help manufacturers in achieving quality assurance at minimal cost. Simulations are performed in ANSYS to optimize the design parameters of machine prototype. Working of machine prototype is categorized in three stages: (i) linearization, (ii) slot allocation, and (iii) image processing. To segregate, the individual nut from the feeder motoring mechanism is adopted for the linearization stage. For defect analysis, nuts are moved with the help of a motor on the vacant slots of rotating disk mounted with camera at fixed location. Images of nuts are captured by camera and processed for defect detection by adopting edge detection and background subtraction algorithms in image processing. The detection of nut dimensions using machine prototype is also reported.

Keywords Linearization · Finite element analysis · Slot allocation · Image processing · Segregation of nuts

1 Introduction

Automobile industry of India is among the fastest growing markets in the world. With advancement in technology and increase in consumer preferences, the automobile industry is changing at a very fast pace. To meet the ongoing demand of the automobile industry, a lot of technological advancement is also required at fastener industries [1]. One of the biggest concerns with fastener industries in India is to

A. Kumar (✉) · B. Mittal · G. Sapra
Electrical and Electronics UIET, Panjab University, Chandigarh, India
e-mail: akhilesh21101997@gmail.com

G. Kaur · G. Joshi
Electronics and Communication UIET, Panjab University, Chandigarh, India

R. Kumar
Mechanical Engineering, UIET Panjab University, Chandigarh, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_23

pace up the production without compromising on the quality. Quality assurance at minimal cost with strict deadlines is the biggest concern for small and medium-scale enterprises (SMSEs). For fastener industry, nut segregation and detection plays a vital role in quality assurance. Various imported advanced nut segregation and detection machinery based on artificial intelligence, machine vision, image processing, neural networks, machine learning, RFID sensing, and software technology are available globally [2–6]. These machines do not provide all the required stages of sorting like linearization, orientation of nut, fault on faces, and fault on edges. The cost of imported machinery is beyond the reach for most SMSEs in our country. To address the concern of industry design and fabrication of cost-effective and reliable, indigenous machine prototypes is proposed in this work.

The working model of the prototype is categorized in three stages: (i) linearization, (ii) slot allocation, and (iii) image processing. In this paper, design and working of nut segregation and defect detection machines are discussed in detail. Also, the detection of nut dimension using the prototype is reported in this work.

2 Experimental Section

A. *Schematic of machine*

Three stages of the proposed machine are divided into four sections: (a) linearization of nuts, (b) rotating disk, (c) image processing, and (d) segregation as shown in Fig. 1.

First, linearization is done by pouring a bulk of nuts in the feeder. From the feeder, few nuts are made to fall on the conveyor. As the conveyor vibrates, nuts get linearized and start moving one by one in line. After linearization, one nut at a time takes the position in the slot of rotating disk as shown in Fig. 2. Further, camera is mounted above the rotating disk. On finding the nut in the field view of the camera, infrared (IR) sensor detects it and sends the feedback signal to camera for image capturing. The image captured is then further processed for fault detection using edge detection [7] and background subtraction algorithm [8] implemented in Python computer language. Final segregation of faulty nuts is done by energizing the pneumatic blower system [9].

Various materials such as wood, plastic, and galvanized aluminum sheets are used to develop the machine prototype. Overall base of the machine structure is made up of wood, and wooden planks are used to develop structures like hopper base, axle support structure, and vibrating mechanism support. For cost effectiveness, plastic bottles are currently being used as a feeder. Due to high tensile strength

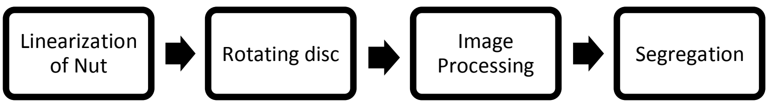


Fig. 1 Schematic of machine

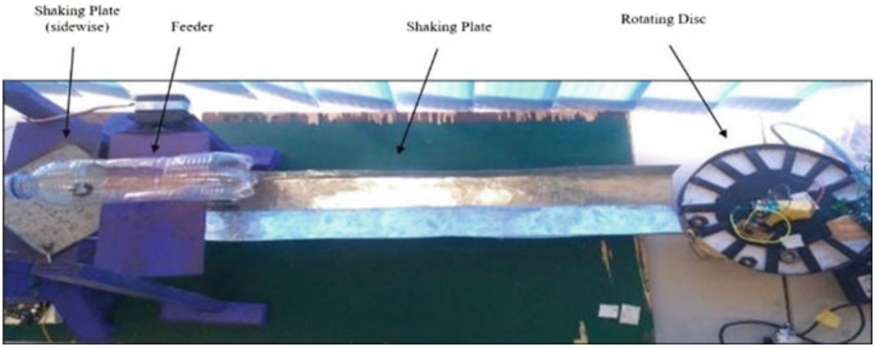


Fig. 2 Photograph of machine prototype. Materials used

and lightweight, galvanized aluminum sheets are selected to develop horizontal and vertical shaking plates. The overall dimension of machine prototype is given in Table1.

B. *Simulations*

To optimize the design parameters of machine prototype finite element, analysis is done using ANSYS software [10, 11]. For making the machine design more robust and reliable, material selection for machine parts is selected using simulations (Fig. 3).

Three different materials, i.e., structural steel, stainless steel, and aluminum alloy, are selected for analysis. The detailed material properties are given in Table 2. From ANSYS simulations, it is found that the stationary components of the model are to be fabricated with structural steel, and for the moving parts of the model, aluminum alloy is to be used for the fabrication.

Table 1 Machine prototype parts with dimensions

Name of parts	Length (cm)	Breadth (cm)	Height (cm)
Hoper base (To attach with stepper motor)	18	8	0.5
Hopper side channel (Stand on which hoper is connected)	3	1.4	38.3
Shaker side channel (Stand on which shaker is connected)	3	1	25
Shaker base	19	7.5	0.5
Swing plate	16.5	8	0.5
Rotating bar	4.5	1	9
Side channel for swinging plate	5	2	9

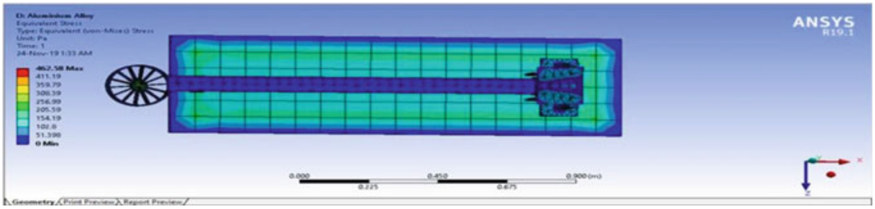


Fig. 3 Simulation model of machine prototype using ANSYS software

Table 2 Materials required

S.NO	Material	Stress (pa)	Strain 10^{-9}	Deformation (m*10^{-11})	Density (kg/m^3)
1	Structural steel	1313.7	6.67	7.80	7800
2	Stainless steel	1296.2	6.82	7.90	8000
3	Aluminum alloy	462.58	6.62	7.47	2770

3 Working Mechanism

A. *Linearization*

For the linearization stage, bulk of nuts is poured in a feeder for fault detection. For pouring a few nuts on a shaking plate from the feeder, it is bent downward to an angle of 30° with the help of a servo motor. Shaking plate is vibrated with the help of DC motor. Speed of the DC motor is controlled by an Arduino UNO microcontroller [12–14]. From horizontal shaking plate, one nut at a time falls on the vertical aluminum channel. A geared DC motor of specification 60 RPM at 12V is connected with a shaft of vertical aluminum channel. Aluminum channels are made to vibrate using a DC motor whose speed is controlled by using Arduino and L293D I.C [15]. Initially, the speed of vibration is kept low, i.e., 40 rpm, and it is increased gradually to get accurate results for linearization as shown in Fig. 4.

B. *Slot Allocation*

After the linearization stage, the nut occupies the vacant slot in the slotted disk one by one. To avoid the overlapping of nuts, the structure of a rotating disk is designed in such a way that it consists of two circular disks placed one over another. Base circular disk is kept fixed, and the slotted disk is connected with a DC motor for circular rotation. The rotation of disk and vibration of channel is synchronized to avoid overlapping of nuts as shown in Fig. 5.

C. *Image Processing*

To segregate the faulty nuts, image processing technology is adopted [16]. For image processing, a camera is mounted at a fixed location over the rotating disk. As soon

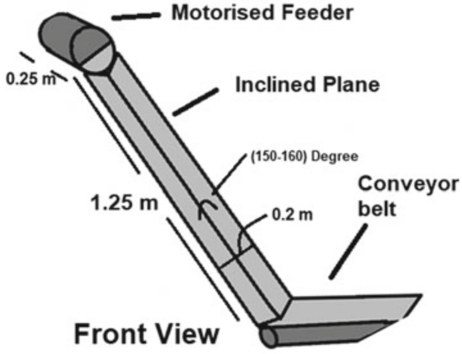


Fig. 4 Linearization stage of machine prototype

Fig. 5 Slot allocation stage of machine prototype

as a nut comes in the vicinity of the camera, the IR sensor detects it and sends the feedback signal back to the camera [17]. On receiving the feedback signal, the camera captures the image of the nut. The captured image is then compared with the image of the correct nut using edge detection and background subtraction algorithm implemented in Python computer language as shown in Fig. 6.

The dimension of the nut is detected by machine prototype using image processing technology. A sample nut of dimension is found to be 0.9 inches with metric scale is detected with the same dimension using machine prototype as shown in Fig. 7.

Fig. 6 Image processing stage of machine prototype

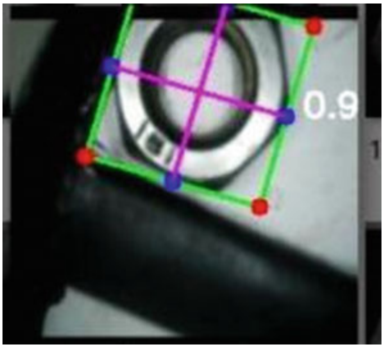


Fig. 7 Detection of nut dimension using machine prototype

Machine prototype is capable of detecting ± 0.1 inches diametrical variations in nut dimension.

4 Conclusion

In this paper, the design and fabrication of nut segregation and defect detection machine are discussed in detail. For optimizing the design parameters of machine, finite element simulations are performed using ANSYS software. Simulation results suggest that the stationary components of the prototype are to be fabricated with structural steel and moving parts with aluminum alloy. The working mechanism is divided into three stages: (i) linearization, (ii) slot allocation, and (iii) image processing. Linearization stage involves the use of two motors, i.e., servo motor and DC-geared motor for vibrating horizontal and vertical shaking plate, respectively. Speed of motors is controlled by an Arduino UNO controller. After linearization, nuts are placed in a vacant slot of slotted disk. In the third stage, IR sensor is used to detect the presence of nuts under the vicinity of the camera and send the feedback signal to the camera for image capturing. Camera is mounted at a fixed location over the slotted disk. The captured image is then processed using edge detection and background subtraction algorithms implemented in Python software language. Machine prototype developed is capable of detecting the diameter of the nut with a tolerance ± 0.1 inches. In future, accuracy and nut detection speed of machines will be improved by adopting high resolution cameras and controllers.

Acknowledgements The authors acknowledge the efforts of Technology Enabling Center (TEC), Panjab University (P.U.), Chandigarh for giving us the opportunity to work on real-world problem of MSME industries. Authors are grateful to Professor Sanjeev Puri, Coordinator, TEC, P.U. and Professor Manu Sharma, Co-coordinator TEC, P.U. for continuous motivation, guidance, and support. Authors also thank Director, UIET for providing necessary facilities and support.

References

1. De Filipi JP, Wright LN (2015) Inspection and sorting machine. Patent Application Publication, 1–2
2. Rosenfeld A (1977) Digital image processing and recognition. In: Nagel HH (eds) Digitale bildverarbeitung digital image processing. Informatik—Fachberichte, vol 8. Springer, Berlin
3. Johan T, Prabuwono AS (2011) Recognition of bolt and nut using artificial neural network. In: Proceedings of the 2011 international conference on pattern analysis and intelligent robotics, ICPAIR
4. Dhenge A, Keskar P, Kuhikar A, Kawadkar P, Chaudhary T, Palasmode P, Neware M (2015) Computer vision based object sorting and fault detection using ann. In: International journal of engineering research in electronics and communication engineering (IJERECE)
5. Li B (2018) Application of machine vision technology in geometric dimension measurement of small parts(IEEE) 127
6. Priya CB, Sudha V (2019) Image processing based fault detection and isolation for mechanical components. Int J Eng Adv Technol (IJEAT)ISSN: 2249–8958
7. Gruenwedel S, Van Hese P, Philips W (2011) An edge-based approach for robust foreground detection. In: Blanc-Talon J, Kleihorst R, Philips WPopescu D, Scheunders P (eds) Advanced concepts for intelligent vision systems. ACIVS 2011. Lecture Notes in Computer Science, vol 6915. Springer, Berlin
8. Herrero S, Bescós J (2009) Background subtraction techniques: Systematic evaluation and comparative analysis. In: Blanc-Talon J, Philips W, Popescu D, Scheunders P (eds) Advanced concepts for intelligent vision systems. ACIVS 2009. Lecture notes in computer science, vol. 5807. Springer, Berlin
9. Khobragade AS, Dhenge AI (2013) Article: recognition of bolt and nut using principle component analysis. In: IJAIS proceedings on 2nd national conference on innovative paradigms in engineering and technology (NCIPET 2013) NCIPET(1), pp 30–32
10. Wang B, McDaid A, Giffney T, Biglari-Abhari M, Kean CA | Duc Pham (Reviewing Editor) Design, modelling and simulation of soft grippers using new bimorph pneumatic bending actuators. Cogent Eng
11. Al-Khazali HAH, Askari MR (2012) The experimental analysis of vibration monitoring in system rotor dynamic with validate results using simulation data. Int Scholarly Res Not vol. 2012:17, Article ID 981010
12. Agung IGAPR, Huda S, Wijaya IWA (2014) Speed control for DC motor with pulse width modulation (PWM) method using infrared remote control based on ATmega16 microcontroller. In: 2014 International conference on smart green technology in electrical and information systems (ICSGTEIS). pp 108–112
13. Syllignakis J, Panagiotakopoulos P, Karapidakis E (2016) Automatic speed controller of a DC motor using arduino, for laboratory applications
14. Chaouch S et al. (2018) DC-motor control using arduino-uno board for wire-feed system. In: 2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM). Algiers, pp 1–6
15. Lutfy OF, Selamat H, Mohd Noor SB (2015) Intelligent modeling and control of a conveyor belt grain dryer using a simplified type 2 neuro-fuzzy controller, Drying Technol 1210–1222
16. Meng X (2018) Digital image processing technology based on MATLAB. In: Proceedings of the 4th international conference on virtual reality (ICVR 2018). Association for Computing Machinery, New York, pp 79–82
17. Karim A, Andersson JY (2013) Infrared detectors: Advances, challenges and new technologies. IOP Publishing. 1–12

IR-Based V2V Communication System to Aid People Suffering from Photosensitive Epilepsy

B. Aadhi Aadhavan and Sumit Kumar Jindal

Abstract Photosensitive epilepsy refers to the condition, where a human, suffers from seizures when exposed to flickering or flashing lights or when said lights appear in a said pattern or such as stripes or bands. It is a well-known fact that the conventional method of road signaling and communication between drivers of non-automated cars is completely reliant on the use of visible lights. Every pattern or position of light communicates a certain message to the drivers around the vehicle. Evidently, we can conclude that the current system that is in place, which enables road traffic to communicate, doesn't take into account, the plight of the masses that suffer from photosensitive epilepsy. This group of people, even though they are a minority, suffers from a huge risk of being victim to a seizure while driving. Thus, emerges the need, to create a new communication system that doesn't work in the region of visible light, but at the same time, is equally accurate and reliable. Thus, this work proposes a method to replace/install a simple, cost-efficient way, to enable photosensitive epilepsy friendly, vehicle-to-vehicle (V2V) communication, with the use of modulated IR rays. This method is equally fast and reliable as the conventional method which uses visible light.

Keywords IR communication · V2V communication · Photosensitive epilepsy · Serial communication · Signal modulation

1 Introduction

Photosensitive epilepsy refers to the triggering of seizures when the victim observes flashing or patterns of light [1, 2]. These seizures may be mild, or might sometimes even be fatal. If such a condition is being triggered while a person is driving, it is certainty for an accident. A busy road is a place full of flashing and patterned light, which are used by people for communicating between each other on the road conventionally. This seemingly harmless mode of communication poses a huge threat

B. A. Aadhavan · S. K. Jindal (✉)
School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: sumitjindal08@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_24

for those suffering from photosensitive epilepsy as it triggers seizures in them, which will prove to be fatal in a busy road [3].

The laws related to photosensitive epilepsy are very hazy and doesn't restrict them from driving. Approximately, 1 in 4000 people has been diagnosed with photosensitive epilepsy. That makes the global count 2.5 million cases. Thus, photosensitive epilepsy isn't a rare phenomenon either. So, it is not uncommon to see, photosensitive epilepsy-related accidents. Research states that some countries prohibit people who suffer from photosensitive epilepsy to drive after 6:00 PM. All these fatalities and inconveniences can be avoided by re-inventing the traditional visible light-based indication system [4] to an infrared-based (IR) system.

An IR-based system not only allows the victims of photosensitive epilepsy, to have a safe and hassle-free travel but it also enables fast, secure, light-free communication which prevents light-triggered migraines and as an added benefit, reduces sound pollution. This will be a boon for those who suffer from photosensitive epilepsy, as IR isn't in the visible region of light.

Further, the use of an IR-based communication system allows us to use more complex indication patterns; as the decoding and processing is done on the microcontroller, the information can be passed to the driver's dashboard [5]. This brings about the end accidents related to misread and misunderstood road indications.

This highly effective, cost-efficient alternative to the conventional indication system, not only enables an entire minority to have a safe travel but claims a higher accuracy as compared to the conventional mode of visible light-based communication or the use of other sensors like light-dependent resistor (LDR) or ultrasonic sensors [6]. The accuracy in an IR-based system is high because the use of IR eliminates human error and is advantageous over LDR which may cause problems due to difference in luminous intensity.

The only limitation with IR-based V2V communication is the interference which could be caused by other objects. This can be solved by modulating the transmitted signal. This helps in reducing noise and interference. Further, a packet of data can be transmitted, which contains ID of the transmitting vehicle, visual cues of transmitting vehicle, and the message to be transmitted [7, 8]. This detailed information is sufficient to prevent any indication-based accidents and would help people suffering from photosensitive epilepsy.

Further, modulation increases range of communication. This is advantageous over visible light-based communication because it can work even in places where there is fog or rain, where visible light-based communication may fail due to loss of visibility. This makes sure that their percentage of accidents that happen because of lack of proper road communication decreases [9].

Additionally, modulation allows multiplexing and also helps in reducing the size of the antenna and also enables multiplexing. This enables us to have multiple vehicles communicating at the same time without experiencing cross-interference.

2 Literature Survey

A combination of IR and RF can be used to replace the conventional signaling system using horns to reduce noise pollution. The aim of the research was to get the attention of the drivers around the given vehicle. The wirelessly transmitted message was displayed on an LCD screen placed next to the driver which enables him to take decisions based on the displayed message, which was transmitted by other drivers in the vicinity. A TSHOP1738 IR receiver was used, and an IR transmitter controlled by a PIC microcontroller. The IR transmitter can be any universal transmitter. Here, a TV remote was used as a transmitter as TV remotes use the same modulation, which can be decoded by the TSHOP [10].

On studying into the accidents that take place in the highways, it can be understood why they are referred to as 'bottle necks'. Considering this, the employment of ultrasonic sensors and IR sensors for object avoidance and detection was suggested. Along with this, an anti-lane drift mechanism/algorithm was also introduced, which prevents the user from switching lanes without proper indication. This was meant to replace the auditory signaling system in everyday vehicles. Further, data security can also be integrated so as to add the model seamlessly to an IOV system. A comparative study was drafted with multiple components and other literature [11].

Research into an automobile braking system using IR results in the conclusion that they are sophisticated enough to replace the current ABS system which is in place. The suggested system calculates the rate at which the distance between two vehicles change. If there is a sudden decrease in the rate of decrease of distance, the onboard SoC triggers the vehicles braking system. This causes the braking system to come into picture and prevent the collision between the vehicles. Further, the vehicle is integrated with traction control and a throttle system. This is kept in place so as to automate throttling system after an emergency break. If the distance between the cars increases again after an emergency brake, the vehicle powers up again and assists in the throttling action [12].

An optical audio transmission system which is capable of converting audio signals into IR pulses to orchestrate an optical wireless audio system was designed. This model improves transmission distance and improves the transmission coverage region as compared to conventional IR transmission. The low-cost factor involved in making this model makes it suitable for everyday use. The device works by initially amplifying the sound and then sampling it and converting it to IR pulses which is then received by the receiver end. Here, the receiver decodes the incoming signals. This is followed by another re-amplification process. This amplified signal is produced as an output. This reduces the loss of data over transmission (attenuation) and also increases the speed at which data are transferred [13].

On studying the working and efficiency of various V2V communication models, we can conclude that V2V models are ready to be adopted into everyday life. It can be seen that the security of the model is of major concern and that serious development should be introduced into the domain of security in V2V communication as the threat of data theft grows every day. V2V communication can take place under various bandwidths, and the use of public key infrastructure (PKI) can be introduced [14].

The working of LDR and IR sensor was compared in the application of finding parking spots and assisting in parking. The tests were conducted at different luminous intensities and also at different intensities. It was found that the LDR sensor was too reliant on the luminous intensity of the environment. This introduced various threshold values which increases the complexity of the code involved. Further, the LDR gives false positives in the presence of shadows. It was proved experimentally that the accuracy of IR far surpassed that of LDR by not being affected by environmental factors, but it consumes more energy which makes it less cost-effective as compared to LDR but definitely better than the existing camera and image processing-based systems [15].

On viewing the performance of IR and ultrasonic sensor, side by side in the application of obstacle navigation and detection/avoidance, it was proved that the ultrasonic sensor was a better performer for specific kinds of obstacles, whereas the IR sensor was more efficient and accurate for a specific set of materials. The ratio of measured distance versus actual distance which was rendered from the sample data procured by the set up was useful to chart the flexibility and use of the said sensors. Further, the use of neural networks/deep learning or fuzzy logic can enhance the quality of the above data by a huge magnitude [16].

On employing and understanding the accuracy and efficiency of IR sensors of variable wavelength detection thresholds, it is evident that they are ideal for detecting and assist the landing of airplanes in the presence of fog/mist. It was identified that the usage of a sub-0.285-micron UV filter shows convincing proof that results are affected by background light intensity and thus shows good results at all times of the day. It was able to precisely deliver a visibility enhancement of 2400 RVR, with high reproducibility, whereas the naked eye is capable of providing only provided a visibility of 700 RVR, thus making this method almost 3.5 times more accurate than the use of visual aid and also more precise than any conventional method [17].

To explore visible light communication, the use of rolling shutter CMOS sensors was employed. Its results have been proven to be promising. The fundamentals of the rolling shutter CMOS sensors were explained, and further, the results were analyzed, keeping in mind, the factors of signal quality and data rate. A new measurement to define the signal quality accurately was defined, and this was the signal interference plus noise ratio. The equations necessary for calculating the data rate and the SINR were defined, and further, based on these results; simulations were done to analyze the output [18].

3 System Model

A. *System Architecture*

Figure 1 explains the basic system architecture involved in modeling the proposed

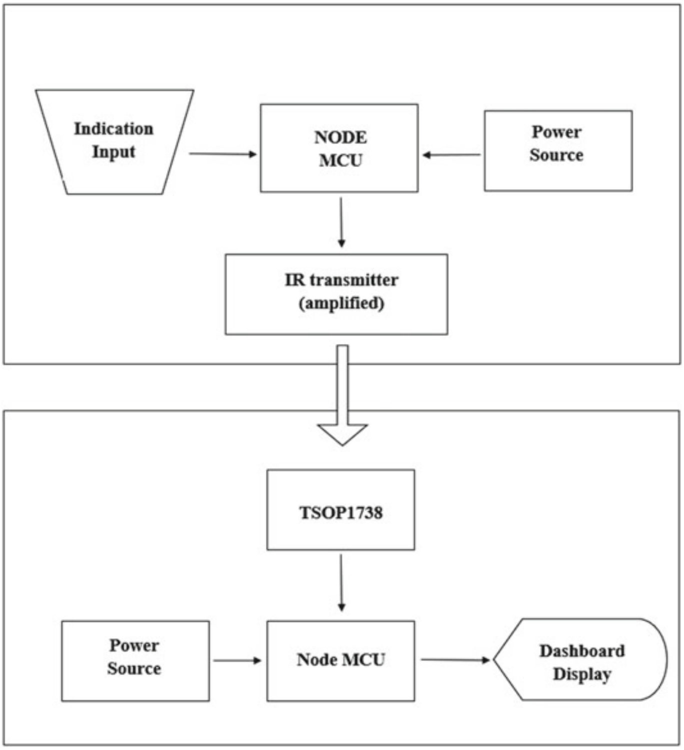


Fig. 1 System architecture of V2V system using IR

system. The architecture was designed by considering constraints like reduced latency, increased transmission distance, and low cost. The hardware choices were also made with the above constraints in mind.

The system is activated when the microcontroller (here NodeMCU) detects an input from the user. The input is the indication that the user wants to be communicated across, to the receiver. The provided input is processed by the NodeMCU, and serial communication is established with those receivers that are in the line of sight (LOS) of the transmitter via which the NodeMCU communicates.

The serially transmitted information is received by the TSOP1738, and the data are decoded. The decoded and processed data are then displayed on the receiving end's dashboard, thus enabling V2V communication via IR.

B. *Flowchart*

Figure 2 depicts the proposed system model. The model is initiated, when the user/driver requires to communicate with his other fellow drivers on the road. This is when he gives the manual input, which is his message to be transmitted. This will be done in the form of a conventional indication lever or could be modified into a button.

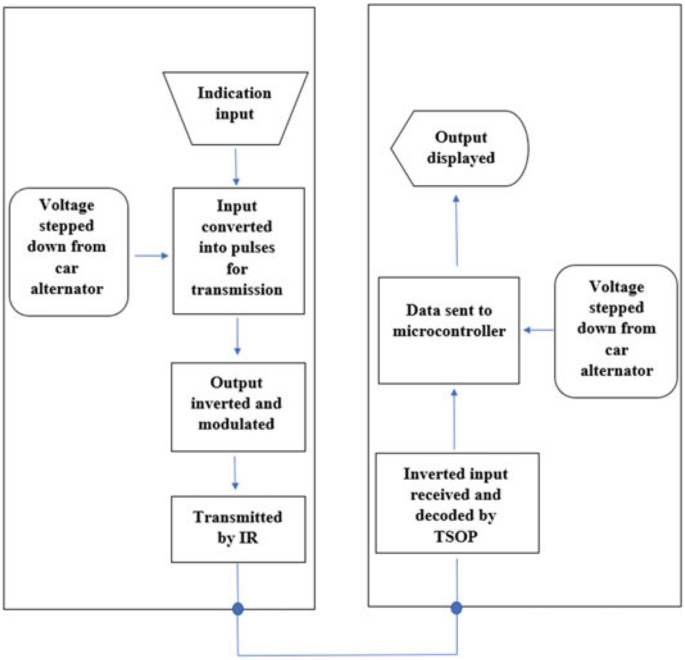


Fig. 2 System implementation flowchart

This triggered input is received by the microcontroller. The microcontroller builds a data packet with information like the car's registration plate and the vehicles color so that the users who receive this data can easily, visually identify the transmitter. This data packet is then modulated to 38 kHz. This is done, so as to increase the transmission range and to further allow multiplexing.

This modulated signal is now ready for transmission. It is limited by the receiver's ability to decode only inverted messages. Thus, the transmission has to be inverted. This is done with the aid of an n-p-n transistor. It is possible to implement inversion with just the code/software, but it has proven to be more efficient and also reduces latency if the inversion is done with the hardware component. Thus, an inverter circuit is designed before transmission. The data are now ready to be serially transmitted.

This transmitted data are now received and decoded by the receiver. The inverted data are decoded/demodulated and then sent to the microcontroller. Micro-controller, then, displays this data in the dashboard of the receiver. This allows the receiver to take a decision on whether to slow down or halt based on the incoming data.

As all the components used are low power consuming devices, the power can be directly taken from the car alternator, which is in turn powered by the car's kinetic energy.

C. *Hardware Specification and Implementation*

C.1. *Hardware Specifications*

- As seen in Fig. 3, the core of the proposed system is a NodeMCU. NodeMCU is cost-efficient IOT development platform which is open sourced. Out of the box, the firmware runs on the ESP8266 Wi-Fi SoC. The NodeMCU's GPIO pins are mapped as shown in Table 1.
- The TSOP receiver is used to detect and decode IR signals which are modulated with up to 38 kHz. It operates at a very low current of 5 mA.
- The IR LED is basically a conventional LED which allows the emission of rays with wavelength up to 940 nm. It has a viewing angle of 30 to 40 degrees with a spectral bandwidth of 45 nm.
- The BC 547 is a versatile transistor. This has a DC current gain of 800. And has a maximum emitter base voltage of 6 V. Its base current is rated at a maximum of 5 mA. An inverter circuit is designed with the BC 547 by providing the base with a resistance of 1 K Ω and the collector with a resistance of 100 Ω

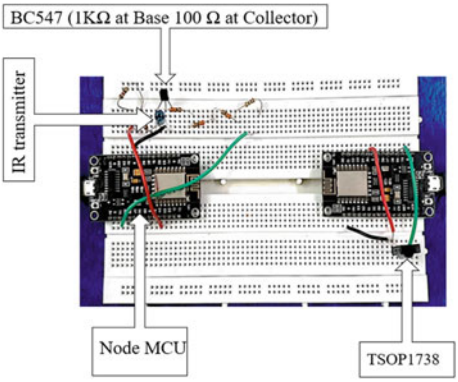


Fig. 3 Experimental setup of the proposed model

Table 1 GPIO pins of NodeMCU

I/O pin	ESP8266 pin
0	GPIO 16
1	GPIO 5
2 3	GPIO 4 GPIO 0
4	GPIO 2
5	GPIO 14
6	GPIO12
7	GPIO 13
8	GPIO 15
9	GPIO 3
10 11 12	GPIO 1 GPIO 9 GPIO 10

C.2. *Hardware Implementation*

The manual indication input in a real-life situation is simulated by software for experimentation purposes. On entering the indication into the serial monitor, for a dedicated NodeMCU, the information is first processed by NodeMCU. This data are then made into a data packet, complete with a start bit and a stop bit. Further, the license plate data and the vehicle color are also added so as to include visual aid to the driver.

This data are then modulated to 38 kHz by the NodeMCU. This allows the transmission to be resistant to external interference and also increases the range of transmission. Moreover, it allows multiplexing too.

The modulated data then undergo inversion. To serve this purpose, a BC547 N-P-N transistor has been used. It is used in its forward-biased configuration. The base is provided with a resistance of 1 K Ω and a collector resistance of 100 Ω. Thus, an inverter circuit is realized. The emitter is connected to the negative pin of the LED and the collector to the positive pin. The inverter circuit helps in inverting the output of the NodeMCU.

The modulated, inverted output is then transmitted via the IR LED. The transmitted signal is received by the TSOP 1738 IR receiver. The received signal is decoded by the receiver and is sent to the NodeMCU. The NodeMCU converts the decoded data into readable data and then displays it on the serial monitor.

4 Results and Discussion

The system has been designed with the help of NodeMCU, and its efficiency has been studied. It was identified that there is seamless transmission of data with little to 0 latency. Further, the latency remains the same for up to 10–15 m; even after which, the rise in latency is very less.

The results are as follows:

Figure 4 shows the indication input that the driver will be providing through an indication lever, or button is simulated via the IDE. Assuming the driver wishes to take a ‘right’ turn, we simulate that by entering ‘right’ in the serial monitor. The above data are now ready for serial transmission across the vehicles.

Figure 5 shows the output on the receiver end. It can be noted that the intention

Fig. 4 Input at the transmitter’s end

Fig. 5 Transmitted input, received at receiver’s end

Fig. 6 Sample input 1 at transmitter's end

of the transmitter along with the license plate number and the color of the vehicle is transmitted so as to provide visual assistance to the receiver to identify the transmitter and act accordingly. This helps the driver to understand his surroundings quicker and also in a more effective way as compared to simple visible light-based communication.

A few more sample results are as follows:

a. **Sample input**

Figure 6 shows the simulation of another common road indication that has been communicated serially via the proposed IR-based system.

Figure 7 depicts an input that cannot usually be communicated via visible light-based systems. This is due to the complexity of light patters that would have to be designed if different messages like these are transmitted. The use of IR eliminates this forthcoming of the conventional system, due to the automation of decoding and displaying of messages in the user's dashboard.

b. **Corresponding output**

Figure 8 depicts the message that will be printed at the receiver's dashboard after successful transmission of the input show by Fig. 6.

Figure 9 shows how even unconventional, complex messages can be communicated with ease with the help of IR bases V2V communication. Figure 9 depicts the output associated with the input shown by Fig. 7.

IV.1 Comparative study with existing literature

Table 2 shows the comparative study of the proposed paper with existing literature.

Fig. 7 Sample input 2 at transmitter's end

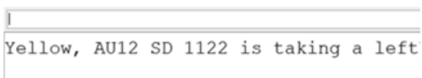


Fig. 8 Sample output 1 at receiver's end

Fig. 9 Sample output 2 at receiver's end

Table 2 Comparative study with existing literature

S. No	Title	Result	Comparison
1	Vehicle-to-vehicle communication using RF and IR technology [10]	IR serial communication was established using a universal transmitter	A dedicated transmitter was designed to establish high-speed communication
2	Design and implementation of infrared (IR) communication system [11]	Implementation of optical audio transmission system with IR	The use of TSOP1738 prevents the need for signal amplification
3	Performance study on IR obstacle sensor for automobile braking system [12]	Implementation of braking system with IR	IR is used to convey a message rather than to detect obstacles
4	Vehicle-to-vehicle communication technology [14]	The security threats in V2V communication is explored	The suggested application doesn't require privacy as information is broadcasted

5 Conclusion

A V2V IR communication system has been implemented which can aid the people suffering from photosensitive epilepsy. This system will allow those unfortunate, to have hassle-free travel without endangering their own safety or that of their fellow drivers.

Further, the modulation of the transmission to 38 kHz allows communication without interference and also increases the effective range for which the message can be transmitted. The proposed system can also help in communication in the case of fog or bad weather.

The proposed model can be upgraded with better components like IR lasers, which could increase the distance for which the message can be transmitter. Further, an array of IR transmitters and receivers might be used, to enable a constant communication system with the vehicles in the front and back, which can automatically detect and inform the user, regarding unexpected indication less lane change or change in speed etc.

Also, IR communication systems can be set all around the vehicle, thus allowing the system to predict, from where the data are being transmitted from and can supply the driver with that information as well.

References

1. Martins da Silva A, Leal B (2017) Photosensitivity and epilepsy: Current concepts and perspectives—A narrative review. Seizure 50:209–218
2. Lopes da Silva FH, Harding GFA (2011) Transition to seizure in photosensitive epilepsy. Epilepsy Res 97(3):278–282

3. Yuen ESM, Sims JR (2018) How predictive are photosensitive epilepsy models as proof of principle trials for epilepsy? Elsevier 490–493
4. Jamali AA, Rathi MK, Memon AH, Das B, Ghanshamdas, Shabeena (2018) Collision avoidance between vehicles through LiFi based communication system. IJCSNS Int J Comput Sci Netw Secur 18(12) (2018)
5. Vibin V, Sivraj P, Vanitha V (2018) Implementation of In-Vehicle and V2V communication with basic safety message format. In: International conference on inventive research in computing applications (ICIRCA 2018)
6. Eckelmann S, Trautmann T, Uller H, Reichelt B (2017) V2V-communication, LiDAR system and positioning sensors for future fusion algorithms in connected vehicles. In: 20th EURO working group on transportation meeting
7. Chauhan K, Chidrawar S, Avhad N, Gore M, Jain A (2018) Vehicular emergency system and V2V communication using IOT. Int Res J Eng Technol (IRJET)
8. Pawar O, Bharade N, Ombale S, Moje RK (2019) V2V communication protocol. IJARIIE 5(2)
9. Meena MS, Prakash V (2018) Vehicle to vehicle communication for collision avoidance. Eng Technol Appl Sci Res 6(5):1380–1386
10. Hossain E, Mamun N, Faisal MF (2017) Vehicle to vehicle communication using RF and IR technology. In: 2017 2nd international conference on electrical & electronic engineering (ICEEE)
11. Kunjir P, Shind B, Kolate S, Samge D, Patil P (2019) A vehicle to vehicle communication system. In: 2019 International research journal of engineering and technology (IRJET)
12. Sreeraman N, Sathyapriya G, Ganesan G, Ajithkumar G, Praveen Kumar S (2018) Performance study on IR obstacle sensor for automobile braking system. In: 2018 international research journal of engineering and technology (IRJET)
13. Al-nassar SI, Hatem HR, Shehab JN (2018) Design and implementation of infrared (IR) communication system. Diyala J Eng Sci
14. Demba A, Moller DPF (2018) Vehicle-to-vehicle communication technology. In: 2018 IEEE international conference on electro/information technology (EIT)
15. Bachani M, Qureshi UM, Shaikh FK (2016) Performance analysis of proximity and light sensors for smart parking. In: 2016 the 7th international conference on ambient systems, networks and technologies (ANT 2016)
16. Adarsh S, Kaleemuddin SM, Bose D, Ramachandran KI (2016) Performance comparison of Infrared and Ultrasonic sensors for obstacles of different materials in vehicle/robot navigation applications. In: 2016 International conference on advances in materials and manufacturing applications (IConAMMA)
17. Norris VJ Jr, Evans RS, Currie DG (1999) Performance comparison of visual, infrared, and ultraviolet sensors for landing aircraft in fog. In: 1999 Proceedings of the SPIE
18. Do T-H, Yoo M (2016) Performance analysis of visible light communication using CMOS sensors. Sensors 16(3):309

Experimental Design and Implementation Model of a Cashless Money Transfer Wristband

Yajat Grover and Sumit Kumar Jindal

Abstract In today's era, traditional money transfer systems are more prone to attacks like forging signatures, theft of debit cards, etc. However most of the debit cards which are now being used are RFID-based Tap to Pay cards which therefore are more vulnerable to security risks. Money Transfer Wristband offers a platform where the chances of such attacks are minimized and money can be transferred easily from one user to another at any time. This system contains a Fingerprint Sensor, LCD display and a Serial Communication port for transferring data between any trusted applications to the microcontroller. After linking microcontroller with the application, wallet balance is displayed along with details which are entered by the user on LCD for simple communication and after that registration of authenticated finger is done. After entering amount and user ID, it again asks the user to place the finger for initiating transaction and if the finger matches with the finger previously stored, transaction is successful otherwise transaction fails. Thus, this work proposes a method where the data is being transferred serially and is therefore faster as compared to the data transmission over the Internet. The system designed is independent, cost efficient and powerful which can perform most desired activities of a particular user.

Keywords Internet of Things · Arduino Uno · Fingerprint sensor · Digital money · Electronic funds transfer

1 Introduction

Transfer of money for either paying electricity bills or for utility services is required as there will always be the need for sending or receiving money from one person to another to sustain in any environment [1]. A new technology of Electronic or Digital money serves as a medium of exchange when it comes to transfer of money from one user to another [2]. It will benefit the community only when people actually start knowing about it and they start using it [3]. Most people generally carry cash

Y. Grover · S. K. Jindal (✉)
School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: sumitjindal08@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_25

when they go outside as they do not consider electronic money safe, reliable and efficient. However, transactions done are protected from third party and are harder to trace [4]. Sometimes people forget to carry cash and something comes up where they have to send money to someone then E-money plays a powerful role [5]. It also saves our time as when a large amount of money is involved, time which is needed to count the cash also increases but with Digital payments, there is no need to count the money [6]. With E-money you can deposit real money at one place and use it anywhere. Nowadays, every other person is using E-money including medical Stores, restaurants, schools, colleges, etc. which results in easy transfer of payments [7]. Some of the Applications which are widely used are GooglePay, PhonePe, BhimUPI, PayTM and which are absolutely safe and does not store any personal data [8]. With electronic money, transferring amount is much easier without involving bank, the amount simply deducts from user's bank account and gets deposited in some other account as this makes it easier for the user to send or receive money [9]. Electronic money enables digital payments which automatically reduce dependency of people on cash and with this, people can achieve broader range of financial services especially when it comes to big organizations as it offers more opportunity for revenue growth [10]. The size of the wearable would be very small and compact and will therefore be easy for wear and tear and would hopefully attract more users [11]. Every single person is now paying money through UPI and usage of Cash has been reduced exponentially in the last 4–5 years. This also saves government's money for printing cash [12]. With this, an emotion of not carrying cash should not be there and people should feel free regarding the same [13]. It also saves time as compared to the traditional systems where payments through cash was done [14, 15]. With electronic money transfer, there would be no complexity and easy transfer of money would take place [16]. This will help in creating awareness among people to switch to new methods for their development [17]. Electronic money system is better for growing and developing economy than the traditional systems [18–20]. Similarly, implementation of Money Transfer Wristband is to ensure safe and easy payments when people step outside. The system responses rapidly when addition or deletion of user's fingerprint is to be done.

2 Proposed System

A 5 V power supply has been given to the Arduino Uno as well as to the LCD display. The Fingerprint Module is connected with Arduino Uno and the Serial Monitor is used for taking the user's input. LCD display makes it easier for the user to understand the system and act accordingly (Fig. 1).

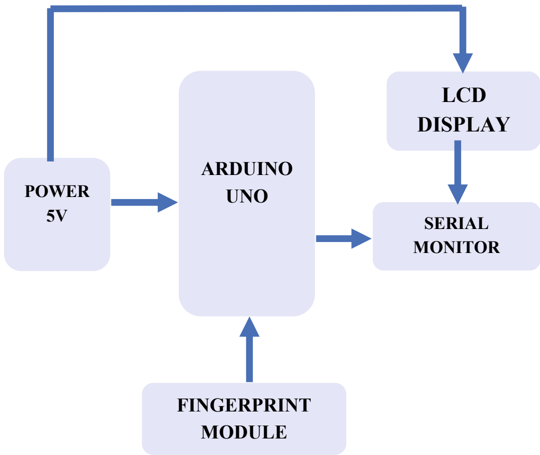


Fig. 1 Block diagram of proposed system

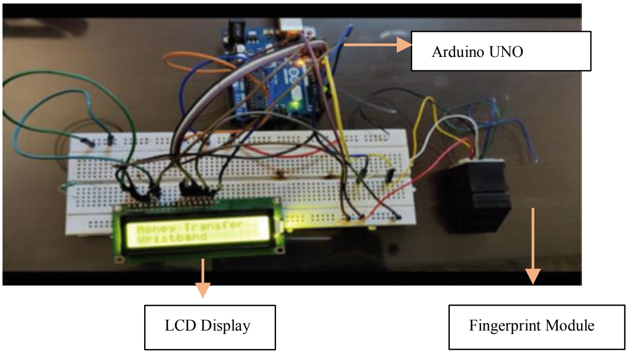


Fig. 2 Experimental setup of proposed model

3 Experimental Setup

Systems which are used nowadays are mostly encrypted with biometrics (either Face or Eye scanner or Fingerprint scanner). Arduino Uno and LCD keeps the user updated about the things which are to be done. The fingerprint sensor takes care about the security of the system (Fig. 2).

4 Hardware Used

An Arduino Uno is used for communication between the system and the user. Fingerprint Module R307 has been used for authenticating the user while completing the transaction thus ensuring full security.16 × 2 LCD display gives better understanding of the system as all the instructions are displayed in front of the user.

5 Results and Discussion

Figure 3 shows as soon as the connections are made and a 5 V power supply is given, LCD display is turned on and displays a welcome message and the product name as "Money Transfer Wristband".

Figure 4 shows that finger print module is connected and is ready to register user's fingerprints. If it is not available then the LCD displays "Finger print sensor not found". This process is done twice so as to ensure proper security and storing of fingerprint data. If the fingerprint stored is blurry or is not recorded due to some error, the system displays the error whether it was due to blurry fingerprints or any communication error.

Figure 5 shows successful registration of user's fingerprint after placing the same finger. If the second finger does not match with the first one, then the system displays "Fingerprint did not match" on serial monitor rather than storing the identity of first finger. This will maintain a level of security in the system.

In, Fig. 6 LCD displays the original balance which is there in user's wallet. If the user wishes to add more amount in the wallet, the same will be reflected on LCD display.

Figure 7 shows that the system asks the user to type the recipients name for initiating the transaction.

Figure 8 shows that the user has entered the recipient's name and the transaction

Fig 3 Welcome message on starting up the system

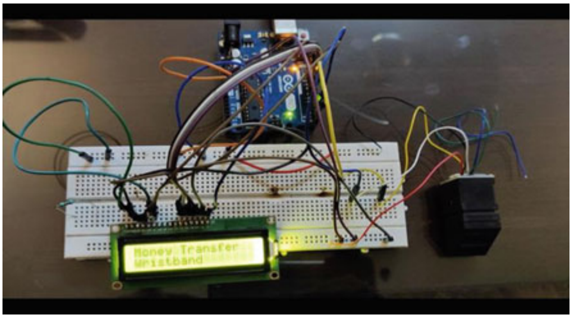
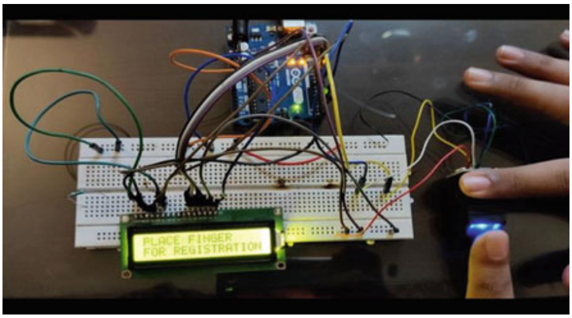


Fig. 4 Enrolling fingerprints of the user



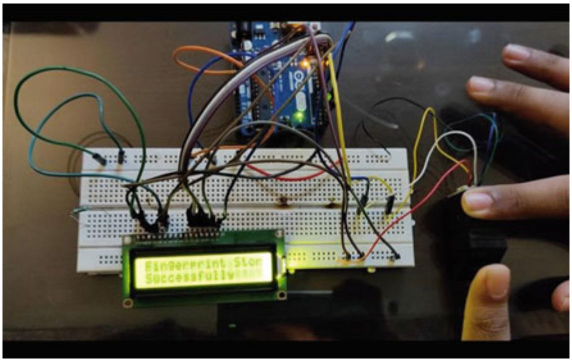


Fig. 5 Fingerprint is successfully registered

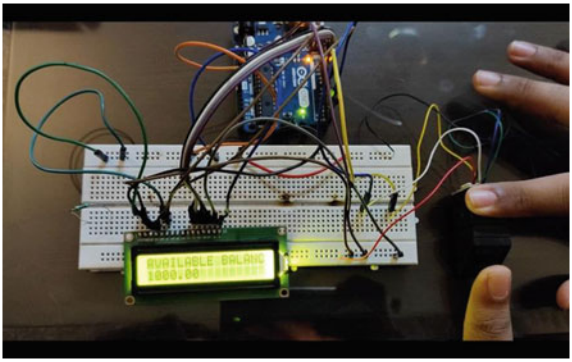


Fig. 6 Wallet balance

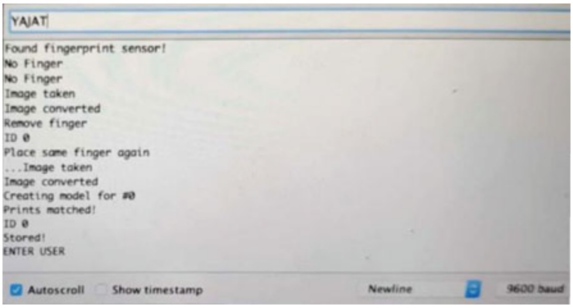


Fig. 7 Entering recipient through serial monitor

has been initiated.

Figure 9 shows the amount being entered which is to be sent to the recipient typed above. This typing of data is being done from users end, through serial monitor.

Alternatively, any communication type can be used for example, a Bluetooth Module HC-05 or an application from which the user will be sending the data.

Figure 10 shows the amount entered by the user for the initiated transaction. Since

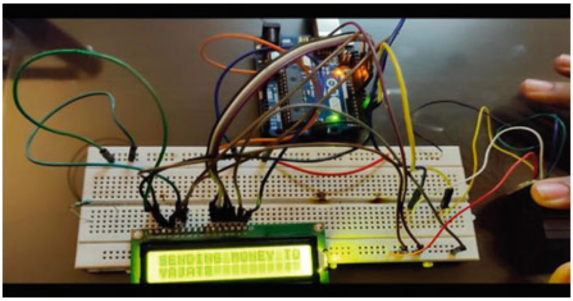


Fig. 8 Recipient displayed on LCD

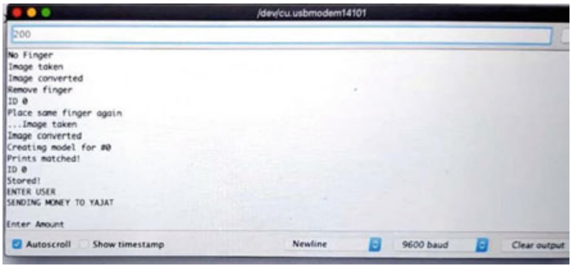


Fig. 9 Entering amount through serial monitor

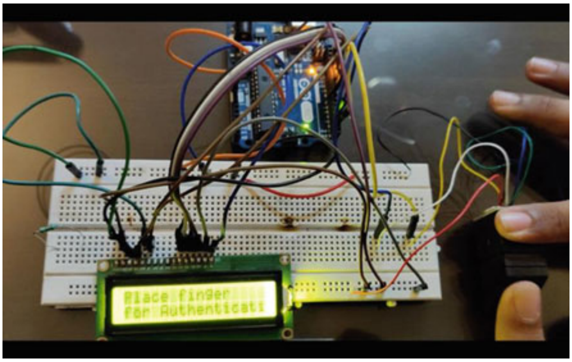


Fig. 10 Amount that is to be sent

the available balance is Rs. 1000, if the entered amount is greater than Rs. 1000, then the system displays “Insufficient Funds” and the transaction fails.

Figure 11 shows the amount that the user entered and after typing in the amount, authentication is required to complete the transaction.

Figure 12 shows that the user’s fingerprints are required for completing the transaction. When the fingerprints get matched with the fingerprints stored previously, the transaction is successful.

Figure 13 shows the remaining balance in the wallet after a successful transaction.

Figure 14 shows that when the fingerprints does not match with the fingerprints stored previously or the amount entered is greater than the available balance, LCD

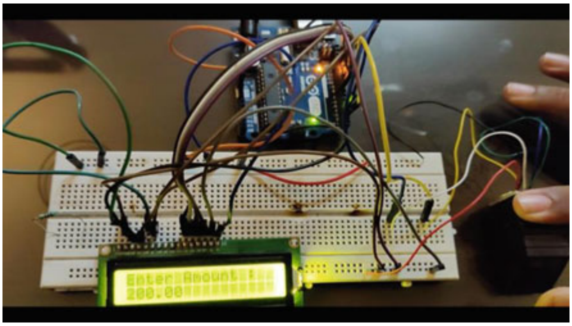


Fig. 11 Entered amount

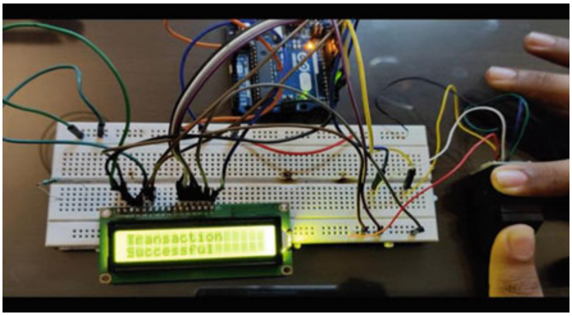


Fig. 12 Authentication successful

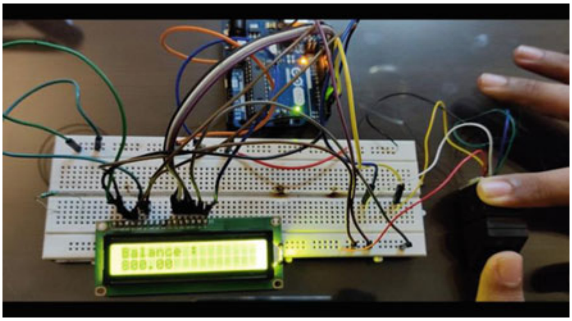


Fig. 13 Remaining balance

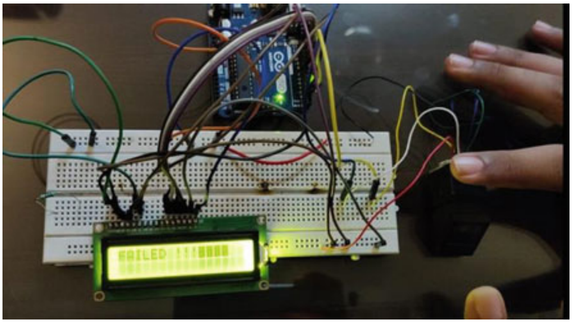


Fig. 14 Failed transaction

displays “FAILED !!!” as the transaction fails. Since, in this system giving any input to the microcontroller has been set to a delay of maximum of 3 s. If that time is exceeded, it automatically fails resulting in a failed transaction. So, the scanning of fingerprint at last step in order to complete the transaction has to be done within three to four seconds otherwise the transaction would fail. This can be altered as per the user’s need. After the process is completed, system goes back to the default page waiting for the same user to initiate another transaction. If the user is changed then system will start again and will ask the user to place fingerprints for registration and the whole process is repeated.

6 Comparitive Study of the Proposed Model with Existing Literature

S. No	Title	Result	Comparison
1	Universal electronic cash [10]	In this system, a smart card is equipped with Rabin scheme chip to store data on the server	The proposed system uses serial data transfer technique which minimizes data traffic and is comparatively faster
2	A secure and privacy-preserving mobile wallet with outsourced verification in cloud computing [11]	Reliable mobile wallet is proposed using cloud computing	This system does not involve cloud and internet as due to network traffic sometimes the commands gets delayed and system takes time to respond
3	Embedding a Digital Wallet to Pay with a Selfie [14]	Payment is done using a face scanner or a selfie	A little more time is consumed using a face scanner as compared with a fingerprint scanner
4	Providing security for e-wallet using e-cheque [20]	Major concern is based on potential security problem regarding misuse of e-cash	The proposed system uses electronic payment methods (EPS) for transferring money which is safe and secured

7 Conclusion

The low-cost, reliable and efficient product has been designed and tested with software and hardware debugging and the data being entered through serial monitor by the user has been transmitted successfully. Placement of each module and sensors has been done carefully thus resulting in the best working of the product. Implementing this system on a real network and platform can be done for future work. It is evident from the comparative study that the proposed methodology minimizes network traffic and is comparatively faster than the existing technologies. Many further improvements can be made in this system such as adding sensors that would display the real time temperature and weather on the LCD display and give alert when the weather is not good. Quick payments can be done using this and when people step outside to grab something. Using certain modules and libraries and growing technology, the product has been successfully implemented.

References

1. Pshenichnikov VV, Babkin AV (2017) Electronic money as a factor in the development of the digital economy Scientific and technical bulletins of the St. Petersburg State Polytechnic University. Econ Sci 10(1):32–42
2. Freedman C (2002) Monetary policy implementation: Past, present and future—will the advent of electronic money lead to the demise of central banking? Int Finance 3(2):211–227
3. Okamoto T, Ohta K (1991) Universal electronic cash. In: Proceeding of the of CRYPTO'91, LNCS 576. Springer, Berlin, pp 324–337
4. Chaudhary T (2017) The critical analysis of cashless transaction. Int J Commer Manage Res 3(3):92–94
5. Lin T, Gao Q (2006) The research of bluetooth protocol stack. Beijing Jiaotong University Press, pp 34–37
6. Xin W, Yan G, Wang W (2010) Study of a wireless power transmission system for an active capsule endoscope. Int J Med Robot 6:113–122
7. Balakrishnan M (2010) Indian payment systems and financial inclusion: Current status and next steps. J Payments Strategy Syst 9(3):242–293
8. Payment Interface (UPI) (2017) A way towards cashless economy. Int Res J Eng Technol 4:762–766
9. Breloff P, Rotman S (2011) An overview of the G2P payments sector in India. CGAP
10. Acharya RN, Kagan A, Lingam SR (2008) Online banking applications and community bank performance. Int J Bank Mark 26(6):418–439
11. Qina Z, Suna J, Wahaballaa A, Zhenga W, Xiongb H, Qina Z (2016) A secure and privacy-preserving mobile wallet with outsourced verification in cloud computing. Elsevier J. Comput Stand Interfaces. http://dx.doi.org/10.1016/j.csi.2016.11.012
12. Varian HR (2010) Intermediate microeconomics: A modern approach. W.W. Norton & Company, 8th Edition, pp 270–286
13. Rivest R, Shamir A (1997) Payword and micromint: Two simple micropayment schemes. In: Proceedings of security protocols workshop, LNCS 1189. Springer, Berlin, pp 69–87
14. Houngbo PJ (2018) Embedding a digital wallet topay-with-aselfie, from functional requirements to prototype, vol. 206. Springer International Publishing
15. Vij A, Namboodiri A (2014) Learning minutiae neighborhoods: a new binary representation for matching fingerprints. In: 2014 IEEE conference on computer vision and pattern recognition workshops

16. Pukkasenunk P, Sukkasem S (2016) An efficient of secure mobile phone application for multiple bill payments. In: Proceeding of the IEEE 30th international conference on advanced information networking and applications workshops (WAINA), pp 487–492
17. Committee on digital payments medium term recommendations to strengthen digital payments ecosystem report. Ministry of Finance, Government of India, Rep. 271216, 09 Dec 2016
18. Shah A, Roongta P, Jain C, Kaushik V, Awadhiya A (2016) Digital payments 2020: The making of a $500 billion ecosystem in India, The Boston Consulting Group, India and Google, India, Rep. TCM21–39245, July 2016
19. Szydlo M (2004) Merkle tree traversal in log space and time. In: International conference on the theory and applications of cryptographic techniques. Springer, Berlin
20. Yahid B, Nobakht MB, Shahbahrami A (2013) Providing security for e-wallet using e-cheque. In: 2013 7th international conference on ecommerce in developing country, pp 1–14

Smart Wheelchair Management System for Disabled People

Ishan Patel, Krish Sethi, Simrit Kaul, Soumya, and Sumit Kumar Jindal

Abstract Over the recent years, the field of integrating human machine interface to power wheelchairs for autonomous driving control has gained great attention by the scientific and engineering communities. The fundamental objective of the human—wheelchair interface is to permit the individual to control the versatility of the seat in least exertion and with more vigour and security. A large variety of electric powered wheelchairs do not meet the needs of a substantial proportion of users because of requirement of muscle force and accuracy. An assistive component for executing well-heeled wheelchair motion for seriously incapacitated individuals is introduced. For making the lives of disabled people easier and less dependent, the wheelchair has been incorporated with eye control interface for the movement and direction control. For the safety of the patient, obstacle avoidance and fall alert mechanisms have been included. Also, body temperature and heart rate are two parameters that are considered for health monitoring of the patient.

Keywords Smart wheelchair · Eye control · Spinal cord injury · Human machine interface · Internet of Things · MEMS sensor

1 Introduction

From the insights of the on-going years, the quantity of paralytic residents is on the ascent. As per the National Census 2011, in India out of 1.2 billion populace, 26.8 million individuals are 'handicapped' which is 2.21% of the absolute populace. These people have either temporary or permanent disabilities due to illness or accidents. It is realized that handicap power is higher in evolving countries. An enormous segment of populace that falls into this category cannot resume normal life in light of detachment to assistive gadgets. This ailment can frame an endless loop and a tiny bit at a time can spoil the life of the person. Inability ought to be not any more an impediment to individual and social improvement in this time. These disabled people must

I. Patel · K. Sethi · S. Kaul · Soumya · S. K. Jindal (✉)
School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: sumitjindal08@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_26

be offered with gadget which may equip them portability alongside some progressive control attributes so they may have at least an average existence. Independent mobility decreases reliance on guardians and relatives, increases vocational and educational opportunities, and advances sentiments of confidence. For infants, self-reliant portability fills in as the establishment for early stage of learning. Youngsters without protected and self-ambulation are denied basic learning opportunity which places them at a formative disservice. For elders, unconstrained movement is a significant device of confidence and takes on a vital job in settling with the age. For instance, if more elder individuals think that its much hard to trundle themselves to the objections, they may dodge or do so less regularly. Versatility challenges are likewise, the core contenders of activities of daily living (ADL) and instrumental ADL handicaps due to the requirement to move to achieve huge numbers of such pursuits. Despite the fact that the requirements of several people with inabilities could be mollified with conventional manual or steerable wheelchairs, a portion of the debilitated network thinks that it is troublesome or strenuous to make use of wheelchairs freely. It is critical to build up an intelligence driven wheelchair whose movement is interceded by an electronic framework which knows about environment and can link with the client to accomplish versatility objectives and maintain a strategic distance from risky circumstances [1–3]. This decreases the client's human exertion to propel the wheels of the wheelchair. Autarchic perception wheelchairs are regulated by the Human UI where the human settles on choices at the most significant intensity of activity and the shrewd control innovation makes the rest of the movement programmed [4, 5]. The insight is added to a wheelchair stage around client control regardless of their incapacities, which makes the investigation of the human–machine interface (HMI) between the client and the wheelchair a significant assistive automated branch of knowledge [6].

2 Literature Review

A huge assortment of electric instigated wheelchairs that utilize distinctive human machine interfaces (HMI), for example, sip and puff control, head movement, speech identification, jaw control and EEG signals have been incorporated since years. Rabhi et al. [7] used fuzzy logic with Vector Field Histogram (VHF) algorithm to develop joystick controlled wheelchair. The signal coming from the VFH refining algorithm and the change in the action of the hands of the debilitated which is identified from the information received from the joystick is acknowledged by the fuzzy logic controller. The fuzzy logic generates output in the form of signals that decide the speed and angle of motor's rotation. LIDAR sensor has been utilized to distinguish the deterrents before the wheelchair. The VFH algorithm processes these signals from LIDAR and gives the input to the fuzzy algorithm which decides the direction to avoid collision from the object coming in path of the wheelchair. Abedan Kondori et al. [8] have instigated an unmediated process for 3-D head pose estimation. The head developments are assessed from a grouping of profundity pictures obtained by Kinect. A succession of pictures are apprehended by Kinect which is put before the client. Head restriction, head pose assessment and direct 3-D movement assessment are finished. To distinguish and confine the client head, image sequence is then sent to the system. Finally, the steer of the wheelchair is controlled by mapping the recovered motion parameters onto the control commands of the wheelchair. Sivakumar et al. [9] have presented voice controlled wheelchair. Speech customisation is used to prepare the system to interpret signals from the sound commands. Voice capture module records speech commands using the microphone found in the speech package. It then transforms the received commands into binary codes, depending on the speaking command frequency. Comparing the ported binary codes with the one stored in the microcontroller is the responsibility of the Voice Recognition sub module. Then if both are remarkably similar, it compares these commands. Achkar et al. [10] have introduced Mobile controlled wheelchair, which allows the clients to locate the position of their wheelchairs and move them to the intended location through the use of special android application. The router was used as a source of direct link between the Mobile App, Arduino, and the relays in this project. An IP camera over the front end of the wheelchair was incorporated to visualize the desired trajectory. The IP camera analyses the data via its unique IP address provided by the router, in order to trace and connect. The router will then send the information packets via Wi-Fi or Internet to the user-installed android device. Thakur and Kulshrestha [11] have presented eye controlled wheelchair. The system consists of a camera, a software for image recognition and a motor controller to control a wheelchair. Shape-adjusted mean shift algorithm is used as mapping algorithm in NI Vision assistant along with NI Vision Acquisition and NI Laboratory View. The outcomes of eyeball monitoring are then used to generate sufficient wheelchair movement. The Chin control interface [12] uses a force-sensing joystick shaft that does not require any precise movement of the face. The person needs to exert force using the neck muscle to provide the input. The necessary force is usually between 0.09 and 0.3 kg. Only people who possess significant strength in their neck muscles will embrace this technology. The Sip-n-puff interface [13, 14] is regulated by the sipping (ingestion) and puffing (exhalation) commands on a pneumatic tube. Sharp sips and puffs are expected to change the trajectory of the wheelchair whilst lower/softer tiers of sips and puffs can achieve the steering. These instruments are mostly designed for speed control and each time require tremendous precision in the sips and puffs. They are still constrained by the fact that they need an external starting point. The tongue touch keypad (TTK) control interface is the only language controller device currently available. It comprises of switches that are installed into the dental mouthpiece and that fit into the mouth roof. The big drawback of such a device is that they must be replaced regularly before feeding and even drinking water.

3 System Overview

The control structure as represented in Fig. 1 for the smart wheelchair comprises the

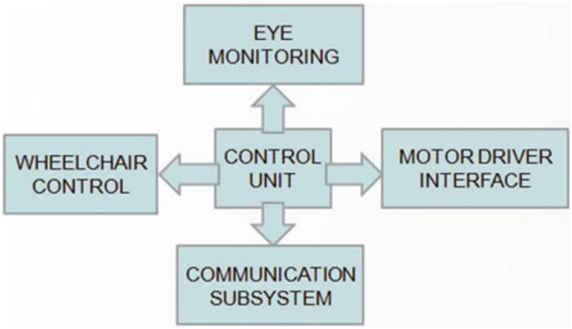


Fig. 1 System overview

following sub-blocks:

- Control Unit—Control unit consists of Arduino UNO. The control unit is responsible for all the decisions and data processing.
- Wheelchair Power Motors—The movement of the wheelchair is controlled via motor driver and DC motors.
- Wheelchair Control Interface—The eyeball control unit, the accelerometer MEMS sensor, Ultrasonic sensor.
- Communication Subsystem—This consists of the GSM module, Wi-Fi module, IoT platform.

A. ***Wheelchair Direction control***

The wheelchair model is attached with wheels at both the ends and at the centre. The DC motors are connected to the wheels that convert the electrical energy to mechanical energy and help to rotate the wheels. DC motor driver drives the motors by providing the necessary current to provide the torque. It uses L293D H-bridge motor driver IC with Voltage regulator 7805. This IC provides a +5 V regulated power supply with provisions to add a heat sink. They for the most part go about as current amplifiers that acknowledge the low current signal from the controller and convert it into a high current signal which assists with driving the motor.

Eye sensor (QTR-1RC) with an operating voltage of 5 V is mounted in front of the patient's eyes. The sensor consists of infrared transmitter and infrared receiver. There is an angle of ±45° between the transmitter and receiver which acts as the directivity of the sensor. The optimum sensing distance for this particular sensor is 3 mm and the maximum is 9.5 mm. The infrared transmitter continuously transmits the infrared rays in the surrounding. When an object appears in front of the infrared rays, it is mirrored backwards and processed by the infrared receiver. In this scenario, when the eye sensor is kept in front of the eyes of the patient, the opening and closing of eyes produce electrical signal of different intensity due to difference in amount of reflected infrared rays. When the eye is open, the black ball in the eye absorbs some amount of infrared rays and rest of the rays are reflected back to the receiver. Whereas, when the eye is closed almost all of the emitted infrared rays are reflected back to the transmitter. This change in the amount of intensity of infrared rays received by the

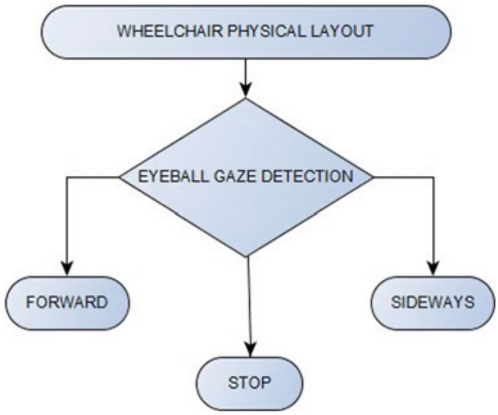


Fig. 2 Wheelchair movement control

receiver of the eye sensor [15, 16], produces a HIGH as well as LOW signal. Two such sensors are used for each eye which produces for combinations of HIGH and LOW signals which instructs the wheelchair for movements in four eye controlled ways: FORWARD, LEFT, RIGHT and STOP. Inputs from both the eye sensors are fed to the Arduino board. When both the eye sensors produce a LOW voltage signal, the wheelchair motor is given command to STOP. When the right eye sensor produces a HIGH voltage signal whereas the left eye sensor produces a LOW voltage signal, a command to turn towards RIGHT direction is given to the wheelchair motor. When the left eye sensor produces a HIGH voltage signal whereas the right eye sensor produces a LOW voltage signal, it turns LEFT.

The following flowchart in Fig. 2 shows the steps to determine the movement of the wheelchair.

B. *Obstacle Detection and Collision Avoidance*

HC-SR04 ultrasonic sensors are used to detect obstacle and measure the distance from the obstacle. The operating voltage of this sensor is 5 V and its range is between 2 and 400 cm with a resolution of 0.3 cm. The sensor is attached in front of the wheelchair which helps to identify any obstacle in the path. The transmitter (trig pin) sends a 340 m/s output signal of high frequency in the local field. If the signal hits an entity, it is mirrored, and it is received by the receiver (echo pin). The echo pin provides the cumulative time of sound wave propagation. Since, the ECHO pin give the total propagation time which includes the time taken by the sound wave to hit the target and the time taken to reach the receiver. The distance considered is double the actual distance between the ultrasonic sensor and the object. Therefore, the distance of obstacle is:

$$d = \frac{\text{speed of the sound wave} \times \text{total propagation time}}{2}$$

Buzzer alarms the patient if the distance of the object is less than 20 cm. The patient is alarmed immediately so that he can stop or change the direction of the

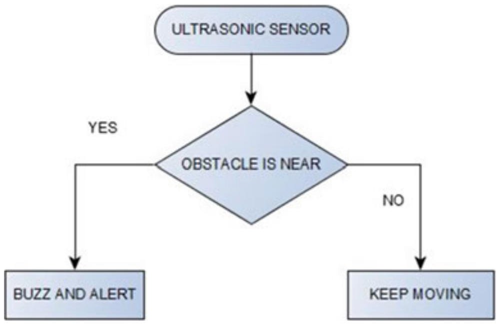


Fig. 3 Collision avoidance

wheelchair as soon as possible. This helps to avoid any collision between the patient and the object. The following flowchart in Fig. 3 shows the steps to determine any obstacle in the path of the wheelchair.

C. ***Health Monitoring***

Body temperature and heartbeat of the patient is continuously monitored. LM35 temperature sensor and heart beat sensor measures the values and relay that in the form of a signal to the Arduino UNO. The LM35 sensor has a wide range of −55 °C to 150 °C with an operating range of 4 V to 30 V. The heart beat sensor operates at 100 mA and the heart beat detection is indicated by LED and Output High Pulse. Wi-Fi module containing the ESP8266 module helps the controller connect to the Internet [17]. It operates at 3.3 V and a baud rate of 9600. A serial connexion is established between the Wi-Fi module and Arduino, as well as the Wi-Fi module and the data storage platform for Internet. This data is sent to the cloud by the Arduino which is assisted by the Wi-Fi Module. These data can be retrieved even in the form of graph for statistical analysis. The following flowchart in Fig. 4 shows the steps to monitor the health of the patient.

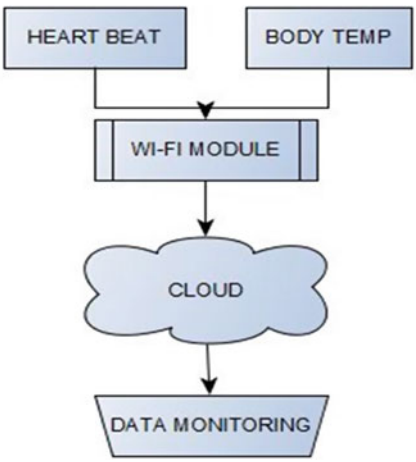
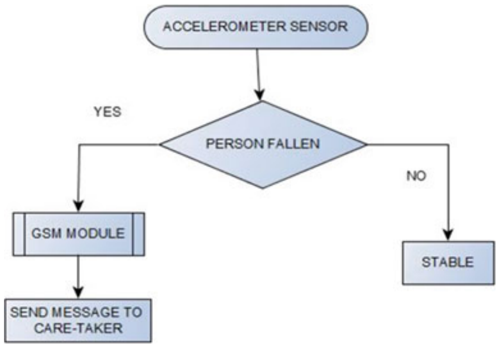


Fig. 4 Health monitoring

Fig. 5 Fall detection



D. *Fall Detection*

The accelerometer sensor measures the orientation [18] of the wheelchair. It operates within the range of 3–6 V (DC) and provides an analogue output. It has a measuring range of ±3 g. For a negative value of X-axis, the wheelchair is considered to be fallen. In this case, the GSM module decides to send an alert message on mobile to the caretaker. The mobile number of the caretaker is provided in the Arduino code. This helps to provide the patient immediate help. If all the three axes values are positive, the wheelchair is considered to be stable and thus, no action is taken. The following flowchart in Fig. 5 shows the steps to detect the fall of the patient.

E. ***Technical Specifications***

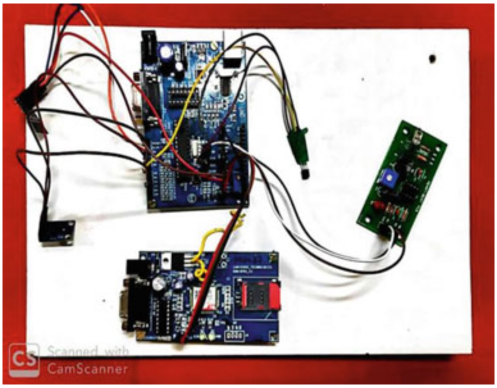
The section below discusses all the technical aspects of the sensors and actuators used in the wheelchair model.

4 Results and Discussion

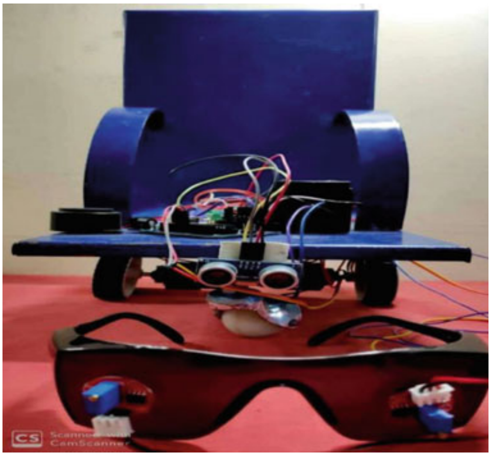
Figure 6a and b present the hardware setup for the proposed design.

(a) The wheelchair direction control is controlled by the eyes. The wheelchair moves forward, left, right and stops according to the input from the eye sensor. When a HIGH signal is received from both the eyes then the wheelchair moves forward. When a HIGH signal is received from only left eye, the wheelchair

Fig. 6 **a** Control unit **b** Complete device arrangement



(a)



(b)

turns left. When a HIGH signal is received from the right eye, the wheelchair turns right. Finally, if LOW signal is received from both the eyes, the wheelchair STOPS.

(b) When an obstacle comes in the way of wheelchair, the patient is alarmed by the buzzer. The ultrasonic sensor successfully detects the presence of any small object in front of it in the range of 2–400 cm with a resolution of 0.3 cm. Thus, obstacle avoidance subsystem works efficiently.

(c) The temperature and heart beat of the patient is continuously measured. The values are displayed on the serial monitor of the Arduino IDE as shown in Fig. 7. This data is successfully uploaded on the Internet and displayed in the form of a graph for data visualization. This can be seen in Fig. 8. Also, the data is updated time to time to keep a hold of the real time data.

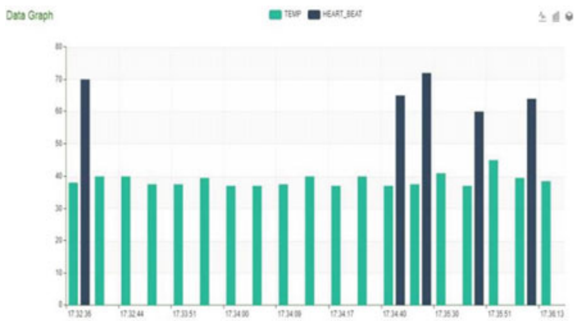
(d) The caretaker is informed if the wheelchair is detected to be fallen. The accelerometer sensor measures the x-orientation of the wheelchair. The GSM module is connected to a 4G mobile network. It sends SMS message, as represented in Fig. 9 immediately to the registered phone number so that the patient can get aid from the caretaker.

Fig. 7 Fall detection on serial monitor

```
TEMPRATURE = 28.81*C
TEMPRATURE = 83.86*f
3.69
/iotmakers/upload.php?id=sssdbkgf&data1=28.81&data2=0TCP connection ready
Sending..
Packet sent

 +IPD,221:HTTP/1.1 200 OK
Content-Txpe: text/html; charset<UTF-8
Soy
-Pa0 n
TEMPRATURE = 27.34*C
TEMPRATURE = 81.22*f
10.03
/iotmakers/upload.php?id=sssdbkgf&data1=27.34&data2=0TCP connection ready
Sending..
TEMPRATURE = 28.32*C
TEMPRATURE = 82.98*f
-41.54
AT
AT+CMGF=1
AT+CMGS="9431437479"
WARNING::wheel chair has fall down
 /iotmakers/upload.php?id=sssdbkgf&data1=28.32&data2=0TCP connection ready
Sending..
Packet sent
```

Fig. 8 Bar Graph showing temperature and heart beat



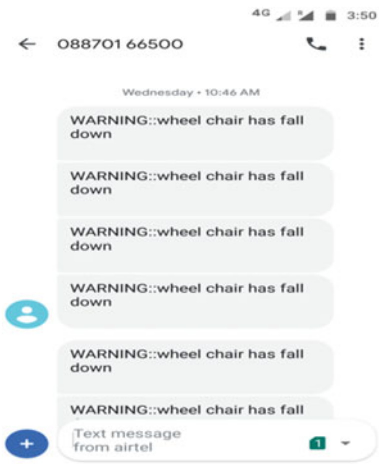


Fig. 9 Alert SMS to caretaker

5 Comparative Study

The proposed work is compared with the existing literature in the tabular format given below (Table 1).

The proposed work focuses on distraction free movement of the wheelchair ensuring more safety than the works suggested in the literature survey. The constant monitoring of the temperature and pulse rate and simultaneous uploading of data to the internet proves to be an additional feature. Automatic communication of a distress signal when a fall is detected evinces an auxiliary functionality.

Table 1 Functional comparison for different features of the electric powered wheelchair (EPW)

Wheelchair type	Features
Multimodular Wheelchair (Ben et al. [1])	HMI to control the movement of EPW by gaze tracking technique using a webcam
Electric Wheelchair for Paralyzed Users (Arai et al. [19])	EPW movement is strictly controlled by motion of the eye using a camera attached on the patient's glasses
Electric Wheelchair control using head gesture (Ohtsuka et al. [2])	The motion of the wheelchair is based on the gestures of the head which is detected by the depth sensor
Touchscreen-based multifunctional wheelchair (Makwana et al. [3])	The portability of the wheelchair can be directed by one finger touch along with obstacle detection
Smart Wheelchair Management System [Proposed Work]	EPW movement is guided using the eye sensor in company with obstacle detection and health monitoring system

6 Conclusion

This mechanism provides an assistive method for the introduction of secure wheelchair motility for chronically disabled individuals. Eye controls four motion of the wheelchair, i.e. FORWARD, LEFT, RIGHT and STOP. Accelerometer sensor detects the orientation of the wheelchair and informs the caretaker immediately using GSM module in case the person falls from the wheelchair. Ultrasonic sensor measures the distance of the obstacle and alarms the person when its distance is less than 20 cm. It continuously monitors health of the patient. The temperature and heart beat sensor upload continuously the data on the Internet through Wi-Fi module for statistical analysis.

References

1. Taher FB, Amor NB, Jallouli M (2015) A multimodal wheelchair control system based on EEG signals and eye tracking. In: 2015 international symposium on innovations in intelligent systems and applications (INISTA). Madrid
2. Ohtsuka H, Kato T, Shibasato K, Kashimoto T (2015) Non-contact head gesture maneuvering system for electric wheelchair using a depth sensor. In: 2015 9th international conference on sensing technology (ICST). Auckland
3. Makwana SD, Tandon AG (2016) Touch screen based wireless multifunctional wheelchair using ARM and PIC microcontroller. In: 2016 international conference on microelectronics, computing and communications (MicroCom). Durgapur
4. Gor NJ, Jeyakumar A (2017) Voice controlled motorized wheelchair with real time location monitoring. In: 2017 International conference on energy, communication, data analytics and soft computing (ICECDS). Chennai
5. Ma C, Li W, Cao J, Gravina R, Fortino G (2016) Cloud-based wheelchair assist system for mobility impaired individuals. In: Li W et al. (eds) Internet and distributed computing systems. IDCS 2016. Lecture notes in computer science, vol. 9864. Springer, Berlin
6. Škraba A, Stojanović R (2015) Speech-controlled cloud-based wheelchair platform for disabled persons, Microprocess Microsyst 39(8):819–828, ISSN 0141-9331, https://doi.org/10.1016/j.micpro.2015.10.004
7. Rabhi Y, Mrabet M, Fnaiech F, Gorce P (2013) Intelligent joystick for controlling power wheelchair navigation, In: *3rd International Conference on Systems and Control, 2013*, pp 1020–1025. https://doi.org/10.1109/ICoSC.2013.6750981
8. Abedan Kondori F, Yousefi S, Liu L, Li H (2014) Head operated electric wheelchair. In: 2014 southwest symposium on image analysis and interpretation, pp 53–56, https://doi.org/10.1109/SSIAI.2014.6806027
9. Sivakumar MS, Murji J, Jacob LD, Nyange F, Banupriya M (2013) Speech controlled automatic wheelchair. In: 2013 Pan African international conference on information science, computing and telecommunications (PACT). pp 70–73. https://doi.org/10.1109/SCAT.2013.7055093
10. Achkar R, Haidar GA, Dourgham H, Semaan D, Araji H (2015) Mobile controlled wheelchair. IEEE Eur Model Sympo (EMS) 2015:429–434. https://doi.org/10.1109/EMS.2015.68
11. Thakur B, Kulshrestha K (2014) Eye controlled electric wheelchair. IEEE Int Conf Comput Intell Comput Res 2014:1–6. https://doi.org/10.1109/ICCIC.2014.7238451
12. Guo S, Cooper RA, Boninger ML, Kwarciak A, Ammer B (2002) Development of power wheelchair chin-operated force-sensing joystick. In: Proceedings of the second joint 24th

annual conference and the annual fall meeting of the biomedical engineering society. Engineering in Medicine and Biology, vol. 3, pp 2373–2374. https://doi.org/10.1109/IEMBS.2002.1053329
13. Jones M, Grogg K, Anschutz JR, Fierman R (2008) A sip-and-puff wireless remote control for the apple iPod. RESNA
14. Ramudu PV, Krishna PPM (2013) Tongue operated robotic wheel chair using tongue drive assistive technology. Int J Eng Res Technol (IJERT) 02(09), Sept 2013
15. Plesnick S, Repice D, Loughnane P (2014) Eye-controlled wheelchair. IEEE Canada Int Humanitarian Technol Conf (IHTC) 2014:1–4. https://doi.org/10.1109/IHTC.2014.7147553
16. Juhong A, Treebupachatsakul T, Pintavirooj C (2018) Smart eye-tracking system. Int Workshop Adv Image Technol (IWAIT) 2018:1–4. https://doi.org/10.1109/IWAIT.2018.8369701
17. Verma P, Sood SK (2018) Fog assisted-IoT enabled patient health monitoring in smart homes. IEEE Internet Things J 5(3):1789–1796. https://doi.org/10.1109/JIOT.2018.2803201
18. Bourke AK, O'Brien JV, Lyons GM (2007) Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. Gait Posture 26(2):194–199. https://doi.org/10.1016/j.gaitpost.2006.09.012
19. Arai K, Mardiyanto R (2011) A prototype of electric wheelchair controlled by eye-only for paralyzed user. J Robot Mechatron 23(1):66–74. https://doi.org/10.20965/jrm.2011.p0066

A 2–10 GHz Common Gate UWB Low Noise Amplifier in 90 nm CMOS

Ashutosh Anand, Soumendra Kumar Dash, Krishna Datta, Deepak Prasad, Srikanta Pal, and Vijay Nath

Abstract The enormous demand for low power devices with low noise invites tremendous research in the last decade. One of the key devices which is in great demand is a low noise amplifier. In this paper, a UWB LNA has been proposed using a cascaded common source and common drain configuration. There has been good matching of input common gate stage with antenna and able to produce the S11 below −20 dB and output matches with common drain output with S22 below −10 dB. The common source configurations have provided a gain of 8.7299 dB. There is excellent reverse isolation of −90 dB, and the noise figure varies between 5 and 10 dB within the targeted frequency range. The 1 dB compression point of the LNA is −27 dBm and has the IIP3 (input third intercept point) as 4.3 dBm at 7.747 GHz at the input voltage of 1 V.

Keywords Common gate · Common source · Common drain · UWB · Low noise amplifier (LNA)

A. Anand (✉)
Department of Electronics and Communication Engineering, Presidency University, Rajankunte, Bengaluru, Karnataka 560064, India
e-mail: ashunitjsr@gmail.com

S. K. Dash · K. Datta · S. Pal · V. Nath
VLSI Design Group, Department of ECE, B.I.T. Mesra, Ranchi, JH 835215, India
e-mail: spal@bitmesra.ac.in

V. Nath
e-mail: vijaynath@bitmesra.ac.in

D. Prasad
Department of Electronics and Communication Engineering, Sarla Birla University, Ranchi, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_27

1 Introduction

The high demand for devices that can process data at a higher speed in the recent past increases the demand of the receivers' needs with large bandwidth and high sensitivity, making high-speed wireless communication possible. The main challenge is maintaining the tradeoff between different performance parameters while maintaining the small size for the wideband topology of LNA. The different performance parameter is gain, noise, bandwidth linearity, power consumption and stability. Many researchers over last decade tried to work on these tradeoffs between these parameters and presented various designs and topologies for LNA with improved bandwidth [1–5]. A low noise amplifier (LNA) is a part of RF device that can amplify the low power signals without degrading the signal to noise ratio (S/N ratio). The primary function of LNA is to amplify the feeble signals that are just above the noise floor [3, 6].

$$NF_{\text{reciever}} = NF_{lna} + \frac{NF_{\text{rest}} - 1}{G_{lna}} \tag{1}$$

Here noise figure, gain and initial noise figure are given by *NF,* G and $NF_{\text{rest,}}$ respectively. The effect of noise can be greatly reduced from the subsequent stages of the receiver chain when LNA is used close to the signal source. This happens because of signal gain created by LNA.

The NF of LNA mostly dominates the NF of the receiver system. The noise figure of subsequent stages can be greatly reduced because of the high gain of LNA. The linearity property of LNA controls the sensitivity of the receiver. These are some of the reasons which promotes the designing of LNA. Equation (1) which is Friss's formula, can easily prove all the above-discussed points. A good LNA should have low NF and high gain with large intermodulation and compression points (IIP3 and P1dB). It also must have good impedance matching.

Figure 1 shows some of different types of distributed amplifier for the wide input matching characteristics [7–9]. All these distributed amplifiers have high power consumption, low gain and have a larger fabrication area. Chen et al. [10] designed a UWB LNA with the input network embedded in 3 section Chebyshev filter. This helped it to achieve good gain, linearity, low noise with good wideband performance. It has the large chip area. Luo et al. [11] designed a 0.2–2.5 GHz active resistive negative feedback LNA using cascade of two different topologies (Fig. 2).

Inductive degenerated common source topology (IDCS) and Common gate (CG) topology of LNA are the most widely used topologies for wideband applications. There are numerous pieces of literature where the design of LNA is based on IDCS [12, 13] and common gate [14] as core topologies. The CG configurations have better noise performance than IDCS configurations at higher frequencies, whereas IDCS performs better at lower frequencies. IDCS also has better gain performance than CG configuration because of better amplification due to input matching network [15]. The source impedance matching of IDCS is done by the method of series matching,

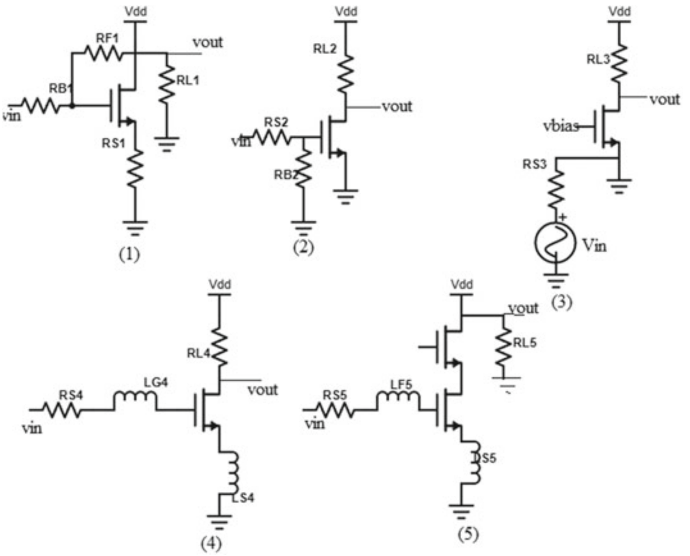


Fig. 1 Fundamental Topologies of LNA: (1) Shunt Series Feedback Common Source (2) Resistive Termination Common Source (3) Common Gate (4) Inductive Degeneration Common Source (5) Cascode Inductor. *Source* Degeneration

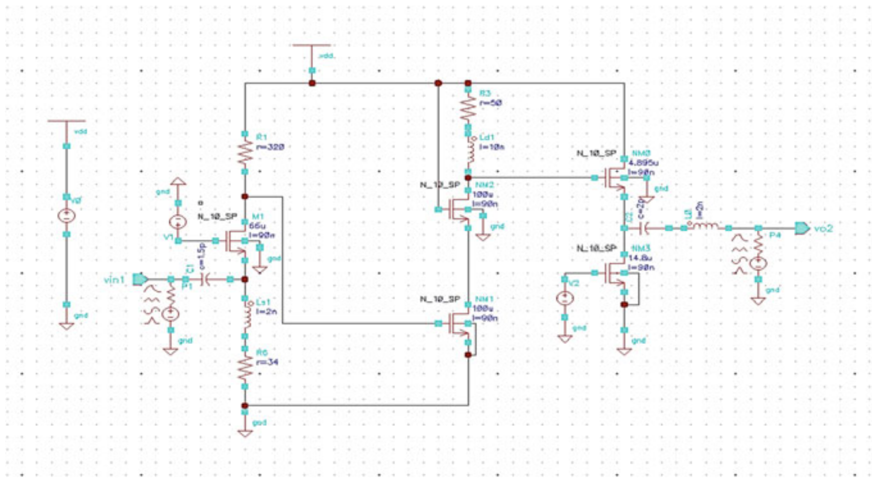


Fig. 2 Schematic diagram of the proposed LNA

whereas in the case pf CG, this is done by the parallel resonance method [15, 16]. Since the quality factor in case of series resonance is more than one, it increases the input transconductance of the transistors. On the other hand, the quality factor for CG topology is equal to one at its resonance. IDCS is more sensitive to the variation in different process and parasitic network, despite the higher value of the quality factor and better input matching network [17]. In case of the CG topology, most of the parasitic inputs get absorbed in the structure and thus reduce its effect on the circuit performance.

The organization of the paper is as follows. Section 2 discusses about the proposed low noise amplifier. The result and discussion of the simulated data is presented in Section 3. Finally at the end conclusion of the present study is discussed along with its comparison with the available literature.

2 Design of Proposed LNA

In this paper, the LNA is designed and proposed using the Cadence Virtuoso UMC 90 nm technology. The proposed LNA have three different stages of operation as shown in Fig. 1. The first stage is common gate stage (CG), which provides the input matching for wideband frequency range. The common source stage (CS) is the second stage which gives the good gain for the LNA. The third and last stage of the proposed LNA is common drain stage (CD) which is also known as source follower stage. This stage acts as a buffer and gives the matching with the output that is output reflection coefficient S22. The input matching filter consists of capacitor C1, inductor Ls1 and resistor R1, and it provides the maximum flow. Series LC filter is in the output matching filter, which mainly consist of inductor L0 and capacitor C2.

In the proposed LNA, CG amplifier has been used in the first stage instead of CS amplifier because the CG amplifier has the lower QF than that of CS amplifier [18–21]. This helps in achieving the wideband input matching. Another advantage of CG amplifier is that it requires fewer number of passive element than to CS amplifier to achieve same bandwidth. Another important parameter is gain of the amplifier. Since the gain of the CG amplifier is generally low and it is difficult to achieve the high gain with CG configuration. So, CS stage is used in the second stage of the proposed LNA, hence the overall gain can be increased.

3 Results and Discussion

The CMOS 90 nm technology has been used to analyses the proposed LNA design. The proposed LNA is designed by cascading the common gate and common source stage, since CG stage performs better at low frequencies and the performance of CS

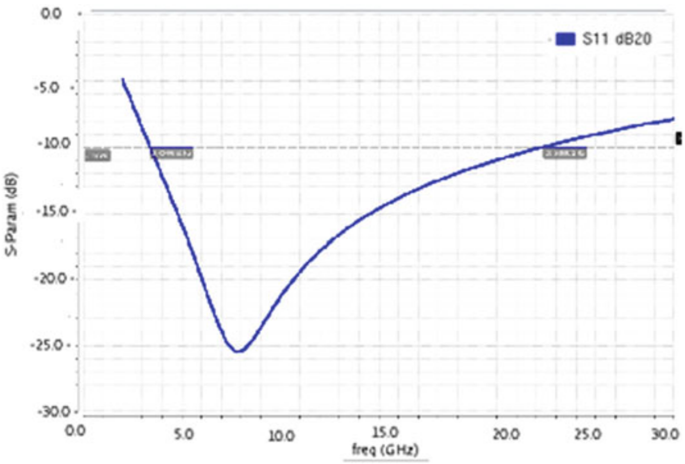


Fig. 3 Variation of S11 with frequency

is better at high frequencies. Therefore, the cascading of these stages with proper interstate matching, helps in acquiring the properties of both of stages.

The input reflection coefficient (S11) of the proposed designed LNA is shown in Fig. 3. It may be observed from the Fig. 3 that the value of S11 is well below −10 dB in the required frequency bandwidth of 3.40476–23.366422 GHz. S11 has the minimum value of −25.45 dB at frequency 7.78461 GHz.

The variation of output reflection coefficient (S22) at different values of frequencies is shown in Fig. 4. It may be observed from Fig. 4 that the value of output reflection coefficient is well below −10 dB in the frequency range of 2–9.16 GHz.

The gain performance of the designed low noise amplifier is shown in Fig. 5. It can be observed from the figure that the designed amplifier has the good gain in the frequency range of 2–9.65 GHz. It produces the maximum gain of 8.73 dB at 4.5 GHz.

The variation of reverse isolation (S12) of the proposed LNA is shown in Fig. 6. It may be observed that the value of S12 is below −85 dB in the discussed frequency band.

The variation of reverse isolation (S12) of the proposed LNA is shown in Fig. 6. It may be observed that the value of S12 is below −85 dB in the discussed frequency band.

Figure 7 shows the variation of the noise figure (NF) with frequency. The minimum value of the NF is 5.89 dB at 6.96 GHz. The NF varies from 5 to 10 dB in the frequency range of 2–10 GHz.

The variation of the stability factors K_f and B_{1f} with frequency is shown in Figs. 8 and 9, respectively. It can be observed from the figures that the $K_f > 1$ in the whole frequency domain, whereas $B_{1f} < 1$ in the frequency range of 4.3–30 GHz.

The properties and characteristics of linearity of the designed LNA is shown in Figs. 10 and 11. The 1 dB compression point (P1) of designed LNA is −27 dBm. The

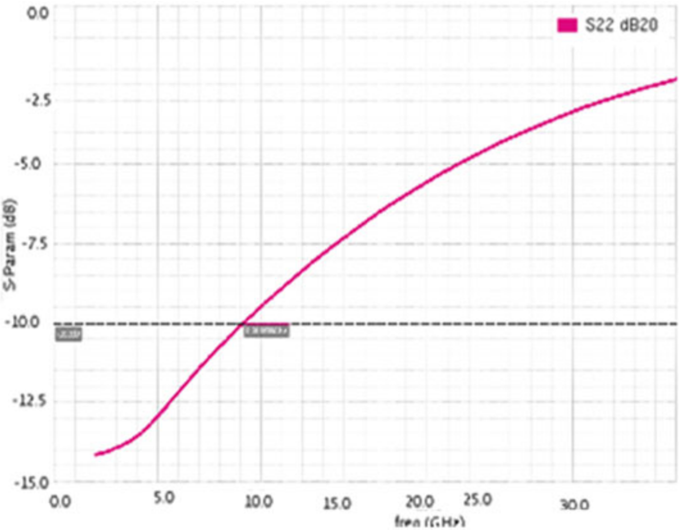


Fig. 4 Variation of output reflection coefficients (S22) with frequency

Fig. 5 Variation of gain (S21) with frequency

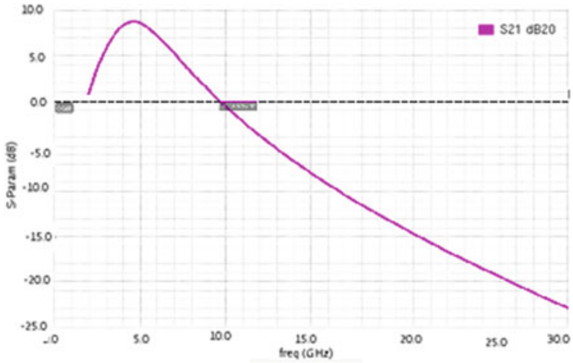
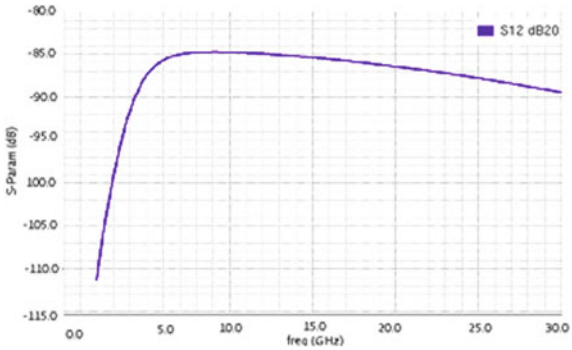


Fig. 6 Variation of reverse isolation (S12) with frequency



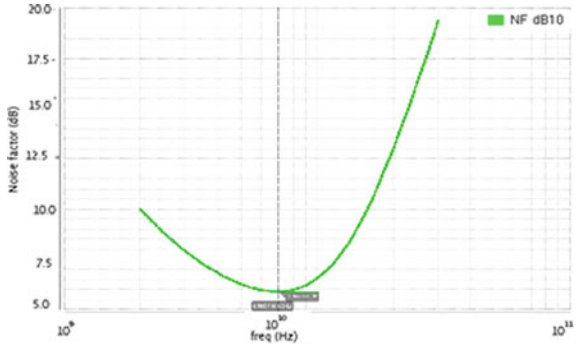


Fig. 7 Variation of noise figure with frequency

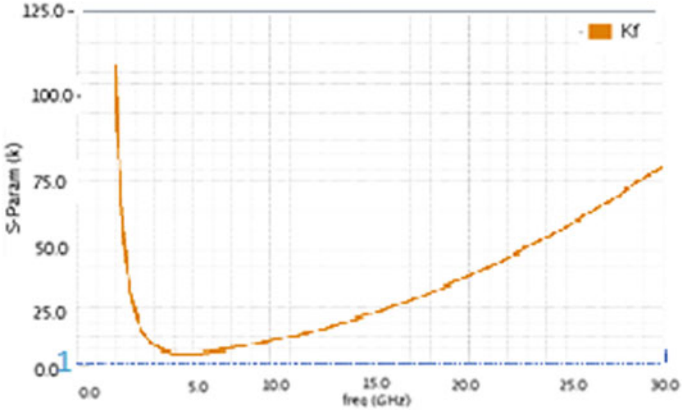


Fig. 8 Variation of stability factor (K_f) with frequency (GHz)

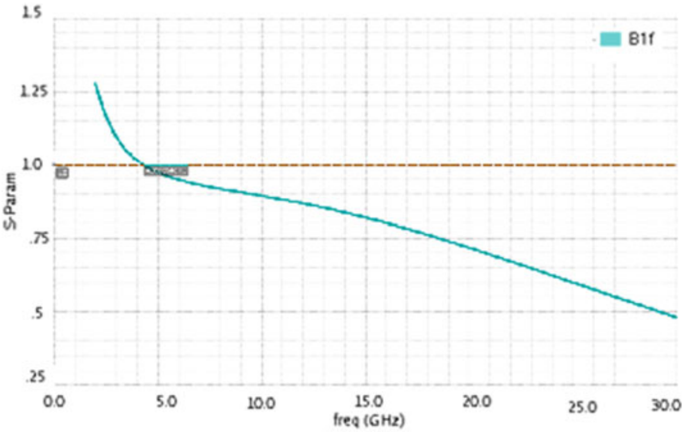


Fig. 9 The stabilty factor (B_{1f}) at different frequencies (GHz)

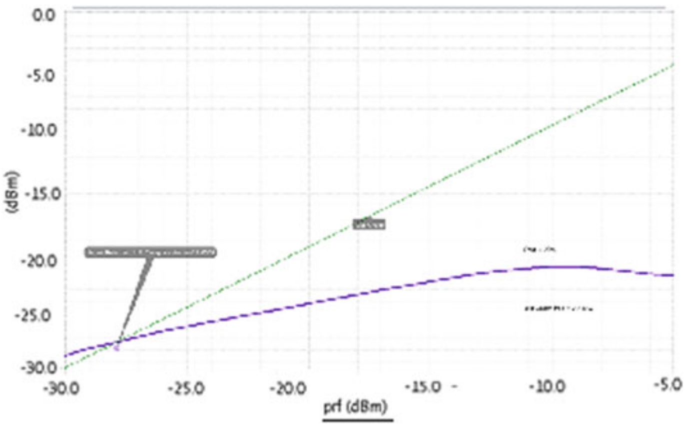


Fig. 10 1 dB compression point (P1 dB) with input power (dBm)

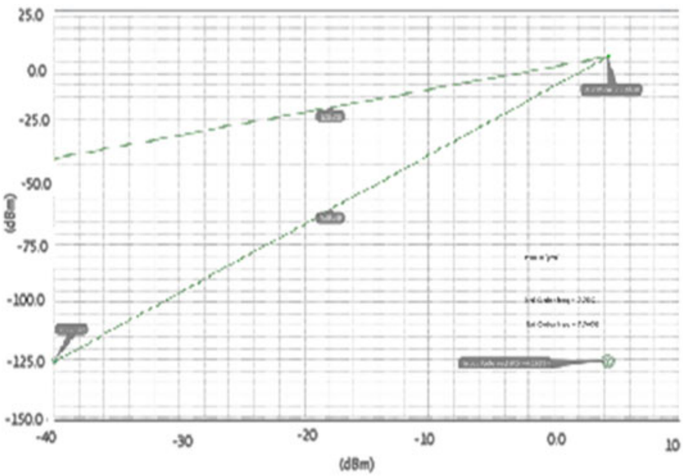


Fig. 11 Input third intercept point (IIP3) with input power (dBm)

input third intercept point (IIP3) is 4.3 dBm while keeping the 10 MHz frequency difference between 1st and 3rd order.

Finally, the present work has been compared with the existing literatures, which is shown in Table 1.

Table 1 Comparison of the present work with existing literatures

Parameter	[6]	[7]	[8]	[9]	[10]	[12]	[14]	This work
CMOS Technology	0.18 µm	0.18 µm	65 nm	90 nm	0.18 µm	0.18 µm	0.18 µm	**90 nm**
LNA Topology used	CG + CS	CG + CS	CG + CS	CG + CS	CG + CS	CG + CS + CD	CG + CS	**CG + CS + CD**
Power Supply (V)	1	1.8	1	1	1.8	1.5	1.5	**1**
Bandwidth (GHz)	8.5–20	1–5	7.6–29	3.1–10.6	0.4–10	3.1–10.6	3.1–10.3	**2–30**
Gain (dB)	11.13	21–25	10.7	17.25 ± 1.25	11.2–12.4	Avg.9.7	9.6–12.71	**Max.8.723**
N.F (dB)	2.1–3.2	2.6–3	4.5–5.6	4.1–9.4	4.4–6.5	5.27–7	2.5–3.9	**19.3–5.6**
S11 (dB)	<−9.44	<−7		<−12	<−10	<−13.5	<−9	**<−15**
S12 (dB)	<−60	–	–	–	–	<−43	<−45	**<−85**
Area (mm^2)	0.116	–	0.3	0.13	0.42	1.0296	0.68	**UC**
IIP3 (dBm)	0.96	–	–	–	−6	−2.23	−3 to 1	**4.30**
1-dB compression point (dBm)	–	–	–	–	–	−10	−12.5	**−27.85**
Power Consumption (mW)	5.4	14.5	12.1	8.5	12	4.5	13.4	**UC**
FOM	15.1–28.4	2.759	4.159	2.255	1.91	2.63	1.873	–

Bold indicates nm = nanometer, Max. = maximum, UC = uncalculated, CG = common gate, CS = common sourse, CD = common drain

4 Conclusion

In this paper, a UWB band LNA has been proposed using the Cadence UMC 90 nm technology. The LNA has been designed using CG, CS and CD cascaded one after the other, each stage fulfilling its characteristic properties. The input common gate stage gives good input match with the antenna, i.e., S11 below −20 dB and the common drain provides the output match, i.e., S22 below −10 dB. The common source provides the gain of 8.7299 dB. The reverse isolation is below −90 dB and the range of NF is from 5 to 10 dB in the discussed frequency range. The 1 dB compression point of the designed LNA is −27 dBm and the input third intercept point (IIP3) of LNA is 4.3 dBm at 7.747 GHz.

References

1. Ballweber BM, Gupta R, Allstot DJ (2000) A fully integrated 0.5–5.5-GHz CMOS distributed amplifier. IEEE J Solid-State Circuits 35:231–239
2. Idris MI, Yusop N, Chachuli SAM, Ismail MM (2014) Design and analysis of low noise amplifier using cadence. J Theor Appl Inf Technol 69(1):151–160
3. Friis HT (1944) Noise figures of radio receivers. Proc IRE 32(7):419–422
4. Sreekumar R, Nasrollahpour M, Hamedi-Hagh S (2017) Cascode stage based LNA for bluetooth applications in 45 nm CMOS technology. In: 2017 new generation of CAS (NGCAS). IEEE, pp 145–148
5. Wang J-J, Chen D-Y (2016) LNA with wide range of gain control and wideband interference rejection. Int J Electron 103(10):1748–1758
6. Pandey S, Gawande T, Inge S, Pathak A, Kondekar PN (2018) Design and analysis of wideband low-power LNA for improved RF performance with compact chip area. IET Microwaves Antennas Propag 12(11):1816–1820
7. Ghadimipoor F, Garakani HG (2011) A noise-canceling CMOS low-noise amplifier for WiMAX. In: 2011 international conference on electronic devices, systems and applications (ICEDSA). IEEE, pp 165–169
8. Qin P, Xue Q (2017) Design of wideband LNA employing cascaded complimentary common gate and common source stages. IEEE Microwave Wirel Compon Lett 27(6):587–589
9. Axholt A, Ahmad W, Sjoland H (2008) A 90nm CMOS UWB LNA. In: 2008 NORCHIP, IEEE, pp 25–28
10. Chen K-H, Lu J-H, Chen B-J, Liu S-I (2007) An ultra-wide-band 0.4–10-GHz LNA in 0.18um CMOS. IEEE Trans Circuits Systems II: Express Briefs 54(3):217–221
11. Luo L, Li Z, Cheng G, He X, He B (2017) A 0.2–2.5 GHz resistive feedback LNA with current reuse transconductance boosting technique in 0.18-μm CMOS. In: 2017 IEEE 15th student conference on research and development (SCOReD). IEEE, pp 424–427
12. Weng R-M, Liu C-Y, Lin P-C (2010) A low-power full-band low-noise amplifier for ultra-wideband receivers. IEEE Trans Microw Theory Tech 58(8):2077–2083
13. Toteva I, Andonova A, Simulation of LNA in 0.18 μm CMOS technology
14. Lo Y-T, Kiang J-F (2011) Design of wideband LNAs using parallel-to-series resonant matching network between common-gate and common-source stages. IEEE Trans Microw Theory Tech 59(9):2285–2294
15. Heydari P (2007) Design and analysis of performance-optimized CMOS UWB distributed LNA. IEEE J Solid-State Circuits 42(9):1892–1905
16. Zhang F, Kinget PR (2006) Low-power programmable gain CMOS distributed LNA. IEEE J Solid State Circuits 41(6):1333–1343

17. Liu R, Lin C, Deng K et al (2003) A 0.5–14 GHz 10.6-dB CMOS cascade distributed amplifier. In: VLSI circuits Symp, vol. 17. Digest, Kyoto, Japan, June 2003, pp 139–140
18. Bevilacqua A, Niknejad AM (2004) An ultra-wide band CMOS low-noise amplifier 3.1–10.6-GHz wireless receivers. IEEE J Solid-State Circuit 39(12):2258–2268
19. Kim CW, Kang MS, Anh PT et al (2005) An ultra-wideband CMOS low noise amplifier for 3–5 GHz UWB system. IEEE J Solid State Circuits 40(2):544–547
20. Li Z, Wang Z, Zhang M et al., A 2.4 GHz ultra-low-power current-reuse CG-LNA with active gm-boosting technique. IEEE Microw Wirel Compon
21. Lin YS, Chen CZ, Yang HY et al (2010) Analysis and design of a CMOS UWB LNA with dual-RLC-branch wideband input matching network. IEEE Trans Microw Theory Tech 57(2):287–296

Graphene FET and MESFET for Opto-Electronic Applications

Jaya V. Gaitonde and Rajesh B. Lohani

Abstract Graphene and other 2-D materials are emerging alternatives to the existing semiconductors for electronic and opto-electronic applications. In this paper, we present a brief review of the graphene-based photodetectors to enlighten the basic knowledge on graphene-based photodetector and the scope for opto-electronic applications. Further, we provide a comparative analysis of the well-known graphene Field Effect Transistor (FET) and the novel graphene-based Metal–Semiconductor Field Effect Transistor (MESFET) over a discrete set of dimensions and carrier mobilities for photodetector, optically-controlled amplifier, and terahertz modulator applications. The simulations have been carried out in MATLAB Software using a theoretical drift–diffusion model developed elsewhere under dark and optical illumination. The simulations reveal that graphene FET exhibits a larger photo and almost the same amplification responses at lower doping levels at mid-infrared wavelengths, whereas the graphene MESFET shows superior photovoltaic and terahertz modulation responses at higher doping levels. The responsivities span between two to six orders of magnitude, whereas RC-limited bandwidths range from tens of gigahertz to terahertz frequencies. The amplification bandwidths are in the sixties of gigahertz to terahertz range. The devices will greatly contribute toward opto-electronic applications.

Keywords Graphene · Review · 2-D · Photodetector · Opto-electronic · Amplifier · Modulator · Bandwidth

J. V. Gaitonde (✉)
Electronics and Telecommunication Department, Goa Engineering College (Goa University), Farmagudi, Ponda, Goa 403401, India
e-mail: jaya.gaitonde@gec.ac.in

R. B. Lohani
Electronics and Telecommunication Department, Goa Engineering College (Govt. of Goa), Farmagudi, Ponda, Goa 403401, India
e-mail: rblohani@gec.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_28

1 Introduction

Graphene has been an excellent topic of interest in the recent years. It has outstanding properties such as good mechanical resilience, high optical transparency, excellent physical/chemical steadiness, low sheet resistance, and bias-dependent tunability of the Fermi level [1]. But in the intrinsic form, the lack of bandgap hinders its application to various fields. The tailoring of the bandgap can be achieved either by chemical modification of graphene, using bilayer graphene under applied bias, or coercing graphene in single dimensions developing quantum dots or graphene nanoribbons [2, 3]. Graphene can act either as a transparent electrode or semiconductor material for photodetectors. Graphene is composed of a honey-comb lattice of carbon atoms. The band structure exhibits linear dispersion with intrinsic graphene showing zero bandgap at the Dirac point and linearly dispersing conduction and valence bands away from the Dirac point. Due to its excellent properties and its robustness, it is also useful in harsh environments. Graphene is a 2-D material but can be made 3-D by stacking more layers of graphene.

With graphene as a transparent gate electrode, the Schottky junctions in GaN and GaAs Optical Field Effect Transistors (OPFETs) show significant photoresponses [4, 5].

Graphene forms a high speed photodetector (PD) due to its ultrafast dynamics such as short recombination and relaxation times (picoseconds to nanoseconds), ultrahigh carrier mobility (~10,000 $cm^2/V\,s$) and high Fermi velocity (~10^6 m/s). The graphene PDs also show broadband absorption from Ultraviolet (UV) to Terahertz (THz) wavelengths.

Although significant research has been administered in standalone graphene-based transistor photodetectors [6–11] in the UV, visible, terahertz and infrared ranges, the responsivity is limited to 980 A/W. This limits its use in practical high gain applications, although the communication bandwidths achievable using graphene Field Effect Transistors (FETs) span from tens of gigahertz to far gigahertz ranges. To overcome this issue, the authors first, investigated the graphene FET photoconductive characteristics at the optimum wavelength of operation, i.e., the wavelength at which there is resonance effect induced by the optical phonon emission (6.3 μm) [12]. This ensured that the maximum photogeneration was obtained to give maximum output. Secondly, the device was biased close to the Dirac point to minimize the dark current and prevent quenching of inter-band absorption, which led to a significant enhancement of photocurrent. Further, the device characteristics could be modulated by varying the device dimensions and considering different mobility samples. However, a trade-off is expected between the gain, bandwidth, unity-gain cut-off frequency, and dark current. This paper, adopts structural and mobility-based optimization to enhance the device gain-bandwidth performance without creating excessive dark current.

Recently, the authors have reported the characteristics of novel graphene-based Metal–Semiconductor Field Effect Transistor (MESFET) (Fig. 1) for use as a photodetector and terahertz modulator along with comparative analysis with

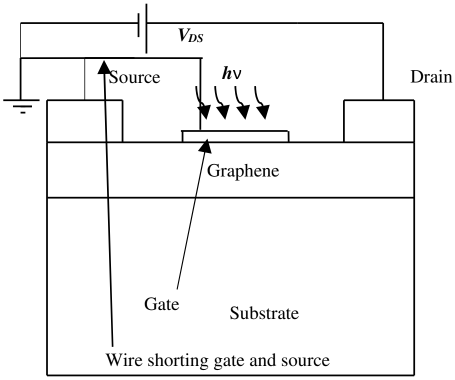


Fig. 1 Schematic sketch of graphene MESFET

graphene FET showing superior performance of graphene FET as a photodetector, whereas the graphene MESFET acting as a good terahertz modulator [13]. The structural and mobility-based optimization specified in the previous paragraph is repeated for graphene MESFET. Graphene MESFET creates greater potential and conductivity gradients between gate-contacted and free space graphene than graphene FET. This is because the dipole capacitance between gate and graphene in the MESFET structure is higher than the oxide capacitance in the FET configuration [13]. Thus, it is expected that the graphene MESFET would produce larger photovoltaic current than graphene FET because it relies on the strength of the built-in electric field at the junction of gate-covered and uncovered graphene when the illumination is allowed to fall within the junction and other factors kept constant. This fact led to the investigation of the factor by which the photovoltaic current in MESFET differs from that in the FET which is the second contribution to the proposed research.

2 Review

A. *Detection mechanisms in graphene photodetectors*

Photovoltaic effect, photoconductive effect, photo-induced thermoelectric effect, field effect doping, and bolometric effect are the various sensing mechanisms in graphene-based photodetectors.

1. *Photovoltaic effect*

The photovoltaic effect results from the generation of photocurrents in the vicinity of metal contact with graphene or in the junctions created by p- and n-regions of graphene [14, 15]. These photocurrents are produced from the generation of excitons in graphene when exposed to light, and the separation of electrons and holes by the built-in electric fields near the junctions.

2. *Photoconductive effect*

This effect contributes due to the inter-band absorption of the photons with energy greater than twice that of the Fermi level resulting in the generation of electron–hole pairs in the neutral region of graphene rather than in junctions [12]. The generated carriers produce a photocurrent depending upon the bias applied to the source and drain contacts of graphene FETs.

3. *Photothermoelectric effect*

The photothermoelectric effect emanates from the induction of temperature difference when exposed to light, eventually leading to a thermoelectric voltage [16, 17]. Hot carriers are generated upon illumination to form a hot Fermion distribution on account of the slow relaxation of electrons and lattice in graphene and the thermal decoupling of photo-carriers from the crystal lattice. The photothermoelectric voltage occurs in the junction of regions with different Seebeck coefficients [17, 18].

4. *Field effect Doping*

This effect occurs in gated photodetectors such as graphene FETs. The gate voltage-induced variation of the channel current and the trapping of charges at certain interfaces due to this variation is called as field effect doping. Under light exposure, the trapping of some of the photo-carriers in the channel or at interfaces will lead to electrostatic production of free carriers in graphene. There is a horizontal shift of the transfer curve on account of the change of the effective gate voltage [19].

5. *Bolometric effect*

This effect occurs through thermally induced heating of the active element under illumination, which alters the resistance of the active element depending upon the temperature [20].

B. *The State-of-Art*

Photodetectors based on traditional semiconductors such as Si and InGaAs suffer from low transmittance, non-resilience, limited spectrum, low resolution, and CMOS-inconsistency. The excellent optical and electronic properties, mechanical resilience, and wafer-scale development and consolidation of 2-D materials viz. Transition metal dichalcogenides (TMDs), graphene, and their unified systems can overcome the above limitations. 2-D heterostructure photodiodes provide ultrafast and broad visible to infrared response. Ultrasensitiveness and broadband photodetection can be registered using 2-D hybrid systems-based phototransistors when integrated with other systems such as perovskites, quantum dots, plasmonic nanostructures, or organic materials. Further, integration of 2-D PDs with Si and CMOS-based systems can lead to high performance, broadband detection, low cost, and imaging modalities [21].

Photodetectors can be classified as photodiodes, photoconductors, and phototransistors. Photodiode either administers the built-in electric field via a *p–n* junction or a metal–semiconductor Schottky junction. Ultrahigh mobility of graphene leads to ultrafast response from picoseconds to nanoseconds, and high bandwidth in the GHz range [7] in graphene photodiodes. However, the External Quantum Efficiency

(*EQE*) of photodiodes is below 100% except in Avalanche photodiodes (APDs), wherein impact ionization occurs near breakdown resulting in multiplication gain, but at the expense of high voltages of 50–100 V. 2-D materials viz. MoS_2, InSe, and BP-based APDs offer multiplication gains of 100–1000 [22]. Photoconductors provide gain by the movement of photoexcited electrons or holes several times across the source and drain ohmic electrodes before their recombination when an external field is applied across the electrodes with the gain defined by the ratio of carrier lifetime to the transit time. While the prolonged lifetime of one type of carriers due to the trapped sites results in high gain, this conversely degrades the temporal and frequency response.

There exist noise-related issues in photodiodes and photoconductors. To address these issues as well as provide high gain and simultaneous large bandwidth, phototransistors are employed. As an extension of the photoconductor, the phototransistor consists of a third gate terminal, wherein a thin dielectric electrically isolates the gate from the channel in 2-D material-based photodetectors [23]. The charge density can be controlled and varied by the gate bias through the field effect and the dark current can be restrained using depletion mode operation.

The 2-D photodetectors without gain can be classified into 2-D PDs with photo-induced bolometric and thermoelectric effects and 2-D photodiodes in heterostructure configuration based on photovoltaic effect. Carrier transport in graphene being dominated by hot carriers and the electron–electron interactions being strong, the photothermoelectric effect is exhibited by p-n junctions in graphene [24]. Certain PDs in metal-graphene-metal (MGM) configuration reported dominance by photovoltaic effect [25]. Graphene photothermoelectric detectors showed high bandwidth of 40 GHz, but limited responsivity of 6.1 mA/W on account of the low absorption and non-existence of gain [7]. Graphene with its feeble phonon-electron coupling and small electronic heat capacity when incorporated in bolometers as bilayer graphene exhibited high sensitivity (33 $fWHz^{-1/2}$) at 5 K and (>1 GHz) very fast speeds at 10 K [26]. 2-D heterojunction photodiodes are grown either by locally doping through chemical modifications or electrostatic interactions or by using transfer processes in both planar and orthogonal directions. These devices are characterized by ultrafast response but without gain. The sandwich structures based on Gr-TMD-Gr out-of-plane configurations were investigated for use as a photodetector. Graphene served as transparent electrode in the top and bottom positions for extracting charges, whereas the TMD absorbed photons and assisted in the transport of carriers. The advantages of the materials utilized were the capability of TMDs of strong interactions between light and matter and the possibility of tuning the Fermi level of graphene based on the bias applied. For example, Gr-Ws_2-Gr structure showed good performance with sensitivity of 0.1 A/W and *EQE* above 30% [27].

Since the earlier type of PDs showed limited responsivities and gains, 2-D hybrid PDs were demonstrated. The photoresponse depends on the photoconductive gain from the difference between the majority carrier electron transit time and the minority carrier whole lifetime. For further enhancement of photoresponse, techniques such as surface doping, sensitizations with quantum dots (QDs), perovskites or metal nanostructures were employed. The cost of photoconductive gain is the 1 kHz order or lower bandwidth. These types of PDs suffice for steady state spectroscopy, sensing, and video imaging applications. The photogating effect in graphene detectors can be enhanced using the combination of optical waveguides, optical microcavities, and the plasmonic field proliferation. Other methods are sensitizing graphene hybrid phototransistors with colloidal quantum dots (CQDs), other 2-D materials, organic materials, perovskites, etc.

Graphene FETs (GFETs) act as excellent transducers on account of the very sensitive channel currents to variations in potential and charge densities at graphene channel or interfaces. This feature is enabled by the shift in the Dirac point depending upon the gate and graphene work function difference, charge density and type at the graphene-substrate interface, and the level of impurity doping [2]. The ultrahigh mobility of graphene results in high speed operation of GFET photodetectors, which outperform Si MOSFETs and exhibit comparable performance to the InP or GaAs HEMTs in terms of cut-off frequency.

3 Results and Discussion

We simulated the Graphene FET (GFET) and Graphene MESFET (GMESFET) devices with drift–diffusion approach and under optical stimulation [12, 28] using MATLAB software. The gold (Au) metal gate is considered to be optically transparent. The gate voltage applied is 0.15 V. The drain bias is varied between 0 and 1 V. The optical wavelength under consideration is 6.3 μm. The device is studied under radiation field of 1 kV/m and at room temperature of 300 K. The parameters used in calculation are presented in Table 1.

The photoconductive characteristics have been evaluated for both the devices as shown in Fig. 2 along with the dark current for the dimensions of gate length (1 μm), gate width (1 μm), and carrier mobility (0.5 $m^2/(V\ s)$). The graphs clearly contrast

Table 1 Parameters used in calculation

Para-meter	Name	Value	Unit
μ	Low field electron mobility	(0.5–1.5)	(m^2/V s)
Φ_m	Gate work function (Au)	(5.54)	(eV)
Φ_g	Graphene work function	(4.5)	(eV)
ε_{ox}	Oxide dielectric constant	(4)	–
L	Gate length	(0.1–1)	(μm)
Z	Gate width	(1–100)	(μm)
d_{ox}	Oxide thickness	0.3	(μm)
v_F	Fermi velocity	1×10^6	(m/s)
C_{it}	Interface trap capacitance	1×10^{-3}	(F/m^2)
d_{eq}	Equilibrium separation distance between Au and graphene	0.331	(nm)

Table 2 Structural and mobility optimization results for GMESFET

Device dimensions and mobility	Bandwidth (Hz)	f_T (Hz)	Responsivity (A/W)	Photovoltaic current (A)	Dark current (A)
$W = 1\ \mu$m, $L = 1\ \mu$m, $\mu = 0.5$ m^2/(V s)	319.3 GHz	72.4 GHz	654.3 A/W	0.11 nA	1.83 μA
$W = 10\ \mu$m, $L = 0.5\ \mu$m, $\mu = 0.5$ m^2/(V s)	63.86 GHz	289.6 GHz	2.62×10^3 A/W	2.17 nA	36.67 μA
$W = 1\ \mu$m, $L = 0.5\ \mu$m, $\mu = 0.5$ m^2/(V s)	**638.6 GHz**	**289.6 GHz**	**2.62×10^3 A/W**	**0.217 nA**	**3.67 μA**
$W = 1\ \mu$m, $L = 1\ \mu$m, $\mu = 1$ m^2/(V s)	319.3 GHz	144.8 GHz	1.31×10^3 A/W	78.31 pA	3.67 μA
$W = 100\ \mu$m, $L = 0.5\ \mu$m, $\mu = 1$ m^2/(V s)	6.386 GHz	579.3 GHz	5.13×10^3 A/W	15.6 nA	0.73 mA
$W = 40\ \mu$m, $L = 0.1\ \mu$m, $\mu = 1$ m^2/(V s)	79.83 GHz	14.5 THz	1.31×10^5 A/W	31.3 nA	1.46 mA
$W = 40\ \mu$m, $L = 0.1\ \mu$m, $\mu = 1.5$ m^2/(V s)	79.83 GHz	21.7 THz	1.96×10^5 A/W	11.1 nA	2.2 mA
$W = 60\ \mu$m, $L = 0.5\ \mu$m, $\mu = 1.5$ m^2/(V s)	10.64 GHz	868.9 GHz	7.85×10^3 A/W	3.32 nA	0.66 mA
$W = 60\ \mu$m, $L = 0.1\ \mu$m, $\mu = 1.5$ m^2/(V s)	53.2 GHz	21.7 THz	1.96×10^5 A/W	16.6 nA	3.3 mA

Significance of bold values are optimum results

Table 3 Structural and mobility optimization results for GFET

Device dimensions and mobility	Bandwidth (Hz)	f_T (Hz)	Responsivity (A/W)	Photovoltaic current (A)	Dark current (A)
$W = 1$ µm, $L = 1$ µm, $\mu = 0.5$ m^2/(V s)	27.2 THz	72.3 GHz	2.9×10^4 A/W	0.11 nA	44.65 µA
$W = 10$ µm, $L = 0.5$ µm, $\mu = 0.5$ m^2/(V s)	5.45 THz	289.3 GHz	1.16×10^5 A/W	2.22 nA	0.89 mA
$W = 1$ µm, $L = 0.5$ µm, $\mu = 0.5$ m^2/(V s)	54.5 THz	289.3 GHz	1.16×10^5 A/W	0.22 nA	89.3 µA
$W = 1$ µm, $L = 1$ µm, $\mu = 1$ m^2/(V s)	27.2 THz	144.6 GHz	5.8×10^4 A/W	84.8 pA	89.3 µA
$W = 100$ µm, $L = 0.5$ µm, $\mu = 1$ m^2/(V.s)	545.1 GHz	578.5 GHz	2.32×10^5 A/W	16.9 nA	17.86 mA
$W = 40$ µm, $L = 0.1$ µm, $\mu = 1$ m^2/(V s)	6.81 THz	14.46 THz	5.8×10^6 A/W	33.9 nA	37.7 mA
$W = 40$ µm, $L = 0.1$ µm, $\mu = 1.5$ m^2/(V s)	6.81 THz	21.7 THz	8.7×10^6 A/W	15.3 nA	53.6 mA
$W = 60$ µm, $L = 0.5$ µm, $\mu = 1.5$ m^2/(V s)	**908.5 GHz**	**867.8 GHz**	**3.48×10^5 A/W**	**4.6 nA**	**16.07 mA**
$W = 60$ µm, $L = 0.1$ µm, $\mu = 1.5$ m^2/(V s)	4.54 THz	21.7 THz	8.7×10^6 A/W	22.97 nA	80.37 mA

Significance of bold values are optimum results

the performance of both the devices. The dark current is in the lower regime for GMESFET as compared to that for GFET. Note that the devices are biased close to the Dirac point (0.133 V) for suppression of dark current. Hence, the Fermi levels and carrier densities are maintained relatively at the lowest level in both cases. Further, the dipole capacitance in GMESFET is much higher than that of the oxide capacitance in GFET, which implies that the diffusion component of current dominates the drift component which suppresses the dark current and vice versa for GFET.

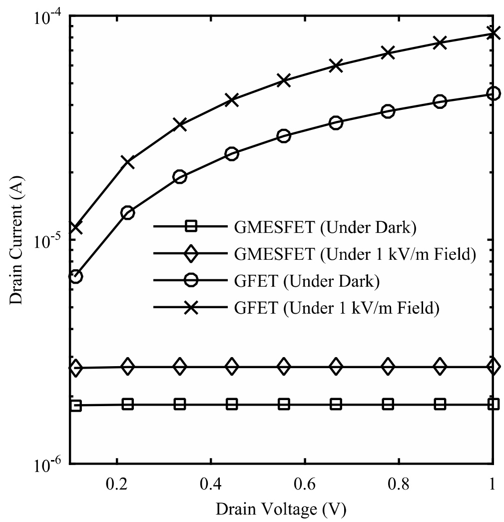


Fig. 2 I-V Characteristics of the GFET and GMESFET devices under dark and illumination

Although the devices are biased close to the Dirac point, the large difference in the said capacitances leads to higher dark electron density in the GMESFET. Thus, under illumination the inter-band absorption is higher in GFET resulting in more photogenerated carriers producing larger photoconductive current. The graphs also depict that high sensitivities are attained in both devices. Another reason for this is that the operating wavelength of 6.3 μm corresponds to the optical phonon emission resonance, at which the photon and phonon absorption and emission processes are exactly balanced at higher quasi Fermi levels.

Further, we calculated the RC-limited photoconductive bandwidth, amplification bandwidth (f_T), photoconductive responsivity, photovoltaic current, and the dark current for both devices over a discrete set of dimensions and mobilities as depicted in Tables 2 and 3 at a drain voltage of 1 V. The RC-limited bandwidth is calculated at a matched load resistance (R_L) of 50 ohms and is given by $(1/(2\pi R_L C_G))$, where C_G is the gate capacitance. Note that the gate capacitance depends upon the quantum capacitance, dipole/oxide capacitance, and the interface trap capacitance. In the present work, the interface trap capacitance is maintained constant. The quantum capacitance in both cases exhibit similar values due to the compensation effects of higher dark carrier density and lower photoconductive effect in GMESFET with the lower dark carrier density and higher photoconductive effect in GFET. Thus, the gate capacitance value is decided by the dipole/oxide capacitance which is larger in the GMESFET thus increasing the gate capacitance. Thus, for similar dimensions and mobilities, the bandwidth is larger in GFET.

The unity-gain cut-off frequency (f_T) is maintained at the same level in both devices since under the present operating conditions, the channel capacitance which gives the transconductance cancels out the effect of gate capacitance due to dominance of the quantum capacitance over interface trap capacitance. This implies that the f_T value is dependent upon only the dimensions, mobility, and the drain voltage which causes same values for both devices.

The responsivities are comparatively larger in the GFET devices due to larger photoconductive currents. The photovoltaic currents are comparable due to the compensation effects as discussed earlier for quantum capacitance which create similar built-in electric fields. The dark currents are much smaller in the case of GMESFET which is the major advantage of the device over GFET when biased close to the Dirac point.

Apart from the above observations, the other observations are that the optimum dimensions and mobility for GMESFET as detector-cum-amplifier with low dark currents are $W = 1$ µm, $L = 0.5$ µm, $\mu = 0.5$ m^2/(V s) wherein it exhibits a detection-cum-amplification bandwidth of 289.6 GHz, a responsivity of 2.62×10^3 A/W, and a dark current of 3.67 µA. On the other hand, the GFET shows optimum bandwidth of 867.8 GHz, a responsivity of 3.48×10^5 A/W, and a dark current of 16.07 mA at the dimensions and mobility of $W = 60$ µm, $L = 0.5$ µm, $\mu = 1.5$ m^2/(V s).

To indicate the terahertz modulation-cum-detection capability of the two devices, the devices were studied at an operating wavelength of 80 µm and at a gate bias of 0.6 V. At this wavelength, the major contribution to optical conductivity in GMESFET is from the intra-band absorption when the gate bias is high so that a large number of free carriers are available for excitation to higher energy bands and the inter-band absorption is quenched. In GFET, the inter-band conductivity is significantly degraded due to comparatively higher electrostatic doping, which in turn, increases the intra-band conductivity but not to match that of the GMESFET. Although the amount of photogenerated carriers generated are lesser at this wavelength, a significant photoconductive responsivity is still registered in GFET. Table 4 depicts the photovoltaic responsivities (as a function of dimensions in both devices) and the

Table 4 Photovoltaic and Intra-band Performance Comparison of GFET and GMESFET at a gate bias of 0.6 V and a wavelength of 80 µm

Device dimensions and mobility	GFET		GMESFET	
	Photovoltaic current (A)	Responsivity (A/W)	Photovoltaic current (A)	Responsivity (A/W)
$W = 1$ µm, $L = 0.5$ µm, $\mu = 0.5$ m^2/(V s)	0.22 nA	0.33 A/W	0.474 nA	0.357 A/W
$W = 100$ µm, $L = 0.5$ µm, $\mu = 1$ m^2/(V s)	23.5 nA	**0.35 A/W**	0.3 µA	**4.6 A/W**
$W = 60$ µm, $L = 0.5$ µm, $\mu = 1.5$ m^2/(V.s)	12.56 nA	**0.315 A/W**	0.326 µA	**8.19 A/W**
	Intra-band Conductivity (S)		**Intra-band Conductivity (S)**	
	7.86 µS		**146.5 µS**	

Significance of bold values are optimum results

Table 5 Comparative study of the State-of-Art GFET PDs with the present work

Device	Wavelength (m)	Responsivity (A/W)	Bandwidth (Hz)/Time constant (s)	Ref
Mechanically exfoliated FLG or SLG	1.55 μm	~0.5 mA/W	40 GHz	[6]
Mechanically exfoliated BLG	300 nm-6 μm	~6.1 mA/W	16 GHz	[7]
Mechanically exfoliated BLG or SLG	1.3–17 μm	0.05 A/W	~18 GHz	[8]
3D Graphene FETs	UV–Visible	> 1 A/W	> 1 MHz	[9]
GFET	532 nm	980 A/W	–	[10]
GFET	(0.1–10) THz	–	< 1 ns	[11]
GFET	**6.3 μm**	**3.48 × 10^5 A/W**	**908.5 GHz/0.175 ps**	**This Work**

Significance of bold values are optimum results

intra-band conductivities. It can be observed that the responsivities, in overall, are more than one order magnitude higher in the GMESFET for the same set of dimensions and mobility. This show the higher strength of the built-in electric field created at the junction of gate-covered and uncovered graphene by the high gate voltage of 0.6 V. This all comes at the expense of much higher dark current (~750 mA). The significant boosting of the dark current is attributed to the higher dark electron density and the enhancement of the Fermi level. The intra-band conductivities are much higher in GMESFET as opposed to GFET as stated earlier and can be tuned to the desired level by the variation of gate voltage as reported in [12]. Thus, the devices can be switched from photoconductive detection-cum-amplification mode at low bias levels and mid-infrared wavelengths to modulation-cum-photovoltaic detection mode at high biases and terahertz wavelengths.

Table 5 presents the performance comparison of the State-of-Art standalone graphene FET photodetectors with the present work. Although this is a theoretical work and most of the reported works are based on experiments, the much higher gain and bandwidth of the presently investigated GFET in the theoretical limit as compared to the reported work means that our device can still surpass the state-of-art photodetectors.

4 Conclusion

We presented a brief review of graphene as a photodetector with consideration of the detection mechanisms of graphene PDs, and the state-of-art. This will enlighten its scope for opto-electronic applications. Further, the GFET and GMESFET devices were modeled using drift–diffusion approach under dark and optical illumination. The GFET emerges as a high performance photoconductive detector-cum-amplifier when biased close to the Dirac point at mod-infrared wavelengths. The GMESFET also showed significant sensitivity. The structural and mobility optimization of the two devices were carried out to feature the optimum dimensions and mobility keeping the dark current to low values. At high bias levels, the GMESFET showed superior terahertz modulation and photovoltaic detection capability but at the cost of high dark current. The analysis was carried out based on the Fermi level modulation with gate bias, the various capacitances involved, the dark current, the photogenerated dark carrier density, and the built-in electric field. The devices show good prospects toward opto-electronic applications.

References

1. Xie C, Zhang X, Ruan K, Shao Z, Dhaliwal SS, Wang L, Zhang Q, Zhang X, Jie J (2013) High-efficiency, air stable graphene/Si micro-hole array Schottky junction solar cells. J Mater Chem A 1:15348–15354
2. Knap W, Kachorovskii V, Deng Y, Rumyantsev S, Lü J-Q, Gaska R, Shur MS et al (2002) Nonresonant detection of terahertz radiation in field effect transistors. J Appl Phys 91:9346–9353
3. Tauk R, Teppe F, Boubanga S, Coquillat D, Knap W, Meziani YM, Gallon C et al. (2006) Plasma wave detection of terahertz radiation by silicon field effects transistors: responsivity and noise equivalent power. Appl Phys Lett 89:253511(1–3)
4. Gaitonde JV, Lohani RB (2016) UV photodetector based on graphene-GaN Schottky junction in MESFET. In: 2016 conference on emerging devices and smart systems (ICEDSS). IEEE, pp 30–33
5. Gaitonde JV, Lohani RB (2019) Visible range characterization of au/Graphene-GaAs Schottky junctions in MESFET. IETE J Res 1–11
6. Xia F, Mueller T, Lin Y, Valdes-Garcia A, Avouris P (2009) Ultrafast graphene photodetector. Nat Nanotechnol 4:839–843
7. Mueller T, Xia F, Avouris P (2010) Graphene photodetectors for high-speed optical communications. Nat Photonics 4:297–301
8. Pospischil A, Humer M, Furchi M, Bachmann D, Guider R, Fromherz T, Mueller T (2013) CMOS-compatible graphene photodetector covering all optical communication bands. Nat Photonics 7:892–896
9. Deng T, Zhang Z, Liu Y, Wang Y, Su F, Li S, Zhang Y et al (2019) Three-dimensional graphene field-effect transistors as high-performance photodetectors. Nano Lett 19:1494–1503
10. Wang Y, Ho V, Henschel Z, Pradhan P, Howe L, Cooney M, Vinh N (2020) Graphene photodetectors based on interfacial photogating effect with high sensitivity. In: Infrared sensors, devices, and applications X, vol. 11503. International Society for Optics and Photonics, p 1150306
11. Viti L, Cadore A, Yang X, Vorobiev A, Muench J, Watanabe K, Taniguchi T, Stake J, Ferrari A, Vitiello M (2021) Thermoelectric graphene photodetectors with sub-nanosecond response times at terahertz frequencies. Nanophotonics 10:89–98
12. Xu W, Dong HM (2011) Photo-induced carrier density, optical conductance and transmittance in graphene in the presence of optic-phonon scattering. In: Mikhailov S (ed) Physics and applications of graphene–theory. Croatia, InTech, pp 475–498, 501–518
13. Gaitonde J, Lohani R (2021) Novel graphene-based OPFET for optoelectronic applications. Mater Today Proc. https://doi.org/10.1016/j.matpr.2021.08.312

14. Lee E, Balasubramanian K, Weitz RT, Burghard M, Kern K (2008) Contact and edge effects in graphene devices. Nat Nanotechnol 3:486–490
15. Peters EC, Lee E, Burghard M, Kern K (2010) Gate dependent photocurrents at a graphene pn junction. Appl Phys Lett 97:193102
16. Gabor N, Song J, Ma Q, Nair NL, Taychatanapat T, Watanabe K, Taniguchi T, Levitov LS, Herrero PJ (2011) Hot carrier–assisted intrinsic photoresponse in graphene. Science 334:648–652
17. Song J, Rudner MS, Marcus CM, Levitov LS (2011) Hot carrier transport and photocurrent response in graphene. Nano Lett 11:4688–4692
18. Xu X, Gabor NM, Alden JS, Zande AM, McEuen PL (2010) Photo-thermoelectric effect at a graphene interface junction. Nano Lett 10:562–566
19. Sun Z, Liu Z, Li J, Tai G, Lau S, Yan F (2012) Infrared photodetectors based on CVD-grown graphene and PbS quantum dots with ultrahigh responsivity. Adv Mater 24:5878–5883
20. Boyd RW (1983) Radiometry and the detection of optical radiation. Wiley, Hoboken
21. Huo N (2018) and Gerasimos Konstantatos, Recent progress and future prospects of 2D-based photodetectors. Adv Mater 30:1801164
22. Sanchez OL, Dumcenco D, Charbon E, Kis A (2014) Avalanche photodiodes based on MoS_2/Si heterojunctions. arXiv preprint arXiv:1411.3232
23. Lopez-Sanchez O, Lembke D, Kayci M, Radenovic A, Kis A (2013) Ultrasensitive photodetectors based on monolayer MoS_2. Nat Nanotechnol 8:497–501
24. Kotov V, Uchoa B, Pereira VM, Guinea F, Neto AC (2012) Electron-electron interactions in graphene: current status and perspectives. Rev Mod Phys 84:1067
25. Park J, Ahn YH, Vargas C (2009) Imaging of photocurrent generation and collection in single-layer graphene. Nano Lett 9:1742–1746
26. Yan J, Kim M, Elle JA, Sushkov AB, Jenkins GS, Milchberg HM, Fuhrer MS, Drew HD (2012) Dual-gated bilayer graphene hot-electron bolometer. Nat Nanotechnol 7:472–478
27. Britnell L, Ribeiro RM, Eckmann A, Jalil R, Belle BD, Mishchenko A, Kim Y-J et al (2013) Strong light-matter interactions in heterostructures of atomically thin films. Science 340:1311–1314
28. Zebrev G (2011) Graphene field effect transistors: diffusion-drift theory. In: Mikhailov S physics and applications of graphene–theory. Croatia, InTech, pp 475–498, 501–518

Design of CMOS Low Noise Amplifier

Shubham Kumar, Shubham Kumar, Avinash Singh, K. Sai Nischay, Sainnudeep Reddy Nayini, Deril Raju, Soumendra Dash, Ravi Teja, and Vijay Nath

Abstract This paper is discussed the design of CMOS Low Noise Amplifier. This paper has used different techniques for improving the performance of the LNAs. These techniques are evolved in steps and are still being evolved to improve the characteristics further. We will further see the trade-offs done in LNA between their various parameters to improve the other as required to do so.

Keywords Conversion gain (CG) · Input intercept point (IIP) · Mixer · Unbalanced base-band signal (UBBBS) · Local oscillator (LOSC) · Radio-frequency (RF) · Differential balanced signal (DBS)

1 Introduction

Low noise amplifiers (LNAs) are the system devices that can amplify enormously poor signals without decreasing their signal to noise ratio whilst also providing voltage levels adequate for analogue to digital conversion or subsequent analogue processing. Amplifiers will increase the power level of both the signal and noise existing at its input. But furthermore it introduces some noise. Therefore our aim is to minimise noise level without degrading other parameters like bandwidth, gain, chip area, linearity, etc. [1–4]. They are used in a variety of applications involving low-amplitude sources, such as transducers and antennas. When interacting with poor sources, the noise and the gain generated by the very first stage dominate the measuring system's performance. As a result, choosing the right LNA is important for

S. Kumar · A. Singh · K. Sai Nischay · S. R. Nayini · D. Raju · S. Dash · R. Teja · V. Nath
VLSI Design Group, Department of ECE, B.I.T. Mesra, Ranchi, Jharkhand 835215, India
e-mail: be10371.17@bitmesra.ac.in

V. Nath
e-mail: vijaynath@bitmesra.ac.in

S. Kumar (✉)
Steradian Semiconductors Private Limited, Bangalore, India
e-mail: shubhamprice99@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_29

the successful operation of the experimental setup. This note gives a quick overview of LNAs to help you grasp their basic characteristics and select the right device [5]. Take a look at a real-world amplifier with gain of G. The intended signal is Si and an undesirable input noise is Ni comprise the input from such an amplifier. Combined signals will be amplified by a factor of G at the output. Nevertheless, there is another source of noise. Due to the obvious noise made by the device's electrical components, Nampl was created. To correctly amplify a low-level signal, noble LNAs must have high gain and a little inner noise Nampl [7].

Ordinary amplifiers have same parameters as LNAs. Nonetheless, because they are designed to offer extremely high gains whilst adding very little noise to the input signal, the users should concentrate on the amplifier's noise response parameters. The noise factor (F) associates an amplifier's noise yield with that of an ideal noiseless device. Because electrical components inherently generate noise, every amplifier's noise factor is always greater than one. The noise figure (NF) is a related metric that is noise factor in decibel notation, i.e. the noise spectral density (NSD) is the arrangement of the amplifier's inherent noise power throughout its bandwidth. Though it is commonly stated as the corresponding input noise (Nampl/G) for a better contrast with noise of input signal, it correlates to Nampl in the image above [6]. The NSD is included. The capacity of an LNA to amplify a signal without distortion is measured by its linearity. Linear device means output power in decibel (db) is sum of the input power and the gain of the device. But in reality during amplification it introduces some noise which distorts the signal. Now behaviour of LNA is no longer linear.

2 Various Techniques Used for Improving Performance of LNAs

Main goal behind any improvement idea of the LNAs is to fulfil two purposes in them (i) Increase the Bandwidth of the circuit. And second one is (ii) Noise reduction. All the efforts made are driven by these two motivations only. Now various techniques such as current reuse technique, use of common gate arrays, distributed amplifier configuration, use of differential cascode structure, addition of RLC networks are used to solve the two problems which we discussed in Fig. 1. Every technique that we used had some advantages and disadvantages of their own and is used separately for different purposes [7].

Distributed amplifier configuration gives good linearity in the device characteristics and good input output matching with ultrawide bandwidth and low noise but suffers from higher power consumption as sequential classes have been used there.

So a degenerative feedback technique using active/passive elements is commonly used technique in wide band amplifiers. But this technique failed to explain noise and power simultaneously.

Uses of resistive feedback gain to improve the noise cancellation technique which is shown in Fig. 2. Initially, we were using inductor in a circuit which improve the

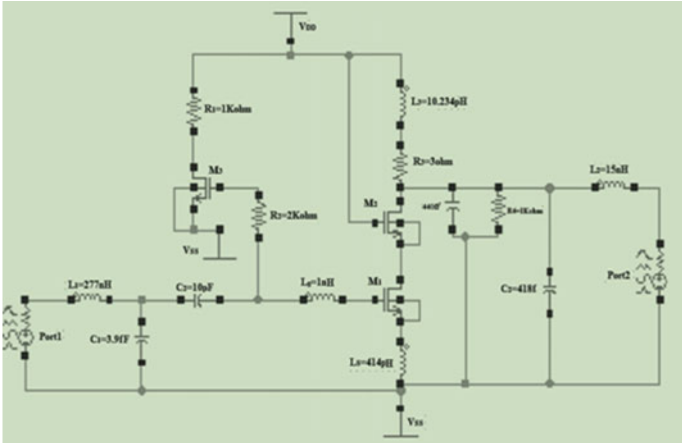


Fig. 1 Conventional CMOS LNA circuit

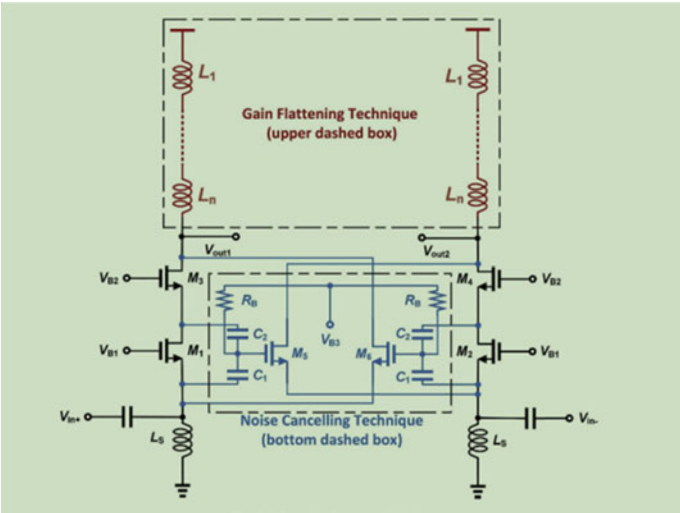


Fig. 2 Proposed CMOS low noise amplifier

bandwidth, linearity and impedance matching but consume more power and more chip area because inductor is bulky in nature. So designer introduced a new methods using resistive shunt feedback technology. Since no inductors are used here so chip area will minimise and also reduces power consumption but degrades noise figure. This can be used if noise figure is not an issue (Fig. 3).

Common Gate arrays are used that overcome the trade-off between noise and power without reducing its stability. So we use noise cancellation techniques which introduce noise in several pathways that are out of phase and by adding this output

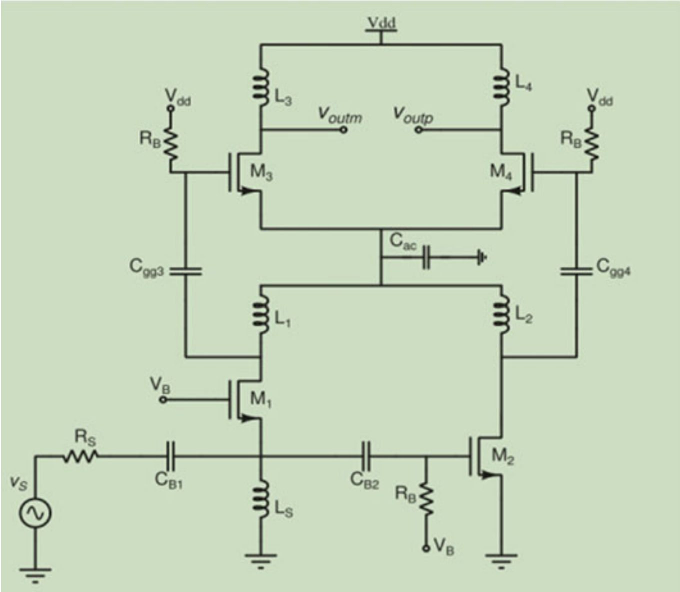


Fig. 3 Proposed CMOS UWB—LNA noise cancellation and current reuse techniques

noise will be attenuated. To cover up problem of the high power consumption current reuse technique is utilised. This exploits low input impedance to achieve lower power consumed in the device [3].

Differential cascode CG structures are used with calculated transconductance values to internally cancel out the noises by perfectly calculating the required Gm values for the various transistors used in the structures. The use of input inductors in this configuration excludes parasitic capacitance and increases the bandwidth of the circuit, but it has a high noise figure attributed to the combination of four inductors and transistors, which is improved by some additions to the configuration and finally achieves the desired performance [4].

Current reusing is a low power consumption strategy, although it struggles from the input impedance matching. A WB balun LNA is used with a well-balanced output which is shown in Fig. 4. Noise and the distortion caused by the input of the transistor are reduced utilising noise cancelling techniques for improved noise performance. The main principle is that the input transistor is the main source of noise, so a node is identified where the signal would seem in the opposite sign whilst noise of input of the transistor would seem in the similar sign, and when the gain of the device is doubled whereas the noise is rescinded, but we need large Gm values for that which means only high power applications. This drawback is also removed by further modifications in the circuit [8–16].

Current reuse with CG-LNA boosting technique—initially we were using current reuse technique to improve although it has a low power consumption; it suffers from

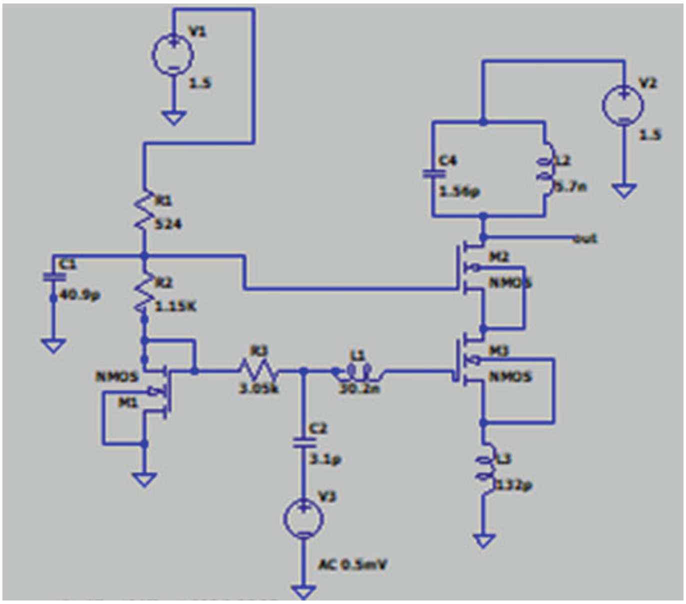


Fig. 4 CMOS low noise amplifier circuit [17]

poor noise figure. Therefore a gm-boosting strategy to further reduce power usage and enhance noise figure has been presented.

Passive implementation of gm-boosting technique only reduce half of operating current but here noise figure is greater than 1 so we further improve our circuit and introduce active implementation of the gm-boosting technique to improve noise, bandwidth and gain of circuit but it suffers from high power which further improved using impedance matching technique.

In this circuit (Fig. 4) which we have taken as our reference a voltage divider circuit is used in which passive elements are used so for our improvement we will change this into a voltage divider such that it comprises active devices, i.e. we are going to use MOSFETs as the resistors in the voltage divider circuit.

Now there are different combinations which we are going to use here, i.e. we will not just change all the resistors in to the MOSFETs rather we are going to separately change the resistors separately one by one, i.e. first we will replace the R1 with a PMOS and keep the R2 intact. In the next step we will change the R2 with PMOS and keep R1 intact.

In further steps we will change both the R1 and R2 with the PMOS and use the simulation using the LT spice tools. Replacement of R1 from Fig. 4 and its responses are shown in Figs. 5 and 6. First we will simulate our reference circuit and find the noise at the output port and also the output voltage.

Circuit of Fig. 7 responses are shown in Figs. 8 and 9.

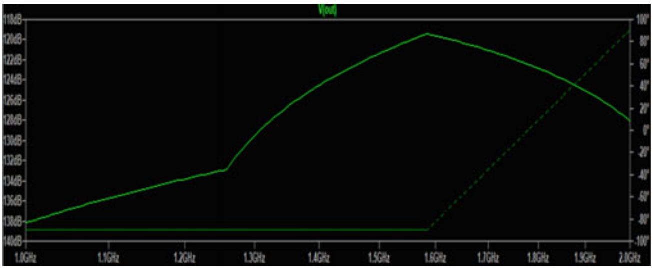


Fig. 5 Simulation results of Fig. 4 in reference Vout with frequency using LTSPICE

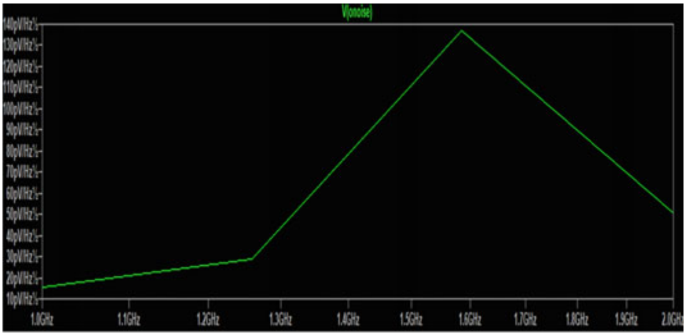


Fig. 6 Reference noise with frequency of Fig. 4

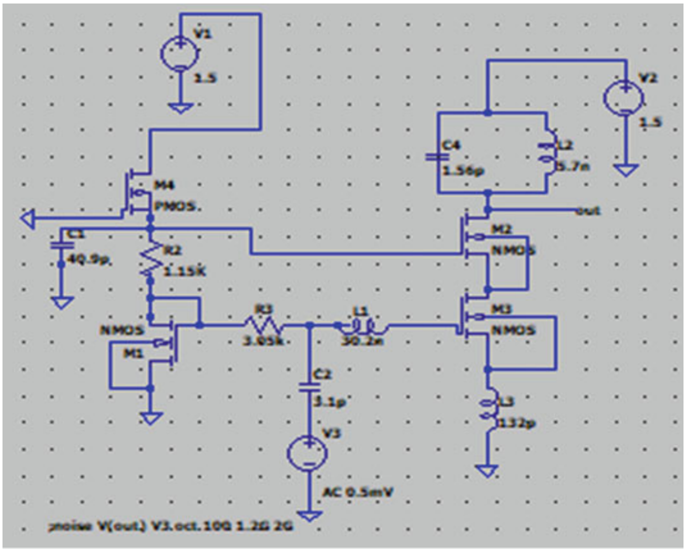


Fig. 7 Replacement of R1 with PMOS

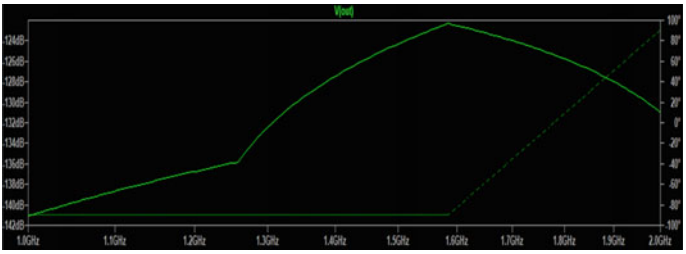


Fig. 8 Vout with frequency after replacement of R1 with PMOS of Fig. 7

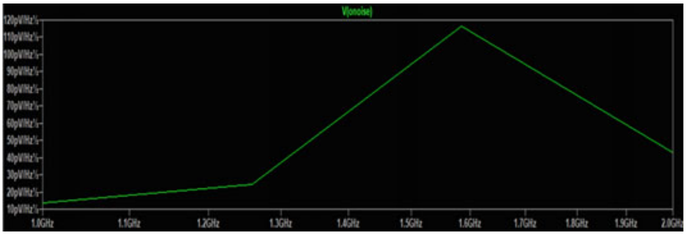


Fig. 9 Noise at the output after replacement of R1 of Fig. 7

Similarly for the other improvements we also need to change the resistors first and then do the simulation using the same LT spice tool and observe the same parameters and finally do a comparison and status check to see if the final circuit performance is improved in comparison with the reference circuit.

Circuit of Fig. 10 it responses are shown in Figs. 11 and 12.

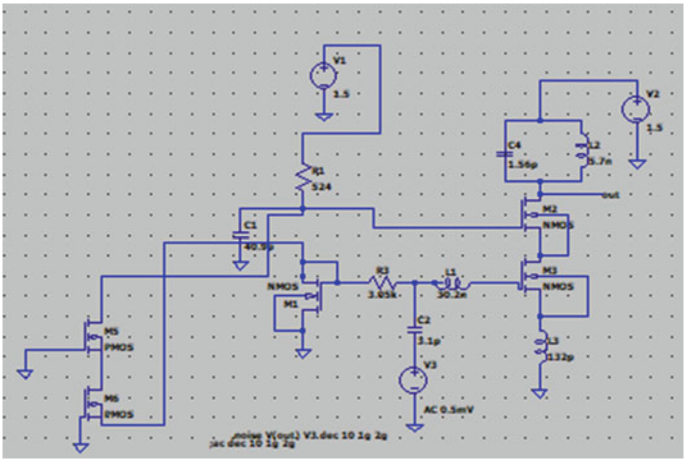


Fig. 10 Replacement of R2 with 2 PMOS in series

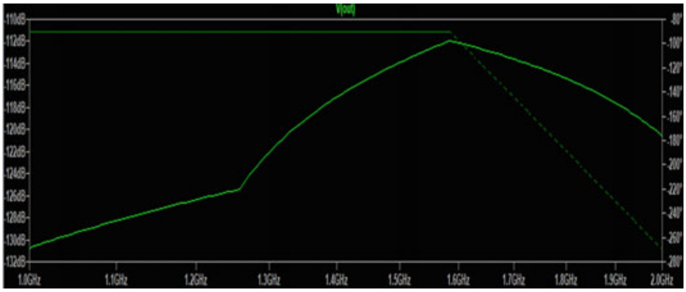


Fig. 11 Vout after replacing R2 with 2 series connected PMOS of Fig. 10

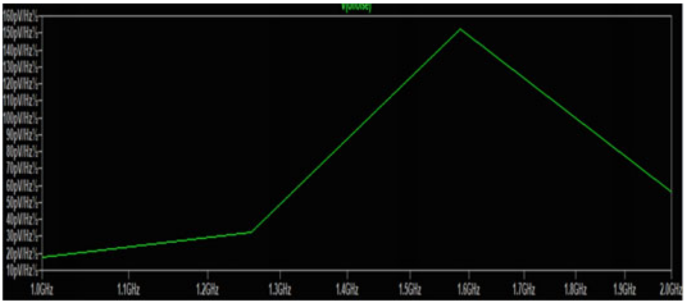


Fig. 12 Noise at the output after replacement of R1 of Fig. 10

In the circuit (Fig. 10) as a change we will replace the R2 with 2 PMOS connected in series whilst keeping the R1 intact, this circuit we will implement and simulate as a next change and next step in our research work. The voltage output for this one we can see that is deteriorated from the earlier case and also the noise value is increased at the output port. This means that this change is making the performance of our circuit degrade not only compared to the reference circuit but also in comparison to the last circuit which we implemented as a change in which removed R1 and replaced that with a PMOS.

So as a comparison between these two circuits we can conclude that this change is more effective and more helpful in our way to gain a circuit with more idealistic nature as compared to the earlier circuit.

In this circuit we have replaced R2 with two PMOS connected in series because if we see in the reference circuit R2 is almost equal to 2x of R1 that's why we are using 2 PMOS in series so that the effective resistance of the new circuit follows the earlier relation. Simulation results for this circuit are as follows:

As we can clearly see that after replacement of R1 the Vout and noise at the output both are changed significantly so, we will continue with the next change in the circuit where R2 is replaced with 2 PMOS connected in series and observe the simulation results for these [18].

As a last change in our reference circuit we have made these replacements to get the following simulation results. We can clearly see the changes we have made with respect to the earlier circuits. Both of the resistors are removed and in their places PMOS are placed but one thing which we can see are that the R1 is replaced by one PMOS, whereas the R2 is replaced by two PMOS which is shown in Fig. 13. Circuit of Fig. 13 output responses are shown in Figs. 14 and 15. There is a reason behind this difference which is that in the reference circuit we can see that the value of R1 is almost half of R2 and the resistance of the MOSFET is dependent on the length of the MOS thus the resistance are used in series to incorporate the effect of the resistance values also. There is also a tank circuit used in this device which is basically used for determining the frequency of operation of the device and changing the value of any one of the L or C will affect the operating frequency of the device so it is basically used to control the frequency [19].

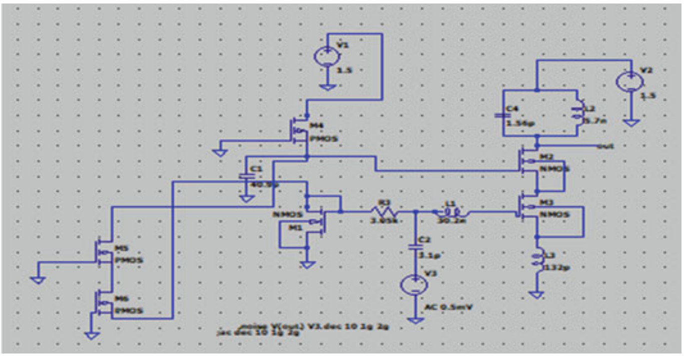


Fig. 13 Replacement of both R1 and R2 by PMOS

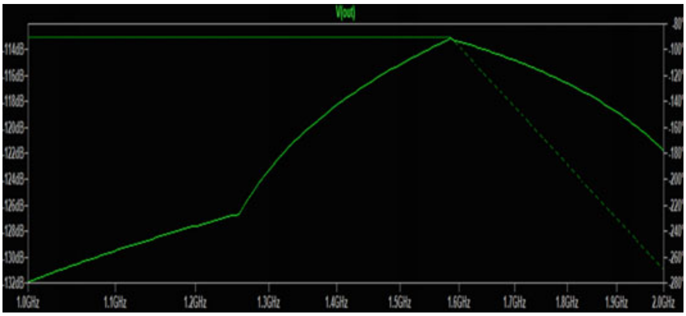


Fig. 14 Vout after replacement of both R1 and R2 with PMOS

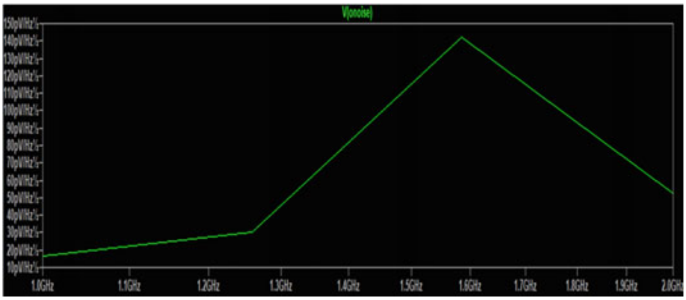


Fig. 15 Noise at the output after replacement of both R1 and R2 with PMOS [20]

3 Conclusions

After reviewing all the papers it can be concluded that large number of efforts are made to reduce noise and increase the bandwidth of the device but there is possibility of many further improvements in terms of low power dissipation and improving noise figure by making some new improvements in upcoming future. This change is giving the results almost same as the reference circuit whereas other two changes either improve the performance or degrade it.

References

1. Vasudeva Reddy K, Sravani K, Kumar P (2017) Low power ultra wide-band balun LNA using noise cancellation and current-reuse techniques. Microelectron J 61:114–122
2. Eskandari R, Ebrahimi A, Sobhi J (2017) A wideband noise cancelling balun LNA employing current reuse technique. Microelectron J 76(2018):1–7
3. Li Z, Wang Z, Zhang M et al (2014) A 2.4 GHz ultra-low-power current-reuse CG-LNA with active gm-boosting technique. IEEE Microw Wirel Compon Lett 24(5):348–350
4. Liu L, Lu Z, Zhang K, Ren Z, Hu A, Zou X (2016) Wideband balun-LNA exploiting noise cancellation and gm compensation technique. Electron Lett 52(8):673–674
5. Aravinth Kumar AR, Sahoo BD, Dutta A (2018) A wideband 2–5 GHz noise cancelling subthreshold low noise amplifier. IEEE Trans Circuits Syst II Express Briefs 65(7):834–838. https://doi.org/10.1109/TCSII.2017.2719678
6. Akbar F, Atarodi M, Saeedi S (2015) Design method for a reconfigurable CMOS LNA with input tuning and active balun. AEU Int J Electron Commun 69(1):424–431
7. Tripathy D, Rout SS, Sethi K (2015) A low power noise cancelling LNA for UWB receiver front end. In: 2015 IEEE power, communication and information technology conference (PCITC). Bhubaneswar, India
8. Shylashree N, Venkatesh B, Saurab TM, Srinivasan T, Nath V (2019) Design and analysis of high-speed 8-bit ALU using 18 nm FinFET technology. Microsyst Technol 25(06):2349–2359. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-018-4112-y
9. Gupta P, Verma VK, Ranjan RK, Appasani B, Priyadarshini B, Nath V (2019) A series expansion method aided design of current mode second generation current conveyor based active control circuit. Microsyst Technol 25(06):2323–2330. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-018-4117-6

10. Verma VK, Ranjan RK, Gupta P, Priyadarshini B, Nath V (2018) A series expansion method aided design of CCII controller for TITO system. Microsyst Technol 24(9):3843–3849. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-018-3869-3
11. Priyadarshi R, Soni SK, Nath V (2018) Energy efficient cluster head formation in wireless sensor network. Microsyst Technol 24(12):4775–4784. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-018-3873-7
12. Chakraborty S, Pandey A, Prasad D, Vedam V, Nath V (2018) Linearity improvement of gain enhanced op-amp using cross-coupled architecture. Microsyst Technol 24(12):4807–4815. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-018-3885-3
13. Priyadarshi R, Soni SK, Bhadu R, Nath V (2018) Performance analysis of diamond search algorithm over full search algorithm. Microsyst Technol 24(06):2529–2537. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-017-3625-0
14. Prasad D, Nath V (2018) An ultra-low power high-performance CMOS temperature sensor with an inaccuracy of −0.3oC/+0.1oC for aerospace applications. Microsyst Technol 24(3):1553–1563. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-017-3564-9
15. Chakraborty S, Pandey A, Nath V (2017) Ultra high gain CMOS op-amp design using self-cascoding and positive feedback. Microsyst Technol 23(3):541–552. Online ISSN 1432-1858. https://doi.org/10.1007/s00542-016-2971-7
16. Pandey A, Nath V (2016) A CMOS temperature sensor and auto-zeroing circuit with inaccuracy of −1/+0.7 °C between −30 °C to 150 °C. Microsyst Technol 23(3):4211–4219. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-016-2968-2
17. Kundu R, Pandey A, Chakraborty S, Nath V (2015) A CMOS low noise amplifier based on common source technique for ISM band application. Microsyst Technol 22(3):2707–2714. Online ISSN: 1432-1858. https://doi.org/10.1007/s00542-015-2550-3
18. Sateesh VA, Surshetty SK, Goel V, Prasad D, Nath V, Pal S (2021) Designing of low-noise amplifier and voltage-controlled oscillator for satellite receiver in Ku band. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. Lecture notes in electrical engineering, vol 692. Springer, Singapore. https://doi.org/10.1007/978-981-15-7486-3_58
19. Datta K, Pal S, Nath V (2021) Design of a 2–30 GHz low-noise amplifier: a review. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. Lecture notes in electrical engineering, vol 692. Springer, Singapore. https://doi.org/10.1007/978-981-15-7486-3_64
20. Rao NS, Prasad D, Nath V (2020) Broad band capacitive coupling antenna for C-band applications. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. NCCS 2018. Lecture notes in electrical engineering, vol 642. Springer, Singapore. https://doi.org/10.1007/978-981-15-2854-5_1

Design of Unmanned All-Terrain Spy Bot

Aditya Kumar, Aradhana, Shristi Kumari, Shipra, Deepak Prasad, Divya Sharma, T. Snehitha Reddy, Bhagwan Singh, Deepika Kotwar, and Vijay Nath

Abstract In the advancement of technological era, the military forces are utilizing these modern technologies to mitigate the risks of casualties and can keep an eye on the enemy. With the development of sophisticated technology, robotics is widely used in disaster management rescue operations and for spy purpose and surveillance in war field. In order to make the bots effective and efficient for war field and disaster-prone areas, several research on different aspect is examined to do the jobs autonomously, without any human intervention. Latest technologies in terms of software and hardware are investigated and implemented to have intelligent and advance robots for different operations in the disaster-prone areas and war field. This paper proposed the robotic technologies being used in disaster management and war field spying. Our robot is equipped metal detector, fire sensor, gas sensor as well as a camera for live streaming and still capture. The robot is controlled manually with the self-design controller wirelessly. In the proposed scheme, the robot will continuously send the GPS coordinates to authenticate mobile phone via SMS. The warning SMS will be sent if metal, fire, or gas are detected along with their coordinates. The camera data are very much secured as it has the password to access the camera network and a particular IP to access the data. The major focus of this proposed paper is on the uses of robot for spying purpose in war field and help NDRF team in the disaster-prone areas and provide a positive impact on the society.

Keywords Spy robot · GPS tracking · GSM services · Payload

A. Kumar (✉) · Aradhana · S. Kumari · Shipra · D. Sharma · T. Snehitha Reddy · B. Singh · D. Kotwar · V. Nath
Department of Electronic and Communication Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: aditya88952@gmail.com

D. Prasad
Department of Electronic and Communication Engineering, Sarala Birla University, Ranchi, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_30

1 Introduction

Every day in the world, many incidents are taking place which are risky and poses life-threatening consequences in case of human intervention. In the proposed scheme, the main aim is to build a robot spy bot which will be useful for disaster management and war field spying. The robot is equipped with a camera which will give real-time video streaming along with still capture. The robot consists various sensors like metal detector, fire sensor, and gas sensor. It can send real-time coordinates to the phone at regular interval of time. The bot will be controlled wirelessly from a remote distance. The spying bot as its name suggests is the one used for the purpose of spying on enemy territories and NDRF team to help the people who are stuck in disaster-prone areas. The plan of action and strategy can be made easily on the basis of the information send by the robot. The wireless communication is done with the help of NRF24L01 module. The robot will serve as an appropriate machine for the defense and rescue sector to reduce the loss of human life and will also prevent illegal activities. The robot consists of two part: payloads and transmitter–receiver unit.

2 Methodology

In the proposed scheme, the robot is designed in two parts; first is payload, and the other one is transmitter and receiver unit. The payload consists of metal detector, fire sensor, and gas sensor along with GPS tracker and GSM sim900A module. The camera plays a very important role in spying purpose. The communication between transmitter and receiver is done via nrf24l01 which operates on 2.4 GHz ISM band and requires very low power, i.e., 3.3 V. The trans-receiver can operate with signaling rate 250 kbps to 2 Mbps. The communication module can use 125 different channels which gives a network of 125 independently working modems in one place.

3 Payloads

The payloads include camera, metal detector, fire sensor, gas sensor, GPS tracker, and GSM module as shown in Fig. 1. The GPS tracker will continuously track the robot and send the real-time coordinates to the phone via GSM module at regular interval of time (3 s). In case any of the sensor detect the signal, the GSM module will send the emergency message along with the coordinates. The camera module will give the live video streaming and can take still capture and store the data on a secured network at a particular IP. The flowchart payload is shown in Fig. 2.

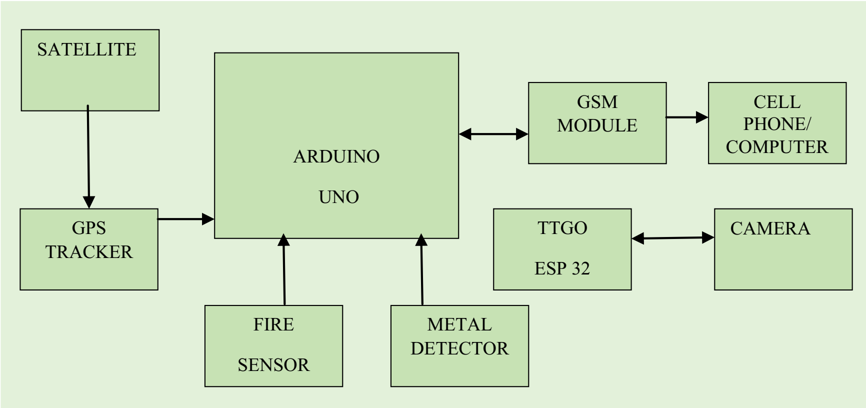


Fig. 1 Block diagram of payload

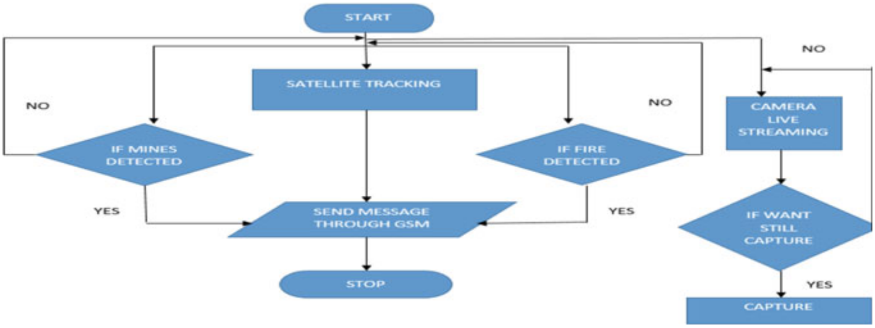


Fig. 2 Flow diagram of payload

4 Receiver and Transmitter

The transmitter consists of a microcontroller, nrf24l01, joysticks, and a potentiometer as shown in Fig. 3. The inputs are fed to the microcontroller through joysticks and potentiometer. Then, the signals are sent through nrf24l01 module to the receiver side. The motors (wheels) of the robot are controlled by joysticks, and the movement of camera is controlled by the potentiometer.

The receiver consists of a microcontroller, nrf24l01, motor driver, DC motors, and servo-motors as shown in Fig. 3. On receiving the signal from transmitter, the controller will give the command to the motor driver to drive the motor according to the signal which is generated by the movement of joysticks. The movement servo-motor is responsible for the angular movement of camera. This movement is due to the change in the resistance in the potentiometer on the transmitter hand side as shown in Fig. 4.

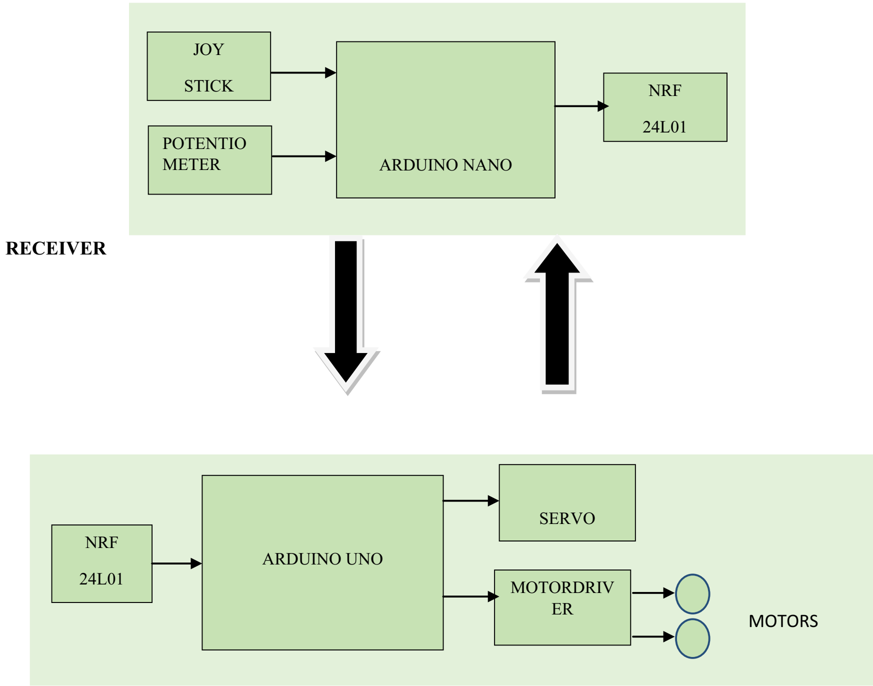


Fig. 3 Block diagram of transmitter and receiver

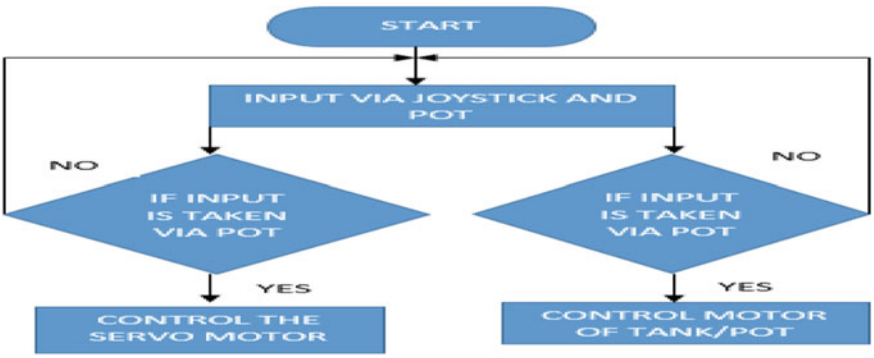
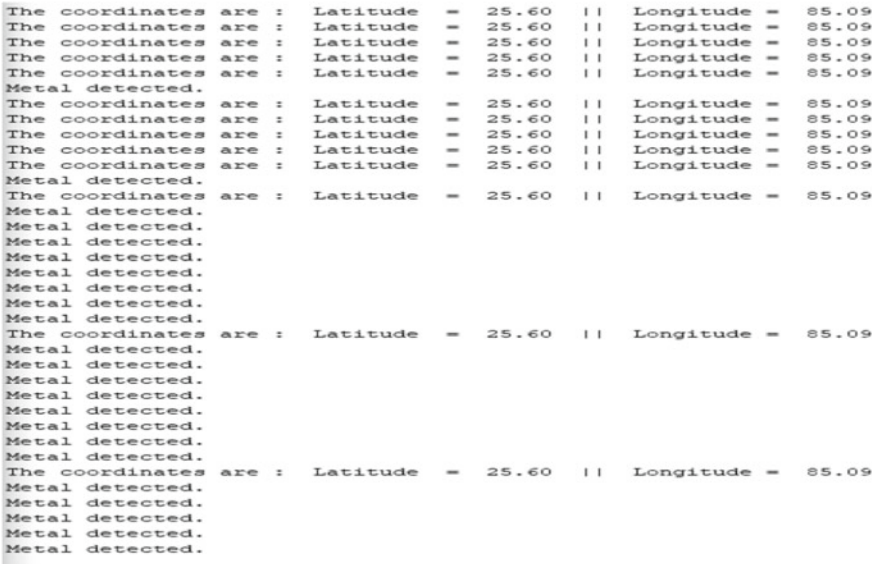


Fig. 4 Flow diagram of transmitter and receiver

5 Result and Conclusion

The GPS will continuously sent the data to the phone, as shown in the serial monitor in Fig. 5. Using the proposed design, a person would be monitored very easily. Such



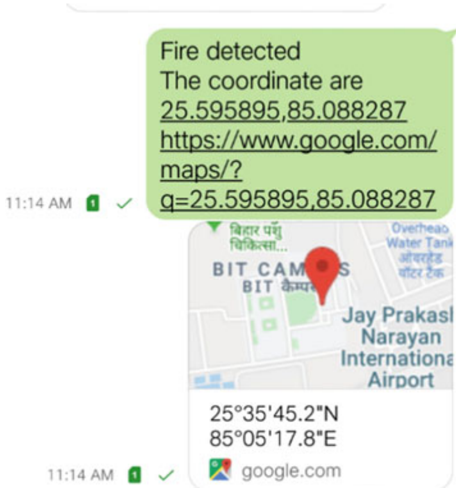
```
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
Metal detected.
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
Metal detected.
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
The coordinates are :  Latitude  =  25.60  ||  Longitude =  85.09
Metal detected.
Metal detected.
Metal detected.
Metal detected.
Metal detected.
```

Fig. 5 Screenshot of serial monitor

designs can be highly useful in defense system, military purpose, disaster-prone areas.

The message with warning alert along with the coordinates is sent to the mobile phone as shown in Fig. 6.

Fig. 6 Screenshot of mobile phone which shows warning message along with coordinates



Bibliography

1. Patel DS (2013) Mobile operated spy robot. Int J Emerg Technol Adv Eng (IJETAE)
2. Mehta L, Sharma P (2014) Spy night vision robot with moving wireless video camera. Int J Res Eng Technol Manage (IJRETM)
3. Samad A, Dasharath JD, Kumar DM (2014) An intelligent combat robot. Int J Adv Res Comput Commun Eng 3(5)
4. War field spying robot using wireless camera. Int J Electr Electron Res 4(1):85–92. ISSN 2348-6988 (online)
5. Jha A, Singh A, Turna R, Chauhan S (2015) War field spying robot with night vision camera. J Network Commun Emerg Technol (JNCET) 2(1)
6. Microchip (2005) PIC16F627A/628A/648A Data Sheet. ISO/TS 16949:2002, Microchip Technology Incorporated, USA
7. Microchip (2001) PIC16F87X Data Sheet 40-pin 8-Bit CMOS FLASH Micro controller. ISO 9001/QS-9000, Microchip Technology Incorporated, USA
8. Ibrahim D (2006) Microcontroller based applied digital control. John Wiley & Sons, Ltd. England. ISBN: 0-470-86335-8
9. Iovine J (2004) PIC robotics: a beginner's guide to robotics projects using the PICmicro. McGraw-Hill
10. Wilmshurst T (2007) Designing embedded systems with PIC microcontrollers. ISBN-10: 0-7506-6755-9, Elsevier
11. Robert P (2012) Introduction to gear design. Continuing Education and Development, Course No: M03-016

Design and Implementation of IoT-Based Medicine Dispensary Box

Disha Agarwal, Rajat Agarwal, Dilip Kumar Choubey, and Vaibhav Shukla

Abstract The health industry and wellness sector are very crucial for the human beings and as this should be one of the important and early sectors to receive the advances of emerging technologies in the field of IoT. In the current health industry, the use of Internet of Things (IoT) makes the whole system more convenient. The Internet of Medical Things (IoMT) is also connected to IoT networks in this scenario to monitor the patients' daily activities and to record them for future analysis. Recently, there have been many attempts to design an entirely new smart medical device which can monitor the medications of the patient and help elderly people for a better life. In this paper, we proposed an improved version of IoT-based medical medicine dispensary box. Basically, we have developed a Host Management System (HMS) which is an efficient of cloud-based installation and monitoring that stores and controls the MEDIBOX functionality for further analysis and future modification in design aspects.

Keywords IoT · IoMT · MEDIBOX · Portable · Monitor · Dispensary box · Medicine

1 Introduction

Today, there are more than a million devices available in to the market which monitor the health issues in different body parts of a person. The Internet of Medical Things

D. Agarwal (✉) · R. Agarwal
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: disha.agarwal9848@gmail.com

D. K. Choubey
Department of Computer Science and Engineering, Indian Institute of Information Technology Bhagalpur, Bhagalpur, Bihar, India
e-mail: dkchoubey.cse@iiitbh.ac.in

V. Shukla
Tech Mahindra Ltd, Mumbai, Maharastra, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_31

refers to that sector which connects these devices and cloud system which gives a proper connectivity between network and devices. In present days, the health and fitness are the biggest concern for humans, and if anyone has any kind of health issue, then it needs medicines to treat them, and here, we come up with our idea. Presently, there is no proper management or any device to monitor the medicines for a particular patient. There are some users who do not have the knowledge like where the medicines, their expiry date quality, or which medicine to take at what time. So, to avoid such situations which may lead to serious issues medicine monitoring model is used. The system is administered by the Raspberry Pi which utilized the IoT and RFID technology. Presently, there are no device which can be installed at home for a particular patient and monitor. In today's world, the health and wellness sector is very important for the human, and this sector should be one of the most prioritized to receive the all advantages of any upcoming skills and advances in IoT. The few Internet of Medical Things (IoMT) are also attached with various such kinds of networks to monitor and observe the everyday activities of patients using the dispensary box. In recent times, there have also been several attempts that can be used to take care of all the medications and also in helping of aged and old people who cannot take care of their own medication to lead a better and healthy life. In this paper, we have made such an attempt to architect a multipurpose, easy to carry and a very smart device named medicine dispensary box which will help the user/patient to take their medication at the right time and have it regularly and keep a record of their dosage. This smart box is very efficient, and it will help monitor the temperature, humidity, and a lot of factors inside the box. Even while traveling we can ensure that the desired environment of the box can be maintained so that the medicines are kept in good condition. In addition to this, we have also worked on the IoT cloud-based systems where all the patient data has been stored into the cloud system. So that all the data is stored in one place and all the records are easily maintained. This also helps in maintaining the suitable environment of the box, and it contains all the records of the patient in one place for easy accessibility and ease of use.

Today, there are more than a million devices available into the market that monitor the health issues in different body parts of a person. The Internet of Medical Things refers to those sectors which connect these devices and cloud system and gives a proper connectivity between network and devices. In present days, environment health and fitness are the biggest concern of human, and if anyone has any kind of health issue, then it required a proper medical treatment, and here, we come up with our idea. Presently, there is no such mechanism or any device by which we can monitor that the patients take their medicines regularly or not. There are some users who do not have the knowledge about the medicines, their expiry date, quality, or which medicine to take at what time. So to avoid such situations which may lead to serious issues, medicine monitoring model is used. In this architecture, we will make use of Raspberry Pi which will administrate the whole system, and it uses the IoT and RFID technology. Presently, there are no device which can be installed at home for a particular patient and monitor.

The rest of the research paper is arranged as follows: Motivation has been introduced in Sect. 2, Literature Review has been discussed in Sect. 3, Proposed Methodology has been presented in Sect. 4, Experimental Details and Implementation are presented in Sect. 5, and Sect. 6 has been committed to Discussion and Future Directions.

2 Motivation

The motivation to develop a box using IoT and cloud came from a lot of factors. But, the most important factor is the dependency of a family or patients on a nurse or caretaker to take medications regularly. Elder age patients generally have problem of remembering which medicine to take on what time and at what dose. This makes them dependent on a nurse or a caretaker which a family has to appoint. This requires a constant monitoring on the patient. There are different kinds of medicine dispensing systems used in hospitals, but the idea of getting a medicine dispensing box at home will take the safety and health benefits to a greater level. It assures the patient take their medicines on time and on proper dose as well as gives them the power to monitor and to control their inventory. In case, if a patient forgets to take their medication, then we can check on what date the patient did not take their medicine and make a proper record on cloud and properly coordinate with doctor for further treatment. The family/hospital which is currently taking care of the patient can immediately take actions when the box is empty. Moreover, the doctors would know if the patient is taking the medicines on time or not, and this would in turn help the patients to be healthy on time.

The proposed box or system is very convenient and helpful for the customers. The health care industry is expanding every day, and people are now busy with their daily routines. At this stage, everything that deduces concern about something can be regarded as a boon and the system plays along exactly the same lines. Patients who have already been discharged from the hospitals require the same level of care as those who are admitted in the hospital. It can reduce the concern of the family members to great extent.

3 Literature Review

Moen et al. [1] have used a convectional analysis method to carry out a survey, and they have used a semi-structured focus group algorithm to complete the analysis on different category of groups. The main focus of this paper was to get information about the people of different age, culture think and get their perception on having multiple dosages of medicine.

Cartwright [2] has divided the work in three categories; the first category gives the readers an idea on the relation between aging, health, medicine consumption, and consultation rates. This paper then gives us an idea on how senior citizen is getting less care and attention or supervision in comparison to what they need.

Pound et al. [3] have given an idea on studies on experiences of medicine taking, and this paper has revealed the importance for taking proper and correct medicines and the authors also highlighted the lay habit of testing pills which mainly focuses on its side effects.

Williams and Manias [4] used patients whose age is greater than eighteen years and who has a diabetic condition, CKD, and systolic hypertension. This paper focused on these types of patients who have either of above stated conditions to carry out the research, most of the patients were gathered from hospitals in Australia from the year 2008–2009. This paper tires to highlight and research the motivation and confidence of the suffering patients from the above stated conditions and how they take their prescribed medication.

Tordoff et al. [5] have done a research on New Zealand has old people taking care of their medication themselves, and it further discusses the issues old people face due to missing out on medication.

Elliott et al. [6], their study suggests to help people having medical conditions related to mental health, dementia so that they do not skip their medication and take is at the right time, to lead a more peaceful life.

Orlu-Gul et al. [7] have focused on how to make the lives of the older people better so that they do not skip medication.

Rogers et al. [9] developed an vacuum-based system for dispensing tablets of oral hard medicine from a big container as storage to the user or patient, in which the medicine is dispensed with the help of a computer controlled system, and the amount or number and category is pre-decided by the user itself.

Ostergaard et al. [10] have done a research on the current technologies and science which is related to developing machines to dispense proper dosages of medicine to the consumer, and this paper particularly deals about medical dispensary machine which has a storage box for storing the medicines into the medicine dispensary box, a medicine segregation area in the box, a system which will pack the medication in a predefined dosages given by the user, it also has a central medication database system and a system in a form of box which will have all these objects.

Zamjahn [12] has developed a calculated medicine dose dispensing machine which consists of a medicine bottle and a calculated dose medicine dispenser area for individual medicine. The calculated dose medicine dispenser provides a medicine box and a system for determining dosage for prescriptive and over the counter medication as well as for providing the proper dose to the patient.

Ratnakar [13] has developed a smart medical dispenser machine which automatically dispenses medicines at particular time intervals and consist of a pill recovery system.

Shepherd et al. [15] have developed a machine which gives dosages of medicine according to a pre-decided scheduled which is also decided by the user, and it also has a container through which plurality of medicine will be provided and accessible to the user.

Hafner [16] has developed a system which stores the medicine and dispense them on the basis of a time-regulated algorithm, and it consist of medicine box which acts as a container which is connected to a secured connection with a signal generator that is active according to the time algorithm.

Braun et al. [8] have designed a system, method, apparatus, and computer program product are provided for detecting that an object has been accessed using tags.

Walker et al. [19], we have studied their articles and reports and found out that in the first section of their study they have discussed about two apparatus, where they take care of taking medicines timely. In the first section, they have discussed whether at least two of the medications were being able to be communicated and dispensed properly, and being able to perform according to the patient needs and dispensing medication according to the information provided to them. Choubey et al. [20] have used IoT and cloud computing approach to enhance the productivity of crop production. The IoT and cloud computing approach have been applied on image dataset. Kumar et al. [21] have utilized IoT and machine learning for the predictions of the cardiac arrest through heart variability analysis.

Table 1 briefly summarizes the existing works, hardware used, advantages, and issues on existing articles.

4 Proposed Methodology

The proposed architecture of our methodology is shown in Fig. 1. It shows the design model of IoT-based medicine dispensary which is efficient and reliable.

Here, we have discussed proposed algorithm which is noted below.

4.1 Proposed Algorithm

Step 1. Define the servo motor, temperature sensor, ultrasonic sensor, LCD, Wi-Fi module, and GSM Module.
Step 2. Check the amount of medicine present in the medicine dispensary box using the ultrasonic sensor, as soon as the level drops low send an alert message.
Step 3. Keep recording and storing the data to the cloud so that all the details are maintained and can be accessed at any point of time by the doctor or the patient.
Step 4. Keep track of the temperature. If it falls low or higher than the desired temperature required, notify by sending an SMS.
Step 5. When RFID tags are tapped, the medicine should be dispensed depending on which RFID tag is tapped. A message stating that the medicine is taken is sent.

All the details are stored on the cloud for maintaining the record and for easy access in the future.

Table 1 Summary existing work on IoT-based dispensary box

Ref. No.	Work done	Hardware used	Advantages	Issues
[1]	A study on 12 particular focused groups consisting of 29 men and 30 women which are from a age gap of 65-years or even more, more than five prescribed medication were studied qualitatively	…	Various age groups are present	Only five samples are used
[3]	In this study, the data is collected and analyzed. A data set was taken from January 1, 1992, to December 31, 2001	…	Proper data analysis is done on the bases of medicines	Only a data set of 10 years was used
[4]	In this paper, there was a study done on Australian patients that were hired from outpatient hospitals. They are minors with diseases like diabetes	…	Various age groups and different medically ill patients were used as the sample data set	Only focused on a particular area and with only 1 year sample cases
[9]	A vacuum operated system for individually dispensing medicine from storage to a user	Several motors, containers, display, keyboard	It dispenses medicines for every individual using ID password	Security breaches
[10]	A research on the current technologies and science which is related to developing machines to dispense proper dosages of medicine to the consumer, and the authors has also developed a dispensary machine	Several motors, containers, GSM, SMS technologies	It has memory for storing data on the machine, and this whole system also consists of a central medication database for getting the maximum efficiency	Proper network is required

(continued)

Table 1 (continued)

Ref. No.	Work done	Hardware used	Advantages	Issues
[11]	A simple device which holds a week's supply of medication and dispenses it at timed increments during the day to remind the user to take the medicine	Electric motors, containers, driving gear	Works automatically according to time for a single user	Can hold medical supplies only up to a 1 week's supply
[12]	A calculated dose medicine dispenser which determines dosage for prescriptive and over the counter medication as well as for providing the proper dose to the patient	Several motors, containers, wires	It calculates dosage or amount of medication required for every medication	Calculation of dosage was not accurate
[13]	Automatically dispensing medicines at particular time intervals and consist of a pill recovery system	Belts, several motors, containers, wires	Dispenses medicines on the basis of time	Automatic system is always not helpful
[14]	A medication dispensing machine which is tamper-proof and is controlled through microprocessor, and dispenses medication for a plurality of days, as well as at a plurality of times per day	Several motors, belts, containers, microprocessor, wires	Use of microcontroller makes the system more efficient	Implementation of microprocessor is tough
[15]	In this study, information was already stored and the dispensary just dispensed the desired medication according to whatever information was provided	Several motors, containers, display, keyboard, speakers	Works on a predetermined medication schedule	Requires a proper a medication program

(continued)

Table 1 (continued)

Ref. No.	Work done	Hardware used	Advantages	Issues
[16]	Device for storing drugs and giving signals to the patient at particular intervals of time	Several motors, containers, display, keyboard, speakers	Gives notification for taking medicine	Signal should be reached to the user
[17]	The system has a disposable canister for storing drugs and delivering the drugs to a medicine delivery device including a medicine dispensing device for manually or automatically dispensing medicine upon a command	Belts, several motors, containers, wires	This system enables a pharmacist to provide a complete prescription in a cost efficient manner without manually handling the drugs	Requires many commands to operate
[18]	Using RFID to detect that an object has been accessed	RFID, scanners	RFID provides easy detection	–
[19]	It maintains two containers and ensures that the communication between the two containers is maintained. So that wirelessly after the first medicine is dispensed the second is ready	Several motors, several containers, communication modules	Wireless communication for achieving adaptive storing of medicines between two machines	Requires a proper communication path to function

5 Experimental Details and Implementation

We are proposing a system that consists of advanced sensors such as to monitor the MEDIBOX. The sensors such as temperature, humidity, motion, fire. The stock of the medicine can be easily updated using RFID system. In this proposed system, we are using PIC controller as the prime which controls the actuators. Here, we are using various sensors such as temperature sensor, RFID reader and tag. The temperature sensor senses the temperature of the MEDIBOX and PIR sensor senses the motion of the person, whenever the patient wants medicine he has to show the RFID tag. After showing the card, all the data will be displayed on the LCD. The each medicine will have tag after showing the card medicine remaining medicine present will displayed.

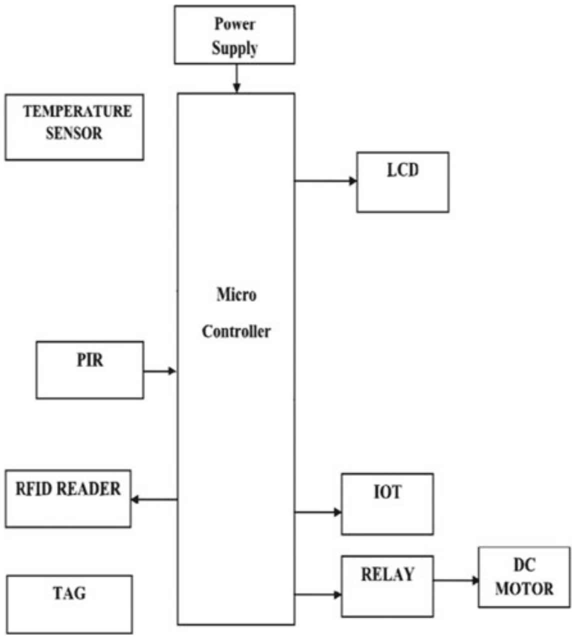


Fig. 1 Architecture of IoT-based medicine dispensary box

Humidity sensor senses the humidity value of the MEDIBOX. The current sensor senses the current value of the power given. All the values are updated to the IOT and displayed on the LCD. It is like a vending machine just the difference is instead of money we scan our RFID tag.

After development, remodeling and testing the project in different situations and conditions, detailed system-wise results are provided. Comments on obtained data, observed behavior, working accuracy, and reliability are also incorporated.

(i) Reminder System
It shows the data obtained through an SMS alert seen by the patient and the caretaker.

(ii) Patient Monitoring System
This is a real-time cloud environment in which the data is uploaded and monitored by the caretaker and the doctor.

(iii) Inventory System
This mechanism notifies the patient to restock the medicine supply before they run out. It is done in advance to make sure that there is adequate time to check the availability of medicines in the market and purchase them.

Figure 2 shows a snapshot of the Arduino IDE where we tried to implement the proposed algorithm to make the medical dispensary box.

Figure 3 shows the demonstration of the proposed architecture. It is a prototype model of the project design and implementation of smart medicine dispensary box: IoT.

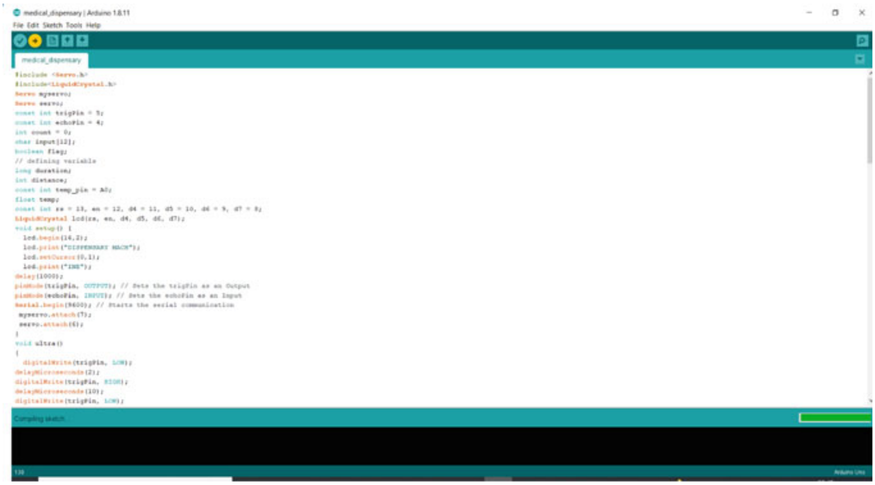


Fig. 2 Arduino IDE—code snapshot

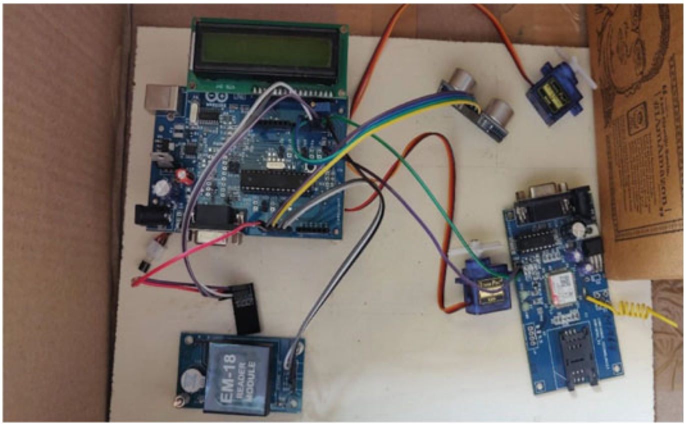


Fig. 3 Implementation of medicine dispensary box

6 Discussion and Future Directions

This paper explains and describes a MEDIBOX which can be used to help patients adhere to their meds on a daily basis. It is intended for senior people and those who have difficulty remembering medicine quantities and times. All of these issues are handled with the use of a variety of sensors, IoT, and cloud services. For data collection, the box contains a temperature module and an ultrasonic sensor. These sensors are mounted to the box and are used to detect whether it is open or closed. All of the collected data is sent to the cloud server using ESP8266 Wi-Fi module as a mode of connectivity. All of the collected data is then stored on a cloud server that can be accessed by the caretaker and the doctor to monitor the medications of the patient in real time. SMS alerts are sent to the patient and caretaker using the SIM800C GSM module. Depending on the action, it sends a SMS to the patient's mobile phone. Then, an SMS alert is also sent to the caretaker regarding if the patient really took the medication. An alarm is sent to the caretaker when the box is opened. The GSM module can also transmit a warning to the patient when the drug supply in the compartment or box is running low. When the ultrasonic detects a critically low level of medication, it triggers I/O and an SMS alert is sent to the patient, so that we can refill the supply of medicines.

The future work of this study is to create a box which can be utilized to enhance the comfort of a patient on a daily basis. It aims at elderly patients and also those who have problems with remembering which medicine to take at what time and at what quantity. All these problems are addressed using sensors, IoT, and cloud service. It should also send an SMS to the mobile number of the patient depending upon the action and also to the caretaker regarding if the patient really consumed the medication. The box should also send an alert to the patient when the supply of medications is low in the compartment. This project will drastically reduce the occurrence of a patient missing his medications. This in turn improves the efficiency of the effect of medicine on the disease. Therefore, after a lot of research work the researchers can say that this medical dispensary box will help us avoid unwanted medical complications caused due to untimely medication and due to any sort of medical negligence.

Here, for future directions in more elaborations authors have also reviewed several existing research papers and patents on IoT-based medicine dispensary box.

Table 2 summarizes future work over existing work for IoT-based medicine dispensary box.

Based on Table 2 and our experimental study, the future work of this research article is to design a multipurpose portable intelligent device in an efficient and effective way which will also help maintain the potency of the medicines. The medical dispensary box can be used for monitoring stores and controlling the functionality of the box for further analysis. This will in turn help in reducing the occurrence of a patient missing his medication which improves the efficiency of the effect of medicine on the disease.

Table 2 Summary of future work over existing work

Ref. No.	Existing work	Future work
[1]	This study focuses on several groups and did a study on them using medicines and qualitative analysis was done	Number of groups and sample cases can be increased for better and accurate results
[3]	In this study, the data is collected and analyzed. A data set was taken from January 1, 1992, to December 31, 2001. (A data set of 10 years is used)	Time span of data set can be increased to have better results
[9]	A vacuum operated system for individually dispensing medicine from storage to a user	Instead of ID password a user can provide other login credentials like fingerprint scanner or face recognition
[12]	A calculated dose medicine dispenser which determines dosage for prescriptive and over the counter medication as well as for providing the proper dose to the patient	For calculating, the dosage of medicine many different techniques can be used, and for example, a user can provide relevant data to help system calculate the dosage
[13]	Automatically dispensing medicines at particular time intervals and consist of a pill recovery system	After dispensing the medicine, the user can receive a notification alarm to take the medicine
[15]	In this study, information was already stored and the dispensary just dispensed the desired medication according to whatever information was provided	User interface for adding the predetermined medication can be improved and can be made more efficient and user friendly
[19]	It maintains two containers and ensures that the communication between the two containers is maintained. So that wirelessly after the first medicine is dispensed the second is ready	Better communication modules can be used to communicate between the two containers

References

1. Moen J, Bohm A, Tillenius T, Antonov K, Nilsson JL, Ring L (2009) Patient Educ Couns 74(2):135–141
2. Cartwright A (1990) Medicine taking by people aged 65 or more. Br Med Bull 46(1):63–76
3. Pound P, Britten N, Morgan M, Yardley L, Pope C, Daker-White G, Campbell R (2005) Resisting medicines: a synthesis of qualitative studies of medicine taking. Soc Sci Med 61(1):133–155
4. Williams A, Manias E (2014) Exploring motivation and confidence in taking prescribed medicines in coexisting diseases: a qualitative study. J Clin Nurs 34(34):471–481
5. Tordoff J, Simonsen K, Thomson WM, Norris PT (2010) “It’s just routine.” A qualitative study of medicine-taking amongst older people in New Zealand. Pharm World Sci 32(2):154–161
6. Elliott R, Goeman D, Beanland C, Koch S (2015) Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 10(3):213–221
7. Orlu-Gul M, Raimi-Abraham B, Jamieson E, Wei L, Murray M, Stawarz K, Stegemann S, Tuleu C, Smith FJ (2014) Public engagement workshop: how to improve medicines for older people. Int J Pharm 459(1–2):65–69

8. Fitzpatrick R (2005) Evaluation of an automated dispensing system in a hospital pharmacy dispensary. Pharm J 274
9. Rogers LW, Kraft TL, Berry JF, Kelley SA, Thompson III JA, Ober CD, Kuchar MC, Mayer Jr RR, Hoskins VW, Weido VC, Henckel MG (1995) Vacuum operated medicine dispenser. US Patent 5,405,048
10. Ostergaard J, Barfoed J, Olsen O, Geisler O (2004) Medicine dispenser. US Patent application 10/477,019
11. Brown JH, Brown John H (1980) Timed medicine dispenser. US Patent 4,207,992
12. Zamjahn DB (2005) Calculated dose medicine dispenser. US Patent 6,904,867
13. Ratnakar N (2010) Novation science holding LLC, "Smart medicine container". US Patent application 12/613,520
14. Christiansen LT, Christiansen Lee T (1988) Medication dispenser. US Patent 4,763,810
15. Shepherd CG, Agnew EG, Silberfeld M (1990) Dispenser. US Patent 4,911,327
16. Hafner D (1993) Device for the storage and time-regulated dispensing of drugs. US Patent 5,181,189
17. Lion N (1998) System for dispensing drugs. US Patent 5,838,575
18. Braun PJ, Sallustio J (2011) Proximity-based inventory management system using RFID tags to aid in dispensing and restocking inventory. US Patent 8,006,903
19. Walker JS, Mik M, Kobayashi M, Gelman GM, Sammon RP, Golden AP (2008) Methods and apparatus for increasing, monitoring and/or rewarding a party's compliance with a schedule for taking medicines. US Patent 7,366,675
20. Choubey DK, Kumar A, Srivastava K, Pahari S (2020) Notification and image analysis in cloud. In: International conference on emerging trends in information technology and engineering, 24–25 February 2020
21. Kumar S, Mohapatra UM, Singh D, Choubey DK (2020) IoT-based cardiac arrest prediction through heart variability analysis. In: Advanced computing and intelligent engineering: proceedings of ICACIE 2018, vol 2, pp 353–363

Emerging Smart Manufactory: Industry 4.0 and Manufacturing in India: The Next Wave

Rahul Priyadarshi, Apoorv Singh, Deepak Agarwal, Umesh C. Verma, and Anurag Singh

Abstract Today design-manufacturing industry is automated and digital, where machines communicate with every other. A clever factory equipped with statistics alternate in industrial, the Internet of Things (IoT) is the forthcoming, and specialists are shouting it uprising Industry 4.0. The future factory is going to comprise a new integrative, where now not only all manufacturing sources (sensors, actuators, machines etc.) are connected and trade data automatically and Industry 4.0 implies a whole communication network will exist between a range of companies, factories, supplier, logistics, resources, customers. Reports peg the clever manufacturing facility enterprise to contact \$215 billion through 2025, and there has been vital financial device in the world that is now embracing it. The contemporary globalization is confronted by way of the usage of the challenge to meet the always-growing global demand for capital and purchaser items by means of using concurrently making certain a sustainable evolve of human existence in its social. Industrial revolutions and future view of Industry 4.0 see the introduction of disruptive new applied sciences, the current developments of automation, and knowledge alternate that is opening to arise in manufacturing technologies. This consists of cyber-bodily techniques IoT, information analytics, cloud computing, and artificial Intelligence (AI). Implementation will advocate radically altering the documents framework and industry organization methods of your entire operation.

Keywords Industry 4.0 · Manufacturing systems · Cloud manufacturing · IoT · Sensor · Deployment

R. Priyadarshi (✉)
National Institute of Technology, Patna, Bihar, India
e-mail: rahul.glorious91@gmail.com

A. Singh · D. Agarwal · U. C. Verma · A. Singh
IET, Dr. Ram Manohar Lohia Avadh University, Ayodhya (Faizabad), Uttar Pradesh, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_32

1 Introduction

Since the principal industrial revolution, ensuing insurgencies have brought about extremist adjustments in assembling, from water and steam-controlled machines to electrical and advanced computerized creation. Assembling techniques have end up increasingly confounded, mechanized, and maintainable, which potential people can work machines basically, practically, and constantly [1–4]. India is amped up for getting Industry 4.0 and has taken a couple of exercises. According to IBEF, the Government of India plans to fabricate the responsibility of gathering region to 25% of gross domestic product (GDP) by 2025, from the current element of 16%. India is furthermore organized to defy overall test by accepted the Make in India programs. It is a great idea to go to lead the world with smart manufacturing. Regardless, with the presenting of monetary changes and the opening up of the Indian market, the circumstance has changed definitely. Head executive Mr. Narendra Modi has laid boost in collecting lead improvement, in his fight to increment manufacturing region "Make in India."

The points of the mission are as follows:

- To intensify fabricating region blast to 12–14% per annum over the medium term.
- To augment in the portion of assembling in the country's GDP from 16 to 25% with the guide of 2022, to make 100 million extra positions with the guide of 2022 in assembling area.
- To make beautiful ability sets among country travelers and the metropolitan negative for comprehensive development.
- To intensify home estimation expansion and innovative profundity in assembling.
- To enrich the global seriousness of the Indian assembling area.

While India's administration area has developed unmistakably over the excess few decades, most assembling areas are still stayed with innovation, hardware, and cycles from the Industry 2.0 phase, for example, guide inputs, absence of ICT combination in assembling, and essential holes in ability. The "Make in India" crusade desires to increment fabricating capacity and mechanical skill up degree. India's rising interest, great arrangement, and the opportunity to arrangement minimal expense plants are bringing Multi National Company) (MNCs), which will make a commitment to the expansion in fabricating area. In excess of 250 worldwide OEMs like GM, Renault Nissan, Boeing, GE, Daimler, and so forth have their exploration and improvement in India, and the full-size achievement of the IT region has made India pertinent to be an assembling center of the world [5–9].

These issues power the improvement of mechanical applied sciences for bringing down the workforce, shorting the developing season of the item, utilizing assets productively, etc., of which the cyber-physical system (CPS) and IoT are two super present day applied sciences progressed inside the leftover decade.

The improvement toward Industry 4.0 has these days an immense impact on the assembling business. It depends on the foundation of shrewd manufacturing plants, sharp product, and brilliant contributions installed in a trap of issue and of administrations also known as modern Web [10]. Moreover, new and troublesome undertaking models are advancing around these Industry 4.0 components [11].

With the improvement of these advancements, another idea, Industry 4.0, was once conveyed via German at some stage in the Hannover Fair match in 2011, which represents the building up of the fourth mechanical insurgency [12]. Since its first distribution, numerous European assembling query organizations and gatherings have delivered work on this subject, which underscores that under Industry 4.0, assembling will comprise of traded records and oversaw machines and creation devices showing up self-ruling and keenly in interoperable. Be that as it may, analysts hold various assessments of the particular prerequisites of Industry 4.0 and its achievement, performing on their number of mechanical innovative expertise capacities [13, 14]. It is evident that contemporary assembling is a summed up subject, which is explained in multi-fields. Hence, the contemporary comprehension of Industry 4.0 can't announce the standards. Likewise, the assembling still up in the air for a progressive strategy for innovative skill application, which will direct individuals to satisfy Industry 4.0.

Explaining industrial upheavals and future perspective on Industry 4.0 sees the presentation of troublesome new advancements, the cutting edge patterns of mechanization, and records substitute that is opening to show in assembling innovations. This incorporates cyber-actual systems, the Internet of Things (IoT), data analytics, cloud computing, and artificial intelligence (AI). Industry four is much of the time alluded to as the "fourth modern upset" and is set to in a general sense adjust the way we live, work, and identify with each other. Execution will propose fundamentally changing the records structure and business venture cycles of your whole activity [15].

2 Modern Manufacturing and Industry 4.0

The savvy plant and a portion of its crucial elements: availability, streamlining, straightforwardness, proactivity, and deftness. Every one of these angles can play a capacity in empowering decisions that are more proficient and can help associations upgrade the assembling system. Know that no two sharp industrial facilities will probably appear to be something very similar, and makers can focus on the number regions and elements generally pertinent to their specific necessities. The worldview of Industry 4.0 is in truth illustrated through three measurements [16–18], level mix all through the whole expense creation organization, start to finish designing across the whole item ways of life cycle, as appropriately as vertical coordination and arranged assembling frameworks. The level reconciliation all through the entire expense creation network portrays the cross-organization and friend's interior keen cross-connecting and digitalization of significant worth appearance modules during the expense chain of an item ways of life cycle and between esteem chains of bordering item ways of life cycles [19]. The start to finish designing all through the total item life cycle depicts the savvy cross-connecting and digitalization over the span of all periods of an item ways of life cycle: from the uncooked texture procurement to assembling framework, item use, and the item end of life [20]. The perceptive cross-connecting and digitalization covers the utilization of a start to finish answer the utilization of realities and correspondence applied sciences, which are installed in a cloud.

In an assembling framework, the clever cross-connecting is acknowledged via the utility of supposed cyber-physical frameworks (CPS), which are running in a self-coordinated and decentralized way [21, 22]. They are fundamentally founded on installed mechatronic perspectives, i.e., applied sensor frameworks for gathering data as appropriately as actuator frameworks for affecting actual techniques [23]. CPS is wisely connected with one another and is ceaselessly exchanging information by means of virtual organizations like a cloud continuously. The actual cloud is executed in the net of things and administrations [24]. Being part of a sociotechnical framework, CPS is the utilization of human–machine interfaces for cooperating with the administrator.

2.1 Industrial Internet of Things

Perhaps, the most hugest assortments of science that will make commitments to Industry 4.0 and smart assembling are the Industrial Internet of Things (IoT). The IoT is another transformation coming about because of the assembly of mechanical frameworks with cutting edge figuring, sensors, and pervasive correspondence frameworks. It is an extraordinary competition where incalculable mechanical gadgets, each noteworthy and new, are starting to utilize Internet protocol (IP) verbal trade advances [25].

The Industrial Internet of Things is a subset of what we have come to perceive as the Internet of Things (IoT). The IoT is a synopsis idea that catches a movement that began when we began coordinating processing and correspondence innovation into a significant number of the "things" that we use at homegrown and work. It started with the considering labeling and checking "things" with low-value sensor innovations, for example, radio recurrence recognizable proof (RFID) gadgets. In any case, the outlook changed as the market started giving over economical processing and Internet-based correspondence advances, at the same time with the vertical push of the omnipresent cell phone. This powerful coincidence of low charge figuring and unavoidable broadband systems administration has permitted the IoT to advance. Presently, the IoT comprises of a wide range of devices going from home machines, gentle bulbs, robotization frameworks, watches, to even our vehicles and trucks. In fact talking, the IoT is a progression of actual ancient rarities that involve implanted designs of electrical, mechanical, figuring, and correspondence components that empower Internet-based discussion and records trade.

The Industrial IoT follows a similar center meaning of the IoT, yet the issue and wants of the Industrial IoT are normally unique. A few instances of the "things" of the Industrial IoT incorporate gadgets like sensors, actuators, robots, fabricating units like processing machines, 3D-printers, and meeting line parts, synthetic blending tanks, motors, medical services gadgets, for example, insulin and imbuement siphons, and even planes, prepares, and vehicles. Surely, it is a huge range of gadgets.

Another term every now and again utilized while examining the Industrial IoT is functional innovation. Functional innovation (OT) alludes to the normal equipment, and programming frameworks found inside mechanical conditions. A few models incorporate programmable sound judgment regulators (PLC), disseminated control structures (DCS), and human–machine interfaces (HMI). These constructions are furthermore viewed as industrial control systems (ICSs) since they "control" the different cycles that occur inside a modern climate. These customary control structures are quickly starting to utilize Internet-based verbal trade advancements, so they can be incorporated into assembling associations' data innovation (IT) constructions and frameworks. This OT/IT incorporation development is as of now occurring in enormous scope across various businesses, and it offers a mechanical arrangement with the requirements of future brilliant assembling frameworks and Industry 4.0 [26–28].

2.2 Cloud-Based Manufacturing

Cloud-based assembling (CBM) is each other rising worldview that will make commitments radically to the accomplishment of Industry 4.0. CBM can be portrayed as an arranged assembling life-sized model that endeavors on-request admittance to a common assortment of different and apportioned assembling assets to shape transitory, reconfigurable digital actual creation lines which enhance proficiency, limit item lifecycle costs, and take into consideration most gainful valuable asset distribution because of variable-request buyer produced entrusting [29, 30]. Attributes of CBM include organized assembling, versatility, nimbleness, universal access, multi-tenure and virtualization, gigantic records and the IoT, everything-as-a-administration (e.g., foundation as-a-administration, stage as-a-administration, equipment as-a-administration, and programming as-a-administration), adaptability, and helpful asset pooling.

A comparable worldview to CBM has become perceived as cloud-based plan and assembling (CBDM) [31]. CBDM alludes to a more complete perspective on the item acknowledgment strategy, whereby the qualified components of the ordinary machine of plan and assembling assets are incorporated into distributed computing model. CBDM used to be toward the start characterized through Priyadarshi et al. [32] as follows: "Cloud-based design and manufacturing alludes to an item advancement model that empowers aggregate open development and fast item improvement with negligible expenses through long-range informal communication and publicly supporting constructions combined with shared help pools of plan and assembling resources and parts." Essentially, CBDM looks to enliven cloud-based assembling through incorporating it with cloud-based plan alongside the possibility of social item improvement.

2.3 The Perception and Methodology of Industry 4.0

There is an essential agreement among numerous analysts that the modern amendments require a long-term time of improvement and cowl the accompanying four viewpoints, considered as the future assembling dreams.

Factory, as one of the fundamental parts of Industry 4.0, the future plant will include a new integrative, where presently not just all assembling sources (sensors, actuators, machines, robots, transports, and so on) are associated and exchange information consequently, yet in addition, the plant will end up being careful and cunning adequate to anticipate and hold the machines, to deal with the creation interaction, and to control the assembling office framework. What is more, many assembling measures, for example, item plan, creation arranging, fabricating designing and creation, and administrations, will be mimicked as particular, and afterward connected intently start to finish, which implies these methods are not exclusively told with the guide of a decentralized gadget yet in addition oversaw reliantly. This assortment of future assembling office is recognized as a smart factory [33].

- Business and management, Industry 4.0 infers an entire correspondence organization will exist between a scope of organizations, production lines, provider, coordination's, assets, clients, and so forth. Each segment streamlines their arrangement progressively depending on the requests and acclaim of related areas in the organization, which makes the greatest pay for all cooperatives with the confined sharing assets. Moreover, the expenses and contamination, crude materials, CO_2 discharges, and so forth will be diminished. As such, the future business endeavor network is impacted through each participating segment, which could accomplish a self-coordinating status and communicate the constant reactions [34].
- Products, benefitting from Industry 4.0, will be another kind of item created in assembling that of shrewd items. These are stock inserted with sensors, recognizable parts, and processors which convey data and mastery to convey the useful direction the clients and communicates the utilizations criticism to the assembling framework. With these components, many capacities might need to be conveyed to the items, for instance, estimating the nation of product or clients, conveying this data, following the items, and examining the outcomes relying upon the data. Furthermore, a full creation data log can be installed with item helping item designer in streamlining the plan, the forecast, and the protection [35].
- Consumer, customers will likewise have a ton of benefits under Industry 4.0. Another buying strategy will be given to clients. It grants clients to arrange whatever components of items, with any assortment regardless of whether just one is. Furthermore, clients could trade their request and thoughts whenever during assembling even at a definitive moment with no charge. On the distinctive hand, the advantage from the savvy items empowers the customer not exclusively to know the creation realities of the item yet moreover to obtain the exhortation of usage relying upon their own practices [36]. Other than these arranged dreams of assembling, numerous scientists and organizations have been chipping away at Industry 4.0 in many fields around these standards [37]. Two average models

display the improvement of the Industry 4.0; which is fabricating line exhibition and shrewd items.

3 The Next Generation of Manufacturing Industry in India

In the most recent surveys presented out with the guide of rumored associations, India has demonstrated predictable upgrade in scores concerning fabricating skills just as various variables. For instance, India's evaluating among the world's 10 biggest assembling global nations has stretched out through three spots to 6th situation in 2015.

3.1 The Area of Market

India's assembling region can possibly contact US$1 trillion through 2025. Business specifications in the Indian assembling district continue to keep on being positive. While foundation and efficiency stay the key concerns, the accessibility of qualified gathering of laborers is a significant resource for India. Almost, 60% of its early stages in the age gathering of 18–35 ensures a segment profit for India, which if appropriately utilized, can launch India's assembling to more prominent statures [38].

As indicated by the World Economic Forum record in November 2016, India has hopped sixteen areas in the World Economic Forum's worldwide intensity positioning, an outcome of the brilliant way, in which the contemporary government is considered through financial backers. India positions 39 out of 138 nations. Israel positions 24. This is a far reaching get some answers concerning with a scope of columns like framework, governance, education, and so forth.

These outcomes are generally because of the new drives dispatched by present government. The previously mentioned report by WEF states that the most outstanding upgrades in India are in the fundamental drivers of seriousness looks good for the future, particularly for the improvement of assembling area [39].

3.2 Procedure for Robust Accomplishment

As indicated by a new article by Mr Arun Maira, ex Planning Commission Member, India will take a gander at following a three-pronged procedure to quickly move from Industry 2.0 to Industry 4.0.

- Strategy to further develop the business climate for assembling: The Make in India program looks to welcome Foreign Direct Investment in manufacturing to give this region a basic lift.

- Strategy to update human resource abilities and catch the “Segment Dividend:” preparing and skilling of employable workforce in the entwining of advanced and assembling regions is basic.
- Strategy of guaranteeing expedient execution: A policy is just comparable to its execution. India will try to jump the following upset of industrialization.

3.3 Government Administration Initiative

Depending on India’s power in Information Technology and an enormous labor force of IT experts, the groundbreaking ride of assembling by means of Industry 4.0 has effectively started in the country. Under the Government of India’s “Savvy Cities Mission,” the errands to fabricate 100 shrewd urban areas across India are being promoted as the trailblazers of the Industry 4.0 climate. With quick improvement in the fields of realities science and equipment, the world is going to observe a fourth modern unrest. The thought of “Industry 4.0” will change the manner in which India makes, plans, and repairs the items. Driven with the guide of the force of huge information, exorbitant registering limit, man-made brainpower and investigation, Industry 4.0 targets to totally digitize the assembling quarter [5]. The Department of Heavy Industry, Govt. of India is releasing many projects among industry groups, the scholarly community industry coordinated effort, just as supporting irreplaceable innovative ability intercessions through industry to produce ahead of its rivals.

As indicated by a document by the India Brand Equity Foundation (IBEF), the Government of India has an eager organization to make locally, upward of 181 product. The pass could help framework areas like force, oil and gas, and vehicle producing that require monster capital use and resuscitate the Rs. 185,000 crore (US$27.42 billion) Indian capital things business. With impulse on creating mechanical hallways and shrewd urban communities, the public authority aspirations to ensure comprehensive improvement of the country. The hallways would comparably help in incorporating, checking, and making favorable environmental elements for the mechanical improvement and will advance foster practices in assembling.

4 An Absolute Structure of Manufacturing for Industry 4.0 and Beyond

Modern groups, whether topographical or area explicit, are critical to speed up the R&D and Innovation. Bunches not just work with various partner collaboration, they additionally quick track the wide basing of development among the SME, which is the driving force of development in any country. Mechanical bunches in India are not overseen viably; however, numerous urban areas, for example, Pune, Bangalore, and Chennai do have phenomenal accumulation of organizations, exploration, and scholastic establishments. Most bunches are, best case scenario, a land recommendation that are run and overseen by private designers, and thus, R&D is certifiably not a focal component in the whole scene. Draconian work laws and an incredibly bulky land procurement framework renders up scaling of industries in India an unviable suggestion.

5 Conclusions and Suggestion

The arising economies need to use their endowments, for example, a bigger workforce, enormous business sectors, quickly developing economies, and beneficial specifications for assembling. A truly spread out technique, enunciation of the National Agenda for Manufacturing on a mission mode, assistance of biological system that speed up Innovation and R&D too, and quick execution of the instruments hence perceived will be very imperative in the new worldview. On the homegrown front, every one of us should recognize and focus on the couple of spaces of assembling where it enjoys an essential benefit and point of convergence on utilizing this addition on the worldwide stage. For instance, India's imposing achievement in the product business can be utilized to viably find its angle in computerized fabricating. The significant partners from academia, industry, government, labor force, and multipliers should be added by and large on a stage, and their commitments and action items really depicted to make new collusions with key accomplices from various worldwide areas should be fashioned. Mechanical and advancement bunches that speed up assembling abilities in all cases should be set up and sustained. Government wishes to make an empowering structure for industry to develop and work freedoms to prosper. At long last, creating with the guide of dominating and sharing is uncommonly significant. Systems administration structures both broadly and worldwide can be exceptionally helpful in encouraging world discourse.

This paper centers around the charm of Industry 4.0 in a creation framework and presents the normal assessments of Industry 4.0 and producing. Summing up in excess of a couple of viewpoints, the dominating standards of future assembling have been recognized to advise the exploration point. In the same manner as the whole business, there is an enormous hole between current industry and the satisfaction of Industry 4.0, which has been in actuality perceived in this paper. Likewise, a structure of Industry 4.0 is introduced, which distinguishes how explicit knowledge level advancements are acted inside three computerization of creation frameworks; from the system, it is obvious that the eventual fate of present day producing is filling toward Industry 4.0.

References

1. Wahlster W (2012) From industry 1.0 to industry 4.0: towards the 4th industrial revolution. In:

Forum business meets research
2. Hewitt PS (2002) Depopulation and ageing in Europe and Japan the hazardous transition to a labor shortage economy. Internationale Politik und Gesellschaft 1:111–120
3. Priyadarshi R, Gupta B, Anurag A (2020) Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues. J Supercomput 1–41
4. Lee J (2013) Industry 4.0 in big data environment. German Harting Mag 1(1):8–10
5. Lee J, Bagheri B, Kao H-A (2015) A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf Lett 3:18–23
6. Drath R, Horch A (2014) Industrie 4.0: hit or hype? [industry forum]. IEEE Industr Electron Mag 8(2):56–58
7. Priyadarshi R, Singh MP, Bhardwaj A, Sharma P (2017) Amount of fading analysis for composite fading channel using Holtzman approximation. In: 2017 fourth international conference on image information processing (ICIIP). IEEE, pp 1–5
8. Moktadir MA, Ali SM, Kusi-Sarpong S, Shaikh MAA (2018) Assessing challenges for implementing Industry 4.0: implications for process safety and environmental protection. Process Saf Environ Prot 117:730–741
9. Priyadarshi R, Rawat P, Nath V (2019) Energy dependent cluster formation in heterogeneous wireless sensor network. Microsyst Technol 25(6):2313–2321
10. CaR Hilger J (2013) Auto-ID integration-a bridge between worlds. German Harting Mag. 14–15
11. Priyadarshi R, Soni SK, Nath V (2018) Energy efficient cluster head formation in wireless sensor network. Microsyst Technol 24(12):4775–4784
12. Schmidt R, Möhring M, Härting R-C, Reichstein C, Neumaier P, Jozinović P (2015) Industry 4.0-potentials for creating smart products: empirical research results. In: International conference on business information systems. Springer, Cham, pp 16–27
13. Priyadarshi R, Singh L, Singh A (2018) A novel HEED protocol for wireless sensor networks. In: 2018 5th international conference on signal processing and integrated networks (SPIN). IEEE, pp 296–300
14. Priyadarshi R, Nath V (2019) A novel diamond–hexagon search algorithm for motion estimation. Microsyst Technol 25(12):4587–4591
15. Priyadarshi R, Soni SK, Sharma P (2019) An enhanced GEAR protocol for wireless sensor networks. In: Nanoelectronics, circuits and communication systems. Springer, Singapore, pp 289–297
16. Priyadarshi R, Rawat P, Nath V, Acharya B, Shylashree N (2020) Three level heterogeneous clustering protocol for wireless sensor network. Microsyst Technol 26:3855–3864
17. Priyadarshi R, Singh L, Singh A, Thakur A (2018) SEEN: stable energy efficient network for wireless sensor network. In: 2018 5th international conference on signal processing and integrated networks (SPIN). IEEE, pp 338–342
18. Priyadarshi R, Gupta B (2021) Area coverage optimization in three-dimensional wireless sensor network. Wireless Pers Commun 117(2):843–865
19. Priyadarshi R, Gupta B, Anurag A (2020) Wireless sensor networks deployment: a result oriented analysis. Wireless Pers Commun 113(2):843–866
20. Stouffer K, Falco J, Scarfone K (2011) Guide to industrial control systems (ICS) security. NIST Spec Publ 800(82):16–16
21. Priyadarshi R, Ranjan Kumar R (2021) An energy-efficient LEACH routing protocol for wireless sensor networks. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 423–430
22. Anurag A, Priyadarshi R, Goel A, Gupta B (2020) 2-D coverage optimization in WSN using a novel variant of particle swarm optimisation. In: 2020 7th international conference on signal processing and integrated networks (SPIN). IEEE, pp 663–668
23. Lucke D, Constantinescu C, Westkämper E (2008) Smart factory—a step towards the next generation of manufacturing. In: Manufacturing systems and technologies for the new frontier. Springer, London, pp 115–118

24. Carvalho N, Chaim O, Cazarini E, Gerolamo M (2018) Manufacturing in the fourth industrial revolution: a positive prospect in sustainable manufacturing. Proc Manuf 21:671–678
25. Priyadarshi R, Gupta B (2020) Coverage area enhancement in wireless sensor network. Microsyst Technol 26(5):1417–1426
26. Schlechtendahl J, Keinert M, Kretschmer F, Lechler A, Verl A (2015) Making existing production systems Industry 4.0-ready. Prod Eng Res Devel 9(1):143–148
27. Rawat P, Chauhan S, Priyadarshi R (2020) Energy-efficient clusterhead selection scheme in heterogeneous wireless sensor network. J Circuits Syst Comput 29(13):2050204
28. Finance ATCC (2015) Industry 4.0 challenges and solutions for the digital transformation and use of exponential technologies. In: Finance, audit tax consulting corporate: Zurich, Swiss pp 1–12
29. Shrouf F, Ordieres J, Miragliotta G (2014) Smart factories in Industry 4.0: a review of the concept and of energy management approached in production based on the internet of things paradigm. In: 2014 IEEE international conference on industrial engineering and engineering management. IEEE, pp 697–701
30. Wittenberg C (2015) Cause the trend industry 4.0 in the automated industry to new requirements on user interfaces? In: International conference on human-computer interaction. Springer, Cham, pp 238–245
31. Priyadarshi R, Yadav S, Bilyan D (2019) Performance analysis of adapted selection based protocol over LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 247–256
32. Priyadarshi R, Yadav S, Bilyan D (2019) Performance and comparison analysis of MIEEP routing protocol over adapted LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 237–245
33. Priyadarshi R, Soni SK, Bhadu R, Nath V (2018) Performance analysis of diamond search algorithm over full search algorithm. Microsyst Technol 24(6):2529–2537
34. Soni SK, Kumar S, Priyadarshi R (2018) Energy-aware clustering in wireless sensor networks. In: International conference on nanoelectronics, circuits and communication systems. Springer, Singapore, pp 453–461
35. Kumar S, Soni SK, Priyadarshi R (2018) Performance analysis of novel energy aware routing in wireless sensor network. In: International conference on nanoelectronics, circuits and communication systems. Springer, Singapore, pp 503–511
36. Policy CF (2010) Department of industrial policy and promotion ministry of commerce and industry government of India
37. Rawat P, Chauhan S, Priyadarshi R (2021) A novel heterogeneous clustering protocol for lifetime maximization of wireless sensor network. Wireless Pers Commun 117(2):825–841
38. Sateesh VA, Kumar A, Priyadarshi R, Nath V (2021) A novel deployment scheme to enhance the coverage in wireless sensor network. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 985–993
39. Singh L, Kumar A, Priyadarshi R (2018) Performance and comparison analysis of image processing based forest fire detection. In: International conference on nanoelectronics, circuits and communication systems. Springer, Singapore, pp 473–479

GIS-Based Rain Water Harvest Considering Population as a Function

Pranab Hazra, Payel Mukherjee, Archishman Nath, Manish Kumar, Isha Bhattacharjee, and Sinthia Mukherjee

Abstract Pure water is the world's first and foremost medicine. So, if crisis happens, our society will be destroyed because a drop of water is more important than a sack of gold for a thirsty man. Not only drinking, but, water is also used day-to-day life in household as well as outdoor field work and various domestic works. In West Bengal, water is very essential in agricultural field. Here, ground water is limited, and now, the water level is gradually decreasing with the increasing of the population. Ground water is the one and only source of our drinking water. For preserving our ground water for the near future, the best alternative way is to utilize our monsoon season and store rain water as much as we can. If rain water is harvested properly, then it can be used in agricultural field as well as many household works so that usage of ground water decreases. There is a large variation in the amount of rainfall in the different parts of West Bengal annually. This is where the concept of rainwater harvesting comes into effect. In this paper, we describe the actual rain fall scenario in monsoon season (including pre- and post-monsoon) of West Bengal with compare to population which shows the requirement of water district wise. From this concept, we will be able to understand how much rain water harvesting is required and why it is so important? Using remote sensing and QGIS technique, we analyze district-wise rainfall data of West Bengal for a particular period of time (2002–2010). We also observed the population scenario of every district for the same period of time. For both the cases, we prepare a thematic maps. Efforts are given to make observation from the thematic that how much this process is beneficial for the society.

Keywords Remote sensing · QGIS · Rainwater · Monsoon season · Population

P. Hazra · P. Mukherjee (✉) · A. Nath · M. Kumar · I. Bhattacharjee · S. Mukherjee
Electronics and Communication Engineering, Narula Institute of Technology, 81, Nilgunj Road, Kolkata 700109, India
e-mail: payelmukherjee1798@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_33

1 Introduction

Water is the main driving force of our nature. The primary source of water in the dry areas and deserts and even in the semi-dry regions is ground water because of its ability of getting less affected by pollution compared to the surface water in most of the areas of West Bengal [1]. The climatic condition of West Bengal ranges from savannah (usually warm climatic condition with temperature from twenty degree to thirty degree) to mid subtropical climate (i.e., summer and winter) [2]. The rainy season in West Bengal is from July to the end of September. The annual rainfall differs in different parts of the state where the highest rainfall of 200 to 400 cm is received by the North Bengal; almost, 200 cm is poured in the coastal area in the Ganga plain, and the middle or the central part of the state receives a rainfall of about 150–200 cm, and an amount of 100–125 cm of rainfall is experienced by the western plateau region [3]. The importance of rainwater harvesting can be understood from the increasing demand for irrigation and the unsustainable practices of groundwater utilization. It has been found that the farmers of the Bankura and Purulia districts would face a difficult problem in the upcoming years due to the scarcity of water, and the only remedy to this major problem is the utilization of rainwater.

In this way, the drops of rain are collected and stored it for the purpose of future uses instead of wasting. Water harvesting is an important factor to preserve the natural resources [4]. The case study in Soankhad, Punjab water shed used GIS to find the appropriate place for water harvesting. The usage of thematic maps, hydrological maps were done for selecting appropriate site to construct the water harvesting structures [4]. The ideal places are rivers or roofs from where rainwater can be collected easily, and it can be stored in deep well, any reservoir, shaft, or in aquifer [5, 6]. The rooftop rainwater harvesting has already gain much popularity in India and has been inbuilt in many housing complexes or institutes to recharge the aquifers. It is done to maintain the water balance by precipitation and evapotranspiration [7]. We can use this water for gardening purposes, for domestic use after some treatment, heating purpose for houses, etc. [5, 6]. For drinking purpose, the harvested water can also be used (after purification) [5]. It is a very judicious process of rainwater utilization.

Making available the harvested rainwater in the dry season plays a major role for agricultural purposes and livelihood of mankind [8]. It enables the farmers to grow a type of crop called Rabi crop during winter season [8]. The funding by MGNREGS focused on the design of hapas to cover 5% of the land to provide irrigation to paddy crops in kharif season as the SC and ST depended on the rain fed agriculture [8].

Ground water is a disguised natural resource [9]. However, to achieve accurate results, integrated studies using satellite image data and geographical information system (GIS) tools are used [9]. Previously, many researches have been done their work regarding remote sensing and GIS tools to define spatial distribution of ground water potential zones on the basis of various parameters [9].

In urban areas, rainwater is collected in paved areas and at the top of the roof [10]. With the expanding population, the freshwater demand is also increased day by day. Only, groundwater source cannot fulfill public demand on regular basis. So, the metropolitan cities specially faced some deficit of ground water supply [10]. Seasonal fluctuations, which is the most common, also affect sometimes [10]. A case study in Tamil Nadu, India shows the benefits of tank system to harvest rainwater through the earthen embankments called bunds [11]. It provides environmental as well as economical benefits to farmers [11].

2 Study Methods

We are using geographical information system (GIS) technique for this project. This technique is very helpful and has many advantages. It helps in producing and providing information in spatial and temporal domains which is very much important and essential for getting accurate result and analysis [12]. It is such a tool which works with different types of data and combines them to work into a single domain so as to analyze and solve the problems and get perfect correct result and observation [13]. The remote sensing technique helps us to improve the characteristics of the land surface to increase the hydrological studies and advancement of remote sensing [14].

2.1 Study Area

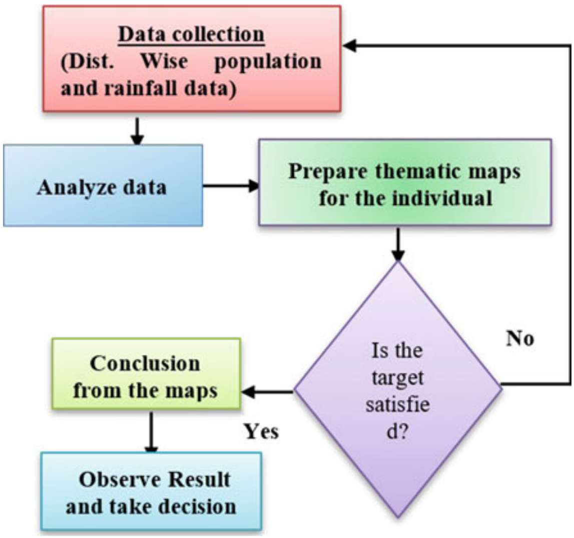
West Bengal lies in between 85° 55′ and 89° 55′ E and 20° 25′ and 27° 13′ N [15].

2.2 Representation

See Fig. 1.

Fig. 1 Study area of India [marked as yellow]

2.3 Workflow of Study



See Table 1.

Table 1 District-wise population data [16, 17]

S. No.	Dist.	Population
1	Bankura	3,596,292
2	Bardhhaman	7,723,663
3	Birbhum	3,502,387
4	Kolkata	4,486,679
5	Darjeeling	1,842,034
6	Haora	4,841,638
7	Hooghly	5,520,389
8	Jalpaiguri	3,869,675
9	Coach Bihar	2,822,780
10	Malda	3,997,970
11	Purba Medinipur	5,094,238
12	Paschim Medinipur	5,943,300
13	Murshidabad	7,102,430
14	Nadia	5,168,488
15	North 24 Parganas	10,082,852
16	Puruliya	2,927,965
17	South 24 Parganas	8,153,176
18	Uttar Dinajpur	3,000,849
19	Dakshin Dinajpur	1,670,931

2.4 Thematic Map of Population

See Fig. 2.

2.5 Thematic Map of Total Rainfall

See Fig. 3.

3 Observation

As we know that population increasing rapidly, so the use of water consumption is also increasing.

As per the availability and reliability of the information, we have taken into consideration the years from 2002 up to 2010 and hence analyzed the data.

Finally, we have plotted the values of rainfall in millimeter for each district (from Table 2). From there, we have found that population of many districts are very high,

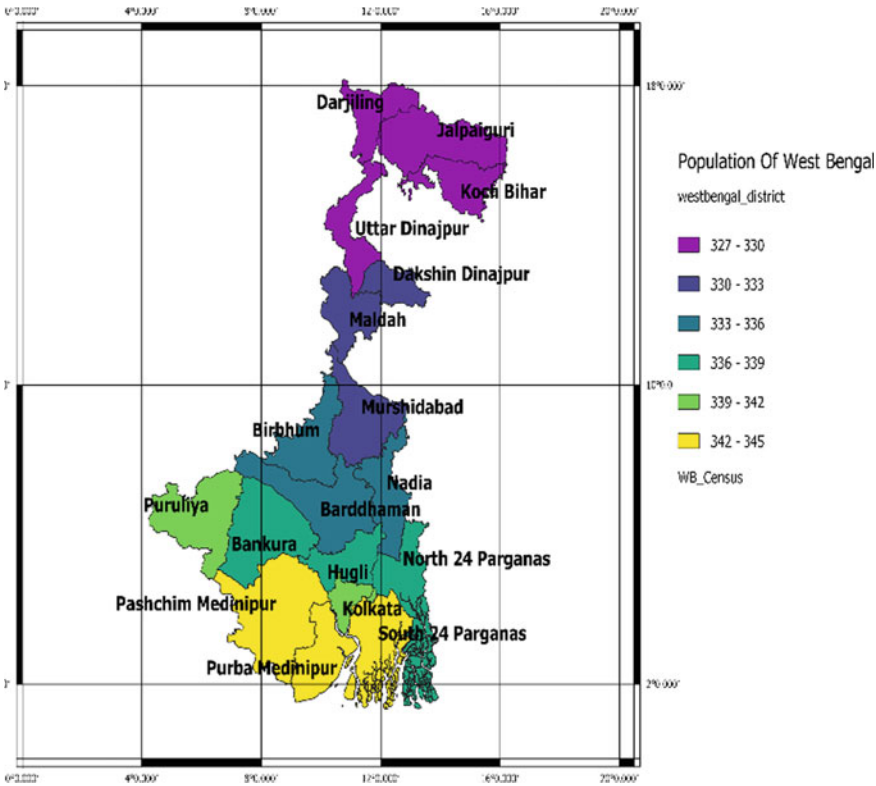


Fig. 2 District-wise population in West Bengal

but rainfall in that particular district is inadequate to fulfill the daily requirement; therefore, people are facing many problems in their daily livelihood instead of having ground water because it is not possible for all specially in remote village to use the ground water. Again, we have got the information from both the tables (Tables 3 and 4) that in some district, population is low, and rainfall amount is high.

We know that groundwater level is decreasing rapidly due to high pressure of population and industrialization. So, people are facing the problems of water shortage for their daily requirements. So, we can try to overcome their problems by proper maintenance of water using rainwater harvesting technique and collecting and supplying excess amount of water from one district to the other district where water is inadequate. In this way it could be maintained a proper balancing of water in the West Bengal districts. We must distribute excess amount of water from one water surplus district to a water deficient district.

We have found that the North 24 Parganas and South 24 Parganas are highly populated districts in West Bengal (from Table 3), and from Table 4, it is found that rainfall in that districts is inadequate to fulfill the daily necessities of the residents.

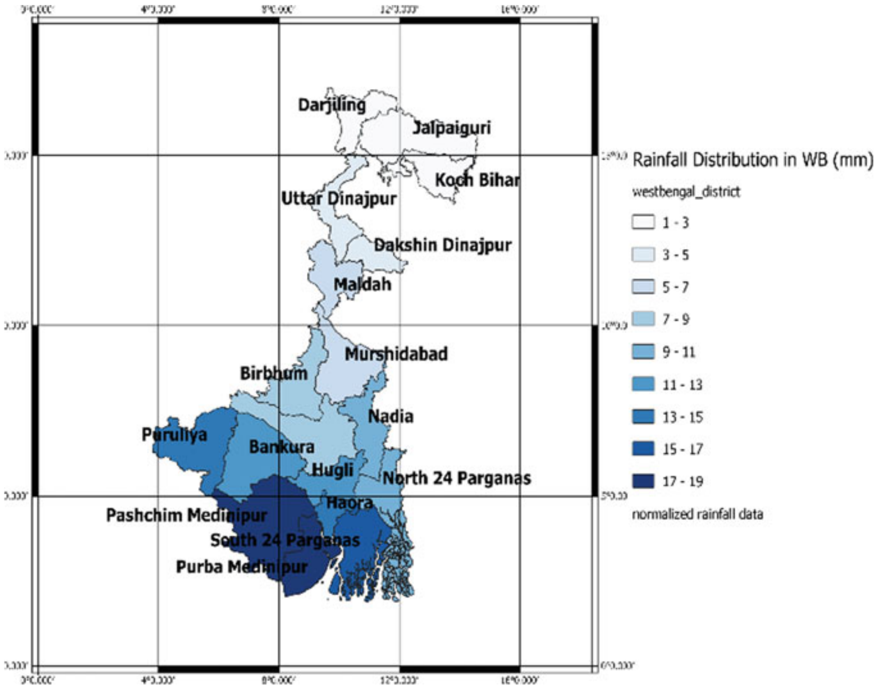


Fig. 3 District-wise rainfall in West Bengal (in normalized condition)

We also found that many districts like Birbhum, Murshidabad, Hooghly, Bankura, Nadia, Puruliya have population much less than the North 24 Parganas and South 24 Parganas (from Table 4), and the rainwater amount in that districts is much higher than North 24 Parganas and South 24 Parganas (from Table 3). So, we can supply excess amount of water from the abovementioned districts to the North 24 and South 24 Parganas for the better daily livelihood. We are also considering the neighboring districts and comparing the water data and population data of the neighboring districts, and also, we can supply extra amount of water from one district to the district which are facing problems of availability of water which will very fruitful for West Bengal population. We also observed that the population of Kolkata is also high (from Table 4), but there is lack of rainfall amount in Kolkata (from Table 3); the water table shows that there is a water shortage problems arising in Kolkata, so we will do supply of water from neighboring district to Kolkata so that this district can also sustain their livelihood without facing the problems regarding water.

Comparing the population and rainfall in Kolkata (refer to Tables 3 and 4), we find that the rainfall in Kolkata is too less compared to the population density. Hence, the neighboring district Howrah can supply water to Kolkata during its need.

Again, we can see Birbhum has much higher amount of rainfall (from Table 3), but its population density is not much (from Table 4 and Fig. 2), so it may supply water to its neighboring districts like Murshidabad and Bardhhaman. But again, these two

Table 2 District-wise rainfall data [16–19]

S. No.	Dist.	Total rain fall (in mm)	Normalized rain fall
1	Bankura	10,273.01	7.321
2	Bardhhaman	14,492.6	10
3	Birbhum	12,262.8	8.584
4	Kolkata	317.04	1
5	Darjeeling	1191.77	1.555
6	Howrah	2005.86	2.072
7	Hooghly	9184.04	6.629
8	Jalpaiguri	3313.83	2.902
9	Coach Bihar	2256.5	2.231
10	Malda	5973.5	4.591
11	Purba Medinipur	6504.06	4.928
12	Paschim Medinipur	7534.08	5.582
13	Murshidabad	10,862	7.695
14	Nadia	6862.37	5.156
15	North 24 Parganas	6530.59	4.944
16	Puruliya	8475.01	6.179
17	South 24 Parganas	5825.08	4.497
18	Uttar Dinajpur	1476.15	1.736
19	Dakshin Dinajpur	1820.46	1.954

districts get good rainfall, so the excess water stored by Birbhum can be supplied to farther districts like Malda.

4 Discussion and Calculation

Throughout our study, we have calculated the total population of each and every district of West Bengal (Fig. 2), plotted them (using QGIS software), and also, the total rainfall of West Bengal's 19 districts has been estimated (Fig. 3). The population density chart is as follows in decreasing order—North 24 Parganas, South 24 Parganas, Bardhhaman, Murshidabad, Paschim Medinipur, Hugli, Nadia, Purba Medinipur, Haora, Kolkata, Malda, Jalpaiguri, Bankura, Birbhum, Uttar Dinajpur, Puruliya, Koch Bihar, Darjeeling, Dakshin Dinajpur (as per Table 4). Hence, we

Table 3 District-wise rainfall

S. No.	District-wise descending order rainfall
1	Bardhhaman
2	Birbhum
3	Murshidabad
4	Bankura
5	Hugli
6	Puruliya
7	Paschim Medinipur
8	Nadia
9	North 24 Parganas
10	Purba Medinipur
11	Malda
12	South 24 Parganas
13	Jalpaiguri
14	Koch Bihar
15	Haora
16	Dakshin Dinajpur
17	Uttar Dinajpur
18	Darjiling
19	Kolkata

observe that Dakshin Dinajpur has the least population among all the districts, and North 24 Parganas has the highest population (from Table 4).

Following that we gathered data on rainfall in all of West Bengal's districts. We have calculated the total rainfall in each district. Then, we have normalized the data in a range of 1–10 (refer to Fig. 3) and arranged sequentially in Table 3. The normalization is done and applied in the data using N_x

$$N_x = a + \frac{(x - A)(b - a)}{(B - A)}$$

A Smallest value in dataset (here $A = 317.04$)
B Largest value in dataset (here $B = 14{,}492.6$)
a Normalized scale minimum (here $a = 1$)
b Normalized scale maximum (here $b = 10$)
x Any number from the dataset whose value needs to be normalized
N_x Normalized value

Based on the data in Tables 1 and 2, we have enlisted the rainfall received by the districts in descending order against the descending order arrangement of the districts in terms of population density.

Table 4 Mapping of population

District-wise descending order population
North 24 Parganas
South 24 Parganas
Bardhhaman
Murshidabad
Paschim Medinipur
Hugli
Nadia
Purba Medinipur
Haora
Kolkata
Malda
Jalpaiguri
Bankura
Birbhum
Uttar Dinajpur
Puruliya
Koch Bihar
Darjiling
Dakshin Dinajpur

5 Result

The process by which we can take our decision which district is most appropriate for harvesting. The rain water is cost-effectively which is shown in Fig. 4.

6 Conclusion

From our observation, it can be said that higher the population density greater is the water demand. Again, a high-water demand and less rainfall may lead to water scarcity. Similarly, a high rainfall and less population imply an area with surplus water availability. The excess water that is harvested can be collected and transported to those areas suffering from water scarcity. Hence, it can provide an independent water supply during regional water restrictions like drought. It can mitigate flooding of low lands. It will help in flourishing of agricultural lands and reduce the crisis of potable water during summer. It will help to sustain ground water levels and thus ensure availability of potable water, clean, and free from salinity.

Our paper is demanding a high-value publication because we are considering the approach of demand side and supply side and also some factors (average rainfall,

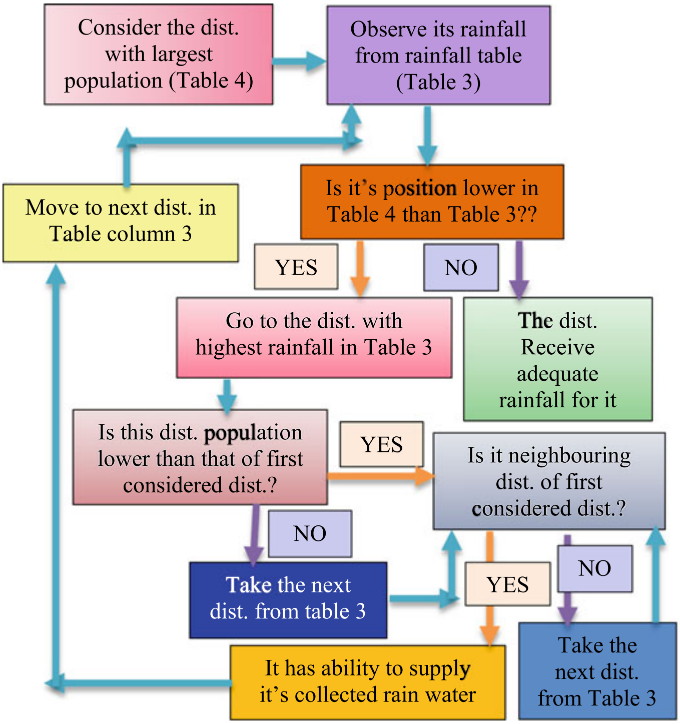


Fig. 4 Process for taking decision to supply harvested rainwater in crisis area

population, and distance between the districts) which are very much important for determining the volume of rainwater storage tanks with optimum capacity at viable cost.

Rainwater harvesting and distribution of harvested rainwater considering population density and rainfall as its key factor is our prime area of concern in this paper. It is very important to glance a light on population while attempting to rainwater harvesting. If we don't consider population, then the crisis and demand of water in an area cannot be judged properly. Hence, this comparative study of rainfall with population contributes an uniqueness to this article.

7 Future Scope

As a future prospect, we have thought of supplying excess water to the neighboring region during urgent need such as drought and flood.

There are many districts in West Bengal such as Dakshin Dinajpur which suffer from lacking of fresh water sometimes, mainly during flood, when the sea water percolates the soil and mixes with the ground water and thus making it unsuitable for human consumption. Then, the harvested rain water will provide some relief.

Again, there are districts having large agricultural fields like Bankura, Purulia, Nadia, Birbhum, and West Bardhaman, but the farmers living there may also suffer from water deficiency during summer. Here, also, harvested rain water will give them some relief.

Now, in the further advancements, we will consider the fresh water bodies present in the districts—be it river or lake, and the ground water level and modify our distribution system of the harvested rainwater accordingly. As we know, the population across the world is increasing with great rate, and it affects the whole availability of rainwater and groundwater on the earth surface [20]. It affects the area of aridity and semi-aridity and the farmers in regions where annual rainfall is relatively low. Also, in this region, the soil moisture is not good for growth of crops and plants [20]. In this region, we can also use the method of rainwater harvesting, based on groundwater recharge, for storage of water [20].

Further, this system of harvesting and inter-district distribution of harvested rainwater will prove highly useful to states like Tamil Nadu. As we all know, Tamil Nadu receives rainfall only twice a year [21]. So, in order to augment ground water resource, it has been made mandatory to provide rain water harvesting structures in its buildings. In Tamil Nadu, out of 2,392,475 buildings in Town Panchayat, 2,294,342 buildings have been provided with rainwater harvesting facilities [21]. In addition to it, if the proper authority incorporates our technique of rainwater harvesting and distribution from one water surplus district to another water deficient district, then they will be able to combat the problem of water scarcity more efficiently.

Acknowledgements The authors thankfully acknowledge the support of the Narula Institute of Technology, Agarpara, and Kolkata-700109, India in carrying out this work. Authors are also grateful to the respected editor and anonymous reviewer of this paper for their valuables comments and suggestions; the inclusion of those has sufficiently improved this revised version.

References

1. Mondal I, Bandyopadhyay J, Chowdhury P (2018) A GIS based DRASTIC model for assessing groundwater vulnerability in Jangalmahal area, West Bengal, India. Springer International Publishing AG, part of Springer Nature
2. Climate (2006) West Bengal: land. Suni System (P) Ltd. 2006-09-05
3. Baisakhi K (2006) Glossary of meteorology. Am Meteorol Soc
4. Singh JP, Singh D, Litoria PK (2009) Selection of suitable sites for water harvesting structures in Soankhad watershed, Punjab using remote sensing and geographical information system (RS & GIS) approach—a case study. J Indian Soc Remote Sens 37:21–35
5. Eccles A (2017) Rainwater harvesting for livestock
6. Uddipta G, Ranajit BK (2016) Study on scope of roof top water harvesting for recharging ground water for combating excess ground water withdrawal for rice cultivation in the district of Nadia in West Bengal 9(6):987–993

7. Jasrotia AS, Majhi A, Singh S (2009) Water balance approach for rainwater harvesting using remote sensing and GIS techniques, Jammu Himalaya, India. Water Resour Manage 23:3035–3055
8. Agricultural water management learning and discussion (2012)
9. Gouri BS, Sailesh S, Dilip PK, Babita P (2012) Assessment of ground water potential zone in Paschim Medinipur district, west Bengal—a Meso-scale study using GIS and remote sensing approach 2(5)
10. Sen P (2012) Implementing rainwater harvesting methods—a study in Baishnabghata-Patuli, Kolkata, India 5(1):01–05
11. Bitterman P, Tate E, Meter KJV, Basu NB (2016) Water security and rainwater harvesting: a conceptual framework and candidate indicators. Appl Geogr 75–84
12. Thapa R, Gupta S, Kaur H (2017) Delineation of potential fluoride contamination zones in Birbhum, West Bengal, India, using remote sensing and GIS techniques. Arab J Geosci 10:527
13. Mondal D, Pal S (2015) A multi-parametric spatial modeling of vulnerability due to arsenic pollution in Murshidabad district of West Bengal, India. Saudi Soc Geosci
14. Fenta AA, Kifle A, Gebreyohannes T, Hailu G (2015) Spatial analysis of groundwater potential using remote sensing and GIS-based multi-criteria evaluation in Raya Valley, northern Ethiopia. Hydrogeol J 23:195–206
15. Groundwater arsenic contamination in west Bengal—India. School of Environmental Studies, Jadavpur University (2006)
16. india.gov.in.
17. data.gov.in.
18. http://cgwb.gov.in/GW-data-access.html
19. https://www.researchgate.net/
20. Rajasekhar M, Gadhiraju SR, Kadam A, Bhagat V (2020) Identification of groundwater recharge-based potential rainwater harvesting sites for sustainable development of a semi-arid region of southern India using geospatial, AHP, and SCS-CN approach. Arab J Geosci 13:24
21. Sarkar AB, Sarkar B, Sengupta K (2020) Madhyamik geography and environment

Motion Correction of MR Images Using Cross-Guided Bilateral Filter

Shiju Samuel, Rohini S. Ochawar, and M. S. S. Rukmini

Abstract Motion artifact has been the foremost concern in the field of medical imaging (MI). These artifacts compromise image quality and interfere in image interpretation, especially in MRI applications with low signal-to-noise ratio. In motion artifact removal process, the most demanding problem is to protect the data bearing structures such as edges and surfaces to get good visual quality while enhancing peak signal-to-noise ratio (PSNR). Application of filters on these motion artifact affected MRI signals yields better performance in terms of removal of artifact and consequently correct interpretation by the radiologist. The paper presents the analysis on the use of various filters/techniques for motion artifact removal. Motion artifact removal using cross-guided bilateral filter (CGBF) is experimented, and the results are discussed with various evaluation parameters. Motion correction using CGBF is compared with the existing techniques, and the performance of CGBF is found to be superior.

Keywords Magnetic resonance imaging · Motion artifact · Medical imaging · Filter techniques

1 Introduction

Different kinds of imaging techniques are introduced in healthcare sector, and therefore, implementation of various methods of image processing algorithms on these biomedical images for both diagnostic and research purpose offers guidance to physicians in personalized care [1]. The techniques used in the healthcare sector to capture the internal structures of the human body are positron emission

S. Samuel (✉) · M. S. S. Rukmini
Department of Electronics and Communication Engineering, Vignan's Foundation for Science, Technology and Research, Vadlamudi, Guntur, Andhra Pradesh, India

R. S. Ochawar
Department of Electronics Engineering, Shri Ramdeobaba College of Engineering and Management Nagpur, Nagpur, Maharashtra, India
e-mail: ochawarrs@rknec.edu

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_34

tomography–computed tomography (PET-CT), photo-acoustic imaging, molecular imaging, computed tomography, magnetic resonance imaging (MRI), ultrasound, and fluoroscopy. Among all imaging techniques, MRI is preferred for capturing the physiological function and anatomical structure, and also it has the benefit of being non-ionizing and non-invasive technique; therefore, while taking MRI, the health of the patient is not affected [2, 3]. Magnetic resonance imaging (MRI) has the capability to visualize the detailed internal structure of the human body [4].

Moreover, the capability of MRI is excellent in visualizing the structure of soft tissues; hence, it is used to analyze the cardiac function, angiography, and brain function too. Likewise, it can deliver the detailed image of bones, human organs, and all other internal body structures [5]. Any abnormality in human body caused due to bone problems can be investigated with the MRI technique along with other imaging methods [6, 7]. The quality of MRI images is affected by various noise sources, for instance, receiver bandwidth, the presence of background tissue, RF coil in MRI scanning machine, breathing motion, ambient noise from the environment, recording medium, transmission medium, measurement errors, acquisition noise from the equipment, body fat, patient motion, RF pulses, and field strength [8]. From the above-mentioned artifacts, motion artifact is considered as the most potential artifact in various applications that can reduce the image quality [9]. Blurring of the image content and aliasing along the phase-encoding direction have been experienced during the long MR examination time because of the physiological behavior and patient motion. Therefore, the reconstruction is vital in MRI to reduce the artifact caused by motion [10]. At present, various de-noising methods are implemented on MRI to enhance both the signal-to-noise ratio (SNR) and image quality [11, 12].

Motion artifact deals with image resolution and reframing, mainly in applications of magnetic resonance imaging (MRI) with small signal-to-noise ratio (SNR) such as diffusion tensor imaging or MRI. High resolution image has high tendency to motion artifact, and scan time repeatedly provokes motion artifacts. Throughout the scan time of high-speed imaging techniques and its progression patient positioning, recipient coils and proper indication can reduce motion artifacts. Physical noise origin is due to motion from inhalation, movement, and pulse, from the devouring reflex and small unplanned head activity.

Motion artifacts for logical MR imaging can be considered with two important categories: macroscopic motion and microscopic motion. Microscopic includes blood circulation and water dispersing like MR in angiography and dispersing MR imaging. Macroscopic refers to patient motion associated with motion of the imaging object (if this involves extremities, head) or physical movement induced by simple functions of body.

Motion for fetal and neonatal MRI is expected due to maternal respiratory motion and unpredictable due to fetal circulation in uterus or neonatal head circulation. Motion may remain and continue during the MR imaging acquisition (example directionless neonate), periodic (example respiratory motion), or irregular (example produced by MRI scanner aural noisy). In MRI motion, neonatal and fetal brain is considered to be rigid body, with nominal or no disfigure and with all measurements of the imaging aim being conserved. Motion can be constricted in a 2D field with revolving and interpreting possible different divisions. Motion is the most difficult to recompense. The fetal and neonatal motion (through amplitude in few centimeter) can be rather greater than that of adults (through amplitude in few millimeter) and scaling better than the original dimensions from the anatomy of interest, and its results on particular image standard are harmful.

2 Generalized Analysis on Different Motion Artifact

Some of the recently published work related to the reduction of motion artifact in MRI is listed below:

In MRI, patient motion is utmost the origin of artifacts. Image blurring caused by patient motion frequently leads to limit MRI image quality. Head movement has become a controlling factor in MR neuroimaging approach. The body motion correction problem is studied by Johnson and Drangova [13] using conditional generative adversarial network (CGAN). The objective is to improve the MRI image quality and to compare with motion artifact affected images. In order to remove the effect of motion in MR images, it is suggested that a set of gray hidden layer with three-dimensional feature vector must be applied on the MRI image patches (example: $3 \times 3 \times 3$ brain image). This coding system consists of three-dimensional convolutional network (example: $3 \times 3 \times 3$ filter) with the three-dimensional maxpooling algorithm; after applying the convolution with up-sampling algorithm, the received MRI image can run with concatenate algorithm; finally once again applying the three-dimensional convolution algorithm (example: $4 \times 4 \times 4$ filter), the data is experimented with the CGAN network.

Saladi and Amutha Prabha [14] conducted a survey of applying de-noise filter algorithms on MRI images. Various de-noising filter algorithms are implemented by researchers like non-local mean, principle component analysis, bilateral, and spatially adaptive non-local means (SANLM) filters. Modified survey of all mentioned algorithm was conducted with the performance evaluation with various metrics like signal-to-noise ratio (SNR), peak signal-to-noise ratio (PSNR), mean squared error (MSE), root mean squared error (RMSE), and structure similarity (SSIM). This evaluation exhibits that SANLM de-noise algorithm gives the finest presentation of PSNR and SSIM.

Middlebrooks et al. [15] analyzed motion artifacts and its noise removal with independent component analysis (ICA). The motivation of this analysis was to estimate the results of data de-noising with independent component analysis in MR images of the patients for preoperative analysis of glioma. FMRI results for preoperative analysis of images related with glioma patients resulted in decreased false positives and widening of true positives.

Haskell et al. [16] have introduced and validated retrospective motion adjustment algorithm for brain MRI images that includes a motion affected image which is reconstructed through multi-coil data. Motion artifact removing algorithm consists various steps like (a) both real component and imaginary component (two-channel complex MRI image) to be applied on motion detected CNN, (b) the uniformity data error reduction from the CNN image using complex voxel, and (c) reconstruction of the complete image data with position coordinates and multi-coil data model. The image creation integrates the motion restriction in a CNN-based model.

Chang et al. [17] designed a parallel computing-based algorithm for brain MR images using bilateral filtering method. This paper addresses an optimization technique to speed up calculation in either memory resources or thrust usages. A fully computerized and restriction-free revamp frame for the speedup bilateral filter related to ANNs is organized. At last, a vast variety of brain MR images are exploited to confirm the noise removal performance of proposed filtering method.

The strategies such as collection of imaging sequences, tracking of motion by sensors like cameras [18], implementation of various algorithms as well as reconstruction techniques [19], prospectively corrected acquisitions as fast pulse sequence designs, applying compressed sensing theory, parallel imaging, motion-resolved imaging and application of motion-robust acquisition schemes, respiratory belts, and MR navigators are mostly applied to correct and reduce the source of motion artifacts [20]. Among those strategies, reconstruction techniques using various algorithms are mostly preferred to reduce the motion artifact [21]. Today, various regularization constraints are combined with the many optimization-based algorithms, for example, sparse a priori in a certain transform domain and total variation (TV). The reconstruction techniques are the most efficient methods than the various interpolation methods; however, the images obtained from the reconstruction techniques have some deterioration at under sampling rates. In augmentation method, a nonlinear revamp algorithm should be stable with acquired data in k-space in opposition to the scattered in the modification domain [22, 23]. Determination of transform domain and appropriate regularization is difficult in the optimization-based methods which are also affected by the time-consuming iterations; therefore, those methods are non-suitable for clinical application. To eliminate the above-mentioned drawbacks, bilateral filtering techniques are well recognized in motion artifact removal of MRI. Optimal parameter tuning of bilateral filtering through bio-inspired optimization techniques delivers the finest way to enhance the convergence of motion correction methods and reduce the computation time [24, 25]. Table 1 summarizes the various algorithms/techniques experimented by the researchers for motion artifact removal. The performance is also enlisted with the help of various evaluation parameters.

3 Summary of Various Techniques Used for Motion Artifact Removal

See Table 1.

Table 1 Analysis on the techniques used for motion artifact removal

Ref.	Technique	PSNR	SSIM/MSSIM	RMSE	NRMSE	EME
[3]	LPA-ICI filter using generic algorithm	19.5934	0.44871	26.7221	–	–
[5]	Multimodal fusion—combination of stationary wavelet transform (SWT)—principal component analysis (PCA) and non-subsampled contourlet transform (NSCT) domains	39.5643	0.8867	–	–	24.7221
[9]	Higher-order singular value decomposition (HOSVD-R) recursive algorithm	30.53	0.887	–	–	–
[19]	DNN architecture for ROI MRI reconstruction (ROIRecNet)	32.17	0.901	–	–	–
[24]	DL algorithm for retrospective MRI motion correction	–	0.846	–	0.11	–
[13]	Conditional generative adversarial network	36.0	0.91	–	–	–
[14]	CNN-based reduction of motion artifacts (MARC)	37.2451	0.9014	3.0025	–	
[15]	Independent component analysis	–	–	–	–	–
[16]	CNN-network accelerated motion estimation and reduction			17.7		
[17]	Bilateral filter and back propagation network (BPN)	30.36—(for 9% noise) 31.46—(for 7% noise)		76.0538—(PCA) 35.8323—(T-test)		

PSNR Peak signal-to-noise ratio
SSIM Structure similarity
MSSIM Minimum structure similarity
RMSE Root mean squared error
NRMSE Normalized root mean squared error
EME Measure of enhancement

4 Generalized Filtering Techniques for Removal of Artifact

Many researchers focused their work on the removal of motion artifact with filtering techniques. Analysis on application of filtering methods on motion artifact affected MR images resulted in summarizing the following, which one has to consider during implementation.

1. **Variation filters**: This filter is used for reconstruction of MRI image from noisy MRI image by reducing the noise with the help of de-noising algorithm. It presents an upgraded non-local mean variation in MRI image 'u' by [26]

$$\min||Su||1 + \beta||u - Qu||_2^2.$$

Here Q indicates value of the non-central pixel of the chosen window, and S indicates set of perpendicular and parallel operators.

2. **Filters based on non-local means**: This filter regenerates same patches. The non-local mean filter preserves the MRI image edges, and it replaces noise pixels with weighted pixels of Gaussian neighbors and uses it in the fixing method [26]. The de-noise value of pixel is:

$$X[f](l) = \sum_{l \in O} T(z,l) f(l)$$

Here X is an evaluated value, f is the discrete noise image, $f = \{d(l)|z\epsilon O\}$, and T is weights $\left(\sum_j T(z,l) = 1\right)$ as:

$$T(z,l) = \frac{1}{z(i)} e \frac{-\left(\left|\left|d(N_i) - d\left(N_j\right)\right|\right|_2^2 \cdot \sigma\right)}{h^2}$$

where $\left(\left|\left|d(N_i) - d\left(N_j\right)\right|\right|_2^2 \cdot \sigma\right)$ is the value of Euclidean distance, σ is standard deviation, N is a fix neighborhood window, h is the stage of the filter, and $z(i)$ is constant.

3. **Wavelet filters**: Incorrect scaling and thresholding parameter chosen with wavelet filter may create artifact in image. The approximate coefficient needs to be considered; otherwise, it may result in information loss.
4. **Filters based on anisotropic diffusion**: This filter reduces image noise and preserves edges and controls the flatten process for little and high thresholds. The anisotropic diffusion [26] is given below:

$$f_t = \operatorname{div}(I \cdot p(||\nabla f|| \cdot \nabla f)), \quad I = 1 - \nabla G_\sigma * p(f)$$

Here, p: diffusion coefficient, I: spatially adaptive phrase, and G_σ: Gaussian kernel.

5. **Filters based on partial differential equation**: This filter is used to enhance and preserve the edge and remove the noise by reducing the Laplacian image's value. It also reduces the blocky artifacts and resolves the ramp edge artifacts [26].

$$f_t = -\left(\beta \frac{f_{xx}}{\left|\nabla^2 k\right|_\varepsilon}\right)_{xx} - \left(\beta \frac{f_{xy}}{\left|\nabla^2 k\right|_\varepsilon}\right)_{yx} - \left(\beta \frac{f_{yx}}{\left|\nabla^2 k\right|_\varepsilon}\right)_{xy} - \left(\beta \frac{f_{yy}}{\left|\nabla^2 k\right|_\varepsilon}\right)_{yy} - \lambda(k - f_0)$$

Here, k: input image and ε: small positive parameter $\left(\left|\nabla^2 k\right|_\varepsilon = \sqrt{|\left(\nabla^2 k\right)|^2 + \varepsilon}\right)$.

Diffusivity function $(\beta) = \frac{1}{\left(1+h|\nabla G * f_0|^2\right)}$ where h: threshold variable and G: Gaussian value.

6. **Morphological filters**: Morphological operations are nonlinear, and it does not consist of analytical values. The morphological operation consists of three steps: (i) finds the corrupted pixels in the chosen window; (ii) finds the corrupted pixels with the borders; and (iii) combines morphological parameter of dilation and erosion with a standardized factor of 3×3.
7. **Convolutional neural network-based filters**: It is a deep and feed forward artificial neural networks (FFANN) which uses a variation of multilayer perceptrons and preprocessing.

$$y_t = y_{t-1} - \left(\left(\sum_{i=1}^{Q_k} Ri * \phi_i^t (Ci * y_{t-1}) + \beta^t (y_{t-1} - 1)\right)\right).$$

Here, y: noise image, R_i: set of linear filters, Q_k: number of filters, β^t: force weight phase implementation, Ci: convolution kernel (rotates 180° in Ci), and ϕ_i^t: radial basis functions.
8. **Filters based on singularity function**: This method subdivides the source image into various sub-image and regenerate by two-dimensional singularity function analysis (DSFA).
9. **Curvelet filters**: Its concept is based on multiscale geometry (like direction, scale, and position). Its performance is studied with induced curvelet artifact production and smooth area.
10. **Contourlet filters**: In both directional and spatial resolutions, this filter gives sparse representation. It elevates computational complexity. It is used to acquire the number of scales and directions and also apply inverse contourlet transform to obtain de-noised image [26].
11. **Bilateral filter**: This filter is used for edge preservation, the corrupted pixels are restored with average of Gaussian neighboring pixels, and it also reduces blurring. The bilateral filter

$$f(i) = \frac{1}{(2R+1)^2} \sum_{j \in \{-RR\}^2} f(i-j)$$

Here, R: controls the measure of smoothing.

5 Experimentation with Cross-Guided Bilateral Filter (CGBF)

Cross-bilateral filter uses one image for finding the kernel and other to filter, and vice versa. In order to form the filter kernel and filter the other image, CBF considers both gray-level similarities and the geometric closeness of neighboring pixels in one image [27]. In guided bilateral filter [28], similar pixels of an image in each neighborhood are indicated by the guide image G. If similar pixels are present, then they are substituted. The link with robust estimation is lost because of this substitution. To re-establish this link, it is proposed to retain w_p and to introduce the guide image G into a third weight, w_g. Weight w_g like spatial weight is an indication of similarity and not a photometric weight. The guide image can be indeed seen as a label image that defines the arrangement of comparable pixels in each neighborhood. This is named as the guided bilateral filter. The hybrid of cross-bilateral filter (CBF) [27] and guided bilateral filter (GBF) [28] called cross-guided bilateral filter (CGBF) and is used to remove MRI motion artifact (Fig. 1).

1. **Study and analysis of motion artifact affected image**: The motion artifact is a blend of both blur and hazy images. With the assistance of radiologist, we can recognize the motion artifact affected MR images. These images need to be analyzed before processing in order to understand the artifacts present in image.

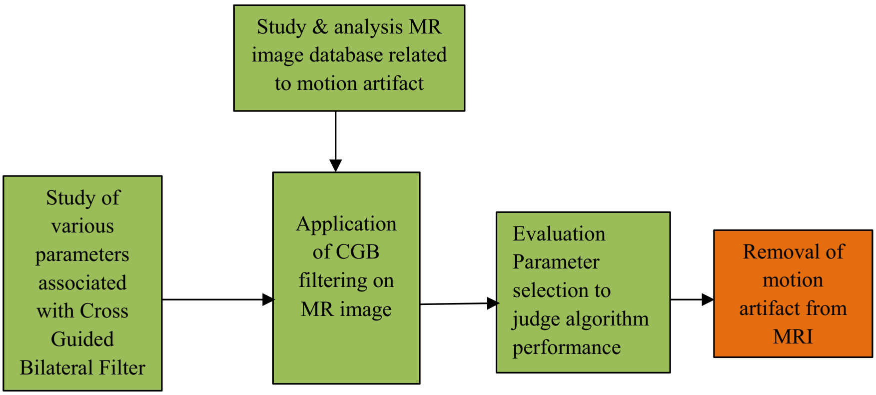


Fig. 1 Block schematic for the application of CGBF filter on motion artifact affected MRI

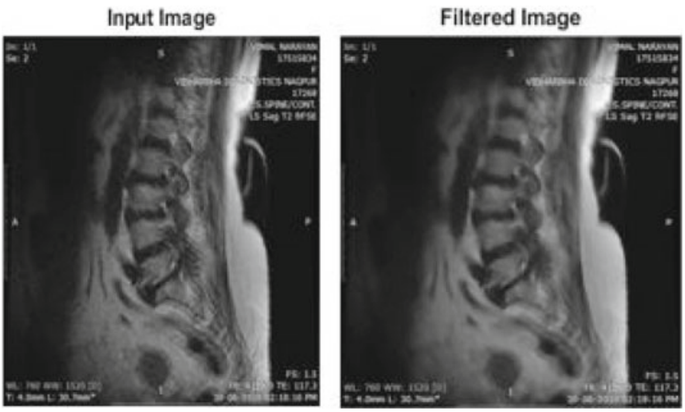
2. **Selection of various parameters in CGBF filter design**: Parameter selection associated with CGBF design is very important in order to process artifact affected MRI image.
3. **Performance evaluation of CGBF algorithm**: In order to evaluate the performance of algorithm, the parameters like MSE, PSNR, SSIM, etc., are required to judge the efficiency of the technique.

CGBF Algorithm:

Step: 1 Select and read the input image.
Seep: 2 Covert the input image into gray.
Step: 3 Resize the input image into 240 $\times$240.
Step: 4 Apply cross-guided bilateral filter by changing the kernel and sigma values.
Step: 5 CGBF performance is evaluated with PSNR, RMSE, SSIM, and VRMSE.

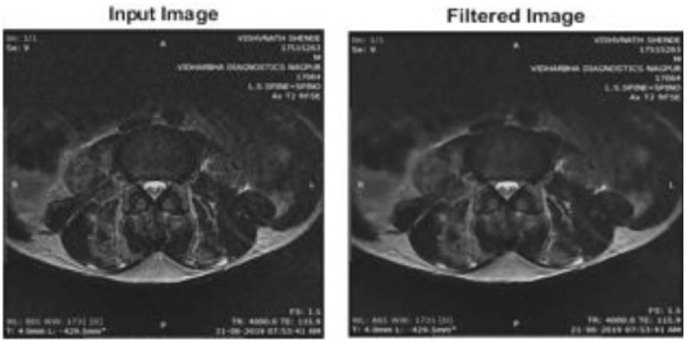
6 Result and Discussion

After application of CGBF filter on to these motion affected MR images of four patients, the results are obtained as shown in Figs. 2, 3, 4, 5. Figure 2a indicates sagittal view of the spinal cord MR image of patient 1, and Fig. 2b indicates its CGB filtered output. Figure 3a represents axial view of the spinal cord MR image of patient 2, and Fig. 3b indicates its CGB filtered output. Similarly, Figs. 4a and 5a indicate the axial view of motion affected MR images of patient 3 and 4, respectively. Figures 4b and 5b represent the CGB filtered output for those corresponding patient 3 and 4, respectively. After CGB filter application on to these motion artifact affected



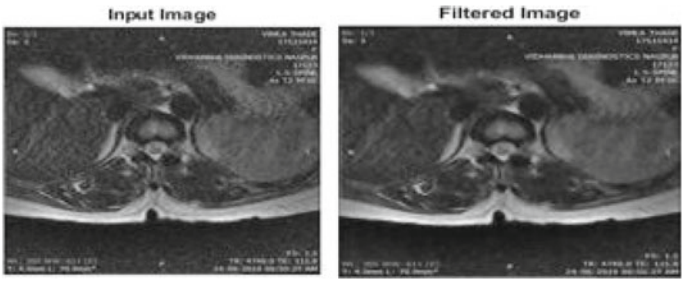
(a) (b)

Fig. 2 **a** Sagittal view of spinal cord image of patient 1, **b** CGB filtered image



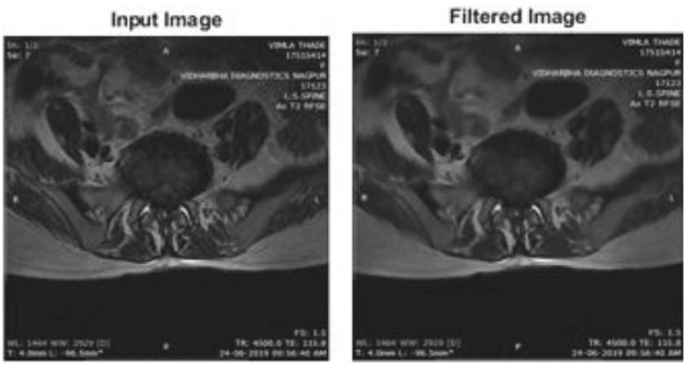
(a) (b)

Fig. 3 **a** Axial view of spinal cord image of patient 2, **b** CGB filtered image



(a) (b)

Fig. 4 **a** Axial view of spinal cord image of patient 3, **b** CGB filtered image



(a) (b)

Fig. 5 **a** Axial view of spinal cord image of patient 4, **b** CGB filtered image

Table 2 Various MRI spinal images showing their PSNR, RMSE, SSIM, and VRMSE values

Spinal cord MR image	PSNR	RMSE	SSIM	VRMSE
Sagittal view of patient 1	49.336	0.0010	0.989	0.7532
Axial view of patient 2	48.170	0.0012	0.970	0.9815
Axial view of patient 3	45.838	0.0016	0.9130	1.6739
Axial view of patient 4	48.824	0.0011	0.981	0.839

Table 3 Performance parameters of existing methods and proposed method

METHODS	PSNR	RMSE	SSIM	VRMSE
Proposed CGBF	48.042	0.00122	0.96325	1.0619
BF-GA	30.0996	0.0020	0.8827	3.45
DCT-ACO	34.63	0.0017	0.997	3.1
DF-MFF	46.22	0.0015	0.9989	1.18

images, the CGB filtered images are analyzed visually by various human observers, and the artifact is seen to be removed. The algorithm performance is evaluated with the help of evaluation parameters like MSE, PSNR, SSIM, and VRMSE, and their corresponding results are presented in Table 2.

Table 3 gives the performance parameter values in terms of PSNR, RMSE, VRMSE, and SSIM of various existing methods and proposed method. Figure 6 depicts the evaluation of our proposed method with the existing method for motion artifacts reduction, in terms of PSNR, SSIM, RMSE and VRMSE. Here, the existing methods considered are bilateral filter-genetic algorithm (BF-GA) [29], discrete cosine transform-ant colony optimization (DCT-ACO) [30], and dragonfly-modified firefly (DF-MFF) [31]. Figure 6a indicates the PSNR estimation of various methods. PSNR value indicates the quality of an image, a higher PSNR indicates a good quality images, and our proposed system shows higher PSNR compared with the existing methods, which means our method removes all the motion artifacts. Figure 6b and c demonstrates the RMSE and VRMSE values, and an image with low value of RMSE and VRMSE gives low-error rate. Figure 6d explains the SSIM comparison of proposed method with the existing methods. Our proposed method achieved high SSIM, which means our filtered output is a good quality image.

7 Conclusion and Future Direction

In this paper, we proposed cross-guided bilateral filter (CGBF) for MRI motion artifact affected image. Here, we applied CGBF and compared with other existing method like BF-GA, DCT-ACO, and DF-MFF. The proposed system achieved better result in terms of VRMSE, PSNR, RMSE, and SSIM. The proposed method increases

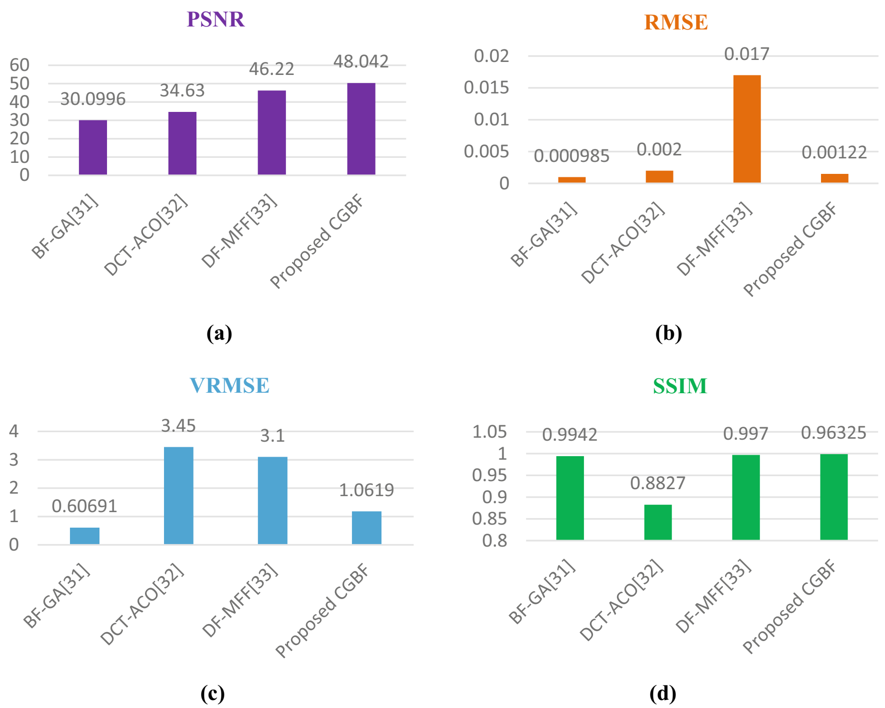


Fig. 6 **a** PSNR comparison of various methods, **b** RMSE comparison of various methods, **c** VRMSE comparison of various methods, **d** SSIM comparison of various methods

the quality of the image. In artifact removal process, proper method selection is very important in order to obtain the expected results. Experimentation with MR images of four patients yields average PSNR value of approximately 48 and average SSIM value of 0.96 indicating the better performance of CGBF technique. In the future work, using bald eagle search method and salp swarm optimization method, the parameters of CGBF can be optimized. Thus, the proposed work can be extended for the better diagnosis of motion artifact affected images.

References

1. Tempany C et al (2014) Multimodal imaging for improved diagnosis and treatment of cancers. Cancer 121(6):817–827
2. Tchagna Kouanou A, Tchiotsop D, Kengne R, Zephirin D, Adele Armele N, Tchinda R (2018) An optimal big data workflow for biomedical image analysis. Inform Med Unlocked 11:68–74
3. Dey N et al (2015) Parameter optimization for local polynomial approximation based intersection confidence interval filter using genetic algorithm: an application for brain MRI image de-noising. J Imag 1(1):60–84

4. Wang Y-X (2015) Current status of super paramagnetic iron oxide contrast agents for liver magnetic resonance imaging. World J Gastroenterol 21(47):13400
5. Bhateja V, Patel H, Krishn A, Sahu A, Lay-Ekuakille A (2015) Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains. IEEE Sens J 15(12):6783–6790
6. Telischak N, Detre J, Zaharchuk G (2014) Arterial spin labeling MRI: clinical applications in the brain. J Magn Reson Imaging 41(5):1165–1180
7. Zhuang X, Shen J (2016) Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI. Med Image Anal 31:77–87
8. Zaitsev M, Maclaren J, Herbst M (2015) Motion artifacts in MRI: a complex problem with many partial solutions. J Magn Reson Imaging 42(4):887–901
9. Zhang X et al (2015) Denoising of 3D magnetic resonance images by using higher-order singular value decomposition. Med Image Anal 19(1):75–86
10. Culvenor A et al (2015) Early knee osteoarthritis is evident one year following anterior cruciate ligament reconstruction: a magnetic resonance imaging evaluation. Arthritis Rheumatol 67(4):946–955
11. Chen J, Glover G (2015) Functional magnetic resonance imaging methods. Neuropsychol Rev 25(3):289–313
12. Rajeswaran N, Gokilavani C (2016) Reduction of FBM noise in brain MRI images using wavelet thresholding techniques. Asian J Inf Technol 15(5):855–861
13. Johnson PM, Drangova M (2019) Conditional generative adversarial network for 3D rigid-body motion correction in MRI. Magn Reson Med 82(3):901–910
14. Saladi S, Amutha Prabha N (2017) Analysis of denoising filters on MRI brain images. Int J Imaging Syst Technol 27(3):201–208
15. Middlebrooks E, Frost C, Tuna I, Schmalfuss I, Rahman M, Old Crow A (2016) Reduction of motion artifacts and noise using independent component analysis in task-based functional MRI for preoperative planning in patients with brain tumor. Am J Neuroradiol 38(2):336–342
16. Haskell M et al (2019) Network accelerated motion estimation and reduction (NAMER): convolutional neural network guided retrospective motion correction using a separable motion model. Magn Reson Med 82(4):1452–1461
17. Chang H, Lin Y, Zhuang A (2018) An automatic parameter decision system of bilateral filtering with GPU-based acceleration for brain MR images. J Digit Imaging 32(1):148–161
18. Kustner T et al (2017) MR-based respiratory and cardiac motion correction for PET imaging. Med Image Anal 42:129–144
19. Sun L, Fan Z, Ding X, Huang Y, Paisley J (2019) Region-of-interest under-sampled MRI reconstruction: a deep convolutional neural network approach. Magn Reson Imaging 63:185–192
20. Sun C, Tang C, Zhu X, Ren H (2015) Exponential total variation model for noise removal, its numerical algorithms and applications. AEU Int J Electron C 69(3):644–654
21. Devi R, Anandhamala G (2018) Recent trends in medical imaging modalities and challenges for diagnosing breast cancer. Biomed Pharmacol J 11(3):1649–1658
22. Chu M, Chang H, Chung H, Truong T, Bashir M, Chen N (2014) POCS-based reconstruction of multiplexed sensitivity encoded MRI (POCSMUSE): a general algorithm for reducing motion-related artifacts. Magn Reson Med 74(5):1336–1348
23. De Martino F, Moerel M, Ugurbil K, Formisano E, Yacoub E (2014) Less noise, more activation: multiband acquisition schemes for auditory functional MRI. Magn Reson Med 74(2):462–467
24. Küstner T, Armanious K, Yang J, Yang B, Schick F, Gatidis S (2019) Retrospective correction of motion-affected MR images using deep learning frameworks. Magn Reson Med 82(4):1527–1540
25. Papież B, Heinrich M, Fehrenbach J, Risser L, Schnabel J (2014) An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. Med Image Anal 18(8):1299–1311
26. Mafi M, Martin H, Cabrerizo M, Andrian J, Barreto A, Adjouadi M (2019) A comprehensive survey on impulse and Gaussian denoising filters for digital images. Signal Process 236–260, 0165-1684

27. Kumar BS (2015) Image fusion based on pixel significance using cross bilateral filter. SIViP 9(5):1193–1204
28. Caraffa L, Tarel JP, Charbonnier P (2015) The guided bilateral filter: when the joint/cross bilateral filter becomes robust. IEEE Trans Image Process Inst Electr Electron Eng 24(4):1199–1208. https://doi.org/10.1109/TIP.2015.2389617
29. Akar SA (2016) Determination of optimal parameters for bilateral filter in brain MR image denoising. Appl Soft Comput 43:87–96
30. Miri A, Sharifian S, Rashidi S, Ghods M (2018) Medical image denoising based on 2D discrete cosine transform via ant colony optimization. Optik 156:938–948
31. Elhoseny M, Shankar K (2019) Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements 143:125–135

Assessment of Transformer Fault Diagnosis and Condition Monitoring Methods

Nasir ul Islam Wani, Anupama Prakash, and Pallavi Choudekar

Abstract Power transformer provides an essential element for transmission and distribution of electrical energy. Such static devices are prone to many harsh conditions during their operation and hence has a maximum chance of deterioration. Since transformers are one of the costly equipment's of a transmission system, their failure can lead to huge losses. If transformers are regularly monitored their failure can be prevented. This results in efficient utilizations of resources. This paper incorporates various monitoring techniques which can be used for fault detection of transformers.

Keywords Condition monitoring · Power transformer · Faults

1 Introduction

Power transformers provide an essential element for the generation and distribution of electrical energy [1]. They form an important component of power system. Such static devices are subjected to harsh operating conditions and hence has a maximum chance of decline in their performance. Thus, their better performance implies high power system efficiency.

Transformer faults are classified as Internal and External faults [2]. Internal faults can be active or incipient in nature. Active faults are solid faults that appear on wending of transformer. It may be phase to phase or phase to ground. These faults lead to insulation breakdown. Incipient faults are thermal or electrical. The paper discusses various types of monitoring techniques for fault detection.

N. I. Wani (✉) · A. Prakash · P. Choudekar
EEE Department, Amity University Uttar Pradesh, Noida, India
e-mail: nasir12wani@gmail.com

A. Prakash
e-mail: aprakash1@amity.edu

P. Choudekar
e-mail: pachoudekar@amity.edu

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_35

The condition monitoring technique of power transformer has been translated into practice in the space of few years. It caters largely for diagnosis of internal faults in transformers caused due to aging and degradation [1]. The condition monitoring allows early identification of developing faults before any failure. Condition monitoring is aimed for not only the detection of fault at early state but also reduces the maintenance cost for transformer.

2 Condition Monitoring Techniques

2.1 Chemical Techniques

It involves Dissolve Gas Analysis and Partial Discharge Analysis.

2.1.1 Dissolve Gas Analysis

It is used to know the condition of an oil immersed transformer [2]. The concentration and production rate of the generated gases is used for various faults detection.

- Arcing fault can be detected by the analysis of H_2, C_2H_2 gases in the oil.
- Corona effect can be analyzed by monitoring the production of carbon monoxide in the transformer oil.
- Sparking can be detected by monitoring the level of methane in the transformer oil.

The Dissolve Gas Analysis is performed by carrying out the gas chromatography in the oil samples taken from the main oil stream to measure the concentration and production of various gases. The process of Dissolve Gas Analysis is shown in Fig. 1.

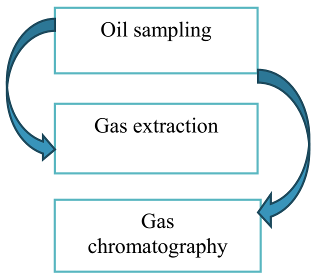


Fig. 1 Process of dissolve gas analysis

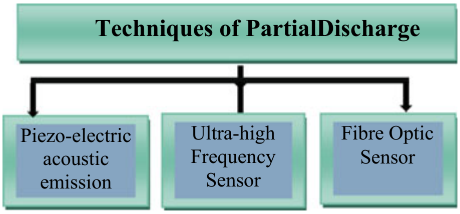


Fig. 2 Different techniques of partial discharge

2.1.2 Partial Discharge Analysis

Partial discharge analysis has evolved as a crucial, unobjectionable, quick and robust diagnostic utility and contains a substantial amount of insulation state statistics of power transformer [3]. All the Ultra-high frequency Partial discharge signals of different frequency band components are recorded and used for making the frequency matrix and then feasible signal region is developed using multidimensional energy parameter and multi dimension sample entropy parameters [4].

The techniques involved in partial discharge are (Fig. 2).

- Piezoelectric Acoustic Emission: In this, the acoustic sensors which converts discharge acoustic emission signal to electrical signal are placed on the experimental transformer tank. These sensors help in identifying source of Partial Discharge within the transformer.
- Ultra-high Frequency Sensor: Detection of any partial discharge activity is an indication of insulation defect. Once detected analysis of data is carried out to identify the type and location of defect.
- Fiber optic sensor: It can detect acoustic pressure and detect partial discharge in small voids and then transmit the information from high voltage equipment to monitoring equipment for analysis.

2.2 Electrical Techniques

It involves following methods.

2.2.1 Vibration Analysis

It is an effective method for detecting internal fault in transformers. There are different types of vibrations associated with transformers.

- Winding vibration: The electric forces that are caused due to interaction between circulating currents in winding and leakage flux associated to the winding. Such forces are generated both in axial and radial direction.
- Core vibration: These vibrations are caused by magneto restriction and magnetic forces. The magnetic force changes the dimension of objects.

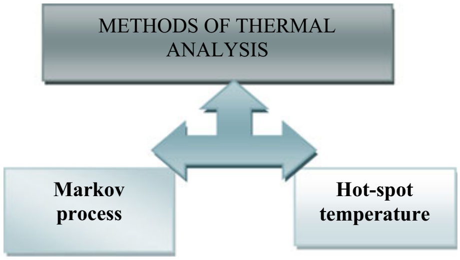


Fig. 3 Different methods of thermal analysis

The accelerometer is placed in the vicinity of transformer to capture these vibrations. The vibrations are converted into electrical signals [3]. The frequency spectrum is plotted and is compared with the reference spectrum for detection of any unusual vibration.

2.2.2 Thermal Analysis

Thermal analysis allows study of heat distribution in the transformer at different operating conditions [5]. It is seen that 80% of failures occur due to coil insulation break down which is as an outcome of hotspots in the winding. It has been established that an increase in the working temperature range of the transformer by 10 K decreases it average lifetime by 2.5% [6]. The thermal analysis is carried by following methods as in Fig. 3 are:

- Markov Process: This process forms a probabilistic approach where the transit plausibility can be determined from existing data base [7]. This method is thus used to optimize periodic inspection so that the cost is minimized, and availability of transformer is maximized.
- Hot-spot Temperature: Hot-spot temperature is one of the most important parameters when defining thermal condition and overloading capacity. The abnormal condition is detected by analyzing Hot-spot temperature [8]. In order to measure the hot-spot temperature, the thermal sensors is placed between the winding and is attached to the optical fiber. The signals are transmitted via optical fiber to the control center.

2.2.3 Winding Movement and Deformation

The transformer winding and deformation are mainly caused due to short circuit event caused in transformers. The high current interacts with leakage flux density and result in extreme force to act on winding ($F = L*I*B$) [9]. These forces have radial and axial components thus leads to both radial and axial deformation.

$$\text{Axial deformation } (F_y) = [50.8 * S]/[Z * H * f]$$

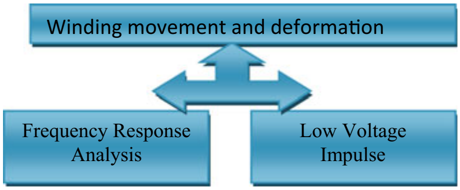


Fig. 4 Methods for detecting winding movements

$$\text{Radial deformation } (F_x) = [\mu \mathrm{NI}^2 / H][*\pi * D]$$

where

F	Force on winding.
L	length of the conductor.
B	Flux Density.
S	Rated power per limb in kVA.
Z	Impedance in per unit.
H	Height of winding in meters.
f	Frequency (Hz).
μ	Absolute permeability.
NI	Ampere turns.
D	Mean diameter.

These deformations not only reduce the short circuit withstand capacity of the transformer but also ruptures the insulation.

Winding movement [10] and deformation analysis is done by different methods as mentioned in Fig. 4:

- Frequency Response Analysis (FRA): This is an effective characteristic technique extensively used for recognition of lateral faults [9]. Frequency Response Analysis is an offline approach where an AC signal with low voltage is introduced to the terminal of winding and reaction is recorded at the other terminal [11]. The transfer function, impedance or admittance of winding is measured over a frequency range and these are used for fault diagnosis. The faults diagnosed by FRA are:

  Mechanical faults.

(a) Winding deformation
(b) Core displacement

  Electrical Faults.

(a) Short circuit or open circuit
(b) Bad ground connection

  - Low Voltage Impulse**:** Low voltage impulse is sensitive technique for detecting winding movement and deformation in transformer caused due to short circuit or short low voltage impulse (300 V) [2]. The impulse is applied, and the coupled currents are recorded. The FFT is carried out on the recorded waveforms and compared with results of the transformer of same type or previously recorded data on same transformer to identify the faults.

2.2.4 On Load Tap Changing(OLTC)

On load tap changers are responsible for voltage regulation in transmission system [12]. They are mainly connected to transformer for maintaining voltage levels under variable loads. The OLTC leads to change in turns ratio and thus level of output voltage. There are primarily two OLTC designs:

- Diverter Design: This is mainly employed for value of high voltage and power. This type of Diverter consists of a separate tap selector to select taps in transformer tank and a separate diverter switch. To switch the load current.
- Non-Diverter Design: This is mainly employed for low values of voltage and combines function of diverter switch and tap selector.

Most of the transformer's failures are associated to tap changing. Thus, we need to carefully review and watch their operating conditions. The methods used for on load tap changing are [2]:

- Winding Resistance: The resistance at various connections are measured and this is compared with the values of factory and any variation from these values are recorded.
- Torque Measurement: This method uses motor parameter for finding any problem due to any mechanical defect due to aging.
- Gas Analysis: In this method the concentration of gas is determined in tap changing compartment. The variation in gas level indicates any abnormal behavior.

2.3 Soft Computing Methods

This methods incorporates approximations to any problem followed by learning and optimization [13] two processes:

2.3.1 Artificial Neural Network

Artificial neural network is an information processing paradigm that is inspired by biological nervous system in human brains. Neural network provides an advantage in following ways:

(a) Adaptive learning: The capability of performing the given piece of work based on information it receives.
(b) Self-organization: They capability to arrange the information it receives during learning time.

(c) Real time operation: Artificial neural network computation are done in a coordinated and special hardware device.

Neural network base algorithm for fault detection in transformers**:** The artificial neural network consists of three operating layers: the input layer, the output layer and the hidden layer [14]. The neural network is used in following procedure:

(a) The neural network is provided with data to train it.
(b) The output is checked and compared with the actual output.
(c) Then the neural network is trained, and the weights and bias are changed.

Some of the neural networks used for fault detection are [15]:

- Multilayer perceptron: In a multilayer perceptron there can be more than one layers. It is a feed forward network consisting of several fully connected layers where the parameters of each unit are independent of rest of units in the layer.
- Back propagation: The back propagation algorithm is the basic method for weight updating. In this gradient-descent method is used to give the error back to the hidden layer.

The different faults techniques in which artificial neural network can be used are:

(a) Dissolved gas analysis
(b) Partial discharge analysis.

2.3.2 Fuzzy logic

The fuzzy logic is a contemporary technique of determining imprecision and unpredictability in information. In fuzzy logic a declaration has certain level of association varying from totally legitimate through partial legitimacy to totally inaccurate which differs from the traditional method [16].

The fuzzy logic analysis consists of three parts:

(a) Fuzzification: It is a method of converting a systematic data set into classes of fuzzy set. This process converts numerical values into linguistic sets and then assigns a membership function value. The Fuzzifier is used to covert systematic data in fuzzy variables.
(b) Fuzzy inference system: The fuzzy inference system is responsible for drawing conclusion from the knowledge-based fuzzy rule set of if–then linguistic.
(c) Defuzzification: It is a process of converting the fuzzy output values back to crisp values. Defuzzification is a process of converting output fuzzy variable into a unique number. Defuzzification methods include

    - Max-membership principle
    - Centroid method
    - Weighted average method
    - Mean-max membership.

The different fault analysis techniques developed on fuzzy logic are [13]:

- Fuzzy Roger's ratio: This includes calculation of different gas ratio codes which are given as input and classified as low, high, medium, very high. The membership boundaries are introduced for gas ratio and if–then rule system is applied.
- Fuzzy IEC ratio: The gas ratios are classified as low, medium and high. The membership function is introduced, and fuzzy rules are defined on type of fault.
- Fuzzy dual triangle: The three gases are converted into fuzzy logic controller with three input and each output with some membership function.
- The fuzzy based monitoring of transformer [17] is a developing field and many works are being carried out to find the techniques controlled by fuzzy systems for transformer condition supervision.

3 Results

The overall transformer management consists of the techniques of condition monitoring, maintenance plans and aging. The overall condition monitoring technique as shown in Fig. 5 are used detection of faults at early stages and hence reduce the overall maintenance cost of the transformers.

For monitoring the transformer the position of PD wrt to phase angle is important. It is noticed that PD pulse has 90° phase angle in positive half cycle and nearly 270° phase angle in negative half cycle of the 5 kV applied voltage which is shown in Fig. 6. The nether voltage rate between the test object is not enough to cause field intensity within the void in excess of PD inception strength. Therefore, PDs are mostly appearing at 90° phase angle and 270° phase angle of the applied voltage where the maximum amplitude of the applied voltage is reached.

Table 1 depicts the level p of the different gases. Level 1 corresponds to the range of normal operation of the transformers and Level 2 corresponds to the abnormal conditions where the excessive decomposition of insulator and oil has occurred in gas analysis. The Level 2 depicts that if operation of transformer will be continued their may be the failure of transformer.

Table 2 shows that any variation of frequency exceeding in range of +3 to −3db the fault has occurred and the range of frequency also shows the type of fault. The traces are to compare with the baseline for diagnosis and if the deviations occur it shows the fault has occurred and transformer requires attention.

4 Conclusion

The survey provides an insight to various fault detection and monitoring techniques for the transformer faults. The paper discusses the major techniques such as Electrical techniques, Chemical techniques, Soft computing techniques for monitoring of the transformer faults. These techniques help in life estimation, reducing maintenance

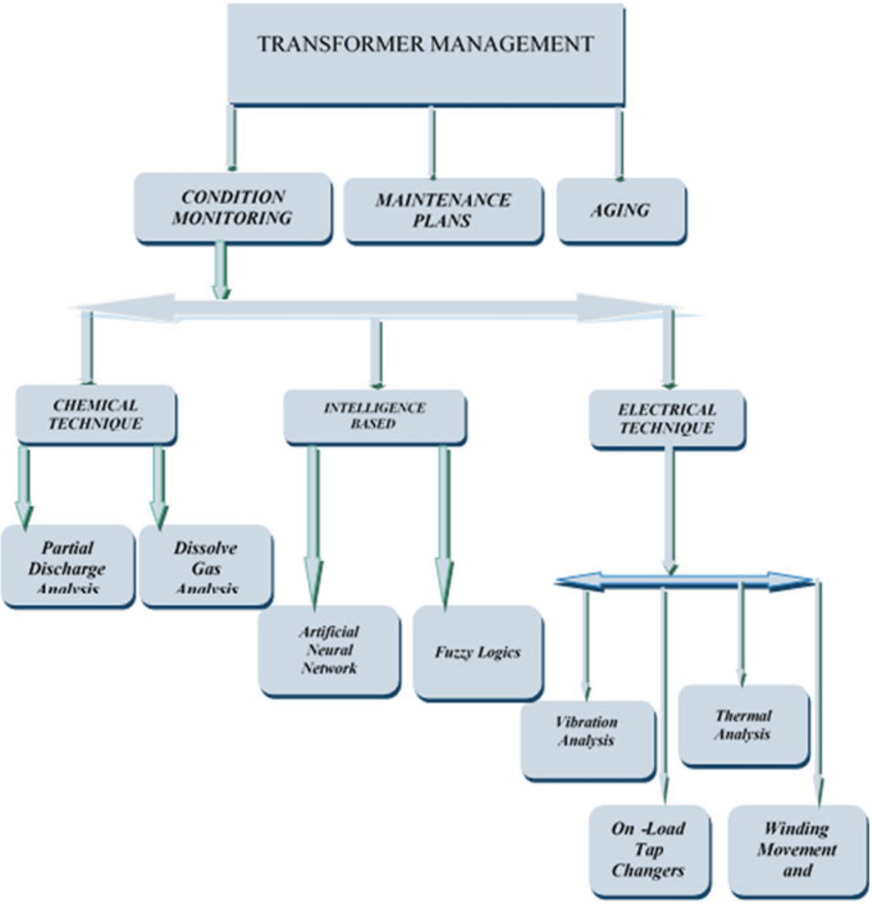


Fig. 5 Overall condition monitoring techniques

cost, enhancing life span and the services of device. Table 3 shows the various methods and the faults diagnosed by them.

There are many researches undergoing on these condition monitoring techniques. Major focus in on development of techniques which can predict any abnormality in transformer before it occurs so that preventive measures can be taken.

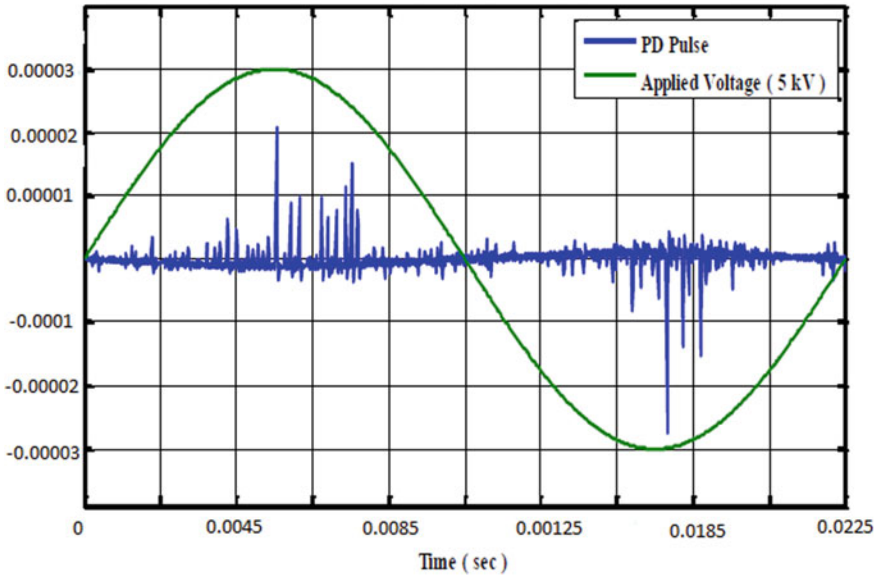


Fig. 6 PD pulse observed at 5 kV voltage

Table 1 Concentration of gases normal and fault operation

Gas	Level 1(ppm)	Level 2(ppm)
Hydrogen	100	> 1700
Ethylene	45	> 195
Carbon monoxide	320	> 1375
Methane	110	> 985

Table 2 Detection of fault in terms of their frequency variations in FRA

Frequency range	Change in frequency (db)	Fault observed
5 Hz–2 kHz	> 6	Open circuit
50 Hz–20 kHz	> 6	Bulk winding movement
500 Hz–2 MHz	> 6	Deformation in winding
25 Hz–10 MHz	> 6	Winding leads problem

Table 3 Various methods and fault diagnosed by the method

S. No	Method	Diagnosis
1	Dissolve gas analysis	Arcing faults, sparking, corona effect
2	Partial discharge gas analysis	Short circuit, insulation failure due to overvoltage
3	Vibration analysis	Partial looseness of windings, Winding deformation
4	Thermal analysis	Insulation breakdown due to excess heating
5	OLTC	Tap changing fault troubleshooting

References

1. Aslam M, Arbab MN, Basit A (2019) A review on fault detection and condition monitoring of power transformer. Int J Adv Appl Sci 6(8):100–110
2. Moravej Z, Bagheri S (2015) Condition monitoring techniques of power transformers: a review. J Oper Autom Power Eng 3(1):71–82
3. Deolanakar VS, Gandhar WZ (2017) Partial discharge analysis in high voltage current transformer. Int Refered J Eng Sci 6:24–30
4. Jia R, Xie Y (2016) Transformer partial discharge fault diagnosis based on multidimensional feature region. Math Prob Eng 2016:11. Article ID 4835694
5. Kravchenko EV, Ivleva D (2015) Thermal analysis of power transformer. MATEC Web Conf 23:01021. https://doi.org/10.1051/matecconf/20152301021
6. Pierce LW (1992) An investigation of the thermal performance of an oil filled transformer winding. IEEE Trans Power Deliv 7(3):1347–1358
7. Yahaya MS, Mohd Selva AN, Ab Kadir A, Jasni MZA, Kadim J (2006) A maintenance cost study of transformers based on markov model utilizing frequency of transition approach. Energies 11:2006
8. Isha MT, Wang Z (2008) Transformer hotspot temperature calculation using IEEE loading guide 2008. In: International conference on condition monitoring and diagnosis, pp 1017–1020. https://doi.org/10.1109/CMD.2008.4580455
9. Alsuhaibani S, Khan Y, Beroual A, Malik NH (2016) Review a review of frequency response analysis methods for power transformer diagnostics. Energies 9:879
10. Mcdowell G, Lockwood L (1994) Real time monitoring of movement of transformer winding. In: IEE colloquium on condition monitoring and remnant life assessment in power transformers, pp 1–14
11. Dick E.P., Erven C.C: 'Transformer diagnostic testing by frequency response analysis', IEEE Trans. Power Appl. Syst., 1978, PAS-97,(6),pp. (2144–2153)
12. Kandagal SS, Santhoshkumar GM, Shivanagutti JG, On-load tap changer fault diagnosis and maintenance of 100 MVA power transformer. J Inf Knowl Res Electr Eng. (ISSN: 0975–6736, GIF: 01.1419, SIF: 02.798)
13. Tang W, Wu QH (2011) Condition monitoring and assessment of power transformers using computational intelligence. Int J Electr Power Energy Syst 33:1784–1785
14. Siva DVSS, Kalyani GNS (2004) ANN approach for condition monitoring of power transformers using DGA. In: TENCON, IEEE Region 10 conference, pp 444–447
15. Kaur A, Brar YS, Leena G (2019) Fault detection in power transformers using random neural networks. Int J Electr Comput Eng (IJECE) 9(1):78–84
16. Hossam-Eldin AA, Refaey M, Ramadan H (2017) New approach to power transformer asset management and life assessment using fuzzy logic techniques. In: 2017 19th international middle east power systems, conference (MEPCON) Cairo, vol 10, 2017, pp 901–908
17. Németh B, Laboncz S, Kiss I (2009) Condition monitoring of power transformers using DGA and fuzzy logic. In: IEEE electrical insulation conference, Montreal, QC, Canada, 31 May–3 June 2009, pp 373–376

Use of Deep Neural Networks in Detecting Breast Cancer Lesion

S. Muthumanickam, R. Saranya, N. M. Perezhil, K. Sangeetha, and S. Vinisha

Abstract This paper presents a deep neural network module for breast cancer lesion detection which is trained and tested over 569 datasets. The model produces an accuracy of 0.95 in predicting the benign and malignant lesions in the breast. This model gives high accuracy to few advancement in technology with multiple input views and optimised amongst many choices. A thorough analysis of the model's performance on various populations is conducted. The python code used in the model uses tensor flow library and a seaborn visualisation library. The data sets used in the code are segregated as training and testing datasets. Training sets contain 80% of data and testing sets contain 20% of data. Histogram images are created to view all the parameters separately for training and testing. Later the accuracy and cost are found so that it can be implemented in real-world applications without any wrong predictions.

Keywords Breast cancer · Deep learning · Breast screening · Mammography · Neural networks

1 Introduction

Cancer is an abnormal cell growth with the potential to attack or spread to other parts of the body. Breast tissues contain fat and connective tissues, lymph's nodes and blood vessels. Breast cancer is the 2nd highest cause of death amongst women. The various diagnostic methods for breast cancer are Mammogram, Digital Mammography, Computer-Aided Detection (CAD), Biopsy, Microwave Imaging and Breast Magnetic Resonance Imaging. Mammography is an X-ray of the breast which aids in the early detection and diagnosis of breast carcinoma in women. This is the most commonly used diagnostic method at present. There are seven stages of

S. Muthumanickam · R. Saranya (✉) · N. M. Perezhil · K. Sangeetha · S. Vinisha
Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Thiruvallur, India
e-mail: saya304@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_36

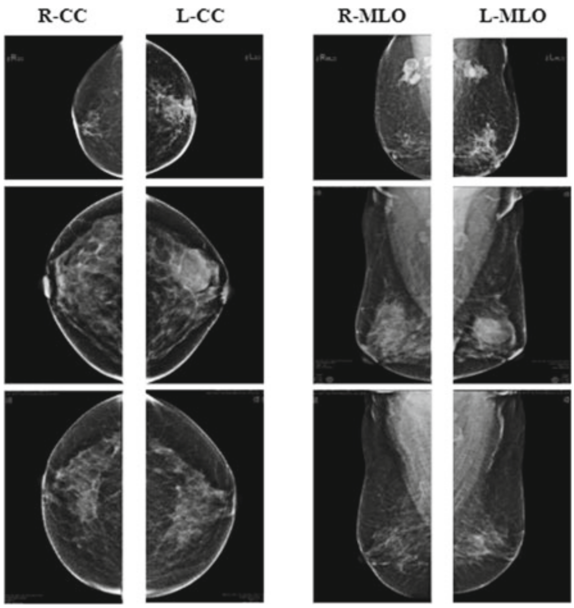


Fig. 1 Craniocaudal and mediolateral mammographic images of malignant, benign and normal lesions in order

Breast Imaging—Reporting and Data System (Bi-Rads). Stage 0 indicates incomplete mammogram, Stage 1 indicates Negative—No cancer cells detected, Stage 2 indicates the presence of benign cyst (Fibroademonas), Stage 3 indicates represents grey zone which says 2% chance of cancer, Stage 4 indicates suspicious case—needs a biopsy, Stage 5 indicates highly suspicious malignant case and Stage 6 indicates a fatal case of cancer. The report is examined manually with the help of CAD which differentiates and indicates the lesions from the normal cells. The disadvantage of the above approach is that it is time-consuming and sometimes detected incorrectly.

The proposed system uses a deep neural network to detect the type of cancer. To make the detection automated, python code is written using deep learning technique. The datasets are fed into the code where 80% of them are trained and the remaining 20% is used for testing the trained module. The accuracy of the model is found. The goal of this paper is to develop neural networks which help radiologists interpret breast cancer screening exams in a better and efficient way (Fig. 1).

2 Material and Methodology

The dataset used in this model is the *Wisconsin datasets*. The mammographic images are analysed and converted to.csv files that consist of numerical parameters taken from the image. A total of 569 datasets out of which 212 cases are malignant and 357 datasets are benign. The model trains 80% of the total datasets and uses the

Table 1 Number of datasets with malignant and benign findings extracted from the mammographic image report and the number of trained datasets and tested datasets

	Malignant	Benign
Training	169	285
Testing	43	72
Total	212	357

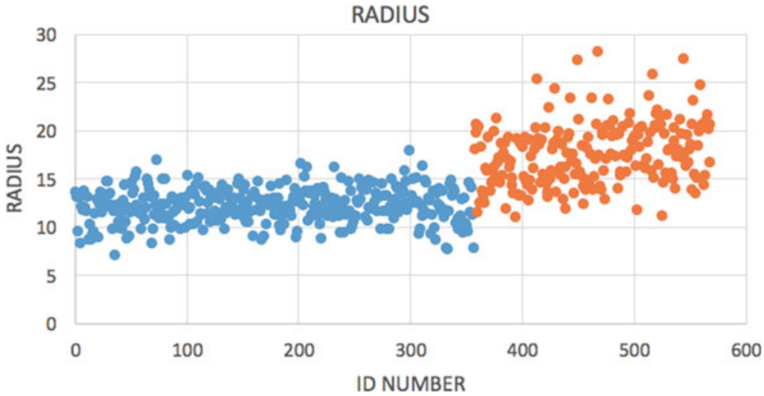


Fig. 2 Scatter plot of the data set considering the parameter, radius

remaining 20% for testing purposes, i.e. 456 of the total 569 datasets are trained and 113 are tested (Table 1; Fig. 2).

Here the radius of the benign cases is less when compared to malignant cases. The benign cases are noted in blue and the malignant cases are noted in orange colour. There is a notable difference in values for malignant and benign cases. The malignant case starts from 358 and so there is a difference in radius for benign and malignant cases (Table 2; Fig. 3).

Table 2 Parameters used

Radius	Sum of the distance between midpoint to point on the boundary
Texture	The standard deviation of grey-scale values
Perimeter	The length measured between the snake points
Area	Total pixel number on the interior of the snake and add half of the pixel in the perimeter
Smoothness	Local variation in radius length
Compactness	Perimeter^2/ area—1.0
Concavity	Acerbity of the concave portions of the Contour
Symmetry	The difference in length between lines perpendicular to the major axis to the cell boundary in both directions
Fractal Dimensions	Coastline approximation-1

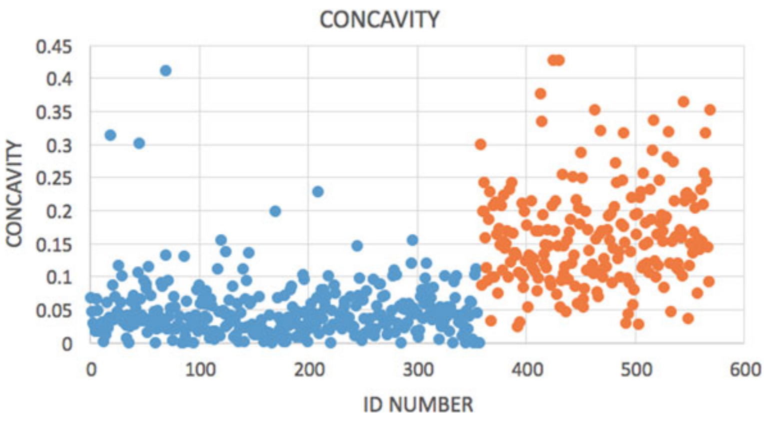


Fig. 3 Scatter plot of the datasets considering the parameter, concavity

The range of values is greatly varied for malignant and benign cases for the concavity in the image. The malignant cases will have the greater concavity since the damage is more in malignant cases. The benign cases are noted in blue and the malignant cases are noted in orange colour. There is a notable difference in values for malignant and benign cases. The lesion will be more deeper in case of malignant whereas it will be less deeper in case of benign cases (Figs. 4 and 5).

For malignant cases, the compactness will be higher in some cases since the perimeter of the image taken will be greater. But in benign cases the compactness is not that high when compared to malignant cases. The benign cases are noted in blue and the malignant cases are noted in orange colour.

There is a notable difference in values for malignant and benign cases. Compactness is basically defined as (Fig. 6)

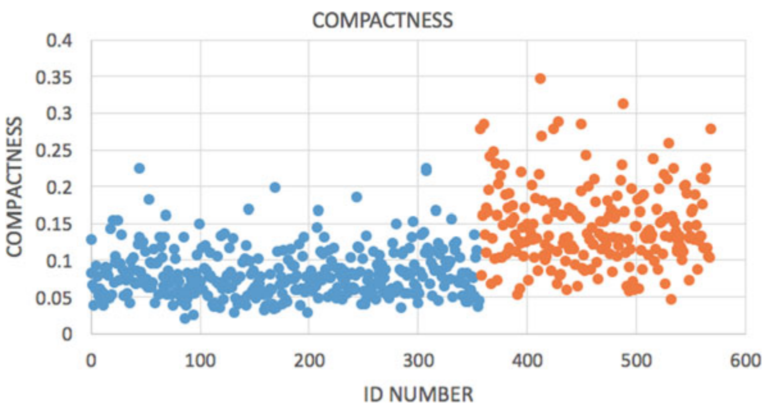


Fig. 4 Scatter plot of the datasets considering the parameter, compactness

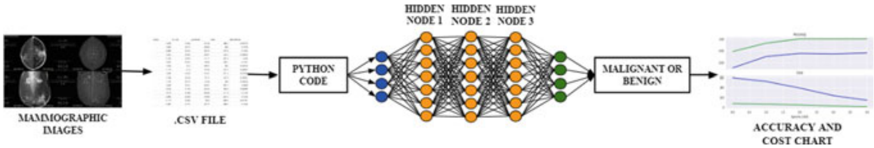
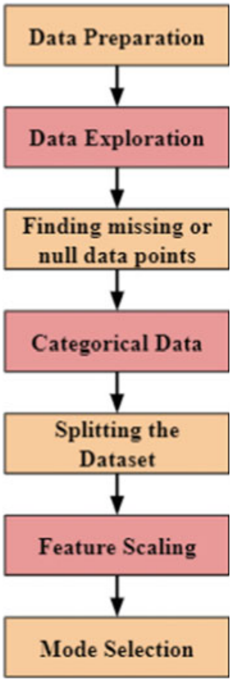


Fig. 5 Architecture of the model

Fig. 6 Flow diagram on processing of dataset



$$\text{Compactness} = (\text{Perimeter})^2/\text{Area} - 1.0$$

The dataset used in this model is openly available and was created by Dr. William H. Wolberg, a physician at the University of Wisconsin Hospital at Madison, Wisconsin, USA. He took samples from patients who possess solid breast masses. He used a convenient graphical computer programme called Xcyt, which analyses the cytological features based on a digital scan used to create this dataset. The programme uses a curve-fitting algorithm, which enumerates ten features from each one of the cells in the sample, then it determines the mean value, standard error of each feature for the image, extreme value and, returning a 30 real-valued vector. After the data is being examined, the missing or null point data sets are removed if any using the pandas library.

Categorical data are variables that contain label values. The number of possible values is generally limited to an established set. Then the data used is split into

```
Epoch: 0 Acc = 0.87719 Cost = 298.80365 Valid_Acc = 0.94643 Valid_C
ost =  36.50434
Epoch: 1 Acc = 0.92544 Cost = 261.54120 Valid_Acc = 0.98214 Valid_C
ost =  31.51702
Epoch: 2 Acc = 0.93860 Cost = 194.79053 Valid_Acc = 1.00000 Valid_C
ost =  22.57339
Epoch: 3 Acc = 0.93640 Cost = 117.68344 Valid_Acc = 1.00000 Valid_C
ost =  11.61066
Epoch: 4 Acc = 0.94079 Cost = 71.27739 Valid_Acc = 1.00000 Valid_Co
st =  4.33530
Optimization Finished!
```

Fig. 7 Epoch values

training and testing sets. The output is known in the training set and so the model learns from this data and later the same outputs are generalised to other datas also. The test dataset (or subset) is in order to test the model's prediction on this subset.

Training set 80%	Testing set 20%

Now the datasets will have features with highly varying magnitudes, units and range. So using scaling, the data sets are brought to the approximately equal magnitude level and transformed to fit a specific scale (0–100 or 0–1). There are only two sets of values that are either malignant or benign. So the classification algorithm of supervised learning is used. This classification algorithm is used to categorise values. Then Logistic Regression is used as the method of model selection. Now the testing set results are predicted along with accuracy of each model (Figs. 7 and 8).

$$\text{Accuracy} = \frac{\text{number of correct predictions}}{\text{Total number of predictions made}} = 92.3666$$

3 Comparison

This module uses a *Logistic Regression Algorithm* and *Ski Kit Learn* as a library. The compared project uses the *Resnet Algorithm.* For the accuracy check, the importation of confusion matrix method of metrics is done. When the number of misclassifications is calculated, then it is called a confusion matrix. It is the number of predicted classes which ended up in a wrong classification bin based on the true classes. This model has an accuracy of 0.95 and the compared model has an accuracy of 0.895. Using machine learning algorithms for breast cancer risk predictions and

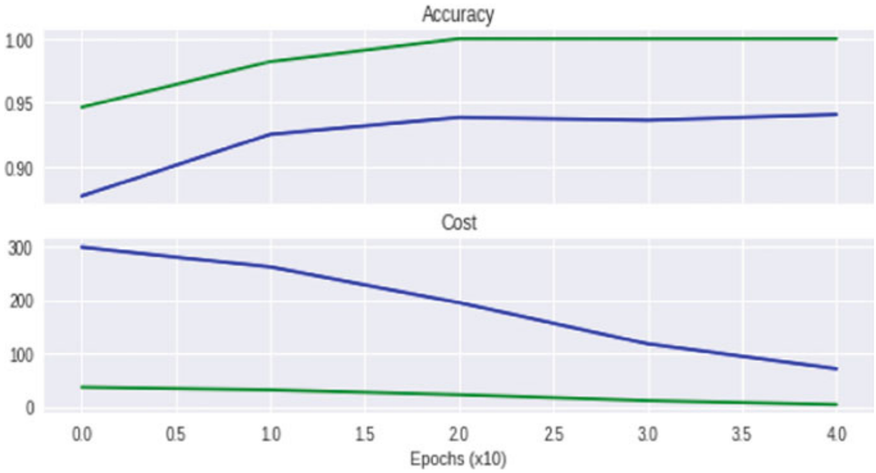


Fig. 8 Accuracy and cost chart

Table 3 Comparison of various models

Parameters	Detection of breast cancer lesions using deep neural networks	Deep neural networks improve radiologists performance in breast cancer screening	Using machine learning algorithms for breast cancer risk predictions and diagnosis
Algorithm	Logistic regression	ResNet algorithm	Support vector machine
Accuracy	0.95%	0.85%	0.99%
Requirements	Adam optimizer	Import the confusion matrix method of metrics class	Standardisation is needed for high accuracy
Time	Less	More	More

diagnosis paper uses *support vector machine algorithms*. But actually to get high accuracy using this algorithm it should be standardised which adds an extra step. This algorithm uses supervised learning. The image fed should be clear enough for further processing. It then maximises the distance of each class from the hyperplane. This avoids overfitting and undercutting using *k*-fold cross validation (Table 3).

4 Conclusion and Discussion

The model helps in the automatic detection of breast cancer lesions since the existing method is manual, time-consuming and has chances of incorrect detection. Deep neural network technique proves to be effective in training the datasets as the accuracy is almost around 95%. The work of the radiologists will be less as the software defines the type of cancer observed in the image fed. In existing methods, biopsy tests have to be done to find the type of cancer which involves incision in the skin which is kind of a complicated method. Another method is giving contrast in IV to find the difference in colour of the tissue which is identified as defected to assure 100% whether it is a cancerous cell or not. So comparatively this automatic detection helps in reducing the risk of the patients. Yet, the model is simple and not clinically implemented. The future development of this project is to classify the type of cancer using BI-RADS classification and proper implementation in real life.

References

1. Fear EC, Meaney PM, Stuchly MA (2003) Microwaves for breast cancer detection? IEEE Potentials 22(1):12–18
2. Hagness SC, Taflove A, Bridges JE (1998) Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Fixed-focus and antenna-array sensors. IEEE Trans Biomed Eng 45(12):1470–1479
3. Bassi M, Caruso M, Khan MS, Bevilacqua A, Capobianco AD, Neviani A (2013) An integrated microwave imaging radar with planar antennas for breast cancer detection. IEEE Trans Microw Theory Tech 61(5):2108–2118
4. Ruvio G, Solimene R, Cuccaro A, Ammann MJ (2013) Comparison of noncoherent linear breast cancer detection algorithms applied to a 2-D numerical model. IEEE Antennas Wirel Propag Lett 12:853–856
5. Joachimowicz N, Conessa C, Henriksson T, Duchene B (2014) Breast phantoms for microwave imaging. IEEE Antennas Wirel Propag Lett 13:1333–1336
6. Wilson TE, Helvie MA, August DA (1994) Breast cancer in the elderly patient: early detection with mammography. Radiology 190(1):203–207
7. Ikedo Y et al (2007) Development of a fully automatic scheme for detection of masses in whole breast ultrasound images. Med Phys 34(11):4378–4388
8. Moon WK et al (2011) Comparative study of density analysis using automated whole breast ultrasound and MRI. Med Phys 38(1):382–389
9. Lo CM et al (2014) Multi-dimensional tumor detection in automated whole breast ultrasound using topographic watershed. IEEE Trans Med Imaging 33(7):1503–1511
10. Tan T, Platel B, Mus R, Tabar L, Mann RM, Karssemeijer N (2013) Computer-aided detection of cancer in automated 3-D breast ultrasound. IEEE Trans Med Imaging 32(9):1698–1706
11. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on, 2005, vol 1. IEEE, pp 886–893
12. de Vos B, Wolterink J, de Jong P, Leiner T, Viergever M, Isgum I (2017) ConvNet-based localization of anatomical structures in 3D medical images. IEEE Trans Med Imaging
13. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp 3431–3440
14. Geras KJ, Wolfson S, Shen Y, Wu N, Kim S, Kim E, Heacock L, Parikh U, Moy L, Cho K (2017) High-resolution breast cancer screening with multi-view deep convolutional neural networks. arXiv:1703.07047
15. Wu N, Geras KJ, Shen Y, Su J, Kim SG, Kim E, Wolfson S, Moy L, Cho K (2018) Breast density classification with deep convolutional neural networks. In: ICASSP, 2018
16. Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A (2017) Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest Radiol 2017

17. Heath M, Bowyer K, Kopans D, Kegelmeyer P, Moore R, Chang K, Munishkumaran S (1998) Current status of the digital database for screening mammography. Digit Mammography
18. Litjens G et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
19. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional net- works for biomedical image segmentation. In: Proceedings of MICCAI, 2015, pp 234–241
20. Dou Q et al (2017) 3D deeply supervised network for automated segmentation of volumetric medical images. Med Image Anal 41:40–54
21. Lin D, Dai J, Jia J, He K, Sun J (2016) ScribbleSup: scribble- supervised convolutional networks for semantic segmentation. In: Proceedings of CVPR, 2016, pp 3159–3167
22. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y (2014) OverFeat: integrated recognition, localization and detection using convolutional networks (Online). Available: https://arxiv.org/abs/1312.6229

Predicting Breast Cancer Using Changing Parameters of Machine Learning Model

Shivam Bagri, Manish Pandey, and Santosh K. Sahu

Abstract Breast cancer is the most common type of fatal ailment seen in females in the world. There are many types of cancer, one of which is breast cancer. Several types of breast cancer are found in women that affecting their lives across the world. Several types include "lobular carcinoma in situ (LCIS), ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC)." How many people are dying due to cancer today, and one of the main reasons is not known in time. Generally, breast cancer may be a malignant neoplasm that begins within the cell of the breast and eventually spreads to the encompassing tissue. Due to breast cancer, a lot of death is happening. The death rate can be reduced by using machine learning techniques. Mammography is a good and effective modality that is used in the detection of breast cancer in today's time. In this paper, we used different machine learning algorithms like Naïve Bayes, *k*-nearest neighbors, logistic regression, support vector machine, decision tree, and convolution neural network. After changing the unique hyperparameter of each model, find the better accuracy within the model and also do the comparison between models. The performance of convolution neural network is found maximum accuracy with minimum loss. The accuracy achieved by convolution neural network is 99.05%.

Keywords Breast cancer · Support vector machine (SVM) · *k*-nearest neighbor (KNN) · Naive Bayes · Logistic regression · Decision tree · Convolution neural network

S. Bagri · M. Pandey (✉) · S. K. Sahu (✉)
Computer Science Engineering, Maulana Azad National Institute of Technology, Bhopal, India
e-mail: contactmanishpandey@yahoo.co.in

S. K. Sahu
e-mail: sks.173112004@manit.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_37

1 Introduction

Breast cancer is cancer that forms in the breast cells. Breast cancer mainly occurs in women. The death rate can be reduced by detecting the disease in an earlier stage. There are mainly three different stages of severity of cancer. This is possible only when patient performing various tests like biopsy, ultrasound, mammography, and MRI. Cancer cells travel to other parts of the body through our bloodstream or lymphatic system. The lymph nodes are also an important carrier of cancer cells. The cancer cells capture and infect other healthy breast tissue. Then, they travel to the lymph nodes under the arms and cause the cancer to spread [1]. This causes rapid growth in cancer cells, and our objective is to identify the malignant tumors in breasts and stop their spread to different parts of body and save the patient.

According to National Cancer Information, there are 90% of women survive to 5 years in breast cancer. Cancer treatment depends on the stage of the cancer. If the cancer detected earlier stage, then 90% chance to patient survive more than 5 years.

Supervised and unsupervised machine learning algorithms that are used to solve a complex problem in the medical field by using machine learning algorithms build a very accurate prediction in the medical field. Machine learning techniques helps to detect the cancer in an earlier stage.

2 Dataset

In these studies, dataset is taken from the breast cancer Wisconsin dataset. There are two classes in this dataset such as malignant and benign. Our dataset has 569 instances. The dataset has computed 30 features with the help of fine-needles aspirate (FNA) of breast mass. Generally, cancer datasets are in the form of images. In our dataset, the features are computed using a digitized image of FNA of a breast mass. These features describe the different characteristics of the nuclei of cells in the image. Ten real-valued features are extracted from the image for each cell nucleus. These features are described below:

(i) Radius which is the mean of distances from center to points on the perimeter
(ii) Texture that is standard deviation of the grayscale values in the image
(iii) Area of the segmented cells
(iv) Perimeter of the segmented cells
(v) Compactness whose formula is perimeter^2/area—1.0
(vi) The number of concave portions of the contour
(vii) Concavity that is the severity of concave portions of the contour
(viii) Smoothness that is local variation in radius lengths
(ix) Symmetry in the images
(x) Fractal dimension which is “coastline approximation”—1.

3 Machine Learning Technique

Our target is to predict the tumor is benign or malignant. We came up with a model that compares with other models. We will find out the numerical results we get from the algorithms. We read the various paper for breast cancer including six algorithms: SVM, KNN, Naive Bayes, logistic regression, decision tree, and convolution neural network.

3.1 Naïve Bayes

Naïve Bayes classifiers are a group that are quite similar to the linear models. They tend to be even faster in training. A classifier under supervised machine learning group is based on the probabilistic logic. Each attribute from each set is the user's probability to make the prediction. Naïve assumption makes feature. Naïve Bayes algorithm is a very strong prediction. Naive Bayes is the mainly classification algorithms. Naïve Bayes classifier is used to conditional probability [2].

$$P(C|A) = P(C) * \frac{P(A|C)}{P(A)} \tag{1}$$

where $P(C|A)$ is posterior probability, $P(C)$ is prior probability, $P(A|C)$ is likelihood probability, and $P(A)$ is evidence probability.

3.2 K-Nearest Neighbor

The k-nearest neighbor algorithm is a very simple technique that is used for both regression and classification. KNN algorithm takes input to k nearest in the training set. The KNN is a nonparametric model that is describing the instance-based learning. In the KNN algorithm, select the k number of the training datasets and find the k number of nearest neighbors of the sample space that we want to classify and then assign the class label with the help of majority votes. KNN is the memory-based approach that means required to store K values is that classifier immediately adopts as we collect new training data [3].

$$\text{Euclidean } D_e = \sqrt{\sum_{i=i}^{k}(x_i - y_i)^2} \tag{1}$$

$$\text{Manhattan } D_m = \sum_{i=1}^{k}|x_i - y_i| \tag{2}$$

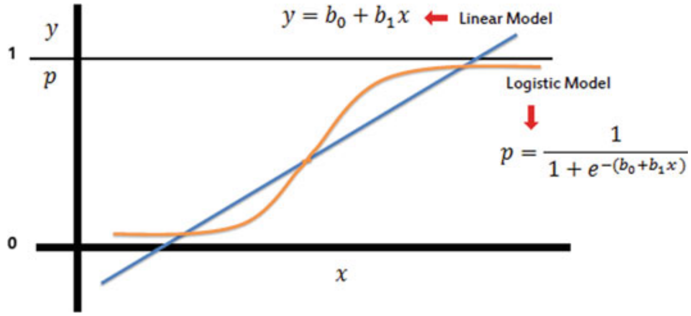


Fig. 1 Logistic regression

$$\text{Minkowski } D = \left(\sum_{i=1}^{k} (|x_i - y_i)^q \right)^{\frac{1}{q}} \tag{3}$$

3.3 Logistic Regression

Logistic regression is one of the most important techniques to predict tumor malignant (1) or not (0). Logistic regression can be used for the binary classification that produces two outcomes: for example, "spam" versus "not spam." The output is typically coded as "1" or "0," if the output is "1" that means the prediction is correct. The logistic function changes the real-value input to outputs 0 and 1 [4] (Fig. 1).

3.4 Decision Tree

Decision trees are widely used in classification and regression tasks. They learn a hierarchy of if/else questions, leading to a choice [5]. We use real systems for mentioning qualities as the source or within a center point. As ought to be clear from that decision tree manages the sum of product structure which is generally called disjunctive normal form. Decision tree classifier is very attractive model as the name suggests that decision tree breaks down our data by making decision based on asking a series of question. In decision tree, each internal node represents as the test and branch of the tree represent as the outcome of the test. The three impurity performance measures that commonly used binary decision tree is entropy, Gini index, and the classification error (Fig. 2).

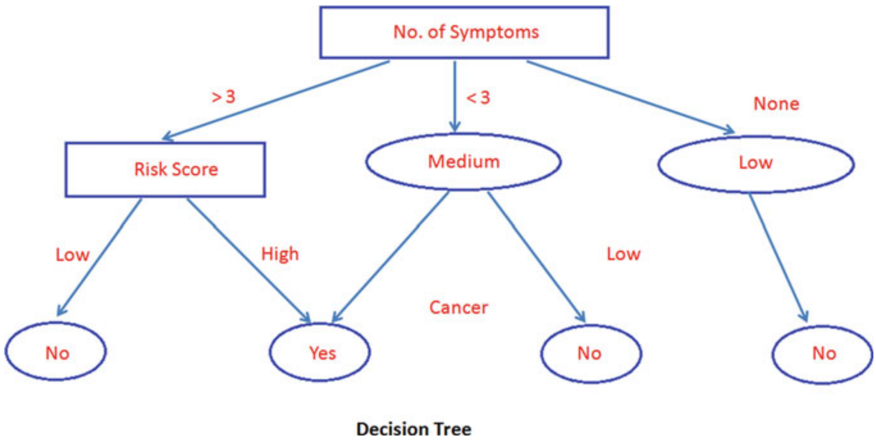


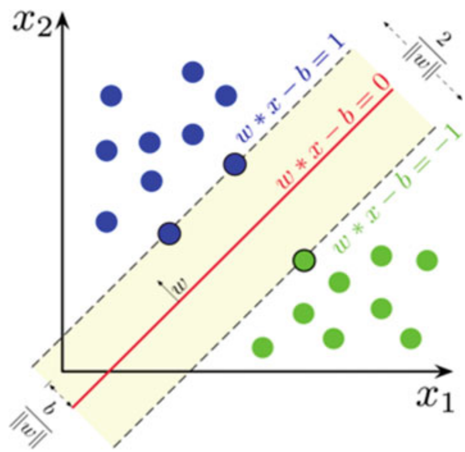
Fig. 2 Decision tree

3.5 *Support Vector Machine (SVM)*

The SVM is very powerful and popular algorithms used for both supervised tasks: classification and regression. Support vector machine is optimize the objective to maximize the margin. Margin in SVM is the distance between the two hyperplanes, and the support vector is closest to this hyperplane. The main reason why SVM is so popular because it also solves nonlinear classification problem by using kernel SVM [6] (Fig. 3).

SVM mainly focuses on finding the hyperplane by dividing the region into separate classes. Hyperplane equation is as follows:

Fig. 3 Linear SVM



$$w \cdot x + b = 0 \tag{1}$$

where W is take the weight vector, X is take the input vector, and b is take the bias.

With help of above equation, we can write:

$$w \cdot x + b \geq 0 \, \text{for} \, d_i = +1 \tag{2}$$

$$w \cdot x + b < 0 \, \text{for} \, d_i = -1 \tag{3}$$

where d_i is the margin of separation.

3.6 Convolutional Neural Network (CNN)

Deep learning is an advanced version of a machine learning algorithm. A CNN is a deep learning technique that solves the nonlinear property of a neural network. CNN is used for image classification. In our model, CNN has one input layer, two hidden layers, and one output layer. Backpropagation algorithms are used to train the neural network, and the Adam gradient descent cost function is used for optimization. ReLU activation function is used in the different hidden layers. Sigmoid activation function is used to convert output into two classes for output layer.

4 Performance Evaluation Results

The experiment has been performed on Google Colab Notebook. In our experiment, import varies library function such as NumPy use for array, and perform the matric operation; Pandas is used to analyze the data; Matplotlib is used for plotting the graphs; TensorFlow and Keras are used for training and testing the model.

Performance of the machine learning models can measure with the help of accuracy (AC), precision (PR), recall (RE), and F1-score($F1$) in the term of positive and negative.

$$\mathrm{AC} = (\mathrm{TP} + \mathrm{TN}) * \frac{100}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}} \tag{4}$$

$$\mathrm{PR} = \mathrm{TP} * \frac{100}{\mathrm{TP} + \mathrm{FP}} \tag{5}$$

$$\mathrm{RE} = \mathrm{TP} * \frac{100}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

Table 1 Performance evaluation of Naïve Bayes

	Precision	Recall	$F1$-score	Support
0	87.21	93.23	91.12	63
1	96.13	93.65	94.32	108

Table 2 Accuracy of KNN with changing parameter K

Neighbors	$K = 1$	$K = 3$	$K = 5$
Accuracy	92.13	94.75	95.64

$$F1 = 2 * 100 * \frac{\text{Recall*Precision}}{\text{Recall } + \text{Precision}} \tag{7}$$

Here

TN True negative
TP True positive
FP False positive
FN False negative.

4.1 Naïve Bayes

Our experiment Bernoulli Naive Bayes classifier used and found the accuracy is 93%.

Table 1 shows the performance evaluation of the Bernoulli Naïve Bayes.

4.2 K-Nearest Neighbor

K-nearest neighbor has been performing with three different values of K show in Table 2. Found the maximum accuracy at $K = 5$.

Figure 4 shows the distribution of a class of begin and malignant class on the train and test subset and predicts the result with help of the majority voting system.

Table 3 shows the performance evaluation of KNN with value of $K = 3$.

4.3 Logistic Regression

An experiment perform for three regularization hyperparameter C is shown in Table 4. Found best accuracy at $C = 100$.

Figure 5 shows the coefficient learned by logistic regression with penalty

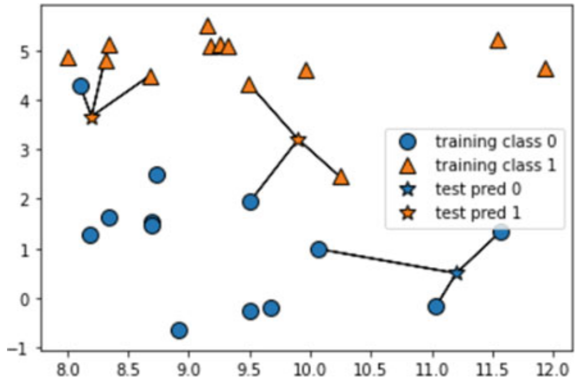


Fig. 4 Class distribution

Table 3 Performance evolution of KNN

	Precision	Recall	$F1$-score	Support
0	98.12	91.34	95.76	47
1	94.32	99.65	96.87	67

Table 4 Accuracy for logistic regression with different regularization parameter

Regularization hyperparameter	$C = 0.01$	$C = 2$	$C = 100$
Accuracy	93.13	95.54	96.27

parameters for different values of C.

Table 5 shows the precision, recall and $F1$-score for logistic regression with regularization parameter $C = 100$.

4.4 Decision Tree

The decision tree applied for two different depths is shown in Table 6 and found the best accuracy at depth is 4.

Table 7 shows the performance evaluation of the decision tree with max_depth = 4.

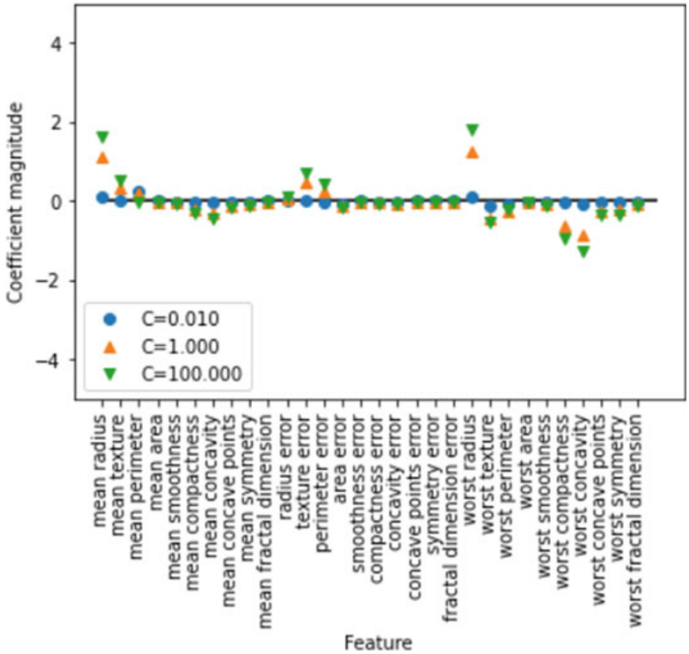


Fig. 5 A graph between coefficient magnitude and feature

Table 5 Performance evaluation of logistic regression

	Precision	Recall	$F1$-score	support
0	91.23	97.54	94.65	63
1	98.54	94.34	96.87	108

Table 6 Accuracy of decision tree

Depth	Max_Depth = 2	Max_Depth = 4
Accuracy	90.96	95.71

Table 7 Performance evaluation of decision tree

	Precision	Recall	$F1$-score	Support
0	94.54	92.76	93.87	53
1	96.43	97.98	96.65	90

Table 8 Result of different parameter C

Parameter	$C = 1$	$C = 10$	$C = 100$
Accuracy	92.96	93.78	94.79

Table 9 Performance evaluation of SVM

	Precision	Recall	$F1$-Score	Support
0	96.32	91.54	93.76	97
1	94.12	97.43	96.45	67

Table 10 Accuracy and loss for epoch 1–100

Epoch	Epoch = 1	Epoch = 100
Accuracy	60.28	99.05
Loss	79.47	2.88

Table 11 Performance evaluation of CNN

	Precision	Recall	$F1$-score	Support
0	1	94.87	97.98	47
1	96.76	1	98.67	67

4.5 Support Vector Machine (SVM)

The performance of SVM has been evaluating for different regularize parameter C is shown in Table 8 with redial basis function and found the best accuracy at $C = 100$.

Table 9 shows the performance evaluation of SVM with regularization parameter value $C = 100$.

4.6 Convolutional Neural Network

CNN has been done for different number of epoch. Our experiment convolutional neural network trains from epoch = 1 to 100 to maximize the accuracy and minimize the loss is shown in Table 10.

Table 11 shows the performance of CNN.

Figures 6 and 7 show the accuracy and loss model.

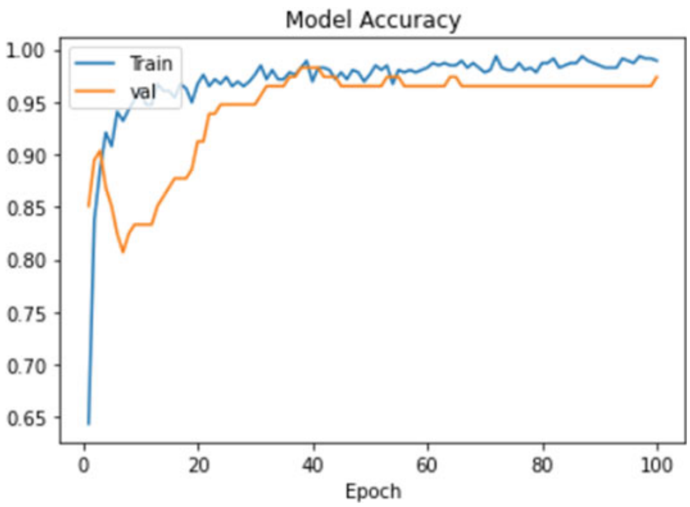


Fig. 6 Accuracy model

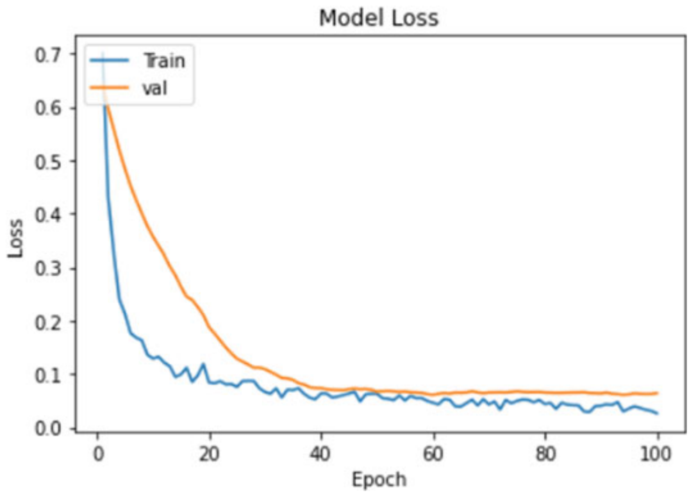


Fig. 7 Loss model

5 Conclusion

In this paper, the performance of six different machine learning classification models is evaluated for the Wisconsin breast cancer dataset. The best accuracy is obtained using the convolutional neural network with the help of Adam optimization algorithm. The accuracy we get using CNN is 99.05%. Medical images play a very important role in breast cancer classification. Hyperparameter tuning on deep learning models and CNN gives us very good result on our dataset and can be used in other medical image classification tasks.

References

1. National Breast Cancer Foundation Inc. http://www.nationalbreastcancer.org/about-breast cancer
2. Medjahed SA, Saadi TA, Benyettou A (2013) Breast cancer diagnosis by using k-nearest neighbor with different distances and classification rules. Int J Comput Appl 62(1): 0975–8887
3. Karplus A (2012) Machine learning algorithms for cancer diagnosis. Santa Cruz County Science Fair 2012
4. Hadush S, Girmay Y, Sinamo A, Hagos G (2020) Breast cancer detection using convolutional neural networks
5. Gupta K, Chawla N (2020) Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained CNN. Procedia Comput Sci 167:878–889. https://doi.org/10.1016/j.procs.2020.03.427
6. Sarkar M, Leong TY (2000) Application of K-nearest neighbor's algorithm on breast cancer diagnosis problem. In: Proceedings of the AMIA symposium. American Medical Informatics Association, p 759

Extended Kalman Filter-Based Position Estimation in Autonomous Vehicle Applications

Manjari Singh, Shikha Lakra, Subhranil Das, Sudhansu Kumar Mishra, Ajit Kumar Sahoo, and Bibhudendra Acharya

Abstract As there is great demand in self-operating driving cars, probabilistic-based model is one of the emerging areas while considering the estimation of position of vehicles. In this paper, two probabilistic models are implemented in application to autonomous ground vehicles. Bayesian-state estimator has been repeatedly used for estimating the position as well as state of the vehicle. In this paper, simulation results are executed for Kalman filter (KF), extended Kalman filter (EKF) for predicting the positions of vehicle. A brief description on improvement of each method is being discussed by drawing comparison table.

Keywords Bayesian estimation · EKF · UKF · SLAM

1 Introduction

An autonomous ground vehicle is defined as intelligent vehicle without any involvement of human operators in all types of situations. For avoidance of any obstacles, a defined path is being formulated where navigation plays an important role. Considering the safety of passengers, the only goal is to reach the destination point without making any collisions. As a result, a number of methods have been applied by various researchers for controlling the motion of vehicle [1]. Planning of path as well as taking

M. Singh · S. Lakra · S. Das · S. K. Mishra
Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India
e-mail: sudhansumishra@bitmesra.ac.in

A. K. Sahoo
Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela, Odisha, India
e-mail: ajitsahoo@nitrkl.ac.in

B. Acharya (✉)
Department of Electronics and Communication Engineering, National Institute of Technology, Raipur, Chhattisgarh, India
e-mail: bacharya.etc@nitrr.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_38

decisions for autonomous vehicle have been regarded as formidable tasks where high level defines as long horizon controls and low level defines as short horizon controls [2]. The algorithms related to AGV implementation pose a great challenge in actual scenario in real world in controlling vehicle. The positioning system is one of the vital components in navigation system is constructed by DSO National Laboratories. In order to track the position of the vehicle in the real time, navigation system plays an important role [3].

Autonomous ground vehicle (AGV) has been grabbing much attention toward safety and proper path deployment. Autonomous navigation system has three basic modules which are: sensing and perception, planning, and controlling [4]. Here, the function of path following which lies in the control module is the most basic rule. This specifies that the vehicle will follow the designed path without any parameterization [5].

In autonomous ground vehicle, path planning has taken a great amount of attention for researchers across the world. For defining the path, a vehicle should organize its own path and without confronting obstacles in its path. AGV should be intelligent enough while traversing the path from its initial position to final position. There were many previous approaches, where the constrained optimization has been utilized. Finite number of optimal points are taken where these points have been proved to be an effective one. While designing any algorithm, there are some inherent noises, uncertainties, disturbances, errors in model, etc., that results in inaccuracies in path definition. Therefore, it is foremost thing to include all of the uncertainties for ameliorating the path accuracy [1].

Since conventional constrained optimizations approach is more difficult to implement, probabilistic-based approach has provided an alternative solution for handling uncertainties in an easier way where it utilizes random learning as well as query method [6].

In probabilistic-based approach, the very first step is to constitute information in the form of probability densities. Second step for explaining the idea is though perception where sensors sense the data. Last, decisions are to be taken by the process of action [7].

Currently, linear chance constraints are being developed while designing succession of continuous inputs of control signals that are applicable mainly on predictive control of probabilistic models [8]. Therefore, linear constraints work in tandem with definite set of probabilities. For detecting any kind of irregularities in the defined trajectory, Campo et al. used the Kalman filter [10] for this purpose. Filter banks are used for complicated dynamic systems and were included in system module.

In previous 30 years, there have been considerable amount of research taken place on AGVs. Considering these successes, it can be estimated that there will be a fully automated vehicle introduced that may replace the human drivers. At present, there are various assistance systems that help the driver in better driving and increase the overall safety. These assistance systems are, namely (a) collision avoidance, (b) lane departure warning, and (c) blind spot detection [11]. Pagac et al. [12] have applied the theory of evidence in their work of map building for the operation of the autonomous vehicle. The uncertainties of the environment were captured by the ultrasonic sensors installed in the autonomous vehicle.

The intention for the origination of AGVs is to improve the efficiency of transportation, utilization, and road security. This is a rapidly changing and complex environment, and AGVs need to have high dependability, efficiency along with robustness [13]. Path tracking is among the most important objectives for the control of the motion of the AGVs. This helps the vehicle to commute on the path desired by the path planner with undefined disturbances, model uncertainties, and unavoidable physical factors [14]. Most effective path-tracking feedback control method mainly depends on the measured state of the vehicle [15]. It is hard to obtain the state of the vehicle with low-cost sensors. For this state measurement, we employ some filters, namely: KF, EKF, and UKF that are used frequently to estimate the state of the vehicle. Moreover, sometimes, probabilistic-based optimization techniques such as SLAM are used.

Estimation techniques used for estimating the state of the vehicle can be classified under two categories, i.e., filter-based approach and optimization-based approach as shown in Fig. 1.

On the basis of different methods used in probabilistic measurement, we have shown mathematical equations of Kalman filter, extended Kalman filter, and unscented Kalman filter in next sections. Simulation studies of abovementioned filters have been carried out, and position error and position accuracy are tabulated.

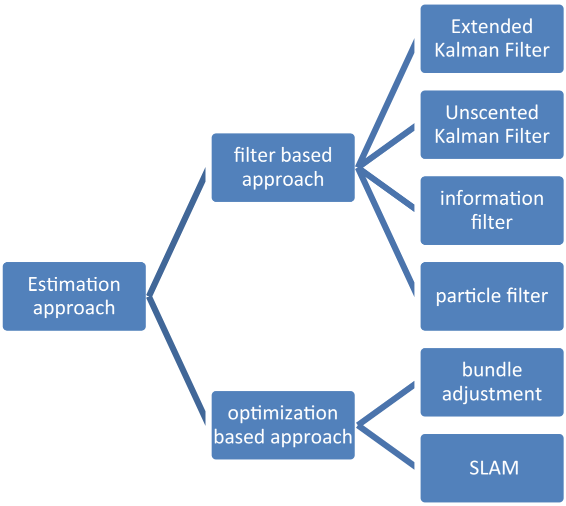


Fig. 1 Classification of estimation techniques used in autonomous vehicle navigation

2 Related Works

2.1 Kalman Filter (KF)

KF was proposed by a scientist named Kalman in the year 1960. KF is a tool that is used to predict a value using some sets of equations taking into assumption that the data being used are in the form of a Gaussian distribution. There is a need to apply a set of linear equations to this Gaussian distribution. Traditional KF depends on the precise vehicle model, and it is assumed that the model and errors or the calculated noise are the Gaussian white noise with its mean value being numerically zero [15]. Its primary objective is to keep record of the dynamic changes in the presence of noise and incomplete values. KF is applied to predict the state, and then, using those predicted measurements, it updates the predictions.

Linear discrete Kalman filter for state dynamic and measurement models:

$$z_k = H_k x_k + v_k \tag{1}$$

x_k is the $(n \times 1)$ state vector, $\emptyset_k$ is the $(n \times n)$ transition matrix, z_k is the observation vector, and H_k is the observation matrix. w_k and v_k are white noises with its mean and covariance as following.

$$E\{w_k\} = E\{v_k\} = 0 \tag{2}$$

$$E\left\{w_k v_k^T\right\} = 0 \tag{3}$$

$$E\left\{w_k w_k^T\right\} = \begin{cases} Q_k, (i = k) \\ 0, (i \neq k) \end{cases}. \tag{4}$$

$$E\left\{v_k v_k^T\right\} = \begin{cases} R_k, (i = k) \\ 0, (i \neq k) \end{cases}. \tag{5}$$

$E\{.\}$ is the expected operator. The Q and R are the covariance matrix of the process noise and the measurement error, respectively.

KF state prediction and state covariance prediction are as follows:

$$\overline{x_k} = \emptyset_{k-1}\widehat{x_{k-1}} \tag{6}$$

$$\overline{P_k} = \emptyset_k \hat{P}_{k-1} \emptyset_{k-1}^T + Q_{k-1} \tag{7}$$

$\hat{x}_k$ denotes the estimated state vector, x_k is the predicted state vector for the next epoch, $\hat{P}_k$ estimated state covariance matrix, and P_k is the predicted state covariance matrix.

The Kalman filter updates are as follows:

$$K_k = \overline{P}_k H_k^T \left(H_k \overline{P}_k H_k^T + R_k\right) \times^{-1} \tag{8}$$

$$v_k = z_k - H_k \overline{x}_k \tag{9}$$

$$\hat{x}_k = \overline{x}_k + K_k v_k \tag{10}$$

$$\hat{P}_k = (I - K_k H_k) \overline{P}_k \tag{11}$$

K_k Kalman gain, used to update the predicted value and error in prediction value obtained previously from the dynamic model of the system.

However, KF had some shortcomings. It needs a pre-planned mathematical model and statistical properties of the desired system. It is difficult to achieve in this complex and dynamic environment [16]. To get over the difficulties of KF, adaptive Kalman filter (AKF) is developed by researches [17]. It uses the measured values to do the recursive filtering along with constantly estimating and modifying the model parameter along with its noise covariance matrix to reduce the error. It makes the system divergence small from the desired system performance [17]. It was seen that the basic Kalman filter and these modifications gave high performance only for the liner system. Since the actual system is highly complex and nonlinear, further modifications in its algorithm are needed.

It has been seen that the autonomous vehicle needs more than one method to perform a single task, for example, its positioning. Shoval et al. [18], in their paper, presented the implementation of Kalman filter for the fusion of the data by equipping various sensors in the autonomous vehicle. It is seen that the filter detected and corrected the position (lateral as well as angular) error and disturbances in a radius of 15 cm within 3–4 s. Sensitivity analysis showed that the error increased linearly with speed. Based on these analysis, vehicle speed and the layout of the landmark are determined in accordance to the process specification, and finally, the sensors data are used by the filter [18]. Sasiadek et al. [19] proposed a method of data fusion based on the adaptive fuzzy Kalman filtering (AFKF). These data were collected using sensors. This method is enforced to combine the position signals from the GPS and INS which is the inertial navigation system. Fuzzy logic adaptive system (FLAS) is implemented for the prevention of the system from divergence. Tang et al. [20] presented the use of multi-model adaptive system (MMAS) unmanned vehicles, where multiple Kalman filter was run in parallel using different dynamic models in GPS-integrated system navigation as the traditional KF has many limitations in the GPS-integrated navigations. Compared to the traditional KF algorithm, their newly developed algorithm had better position and velocity accuracy. Despite recent advancements in the autonomous vehicles, it is still a challenging task to keep the vehicle model in the right lane using the image sensing. Reyes et al. proposed an algorithm that could be tested on the GPS-denied environments or where the environment did not change rapidly [21]. This algorithm is tested in the model designed in the University of Berlin. This model used sensors and processors, such as gyroscopes, accelerometers, motor encoders, camera pointing to the front of the vehicle, one fish eye camera pointed to the roof, an Odriod XU-4 companion computer, and an Arduino nano.

2.2 EKF and UKF

Traditional KF is used for linear systems. However, in real-world scenario, nonlinear systems have to be handle that need to apply nonlinear equations. The EKF uses Taylor series or Jacobian matrix to approximate a nonlinear system linearly around the mean of the Gaussian. Then, it predicts the values. EKF is the most favorable estimation method for the nonlinear systems [22]. Estimation of the state of the nonlinear dynamic system also includes parameter identification and dual estimation in which both state and parameter are simultaneously estimated [23]. For strong nonlinear systems, it is very complicated to compute the Jacobi matrix. UKF was proposed to overcome this difficulty.

State estimation: EKF involved estimating state in the discrete time nonlinear system

$$x_{k+1} = F(x_k, v_k) \tag{12}$$

$$y_k = H(x_k, n_k) \tag{13}$$

where x_k is unobserved state of the system, y_k is the observed signal, v_k is process noise which drives the dynamic system, and n_k is the observation noise. F and H are system dynamic models which are the assumed knowns parameters.

Parameter estimation: It involves determination of a nonlinear mapping.

$$y_k = G(x_k, w) \tag{14}$$

where x_k is the input, y_k is the output, and the vector w parameterizes the nonlinear map G.

A training set is provided with pairs of samples containing known inputs and desire outputs, x_k, d_k. The error is defined as:

$$e_k = d_k - G(x_k, w) \tag{15}$$

The new state space representation:

$$w_k = w_{k-1} + u_k \tag{16}$$

$$y_k = G(x_k, w_k) + e_k \tag{17}$$

where w_k is a stationary process with identity state transition matrix, and u_k is the process noise.

Dual estimation

$$x_{k+1} = F(x_k, v_k, w). \tag{18}$$

$$y_k = H(x_k, n_k, w). \tag{19}$$

where the system state 'x_k' and the set of model parameters 'w' for the dynamic system are calculated simultaneously from the observed noise signal y_k.

However, it is seen that the increasing amount of nonlinearity and wrong modeling of the system greatly affected the performance of the EKF, and it diverged from the true desired dynamics.

Jetto et al. [24] presented in their paper a new method for the accurate localization of the automatic vehicle. Here, the data provided by the sonar sensors and the odometer are combined together though EKF to approximate the position of the vehicle. The newly developed algorithm improved the filter performance, by an online adjustment of the inputs and the measured noise covariance. Adaptive mechanism is provided to minimize the divergence of the system and cope with the realistic environment. The study of the literature by Julier et al. [25] concludes that the vehicle model plays a significant role in the performance of the sensor-based navigation system of the autonomous vehicle. The KF and the EKF are the most widely used localization methods for the autonomous vehicle. Since the control system uses a prior information for the online localization, the accuracy of the vehicle model can determine how the initial information is utilized effectively. Rusdinar et al. have proposed localization and navigation of the automatic vehicle based on the EKF and real-time image processing [26]. They proposed the distinguishable landmark recognition using image processing and navigation algorithms. It used only odometer sensors and a low-cost camera for the indoor environment. Any disturbance from static or dynamic obstacle is eliminated using ceiling landmarks.

Almeida et al. have developed autonomous navigation system for the small-scale vehicle whose sensory system comprised of a low-cost inertial measuring unit and a standard GPS receiver. Onboard processor send data to the ground control station

(GCS) from the onboard sensors, and vehicle then receives commands for navigation from it. The GCS software used EKF for combining the sensor data [27].

UKF is one of the well-known filter used in countless practical systems where some of the states of the system cannot be evaluated, achieved, or measured. Missing status of the state can be estimated with the help of state observer. The fact that this filter requires a prior knowledge of the noise covariance is its major drawback [28]. It is based on unscented transform theory and statistical linearization techniques. It is used for linearization of nonlinear function of a non-specific variable through a linear regression between n points taken from the previous distribution of the random variable [29]. In terms of speed of convergence and its robustness, UKF is preferable than EKF [30]. It has been seen that UKF has better accuracy than EKF at the same level of complexity. The information about the vehicle's state plays a very important role in the vehicles active safety system. In the paper by Ren et al. [22], the simulation results were compared to the numerical results using the Carsim software, which concluded that the UKF estimates the information of the state of the vehicle more accurately and in real time.

Qasem and Reindl [30] in their study analyzed the algorithms and mathematical modeling of the EKF and the UKF. The results showed that the UKF is more accurate and reliable than the EKF whose implementation is easier when it had better prior information of the estimate of the vehicle. These algorithms were programmed in MATLAB. This concluded that UKF should be preferred in the nonlinear and dynamic environment.

3 Simulation Results

For understanding the performance of Kalman filter and extended Kalman filter, a short simulation study is carried out by taking the required number of sample observations for estimating the positions of vehicle. For measuring the positions of next time, i.e., $(t + 1)$ interval, the previous position of vehicle is governed by a set of equations which has been previously described above. A sample set of next 30 positions has been considered for execution as displayed in Fig. 2.

From the above figure, there difference is observed between the actual position and predicted positions of vehicle measurement. There are certain positions in vehicle where measurements are accurately being measured. Difference in positions of error is displayed in Fig. 3.

In this figure, error is largely fluctuating with respect to position. For reducing error while predicting the next positions of vehicle, extended Kalman filter (EKF) is applied which is shown in (Fig. 4).

In this figure, the variations occurred in different positions of vehicle positions in next time $(t + 1)$ position are very small as compared to Kalman filter. The error measured between actual position and predicted position is being shown in Fig. 5.

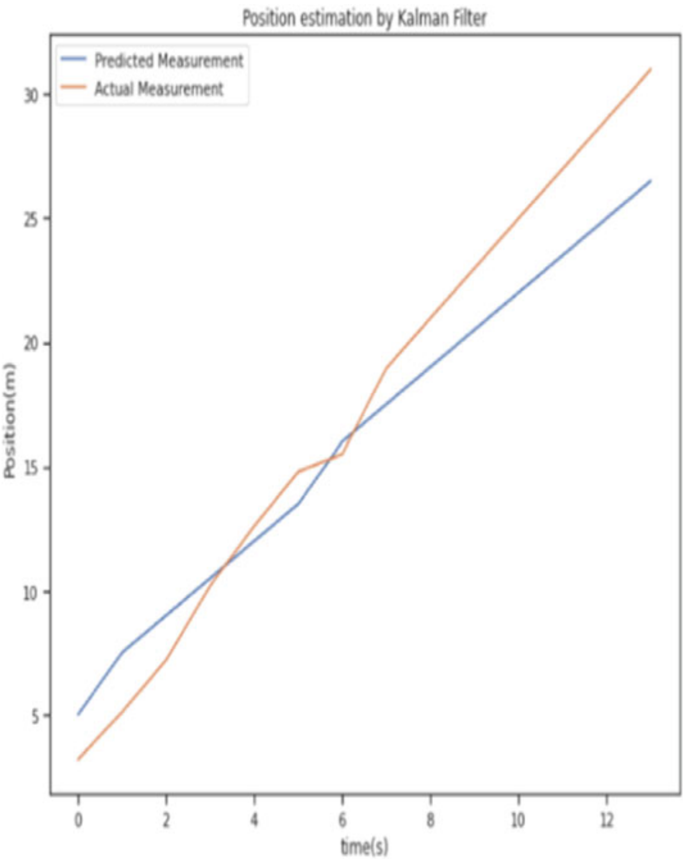


Fig. 2 Position estimation by Kalman filter

In order to sum up, various simulations executed by Kalman filter and extended Kalman filter (EKF), EKF has produced better accuracy position results and low error minimization while finding the next position of vehicle. The comparison between the different filters is explained in (Table 1).

4 Discussion

In this section, the major outcomes of work done are highlighted. Firstly, vehicle localization methods apply Kalman filter for estimating the location of the vehicle using local sensors or the GPS, which can estimate the probabilistic distribution over the possible locations. Secondly, the estimated value of the disturbance is determined. Lastly, the desired level of change in the planning of the path can be successfully

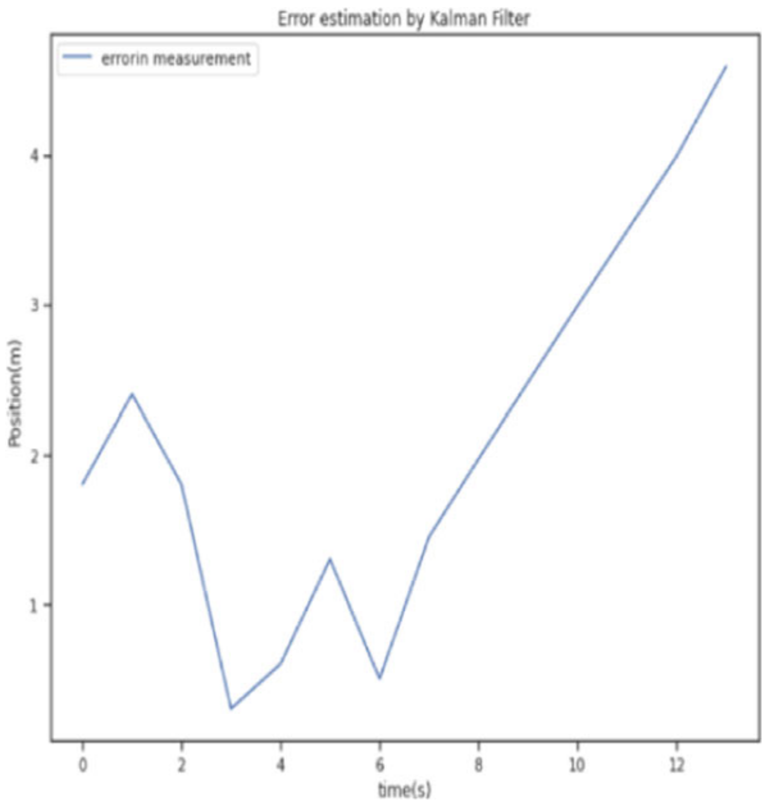


Fig. 3 Error estimation by Kalman filter

specified, by stipulating the probability of the plan being successfully executed. Also, we have executed a sample test on 50 next positions of vehicle. From this simulation study, it is observed that extended Kalman filter (EKF) made better approximations in position accuracy as well as reduction in error variations. EKF is found to have high efficiency but only for the linear systems. To increase the efficiency and accuracy in linear system, combination of EKF and Kalman filter is used. When compared at the same level of complexity, it is observed that the EKF proved to be more efficient than KF.

5 Conclusion

In this paper, three different algorithms, such as Kalman filter and extended Kalman filter, are quite influential in estimating the states in controlling the position. It has been concluded that the extended Kalman filter results are most apprising one

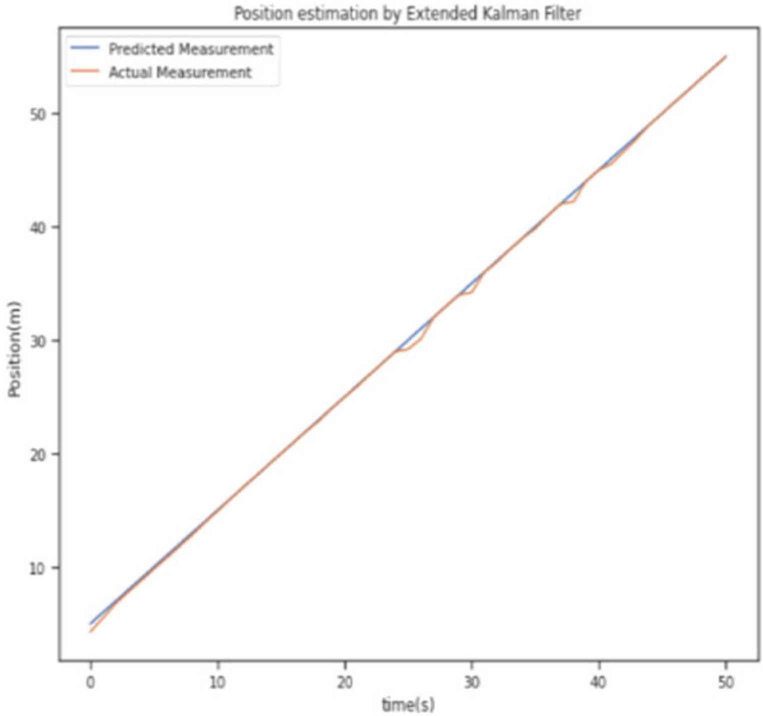


Fig. 4 Position estimation by EKF

and better performance while comparing with the output obtained from Kalman filter. Apart from these three important probabilistic approaches, some of the recent machine learning techniques may be hybridized with them and can be suitably applied to the challenging field of autonomous vehicles. These hybridized techniques will not increase the accuracy but also reduction in calculation of error applicable to the field of autonomous vehicles.

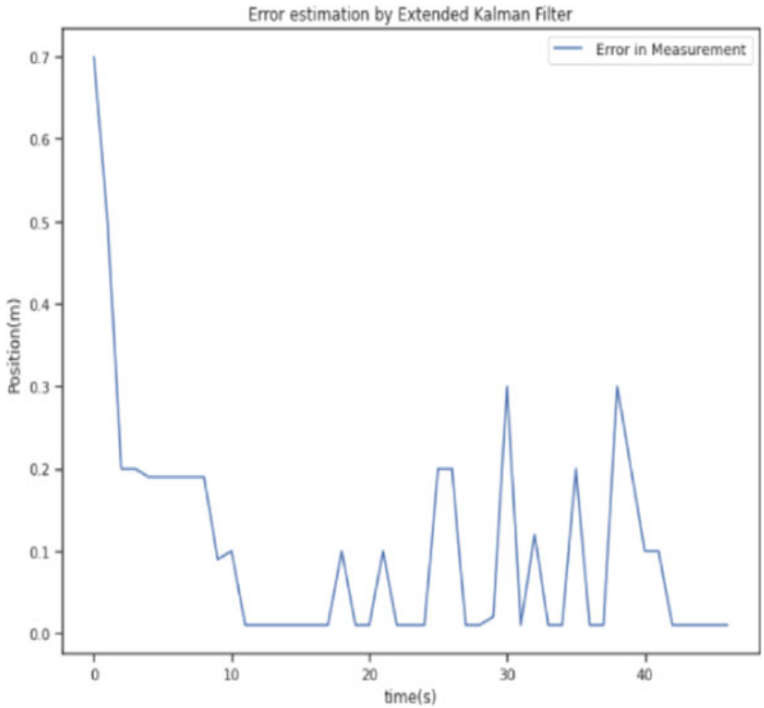


Fig. 5 Error estimation by EKF

Table 1 Comparison of position error and accuracy with other filters

Filter	Position error percentage	Position accuracy percentage
KF	65	63
UKF	60	32
EKF	15	17

References

1. Blackmore L, Li H, Williams B (2006) A probabilistic approach to optimal robust path planning with obstacles. In: American control conference. IEEE, 2006 p 7
2. Pouya P, Madni AM (2020) Leveraging probabilistic modeling and machine learning in engineering complex systems and system-of-systems. In: AIAA Scitech forum. 2020, p 2117
3. Teo TY, Chai JS, Yao WL (2002) Design of a positioning system for AGV navigation. In: 7th International conference on control, automation, robotics and vision, ICARCV, vol 2. IEEE, 2002, pp 637–642
4. Levinson J, Askeland J, Becker J, Dolson J, Held D, Kammel S, Kolter JZ, Langer D, Pink O, Pratt V, Sokolsky M (2011) Towards fully autonomous driving: Systems and algorithms. In: IEEE intelligent vehicles symposium (IV). IEEE, 2011, pp 163–168
5. Mac Thi T, Copot C, De Keyser R, Tran TD, Vu T (2016) MIMO fuzzy control for autonomous mobile robot. J Autom Control Eng 4(1): 65–70

6. Aguiar AP, Hespanha JP (2007) Trajectory-tracking and path-following of underactuated autonomous vehicles with parametric modeling uncertainty. IEEE Trans Autom Control 52(8): 1362–1379
7. Kavraki LE, Svestka P, Latombe JC, Overmars MH (1996) Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Trans Robot Autom 12(4): 566–580
8. Thrun S (2000) Probabilistic algorithms in robotics. Ai Mag 21(4): 93–93
9. Yan J, Bitmead RR (2005) Incorporating state estimation into model predictive control and its application to network traffic control. Automatica 41(4): 595–604
10. Campo D et al (2020) Learning probabilistic awareness models for detecting abnormalities in vehicle motions. IEEE Trans Intell Transp Syst 21(3):1308–1320
11. Vitus MP, Tomlin CJ (2013) A probabilistic approach to planning and control in autonomous urban driving. In: 52nd IEEE conference on decision and control. IEEE, 2013, pp 2459–2464
12. Pagac D, Nebot EM, Durrant-Whyte H (1998) An evidential approach to map-building for autonomous vehicles. IEEE Trans Robot Autom 14(4):623–629
13. Kim J, Kum D (2017) Collision risk assessment algorithm via lane-based probabilistic motion prediction of surrounding vehicles. IEEE Trans Intell Transp Syst 19(9):2965–2976
14. Hu C, Wang R, Yan F, Chen N (2015) Should the desired heading in path following of autonomous vehicles be the tangent direction of the desired path? IEEE Trans Intell Transp Syst 16(6):3084–3094
15. Hu C, Wang Z, Taghavifar H, Na J, Qin Y, Guo J, Wei C (2019) MME-EKF-based path-tracking control of autonomous vehicles considering input saturation. IEEE Trans Veh Technol 68(6):5246–5259
16. Wang R, Liu MS, Zhou Y, Xun YQ, Zhang WB (2016) A deep belief networks adaptive Kalman filtering algorithm. In: 7th IEEE international conference on software engineering and service science (ICSESS). IEEE, 2016, pp 178–181
17. Liu S (1998) An adaptive Kalman filter for dynamic estimation of harmonic signals. In: 8th international conference on harmonics and quality of power. Proceedings (Cat. No. 98EX227), vol 2. IEEE, 1998, pp 636–640
18. Shoval S, Zeitoun I, Lenz E (1997) Implementation of a Kalman filter in positioning for autonomous vehicles, and its sensitivity to the process parameters. Int J Adv Manuf Technol 13(10):738–746
19. Sasiadek JZ, Wang Q (1999) Sensor fusion based on fuzzy Kalman filtering for autonomous robot vehicle. In: Proceedings 1999 IEEE international conference on robotics and automation (Cat. No. 99CH36288C), vol 4. IEEE, 1999, pp 2970–2975
20. Kang-hua T, Mei-ping W, Xiao-ping H (2007) Multiple model Kalman filtering for MEMS-IMU/GPS integrated navigation. In: 2nd IEEE conference on industrial electronics and applications. IEEE, 2007, pp 2062–2066
21. Reyes S, Tokunaga OF, Espinoza ES, Salazar S, Romero H (2018) Autonomous ground vehicle navigation using a novel visual positioning system. In: 2018 22nd international conference on system theory, control and computing (ICSTCC). IEEE, 2018, pp 342–347
22. Ren H, Chen S, Liu G, Zheng K (2014) Vehicle state information estimation with the unscented Kalman filter. Adv Mech Eng 6:589397
23. Wenzel TA, Burnham KJ, Blundell MV, Williams RA (2006) Dual extended Kalman filter for vehicle state and parameter estimation. Veh Syst Dyn 44(2):153–17
24. Jetto L, Longhi S, Venturini G (1999) Development and experimental validation of an adaptive extended Kalman filter for the localization of mobile robots. IEEE Trans Robot Autom 15(2):219–229
25. Julier SJ, Durrant-Whyte HF (2003) On the role of process models in autonomous land vehicle navigation systems. IEEE Trans Robot Autom 19(1):1–14
26. Almeida HP, Júnior CL, dos Santos DS, Leles MC (2019) Autonomous navigation of a small-scale ground vehicle using low-cost IMU/GPS integration for outdoor applications. In: IEEE international systems conference (SysCon), pp 1–8. IEEE, 2019
27. Asl RM, Palm R, Wu H, Handroos H (2020) Fuzzy-based parameter optimization of adaptive unscented Kalman filter: methodology and experimental validation. IEEE Access 8:54887–5490

28. Meng L, Li L, Veres SM, Veres SM (2011) Comparison of linear and nonlinear aerodynamic parameter estimation approaches for an unmanned aerial vehicle using unscented Kalman filter. J Beijing Inst Technol 20(3):339–344
29. Antonov S, Fehn A, Kugi A (2011) Unscented Kalman filter for vehicle state estimation. Veh Syst Dyn 49(9):1497–1520
30. Qasem H, Reindl L (2007) Unscented and extended Kalman estimators for non linear indoor tracking using distance measurements. In: 4th workshop on positioning, navigation and communication, pp 177–181. IEEE, 2007

Automatic Brain Tumor Detection and Segmentation from MRI Using Fractional Sobel Filter and SOM Neural Network

Kamlesh Kumar Sharma, Janki Ballabh Sharma, and Vijay Nath

Abstract In the presented paper, the narrative automatic brain tumor detection and segmentation technique utilizing T1-weighted MR image is presented. Sobel filter is used to avoid the effects of disturbance through noise, and it will help to intensify quality of brain image. Fractional Sobel filter gives supplementary versatility in overhauling the segment results. The feature extraction organizes upon pixels power initial weights of mean, median filter and open–close morphological operation, then the extents of the component space is declined with the help of PCA. Then, the self-organizing map neural network assembles component space, after that the clustering output is utilized as the introduction of an edge detector prompting an edge map. At last, watershed transform on the edge map gives the segmentation result, and the tumor area of the segmented image is obtained by region of interest. Simulation is practiced on the pictures, taken from BRATS-2013 database, execution criteria, for example accuracy, sensitivity and specificity estimation at $\alpha = 0.3$ is calculated. The simulation output displays the exhibition of the propounded plans relates to the closest plans analyzed.

Keywords Brain tumor · Segmentation · Fractional Sobel filter · Watershed transform · Morphological operation · SOMNN

K. K. Sharma (✉) · J. B. Sharma
Department of Electronics Engineering, Rajasthan Technical University, Kota, India
e-mail: sharma.kamlesh4@gmail.com

J. B. Sharma
e-mail: jbsharma@rtu.ac.in

V. Nath
Department of Electronics and Communication Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: vijaynath@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_39

1 Introduction

Brain tumor plays significant role in medical imaging field. Tumor is a prompt differentiation development of irregular cells in human mind. Brain tumor consists of two categories which are recognized as benign and malignant tumors. MRI is the clinical picturing method, which helps in equipping the demonstrating anomalies in tissues at high goals, furthermore, more prominent differentiation pictures of the cerebrum and carcinogenic tissues. Subsequently, MRI assumes a persuasive job in finding of mind tumors [1–3].

The Brain Tumor division summons [4] center on empowering advancement of cutting-edge procedures for tumor division through giving an enormous informational index of commented on LG and HG. For every patient, there are four types of MR images. The images are as follows: T1-weighted, a T2-weighted, a post-differentiate T1-weighted, and a fluid attenuated inversion recovery MRI was provided. Because of some complications in segmentation and detection of tumors through magnetic resonance imaging, automatic brain tumor division has created a region of dynamic research [5].

Yang et al. [6] proposed a wavelet-vitality form of methodology for mind magnetic resonance image characterization. 2D DWT has been utilized by them for feature extraction. SVM classifier was utilized for brain image classification, and for streamlining the loads (weight), BBO strategy was utilized. The researchers saw that the KSVM, PSO-KSVM, and BPNN were at the lower level from the plan given by them. Selvaraj et al. [7] recommended the brain magnetic resonance image characterization framework. Numerous classifiers have been utilized by them for characterization; for example, neural classifier, SVM classifier, and statistical classifier. LS-support vector machines outflanked 98% accomplishment standard from all the classifiers which were utilized. El-Dahshan et al. [8] recommended strategy which includes three phases that are feature extraction, feature reduction, and classification. The feature extraction is reviewed for the estimation sub-band of discrete wavelet transform. The feature decrease utilizes PCA. FP-ANN and *k*-NN were used as classifier for the characterization, and they attained 97% and 98% accuracy, respectively. The texture gradient-based watershed transform has been accounted for [9] utilizing image segmentation algorithm. Pereira et al. [10] play out a customized division methodology taking into account convolution neural system, exploring minimal 3 parts. These system create discrete systems for dividing LG as well as HG glioblastomas. Four layers of convolutional, trailed by two thick and a classification network are comprised by LG network. Seven convolutional layers are made out of HG network, by utilizing 2D convolutions.

Not a bit of plan for segmentation of cerebrum tumor with the help of MR imaging, a hybrid technique utilizing fractional Sobel filter and self-organizing map, watershed transform has been accounted for. Hence, in this paper, another plan of segmenting of cerebrum tumor from MR imaging with the help of fractional Sobel filter, SOM neural network through watershed transform is proposed.

Firstly, direct and non-straight noise removes, and the feature extraction organizes upon pixels power, nearby initial weights of mean, median filter, and open–close morphological activity. Consequently, element of feature space is diminished utilizing PCA [18].

At that point, self-learning network, i.e., self-organizing map network clusters feature space. Following to this part, it contains the introduction of an edge locator prompting an edge map which is utilized by the clustering output. At last, the clustering output is introduced on the edge map by watershed transform and obtained segmented image of tumor. Fractional-order Sobel filter produce more keen inclined magnitude picture, reason being the Sobel filter improves the moderate frequency quality attributes and not so much as delicate to clamor. Simulation aftereffects of the propounded scheme are likewise contrasted with [11] and show the efficiency of the plan.

A remnant part of the paper presented is sort out as accompanied by: area II portrays the Material and Procedure, a propounded technique is portrayed in Area III, the simulation outputs are introduced in Area IV, and finally, in Area V, there is an introduction of conclusion.

2 Materials and Procedure

2.1 Morphological Activity

The morphological reconstruction is trailed by erosion and dilation individually. Erosion expels little articles in morphological opening, and the resulting dilation will in general reestablish the state of the items. Dilation also performs the same procedure as the erosion in morphological closing activity [5]. The morphological reconstruction by opening and closing of a picture I utilizing structuring component S is characterized by $R_I(I\theta S)$ and $R_I(I \oplus S)$, respectively.

2.2 Watershed Transform

Watershed transform division is a substantial unaided instrument for morphological picture division. Watershed transform is very well may be checked on as a geological area developing strategy. The geographical model of a landscape, where the pixel esteem is indicated by the height of the point is designed by the dark scale picture [12].

Since the reason of the executing watershed transform is to place picture objects into catchment bowls [13], edge identification appears, by all accounts, to be a smart thought to assist watershed with doing this activity. We engage in the precipitation reenactment to be carried out on the edge-detected image.

Picture slope closely resembles the slopes and troughs of a scene; the relationship is proceeded by envisioning precipitation streaming upon the scene, where the water pouring over the scene will drift toward the base. Consequently, the scene will have the drenching recreation where the water will overflow from drainage bowls as the height arrives at the nearby most extreme. When two floods beginning from various catchment bowls meet, a dam subsequently works to keep the bowls from joining [14].

The precipitation with a sharp drop is applied utilizing the 3 × 3 window fixated at every pixel in the angle map. The sharp slope course from the eight associated neighboring pixels is figured. The neighboring pixel with the sharp drop bearing is checked, and then the window moves along that heading. Until the way reaches the base, the procedure of checking pixels and moving window is replicated. The way which is established by the pixels embraces the name of the base. The precipitation reenactment is duplicated as it follows the way of the sharp plunge of the pixels which are non-labeled.

The ways coming toward typical least receive a similar name as that base and establish a catchment bowl, which alludes to a segment in the picture. At long last, it must be referenced that not at all like watershed-based segmentation calculations, for example, [2, 3, 8], a locale merger is not used in our methodology.

The two categories of watershed transformation:

- Distance transformation based on watershed transformation and
- Fractional gradient based on watershed transformation [9].

2.3 Fractional-Order Sobel Filter

The fractional-order derivative is utilized to locate the maximal point in gradient-based detection techniques by the virtue of the assistance of first-order derivative [3]. The image with the gradient of an image $G\,(p, q)$ on the position (p, q) based on first-order derivative is characterized by

$$\nabla G_{(p,q)} = \sqrt{G_p^2 + G_q^2} \tag{1}$$

Here, G_p and G_q represent the gradient magnitude in p and q direction, respectively.

Here, Grunwald–Letnikov (*G-L*) definition of fractional-order Sobel filter is utilized. A fractional-order Sobel filter can be obtained from the first-order Sobel filter as

$$G_p^\alpha = \frac{1}{2}\left(\frac{\partial^\alpha G(p+1, q-1)}{\partial p^\alpha} + 2\frac{\partial^\alpha G(p+1, q)}{\partial p^\alpha} + \frac{\partial^\alpha G(p+1, q+1)}{\partial p^\alpha}\right) \tag{2}$$

$$G_q^\alpha = \frac{1}{2}\left(\frac{\partial^\alpha G(p-1, q+1)}{\partial q^\alpha} + 2\frac{\partial^\alpha G(p, q+1)}{\partial q^\alpha} + \frac{\partial^\alpha G(p+1, q+1)}{\partial q^\alpha}\right) \tag{3}$$

$(-1)^k\, C_k^\alpha/2$	$(-1)^k\, C_k^\alpha$	$(-1)^k\, C_k^\alpha/2$
$\vdots$	$\vdots$	$\vdots$
$(\alpha^2-\alpha/4)$	$(\alpha^2-\alpha/2)$	$(\alpha^2-\alpha/4)$
$-\alpha/2$	$-\alpha$	$-\alpha/2$
$1/2$	1	$1/2$

(a)

$(-1)^k\, C_k^\alpha/2$	……	$(\alpha^2-\alpha/4)$	$-\alpha/2$	$1/2$
$(-1)^k\, C_k^\alpha$	……	$(\alpha^2-\alpha/2)$	$-\alpha$	1
$(-1)^k\, C_k^\alpha/2$	……	$(\alpha^2-\alpha/4)$	$-\alpha/2$	$1/2$

(b)

Fig. 1 **a**, **b** Fractional Sobel filter

Here, α is the order of fractional-order Sobel filter. With the assistance of fractional gradient derivatives, fractional-order Sobel convolution filter portrayed as follows (Fig. 1).

Here,

$$C_k^\alpha = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\,\Gamma(\alpha-k+1)} \quad (4)$$

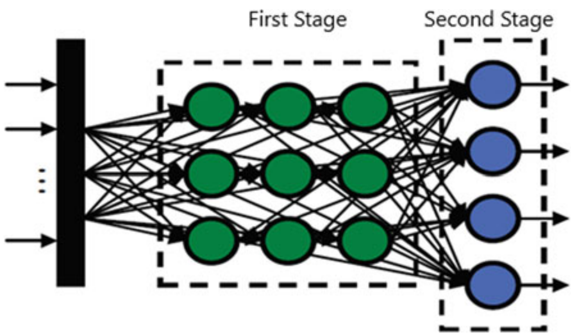
where Γ is the gamma function and $k = 0, 1, 2, 3\ldots$.

2.4 SOM

Kohonen's self-organizing map is a self-learning network emulative neural system. The surmised horizontal neural connection utilizes the local communication set, then it covers up the information to find the topological structure for seeing the showcase in a couple of dimensional space. SOM have two layers; the first one is the input layer, and second one is the output layer or the feature map. In Fig. 2 self-organizing map network topology is appeared.

Let Z be the input vector for training the SOM network, for example,

Fig. 2 The SOM network topology



$$Z = [Z_1, Z_2, \ldots, Z_i, \ldots, Z_r]^T \tag{5}$$

where Z_i is the component vector which is accompanied by $P \times Q$ components. The Euclidean distance among Z, the weight vector w_i relating to the ith neuron is figured in every iteration of the preparation procedure, for example,

$$\|Z - w_i\| = \left\{(Z_1 - w_{i1})^2 + (Z_2 - w_{i2})^2 + (Z_3 - w_{i3})^2 + \cdots + (Z_n - w_{in})^2\right\} \tag{6}$$

The winner neuron V has the weight vector W_v is the nearest to Z,

$$\|Z - w_v\| = \arg\min_i\{\|Z - w_i\|\} \tag{7}$$

The winner neuron's weights and all of its nearby N_v are refreshed by accompanying preparing rule:

$$w_i(r+1) = \begin{cases} w_i(r) + \alpha(r)[Z(r) - w_i(r)] & \text{if } i \in N_v, i = v \\ w_i(r) + \beta(r)[Z(r) - w_i(r)] & \text{if } i \in N_v, i \neq v \\ w_i(r) & \text{if } i \in N_v \end{cases}$$

where $\alpha(r)$ and $\beta(r)$ decide the preparation learning rate parameters for the winner neuron and its neighbors in the rth iteration, i.e., $\alpha(r)$ greater than zero and $\beta(r)$ less than one. The above parameters could be linear, exponential, or inversely proportional to r.

3 The Proposed Work

In the suggested segment, brain tumor segmentation plan dependent on fractional-order Sobel Filter as well as SOM utilizing watershed transformation is introduced.

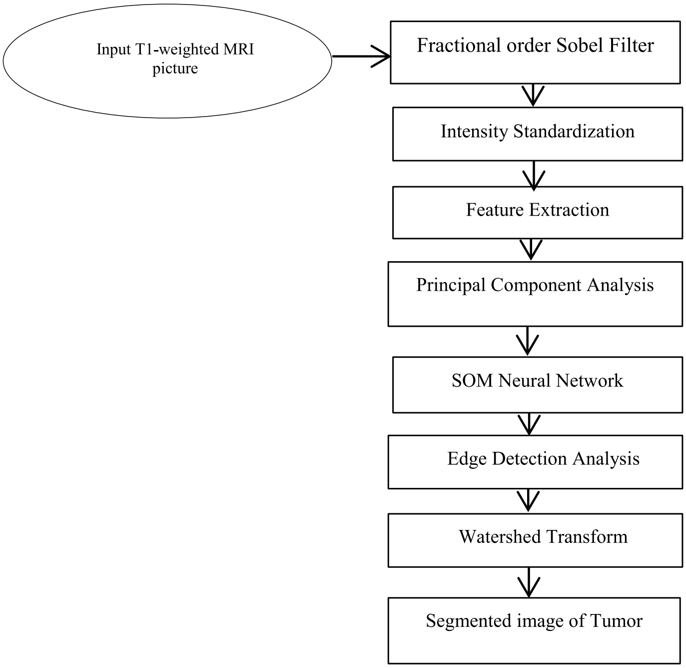


Fig. 3 Flowchart of the suggested conspire

In Fig. 3, the block diagram is appeared of the suggested scheme. Because of less calculation intricacy, 3×3 fractional Sobel Filter for $k = 2$ (shown in Fig. 1) is utilized; this α represents the fractional order.

In T1-weighted MRI mind picture tissues, for example, cerebrospinal liquid, fat, irritation (edema, localized necrosis, and sore) speak to high power (splendid), and white matter, moving blood and bones speak to less power (dull) [9].

However, the tumor is additionally having high power (intensity), in this way, most extreme attribute is utilized for distinguishing the tumor as well as help in discovering greatest powerful region for utilizing to find the tumor area. Morphological activities make maximum level inside every locale of cerebrum picture.

Intensity standardization needs quantitative surface examination. The plan, histogram standardization strategy will utilize that will in general move and stretch the separated picture histogram for incorporating entire dark level qualities which are present in picture [15].

Originally, morphological opening is put in to the $T1$-weighted magnetic resonance image with the help of recreation activity. In the morphological opening, erosion trails morphological reformation. The structuring element (SE) helps in choosing the dimensions of erosion. The dimensions of the structuring component choose the experimentation for accomplishing upgraded outcomes. In the presented activity, the finer outcomes are obtained by utilization of disk-shaped structuring element of size 1. The restrictive dilation activity helps in the performance of the morphological recreation until there is no requirement of the alteration in the reconstruction of the picture. The restrictive dilation has a combining feature of dilation with genuine filter picture by utilizing reasoning operation [5]. Furthermore, morphological closing with the help of remaking is done on the subsequent image with similar dimensions of structuring component. The above activities are utilized to evacuate splendid and dull articles that are littler than the structuring component; they likewise expel the impulse noise.

Now, we will concentrate our technique of extracting features from magnetic resonance imaging. It consists the following-initial weights of mean, median filter, and open–close morphological activity when distinctive neighborhood dimensions and extraction of feature is prominent on picture element intensity. Thereafter, the feature space size is decreased utilizing PCA [18].

At that point, the SOM system clusters the feature space. In the wake of completing the preparation step, prepared SOM system could be utilized into the test stage for clustering test pictures. SOM characterized in two phases, where the nine neurons are comprised in the principal stage, and the subsequent stage comprises four neurons. The principal stage carries out feature mapping to a lower measurement, while subsequent one decides the yield cluster. Following to this magnetic resonance imaging, includes the bone, delicate tissue, fat, and foundation, the yield stage utilizes four neurons for every class [18]. The cluster is determined from the winner neuron in the second SOM. After clustering, iterative median filtering is used for slightly improving the outcome. In our work, topology is utilized hex top and for compute good ways from a home neuron to its neighbors are utilized dist methodology. In the following portion, the section of an edge identifier utilizes the clustering outcome for prompting an edge map.

At long last, watershed transform on the edge map delivers the clustering outcome. Edge recognition is utilized on the self-organizing map clustering outcome.

The segmentation of tumor locale is carried out by putting in marker form watershed transformation with respect to altered edge detection picture. Altered edge picture is contributed to the watershed transformation. A topographical plane observes a picture in slope-based watershed transform within which gradient magnitude indicates peak power edges, and markers are the short power esteem.

The segmented marked picture is the result of the watershed transformation and the tumor area of the segmented image is obtained by region of interest.

The summed up stages of the proposed conspire are:-

Stage 1: First, T1-weighted MRI Brain picture is drawn for filtration by fractional Sobel filtration.
Stage 2: Filtered image is intensity standardization and finding the fringe.
Stage 3: Now, features are extracted from the MRI picture (mean, median, and morphological features extracted).
Stage 4: Dimension reduction using PCA.
Stage 5: The feature is clustered by SOMNN.
Stage 6: Edge identification is then utilized on the SOMNN result.

Stage 7: Watershed transform is utilized on the edge identification result.
Stage 8: Finally, segmented tumor image obtained.

4 Simulation Results

MATLAB toolbox is utilized for implementation of the suggested plans for simulation results. T1-weighted brain magnetic resonance imaging of brain patients 006, 009, 015, and 024 with HGG is taken from BRATS-2013 database. The suggested conspire is contrasted with [11, 16, 17], and for estimation of α at 0.3, the execution parameters like accuracy, sensitivity, and specificity are assessed. Ground truth recommended image are likewise extracted by similar informational index.

$$\text{Sensitivity} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \times 100\% \tag{9}$$

$$\text{Specificity} = \frac{(\text{TN})}{(\text{FP} + \text{TN})} \times 100\% \tag{10}$$

$$\text{Accuracy} = \frac{(\text{TN} + \text{TP})}{(\text{FP} + \text{TP} + \text{FN} + \text{TN})} \times 100\% \tag{11}$$

where true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP) are known as truly characterized positive pixels, truly arranged negative pixels, falsely arranged positive pixels, and erroneously grouped negative pixels, respectively.

Prime T1-weighted MRI brain images for patient 006, 024, 009, 015 are appeared in Figs. 4a, 5a, 6a, and 7a, respectively. Fractional Sobel filter of size 3 * 3 and fractional orders $\alpha = 0.3$ are applied on the prime image appeared in Figs. 4b and 5b. Intensity standardization images are appeared in Figs. 4c and 5c. Figures 4d and 5d show searching the fringe of the brain image, which is acquired by erosion and subtraction. Figures 4e and 5e show the 9*9 SOM feature map for the primary SOM systems. Feature space is clustered by two-phase SOM network, i.e., 300 epochs for training the first SOM and 200 epochs for the second one have been considered; the SOM clustering results are appeared in Figs. 4f and 5f. Figures 4g and 5g show the edge identification image. Figures 4h and 5h show the segmented image after watershed transform and Figs. 4i, 5i, and 6b and 7b show the segmented color image after watershed transform. Figures 4j, 5j, 6c, and 7c show the segmented image of tumor, and Figs. 4k, 5k, 6d, and 7d show the ground truth image.

Table 1 shows the exhibition analysis on BRATS database with various existing techniques and proposed strategy at $\alpha = 0.3$. The proposed strategy accomplishes preferable exhibition over the current technique. In this paper, we take randomly four images from BRATS 2013 dataset, comparing the parameter of accuracy, sensitivity, and specificity with [11, 16, 17] existing method (Fig. 8).

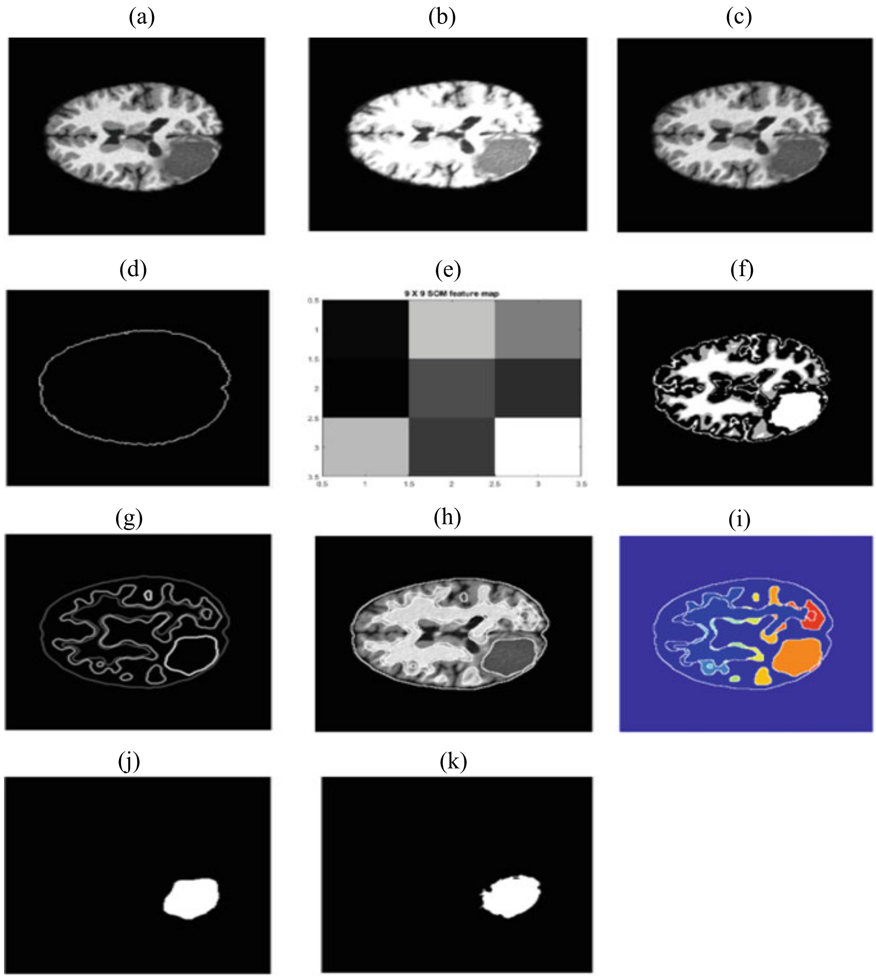


Fig. 4 Patient 006 T1-weighted MRI image, **a** prime MRI image, **b** fractional Sobel filter image, **c** intensity standardization image, **d** searching the fringe (border), **e** 9 × 9 SOM feature map, **f** SOM clustering result, **g** edge identification image, **h** segmented image after watershed transform, **i** segmented color image after watershed transform, **j** segmented image of tumor, and **k** ground truth image

5 Conclusion

The presented paper introduce a fractional Sobel filter as well as SOMNN with watershed transformation which is build up on a hybrid tumor segmentation scheme complex T1-weighted brain magnetic resonance imaging. The simulation results exhibit the efficiency in the suggested plan. The fractional-order α in fractional Sobel filter give extra level of opportunity in enhancements of results; for example,

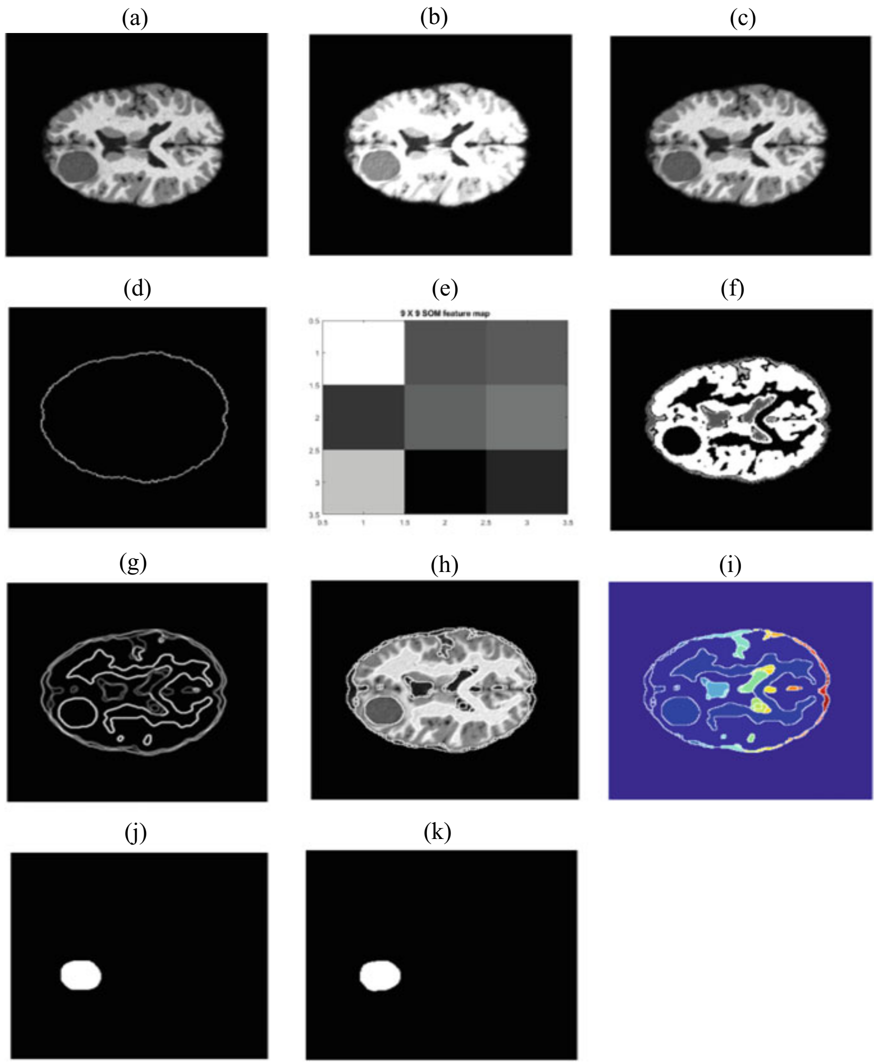


Fig. 5 Patient 024 T1-weighted MRI image, **a** prime MRI image, **b** fractional Sobel filter image, **c** intensity standardization image, **d** searching the fringe (border), **e** 9 × 9 SOM feature map, **f** SOM clustering result, **g** edge identification image, **h** segmented image after watershed transform, **i** segmented color image after watershed transform, **j** segmented image of tumor, and **k** ground truth image

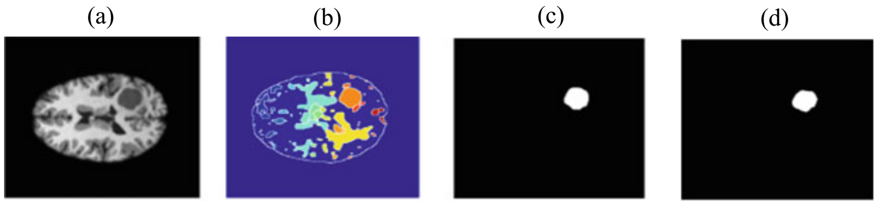


Fig. 6 Patient 009 T1-weighted MRI image, **a** prime MRI image, **b** segmented color image after watershed transform, **c** segmented image of tumor, and **d** ground truth image

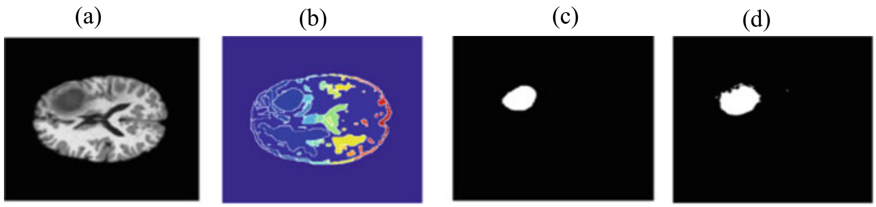


Fig. 7 Patient 015 T1-weighted MRI image, **a** prime MRI image, **b** segmented color image after watershed transform, **c** segmented image of tumor, ground truth image

Table 1 Contrasting of execution parameters at $\alpha = 0.3$ with [11, 16, 17]

Patient	Accuracy	Sensitivity	Specificity
006	99.64	94.34	99.82
009	99.80	95.96	99.85
015	99.21	70.37	99.99
024	99.79	97.89	99.82
[16]	99.19	69.91	99.92
[17]	98.48	83.84	99.34
[11]	96 ± 0.7	94 ± 0.3	95 ± 0.3

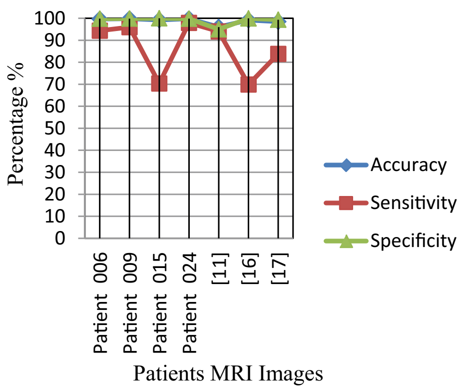


Fig. 8 Scatter diagram of comparison of performance parameters of the BRATS 2013 dataset

fractional Sobel filter gives various decisions to pick the request to get the almost exact aftereffect of tumor segment. Suggested plan could likewise be utilized for segmenting another kind of tumors (malignancy) and computed tomography pictures. Various performance parameters feature space is set up, in view of neighborhood initial weights of mean, median filter, and open–close morphological activity. Consequently, size and shape of the feature space is decreased with the help of principal component analysis. Self-organizing map clustering is utilized on to the feature space. The clustering result determines the edge map. At long last, watershed transform on the edge map delivers the segmentation output, and the tumor area of the segmented image is obtained by region of interest.

References

1. Zulkoffli Z, Shariff TA (2019) Detection of brain tumor and extraction of features in mri images using K-means clustering and morphological operations. In: 2019 IEEE international conference on automatic control and intelligent systems (I2CACIS)
2. Bhanothu Y, Kamalakannan A, Rajamanickam G (2020) Detection and classification of brain tumor in MRI images using deep convolutional network. In: 6th international conference on advanced computing and communication systems (ICACCS)
3. Tian D, Wu J, Yang Y (2014) A fractional-order edge detection operator for medical image structure feature extraction. In: Control and decision conference, pp 5173–5176, May2014
4. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R et al (2015) The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34(10):1993–2024
5. Benson CC, Lajish VL, Rajamani K (2015) Brain tumor extraction from MRI brain images using marker based watershed algorithm. In: International conference on advances in computing, communications and informatics, pp 318–323, Aug2015
6. Yang G, Zhang Y, Yang J, Ji G, Dong Z, Wang S, Feng C, Wang Q (2015) Automated classification of brain images using wavelet-energy and biogeography-based optimization. Multimedia Tools Appl 1–17
7. Selvaraj H, Selvi ST, Selvathi D, Gewali L (2007) Brain mri slices classification using least squares support vector machine. Int J Intell Comput Med Sci Image Process 1(1):21–33
8. El-Dahshan E-SA, Mohsen HM, Revett K, Salem A-BM (2014) Computer- aided diagnosis of human brain tumor through mri: a survey and a new algorithm. Expert Syst Appl 41(11):5526–5545
9. Hill PR, Canagarajah CN, Bull DR (2003) Image segmentation using a texture gradient based watershed transform. IEEE Trans Image Process 12(12):1618–1633
10. Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging 35(5):1240–1251
11. Nabizadeh N, Kubat M (2015) Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. Comput Electr Eng 45:286–30
12. Tzortzis G, Likas A (2008) The global kernel K-means clustering algorithm. In: IEEE international joint conference on neural networks. IJCNN, pp 1977–1984
13. Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. Pattern Anal Mach Intell 13(6):583–598
14. Kuo W, Lin C, Sun Y (2008) Brain MR images segmentation using statistical ratio: mapping between watershed and competitive Hopfield clustering network algorithms. Comput Methods Programs Biomed 91(3):191–198
15. Cortes C, Vapnik V (1995) Support vector networks. Mach Learn 20:273–297

16. Padlia M, Sharma J (2019) Fractional Sobel filter based brain tumor detection and segmentation using statistical features and SVM. In: Nanoelectronics, circuits and communication systems, pp 161–175
17. Padlia M, Sharma J (2017) Brain tumor segmentation from MRI using fractional sobel mask and watershed transform. In: International conference on information, communication, instrumentation and control (ICICIC), Aug 2017
18. Emambakhsh M, Sedaaghi MH (2009) Automatic MRI brain segmentation using local features,self-Organizing Maps, and watershed. In: 2009 IEEE international conference on signal and image processing applications, Kuala Lumpur, 2009,pp 123–128

Image Sensor—CCD and CMOS

R. Tejas, Pavan Macherla, and N. Shylashree

Abstract In recent times, the image sensor has become a crucial part of smartphones and have always been critical part of professional cameras, telescopes used in the astronomical observatory and space telescopes. The development of the images sensor in the present day is much particular to the application which it is used for. The image sensor are of two types CCD image sensor and CMOS image sensor. The CMOS image sensor is favoured over the CCD sensor due to the increased speed of processing and comparatively, good pixel size the CMOS sensor is having a fill factor of 30–60% with HDR capabilities and 0.37–0.4 μm technology. The below paper gives the types of an image sensor, the individual architecture of the image sensor and the applications.

Keywords Charge-coupled devices (CCD) · Complementary metal-oxide semiconductors (CMOS) · Active pixel sensor (APS) · Dynamic range · Power consumption · Metal-oxide semiconductor capacitors (MOSCAP) · Pixel · Photodiode

1 Introduction

This section gives the introduction to the image sensor and explains the types of image sensor by giving the history of the image sensors along with the pros and cons of the image sensor.

R. Tejas · P. Macherla · N. Shylashree (✉)
Department of Electronics and Communication Engineering, RV College of Engineering, Bengaluru 560059, India
e-mail: shylashreen@rvce.edu.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_40

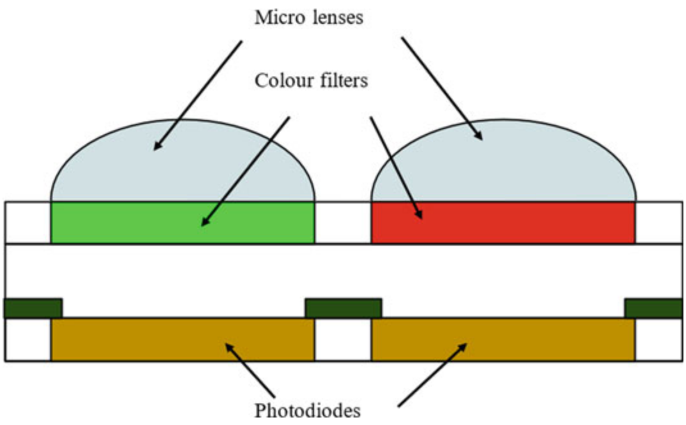


Fig. 1 Basic components of image sensor

1.1 Introduction of Image Sensors

An image sensor is an electronic device that captures the light falling on its pixel layer and conveys them for processing ultimately to make an image. It captures the different amounts of lights falling on its pixels and converts them into equivalent current values and processing them to produce an image as seen through the lens. Basic structure of image sensor consists of microlenses, colour filters and photodiodes as shown in Fig. 1.

- **Microlenses**: Lenses are placed accordingly to concentrate and focus the light in a proper way.
- **Colour filters**: These are used to filter out the RGB lights.
- **Photodiode**: These are the devices which capture the light for further processing.

1.2 Types of Image Sensor

There are two types of image sensors—charge-coupled device (CCD) and complementary metal-oxide semiconductor (CMOS). Both of them use metal-oxide semiconductors, but CMOS is based on MOSFET amplifiers whereas CCD is based on MOS capacitors and both of them use different processing methods.

CCD sensor: Photons are captured in each pixel. After the exposure time, the charge moves to the amplifier located at the end of the sensor. Externally analogue signals are converted to digital form and stored in memory.

CMOS sensor: Every pixel has its circuitry. Each pixel can modify the data right in the sensor. Every pixel can be accessed individually.

Table 1 Characteristic comparison of CMOS and CCD sensor

Characteristics	CCD	CMOS
Fill factor	High	Low
Dynamic range	High	Moderate
System complexity	High	Low
Speed	Moderate	High
Power consumption	Moderate to high	Low
Clocking	More voltage	Less voltage

1.3 Pros and Cons of Image Sensor

CCD image sensors have high sensitivity and lower noise due to the enhanced surface use of the pixels [1]. As CCD is based on MOSCAP it has simpler structure and thus fewer defective pixels. The centralized A/D converter gives better image homogeneity. CCD sensors are used for their better image quality and can reduce the SNR in the sensor. Although even with these advantages the CCD has disadvantages, the centralized ADC causes the readout to slow down and the readout method does not give direct access to the pixels. The camera is much more expensive due to the higher cost of electronics and supporting equipments. Also, the energy consumption of the CCD image sensor is much more as the working is based on the charge accumulation. Due to overexposure of the CCD image sensor while readout can cause smearing and blooming effect which is to be overcome using post processing methods.

CMOS image sensors consume very less power. They work on low voltage value, and need 1/4th of the supply required for CCD [2]. CMOS image sensors support on-chip functionalities. Many fabrication techniques make it possible to have supporting circuits for image processing on a single chip, this makes the processing of the images very fast [3]. CMOS image sensor allows different ways of selection of the pixels which makes the processing of the image easy. These image sensors avoid blooming (charge leaking of a pixel to neighbour pixels) and smearing effects which are the major disadvantages of the CCD sensor [3]. Although it has so many advantages, CMOS image sensors have disadvantages like noise, dynamic range, etc. CMOS image sensors have lower fill factor, thus have less efficiency. These image sensors produce a lower quality image compared to the CCD image sensors. It is always a trade-off between quality of image and image processing time. Table 1 shows the characteristics in a comparative way.

1.4 History of Image Sensors

Earlier, films were used for photography. In the late 1880s, Kodak produced the first film and made photography available to the people. Since then many new methods for capturing pictures and manufacturing cameras came into existence.

1.4.1 History of CCD Image Sensor

The sensor was invented in the year 1969 by Willard Boyle and George E. Smith at the AT&T Bell Labs. The charge-coupled device is a planar collection of capacitors that are etched into a silicon surface which forms an integrated circuit with light-sensitive elements [4]. This forms a pixel that is used to acquire the light source. In the year 1980, Nobukazu Teranishi, working at the NEC institution of Japan who also created the pinned photodiode. This invention made it possible to increase the resolution of the image and improve its quality by improving the signal-to-noise ratio. Even though the CCD was invented for a different purpose for the use in computer memory. Before the year of 1970 the relays, delay lines, vacuum tubes, ferrite cores, tapes and magnetic drums were used in the computing circuits for storage of data. Until the 1970s, the RAM-based on the use of semiconductors for the circuits development was not a common practice. Before the invention of the CCD's, the Bell Labs wanted to develop for the purpose of memory circuits [5] as per the purpose in the year 1966 Bell Labs engineer Andrew H. Bobeck invented the magnetic bubble memory. Shortly after that George Smith, Willard Boyle had the purpose of developing the kind of bubble memory which is much cheaper and easier to manufacture also more easily integrated into circuits than the magnetic bubble memory. All these led to the development of the charged-coupled device and that can also be used in the image sensor.

1.4.2 History of CMOS Image Sensor

First, CMOS imager came into existence in late 60s, but due to large pixel size and poor performance, it could not compete with the CCDs which had lower fixed pattern noise (FPN) and high dynamic range [6]. Research was carried on and because of advantages like high-speed imaging and avoid blooming effects, they emerged back in the 90s. Passive pixels were first introduced and with improvements to reduce noise using amplifiers, a second generation of pixels, called active pixel sensors were introduced [7]. These pixels have high readout, lower noise levels compared to passive pixels.

As APS has shown better performance, research has been done, mainly in this area. Several improvements like larger fill factor and high dynamic range have been done. New applications have emerged and accordingly different architectures of APS have been proposed to meet the requirements.

2 Charge-Coupled Device (CCD) Image Sensor

Charge-coupled devices use the main pixel elements as the MOSCAP, and light is stored as the charge in the capacitors. The below section gives different types of architecture of CCD, pixel of CCD along with the device enhancements and channel types.

2.1 Architecture

CCD sensor has an architecture made of semiconductor where the charge transfers happen through the storage areas. The CCD architecture is used for the transfer of charge vie storage area by many sensors which lie in the visible region. The three basic operations that the CCD architecture sensors follow is—charge collection that happens when exposed to light source, charge transfer that happens within the sensor, at last converting the acquired charge into a voltage that can be measured for interpretations in digital form.

The charges which are created in a single pixel of the CCD array is proportional to the incident light source levels which it is exposed to. The MOS capacitors are the technology that are used in the CCD sensor [8].

The photoelectric effect is the cause for generating the charges as shown in Fig. 2 at the MOS gate also known as the "photogate". The charges are generated by the photodiodes for the systems such as interline-transfer structure. Once the charge is generated it is transferred to the conversion node that occur within the MOSCAP for all systems; i.e. conversion of charge to voltages occurs here. The CCD has different types of architecture those include—full frame CCD sensor, frame transfer sensor and interline-transfer CCD [8]. The general frame structure of CCD is as shown in Fig. 3 with a centralized amplifier.

Full frame CCD—it is a form of sensor where the incoming light source comprising the photos fall onto the fully light-sensitive array of the sensor [11]. The accumulated charges are vertically shifted for readout and this is done row by row into the serial output register, and then horizontal shifting is done; this is known as "progressive scan". The structure is as shown in Fig. 4 for full frame CCD.

Frame transfer CCD—this is a two-part type of architecture sensor where one of the halves is the array of pixels (as shown in Fig. 5) used for storage of charges, and this is the region which is protected from light using the reflective materials like

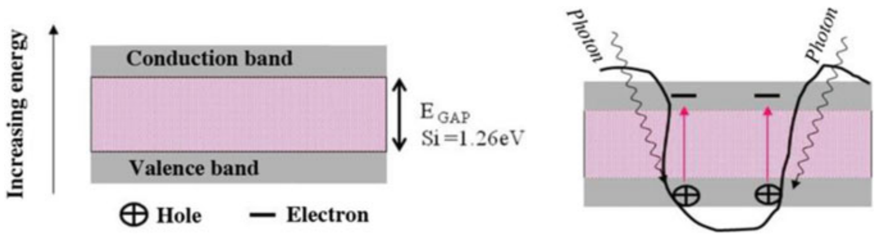


Fig. 2 Energy bands in silicon CCD pixel [9]

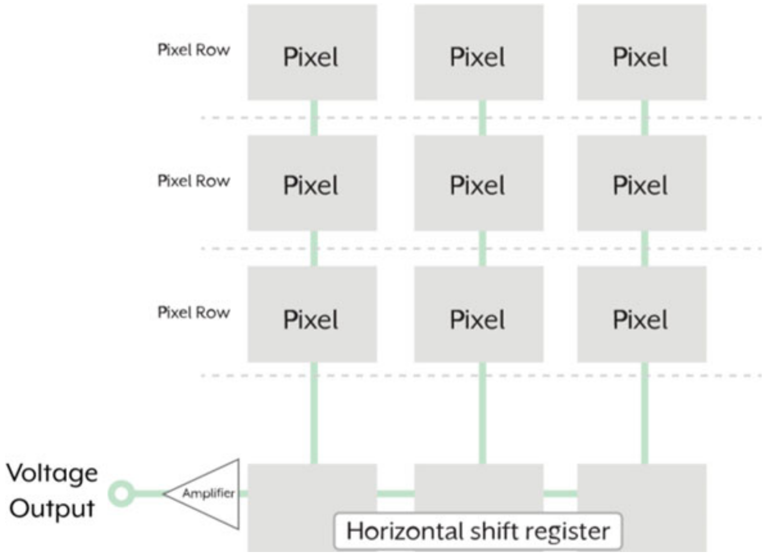
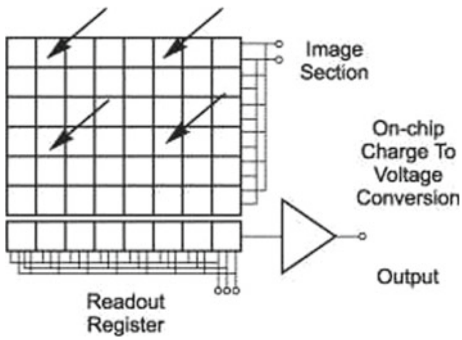


Fig. 3 General frame structure of CCD [10]

Fig. 4 Full frame structure for CCD [12]



aluminium [13]. The part of the sensor which is not covered by the reflective surface is the part which is exposed for charge accumulation, and then it is rapidly shifted into the masked storage area. Simultaneously when the charge is being shifted, the part of the CCD is still exposed for charge accumulation, and this is done so that there is no delay in successive exposures.

Interline architecture—this architecture extends the concept of the frame transfer architecture, and it covers every alternative column of the image sensor which acts as the storage for charge that are shifted from the exposed part of CCD as shown in Fig. 6 as in this architecture, there is only shifting of one pixel for storage so it is faster the frame transfer [13].

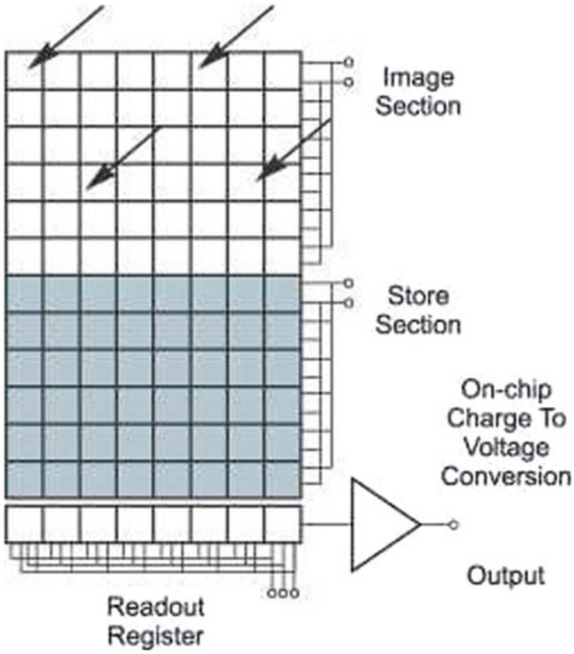


Fig. 5 Frame transfer CCD [13]

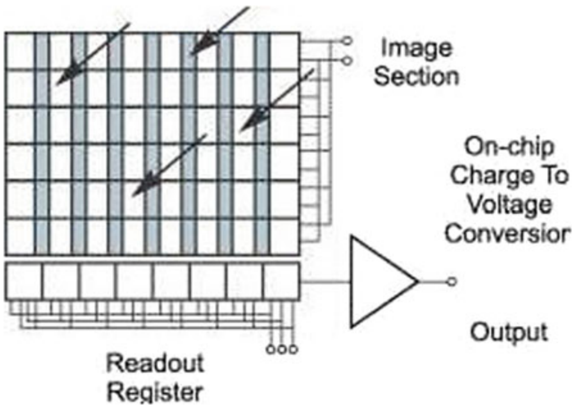


Fig. 6 Interline architecture of CCD [13]

2.2 Pixel of CCD

The CCD sensor technology is based on the MOS capacitors as shown in Fig. 7. The electron–hole pair is created when a photon energy exceeds a energy gap which is absorbed in the depletion zone. Earth electrode attracts the hole towards it while the electron remains in the depletion zone [14]. The voltage applied to the sensor, oxide thickness and the surface of the gate electrode is proportional to the amount

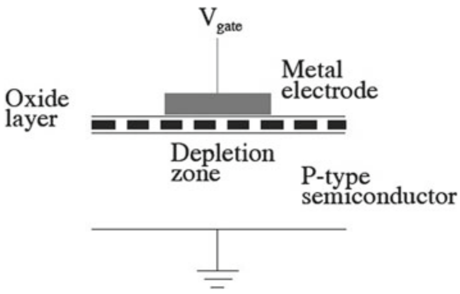


Fig. 7 MOS gate of silicon [14]

of negative charge. The photons are absorbed at the increasing depths when the wavelength increases.

Figure 8 shows the working of the CCD charge transfer between gates (a) shows the overlap of depletion zones. The charge transfer can occur only when the gradients of depletion zones overlap. The photo electrons are collected at $Well_1$ when the voltage is applied at $Gate_1$ (b). The electrons move to $Well_2$ similar to a waterfall as soon as the voltage is applied to $Gate_2$ (c). The charge balances rapidly between the two wells as the process is fast (d). The electron flow cascades into $Well_2$ as well as the potential well decrease due to the reduction of voltage at $Gate_1$ (e). At last all electrons are in $Well_2$ when the voltage applied at $Gate_2$ approaches nil (f).

The charge transferred across all shift registers by repeating the above process several times until the completion of shifting. The voltage levels of the gate determine the behaviour when the voltage is low, it acts as barrier and when it is high, charge can be stored.

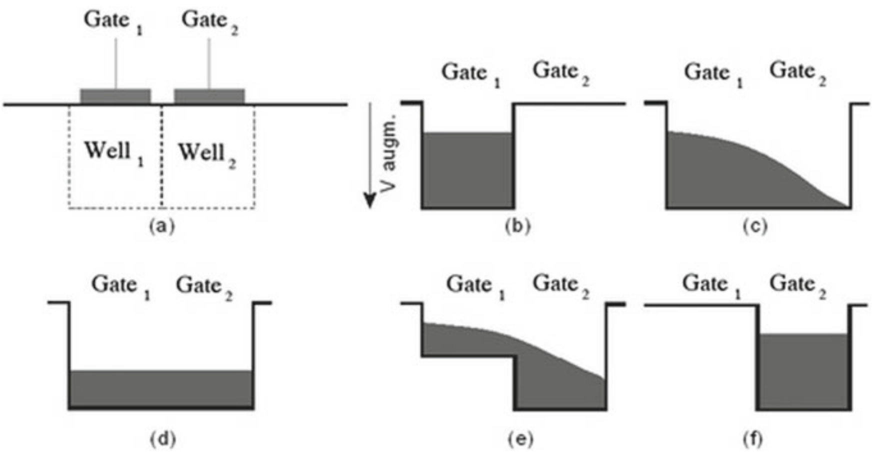


Fig. 8 Charge transfer between gates [14]

2.3 Device Operation and Enhancements

CCDs operate in a state referred to as the deep depletion since, after initial biasing, the holes are forced to move farther towards the substrate while no free electrons are at or close to the silicon surface [15]. Operation in the deep depletion region is necessary since it might take up to one hour in some devices (such as CCDs in high-end scientific cameras) to reach the thermal equilibrium required for strong inversion [16]. A comparison between an ideal and non-ideal NMOS photogates is shown in Fig. 9, whereas the effect of the gap between gate electrodes on the charge transfer process is illustrated in Fig. 10.

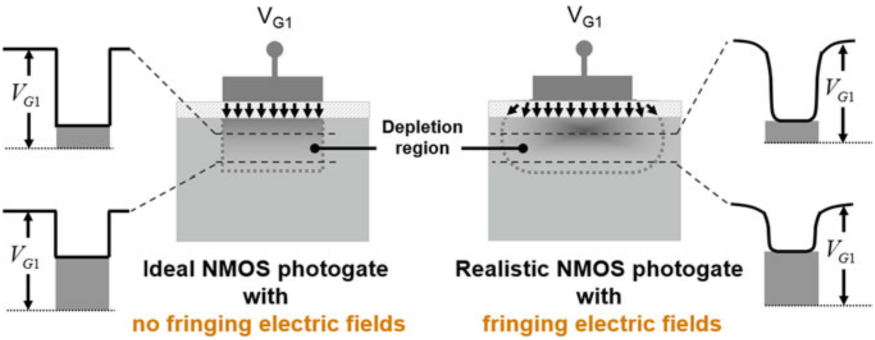


Fig. 9 Electric fields in NMOS photogates [17]

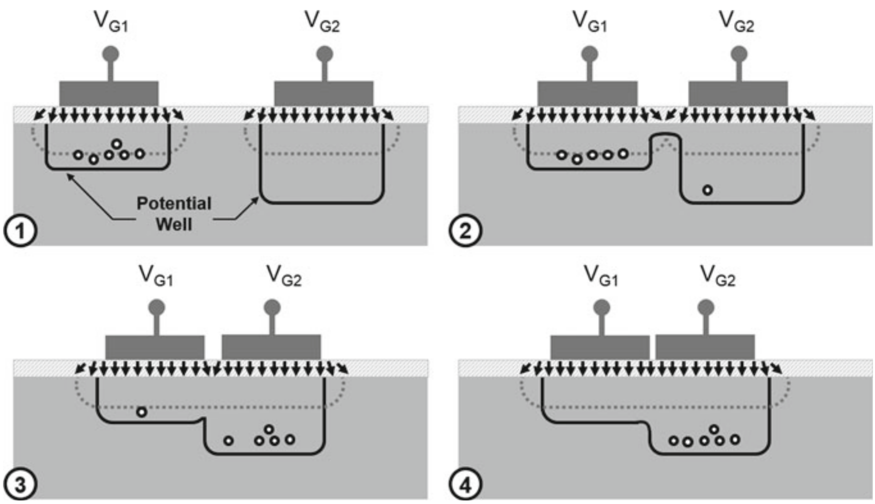


Fig. 10 Effect of shortening the gap between two adjacent gate electrodes [17]

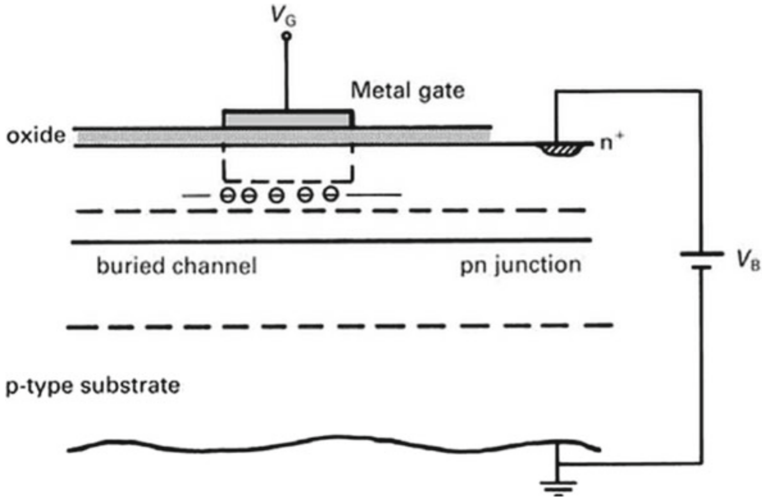


Fig. 11 Cross-sectional view of a storage site of buried channel CCD [18]

Buried channel CCD—the surface channel/thin oxide interface layer usually suffers from crystal defects that may trap charges, resulting in charge losses and image smearing effects when charge packets are transferred along with this interface. In buried channel CCDs, a layer of n-doped silicon over the p-doped silicon layer with a voltage bias applied between them, as illustrated in Fig. 11, the charge storage region is buried within the depletion region.

Backside-illuminated CCD—illuminating the CCD from its backside will allow light to strike the photocathode layer without passing through the electrodes and wiring layer therefore avoiding losses. Backside illumination involves thinning the CCD, either by dry or by wet etching, to about 15 microns. In a back-illuminated CCD (Fig. 12), the wiring is oriented behind the photocathode layer by performing backside etching of the silicon wafer during the fabrication process so that incident light can strike the photocathode layer without passing across the electrodes and wiring layer.

2.4 CCD Processing Unit

Due to the way the CCD image sensor stores the image, while the sensor or the object is in motion, it can have smearing and blooming effect caused due to the charge accumulation delay in capacitors [20]. The processing chip is used to compensate for the effects of the delay for getting better images.

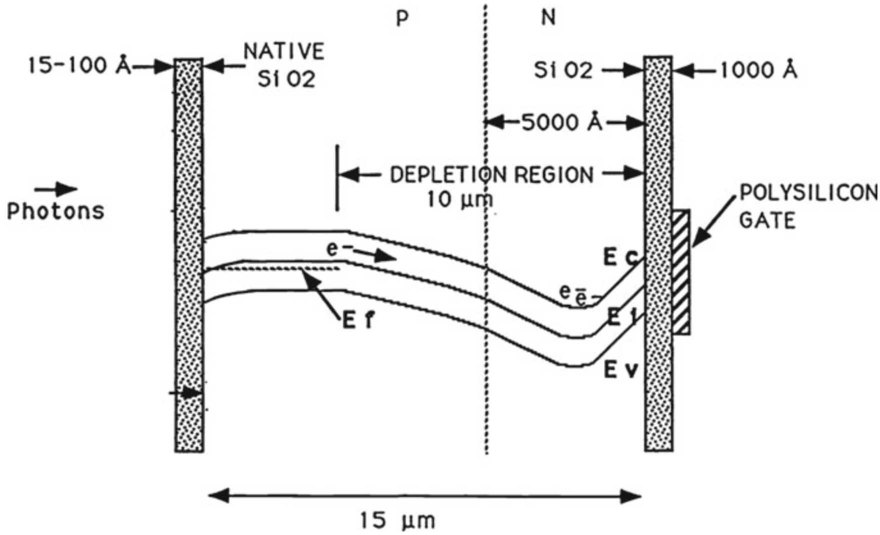


Fig. 12 Cross-sectional view of a backside-illuminated thinned CCD [19]

3 Complementary Metal-Oxide Semiconductors (CMOS) Image Sensor

CMOS image sensor uses MOS gates transistors circuits to store the light from the source/object in its pixels. This section demonstrates the working of CMOS images sensor using a generalized architecture, pixel circuits, types of pixels and another supporting circuitry for the CMOS image sensor.

There are many technologies which are used in the manufacturing of CMOS image sensor. Newer CMOS image sensor uses backside illumination (BSI) which is an alternative to frontside illumination (FSI) as shown in Fig. 13 [21]. In this method, the sensor is turned upside down and microlenses are attached at the backside of the

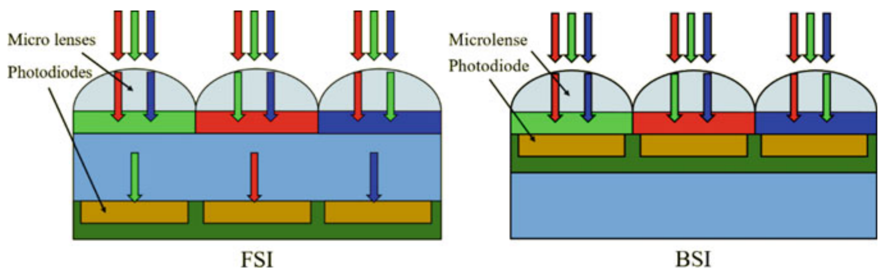


Fig. 13 Frontside illumination versus backside illumination

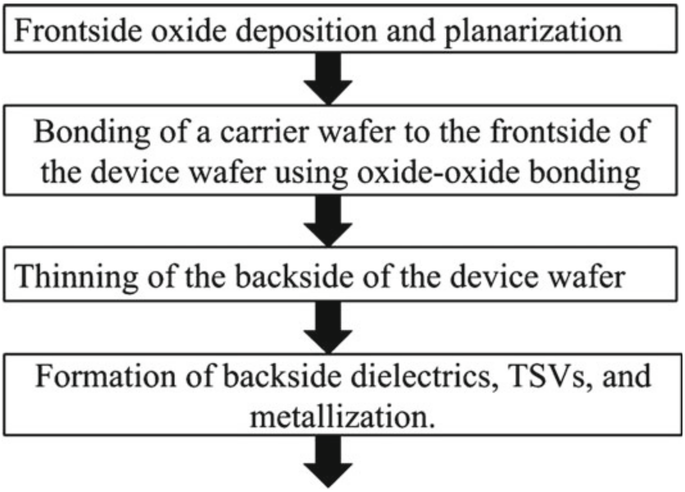


Fig. 14 Flow diagram of manufacturing BSI image sensor [21]

structure. Sensors manufactured using BSI technology offer higher quantum efficiency and less crosstalk. These sensors are especially used in the low light applications as the backside illumination sensors can capture more light. The general flow of the backside illumination CMOS image sensor in shown in Fig. 14 [21].

3.1 Architecture

When the sensor is exposed to the light source or subject, pixels capture the light. The accumulated photons in pixels are read in a manner similar to that of reading data present at a point in 2D coordinate system. Each pixel is represented as a (x, y) coordinate. A row selector is used to pass "y" value, and a column selector is used to pass "x" value. CMOS fabrication techniques allow the image sensor to be a single-integrated chip. Many supporting circuits, like control circuit for pixel (charge reset, signal readout and charge integration), memory components, bias generators, signal processing unit, etc. can be integrated [22]. Recent chips come with on-chip A/D converters and noise suppression modules which helps in realizing a full digital interface and it also gives low noise performance. Architecture of the sensor depicting row selector, column selector and pixel array is shown in Fig. 15.

There are a variety of signal processing architectures possible for CMOS image sensors [22]. Three types of on-chip architectures are shown in Fig. 16 [22, 23].

Figure 16a shows an architecture where signal processing is done at the chip level. This signal processing architecture is called serial type architecture. The required bandwidth for processing is the product of the number of vertical pixels, number

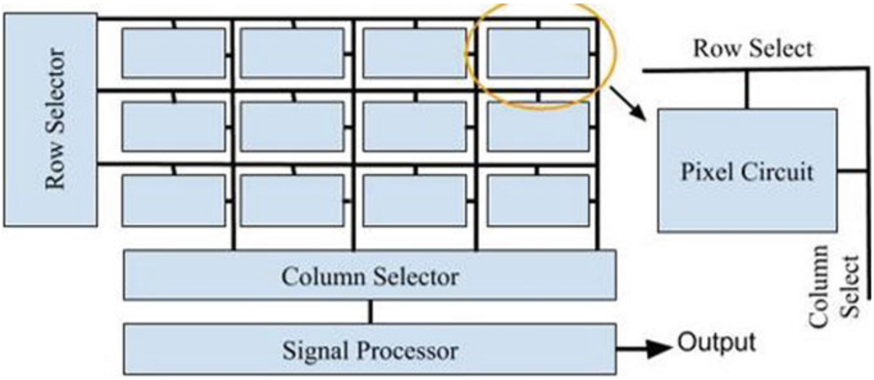


Fig. 15 CMOS image sensor

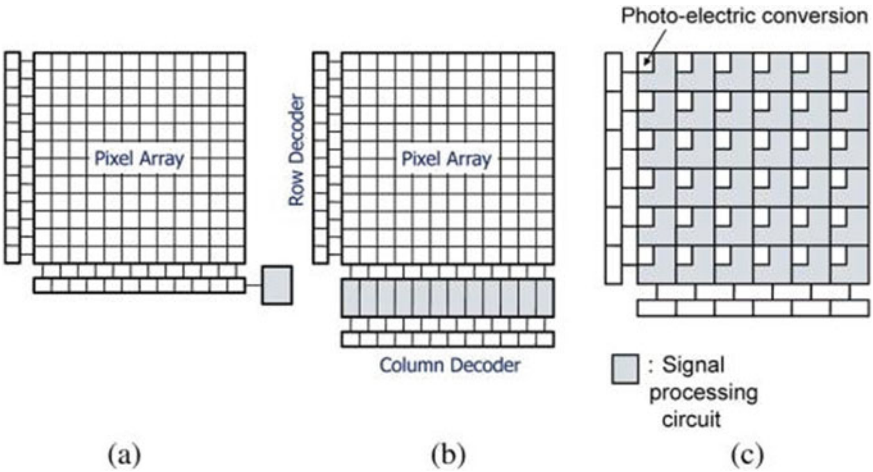


Fig. 16 Three signal processing architectures, **a** serial, **b** column parallel and **c** pixel parallel processing architecture [22]

of horizontal pixels and the frame rate. As pixels do not have the processing circuit inside them, these types of sensors are desirable for small pixel CMOS image sensors.

Another type of architecture, known as column parallel architecture as (shown in Fig. 16b) performs signal processing based on the rows in the pixel array. One row at a time is processed. The required bandwidth for processing depends only on the frame rate and number of vertical pixels. Compared to the serial architecture, the processing bandwidth is reduced by a factor N, where N represents the number of horizontal pixels. For medium size sensors, the processing circuit is arranged as a single-pixel pitch, but for small CMOS image sensors, it is arranged in multiple pixel pitches.

The last type of architecture is called pixel parallel architecture as shown in Fig. 16c. In this type of architecture, signal processing is done inside the pixel itself all over the pixel array. There is a trade-off in these types of CMOS image sensors. It offers high processing throughput, but as the pixel size becomes large, this architecture may not be useful for the high-resolution CMOS imagers.

3.2 Pixel

A simple pixel is made up of a photodiode and a transistor for switching. A pixel can have more transistors based on the design of the pixel, the application of the sensor and many other factors. The pixels for CMOS image sensor can be categorized into two types: active pixel sensor (APS) and passive pixel sensor (PPS). Conceptual diagram is shown in Fig. 17 [22] (Fig. 18).

Earlier CCD and CMOS image sensor had "Passive pixel" as they did not have internal amplification in contrast of "Active pixel" used in the modern CMOS image sensor [25]. Active pixel sensor uses source follower which increases the performance of pixels. Active pixel sensor gives higher SNR [26]. A fixed pattern noise (FPN) suppression circuit is used to negate the noise generated due to varying threshold voltages. The working of a passive pixel [24, 27]:

1. The reverse diode photodiode is set to a high voltage, before exposing the pixel to light.
2. During the time of exposure, the incoming photons decrease the voltage of the photodiode.
3. The voltage across the diode is measured at the end of the exposure time, the difference between this value and original value gives the number of photos falling during exposure time on the diode.
4. The photodiode has to be reset for a new exposure cycle.

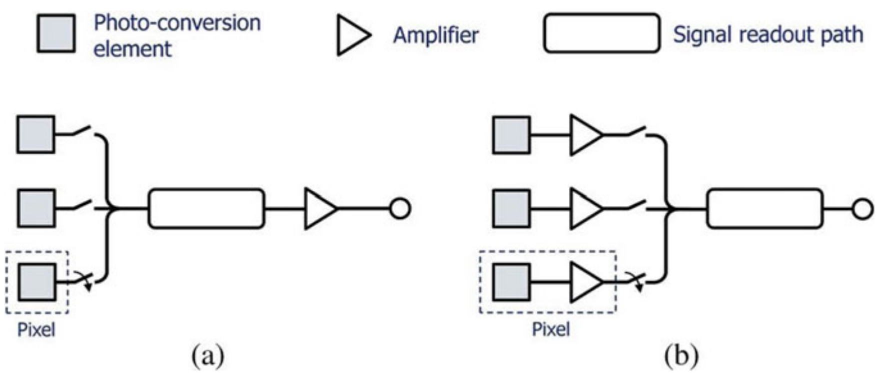


Fig. 17 Conceptual diagram of **a** PPS and **b** APS [22]

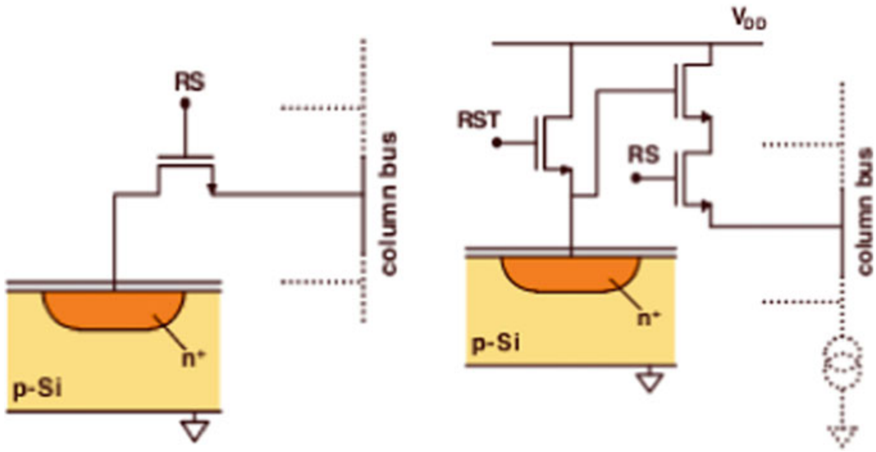


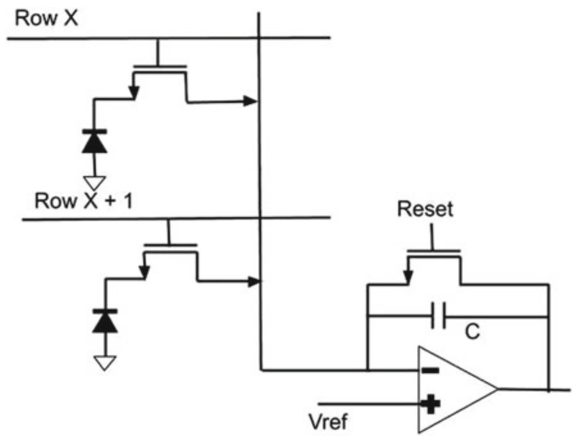
Fig. 18 Basic passive pixel and active pixel structure [24]

Fill factor of a pixel is the ratio of diode area to the total pixel area. Passive pixels have a large fill factor but because of mismatch in the large bus capacitance and small pixel capacitance, passive pixels suffer with a large noise level.

To overcome this capacitance mismatch problem, modern passive pixel sensors use a charge integration amplifier (CIA) at the end of each column bus along with one access transistor as shown in Fig. 19 [28, 29]. This version of PPS still suffers from FPN and speed issues, but it gives the highest fill factor for a given pixel size, leading to very high quantum efficiency. Quantum efficiency tells the number of electrons collected per photon incident [29].

To improve the noise performance, an active pixel is introduced [24, 30]. The working of an active pixel:

Fig. 19 Passive pixel sensor with charge integration amplifier [29]



1. The reverse diode photodiode is set to a high voltage, before exposing the pixel to light.
2. During the time of exposure, the incoming photons decrease the voltage of the photodiode.
3. Using pixel address, voltage across the diode is drawn using a source follower circuit at the end of the exposure.
4. Photodiode is reset for a new cycle.

Active pixel sensor uses MOS buffer amplifier to use the voltage sensing mode for signal readout. Buffer amplifier is used as a source follower in the pixel. The use of a buffer amplifier gives higher readout speed and lower noise level. Power dissipation is minimum in active pixels as amplifiers are switched on during readout. Modern active pixel sensors have managed to reduce the variation and fixed pattern noise [29, 31].

(a) **Three transistor APS**: As shown in Fig. 20, with one photodiode and three transistors readout circuit, 3T pixel is the most used form in present day CMOS active pixel sensor [29]. The three transistors are row-select (M1), source follower (M2) and reset transistor (M3). Light intensity is determined by double sample circuitry. Due to the fact that temporal noise in 3T APS is high, double sampling actually increases the noise power. Yet it is majorly used APS because of its reasonable fill factor (around 30%).

(b) **Four transistor APS**: An additional transistor called transfer gate (TG) is added between diode and the reset transistor (M3) as shown in Fig. 21. It helps in clearing the dark current charge and leakage current [29]. The voltage is subtracted from reset value using correlated double sampling [29, 32].

(c) **Pinned photodiode pixel**: Though these pixels and its different versions reduces noise level, the KTC noise component is still present [24, 33]. To overcome this issue, an improvement in-pixel sensor is done using the pinned photodiode principle and pinned photodiode pixel (PPD) was introduced as shown in Fig. 22 [24]. It has a diode that is fully depleted with mobile charges [29, 34]. It has several features like:

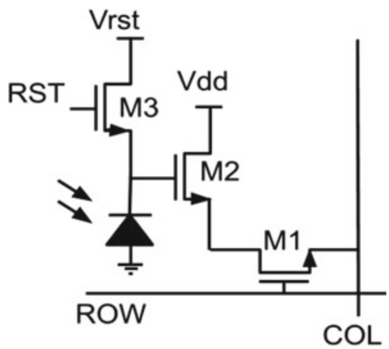


Fig. 20 Three transistor APS

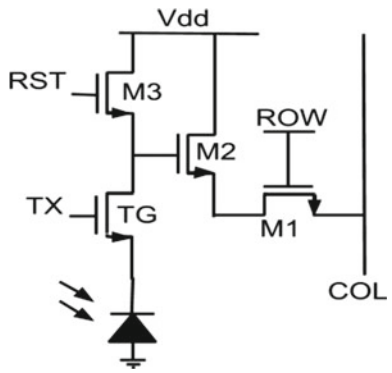


Fig. 21 Four transistor APS

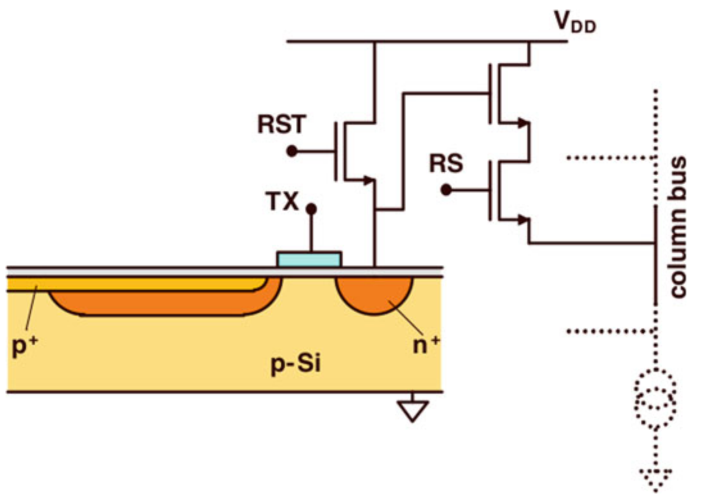


Fig. 22 CMOS imager pixel with pinned photodiode [24]

1. kTC noise is completely cancelled by means of correlated double sampling [35].
2. The intrinsic charge storage capacitance is higher because of $(p+)-n$ and n-p substrate junction, which also provides large dynamic range.
3. The intrinsic charge storage capacitance is higher which gives a large dynamic range.

(d) **Other pixels**: Depending on the application, there are various other pixel structures like for high dynamic range, logarithmic APS is suitable as shown in Fig. 23.

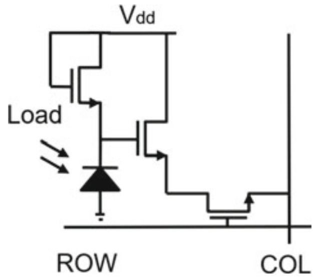


Fig. 23 Logarithmic APS schematic

3.3 Signal Processor

This block on the chip carries circuits which help in improving functionality of the CMOS image sensors. This addition of signal processing circuitry on-chip helps in reducing both fixed pattern noise and temporal noise; e.g. correlated double sampling negates the FPN. It does the analogue to digital conversion and also helps in providing higher SNR. It may contain additional circuitry based on the application like circuitry for motion detection, compression, etc.

4 Developments in Sensor

Each of the image sensor types have developed from the time to time, since they were created. The below section gives some of the improved versions of these image sensors.

4.1 Developments in CCD Image Sensor

CCD image sensor is one of the earliest step towards image sensor developments, capturing and storing images in digital formats. These image sensor are even preferred today for scientific equipments (like microscope, telescope) due to their high spatial resolution and high sensitivity with low SNR. CCD image sensor have improved a lot in terms of material, fabrication technology and device design. They have decreased in dark current, power reduction in pixels and improved in quantum efficiency along with improved signal handling. The skipper CCD image sensor is one of such development made to improve **Charge Transfer efficiency (CTE)**. The design shows the output region and the floating gate electrode used to detect signal charge in the channel near the end of a three-phase horizontal register shown in Fig. 24. CCD image sensor requires a bias to transfer charge packets. This increases the overall noise floor, this noise is higher than that from the amplifiers on chip causing lower efficiency of charger transfer. The CCD Skipper was invented to circumvent the 1/f noise problem

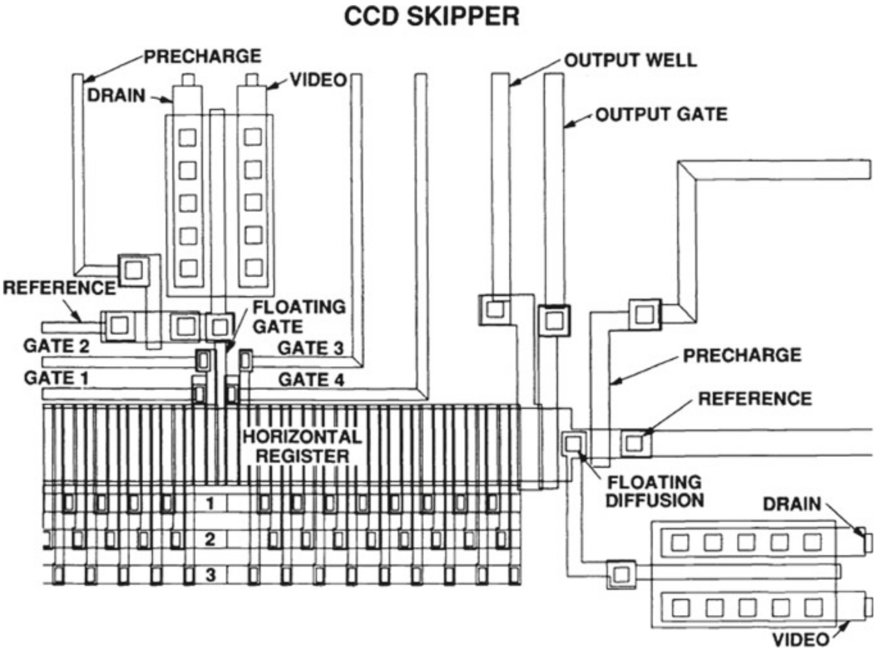


Fig. 24 Design layout of a Skipper CCD [36]

and realize a square root reduction in noise with increasing sample time thereby allowing sub-electron noise floors to be achieved. Figure 24 shows the design layout of the skipper that was fabricated in Ford Aerospace.

The design illustrates the output region and the floating gate electrode used to detect signal charge in the channel near the end of a three-phase horizontal register. The **Quantum efficiency (QE)** is one of the main feature where CCD image sensor is improving in present time to reduce the loss. The increase in QE can be achieved by thinning CCD for use in back side illuminated mode. The methods used to improve QE of back illuminated CCDs are to reduce reflection losses by the deposition of backside thin film coatings, several aspects of device packaging and surface charging to eliminate the backside potential well [37]. Similarly specific features are required to be more accurate than the other based on requirements.

4.2 Developments in CMOS Image Sensor

Every application has different requirements. Based on the requirements, one feature will have more priority than the other. Research has been mainly on features like high dynamic range, reducing noise and accordingly different sensor architectures

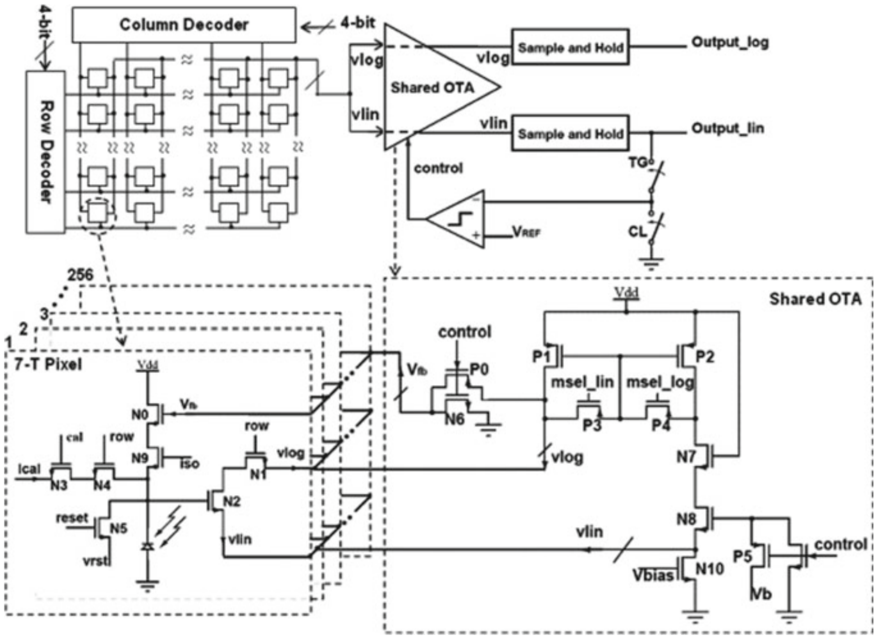


Fig. 25 CMOS image sensor for high dynamic range [39]

are proposed. Following the sensor architectures proposed for a specific application improving a particular feature.

High dynamic range (HDR)—dynamic range is a ratio of two signals of a certain quality. Definition of dynamic range changes from the factor to factor like audio, video, etc. In the case of photography, it is the ratio of saturation signal-to-noise floor. Dynamic range plays a crucial role in scientific imaging applications. A CMOS imager architecture for scientific application is shown in Fig. 25, which provides high dynamic range. This architecture uses combined linear-log technique [38] which improves the readout operation. Linear-log switching is used which reduces post-signal processing [39]. This image sensor is suited for scientific imaging which requires detection of low light scattering.

Noise—the normal readout architecture and use of correlated double sampling reduces the kTC noise, but it also brings out the effect of random telegraph signal noise (RTS noise) [40] which is a main issue in low light applications. As shown in Fig. 26b, a, readout circuit is proposed which reduces the RTS noise [41]. The principle is to switch an nMOS transistor from accumulation mode to inversion mode. This switch biasing technique is the main principle in this circuit. It decreases noisy pixels without affecting output readout speed.

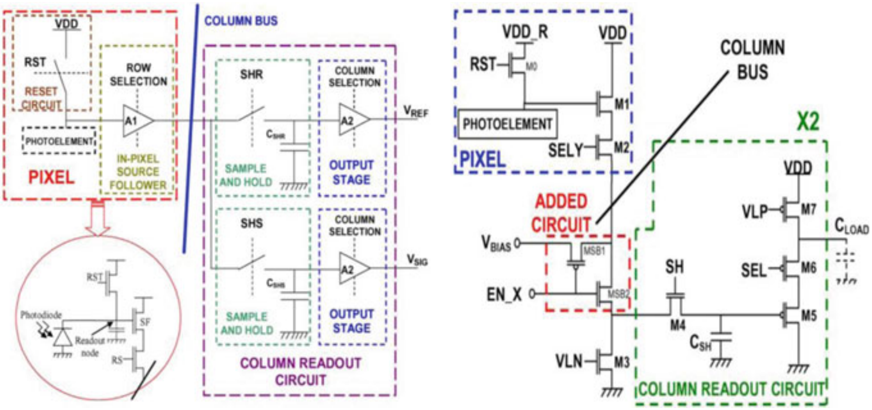


Fig. 26 **(a)** General architecture of image sensor, **(b)** architecture of image sensor with new readout circuitry for reducing RTS noise [41]

5 Applications

CMOS image sensors are used in a wide-range of application areas. Due to its various advantages and features, it is suitable for digital photography and medical applications. CCD image sensor is used for cameras that are used in the field of astronomy, colour cameras which have more clarity than CMOS for static objects [42]. The application of the image sensor is understood by the outcomes from the recent developments in the field. Few of the papers along with their outcomes are shown in Table 2.

The biological laboratories are using the CCD camera for imaging methods for its resolution, ease of using and flexibility. For making a best choice of the CCD to be used for an application, it requires the basic knowledge of the parameters that affects the performance.

5.1 *Photography*

From consumer electronics to security purposes, cameras are used everywhere. CMOS image sensors play a major role in many applications like portable cameras, camcorders, surveillance cameras [43], etc. From the invention of CMOS image sensors till now, there has been much improvement in CMOS image sensors. Various resolutions of cameras are coming into market with many new image correcting algorithms for a high resolution, proper and sharp image. Most of the mobile phones use CMOS image sensors in their camera system.

Table 2 Comparison table

References	Paper title	Outcome
[1]	e2v new CCD and CMOS technology developments for astronomical sensors SPIE digital library	• Methods to manufacture high performance sensor for astronomical and scientific use • Use of sensor in space technologies. Novel CCD and CMOS design for low noise and high sensitivity • Back-thinned CMOS sensor for astronomy
[2]	CMOS image sensors: ECLIPSING CCDs in visual information	• Effects of development in CMOS image sensor • Comparison of CMOS and CCD image sensors
[3]	High-speed camera system using a CMOS image sensor	• Camera system for high-speed imaging is presented which is capable of recording images with a resolution of 256 × 256 pixels and frame rates in excess of 1000 frames per second • Uses an image sensor with on-chip electronic shutter • The camera system contains an image memory for sequence recording and does not require external processing, provides fast interface between image acquisition and image processing unit
[6]	Charge-coupled semiconductor devices	• Describes the CCD semiconductor concept which consists of storing charge in potential wells created at the surface of a semiconductor and moving the charge • Minority carrier charge storage at the SiSiO2 interface of a MOS capacitor
[18]	A review of charge-coupled device image sensors	• Gives the overview of the charge-coupled devices used as image sensor • Physical principles, device operation, enhancements and performance of the CCD
[22]	High-resolution CMOS video image sensors	The basics of complementary metal-oxide semiconductor (CMOS) image sensors and the architectures/technologies used to obtain such high resolutions, and frame rates are reviewed and the technologies to obtain the highest possible quality video images are explained

(continued)

Table 2 (continued)

References	Paper title	Outcome
[23]	Architectures for focal-plane image processing	• On-chip image pre-processing for solid-state imagers using analogue CCD circuits is described for low, medium and high-density detector arrays A spatially parallel architecture for low density, high throughput applications is described
[24]	CMOS image sensors: state-of-the-art	Overview of CMOS image sensor. main focus on shrinkage of pixels. Along with comparison of CCD and CMOS technology
[27]	A random-access photodiode array for intelligent image capture	• The basic requirements of the chip which random access and separation that exists between reading and sampling process are shown • The chip includes an 80*80 matrix of basic cells • Each cell consists of two stages: the first is based on a switch, while the second includes a buffer
[29]	CMOS image sensors: recent innovations in image technology	The earlier designs of pixels with CIA suffer from high parasitic capacitance which results in lower output. Modern pixels use column-wise CIA along with access transistor which improves the performance and leads to high quantum efficiency. Discuss the development, design and working of typical 4T APS
[39]	A high dynamic range CMOS image sensor for scientific imaging applications	A CMOS sensor is proposed which uses linear-logarithmic readout scheme. This scheme helps in operating each individual pixel based on the input levels. This scheme uses switching point detection for the application. The readout architecture is designed using 0.5um CMOS technology, and the sensor contains 7 T pixel array. Dynamic range of 121 dB has been observed on evaluating the sensor chip
[41]	Novel readout circuit architecture for CMOS image sensors minimizing RTS noise	To reduce random telegraph signal (RTS) noise in pixels, a readout architecture and readout sequence for CISs is proposed. The proposed architectures reduce the noise effect without much reduction in performance of the pixels. 0.35um CIS technology image sensor is used in the work

(continued)

Table 2 (continued)

References	Paper title	Outcome
[43]	Review of CMOS image sensors	• This gives the use of CMOS image sensor compared to the CCD sensor • Gives the noise and sensitivity which are the key factors for the sensor. Types of CMOS sensor along with the applications
[44]	A smart CMOS image sensor with on-chip hot pixel correcting readout circuit for biomedical applications	A detection function with correlated double sampling is used in single slope A/D convertor. While saving image sharpness, peak signal-to-noise ratio generated is about 30 dB for a hot pixel density as high as 10%
[21]	Device reliability for CMOS image sensors with backside through-silicon vias	This study uses backside illumination (BSI) CIS to characterize the device reliability and explains about backside process optimization to maintain the required device reliability. Frontside illumination is used in early days of CMOS image sensor; however, as BSI provides high quantum efficiency and high sensitivity, it replaced the FSI technology used for image sensors. 0.18um CMOS image sensors are used to show the reliability of image sensor with BSI technology

5.2 Space Applications

Space science includes study of bodies such as planets and stars. Observing them carefully is a very important task. Image sensors play a crucial role in space technology for different applications. CMOS image sensors are used in space technology because of its various advantages. From navigation cameras, lander and rover image cameras to star tracking, constellations cameras, they are used everywhere. CMOS image sensors with some specific changes are very tolerant to radiation which makes it a very preferable device in space technology [45].

In case of the CCD images sensor in the space application [46] is vital due to the feature of absorbing more amount of light source especially static objects. The two capabilities which make most of the CCD sensor much versatile for scientific imaging is binning and subarray. Properties like readout speed increases, reduces the exposure time, i.e. the total information that has to be transferred to computer.

5.3 Medical Applications

CMOS image sensors are used in radiology, 3-D X-ray microtomographic system, X-ray imaging and X-ray imaging Spectroscopy. A bio-sensor employs CMOS image sensor for extracting information about molecular transformations [44]. CMOS image sensor can also be used to derive blood velocity.

5.4 Automobile Applications

CMOS image sensors are also used in automobile industries [47]. They are used from navigation, object detection, airbag system, crash sensing system and vision system for low light or night driving cases.

5.5 Military Applications

The CCD is used in a special technique to reduce the SNR value of the image so that there is more image quality [48]. Military uses this for remotely piloted vehicles (RPV) [49], aircraft imaging systems, high-resolution camera systems, missile guidance systems and multispectral scanner systems [50].

6 Methodology

Steps for manufacturing image sensors is as follows - first, a metal pad is framed over a semiconductor substrate. Next, a film is formed over the semiconductor substrate for protection and the metal pad. At that point, the protection film is specifically taken out to uncover a surface of the metal pad. Next, a first planarization film is framed over the protection film. At that point, a colour filter layer is formed over the planarization film formed initially. Next, a subsequent planarization layer is framed over the colour filter layer. At that point, a first material layer is framed throughout the subsequent planarization layer. Next, a subsequent material layer is shaped over the material layer formed initially. At that point, a microlens is shaped out of the first and second material layers.

In CCD image sensors, a majority of metal-oxide silicon (MOS) capacitors is situated neighbouring one another. Charge transporters can be put away in every capacitor and moved to its adjoining neighbours (Fig. 27).

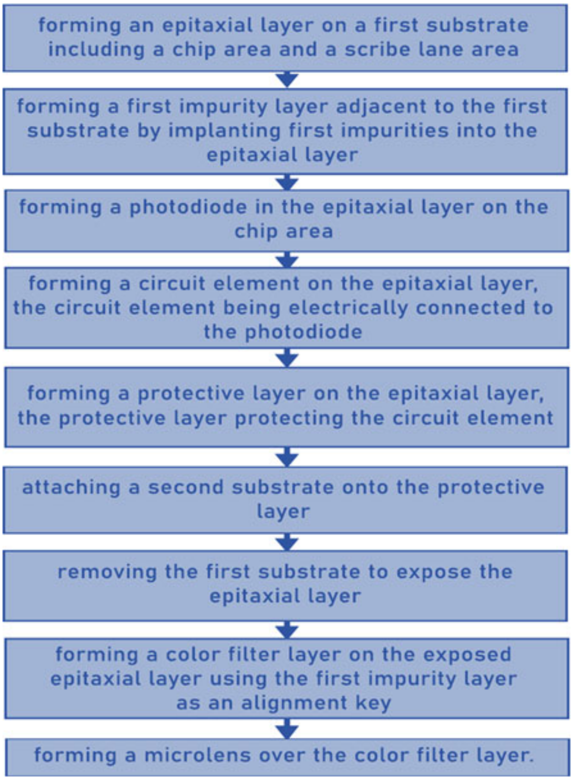


Fig. 27 Image sensor manufacturing flowchart

7 Results and Discussion

Performance of various CMOS image sensors Table 3 here, the types of CMOS sensor is given along with the fill factor, pixel size, dynamic range and transistor per pixel value of each of them are 20–60%, 10–30, 96–140 dB and 4–18 respectively. Performance of CMOS image sensors is evaluated under various factors like fill factor and dynamic range. It is observed that stepped reset-gate CMOS image sensor has higher fill factor, whereas linear-logarithmic CMOS image sensor has high dynamic range. Image sensors are developing to get much better imaging at lower lighting conditions as well as to get much better video recording through those image sensors.

8 Conclusion

The CMOS image sensor is favoured over the CCD sensor due to the increased speed of processing, and comparatively, good pixel size. The CMOS sensor is having a

Table 3 Performance of various CMOS image sensors

CMOS type	Stepped reset-gate CMOS	CMOS with logarithmic switch points	Snapshot CMOS	Linear-logarithmic CMOS	Linear-logarithmic CMOS	
					With injected charge signal	With in-pixel noise reduction
Technology (um)	0.8	0.5	0.35	0.18	0.35	0.25
Pixel size	–	23.4 × 27.15	18 × 18	5.6 × 5.6	7.5 × 7.5	10 × 10
Pixel level transistor	4	7	18	7	4	5
Fill factor (%)	49	24.56	15	33	37	60
Dynamic range (dB)	96	121	97	143	124	140
HDR technique	Well capacity adjustment	Combined linear log	Multiple reset	Combined linear log	Combined linear log	Logarithmic

fill factor of 30–60% with HDR capabilities and 0.37–0.4 μm technology. Charge-coupled device (CCD) and complementary metal-oxide semiconductor (CMOS) devices are the two main categories of image sensors. CCD sensors are the devices where all pixels data is collected and processed at the end, whereas in CMOS sensors, each pixel has its circuitry for processing. Each of them has its own advantages and disadvantages. Based on many factors like noise, power consumption, voltage operation and sensitivity, these advantages are focused. This paper starts with the history of images sensors, which demonstrates the early inventions of CCD and CMOS image sensors. Following, this paper explains the architecture of each sensor, their pixel designs and factors affecting their performance. This paper demonstrates how image sensors have developed over years and their importance in different application areas. In some cases, CMOS image sensors are preferred over CCD image sensors because of their fast processing nature, especially in daily life photography as CMOS imagers are good at negating the blooming and smearing effect. CCD image sensors have their own predominating applications like military, space applications and professional camera because of its better image quality. This paper also demonstrates the latest image sensor architectures for negating and getting a better output based on a particular feature.

Recent image sensors have established their position in applications like high-resolution image capturing, producing sharper and low noise images. However, there are still issues to improve on. CMOS image sensor for high dynamic range, presented in this paper, reduced the complexity of post-processing signal, but it is also observed that due to smaller voltage input to the sensor, the FPN noise is larger compared to the sensor used for traditional consumer application. Image sensors are having wider application in many fields due to the wide variety of environmental conditions in which all these image sensors are exposed so the pixels which are used in the CCD and CMOS have to maintained in certain conditions for better SNR and good quality images so the development of the pixel which can be used in wide variety of conditions is one of the major scope of development. The development of the image sensor which has high zoom capability without losing the quality of the image to be captured.

References

1. Jorden PR, Jordan D, Jerram PA, Pratlong J, Swindells I (2014) e2v new CCD and CMOS technology developments for astronomical sensors. In: High energy, optical, and infrared detectors for astronomy VI, vol 9154, p 91540M. International Society for Optics and Photonics
2. Kempainen S (1997) Cmos image sensors:eclipsing ccds in visual information. EDN 42(21):101–102. (See also pp. 105–6, 108, 110)
3. Hillebrand M, Stevanovic N, Hosticka B, Conde J, Teuner A, Schwarz M (2000) High speed camera system using a CMOS image sensor. In: Proceedings of the IEEE intelligent vehicles symposium 2000, pp 656–61
4. Smith GE (2010) Nobel lecture: the invention and early history of the CCD. Rev Mod Phys 82(3):2307
5. Boyle WS (2010) Nobel lecture: CCD—an extension of man's view. Rev Mod Phys 82(3):2305
6. Boyle WS, Smith G (1970) Charge-coupled semiconductor devices. Bell Syst Tech J 49:587–593
7. Zarnowski J, Pace M, Joyner M (1999) Active-pixel CMOS sensor improve their image. Laser Focus World 35(7):111–114
8. Lecture notes 2 charge-coupled devices (CCDs)—part I—Stanford Education http://isl.stanford.edu/~abbas/ee392b/lect02.pdf
9. Basic principles of Image sensor. http://90.147.102.78/fisica-iaselli/wp-content/uploads/sites/34/2018/01/Fisica-moderna.pdf
10. Understanding Digital Image sensor—LUCID vision labs—resources https://thinklucid.com/tech-briefs/understanding-digital-image-sensors/
11. Dyck RH, inventor; Lockheed Martin Corp, assignee. Rapid-sequence full-frame CCD sensor. United States Patent US 5,754,228. 1998 May 19
12. Lecture 03: astronomical Detectors—Colorado Education http://lasp.colorado.edu/~cranmer/ASTR_7500_2016/Lec_Wenda/wcao_lecture03.pdf
13. Allan AJ (2016) Examination of myocardial electrophysiology using novel panoramic optical mapping techniques (2016)
14. Bases de Radiométrie Optique, J.L. Meyzonnette, Th. Lépine, 2ème Édition, Éditions Cépadues, 2001
15. Wikipedia. Charge-coupled device Wikipedia, The free encyclopaedia, 2015 (Online) https://en.wikipedia.org/w/index.php?title=Charge-coupleddevice&oldid=667558930. Accessed 4 July 2015
16. Sze SM, Ng KK (2006) Physics of semiconductor devices. Wiley
17. PowerPoint lecture notes from King Abdullah University of Science and Technology. Integrated Image Sensors. Course EE391A, lecture 12
18. A review of charge-coupled device image sensors—scientific figure on researchGate. Available from: https://www.researchgate.net/figure/A-cross-sectional-view-of-a-storage-site-in-a-buried-channel-CCD-Adapted-from_fig8_313504888. Accessed 14 Apr 2020

19. Blouke MM (1991) Model of a thinned CCD, pp 136–143 (Online). Available: https://doi.org/10.1117/12.50465
20. Bortoletto F, D'Alessandro M (1986) Universal CCD-controller system. Rev Sci Instrum 57(2):253–258
21. Gambino JP, Soleimani H (2018) Device reliability for CMOS image sensors with backside through-silicon vias. In: 2018 IEEE international reliability physics symposium (IRPS), 2018
22. Takayanagi I, Nakamura J (2013) High-resolution CMOS video image sensors. Proc IEEE 101(1):61–73
23. Fossum ER (1989) Architectures for focal-plane image processing. Opt Eng 28(8):865–871
24. Theuwissen AJP (2008) CMOS image sensors: state-of-the-art. Solid-State Electron 52:1401–1406
25. Weckler G (1967) Operation of p-n junction photodetectors in a photon flux integrating mode. IEEE J Solid-State Circ SC-2(3):65–73
26. Nakamura T, Matsumoto K (1991) Active pixel image sensors: recent developments of the CMD. Tech Dig Int Symp Adv Image-Acquisition Technol 70–79
27. Yadid-Pecht O et al (1991) A random-access photodiode array for intelligent image capture. IEEE Trans Electron Dev 1772–80
28. Peter JW (1968) Noble. IEEE Trans Electron Dev 15–4:202–209
29. Khanduri G (2019) CMOS image sensors: recent innovations in image technology. Int J Eng Res Technol 7(12). ISSN: 2278-0181
30. Weckler GP (1967) Operation of p–n junction photodetectors in a photon flux integrating mode. IEEE J Solid-State Circ 1967:65–73
31. Mendis et al (1997) IEEE J Solid-State Circ 32:187–1970
32. Fowler B et al (2001) Proceedings of SPIE 4306:68–77
33. Guidash RM et al (1997) A 0.6 lm CMOS pinned photodiode colour imager technology. IEDM Tech Digest 1997:927–9
34. Lule T et al (2000) IEEE Trans Electron Dev 47:2110–2122
35. White MH et al (1973) Characterization of charge-coupled device line and area-array imaging at low light levels. ISSCC digest of technical papers. pp 134–135
36. Janesick JR, Elliott TS, Dingiziam A, Bredthauer RA, Chandler CE, Westphal JA, Gunn JE (1990) New advancements in charge-coupled device technology: sub electron noise and 4096 x 4096-pixel CCDs. In: Charge-coupled devices and solid-state optical sensors, vol 1242. International Society for Optics and Photonics, pp 223–237
37. Lesser MP (1994) Improving CCD quantum efficiency, pp 782–791 (Online). Available: https://doi.org/10.1117/12.176777
38. Storm G, Henderson R, Hurwitz JED, Renshaw D, Findlater K, Purcell M (2006) Extended dynamic range from a combined linear-logarithmic CMOS image sensor. IEEE J Solid-Sate Circ 41:2095–2106
39. Guo J (2009) A high dynamic range CMOS image sensor for scientific imaging applications. IEEE Sens J 9(10):1209–1218
40. Findlater KM, Vaillant JM, Baxter DJ, Augier C, Herault D, Henderson RK, Hurwitz JE, Grant LA, Volle JM (2004) Source follower noise limitations in CMOS active pixel sensors. In: Proceedings of SPIE—detectors and associated signal processing, vol 5251, pp 187–195
41. Martin-Gonthier P, Magnan P (2011) Novel readout circuit architecture for CMOS image sensors minimizing RTS noise. IEEE Electron Devices Lett 32(6):776–778
42. Durini D (2014) High performance silicon imaging: fundamentals and applications of CMOS and CCD sensors. Elsevier
43. Habibi M (2010) A low power smart CMOS image sensor for surveillance applications. In: 2010 6th Iranian conference on machine vision and image processing, 2010
44. Cao Y, Tang F, Bermak A, Le T (2010) A smart CMOS image sensor with on-chip hot pixel correcting readout circuit for biomedical applications. In: 5th IEEE international symposium on electronic design, test and applications, 2010
45. Photobit Corporation, USA and Alexandria University, Egypt (2000) Design of radiation hard CMOS APS image sensors for space applications. In: Proceedings of the 17th national radio science conference, 17th NRSC'2000 (IEEE Cat. No.00EX396)

46. Jerram P, Stefanov K (2020) CMOS and CCD image sensors for space applications. In: High performance silicon imaging 2020 Jan 1. Woodhead Publishing, pp 255–287
47. Sheng Zhang S (2013) On-screen-display (OSD) and SPI interface on CMOS image sensor for automobile application. In: 5th international conference on computational intelligence, communication systems and networks, 2013
48. Jain SC, Bist KS (1986) Military applications of CCD image sensors. IETE Tech Rev 3(12):591–596. https://doi.org/10.1080/02564602.1986.11438043
49. K Fcssis, R Ely & L Z Mau, Application of CCO imagers in high shock environments, 3rd Int'l Conf. on the Tech and Application of ceo, pp 92–95, 1976.
50. Bist KS, Jain SC, Kumar V (1986) CCO based multispectral scanner. JETE Tech Rev 3:220–223

Swift Double-Tail Dynamic Comparator

K. B. Sowmya and Meghashree Doddamani

Abstract The requirement for analog to digital with radical less power, more speed and limited area requirement has made the converters to move in the direction of the use of dynamic regenerative comparators in order to increase the power efficiency and the throughput. In the work, a study has been carried out on the delay and utilization of power in dynamic comparators. Designers will acquire an insight of the key suppliers to the comparator delay from the analytical expressions and thoroughly explore the trade-offs in dynamic comparator design. A novel dynamic comparator has been proposed based on the presented research, in which the predictable double-tailed comparator circuit is optimized for fast operation and lower power even in low supplies. Without obscuring the process by adding limited transistors resulting in substantially reduced delay time and applying domino logic technique at the output nodes to reduce the power. It is shown that both the delay time and power intake are reduced considerably in the proposed dynamic comparator. The Cadence virtuoso tool is used to design the comparator.

Keywords Comparator · Dynamic comparator · Domino logic · Double-tail dynamic comparator (DTDC) · High-speed ADCs

1 Introduction

Comparators are the basic constructing blocks of maximum digital circuits that require small chip area with low-power, high-speed comparators. CMOS devices suffer from low supply voltages. Hence, they move to the FinFET devices, a low-cost and high-performance transistor. Instead of conventional flat design, it uses a fin-like structure, which may allow engineers to more compact circuits and build faster operation and computer chips. The word "FINFET" represents a non-planar, double-gate transistor based on the design of a single-gate transistor, designed on an SOI substrate.

K. B. Sowmya (✉) · M. Doddamani
Department of Electronics and Communication Engineering, R V College Of Engineering, Bangalore, India
e-mail: kb.sowmya@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_41

Hence, it is more difficult to build high-speed comparators with less supply voltage. In other words, to attain more speed in a specified technology, superior transistors are prerequisite to recompense for the reduction in supplies. This indicates that there is an additional power and die space. In addition, lower-voltage process results in restricted input assortment for common mode, which is essential in countless higher speed ADC architectures, and example is flash ADCs. Different techniques are like body-driven transistors [1, 2] techniques, supply boosting methods [3, 4], current-mode strategy [5] and dual-oxide progressions. In spite of the advantages, very less trans-conductance will be suffered from the body-driven transistor (which is equal to the transistor's gmb) related to its gate-driven counterpart, whereas distinct manufacturing processes, like deep n-well, require equally nMOS transistors and pMOS transistors to work in body-driven conformation. In the proposed comparator circuit, static power dissipation is abridged. The static power obtained is minimal so that the circuit becomes more effective and takes a smaller amount of power and energy. Compared to existing comparators, the performance of output swing in the proposed comparator is high. For dual-rail comparator, common mode (Cm) voltage is high due to dual rail property; simulation result shows differences in the value of delay and in the value of average power considering input voltage in common mode. In the work of comparator, it is seen that comparatively the digitized output is high. There are two Voltage+ and Voltage− input terminals and single digitized V_o output. The output V_o equation is depicted in Eq. (1).

$$V_o \begin{cases} 1, \text{ if } V_+ > V_- \\ 0, \text{ if } V_+ < V_- \end{cases} \tag{1}$$

2 Swift Double-Tail Dynamic (DTD) Comparator

The application of comparator is most commonly found in real-time applications which require low voltage. Hence, the design of proposed comparator is done using double-tail module. The proposed comparator's significant notion is to advance the limitation $\Delta V_{\mathrm{fn/fp}}$ to advance the velocity of latch renaissance. In a cross-coupling manner, two control MOS transistors (MC1 along with MC2) were connected to the first stage in equivalent to M3 along with M4 MOS transistors.

During the reset process, it is necessary to set CLK as low, Mtail1 and Mtail2 as low to avoid static power. By doing this procedure, MOS transistor M3 along with MOS transistor M4 pulls altogether fn node and fp node to high voltage and then the MC1 MOS transistor and MC2 MOS transistor goes to low state. The in-between stage transistors, MOS-MR1 and MOS-MR2 rearrange all latch outputs to ground. Transistors M3 and M4 switch off during executive process (clock is considered to be VDD, and Mtail1 and Mtail2 are taken to VDD). Turning on transistor as long as fn continues to fall, pulling fp node back to the VDD; an additional switch MOS transistor MC2 halts off letting fn designate completely discharged. Distinct from conventional DTDC, a function $\Delta V_{\mathrm{fn/fp}}$ of input MOS transistor the trans-conductance and the change in input voltage, a PMOS transistor-MC1 switches on, dragging the additional node fp hindmost to the upper voltage; in the modified structure as soon as the comparator senses, the node fn discharges faster, for example. Therefore, as time goes by, the difference between fn and fp ($\Delta V_{\mathrm{fn/fp}}$) increases exponentially, leading to a decrease in latch revival time [6–10]. Regarding the efficiency of the modified concept, one of the points to note is that in this circuit, when anyone of the control MOS transistors (e.g., MOS-MC1) switches on, a VDD is haggard to the ground through contribution transistor and extremity transistor (e.g., MOS-MC1, MOS-M1 and MOS-Mtail1); consequentially, it results in static power congestion. To overawe this problem, two NMOS switches are cast off under the [Msw1 and Msw2] contribution transistors, as shown in Fig. 1. Due to the fact that both fn and fp nodes were pre-loaded to VDD (during the reset process) at the beginning of the decision-making process, both MOS switches that are closed and fn and fp begin falling with unlike discharge rates. The instant the comparator senses that one of the fn/fp nodes discharges more quickly, the control transistors will behave in a way that increases their difference in voltage. The conventional DTC can be adapted as proposed DTD comparator. Figure 1 shows the proposed DTDC. The proposed DTD comparator is faster than the other conventional comparators and consumes a less power [11–20].

Applying domino logic at the output of the comparator which reduces the power consumptions, two PMOS transistors are used at the latch block which provide better results in terms of output swing to the circuit.

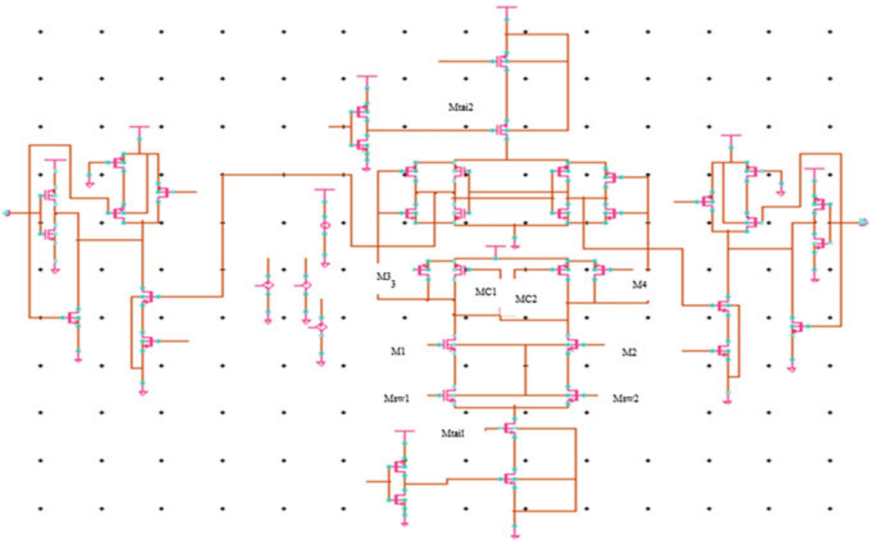


Fig. 1 Swift DTDC

At the output side, two PMOS transistors give better amplification of the output signal. If MR1 is high, then MR2 is low due to pairing in block, then MR4 is pulled to the ground and OUTp to VDD and OUTn to low as MR3 is in off state.

3 Simulation and Results

To compare performance of the proposed DTDC with the conventional dynamic comparator and modified dynamic comparator, the proposed DTD comparator is simulated in the Cadence virtuoso using 18 nm technology and FinFET device with supply voltage 0.8 V.

Figure 2 shows the transient response of the proposed DTD comparator. Here, the output nodes OUTp and OUTn show the result of the proposed DTD comparator. Where INP is high, the OUTp node starts to charging, and where INN is high, the OUTn node starts to charging. The operation is happened only when there is high clock input.

Table 1 shows the performance of the different comparator. Here, it shows the delay and power consumption of the different comparators. By the analysis and simulated result of different comparators, it demonstrates that the proposed comparator is efficient and reliable for low-power and high-speed applications.

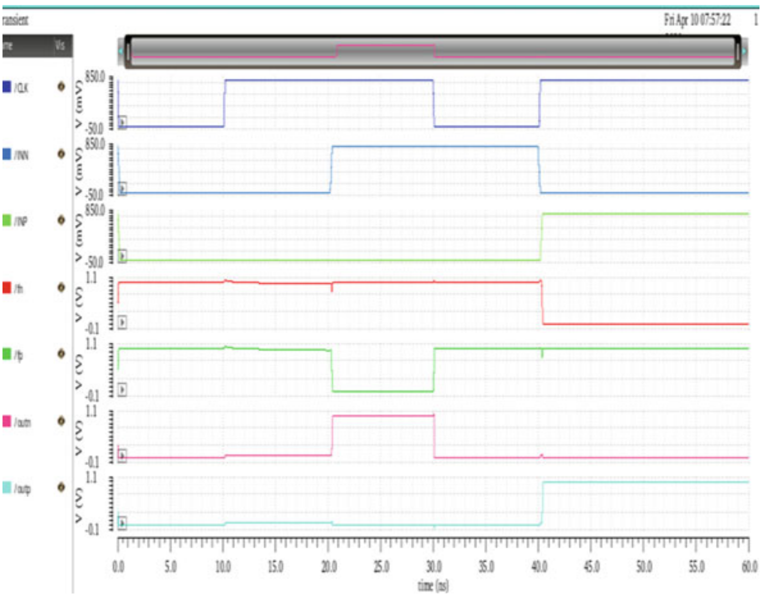


Fig. 2 Transient response of proposed DTDC at VDD = 0.8 V

Table 1 Performance comparison of the comparators

Comparators	Delay (ps)	Power (us)
Dynamic comparator	117.6	254
Double-tail dynamic comparator	78.064	115
Modified double-tail dynamic comparator (main)	70.9	50.26
Modified double-tail dynamic comparator (final)	60.23	43.6
Swift double-tail dynamic comparator	51.55	11.47

4 Conclusion

This paper presented the delay and power analysis of the various comparators and which one is the better performance compare to the other comparators. The conventional dynamic comparator has more delay compare to the double-tail comparator. For increasing the power efficiency, it introduced modified comparator. For reducing the power and to increase the speed reduction of the comparator, this paper has implemented new comparator using FinFET. Using Domino logic to reduce the leakage power, FinFET devices take less supply voltage.

References

1. Nuzzo P, De Bernardinis F, Terreni P, Van der Plas G (2008) Noise analysis of regenerative comparators for reconfigurable ADC architectures. IEEE Trans Circ Syst I Reg Pap 55(6):1441–1454
2. Nikoozadeh A, Murmann B (2006) An analysis of latched comparator offset due to load capacitor mismatch. IEEE Trans Circ Syst II Exp Briefs 53(12):1398–1402
3. Goll B, Zimmermann H (2007) Low-power 600MHz comparator for 0.5 V supply voltage in 0.12 μm CMOS. IEEE Electron Lett 43(7):388–390
4. Schinkel D, Mensink E, Klumperink E, Van Tuijl E, Nauta B (2007) A double-tail latch-type voltage sense amplifier with 18ps Setup+Hold time. In: Proceedings of IEE international solid-state circuits conference. Digest of technical papers, Feb. 2007, pp 314–315
5. Babayan-Mashhadi S, Lotfi R (2012) An offset cancellation technique for comparators using body-voltage trimming. Int J Analog Integr Circ Sign Process 73(3):673–682
6. Goll B, Zimmermann H (2009) A comparator with reduced delay time in 65-nm CMOS for supply voltages down to 0.65. IEEE Trans Circ Syst Exp Briefs 56(11):810–814
7. Ay SU (2011) A sub-1 volt 10-bit supply boosted SAR ADC design in standard CMOS. Int J Analog Integr Circ Sign Process 66(2):213–221
8. Mesgarani A, Alam MN, Nelson FZ, Ay SU (2010) Supply boosting technique for designing very low-voltage mixed-signal circuits in standard CMOS. In: Proceedings of IEEE international Midwest symposium on circuits and systems. Digest of technical papers, Aug 2010, pp 893–896
9. Blalock BJ (2000) Body-driving as a low-voltage analog design technique for CMOS technology. In: Proceedings of IEEE southwest symposium on mixed-signal design, Feb 2000, pp 113–118
10. Maymandi-Nejad M, Sachdev M (2003) 1-bit quantiser with rail to rail input range for sub-1V—modulators. IEEE Electron Lett 39(12):894–895

11. Okaniwa Y, Tamura H, Kibune M, Yamazaki D, Cheung T-S, Ogawa J, Tzartzanis N, Walker WW, Kuroda T (2005) A 40Gb/ s CMOS clocked comparator with bandwidth modulation technique. IEEE J Solid-State Circ 40(8):1680–1687
12. Goll B, Zimmermann H (2007) A 0.12 μm CMOS comparator requiring 0.5V at 600MHz and 1.5V at 6 GHz. In: Proceedings of IEEE international solid-state circuits conference. Digest of technical papers, Feb 2007, pp 316–317
13. Goll B, Zimmermann H (2009) A 65nm CMOS comparator with modified latch to achieve 7GHz/1.3mW at 1.2V and 700MHz/47μW at 0.6V. In: Proceedings of IEEE international solid-state circuits conference. Digest of technical paper, Feb 2009, pp 328–329
14. He J, Zhan S, Chen D, Geiger RL (2009) Analyses of static and dynamic random offset voltages in dynamic comparators. IEEE Trans Circ Syst I Reg Pap 56(5):911–919
15. Kim J, Leibowits BS, Ren J, Madden CJ (2009) Simulation and analysis of random decision errors in clocked comparators. IEEE Trans Circ Syst I Reg Pap 56(8):1844–1857
16. Babayan-Mashhadi S, Lotfi R (2013) Analysis and design of a low-voltage low-power double-tail comparator. IEEE Trans Very Large Scale Integr (VLSI) Syst 34(4):1063–8210
17. Sah N, Mittal E. An improved domino logic. In: 2017 international conference on energy, communication, data analytics and soft computing (ICECDS-2017), vol 21
18. Van de Plassche RJ, Baltus P (1988) An 8-bit 100-MHz full-Nyquist analog-to-digital converter. IEEE J Solid-State Circ 23(6):1334–1344
19. Jensen JC, Larson LE (2003) A 16-GHz ultra-high-speed Si-SiGe HBT comparator. IEEE J Solid-State Circuits 38(9):1584–1589
20. Goll B, Zimmermann H. A 65nm CMOS comparator with modified latch to achieve 7GHz/1.3 mW at 1.2 V and 700MHz/47μW at 0.6 V. In: 2009 IEEE international solid-state circuits conference, San Francisco, CA, USA, 2009, pp 328–329

Role of Thermal Sensor in Detection of Early Stage of COVID-19

Deepak Prasad, Sanjay Kumar Surshetty, Vidushi Goel, and Vijay Nath

Abstract COVID-19 is the common enemy of all of us in this world. It created lot of deaths, loss of economy and many more. The pandemic has created a lot of innovation to bring solutions in helping fight the spread of novel coronavirus and other diseases. In this paper, various thermal sensors for detection of COVID-19 in early stage are discussed.

Keywords Thermal detection · IR detection · Smart camera · Artificial intelligence · Sensors

1 Introduction

Nowadays, this pandemic has become the common topic of research interests, and it also became very challenging to find a way to get out of it [1]. In this paper, some sensors were discussed which help in detecting if a person becomes victim of COVID-19. This helps in resuming our work in daily life without risk after a long gap.

Three non-invasive techniques have been emerged for this purpose like we can do analysis of air, waste water samples and temperature sensors (Fig. 1).

2 Methodology

The various ways of detection include.

D. Prasad (✉)
Department of Electronics and Communication Engineering, Sarala Birla University, Ranchi, India

S. K. Surshetty · V. Goel · V. Nath
Department of Electronics and Communication Engineering, Birla Institute of Technology Mesra, Ranchi, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_42

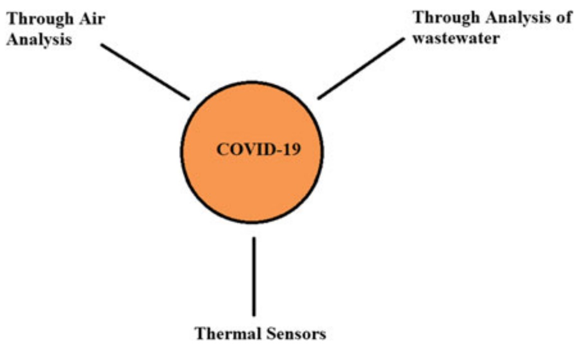


Fig. 1 Ways of detection of COVID-19

2.1 Analysis of Air

For measuring concentration of virus floating in air, optical biosensors are in development stage [2]. This work on two affects mainly optical and thermal one. Generally, this works on basis of specific RNA structure with respect to virus. Its detection is similar to analogy of two complementary objects inserted to become one full object. Same technique was used before in pollution detection devices [3]. Once this is completely developed, it could be used in detection of future pandemics well in advanced [4].

2.2 Analysis of Waste Water Samples

This helps whether particular area is safe are not but cannot point out particular person specifically who is infected with virus [5]. This works based on speed with which wastewater moves and gives snapshot of actual viral content in sewer community showing large group of people who are infectious [6–9].

2.3 Thermal Sensors

This is deployed near various workplaces, airports, shopping malls, industries, general stores, etc. nowadays. There are many ways to check temperature with various devices, but, however, WHO has warned that although temperature measurement does not completely detect COVID-19, fever is not only major symptom but also can cause pneumonia, etc. [10]. Also, temperature measured should be accurate. There are also some limitations. There are non-contact sensors which are safe even for staff.

Non-contact sensors include infrared spot sensors, thermal camera sensors. Contact type are thermometers which we generally keep in mouth to measure temperature, but it requires continuous cleaning which is not possible during this pandemic due to large group of people. Actually, pandemic could have been prevented if we have replaced sensors planted near borders which were installed 10 years back. As we know, everything is smart nowadays, the requirements of accurate AI-driven cameras need to be installed wherever required [11]. By combining various sensor outputs, we can detect various other symptoms beyond temperature. There is elevated skin temperature (EST) which is outer temperature of body. Generally, outer skin temperature is lesser than inner body temperature.

2.3.1 Infrared Spot Sensor

This does not work on all individuals at a time but work on every individual separately. It is non-contact type which compares EST with threshold temperature and responds with green or red signal. These devices are added with additional compensation circuitry to account the effect of room/surrounding air temperature. Also, thermal infrared cameras are developed which are more accurate. These thermal sensors can scan large group of people and also detect temperatures of them individually [12].

2.3.2 Smart Cameras

As contact sensors require manpower involved at various places, it is quite difficult to find those many workers. So, equipment which cannot work without manpower is useless. So, AI-driven automated cameras which also connect to cloud help in remote monitoring without manpower required. These sensors also monitor various persons at a time. It is cost-effective, easy to operate and also helps in tracking the person who is infected. This also helps in reduction of human error due to lack of attention. As it can be connected to cloud, information can be easily shared and controlled from command center itself. These devices are also well trained to detect sneezing and coughing. But, whatever info we get is not sufficient; also, the suspected victim should be sent to authorities [13]. If there are smart connected cities, that would also help in tracking individuals who were in contact with infected victim. Some places are also installed with thermal station in which person is required to enter the station. These stations also give alarm signals if any part of the system is damaged [14]. Some sensors are developed like which can scan large group of people at a time and then can mark as green or red. Then, this data can be linked to map to be able to track individuals and can also trigger alarms if person tracked with green is found near person tracked with red. There is limited accuracy and data privacy concern also. Accuracy is highly important despite of urgency during this pandemic. If people detected incorrectly for fever, then there is more requirement of staff and testing kits. If people are missed in detection, then it is a hazard for future COVID-19 attacks. Also, there is more requirement of training and cleaning of rectal thermometers, etc. more often. So, thermal cameras are being installed everywhere. Some of the limitations include that temperature of whole body is not uniform. Also, person forehead temperature can vary due to outside climate which could cause incorrect detection. The most accurate temperature can be recorded if the camera can focus on inner corner of the eye. Since small change in angle could cause much variation, it is difficult to be measured. Privacy is a concern because whatever information they get by scanning a person should be stored in highly protected place. Those who do not comply with rules laid by ISO should be discarded. These is vast area of development being done to detect like accuracy, speed, number of people scanned at a time and their distance from equipment.

3 Conclusion/Remarks

Here, various ways of detection of COVID-19 are discussed. But, still these do not guarantee 100% accuracy as already said, fever is not the only symptom of this virus. So, there is still development going on to overcome challenges for making better world for future as well as present.

Acknowledgements The authors are sincerely thankful to Vice Chancellor Dr. I. Manna for providing such great infrastructure and laboratories. A piece of thanks also goes to head of department Dr. S. Pal for his constant encouragement.

References

1. https://serverscheck.com/solutions/corona-covid-19.asp. Accessed on 12 Feb 2020
2. Rabl A (2003) Interpretation of air pollution mortality: number of deaths or years of life lost? J Air Waste Manage Assoc 53:41–50
3. Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL (2000) Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. N Engl J Med 343:1742–1749. https://doi.org/10.1056/NEJM200012143432401
4. Pope CA3, Schwartz J, Ransom MR (1992) Daily mortality and PM10 pollution in Utah Valley. Arch Environ Health 47:211–7
5. World Health Organization (WHO) (2011) Guidelines for drinking-water quality, 4th edn. WHO Press, Geneva
6. Dissmeyer GE (2000) Drinking water from forests and grasslands, South Research Station, USDA Forest Service, Ashville, NC, USA
7. Nollet LML (2000) Handbook of water analysis. Marcel Dekker, New York
8. DeZuane J (1997) Handbook of drinking water quality. Wiley
9. Fawell JK (1993) The impact of inorganic chemicals on water quality and health. Annali dell'Istituto Superiore di Sanita 29(2):293–303
10. Smart optical sensors, smart systems, Austrian Research Centers GmbH. http://www.smart-systems.at/products/products smart optical sensors en.html. Accessed on 31 Jan 2020
11. SmartCam Project (2020) Graz University of Technology—Institute for Technical Informatics.http://www.iti.tu-graz.ac.at/en/research/smartcam/site/. Accessed on 31 Jan 2020
12. Sanjay Kumar S, Khalkho A, Agarwal S, Prakash S, Prasad D, Nath V (2019) Design of Smart security systems for home automation. In: Nath V, Mandal J (eds) Nanoelectronics, circuits and communication systems. Lecture notes in electrical engineering, vol 511. Springer, Singapore,

pp 599–604 (Online). ISBN: 978-981-13-0776-8. https://doi.org/10.1007/978-981-13-0776-8_56
13. Prasad D, Nath V (2019) An overview of temperature sensor. In: Nath V, Mandal J (eds) Proceeding of the 2nd international conference on microelectronics, computing and communication systems (MCCS 2017) lecture notes in electrical engineering, vol 476, pp 777–786, Springer, Singapore (Online). https://doi.org/10.1007/978-981-10-8234–4_62
14. Shi Y, Tsui T (2007) An FPGA-based smart camera for gesture recognition in HCI applications. In: Proceedings of 8th Asian conference on computer vision, 2007

Generative Adversarial Network-Based Satellite Image Enhancement

Rudra Narayan Pandey, Shreyas Shubhankar, Bibhudendra Acharya, and Sudhansu Kumar Mishra

Abstract In recent times, satellite images are gaining popularity in the monitoring of agricultural fields, forest ranges, surveillance, and public commute convenience. In many places, it is difficult to implant cameras; thus in that scenario, satellite images establish to be very helpful. However, the raw image taken by the satellite needs to be preprocessed to remove cloud cover, improve contrast, perform color correction, and improve inter-image contrast. In this paper, a method is devised to remove cloud cover and improve inter-image contrast. Since deep learning is gaining immense popularity, a deep learning approach utilizing generative adversarial networks (GANs) has been used for the enhancement of satellite images. To improve inter-image contrast, histogram matching technique has been used. The technique proposed in this paper has been investigated by considering the satellite images collected from SKYMAP Global.

Keywords Convolutional neural network (CNN) · Generative adversarial networks (GANs) · Spatial–temporal–spectral (STS)

1 Introduction

Several emerging applications utilize satellite images to solve problems ranging from monitoring of agricultural fields [1] to public navigation systems [2]. Satellite images play an important role in meteorology [3], remote sensing, and remote surveillance. These applications take the help of computer vision, remote sensing techniques to create clear images required for further processing. Satellite images

R. N. Pandey · S. Shubhankar · S. K. Mishra
Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: sudhansumishra@bitmesra.ac.in

B. Acharya (✉)
Department of Electronics and Communication Engineering, National Institute of Technology, Raipur, Chattisgarh, India
e-mail: bacharya.etc@nitrr.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_43

are usually composed of various channels that extend beyond the visible band of light. Three common categories of satellite images are infrared, visible, and water vapor images. Visible imagery is used for applications like monitoring of forest ranges, infrared imagery is used for meteorology, and water vapor imagery is mostly applied for thunderstorm prediction. Some satellites are capable of capturing multiple bands of light simultaneously, and satellite images, hence, captured are referred to as multi-spectral images. They are generally composed of four different channels: blue, green, red, and infrared channels. The composition of the channels varies according to the intended application of the satellite imagery and may include bands, such as radio, near-infrared, and microwave.

There are many hindrances encountered before using satellite images for terrestrial analysis. Aerosol present in the atmosphere scatters light along with the oceans' surface and makes the analysis of subsurface oceans difficult [4]. The presence of clouds in the image obscures the surface and makes it difficult to observe the terrestrial features. Thick cloud cover completely covers the surface and may require different electromagnetic bands to extract limited terrestrial features. In contrast, thin cloud cover poses a smaller difficulty and can recover most surface features using mathematical and artificial intelligence techniques. Satellite images usually only capture a small amount of surface at a time, and these images need to be stitched together to form an image mosaic that can be used for analysis. These mosaics generally suffer from color differences among different images, and a clear boundary between two images is sometimes visible. These color differences usually result from different sunlight reflectance of surfaces, atmospheric scattering, and temporal differences (Table 1).

Many of the problems discussed above are solvable in some aspects, and possible solutions of cloud cover along with inter-image contrast enhancement will be briefly touched here before being fully discussed in the upcoming sections. Qing Cheng et al. [6] proposed a technique to substitute missing pixels with the help of neighboring pixels using Markov random fields (MRF) global function to remove cloud cover. Aldo Maalouf et al. [7] used Bandelet transform and the multiscale geometrical grouping to perform cloud removal. Martinuzzi et al. [8] proposed a method in which cloudy images are detected and substituted with images of the same place clicked on a different date without any cloud cover. In the artificial intelligence domain, Kenji Enomoto et al. [11] proposed a technique to generate RGB images from multispectral images with no clouds. Praveer Singh and Nikos Komodak in [12] used a technique called Cloud-GAN that learns the relationship between cloudy and non-cloudy images using the adversarial loss to constrain the distribution.

Two images taken side by side can have color differences which will result in a visible boundary where one image ends, and the other begin. Hence, the removal of this difference is essential before creating an image mosaic. Histogram matching is a popular technique, whereby the histogram of the query image is made to match the histogram of the reference image [5]. In the field of artificial intelligence, Guo et al. [13] proposed a variant of CNN that was used for contrast enhancement by extracting multiscale features from the intermediate layers of the CNN model. The cost function

Table 1 Comparative study of different papers

Reference No.	Applied methodology
Cloud removal techniques	
[5]	Cloud-free images were constructed by using image mosaic techniques on multitemporal images
[6]	Substituted missing pixels with the help of neighboring pixels using Markov random fields (MRF) global function
[7]	Used Bandelet transform and multiscale geometrical grouping to perform cloud removal
[8]	Cloudy images were detected and substituted with images of the same place clicked on a different date without any cloud cover
[11]	Generated RGB images from multispectral images with no clouds
[12]	Used a technique called Cloud-GAN that learns the relationship between cloudy and non-cloudy images using the adversarial loss to constrain the distribution
[14]	Performed an analysis of visible band space characterization to rectify the RGB band of cloudy Landsat images using haze-optimized transformation (HOT)
[15]	Used shortwave infrared imagery to dehaze visible bands of image
[16]	Restored Landsat ETM+ images via data captured by MODIS
[17]	Proposed a method to fill gaps in data using neighborhood similar pixel interpolator (NSPI) technique
[20]	Proposed a spectral–spatial–temporal technique which used the spatial correlations to non-local regions from local regions for the restoration of satellite images
[21]	Used CNN to remove cloud cover from an image using a different temporal image
Color correction techniques	
[1]	Aims to correct the color using a reference image to join adjacent images to make a mosaic
[13]	A variant of CNN was used for contrast enhancement by extracting multiscale features from the intermediate layers of the CNN model

was calculated by using Wasserstein distance to measure the inconsistency between two color distributions.

In this paper, a generative adversarial network (GAN) and histogram matching-based technique has been proposed for satellite image enhancement. Also, several techniques have been investigated and compared, which are being used to solve the problem of cloud removal and inter-image contrast enhancement to obtain cloud-free and even-colored stitched images. The advantages and disadvantages of different techniques are also discussed. This paper dives in depth in the recent advancement of artificial intelligence techniques that claim to solve the issues found in satellite imagery. The satellite image has been collected from SKYMAP Global.

2 Cloud Removal

2.1 Broad Cloud Removal Approaches

Removal of clouds is basically a reconstruction process of images and the approaches for the same can be categorized into four parts:

(1) *Non-complementation Approaches*

The remotely sensed image is used to reconstruct cloud contamination region information by utilizing the remaining image sections without the help of any complementary data. A typical technique in this approach is missing pixel interpolation. It should be noted that interpolation works best only for small pixel gaps and is not suitable for huge data gaps caused by the presence of a cloud. New techniques that have emerged in this approach are geometry wavelet, patch filling, and maximum a posteriori.

The above techniques reconstruct missing regions by using the geometrical structure of the remaining image sections around the image sections. Radiometric information gathered from the satellite along with the image can be used to synthesize more realistic results to generate cloud-free images.

(2) *Multispectral Complementation*

In this approach, multispectral image data is used for image reconstruction. Since clouds block visible band of light, the relationship between visible and auxiliary bands can be used to reconstruct missing regions in the visible band. Zhang et al. [14] performed an analysis of visible band space characterization to rectify the RGB band of cloudy Landsat images using haze-optimized transformation (HOT) method. They selected a spectral space containing spectral response of different land cover types in a clear atmosphere which they claimed to be highly correlated. The resultant surface response vector was termed as 'clear line.' The direction of clear line can be showed in terms of its slope angles α, and HOT can be expressed as:

$$\mathrm{HOT} = \mathrm{B1}\sin(\alpha) - \mathrm{B3}\cos(\alpha) \tag{1}$$

where B1 and B3 are the digital numbers (DNs) of the pixel's bands 1 and 3, respectively.

Their method cannot be used in absolute sense since absolute information over many scenes will be unavailable. Therefore, 'α' must be approximated form clearest scenes with 'α' varying from scene to scene. Li et al. [15] used shortwave infrared imagery in their method to dehaze visible bands of image. Similarly, a method to use data captured by MODIS to restore Landsat ETM+ images is proposed by Roy et al. [16]. These techniques work only for thin cloud cover and face difficulties when thick clouds are present.

(3) *Multitemporal Complementation*

This technique is more robust than the above-mentioned techniques and is able to remove thick cloud covers from the image. It is easier to obtain multitemporal images since the satellites are capable of acquiring images of same area multiple times a day. In this method, cloud-free sections of the image from a certain time can be used to fill in the lost data from the image section taken at a different time.

Cloud-free images can also be constructed by using image mosaic techniques on multitemporal images (Helmer and Ruefenacht [5]). Chen et al. [17] proposed a method to fill gaps in data induced by the Scan Line Corrector (SLC)-off problem and Landsat ETM using neighborhood similar pixel interpolator (NSPI) technique. Thick cloud cover was removed by using a modified NSPI approach. Other techniques, such as the local linear histogram matching (LLHM) approach (Scaramuzzaet al. [18]) and the weighted linear regression (WLR) approach (Zeng et al. [19]) were also developed to tackle the problem. With these methods, spatial discrepancies may arise in the reconstructed image since it is not possible to accurately substitute the missing pixels.

(4) *Multispatial–Temporal–Spectral Complementation*

In this method, a combination of spatial, temporal, and spectral data is used to recover missing information from the image. Li et al. [20] proposed a spectral–spatial–temporal technique which used the spatial correlations to non-local regions from local regions for the restoration of satellite images. The concept of CNN can also be used for recovering degraded image. This technique is further described in the next sub-section.

2.2 Convolutional Neural Network Approach

Convolutional neural network or CNN is the combination of convolutional processing and fully connected artificial neural network. The convolutional part of CNN comprises of convolutional, max pooling and activation layer. The output of convolutional layer passes through activation, and then max/average pooling layer. The convolutional part gets repeated, and after some repetition, the output of convolutional part is made to pass through a fully connected layer which itself consists of input, hidden, and output layer. In convolutional layers, kernels are used to extract useful information needed for mapping input to output. Most common kernel sizes are 3×3, 5×5, and 7×7 grid pixels. Many CNN structures have already been proposed that were designed for specific applications. Some of the examples are AlexNet, ResNet, and InceptionNet and can also be used in the cloud removal application (Figs. 1, 2 and 3).

Zhang et al. [21] have used CNN to remove cloud cover from an image using a different temporal image. Their method learns the complex non-linear relationship between input $\times$ 1 representing spatial data with gaps and temporal data, respectively.

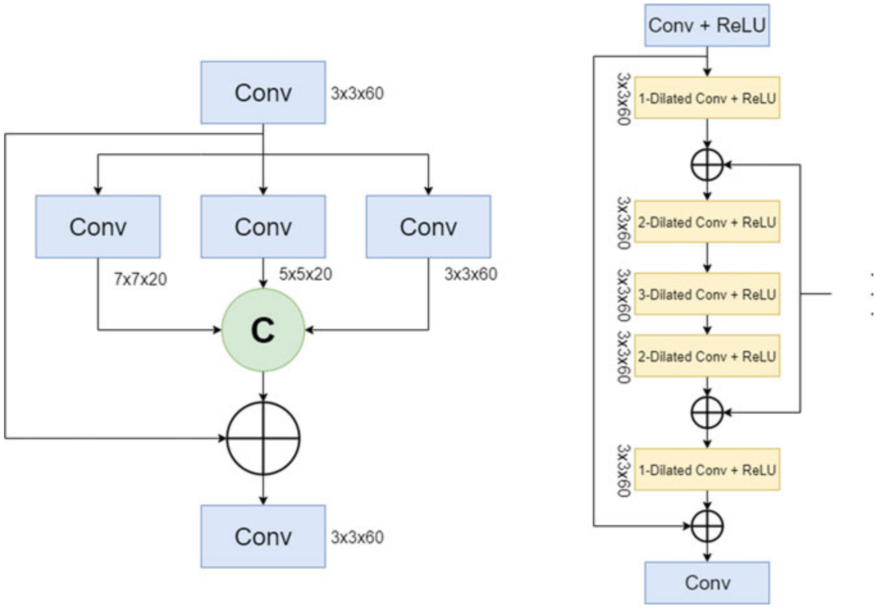


Fig. 1 The conv structure

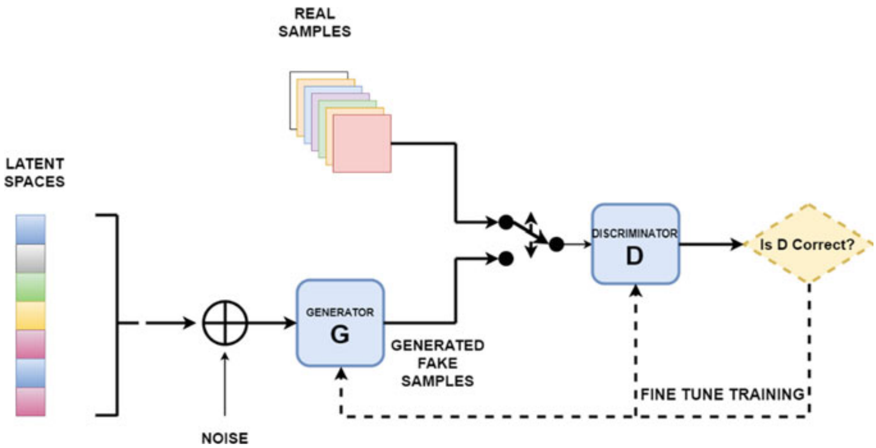
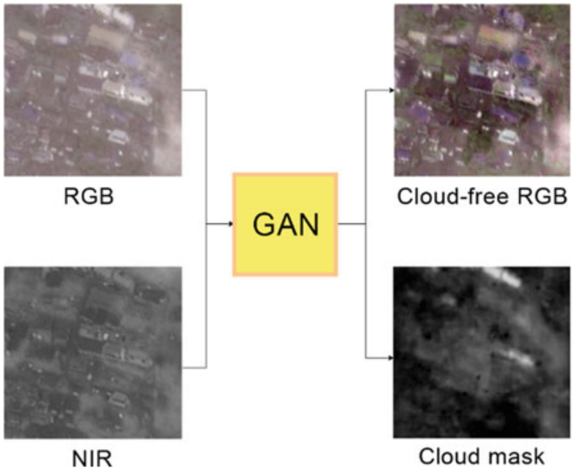


Fig. 2 Skip connection structure as proposed by Zhang et al. [21] for multiscale convolution and dilated convolution, respectively

As explained above, spectral and temporal data can vastly enhance the precision of the reconstruction as the image data usually has a high degree of correlation with the missing regions in the surface features. In their method, they have taken two images as input, one in which cloud is present and one in which cloud is absent. Both the images are passed through a layer of convolutional operation with a 3×3

Fig. 3 Flowchart for generative adversarial network



kernel size generating an output of 30 feature maps. The output of both images when concatenated results in 60 feature maps. This output was then passed to 3 × 3, 5 × 5, and 7 × 7 kernels simultaneously to extract multi-context information, and then, the output of all these kernels were concatenated. Also, convolution parameters can be introduced by increasing the kernel filter size, which can increase the complexity of the training model. Dilated convolutions are being used to solve this issue. The common convolution receptive field and the layer depth has a linear correlation, where the receptive field size is $F_{\text{depth-}j} = (2j + 1) * (2j + 1)$. The dilated convolution receptive field and the layer depth has an exponential correlation, which is equal to

$$F_{\text{depth}-j} = (2^{j+1} - 1) * (2^{j+1} - 1) \tag{2}$$

Further, they have defined residual mapping, i.e.,

$$r_j = y_j - x_j \tag{3}$$

where x_j and y_j (the input 1 in Fig. 4) are the original undamped image and the image with missing data, respectively. Residual mapping results in a decreased training loss by employing a more effective learning status when passed through a multi-layer network. If the number of layers in CNN is increased, then more features can be extracted, but it also results in gradient vanishing or exploding problem. So, in order to mitigate this problem, skip connection was employed as it maintains the image details while passing the feature information of previous layer to its posterior layer.

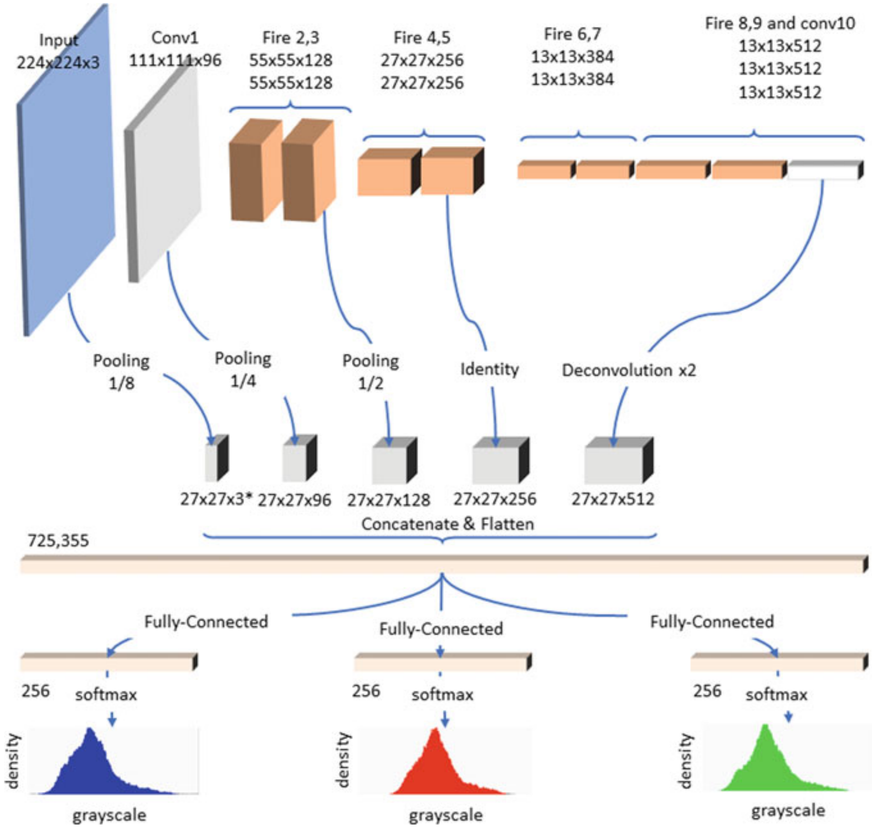


Fig. 4 GANs framework for cloud removal proposed by Enomoto et al. [11]

2.3 Generative Adversarial Network

Cloud removal employs deep learning to generate images free of cloud. CNN requires labelled data, while GAN does not need any cloud-free images to train the network. In deep learning, there are basically two methods to generate new images: variational auto encoder (VAE) [9] and generative adversarial networks (GANs). Generally, GANs are used because they are very robust way of creating artificial images that closely mimic the desired image [10]. There are many variations of GANs which are used for generating images, and hence, GANs find its application in cloud removal. Kenji Enomoto et al. [11] proposed a technique to generate RGB images from multispectral images with no clouds. Singh and Komodakis [12] have used a technique called Cloud-GAN that learns the relationship between cloudy and non-cloudy images using adversarial loss to constrain the distribution.

Enomoto et al. [11] have taken two different band of image as input: RGB and infrared image. Infrared light can penetrate inside thin cloud covers, and hence, it is

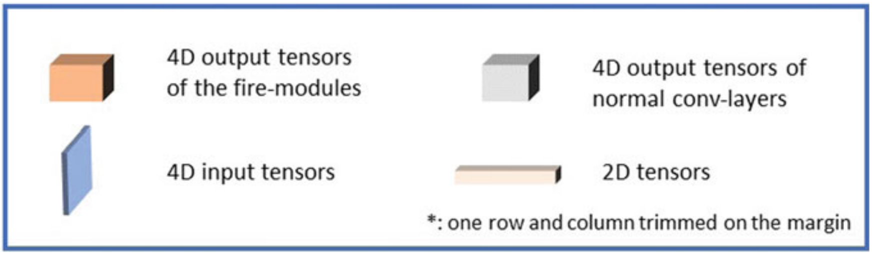


Fig. 5 CNN model structure proposed by Guo et al. [13]

used in this method. In GANs, there are two networks discriminator and generator. The generator generates fake images, and the discriminator rejects those images unless a similar image to the input of discriminator is generated. In GAN, two loss functions as mentioned in Eqs. 4 and 5 are optimized simultaneously to reach a Nash equilibrium (Fig. 5).

At discriminator D,

$$\begin{aligned} D\_\text{loss}_{\text{real}} &= \log(D(x)) \\ D\_\text{loss}_{\text{fake}} &= \log(1 - D(G(z))) \\ D\_\text{loss} &= D\_\text{loss}_{\text{real}} + D\_\text{loss}_{\text{fake}} \\ D\_\text{loss} &= \log(D(x)) + \log(1 - D(G(z))) \end{aligned}$$

The total cost is

$$\frac{1}{m}\sum_{i=1}^{m}\log\left(D\left(x^{i}\right)\right) + \log\left(1 - D\left(G\left(z^{i}\right)\right)\right) \tag{4}$$

At generator G,

$$G\_\text{loss} = \log(1 - D(G(z)))\text{or } -\log(D(G(z)))$$

The total cost is

$$\begin{aligned} &\frac{1}{m}\sum_{i=1}^{m}\log\left(1 - D\left(G\left(z^{i}\right)\right)\right) \\ &\text{Or } \frac{1}{m}\sum_{i=1}^{m}-\log\left(D\left(G\left(z^{i}\right)\right)\right) \end{aligned} \tag{5}$$

3 Inter- and Intra-Image Enhancement

There are multiple variations of this technique used by researchers. One of the techniques used by Burgan et al. [1] is referred to as Image Match, aims to correct the color using a reference image to join adjacent images to make a mosaic. Their histogram matching technique works by comparing cumulative distribution functions. Their method used only image overlap areas to create a lookup table, and the histogram matching required no scaling. Their methodology requires no manual interpretation; however, it depends on finding overlap regions which may not be present in many cases and recreating images of the overlap region may result in loss of data.

In [13], the authors have used Wasserstein CNN to output histograms or distributions by analyzing the input images. Reference images were used only for training, and not to create corrected images after the model was successfully trained. After obtaining the required histogram, histogram matching is performed to obtain the output image. This method does not require any reference image to create color corrected image, and hence, this technique shows promise in removing atmospheric scattering-related digital value variation and obtaining color-corrected image.

Guo et al. [13] have developed a CNN model that uses Wasserstein distance as a loss function, and to reduce memory and computational complexity, modified version of SqueezeNet is used. SqueezeNet is a lightweight CNN, and consists of two basic modules: squeeze layer and expand layer. In squeeze layer, 1×1 size kernel is used while 3×3 size kernel is used as an unexpanded layer. This model is computationally efficient since it helps to reduce the overfitting problem. They have removed final global average pooling and SoftMax layer of SqueezeNet, and the rest part has been utilized to extract useful information from the raw image. The authors have avoided overfitting and color discrepancy problem by employing data augmentation on the original input. Three augmentation operations are: random cropping, random flipping, and random color augmentation. In random cropping, they have extracted an image patch of random size from the original image. In random flipping, each patch is randomly flipped vertically or horizontally. In random color augmentation, brightness, saturation, and gamma values of the patch are changed randomly.

4 Simulation Result

Specification of the system on which tests were executed is Dell Inc. Inspiron 7560; ×64 based Windows 10 operating system; Intel(R) Core (TM) i5-7200U CPU @ 2.50 GHz, 2701 MHz, 2 Core(s), 4 logical processors; 8 GB RAM; 250 GB SSD; and GeForce 940MX 4 GB graphics card.

To evaluate the proposed method, experimental results are listed and discussed in this section. From the experimental results, it is shown that the GANs can improve visibility by cloud removal with RGB and NIR satellite images, and the ability of histogram matching to improve inter- and intra-image contrast has been shown. An image is split into four parts. One of the parts is used as reference (first image in row 1) of Fig. 6. The other three parts are manipulated to synthesize fake contrast changed images. The four images are joined and illustrated in Fig. 7. Synthetic cloud is then generated by using Perlin algorithm and then added to the four images using alpha blending and has been illustrated in Fig. 8. Clouds are then removed using GAN model of the four images individually. These images then pass through a histogram matching algorithm which uses the reference image to correct the three images. These images have then been combined to form Fig. 9. This Fig. 9 is then compared with Fig. 10, using different image evaluation metrics as shown in Table 2. Sample results of the synthesized cloud-obscured images are shown in Fig. 6. The columns represent the NIR images, synthesized cloud-obscured RGB images, RGB images predicted by GANs, and the mask images of the clouds predicted by GANs, from left to right.

The final corrected image is very similar to the original image which can be proved by the metrics shown in Table 2.

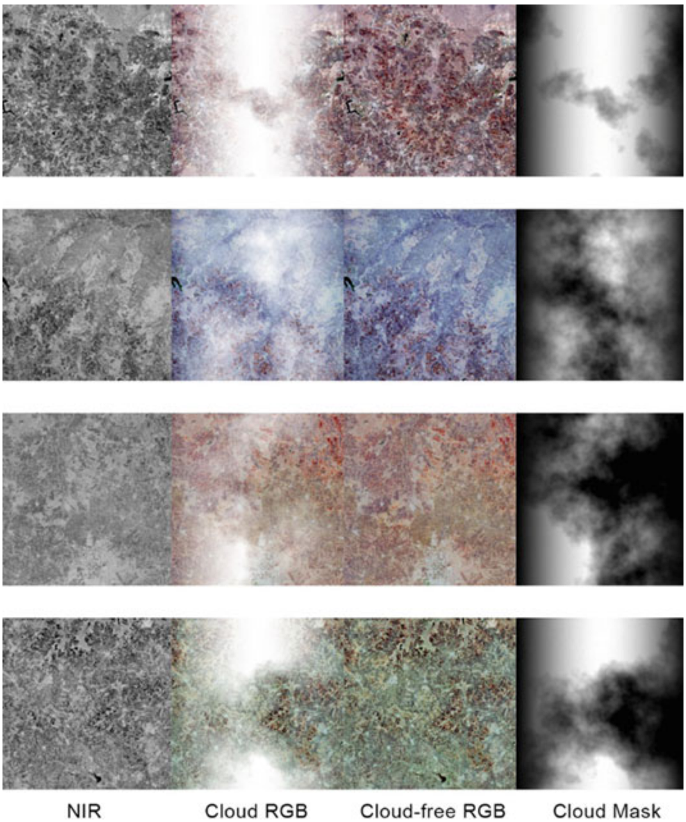


Fig. 6 Prediction results by GANs with the synthesized cloud images

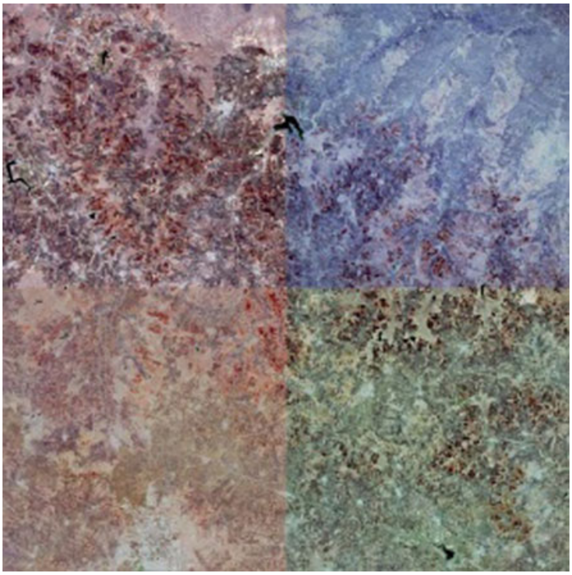


Fig. 7 Synthesized contrasted image

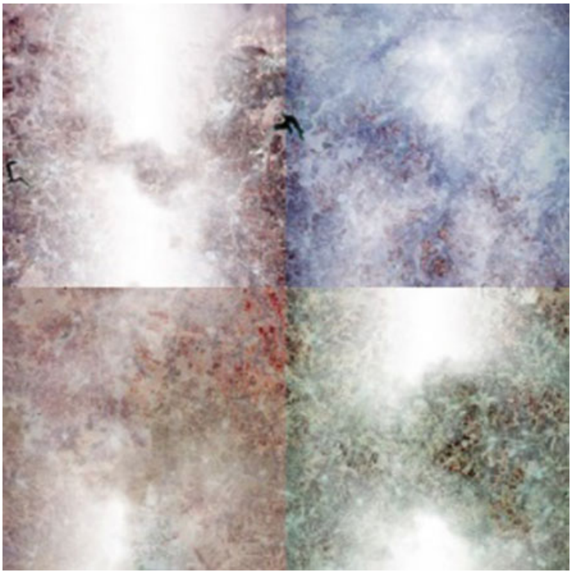


Fig. 8 Synthesized cloud image

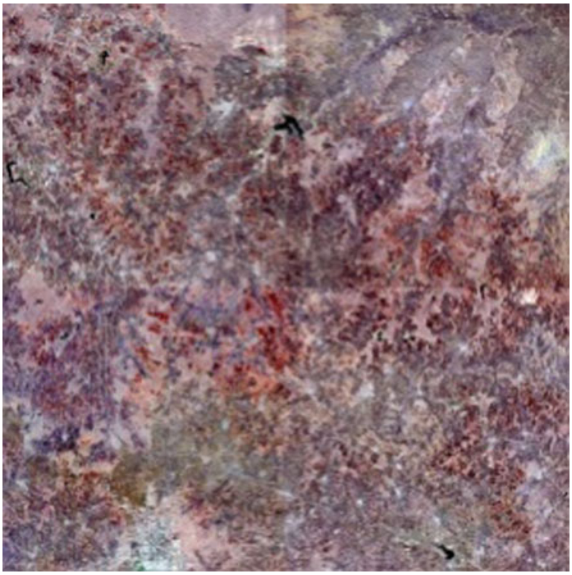


Fig. 9 Corrected image

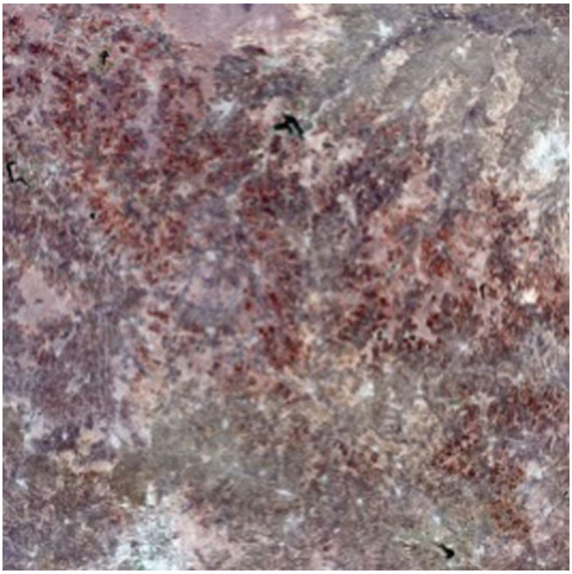


Fig. 10 Reference image

Table 2 Image evaluation metrics comparing Figs. 7, 8, 9, and 10 (PSNR: higher is better, SSIM: higher is better (max value is 1), and MSE: lower is better)

Reference image	Compared image	PSNR	SSIM	MSE
Original image	Contrast-manipulated image	17.28	0.872	1216.35
Original image	Contrast-manipulated image + cloud-added image	11.36	0.58	4752.08
Original image	Final corrected image	25.58	0.97	179.56

5 Conclusion

Great progress has been made in cloud removal and contrast enhancement of satellite images over the years. Earlier rigorous mathematical computations were needed to be performed for cloud removal, but since the advent of neural networks, the computational complexity has decreased that led to the increase in efficiency of the models. CNN was earlier used to remove cloud cover, but they need a training batch consisting of images having cloud cover and no cloud cover. But, GANs have addressed this problem by only utilizing cloud cover images in the training period. GANs have shown promise in solving this problem, but they require large training time and tuning of hyperparameters by trial and error.

In this paper, a method to remove thin clouds from satellite images formed using visible light by extending GANs to multispectral images has been used. The dataset for training networks was constructed by synthesizing simulated clouds with Perlin noise over images without clouds, which made it possible to generate cloud-obscured training images and ground truth of the same area. In addition, to avoid overfitting to certain categories caused by biased datasets, t-SNE has been used to sample images uniformly in each category. Finally, the experimental results evaluated on the constructed data prove that the clouds in the visible light images can be removed if they are penetrated in NIR images. In other words, NIR images should not have clouds if this model is to be used.

Histogram matching technique has also been applied to improve inter-image contrast using a popular and efficient algorithm. Two variants were tested, one in which a different image was used as a reference and the other in which an image of the same area captured at a different time was used. The latter variant produces better images but require an image captured before time while the former variant produces slightly worse result but does not require any prior image.

In the future, a different band of image could be used instead of NIR which can penetrate thicker clouds so that the fourth channel could capture cloudless images albeit without color. Also, the GAN model would be able to provide better accuracy if it is trained with more real images. So, expanding the dataset is of paramount importance and could be done in the future.

Acknowledgements The authors would like to express gratitude towards SKYMAP Global for their financial support in this research. The authors would also like to acknowledge Birla Institute of Technology, Mesra for providing a platform for performing this research work.

References

1. Burgan RE, Hartford RA (1993) Monitoring vegetation greenness with satellite data. Gen. Tech. Rep. INT-GTR-297. US Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT, 13p, 297
2. Prasad R, and Ruggieri M (2005) Applied satellite navigation using GPS, GALILEO, and augmentation systems. Artech House
3. Kidder SQ, Kidder RM, Haar THV (1995). Satellite meteorology: an introduction. Gulf Professional Publishing
4. Gordon HR (1978) Removal of atmospheric effects from satellite imagery of the oceans. Appl Opt 17(10):1631–1636
5. Helmer EH, Ruefenacht B (2005) Cloud-free satellite image mosaics with regression trees and histogram matching. Photogramm Eng Remote Sens 71(9):1079–1089
6. Cheng Q, Shen H, Zhang L, Yuan Q, Zeng C (2014) Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model. ISPRS J Photogramm Remote Sens 92:54–68
7. Maalouf A, Carré P, Augereau B, Fernandez-Maloigne C (2009) A bandelet-based inpainting technique for clouds removal from remotely sensed images. IEEE Trans Geosci Remote Sens 47(7):2363–2371
8. Martinuzzi S, Gould WA, González OMR (2007) Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal. US Department of Agriculture, Forest Service, International Institute of Tropical Forestry. Gen. Tech. Rep. IITF-32, 32
9. Van den Oord A, Kalchbrenner N, Espeholt L, Vinyals O, Graves A (2016) Conditional image generation with pixelcnn decoders. In: Advances in neural information processing systems, pp 4790–4798
10. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680
11. Enomoto K, Sakurada K, Wang W, Fukui H, Matsuoka M, Nakamura R, Kawaguchi N (2017) Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 48–56
12. Singh P, Komodakis N (2018) Cloud-Gan: cloud removal for sentinel-2 imagery using a cyclic consistent generative adversarial networks. In: IGARSS 2018–2018 IEEE International geoscience and remote sensing symposium, pp 1772–1775. IEEE
13. Guo J, Pan Z, Lei B, Ding C (2017) Automatic color correction for multisource remote sensing images with wasserstein CNN. Remote Sens 9(5):483
14. Zhang Y, Guindon B, Cihlar J (2002) An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. Remote Sens Environ 82(2–3):173–187
15. Li H, Zhang L, Shen H, Li P (2012) A variational gradient-based fusion method for visible and SWIR imagery. Photogramm Eng Remote Sens 78(9):947–958
16. Roy DP, Ju J, Lewis P, Schaaf C, Gao F, Hansen M, Lindquist E (2008) Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. Remote Sens Environ 112(6):3112–3130
17. Chen J, Zhu X, Vogelmann JE, Gao F, Jin S (2011) A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. Remote Sens Environ 115(4):1053–1064
18. Scaramuzza P, Barsi J (2005) Landsat 7 scan line corrector-off gap-filled product development. In: Proceeding of Pecora, vol 16, pp 23–27
19. Zeng C, Shen H, Zhang L (2013) Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method. Remote Sens Environ 131:182–194

20. Li X, Shen H, Li H, Zhang L (2016) Patch matching-based multitemporal group sparse representation for the missing information reconstruction of remote-sensing images. IEEE J Select Top Appl Earth Observ Remote Sens 9(8):3629–3641
21. Zhang Q, Yuan Q, Zeng C, Li X, Wei Y (2018) Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network. IEEE Trans Geosci Remote Sens 56(8):4274–4288

Accident Avoidance and Detection on Highways

Deepak Prasad, Ashutosh Anand, Vishnu Anugrahith Sateesh, Sanjay Kumar Surshetty, and Vijay Nath

Abstract Methods and technologies to monitor fatigue levels of driver are being developed, tested and implemented. This project makes use of such technologies which make the car more intelligent, thereby preventing road accidents. There are very few systems existing that record and process data related to the human behavior, and even fewer that are implemented in commercial vehicles. In the current article, a real-time speed regulator in which the vehicle speed is a function of driver fatigue levels, location of the vehicle and its surrounding environment have been proposed. The proposed model uses an advanced method to identify signs of tiredness in drivers and adjust the vehicle speed accordingly to evade road accidents. The key apparatuses of the system consist of real-time sensors like gas sensors, eye blink sensors, blood alcohol level sensors, fuel gauges, impact sensors and a software interface with GPS for finding the location. The proposed project is an attempt to take a step toward safe vehicle navigation. The system automatically senses certain predefined zones and signboards, like school zone, residential zone, Ghat section, blind turn, sharp curve etc. For example, drivers usually over-speed on hilly roads, causing inconvenience and danger to the oncoming vehicles. Though there are safety signs along the road, they are violated by the vehicle drivers. This objective is achieved by using embedded systems to design a vehicle controller that controls speed by sensing the different zones or by defining specific color stripes for different zones and making use of a color sensor to sense the respective zones. The project comprises of the transmitter unit and the receiver unit. The transmitter unit is fitted with a transducer that senses the zone in which the vehicle is present. It then sends this zone status to the receiver unit. The receiver unit is fitted with a microcontroller for the embedded system implementation that takes the zone status from the transmitter and sends out

D. Prasad (✉)
Department of ECE, Sarala Birla University, Ranchi, JH 835103, India
e-mail: deepak.prasad@sbu.ac.in

A. Anand
Department of ECE, Presidency University, Bangalore, Karnataka 560064, India

V. A. Sateesh · S. K. Surshetty · V. Nath
Department of ECE, Birla Institute of Technology, Mesra, Ranchi, JH 835215, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_44

a warning signal to the driver. After a few seconds, if the vehicle speed persists above a set threshold, the receiver unit adjusts the vehicle speed accordingly. Speed control is the need of the hour due to the increased rate of accidents reported every day. Road accidents can be reduced by taking measures like improved road traffic control and management, making the quality of road infrastructure better and making use of safer vehicles. The current safety systems are not sufficient to decrease the number of road crashes. It is necessary to make use of smart speed control in which vehicle speed limit is set by vehicle itself, based on zones, traffic density, etc. In case the color sensor is used, the developed system detects color of specific intensity and accordingly activates/deactivates the speed control system inside the color strips marked on the road. When the vehicle enters roads where the speed is required to be limited, the automobile transducer detects the color and the position of throttle/fuel pump/engine valve which determines vehicle speed control within the established limit is adjusted accordingly. When the system is triggered, the vehicle is controlled within that limiting speed and until the next color strip is crossed. Both the methods have their own advantages and disadvantages. The embedded system approach is futuristic but difficult to implement while the color strip approach is simple and economical. The systems can also be used to alert traffic authorities if rash driving persists, which can be deduced by analyzing the speed control system operation data. The existing devices can only detect rash driving on highways but cannot control it. Also, they involve a lot of human effort, which makes them inaccurate. In this paper, we design a system that not only detects but also regulates dangerous vehicle driving patterns. With the quick development of automobile industry, road safety is becoming a serious problem in recent decades. Major accident-prone areas are sharp turnings, steep slopes, pot-holes, a concealed crossroads, covered safety signs or conditions where the oncoming traffic is obscured. However, it is difficult to find a perfect solution to the problem as a vast majority of vehicles are manned, not automated. Driver weariness is a major reason behind a most accidents. Majority of the traffic accidents occur due to human error and can be avoided by the use of the above-mentioned systems.

Keywords Accident avoidance · Speed limit sensor · Color sensor · Curve detection · Auto braking

1 Introduction

Reckless driving is the root of many road crashes all over the world. A total of about five lakh accidents were recorded in 2001 in India. The traffic populace has risen considerably in India and the existing systems can only monitor the speed of vehicles but cannot control it. The project presented automatically senses the areas like school zone, highway, steep slope or curve zone. The system developed senses traffic signs and sign boards and accordingly assists the driver in controlling the vehicle. The proposed project consists of mainly two units: zone unit and vehicle unit [1]. These special designed zones indicated on road sign boards. For example, near blind turns, the sign board displays “Curve Ahead-Go Slow”. Driver goes at very high speed as usual near the curve causing accidents. As a result, the purpose of displaying warning sign on the roadside boards is defeated. The proposed system automatically senses such traffic signs and accordingly informs the driver. Added features like anti-collision, auto braking with curve detection, Driver fatigue detection system and auto speed limit sensor can be integrated with this system.

This method is very effective in sensing and control of overspeeding. Accidents are not always caused by drunk driving. Even a sober person can drive in a reckless manner. Accidents not only cause a serious damage to the driver but also pose a threat to the surrounding vehicles and people. The current methods of detection by traffic patrols lack efficiency. This is because the patrol officer strength is not sufficient to monitor all the driveways. Moreover, monitoring the roads at night time or in bad weather conditions becomes tough and the proposed technology assists in providing a better alternative to reduce accidents. Additional features like anti-collision, auto braking with curve detection, driver fatigue detection system and auto speed limit sensor can be integrated with this system.

Accidents not only cause a serious damage to the driver but also pose a threat to the surrounding vehicles and people. The current methods of detection by traffic patrols lack efficiency [2]. This is because the patrol officer strength is not sufficient to monitor all the driveways. Moreover, monitoring the roads at night or in poor weather becomes even more difficult. Currently, the police use a potable radar gun and point at the vehicle to note its speed. If the vehicle speed exceeds the speed limit, the nearest traffic check post is alerted to stop the vehicle. This methods proves to be ineffective as a lot of time is wasted between detection and control of the overspeeding vehicle. With the regular increase in road traffic, this method becomes futile and fails to protect the lives of people [3].

Although the suggested model can also be designed using a microcontroller, the drawback of high complex and cost is more than overcome by the efficient way in which the speed of vehicle is controlled [4]. A wide range of sensors can be used, such as inductive loops, video recorders, ultrasonic and microwave transducers and radar-centered sensors. The improvement of the suggested system over conventional cautionary method is that it will help highway traffic police monitor vehicular speed but also regulates it in accordance with the set speed limit. This model can be further developed in future during the introduction of commercial self-driving cars to ensure the safety of the passengers (Fig. 1).

Research is being done to make fully automated roadway system across the world. But will take up to two decades for the full-fledged implementation of these major technological systems on the highways. These systems efficiently control vehicle speeds and movements in order to avoid accidents and utilize sophisticated features built into the roadway and vehicle which are currently not feasible. Rash driving, collision due to obstacles and overspeeding are just some causes of accidents. For the prevention of accidents, government made traffic rules which have to be followed at all times.

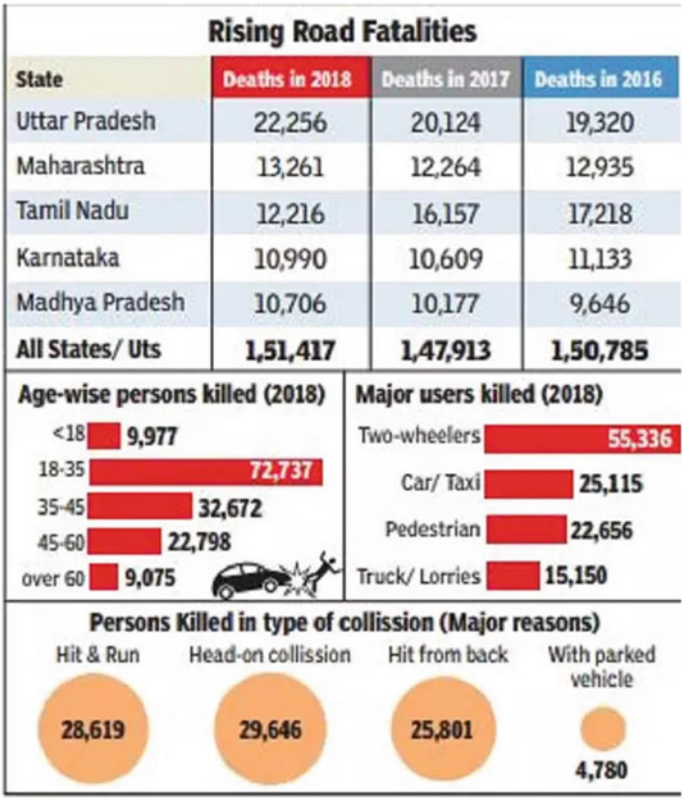


Fig. 1 Road accident deaths by type of vehicle [5]

Statistics show that 1200 deaths and 76,000 injuries annually are caused due to fatigue-related crashes [6]. Developing technologies to detect or prevent sleep drowsiness is a major challenge in designing accident prevention systems. The various ways to detect the tiredness of a person are by measuring the pulse, brain activities, eye movements, etc. This will require several transducers attached to the driver, thereby making driving uncomfortable. The monitoring of the eye movements can sense the driver's weariness to a considerable extent. Fatigue detection involves a sequence of images and observation of eye movements and intermittent patterns. The study of images of the eyes helps perceive if the eyes are shut or not and thus senses tiredness. MATLAB and OpenCV-Python can be used for image processing. MATLAB will be able to capture 2–4 frames per second, while OpenCV-Python will be able to capture 8–10 frames per second, thereby leading to more precise and accurate results [7].

2 Existing System

One of the biggest hazards to safety is on the roads. A mishap occurs every 18 s, and about 20 lakh of them cause temporary or permanent disability. One severe injury could change your life, and it becomes difficult to recover from the post-accident traumatic stress (PATS). Before a new car model can be sold to the public, it must first be tested to ensure it is safe and reliable to drive on the roads. The main components of the vehicle are manufactured by the same automobile companies to ensure that they meet specific performance standards, fuel economy, reliability and comfort, but all of those tests certify the vehicle to be safe presuming that the driver operated the vehicle in a proper manner (Fig. 2).

From the above chart, we can conclude that the major causes of road crashes are carelessness, overtaking and consumption of alcohol among others, which are associated with the driver. The main reason to drive while being intoxicated is that road police cannot test all vehicles circulating on the roads. So we need an effective system to control drunk drivers.

Existing Advance Technological Systems Found in High-End Cars

- ABS (Anti-Locking Braking System):

An anti-lock braking system or anti-skid braking system (ABS) is an automotive safety system that allows motor vehicle's wheels to maintain contact with the road while braking, inhibiting wheels lock, thereby preventing uncontrolled slippage. ABS automatically pumps the normal braking system instead of manual brake pumps to inhibit wheel lock. This helps the driver focus on safety in case of a mishap [9].

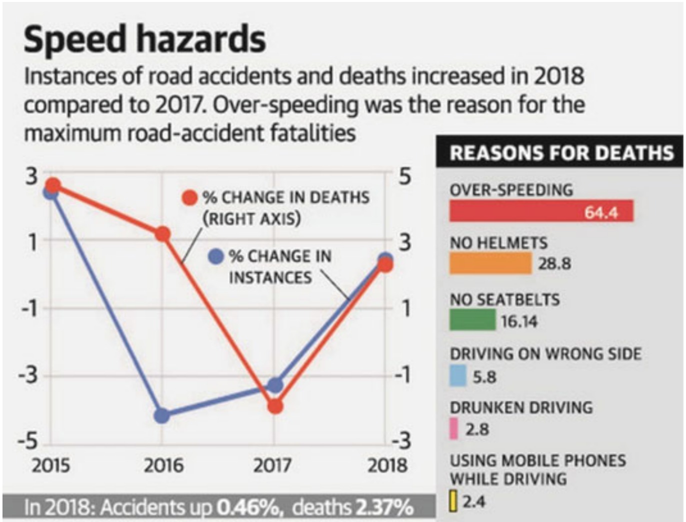


Fig. 2 Causes of road accidents [8]

- EBD (Electronic-brake-force-distribution):

It is an automotive braking technology that automatically changes the applied force to each of the vehicle's brakes, depending on condition of the roads, speed and load together with the ABS system. Generally, the front-end supports more weight and EBD distributes less braking pressure to the rear brakes so that rear brakes do not clog and cause slippage.

- SRS Air Bags (Supplemental Restraint System Air Bags):

An airbag is a vehicle safety equipment. It consists of an elastic casing designed to swell quickly during a crash to prevent occupants from hitting internal objects such as a steering wheel or a window by restraining their movement. Sensors may distribute one or more airbags at different speeds depending on the type and seriousness of impact. The airbag inflates only in case of mediocre to severe accident.

- Immobilizer:

An immobilizer is an electronic device installed in a car to prevent the engine from operating unless the correct key is present.

- Parking Sensors:

Parking sensors are proximity sensors that alert the driver about hidden obstacles while parking the vehicle. Parking sensors are of two types—electromagnetic parking sensors and ultrasonic parking sensors—and are predominantly fitted at the back of the vehicle. They also come with a video display in the rearview mirror to assist in parking the vehicle.

- Cruise Control:

Cruise control is a system that automatically controls the speed of a motor vehicle. The system controls the car accelerator to maintain a constant speed as determined by the driver.

Existing System and its Drawback

After studying the statistical data pertaining to accidents, it can be concluded that majority of the accidents take place due to unsafe driving. Drunk driving is the cause of about one-third of all road accidents, and there is not much the existing safety systems can do to control the driver. This is because none of the existing systems detect driver or passenger misconduct, the driver's skill to judge distances, reaction times, decision-making and vision. They can monitor or at the most assist the driver in avoiding accidents but cannot forcibly take control of the vehicle to drive it to safety (Fig. 3).

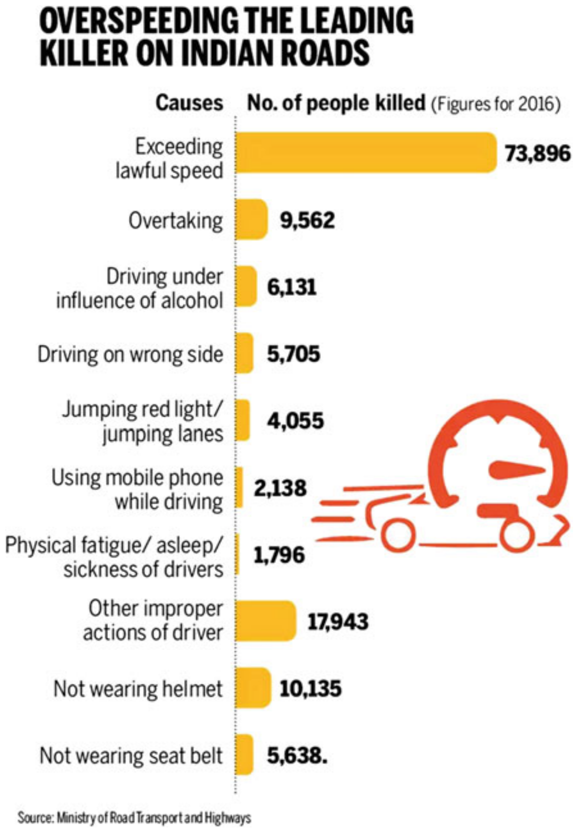


Fig. 3 Reckless driving as a cause for road accident deaths [10]

3 Proposed System

Here, particular zone is detected by system, and accordingly, actions are carried out. The project proposed here consists of zone unit and vehicle unit. Based on traffic sign, zone unit transmits the signal and the appropriate action is done by vehicle unit. In this way, speed of the vehicle is controlled, and accident is prevented. The system includes auto braking, auto speed limit control and curve detection (Figs. 4 and 5).

Establishment of wireless communication by RF transmitter and receiver, overall functioning and controlling by microcontroller, vehicle control system for movement controller, IR trans receiver and sensor for avoidance of collision. RF receiver module and DC motors in microcontroller control whole system. It verifies the data with program in it and according required actions are taken. Transmitted signal controls the vehicle speed, i.e., speed of vehicle decreases by some cut-off and maintained constant until vehicle is in transmitter zone [11].

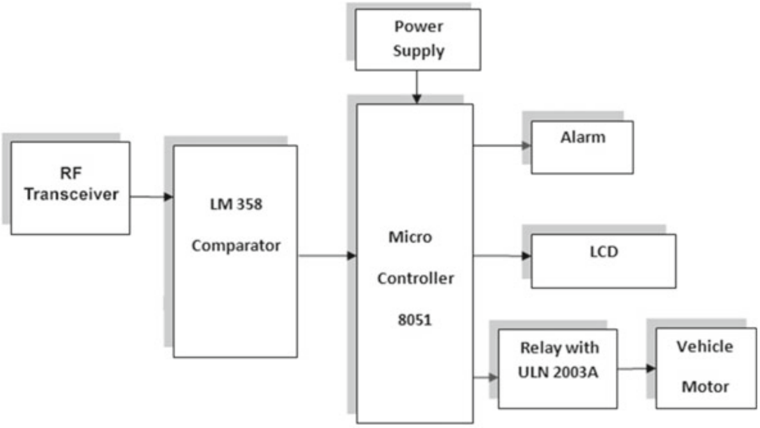


Fig. 4 Speed reduction and curve detection

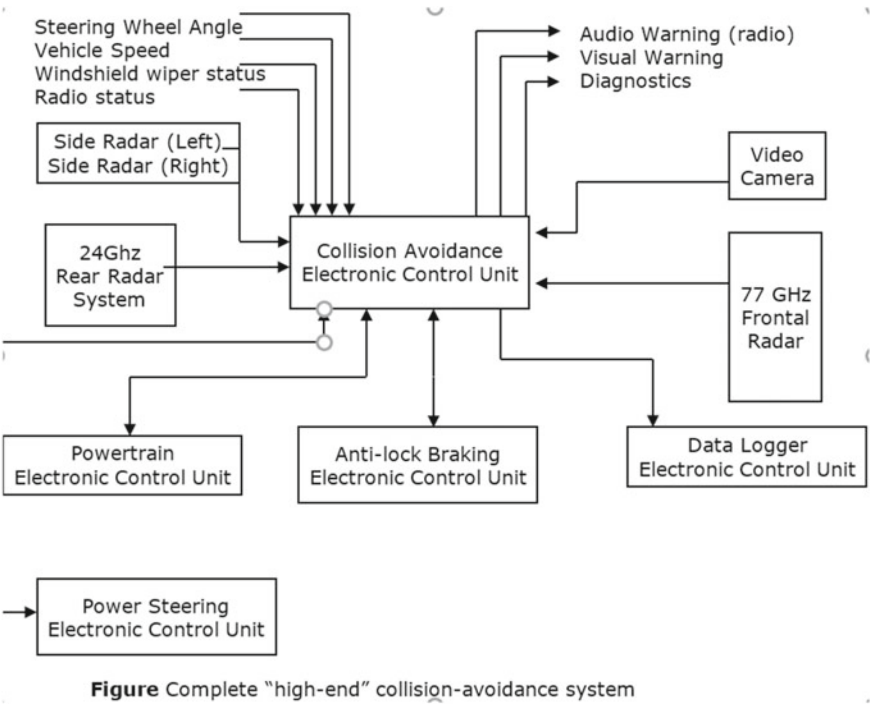


Fig. 5 Collision avoidance

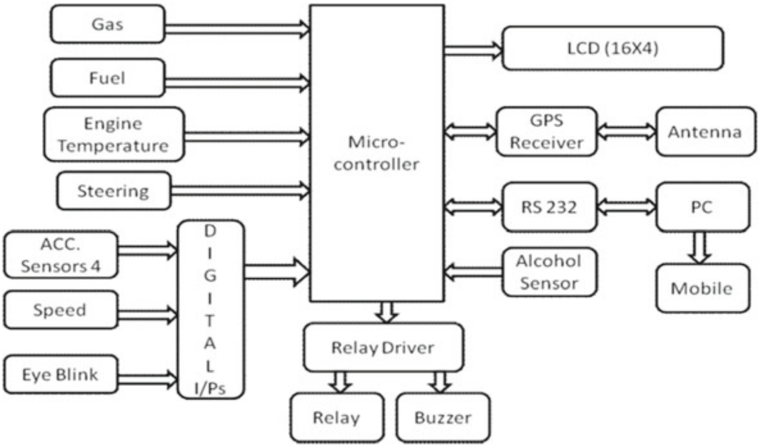


Fig. 6 Proposed system block diagram (CAR unit)

Selection of Color Sensor

Here, color is identified by color sensor and gives corresponding serial output (RBG value). It can identify many color shades giving corresponding RGB value. The detected color is identified as these primary colors (red, blue and green) with accuracy of 8 bit for each primary color. Separation and combination of these colors can be done by RBG values.

Intelligent systems are used in all aspects of the system, CARs are a system that deals with human lives and reality. This includes system behavior, behavioral analysis and coordination. The flashing sensor and the detection of alcohol are important to ensure human life (Fig. 6).

4 Potential Future System

Hardware Description

The microcontroller that accesses data is the heart of the system. The ARM driver is used here. The temperature of the car is measured by the temperature sensor. The signal is transformed into electrical form by conditioning the signal. The conversion from analog to digital is via ADC. Since the ARM processor has an embedded ADC, the signal conditioning unit output is connected to the ARM processor. The LPG gas sensor is connected to the ARM processor by the signal conditioning unit. The temperature indicator helps to keep track of engine temperature. If an overheating occurs, a warning is issued as a beep. RPM controller is used to measure distance and speed. Due to advancement of technology, the new methods are being invented. Mechanical speedometer which is based on digital technique is used in vehicle.

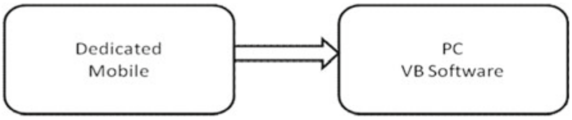


Fig. 7 Block diagram of base unit

The speed is expressed in km/h. The disc rotates through the optical array having a phototransistor and infrared LED. RPM and speed of vehicle are derived from pulses generated by optical assembly. Gas detector keeps track of gas leakage; if any leakage is detected, then alert is given in form of buzzer. If alcohol is found in the driver's breath, he notes it and then turns it off, and then avoids the possibility of an accident. If driver is found to consume alcohol, then it turns ignition off. If decrease in eye blink count is detected by eye blink sensor, then it implies driver is sleepy, and alert is given in form of buzzer. Impact sensors are used to find out place at which impact occurs in case accident is taken place. After complete collection of information, data is stored in internal memory, and by using GSM, message is sent to surveillance unit. This data is now stored in EEPROM. When the switch on car is pushed, then this data is displayed (Fig. 7).

Software Overview

The mainstay of entire project is software overview. Software improvement deals with programs written for microcontroller's interfacing for ADC, keyboard and mobile interface. The best explanation of microcontroller codes can be done by using flowcharts. Next to software and hardware, rudiments of system software programming is incorporated to achieve a given task. When the microcontroller performs calculations or control, its data and display should be easily understood by humans. So, firstly, main program is to divide into subroutines which is further divided to sub-sub routines. After complete assembling of PCB is finished, then software part occurs where C language and assembly language are used. ARM controller performs many operations. The coding is basically divided into modules and sub-modules. Modules deals with initialization of all ports, whereas sub-modules deals with initialization of ADC, LDC, inscription and reading of memory. GPRS helps to locate area of the accident, and message to relatives is sent by GSM.

Rash driving causes much hindrance for people. So, detection of it is necessary which is done by highway speed checker circuit which uses counters, timer, 7-segment display, etc. Working of this circuit is shown in the below diagram. The additional logic module includes timers, NAND ports and ten-year counters. A photodiode (sensor) is a type of photodetector capable of converting the light into current or voltage, depending on the operating mode. Photodiodes use PIN junction rather than PN junction, unlike semiconductors. Photon strikes the diode with energy sufficient to excite electron, thereby creating a hole and mobile electron. The carriers are thrown away from depletion region by built-in-potential if absorption occurs in depletion region of junction or one diffusion length away from it. Therefore, the holes move to the anode, electrons toward the cathode and a photocurrent produced

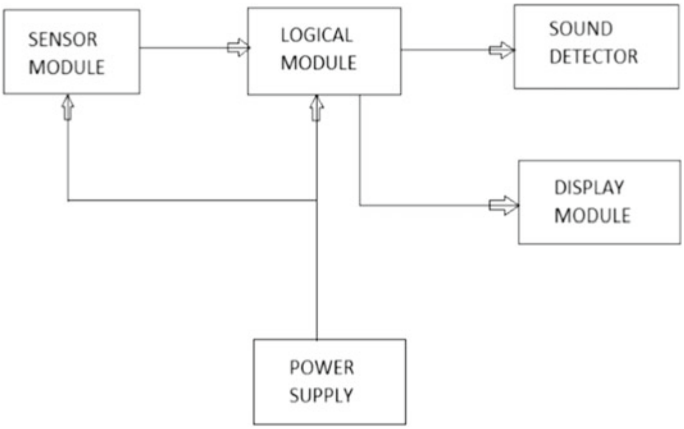


Fig. 8 Block diagram of vehicle speed detector using timer

goes to the timer. Here, we use 555 IC Timer, which is an integrated circuit that implements a diversity of multiple vibrator applications (Fig. 8).

5 Methodology

Rash driving is the main reason for occurring of accidents. The reasons may be poor judgment on part of the driver or poor driving by the driver. After losing the control, it is difficult to drive the vehicle. So speed control of vehicle should be the responsibility of government. Some safety system is needed for prevention of loss of valuable property (which is permanent solution for above problems). Innovative concept is suggested in which speed of vehicle can be controlled automatically at a given limit at particular distance. The proposed concept has potential to control the problems related to traffic. The proposed system here is based on color strips marked on highway.

Operating Principle

Activation or deactivation of speed control is done by color detecting sensor of specific intensity. Accurate speed measurement of the car can be measured with the Hall Effect sensors positioned on the wheels of the vehicle. The longitudinal speed of the vehicle can be controlled with the help of sensor fusion, which applies to information received from subsystems. In practice, the color sensor in vehicle detects the color of the strips on the road which activates the system by sending signals to the MCU whenever a vehicle enters the road with a speed of around 80 km/h. MCU controls the position of the throttle valve which results in engine speed control at the conventional limit. When the system is initiated, the vehicle is controlled at a certain speed or below that limit until it crosses another color strip on the road. Now, in this system, when driver wants to exit from highway to off-road, this speed limit is taken out, the system is deactivated and the vehicle runs under complete control of the driver. RF transmitter units with unique identity code are fitted in monitoring zones for prevention of accidents in smart zone sensing system. Vehicle's number plate should contain RF receiver and respective circuitry. Alert messages sent by zone transmitters will be displayed on dash board.

Transmission of zone code to receiving units is done by radio frequency transmitter. The transmitted zone code signal is received by radio frequency receiver. Based upon the frequency of transmitter, the output pins of transmitter go high. The produced output is fed to buffer. By passing relay for further feeding unit, gain amplification is produced. This gets the status of speed limit and drives two more blocks. If the construction system contains IR transmitters and two receivers, then vehicle can stop automatically if any obstacle comes in front of it. This implies that ignition key of vehicle should be in series with power supply line, and when rider switches on ignition key, vehicle starts.

6 Implementation

MQ-3gas-sensor:

The MQ-3 gas sensor offers good resistance to distribution of gasoline, smoke and vapor. It also has high sensitivity toward alcohol. Percentage of alcohol consumed by the driver can be detected using this sensor; the sensor also has a wide range of applications and is relatively cheaper. Another advantage is that it has a simple drive circuit and has a long lifetime.

MQ-6gas-sensor:

MQ-6 responds to and detects natural gas, LPG, butane and propane. The sensor is of low cost and can be used to detect combustible gases like methane.

GPS Receiver:

The GPS smart receiver consists of 16 channel. The GPS receiver provided is completely enabled to provide excellent sensitivity along with tracking capabilities. Along with that, the receiver provides accurate velocity and time accuracy performances.

Benefits to User:

(1) Very low power consumption
(2) Easy installation
(3) Superior urban canyon performance
(4) Financially feasible and high performance.

This device works ideally when we assume that the maximum speed on the highways is around 60 kmph. To ensure that the power supply output is correct, a multimeter must be used before we start utilizing the system. If the multiplier verifies proper output, switch the supply on. Two photo diodes are utilized here, which are connected with the help of long wires. The speed limit can be selected using the switch SI. As soon as the vehicle crosses the first IR diode, the first photo diode will trigger IC-1. The IC-output is high for the time set to cross the 100 m with the speed that is selected. During this period, the first LED glows. Upon crossing the second IR diode, the output of IC-2 is high, making the second LED glow. The piezo-buzzer sets off an alarm whenever the vehicle crosses the selected speed over given distance. As soon as the first IR diode is intercepted, the counter starts counting, and it comes to a stop after the second IR diode is intercepted. The time taken by the vehicle to cross both the IR diode beams is given by a seven segment display. For example, if the selected speed is 40 kmph, the timer is set at 100 Hz, and if the display count does not exceed 400, it means that the vehicle's speed has exceeded the speed limit. This is accompanied by the buzzer's sound. The circuit is reset for the next passing vehicle. The system is efficient because it utilizes two LDR sensor pairs, which are installed 100 m apart on the highway; each LDR sensor pair on the opposite side of the road.

7 Results

In this project, the speed of a vehicle was controlled whenever it entered a school zone, work zone, curve zone, highway or a U-turn. Accidents on the road can be prevented by detecting the collision distance and speed using an IR sensor. Various parameters of the vehicle such as the engine temperature and steering positions were monitored. As soon as the car was impacted, the impact sensors came into the picture. All of the data is stored in a component known as the μC. If traces of alcohol are detected in the breath of a driver, he/she is immediately warned by the sensor which turns the ignition off. We have designed also an eye blink sensor, and it helps in monitoring the number of times the driver's eye blinks. If the number of eye blinks decreases, it implies that the driver is sleepy, and in that case, a buzzer is operated. In case an accident does occur, impact sensors were utilized to find out where the impact occurred. The information is sent to a surveillance unit or a base through an SMS from the μC component, which uses a GSM modem to accomplish this. On the base side, data such as the engine temperature, eye blink status and level of alcohol is received. The GPS coordinates are updated to the online Google Maps. The reasons of crash or accident can be better understood by the analyst with the help of the PC unit, which has an online VB software. The software displays all of the data and has a feature to show information graphically.

Vehicles on the highway must not exceed the speed limit. However, in spite of all the restrictions, accidents frequently keep occurring because the drivers trust their instinct and reduce/increase speeds according to their will, without paying heed to the rules.

The system was subjected to different conditions. Efficiency at different conditions is calculated. The test is conducted for duration of 60 s to detect the eyes in opened state. The numbers of frames in which eyes are found to be open in this duration are noted. Thus, the efficiency is calculated. The system is tested on different ambient lighting conditions with different users. The system gives accurate result on ideal lighting conditions. The system worked effectively when tested with users of different eyes structure at preferred lighting conditions. The system works efficiently at all lighting conditions except for dark or very little lighting. In low lighting conditions, once the face is detected, eyes are detected with 43% accuracy. The performance is not degraded even when there is slight head movement. But the heavy constant head movement results an accuracy of 47%. The major drawback is when the users wear spectacles which results in multiple reflections on the camera and thus degrades the performance. The system for users with anti-reflective spectacles or spectacles with very little reflections gives an accuracy of 52%. The system does not work for users wearing eyes shades since it is not possible to detect the eyes, resulting in 0% accuracy. The system fails to detect the face if the user is beyond the distance of 40–50 cm. The alarm is triggered instantly without any delay once the drowsiness is detected. It is made sure that alarm is triggered only if both the eyes are closed. Hence, the system works accurately and precisely with 98.5% accuracy at optimum lighting conditions when the driver is at a considerable distance from the camera. The accuracy of the algorithm is checked with different eye lengths to find out the optimum value. It has been found that for eye lengths less than 3, the accuracy will be less, and it will detect eye for only good lighting conditions. The value of eye length with 3 and more will result in the detection of eye status under other conditions like bad lighting, constant head movement and wearing spectacles. It has been found that efficiency gradually increases as per the increase in the value of eye length.

Comparison is carried out based on the metrics like part used to measure drowsiness, technique(s) used for detection, the feature extraction mechanisms, algorithms used for classification and the maximum possible accuracy obtained from the various techniques. It shows the accuracy (i.e., efficiency) of the different techniques utilized to detect the drowsiness status of the driver while they are on the wheel.

Benefits and limitations

Benefits

The eye detection mainly depends on the quality of the camera that is being used. The system gives appropriate results even under low-lighting conditions. Expected results are achieved even when the face is about 40 cm away from the camera. The system works fine even when there are slight head movements, which is a major advantage when the driver is driving the vehicle, considering the road conditions. Capturing of frames form the video and further processing of the frames to detect the status of the eyes are very fast. No delay is incurred to on the alarm when fatigue is detected.

Limitations

The major limitation of all eye detection systems is when the user is wearing spectacles. Although the FDAS provides much resistance to this limitation, the result is found to be inefficient. The system fails to detect the face when the face is very much titled from the vertical axis. Therefore, the eyes cannot be detected. System cannot work under dark or very little lighting conditions.

8 Conclusion

In this project, heavy traffic zone accident prevention techniques are presented, and their feasibility in the real world is discussed. It has been mainly designed in order to avoid accidents and to alert the drivers about the speed limits for safe traveling. These systems can be used in a variety of areas, such as roads with sudden sharp and high curves, thus preventing accidents in Ghats sector and on old bridges. These systems are also used to regulate and control the speed of the vehicle in hospital, school and work zones. Thus, accidents caused by the negligence of drivers can be easily prevented. These systems put emphasis on controlling the vehicle speed automatically without the help of the driver. At the same time, it detects obstacle and stops the vehicle so that accidents can be prevented.

A driver's fatigue is the primary reason for the numerous accidents taking place on highways. We can also see that the rate of traffic accidents keeps alarmingly increasing every year. This project showcases the new fatigue detection algorithms and new systems that use alcohol and eye blink sensors. Fatigue can be detected immediately and regularly alarms the driver and third party. By the means of research presented in this report, we can propose a system to prevent accidents, thus making roads and highways safer places to travel on. Since number of accidents on highways increases at a high rate, it is necessary to keep the speed of the vehicles on highways in check, in order to reduce accidents. This system also minimizes the difficulties faced by the traffic police department, making it relatively easy to control the rash driving on highways. This system can be very useful to the police, as they can sit in control rooms and regulate traffic with ease. Their services will become more accurate, and this could significantly reduce the number of mishaps occurring on highways on a day-to-day basis. This concept can be extended to the future by integrating a camera into the system so that the image of the number plate of offenders can be captured and fed to the control room.

Acknowledgements The authors would like to thanks Vice-Chancellor Prof (Dr.) Indranil Manna, BIT Mesra, Ranchi and Vice-Chancellor Prof. (Dr.) Gopal Pathak, Sarala Birla University, Ranchi for providing the infrastructure and laboratories facility.

References

1. Shiny L, Raja Kumaran A, Vijay S (2015) Vehicle control system with accident prevention by using IR. Int J Innov Res Sci Eng Technol 4(6)
2. Kewate SR, Karmare SV, Sayankar N, Gavhale S (2014) Automatic speed control system by the color sensor for automobiles—an innovative model based approach. Int J Adv Mechan Eng 4(2)
3. Bhumkar SP, Deotare VV, Babar RV (2012) Accident avoidance and detection on highways. Int J Eng Trends Technol 3(2)
4. Jain M, Kumar P, Singh P, Arora CN, Sharma A (2015) Detection of over speeding vehicles on highways. Int J Comput Sci Mob Comput 4(4)
5. Wu Y-S, Abtahi A, Fatigue detection alarm systemz, unpublished
6. Article on 555 Timers—bmet.wikia.com/wiki/555_Timers
7. Article on Drive Wheel Motor Torque Calculations—www2. mae.ufl.edu/designlab/motors
8. Articles on Photodiode, Electronic brakeforce distribution, ABS and EBD systems—wikipedia.org
9. Article on Electronic brakeforce distribution and brake assist—forums.nicoclub.com
10. Bhuta P, Desai K, Keni A (2015) Alchohol detection and vehicle controlling. Int J Eng Trends Appl 2(2)
11. Poddar S, Roy S (2013) Driver's drowsiness detection using eye status to improve the road safety. Int J Innov Res Comput Commun Eng 1(7)

Process Corner Analysis for Recyclic Folded Cascode Operational Amplifiers at 180 nm SCL Technology

Jasdeep Kaur, Maninder Kaur, and Rekha Solanki

Abstract The paper presents a recyclic folded cascode amplifier which is an improved architecture of the conventional folded cascode amplifier. An improvement in DC gain, unity-gain bandwidth, slew rate, and common-mode rejection ratio is observed when compared to the conventional folded cascode amplifier with same power. The recyclic folded cascode architecture delivers gain as 75 dB, slew rate as 10.3 V/μs, unity-gain bandwidth as 8.3 MHz, and common-mode rejection ratio as 123.9 dB in simulation. The power consumption of the recyclic folded cascade is 12 μW with a power voltage of 1.8 V. Further, we designed a bandgap reference circuit using the low-power recyclic folded cascade amplifier. The circuit simulation delivers 633 mV of reference voltage with a coefficient of temperature 27 ppm/°C for the temperature range of 106 °C at 1.8 V supply voltage. The bandgap reference circuit consumed 83 μW of power. The simulation results are obtained in 180 nm SCL technology.

Keywords Operational transconductance amplifier (OTA) · Voltage-controlled current source · Folded cascode (FC) · Recyclic folded cascode (RFC) · Unity-gain bandwidth · Current mirror · Common-mode rejection ratio (CMRR) · Process voltage temperature (PVT)

J. Kaur (✉)
ECE Department, IGDTUW, Kashmere Gate, Delhi, India
e-mail: jasdeep@igdtuw.ac.in

M. Kaur
ECE Department, GTBIT, G-8 Area, Rajouri Garden, New Delhi, India
e-mail: mindersaini82@gmail.com

R. Solanki
Student, ECE Department, IGDTUW, Kashmere Gate, Delhi, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_45

1 Introduction

Operational transconductance amplifier (OTA) is the main building block of various analog circuits such as switched capacitor filter, analog to digital converters, digital to analog converters, and bandgap voltage reference. An OTA is defined as a voltage-controlled current source (VCCS) which produces an output current proportional to differential input voltages with a transconductance of the amplifier as the constant of proportionality. It is similar to an operational amplifier (op-amp) and consumes high power. A single-stage op-amp is an OTA but with limited gain, output impedance, and thereby UGB. Typically, a multistage OTA is designed. A multistage OTA delivers high output resistance and high gain but at the expense of low speed and high-power consumption when compared with a single-stage OTA. The cascading of one or more gain stages in OTA familiarizes a low-frequency pole that limits the bandwidth of the amplifier and increases the risk of instability [1–3]. To further improve the parameters of the multistage, we use compensation techniques which require connection of large capacitance which further increases the power consumption and the area required. For low-power designs, use of single stage cascode structures OTA such as telescopic OTA and folded cascode OTA is preferred.

Folded cascode (FC) is the most commonly used OTA. It has large output signal swing and high gain. In FC OTA, it is possible to short the input and output terminal for calculating closed-loop properties. For achieving higher gain and bandwidth in FC OTA, the bias current is typically increased. A rise in bias current leads to a rise in total power consumption, which is not preferred in low-power devices. In literature [4, 5], many circuit topologies have been implemented to improve the features of an FC OTA. Assaad et al. [6–8] proposes recyclic folded cascode (RFC) structure used to enhance the slew rate, bandwidth, and gain of the FC OTA with same power consumption. The RFC amplifiers use idle tail transistors to recycle the current back into the differential inputs. This recycling of tail current improves gain, bandwidth, and slew rate of the amplifier. Akbari and Prasula [9, 10] discuss the design of an RFC OTA.

Since the dimensions of the device are reduced to nanometers, many differences in manufacturing and process changes can create variability in the physical properties of the manufactured devices. Additionally, the performance of the transistor gets buzz results through voltage and temperature changes while the chip is in operation. Process–voltage–temperature (PVT) differences among the formidable challenges of the transistor scale pathway were set up and, thus, designs with variability have always been a problem in designing analog-integrated circuits [11]. With each new technical generation, various non-intellectual effects begin to have a perceptible effect on the properties of the transistor, and thus on the circuit's performance. In order to assess this, different process corners are designed for the respective transistor models which limit the variations of process to analyze environmental fluctuations, temperature, and supply voltage are varied in a wide range, and the corresponding performance of the circuit is analyzed. In this paper, proposed recyclic folded cascode op-amp has also been tested for different process corner with different temperature values.

A bandgap reference circuit (BGR) [12, 13] generates a DC current or voltage which is constant with temperature, process, and power supply variation. In literature, a number CMOS bandgap reference circuits have been proposed [12, 14–17]. Reference voltage in a BGR circuit is generated by passing positive and negative slope current through a resistor. The positive slope current is referred to as proportional to absolute temperature (PTAT). The current with negative slope is referred to as complementary to absolute temperature (CTAT). The generation of currents in a BGR circuit uses op-amp as a basic block. In this paper, we designed a bandgap reference circuit using low-power recyclic folded cascade op-amp.

This paper presents a low-power RFC OTA and process corner analysis. Sect. 2 describes the folded cascode amplifier; the recyclic folded cascode amplifier is presented in Sect. 3. Sections 3.3 and 4 outline the simulation results and bandgap reference circuit using recyclic folded cascade op-amp, respectively. We conclude our paper in Sect. 5.

2 Folded Cascode OTA

Figure 1 shows a conventional folded cascode amplifier [18–20] with differential inputs and single-ended output. The p-channel MOSFETs (pMOS) MP_0 and MP_1 form the input differential pair. pMOS has lower flicker noise, high-frequency non-dominant poles, lower input common mode level, and lower gain. The n-channel MOSFETs (nMOS) MN_2 and MN_3 form the folded structure. MN_0 and MN_1 are tail current sources. $MP_2 - MP_5$ forms the cascode current mirror load. FC amplifiers can be characterized by its low-frequency gain, slew rate, unity-gain bandwidth (UGB), and common-mode rejection ratio (CMRR). The small-signal transconductance (G_m) of the FC structure is [19]

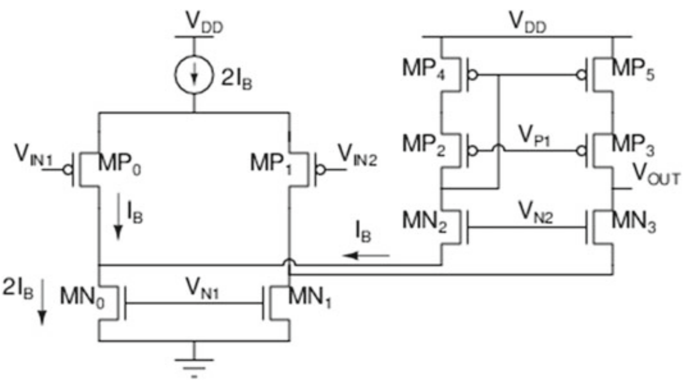


Fig. 1 Schematic of conventional folded cascode amplifier

$$G_m = g_{m_{p0}} \tag{1}$$

$g_{m_{p0}}$, is the transconductance of the MP_0 (or MP_1) the input MOSFETs. The output resistance (R_{out}) is [18–20]

$$R_{\text{out}} = g_{m_{n3}} r_{ds_{n3}} (r_{ds_{n1}} r_{ds_{p1}}) g_{m_{p3}} r_{ds_{p3}} r_{ds_{p5}} \tag{2}$$

$g_{m_{n3}}$, $g_{m_{p3}}$ is the transconductance of MN_3 and MP_3. $r_{ds_{n3}}$, $r_{ds_{n1}}$, $r_{ds_{p1}}$, $r_{ds_{p3}}$, $r_{ds_{p5}}$ are the drain-source resistance of MN_3, MN_1, MP_1, MP_3, and MP_5.

The small-signal differential voltage gain is [19, 20]

$$A_v = \frac{v_{\text{out}}}{v_{\text{in}}} = g_{m1} R_{\text{out}} \tag{3}$$

For a load capacitance C_L, the dominant pole for the amplifier lies at $\frac{1}{R_{\text{out}} C_L}$. Unity-gain bandwidth (UGB) of the amplifier is [19]

$$\text{UGB} = \frac{A_v}{R_{\text{out}} C_L} \tag{4}$$

The slew rate of the FC structure when a large signal is applied at the differential inputs is given by [19]

$$\text{SlewRate} = \frac{2I_B}{C_L} \tag{5}$$

$2I_B$, is the current which either charges or discharges C_L during low-to-high or high-to-low transition at the inputs [21].

The folded structure in FC op-amp gave advantage of shorting input and output terminals and increase in the output swing as there is a smaller number of MOSFETs stacked. MN_0 and MN_1 conduct maximum current and have the largest transconductance. The use of MN0 and MN_1 is limited to current sinking devices only. To improve the use of MN_0, MN_1 the concept of recycling current is introduced. This also improves the performance of the amplifier.

3 Recyclic Folded Cascode OTA

3.1 Small-Signal analysis of Recyclic Folded Cascode OTA

The basic idea of an RFC op-amp is to create multiple paths for the flow of small-signal current, as shown in Fig. 2. Division of input MOSFETs into two reduces the

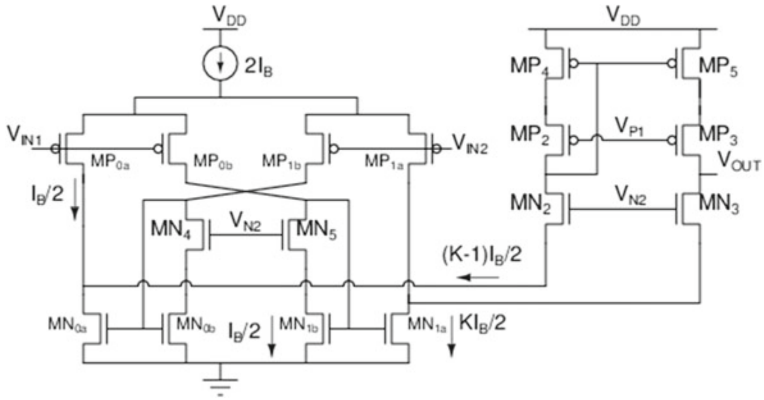


Fig. 2 Schematic of recyclic folded cascode amplifier

branch current to $\frac{I_B}{2}$ instead of I_B. Further, each idle tail transistor (MN_0 and MN_1) is divided in the ratio of 1:K. The value of K is chosen to be 3 for the power consumption to be the same in both FC and RFC op-amp. The cross-over connection also ensures that the currents at sources of MN_2 and MN_3 are added in the same phase during small-signal analysis. The current mirror load, MN_2, and MN_3 carry $(K-1)\frac{I_B}{2}$ current. MN_4 and MN_5 are added to maintain drain voltages of MN_{0a}:MN_{0b} and MN_{1a}:MN_{1b} same for matching. The small-signal analysis short-circuit current and transconductance is given by Eqs. (6) and (7) [18]

$$i_{sc} = g_{m_{p0a}}(1+K)V_{in} \tag{6}$$

$$G_m = g_{m_{p0a}}(1+K) \tag{7}$$

$g_{m_{p0a}}$ is the transconductance of MP_{0a}, and it is half of the transconductance of MP_0. The transconductance of RFC (Eq. (7)) op-amp is twice the transconductance of FC (Eq. 1) op-amp because of an additional $(1+K)$ factor. The small-signal output resistance of the circuit is [18]

$$R_{out} = g_{m_{n3}} r_{ds_{n3}} (r_{ds_{n1a}} r_{ds_{p1a}}) g_{m_{p3}} r_{ds_{p3}} r_{ds_{p5}} \tag{8}$$

$r_{ds_{n1a}}$ and $r_{ds_{p1a}}$ are increased since MN_{1a} and MP_{1a} are carrying $K\frac{I_B}{2}$ and $\frac{I_B}{2}$ current. This increases the output resistance of RFC op-amp in comparison to FC op-amp. The increase in small-signal transconductance and output resistance leads to increase in the low-frequency gain. The gain is expected to increase by 8–10 dB. The dominant pole for the RFC op-amp is also at $\frac{1}{R_{out}C_L}$. Increase in R_{out} shifted the dominant pole to the left. The overall increase in the low-frequency gain increases UGB and common-mode rejection ratio (CMRR) of the RFC op-amp in comparison with FC op-amp [19, 21, 22].

3.2 Large-Signal Analysis of Recyclic Folded Cascode OTA

During low-to-high transition at VIN2 and high-to-low transition at VIN1, MP1a and MP1b are OFF, and MP0a and MP0b are ON. Consequently, MN0b will not set to the desired bias, and hence, no current flows through MN0a. The source potential of MN3 rises, and MN3 is switched off. MN1b gate sets at the desired potential, and $2I_B$ current flows through MN1b. $2KI_B$ current flows through MN1a, cascade current mirror load (MP2–MP5) and C_L. During high-to-low transition at VIN2 and low-to-high transition at VIN1, MP1a and MP1b are ON, and MP0a and MP0b are OFF. MN0b sets at the desired potential and $2I_B$ current flows through MN0b. $2KI_B$ current flows through MN0a, cascode current mirror load (MP2–MP5) and C_L[15]. The slew rate of the RFC OTA is given by Eq. (9). The slew rate of RFC OTA is K times the slew rate of FC OTA (Eq. 5).

$$\text{SlewRate} = \frac{2KI_B}{C_L} \tag{9}$$

3.3 Simulation Results of FC and RFC OTA

The low-power FC and RFC OTA circuits are designed using SCL 180 nm technology with 1.8 V supply, 0.5 pF of load capacitance (C_L), and 3 µA biasing current in Cadence Virtuoso tool. Table 1 shows the sizes of all devices in FC and RFC amplifiers. Long-length devices have been chosen for reducing short-channel effect, high

Table 1 Device sizes (µm) of FC and RFC OTA

Devices	FC OTA	RFC OTA
MP_0/MP_1	6.3/3	–
$MP_{0a}/MP_{0b}/MP_{1a}/MP_{1b}$	–	3.2/3
MN_0/MN_1	2/3	–
MN_{0a}/MN_{1a}	–	1.5/3
MN_{0b}/MN_{1b}	–	0.5/3
MN_2/MN_3	2.2/3	2.2/3
MP_2/MP_3	3.3/3	3.3/3
MP_4/MP_5	3.3/3	3.3/3
MN_4/MN_5	–	1.1/3

output resistance, high small-signal gain, and improved power supply rejection ratio (PSRR). FC amplifier device sizes have been calculated using standard design equations. RFC amplifier's device sizes have been modified accordingly. The bias current and voltages for the amplifier are generated using conventional beta-multiplier circuit or constant-g_m circuit.

Figure 3 shows the open-loop gain of FC OTA is 63.5 dB, and the RFC OTA is 75 dB. An improvement of 11.5 dB is observed in the open-loop gain of the RFC. UGB is also observed to be 4.4 MHz for FC and 8.3 MHz for RFC. UGB of RFC is almost double of FC. Enhancement in the gain by 8–10 DB in RFC op-amp over FC op-amp can be observed due to increase in output resistance and transconductance (almost double). The phase margin of FC obtained is 72.99°. The phase margin is observed to be 45.73° of RFC. It is degraded by 28.26° due to the introduction of extra pole-zero pairs at the gate terminals of the current mirrors.

To measure the slew rate, a step input of 850 mV is applied to unity-gain feedback circuit with capacitive load. Transient response is shown in Fig. 4. The slew rate of FC OTA is 4.1 V/μs and of RFC is 10.3 V/μs. The simulation results show more than twice improvement in the slew rate because bias current in RFC op-amp is mirrored by factor k. The common-mode rejection ratio (CMRR) for FC and RFC is observed to be 104.9 and 123.86 dB as shown in Fig. 5. An improvement of 18.96 dB is obtained in CMRR of RFC compared to FC due to increase in differential gain causing dominant pole shifted to left which leads to the reduction of static settling error. Table 2 shows the simulation-based performance comparison of FC and RFC op-amps [23]. The supply voltage and load capacitance were the same. The improvements in the result are obtained by recycling the large current which is conducted by the MN_0 and MN_1 of FC.

The RFC OTA has also been simulated for various process corners [12, 24] such as typical nMOS typical pMOS (TT), fast nMOS fast pMOS (FF), fast nMOS slow

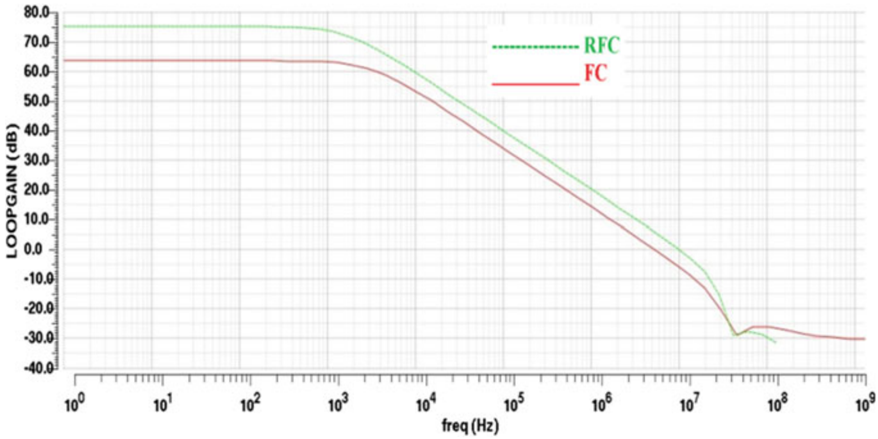


Fig. 3 Simulated open-loop gain plot of FC and RFC OTA

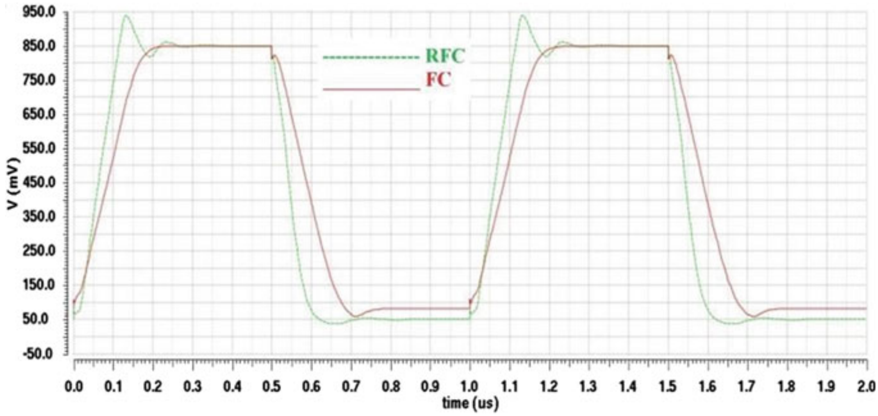


Fig. 4 Simulated transient response of FC and RFC OTA

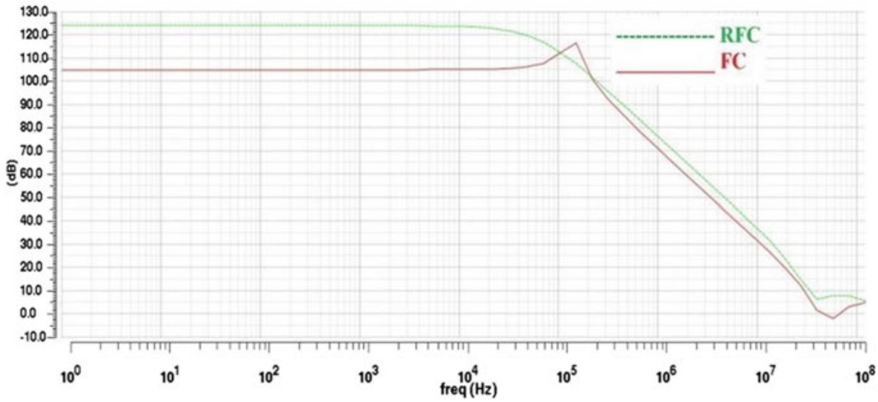


Fig. 5 Simulated CMRR plot for FC and RFC OTA

Table 2 Simulation-based performance comparison of FC and RFC

Parameters	FC OTA	RFC OTA
Bias current (μA)	3	3
Supply (V)	1.8	1.8
OL GAIN (dB)	63.5	75
UGB (MHz)	4.4	8.3
Phase margin (°)	72.99	45.7
Capacitive load (pF)	0.5	0.5
Slew rate (V/μs)	4.1	10.3
3 dB Frequency (kHz)	2.8	1.5
CMRR (dB)	104.9	123.9
Power (μW)	13	12

pMOS (FS), slow nMOS fast pMOS (SF), slow nMOS- slow pMOS (SS), with temperature value set to −40, 27 and 60 °C. Tables 3, 4, and 5 present the results obtained after simulating the RFC OTA at 1.8 V supply voltage across different process corners and temperature.

As temperature increases, the amplifier output resistance increases and transconductance decreases. This leads to decrease in open-loop gain (OL GAIN) and CMRR. Equation (4) shows a reduction in unity-gain bandwidth with reduction in gain with increase in resistance. 3 dB frequency is observed to increase with temperature increase.

Table 3 RFC OTA performance at −40 °C across process corners

Parameters	TT	FF	SS	FS	SF
OL GAIN (dB)	84.4	84	85	84	86
PM (°)	44.5	43.4	45	45	43.6
UGB (MHz)	9.5	10.4	9.4	9.4	9.9
3 dB (kHz)	0.58	0.7	0.54	0.65	0.56
Power (μW)	12.58	12.67	12.52	12.7	12.46
CMRR (dB)	140.9	129.5	142.3	140.6	138.3

Table 4 RFC OTA performance at 27 °C across process corners

Parameters	TT	FF	SS	FS	SF
OL GAIN (dB)	75	73	77	73.5	76
PM (°)	45.7	43	45.5	45.4	43.5
UGB (MHz)	8.3	9	8	8.14	8.6
3 dB (kHz)	1.6	2	1.2	1.8	1.5
Power (μW)	12	12	12	12	12
CMRR (dB)	123.9	119.4	127.4	123.6	123.5

Table 5 RFC OTA performance at 60 °C across process corners

Parameters	TT	FF	SS	FS	SF
OL Gain (dB)	69	67	72	67	71
PM (°)	44	43	45.7	45.8	43.7
UGB (MHz)	7.7	8.23	7.5	7.6	8
3 dB (kHz)	2.9	3.9	2	3.4	2.5
Power (μW)	11.8	11.9	11.8	11.9	11.7
CMRR (dB)	117.7	113.1	121.5	115.4	117.7

4 Bandgap Reference Circuit

Conventional bandgap reference circuits used to generate 1.25 V equivalent to the bandgap of Silicon. In [14–16], a bandgap reference is proposed which generates DC voltage below 1 V. Figure 6 shows the circuit of bandgap reference. The op-amp is implemented using pMOS differential input-based recyclic folded cascade op-amp. PTAT current is generated with the combination of two diodes (D_0 and D_1) and resistor R_P. CTAT current is generated by addition of R_C in shunt with PTAT circuit at the inputs of op-amp. The reference voltage (V_{REF}) is generated by the sum of PTAT and CTAT currents through the resistor, R_N. Table 6 shows the sizes of all the devices used in Fig. 6. BGR device sizes have been calculated using standard design equations [21].

Temperature coefficient is an important parameter for bandgap reference design [12, 17]. It is defined as variation of voltage over a temperature range per reference voltage and is given by Eq. (10). Its units are ppm/°C.

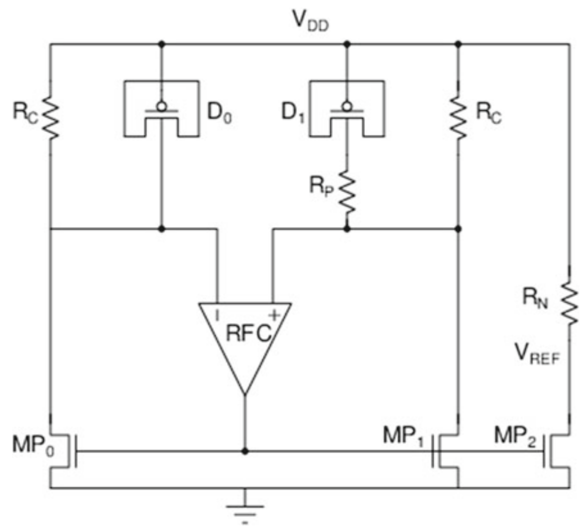


Fig. 6 Schematic of bandgap reference circuit [18]

Table 6 Device sizes of BGR

Devices	Sizes
$MP_0/MP_1/MP_2$	4.4/2 (μm/μm)
R_P	20 KΩ
R_C	126 KΩ
R_N	122 KΩ
D_0	0.42/0.18 (μm/μm)
D_1	8 × (0.42/0.18) (μm/μm)

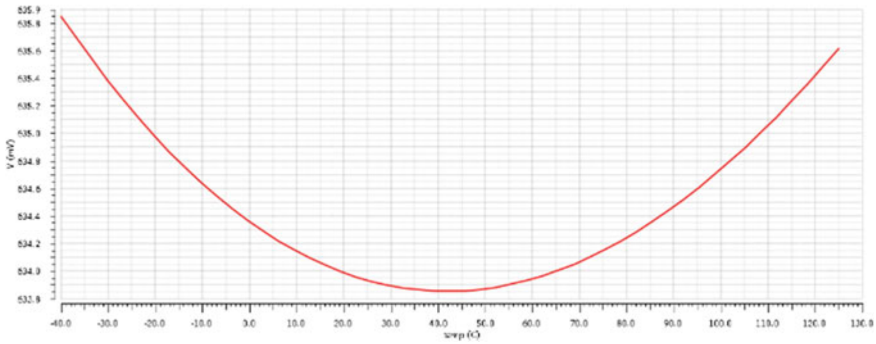


Fig. 7 Simulated response of BGR showing variation in reference voltage with temperature

$$\text{Temp Coeff.} = \frac{(V_1 - V_2)}{(T_1 - T_2)V_{\text{REF}}} \tag{10}$$

4.1 Simulation Results and Performance Comparison

Figure 7 shows the simulation response of the bandgap reference circuit to temperature variation from −40 to 125 °C. 0.18 µm standard SCL process was used for simulating the design. 633 mV reference voltage was generated with 1.8 V supply voltage. The temperature coefficient calculated using Eq. (10) is 27 ppm/°C over a temperature range of 106 °C.

The circuit consumed 83 µW of power. Table 7 shows a simulation-based comparison of proposed bandgap reference circuit with existing BGR designs.

Table 7 Simulation-based performance comparison of this work with existing BGR designs

Parameters	[24]	[13]	This work
Supply voltage (V)	1	2.5	1.8
Ref vol (mV)	632.9	617.7	633
Temp Coef. (ppm/°C)	19.8	3.9	27
Temp range (°C)	−20 to 100	−15 to 150	−40 to 66
Power (µW)	33.5	–	83
Process (µm)	0.18	0.35	0.18 (SCL)

5 Conclusion

The paper presents RFC OTA which has an improved performance over FC OTA. The low-power RFC OTA has 75 dB open-loop gain, 10.3 V/μs (more than double) slew rate, 8.3 MHz (double) unity-gain bandwidth. and double transconductance without using extra power. The phase margin and power of RFC was 45.7° and 12 μW. The performance of RFC OTA was verified by simulation across process corners, at different temperatures. The results of the proposed RFC OTA is simulated in 180 nm standard SCL technology using Cadence virtuoso tool at 1.8 V supply voltage. A bandgap reference circuit was designed using an RFC OTA to generate 633 mV. BGR circuit has 27 ppm/°C as its temperature coefficient and a temperature range of 106 °C. The circuit consumed 83 μW of power with 1.8 V of supply voltage.

References

1. Silva D, Rogerio D, Oki N (2015) An amplifier and capacitor sharing technique using recycling folded cascode operational amplifier with applications in MDAC of CMOS pipelined ADC. In: 2015 conference on design of circuits and integrated systems (DCIS), pp 1–4. IEEE
2. Nageshwarrao D, Suresh Kumar K, Rajasree Rao Y, Jyothi G (2013) Implementation and simulation of CMOS two stage operational amplifier. Int J Adv Eng Technol 5:162–167
3. Yu C-M, Lin Z-M, Chen J-D (2004) A low voltage high unity-gain bandwidth CMOS op-amp. In: Proceedings. The 2004 IEEE Asia-pacific conference on circuits and systems 2004, vol 1, pp 45–48
4. Roh J (2006) High-gain class-AB OTA with low quiescent current. Analog Integr Circ Sig Process 47(2):225–228
5. Nakamura K, Richard Carley L (1992) An enhanced fully differential folded-cascode op amp. IEEE J Solid-State Circ 27(4):563–568
6. Assaad RS, Silva-Martinez J (2009) The recycling folded cascode: A general enhancement of the folded cascode amplifier. IEEE J Solid-State Circuits 44(9):2535–2542
7. Assaad R, Silva-Martinez J (2007) Enhancing general performance of folded cascode amplifier by recycling current. Electron Lett 43(23):1243
8. Patra P, Kumaravel S, Venkatramani B (2012) Design of low power enhanced fully differential recyclic folded cascode OTA. In: Das VV, Stephen J (eds) Advances in communication, network, and computing. CNC 2012. Lecture notes of the institute for computer sciences, social informatics and telecommunications engineering, vol 108. Springer, Berlin
9. Akbari M (2015) Single-stage fully recycling folded cascode OTA for switched-capacitor circuits. Electron Lett 51(13):977–979
10. Prasula TV, Meganathan D (2017) Design and simulation of low power, high gain and high bandwidth recycling folded cascode OTA. In: 2017 fourth international conference on signal processing, communication and networking (ICSCN), pp 1–6. IEEE
11. Bendre VS, Kureshi AK (2019) Design and PVT analysis of robust, high swing folded cascode operational amplifier. Int J Eng Adv Technol (IJEAT) 9(2): 114–118
12. Lee EKF (2014) A low voltage auto-reconfigured power-on-reset/bandgap reference circuit. In: 2014 IEEE international symposium on circuits and systems (ISCAS), pp 918–921
13. Andreou CM, Koudounas S, Georgiou J (2012) A novel wide-temperature-range, 3.9 ppm/°C CMOS bandgap reference circuit. IEEE J Solid State Circuits 47(2):574
14. Yu C, Siek L (2015) An area-efficient current-mode bandgap reference with intrinsic robust start-up behavior. IEEE Trans Circuits Syst II Express Briefs 62(10):937–941

15. Francisco KR, Hora JA (2012) Very low bandgap voltage reference with high PSRR enhancement stage implemented in 90 nm CMOS process technology for LDO application. In: Proceedings of IEEE ICEDSA, pp 216–220
16. Cho G (2009) Assessment of CNTFET based circuit performance and robustness to PVT variations. In: 52nd IEEE international midwest symposium on circuits and systems, 08/2009
17. Banba H, Shiga H, Umezawa A, Miyaba T, Tanzawa T, Atsumi S, Sakui K (1999) A CMOS bandgap reference circuit with sub-1-V operation. IEEE J Solid-State Circuits 34(5):670–767
18. Baker RJ (2008) CMOS: circuit design, layout, and simulation. Wiley, vol 1
19. Razavi B (2001) Design of analog CMOS integrated circuit. McGraw-Hill
20. Allen PE, Holberg DR (2002) CMOS analog integrated circuit design
21. Pourabdollah M (2016) A new gain-enhanced and slew-rate-enhanced folded-cascode op amp. Analog Integr Circ Sig Process 88:43
22. Nayak A, Bonthala S, Uppoor Y, Bhat MS (2019) Design of high gain operational transconductance amplifiers in 180 nm CMOS technology. In: 2019 IEEE international conference on distributed computing VLSI electrical circuits and robotics (DISCOVER), pp 1–4
23. Nischal S, Kaur J (2019) Study of a self biased high swing cascode current mirror based folded cascode operational amplifier. In: 2019 4th international conference on information systems and computer networks (ISCON), pp 427–431
24. Lee EKF (2013) A low voltage low output impedance CMOS bandgap voltage reference. In: IEEE international symposium on circuits and systems (ISCAS), pp 1480–1483. IEEE

Performance Analysis of Array Multiplier Using Reversible Logic

K. N. Hemalatha and B. G. Sangeetha

Abstract Minimization of power and area miniaturization is the most required criteria in today's world. Reversible logic is one of the evolving ways out for low-power computing. Multiplication is the most fundamental operation since the existence of multiply and accumulate unit in most of the computing systems. Paper provides an analysis of latest advancements in designs of unsigned parallel multiplier using reversible logic.

Keywords Reversible logic · Array multiplier · Garbage output · Constant input

1 Introduction

Dissipation of power is a most vital factor in designing circuit. Power dissipation can be divided into dynamic and static. Loss of static power is due to non-ideal components existing in the circuit. One more part is loss of information. Digital circuits which are irreversible consume power for each bit of information lost. Landauer confirmed in 1960s, each lost bit is a cause for information lost. Research reveals that usages of irreversible logic gates dissipate power. Energy in terms of power dissipated on each bit of irreversible operation is given by kT ln2, where T—absolute temperature and k—Boltzmann's constant [1]. Bennett presented that in order to avoid power loss, it is essential that all operations need to be carried using reversible logic [2]. Logic circuits must be made from reversible gates to reduce power dissipation.

Multiplication is one of the most basic operations that are performed on arithmetic circuits and to realize DSP functions. Logic circuits are realized by using different multiplier architectures such as serial multiplier, parallel multiplier, Wallace tree, booth multiplier, Dadda multiplier, and Baugh–Wooley multiplier. Signal processing

K. N. Hemalatha (✉)
Department of ECE, Dr. AIT, Bengaluru, India
e-mail: knhemalatha@gmail.com

B. G. Sangeetha
Department of ECE, RNSIT, Bengaluru, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_46

is the basic operation carried on almost all applications in engineering. Speed performance of additions and multiplications is the requirement.

Well-organized multiplier must have high speed and occupy a smaller area and low-power consumption. Different methods are used to apply internally and externally in the previous, to design a cost-effective multiplier. External practices are associated with the variation of input bit structures, and internal method takes care of circuit, architecture, and system technology [3].

The key objective of the paper is to study the designs of unsigned array multiplier and analyze and comparatively study the design parameters in reversible logic-based multipliers designs. A multiplier is a circuit which can multiply two N bit numbers. Multipliers are basically classified into two methods: serial and parallel multiplier. Array multipliers come under parallel multiplier, and it is the most important subgroup of parallel multiplier because of its regular structure. Different designs of multipliers are realized based on different requirements and applications for which is used. When low-cost efficient design is required, multipliers using serial methods are better. On the other fact, parallel multipliers are used when multiplication with high speed is required.

The paper is planned as explained: Sect. 2 introduces preliminary of the reversible logic; Section 3 is presented with fundamentals of reversible gates; Sect. 4 presents the study on reversible multiplier and comparison of performance parameters; conclusion is provided in Sect. 5.

2 Reversible Logic

No information about the computational states can ever be lost in reversible computation. Thus, previous phase can be retraced by computing back or uncomputing the results. This process in reversible logic is called as logical reversibility. A logic system is reversible when the inputs can be uniquely retraced from the output. Reversible circuits have one-to-one mapping between input vector and output vector, i.e., not only the outputs can be retraced from inputs uniquely but the inputs can be retrieved from the outputs [4].

Landauer [1] mentioned the amount of energy dissipation for each bit to erase is at least kTln2 (where *k*—Boltzmann constant and *T*—room temperature), and during each operation, the in-between bits are used to compute the last result. Removal of bits is the main reason for power dissipation in irreversible circuits.

Bennett [2] proved that power dissipation in every device can be made negligible or ideally zero if the computations are carried through reversible logic. Operations that are done by using irreversible logic can be done with equal efficiency on the circuits using reversible logic.

Reversible logic gate block diagram is shown in Fig. 1. Reversible logic gates are having equivalent number of input and output lines and have one-to-one mapping between input and output. Reversible logic circuit should have the following [3–5]:

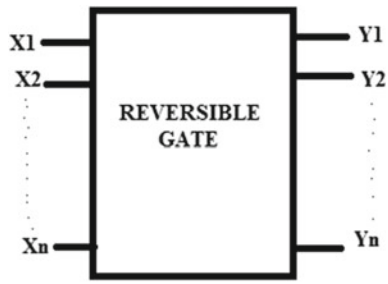


Fig. 1 Reversible gate block diagram

- Reversible logic gates that are used should be less in number.
- Less number of unused outputs, i.e., garbage outputs.
- Less number of inputs that are held at constant inputs.
- Logic circuit should have low-quantum cost.

Following are the few basic terms used in reversible logic [6–9].

Garbage output (GO): Outputs that are not used to compute further design are called garbage output.

Constant Input (CI): To achieve reversibility, inputs that are kept constant at logic 0 or logic 1 are termed as constant input.

Gate count (GC): Total number of reversible gates that are used to realize the reversible design using reversible gates.

Quantum cost: Amount of elementary quantum gates essential for the construction of reversible function.

3 Reversible Gate

Figure 2 shows a Feyman gate which is a reversible 2 × 2 gate with elementary cost, i.e., quantum cost of 1 [10]. In reversible logic, fan-out is prohibited. Feyman gate is normally used to avoid fan-out and also to produce the complement of an input.

Figure 3 represents Toffoli gate having three inputs lines, i.e., $I(A, B, C)$ and three outputs lines, i.e., $O(P, Q, R)$ [11]. First two input lines act as control inputs, and the last input line acts as target input. Without any change, first two inputs are directly transferred to the corresponding outputs and toggle the last output if both first two

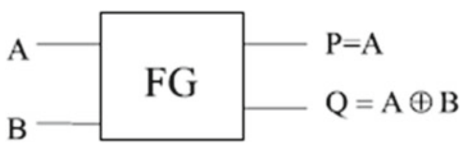


Fig. 2 Feyman gate

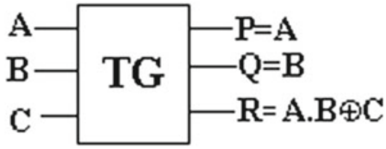


Fig. 3 Toffoli gate

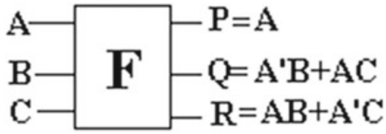


Fig. 4 Fredkin gate

input lines are held at logic 1. Toffoli gate is having the quantum cost of 5 and can be used as universal gate [6–8].

Figure 4 represents 3 × 3 reversible Fredkin gate [12]. Fredkin gate is having the quantum cost of 5. The two output bits are swapped if the control input A is 1, and hence, it is called as swap gate.

A new reversible gate is introduced by Peres [13] known as Peres gate. Peres gate has three inputs and three outputs as shown in Fig. 5. Peres gate cannot be used as universal gate but commonly used in most of the designs in reversible circuit, since it is having quantum cost less with respect to the other reversible gate which are used as universal. Peres gate has quantum of 4 and is also known as new Toffoli gate because it combines Feyman gate and Toffoli gate. Peres gate has equal number of transformation formed by a Toffoli gate and Feyman gate.

HNG gate is 4 × 4 reversible logic gate as shown in Fig. 6. Most important operation of the HNG gate is that it can be independently realized as a reversible full adder [14]. Most of the designs use HNG gate as full adder circuit in reversible logic.

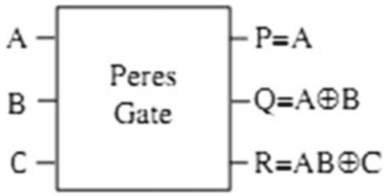


Fig. 5 Peres gate

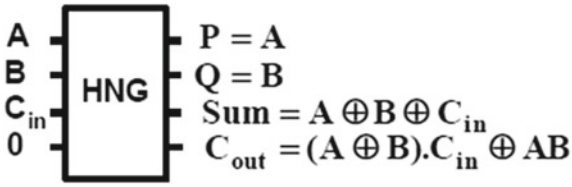


Fig. 6 Haghparast Navi gate

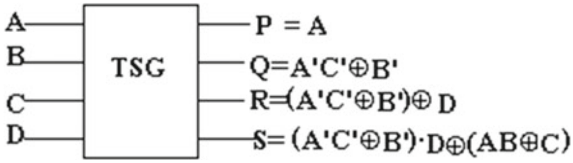


Fig. 7 TSG gate

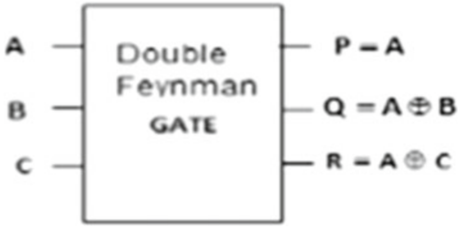


Fig. 8 Double Feyman gate

Ranganatha and Thapliyal [15] designed a 3 × 3 logic gate which is reversible as shown in Fig. 7 known as TSG gate. TSG gate is used to realize all Boolean functions in reversible logic. Main important operation of this gate is it can work independently as a reversible full adder.

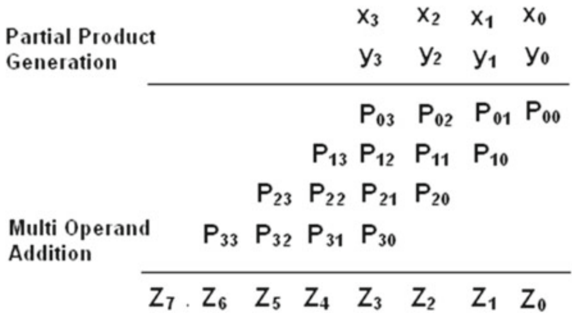
Double Feynman gate (F2G) is a 3 × 3 reversible gate with parity preserving, and quantum cost of the gate is 2 as shown in Fig. 8. One of the important functions of this gate is that it can be used for fan-out generation in circuit designed. Parity of the input bits is equal to the parity of the output bits in parity-preserving gates [10].

4 Reversible Unsigned Array Multiplier

A noble multiplier design is to physically offer a compact, low-power consumption, and high-speed operation. Multiplier is the major part of arithmetic processing unit and is tremendously in high need for high-speed and low-power consumption. In broad, the choice of multipliers is based on issues such as speed, occupancy, latency, and design complexity [11]. Optimized multiplier architecture is in need to reduce power consumption. It is better to shrink the number of operations carried out, hence reducing a dynamic power which is a main part of overall power dissipation [12]. There are number of methods to accomplish multiplication like serial multiplication and parallel multiplication.

The well-known regular structures in multipliers are array multipliers. Design of array multiplier is done in two units: partial product generation and multi-operand addition (MOA) [16–21]. Each partial product bit is produced by multiplying one bit from multiplier and one bit from multiplicand. Partial product produced is moved according to the bits produced and then finally added. Addition operation can be done by using adder of any design like ripple carry adder, carry look ahead adder, carry save adder, and carry select adder [23–26].

Fig. 9 Operation of the 4 × 4 multiplier



Computation of partial products in array multiplier is in the form of radix-2 method, i.e., by considering one bit of the multiplier at a time. Figure 9 shows the operation of 4 × 4 reversible multiplier [27–34]. It consists of *X XY* number of bits of partial products, i.e., 4 × 4 multiplier will have 16 partial product bits, though it may be extended to any order as $n * n$ multiplier.

5 Existing Array Multiplier

Himanshu Thapliyal et al. in 2006 proposed a MXN reversible multiplier design, an advance in array multiplier design [15]. Partial products are produced in parallel with delay of d using Fredkin gates, and next the addition operations can be reduced by log2N by using adders which are paralleled designed from reversible TSG gates. Parallel adder is used to add every two adjacent partial products that are generated. Many improved parallel adder designs are proposed [35]. Addition of partial products which are adjacent will produce the first computation level with N/2 partial product sums. These intermediate sums are added again in the aforesaid method to form a second level of computation with N/4 partial product sums. The final result of the product will have log2N level.

Using Fredkin gates, partial products are produced for 4 × 4 reversible multiplier in parallel as shown in Fig. 10. Each two bits of partial products produced is added by 4-bit reversible parallel adder generating the first level of addition which has sum and carry of partial sums. These two bits of partial sums which are sum and carry are again fed to the next level of four-bit parallel reversible adder, resulting in the realization of final product. Proposed design multiplier efficiency significantly depends on the kind of reversible parallel adders used in multi-addition process.

Haghparast et al. in 2008 designed unsigned array multiplier of size 4 × 4 [14] which includes 16 bits of partial product of the form $x_i y_i$. Partial products are generated by using Peres gate in parallel which are in parallel. Multi-operand addition is done by using HNG and PG gates as shown in Fig. 11. HNG gate is used as full adder, and PG gate is used as half adder. Overall, the design uses eight HNG gates and four PG gates.

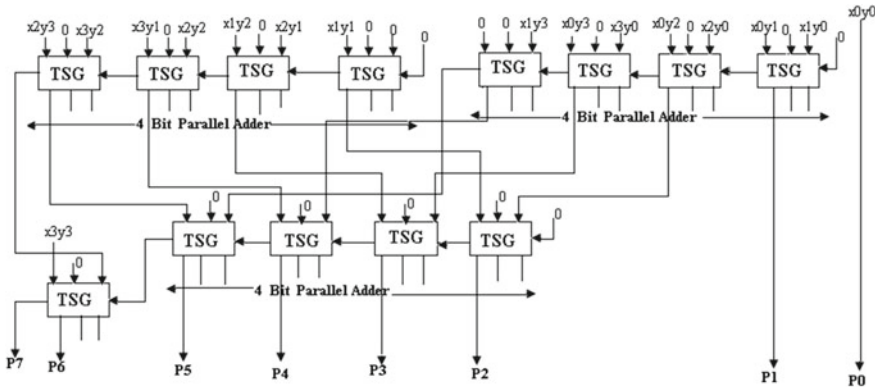


Fig. 10 Novel 4 × 4 reversible multiplier

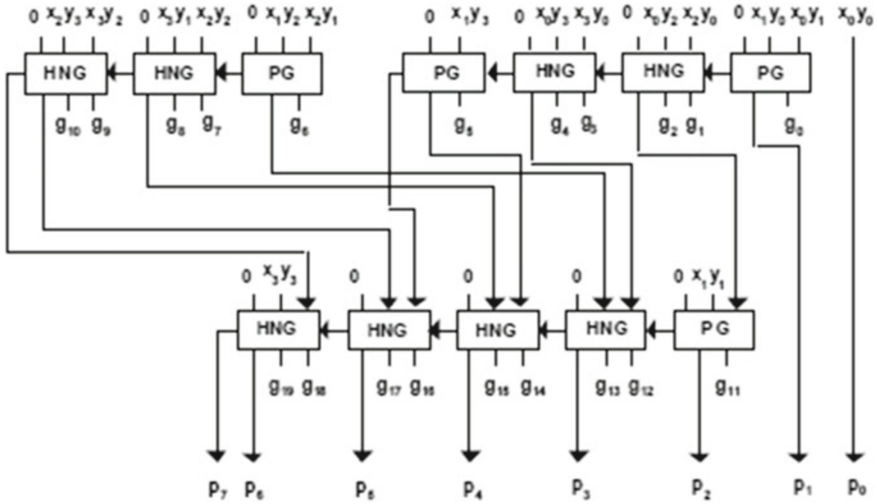


Fig. 11 Reversible 4 × 4 multiplier circuit using HNG and PG gate

Design has four reversible half adders realized by using Peres gate because it has low-quantum cost and hardware complexity compared to HNG gate. Design uses 28 gates totally for partial product generation, half adder and full adder, 52 constant inputs, and 28 garbage outputs.

Haghparast et al. in 2009 [36] proposed a 4 × 4 multiplier designed using two ways as shown in Fig. 12. Partial products are produced using Toffoli and Peres gates in parallel. Most fundamental block for a multiplier is full adder circuit. HNG gate is used as reversible full adder in first approach, and in second design approach, Peres gate is used as full adder.

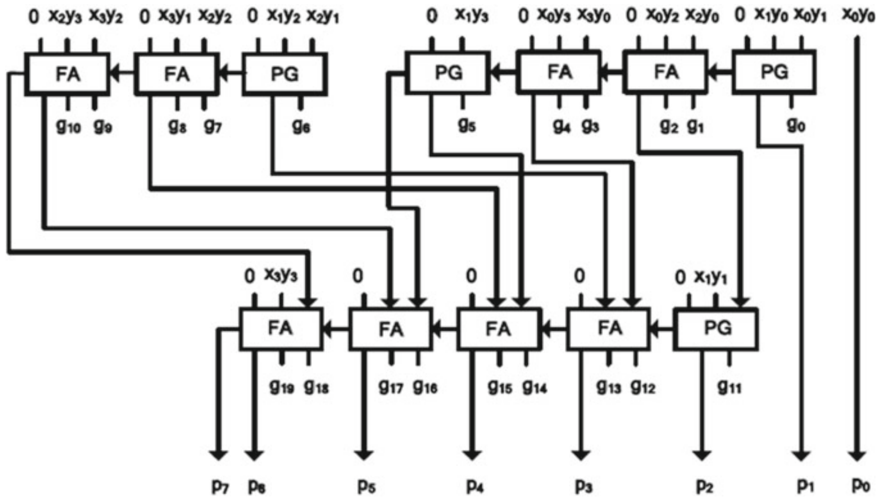


Fig. 12 HNG gate as full adder in design 1 and Peres gate as full adder in design2

Design 1 uses 28 gates, 28 ancilia inputs, 28 unused output, and quantum of the design is 137. Design 2 has 36 gates, 28 ancilia inputs, 28 unused outputs, and quantum cost of the multiplier is 153 (Fig. 13).

Haghparast et al. in 2012 [37] proposed 4 × 4 reversible multiplier using Peres gate and Toffoli gate for the generation of partial products which act horizontally. Reversible summation network uses eight HNG gates and four Peres gates, where HNG gates are realized as reversible full adder and Peres gates are realized has reversible half adder. Compared to the previous design [36], only the structure of the circuit is changed, but the number of gates has not been altered with the previous design.

Layers of the circuit are less in number because its critical path (longest input to output path) is shorter. Shorter the path, the propagation delay is shorter, and the response time is proportionally shorter. Design [37] uses 28 total gates, 28 ancilia inputs, 28 unused outputs, and quantum cost of the design is 137.

Bhagyalakshmi et al. in 2012 proposed [38] a four-bit parallel multiplier. BVPPG gates are used for the partial product generator. BVPPG has a unique feature that it can pass through three inputs and produce two product terms simultaneously with two constant inputs. Eight BVPPG gates are used for partial product generator circuit. As shown in Fig. 14, four operand addition is done by using four PG gates and eight DPG gates. Generation of partial products and addition is realized using Revkit [39].

Rangaraju et al. in 2013 proposed a design [40] using RAM gate for replication of the signals since fan-out of one is permitted in reversible logic. Partial products are generated using PG gate. Proposed multiplier circuit is shown in Fig. 15. Full adder is realized by using eight DPG gates, and half adder is realized by using PG gates. Partial product generation is done by using PG gate, and four RAM gates are used for fan-out realization [14].

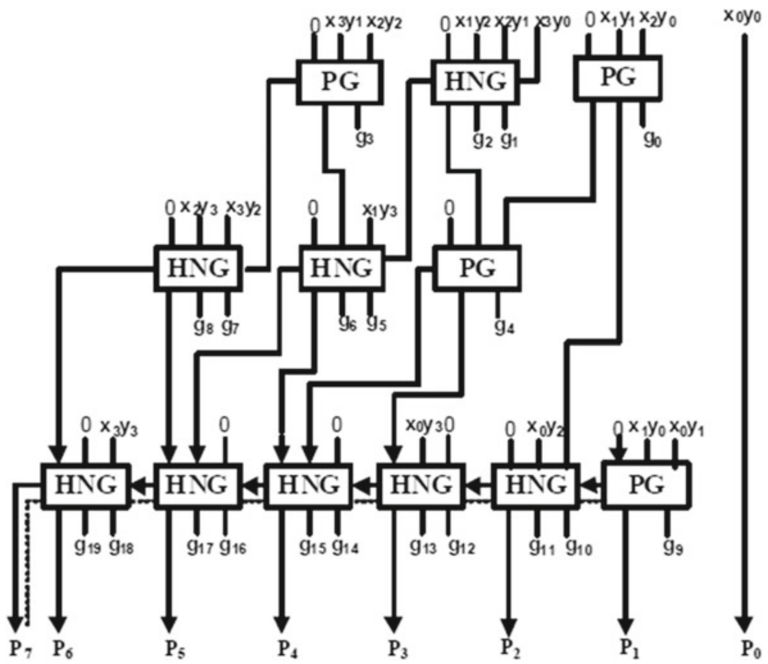


Fig. 13 Reversible summation network

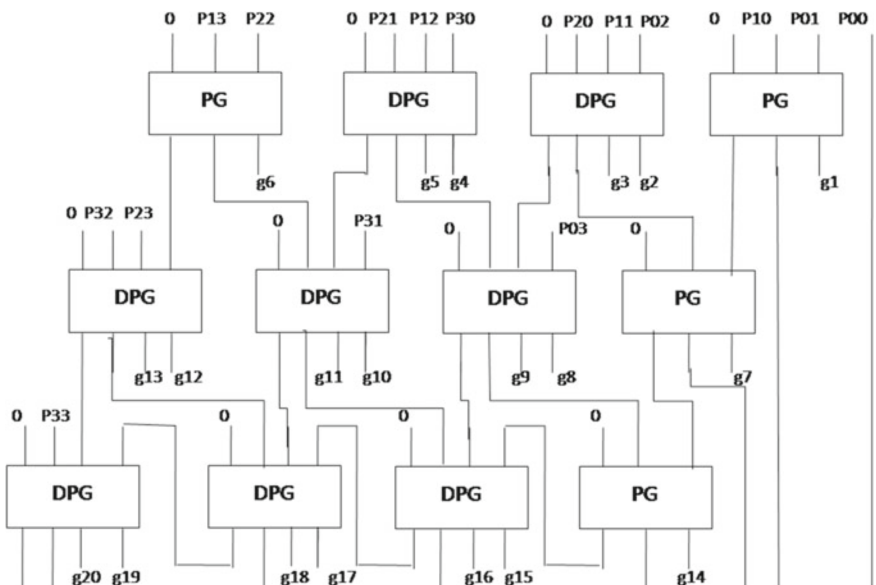


Fig. 14 Four operand addition circuit MOA

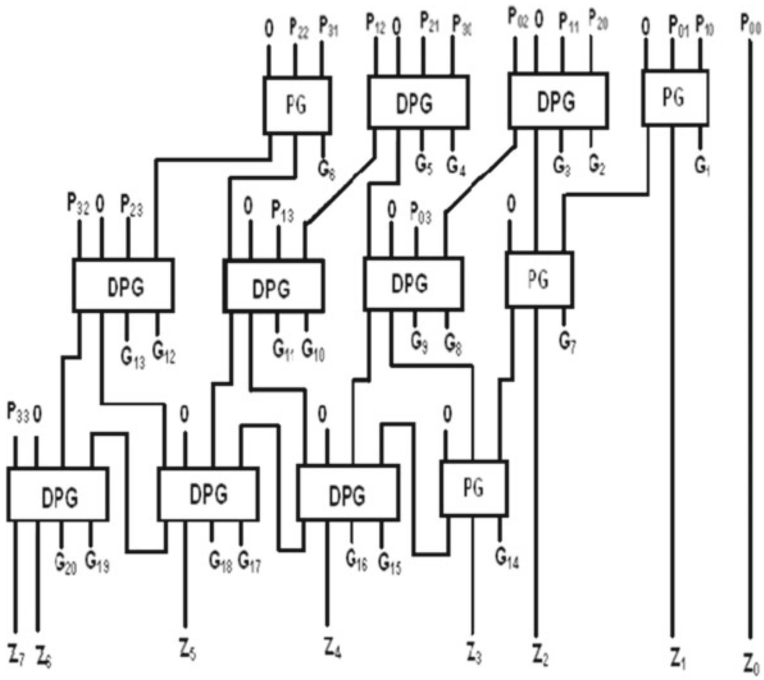


Fig. 15 MOA for reversible multiplier circuit

Valinataj in 2017 proposed design [41] a parity-preserving array multiplier for both and numbers, i.e., signed and unsigned. Signed multiplication is performed using Baugh–Wooley technique. Parity of the input bits is same as the parity of the output bits in parity conserving gates which are reversible [42–48]. Fault-tolerant operation does not require an extra circuit in design or verification [8]. Partial products are produced by using LMH and FRG gates. Partial products are produced by four different designs by using the same reversible gates. ZPLG and ZCG gates are used for multi-operand addition as shown in Figs. 16 and 17. Three designs are discussed for signed and unsigned multi-operand addition.

Pourali Akbar et al. in 2018 proposed design [49] a Wallace tree-based technique of 4 * 4 parallel multiplier. Partial products are generated by using two different methods. In first method, FG and TG gates are used to compute the partial products. Toffoli gates and Peres gates are used for partial product production. The difference between these two methods is with gate count and quantum cost. Wallace-based multiplier requires five FAs, three HAs, and single 4-bit adder circuit. Four-bit ripple carry adder requires three FAs and one HA, so totally eight FAs and four HAs are required as shown in Fig. 18.

Radha et al. in 2018 proposed design [50] 4 * 4 high-speed efficient multiplier realized by HNG gate only. Partial product as well as addition is performed by using HNG gate. Reversible full adder shown in Fig. 19 is done by HNG gate having input

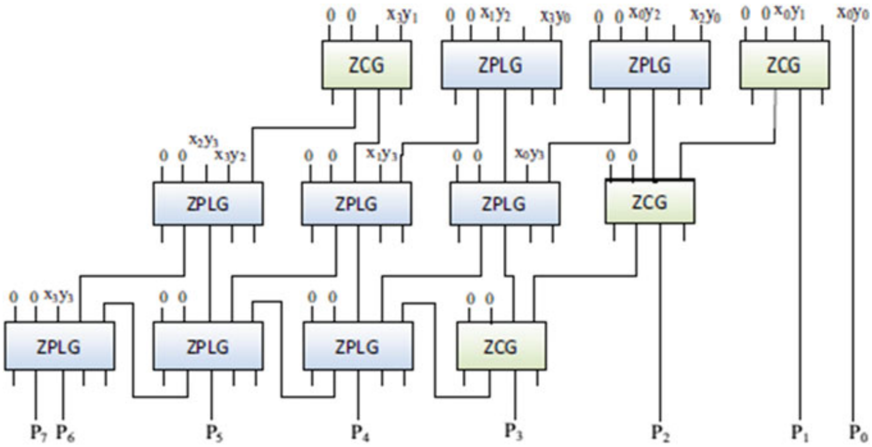


Fig. 16 Multi-operand addition for unsigned array multiplier

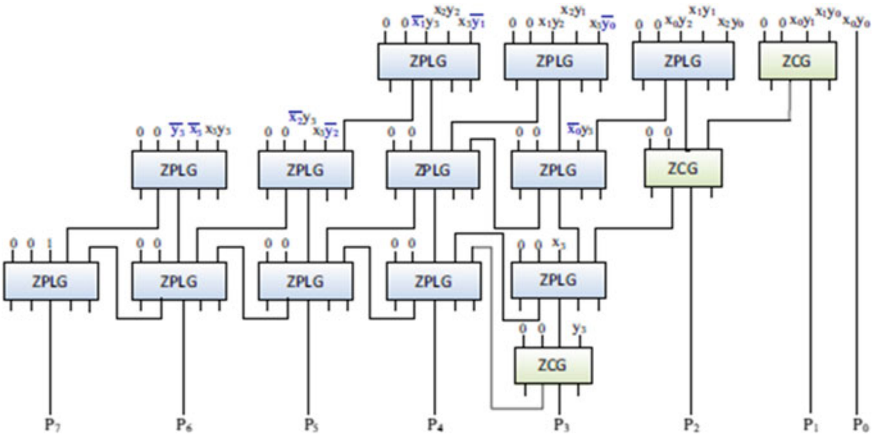


Fig. 17 Multi-operand addition for signed array multiplier

bits as (*A*, *B*, *C* in, 0), and then the output bits are (*A*, C_{in}, sum, cout). HNG gate can be realized as AND gate by having input vector (*A*, *B*, 0, 0), and the output bits are (*A*, *B*, *AB*, *AB*). Proposed multiplier is coded using Verilog and implemented using Cadence tool [17].

Vandhana et al. in 2020 proposed design [50] array multiplier using BME gate and WG gate. BME gates are used for partial product generation as shown in Fig. 20, and two bits of partial product are generated using one BME gate. WG gate as shown in Fig. 21 is used as 4-bit RCA. Comparing with the other existing designs, the recent design is optimized with design parameters like gate count, unused outputs, ancilia inputs, and quantum cost [18] (Table 1).

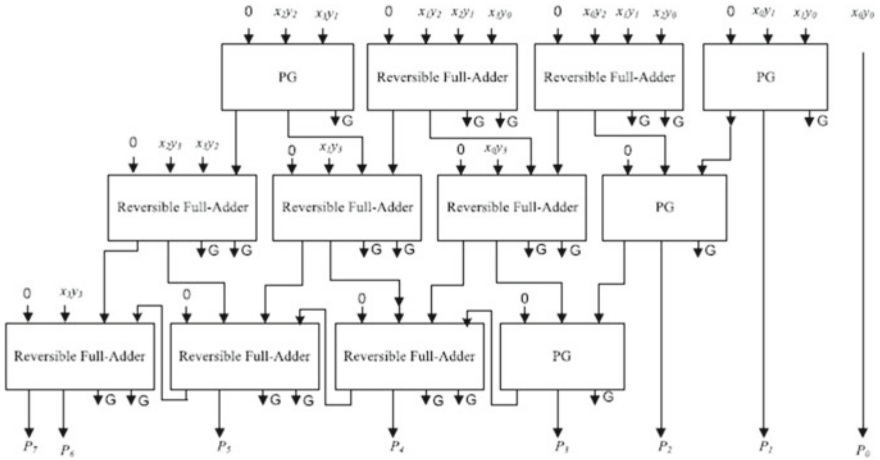


Fig. 18 Multi-operand addition

6 Proposed Array Multiplier

Proposed multiplier design uses BME gates for partial product generation, HNG, and Peres gate for multi-operand addition as shown in Fig. 23.

Figure 24 shows the simulation result for 4-bit unsigned array multiplier.

7 Conclusion

An attempt has been made in this paper to review parallel multiplier in terms of array multiplier using various gates for evolving the combinational logic circuits by means of the reversible logic. Functional working of the designs are verified using EDA tools like Xilinx software or Cadence. Array multiplier design is well known because of its regular structure. Array multipliers are realized by direct transformation with the manual multiplication process. Multiplier circuit is constructed by simple add and shift process. Review on array multipliers is presented. Comparative study is done for existing array multipliers in terms of ancilia inputs, unused outputs, number of gates used, and types of reversible gates used and their quantum cost.

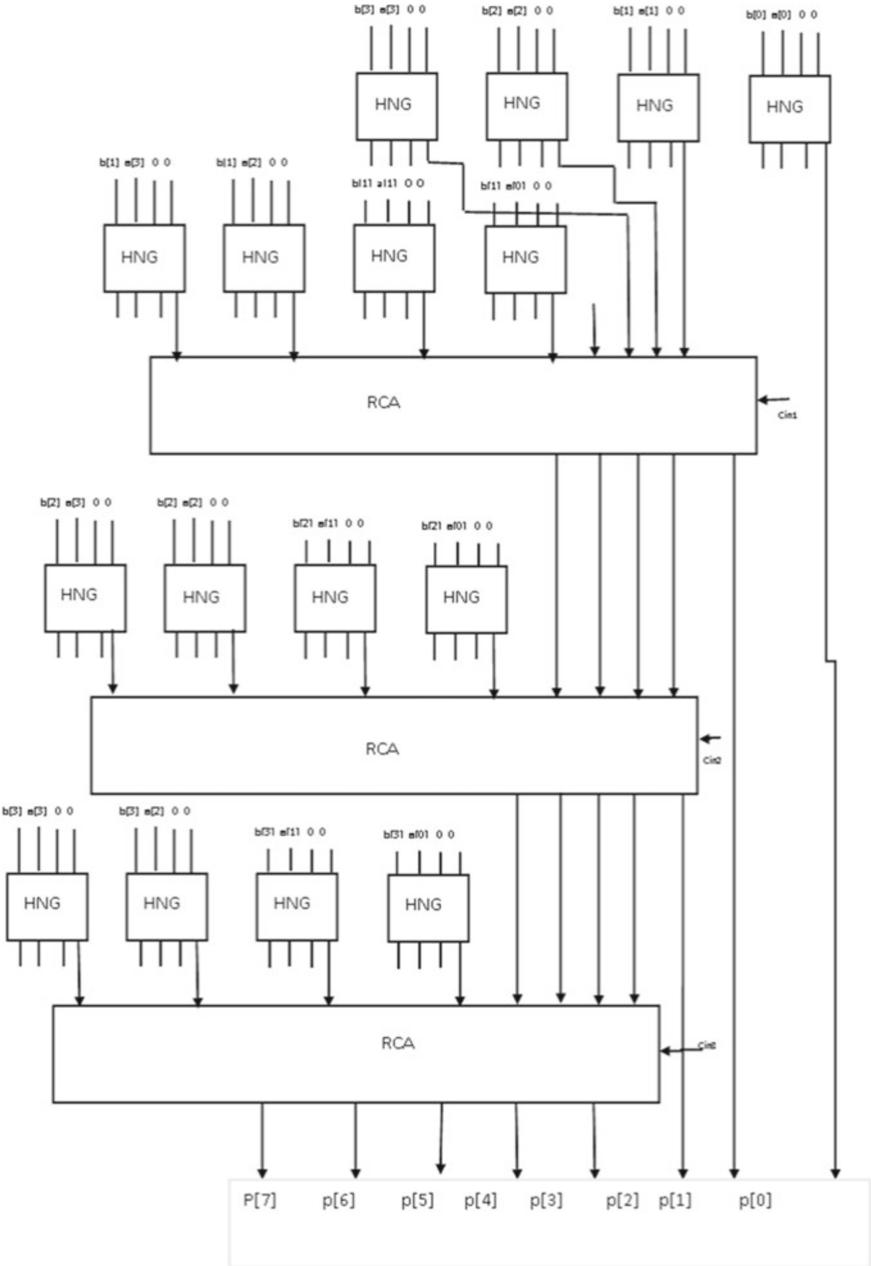


Fig. 19 Partial product and multi-operand addition

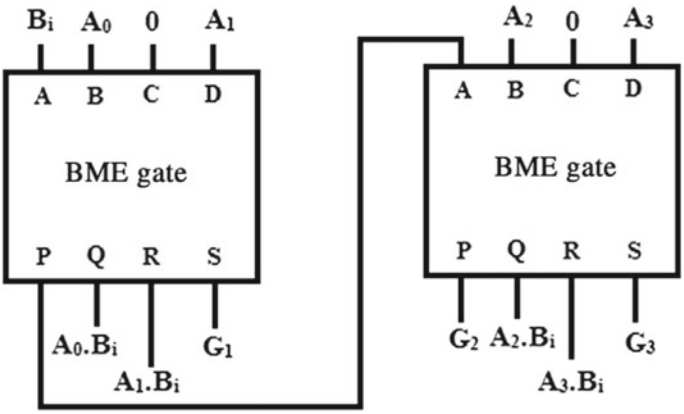


Fig. 20 AND operation for *i*th partial product

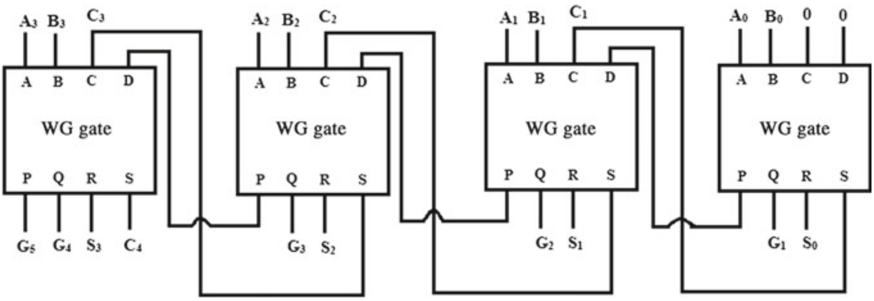


Fig. 21 4-bit ripple carry adder

Table 1 Comparison of design parameters of existing array multiplier

Array multiplier	Constant inputs	Garbage outputs	Number of gates	Quantum cost
Existing design [15]	Unspecified	Unspecified	29	Unspecified
Existing design [14]	52	28	28	Unspecified
Existing design [36]	28	28	28	137
	28	28	36	153
Existing design [37]	28	28	28	137
Existing design [38]	28	28	20	144
Existing design [40]	40	40	32	140
Existing design [41]	52	52	28	238
Existing design [49]	28	28	28	137
Existing design [50]	44	72	28	Unspecified
Existing design [51]	16	25	20	120
Best findings	Existing design [51]	Existing design [51]	Existing design [38] and [51]	Existing design [51]

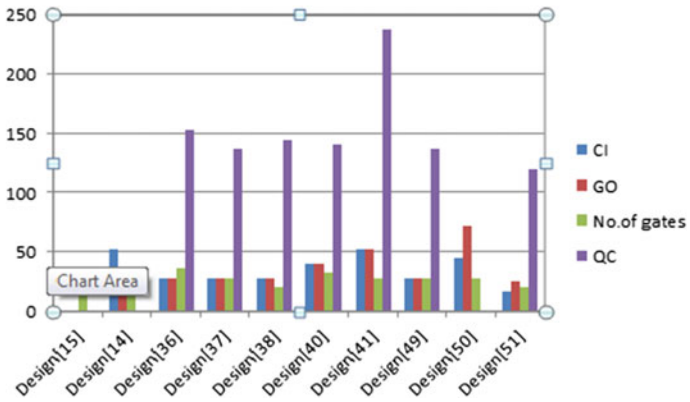


Fig. 22 Comparative analysis of array multiplier

Fig. 23 RTL schematic of array multiplier

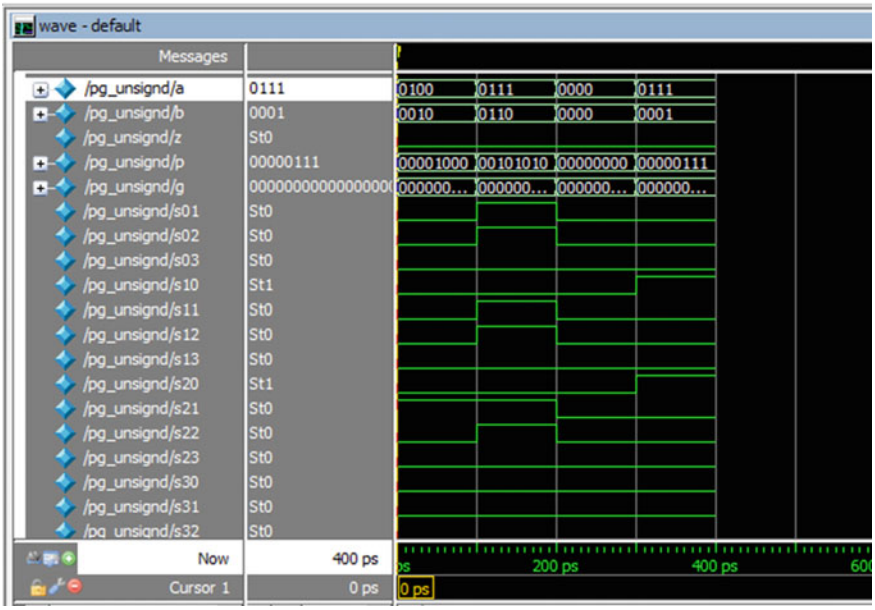


Fig. 24 Simulated waveform

References

1. Landauer R (1961) Irreversibility and heat generation in the computational process. IBM J Res Dev 5:183–191
2. Bennett CH (1973) Logical reversibility of computation. IBM J Res Dev 17:525–532
3. Kerntopf P, Perkowski MA, Khan MHA (2004) On universality of general reversible multiple valued logic gates. In: IEEE proceeding of the 34th international symposium on multiple valued logic (ISMVL'04), pp 68–73

4. Perkowski M, Al-Rabadi A, Kerntopf P, Buller A, Chrzanowska-Jeske M, Yanushkevich VS, Jozwiak L (2001) A general decomposition for reversible logic. In: Proceedings of RM, Starkville, pp 119–138
5. Perkowski M, Kerntopf P (2001) Reversible logic. Invited tutorial. In: Proceedings of EURO-MICRO, Warsaw, Poland
6. Alexis De V (2010) Reversible computing: fundamentals quantum computing and applications. Wiley, Hoboken
7. Frank MP (2005) Approaching the physical limits of computing. In: International symposium on multiple valued logic, pp 168–185
8. Parhami B (2006) Fault tolerant reversible circuits. Pacific Grove, pp 1726–1729
9. Kerntopf P (2006) A comparison of logical efficiency of reversible and conventional gates. In: Workshop on logic synthesis, IEEE, pp 261–269
10. Feynman R (1985) Quantum mechanical computers. Optics News 11:11–20
11. Toffoli T (1980) Reversible computing, Tech Memo MIT/LCS/TM-151. MIT Lab for Computer Science
12. Fredkin E, Toffoli T (1982) Conservative logic. Int'l J Theor Phys 21:219–253
13. Peres A (1985) Reversible logic and quantum computers. Phys Rev A 32(6):3266–3276
14. Haghparast M, Jafarali Jassbi S, Navi K, Hashemipour O (2008) Design of a novel reversible multiplier circuit using HNG gate in nanotechnology. World Appl Sci J 3(6):974–978
15. Thaplyal H, Srinivas MB (2006) Novel reversible multiplier architecture using reversible TSG gate. In: IEEE international conference on computer systems and applications, pp 100–103
16. Islam MS et al (2009) Low cost quantum realization of reversible multiplier circuit. Inf Technol J 8:208–213
17. Shams M, Haghparast M, Navi K (2008) Novel reversible multiplier circuit in nanotechnology. World Appl Sci J 3(5):806–810
18. Naderpour F, Vafaei A (2008) Reversible multipliers: decreasing the depth of the circuit. ICECE
19. Ehsanpour M, Moallem P, Vafaei A (2010) Design of a novel reversible multiplier circuit using modified full adder. In: International conference on computer design and applications, 1389, ICCDA 2010
20. Islam MS, Rahman MM, Begum Z, Hafiz MZ (2009) Low cost quantum realization of reversible multiplier circuit. Inform Technol J 8(2):208–213. ISSN 1812-5638
21. Thapliyal H, Ranganathan N (2017) Design of efficient reversible logic based binary and BCD adder circuits. ACM J Emerg Technol Comput Syst
22. Thaplyal H, Srinivas MB, Arabnia HR (2005) A reversible version of 4x4 bit array multiplier with minimum gates and garbage outputs, (ESA'05), Las Vegas, USA, pp 106–114
23. Han C-Y, Park H-J, Kim L-S (2001) A low-power array multiplier using separated multiplication technique. IEEE Trans Circuit Syst II: Analog Digit Signal Proces 48:866–871
24. Shah S, Al-Khalili A, Al-Khalili D (2000) Comparison of 32-bit multipliers for various performance measures. In: Proceedings of the 12th IEEE international conference on microelectronics
25. Thakre AK, Chiwande SS, Chafale SD (2014) Design of low power multiplier using reversible logic gate. In: 2014 international conference on green computing communication and electrical engineering (ICGCCEE), Coimbatore, pp 1–6
26. Anitha M, Rajani A, Pushpalatha N (2014) Optimized multiplier using reversible logic gates: a Vedic mathematical approach. IJARCET 3(10)
27. Syal N, Sinha HP (2012) Design of fault tolerant reversible multiplier. Int J Soft Comput Eng (IJSCE) 1(6). ISSN: 2231-2307
28. Babazadeh S, Haghparast M (2012) Design of a nanometric fault tolerant reversible multiplier circuit. J Basic Appl Sci Res Text Road Publ 2(2):1355–1361
29. Frank MP (2005) Introduction to reversible computing: motivation, progress, and challenges. In: Proceedings of the 2nd conference on computing frontiers, pp 385–390
30. Maaz MB, Abu-Shama E, Bayoumi M (1996) A fast and low power multiplier architecture. In: Proceedings of the 39th midwest symposium on circuits and systems, Iowa, pp 53–56

31. Babu YR, Syamala Y (2011) Implementation and testing of multipliers using reversible logic. In: Proceeding of international conference on advances in recent technologies in communication and computing IET, pp 171–175
32. Sharath BP, Suhas KV (2013) A new approach to the design and implementation of multipliers using reversible logic. Int J Eng Sci Invent 2(10):20–23
33. Baneriee A, Pathak A (2010) Reversible multiplier circuit. In: Proceeding of 3rd international conference on emerging trends in engineering and technology, IEEE, pp 781–786
34. Islam MS, Rahman MM, Begum Z, Hafiz MZ (2013) Low cost quantum realization of reversible multiplier circuit. Inf Technol J 208–213
35. Thapliyal H, Srinivas MB (2005) Novel reversible TSG gate and its application for designing reversible carry look ahead adder and other adder architectures. In: Tenth Asia-Pacific Computer Systems Architecture Conference (ACSAC05), pp 24–26, 805–817
36. Haghparast M, Mohammadi M, Navi K, Eshigi M (2009) Optimized reversible multiplier circuit. J Circuit Syst Comput World Sci Publ Co 18:311–323
37. Assarian M, Haghparast M, Navi K (2012) Delay reduction in optimized reversible multiplier circuit. Res J Appl Sci Eng Technol 4(1):27–32. ISSN: 2040-7467
38. Bhagyalakshmi HR, Venkatesha MK (2012) Optimized multiplier using reversible multicontrol input Toffoli gates. Int J VLSI Des Commun Syst (VLSICS) 3(6)
39. Soeken M, Frehse S, Wille R, Drechsler R, RevKit: a toolkit for reversible circuit design
40. Rangaraju HG, Babu Suresh A, Muralidhara KN (2013) Design of efficient reversible multiplier. Adv Comput Inf Technol AISC 178:571–579
41. Valinataj M (2017) Novel parity-preserving reversible logic array multipliers. Springer, Berlin
42. Hatkar A, Narkhede NP (2014) ASIC design of reversible multiplier circuit. IEEE Comput Soc 47–52
43. Parhami B (2006) Fault-tolerant reversible circuits. In: 40th Asilomar conference on signals, systems and computers (ACSSC), pp 1726–1729
44. Valinataj M (2017) Design of parity-preserving reversible logic signed array multipliers. World Acad Sci Eng Technol Int J Phys Math Sci 11(7)
45. Ahmad N, Mokhtar AH, Othman NB, Soon CF, Hadi Ab Rahman AA (2017) VLSI implementation of fault tolerance multiplier based on reversible logic gate. In: IOP Conference Series: Materials Science and Engineering, p 226
46. Jamal L, Mushfiqur Rahman Md, Hasan Babu HMd (2013) An optimal design of a fault tolerant reversible multiplier. IEEE
47. Poornima M, Suma MS, Palecha N, Malavika T (2013) Fault-tolerant reversible logic for combinational circuits: a survey. In: Proceedings of international conference on VLSI, communication, advanced devices, signals & systems and networking. Springer, Berlin
48. Akbar EPA (2018) An efficient design for reversible Wallace unsigned multiplier. Theoretical computer science. Elsevier, Amsterdam
49. Radha N, Maheswari M (2018) High speed efficient multiplier design using reversible gates. In: International conference on computer communication and informatics (ICCCI-2018), Jan 04–06 2018
50. Shukla V, Singh OP, Mishra GR, Tiwari RK (2020) Design of array multiplier circuit using reversible logic approach with optimized performance parameters. Elsevier, Amsterdam

Genetic Algorithm and Naïve Bayes-Based (GANB) Diabetes Mellitus Prediction System

Dharmveer Kumar Yadav, Chandrashekhar Azad, Kanchan Bala, Pradeep Kumar Sharma, and Sujeet Kumar

Abstract Today, machine learning plays a significant role in the classification of healthcare conditions. Machine learning is the process of finding, discovering, and modeling massive amounts of data in order to identify unknown relationships and patterns that are useful to the decision makers. Medical data mining has begun to emerge exemplary, with the potential to discover interesting insights from medical domain datasets. These patterns may be used in the making of clinical decisions. In this research work, we proposed genetic algorithm and Naïve Bayes-based (GANB) diabetes mellitus prediction system. In which, genetic algorithm (GA) is used for feature selection and Naïve Bayes is used for prediction. The proposed system is trained using Pima Indian diabetes dataset (PIDD). The dataset is preprocessed using synthetic minority oversampling technique (SMOTE) for solving the issue of class imbalances. The effectiveness of the system is evaluated using classification accuracy (CA) and error. The simulation results show that the proposed GANB system achieves better accuracy as compared to the results of related earlier studies.

Keywords Naïve Bayes · SMOTE · Genetic algorithm · Machine learning · Health · PIMA Indian diabetes dataset

D. K. Yadav (✉)
Computer Science and Engineering, Katihar Engineering College, Katihar, India
e-mail: Kumar.dharmveer@gmail.com

C. Azad
Computer Application Department, Computer Science and Engineering, NIT, Jamshedpur, India
e-mail: csazad.ca@nitjsr.ac.in

K. Bala
Sitamarhi Institute of Technology, Sitamarhi, India

P. K. Sharma · S. Kumar
Computer Science and Engineering, Katihar Engineering College, Katihar, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_47

1 Introduction

The most common chronic disease in India is diabetic mellitus. Diabetes mellitus is a disorder that influences the insulin receptor, prompting an unpredictable starch digestion, and increases glucose levels. This high blood sugar affects various human body organs which, in effect, complicates many of the body's causes, particularly the blood veins and nerves. Diabetes occurs as a long-lasting chronic disease due to insulin-producing pancreas malfunction, which controls the level of blood sugar. Late diagnosis of diabetes could lead to increased risk of macro-vascular and capillary complications, kidney failure, etc., which could increase the cost of health care and/or even endanger the patient's life. It is a disease characterized by metabolic disorder in which the patient has high blood glucose level which may be due to insulin production being low, or because the body is not responding properly to insulin, or both. Indian Medical Research Council (ICMR, Section 1, 2005) describes diabetes as the metabolic cum vascular syndrome [1]. Type 1 diabetes is a situation when pancreas fail to create insulin insufficient. The reason for type 1 diabetes is unsure, and with current information, it is now not preventable. The symptoms include the urinary discharge, weakness, weight reduction, changes in vision, and tiredness. Some signs and symptoms can happen suddenly. Type 2 diabetes occurs when body fails to produce sufficient insulin. This is also called adult onset diabetes occurs in people whose age is more than 40. But, now a days, it is increasing in child also. Type 2 diabetes is a result of extra physique weight and physical idleness. The pancreas releases some insulin usually in type 2 diabetes. But, body cells don't accept insulin. Type 2 may increase the chance of heart disease. Third type is called gestational diabetes. In this type of disease, glucose levels become high during pregnancy. Women with gestational diabetes are at increased risk of pregnancy and delivery complications. In the future, they and their children also have the risk of type 2 diabetes [2, 3]. In health care, machine learning algorithm is used to predict the disease [4, 5].

The organization of paper is arranged as follows: Related work is presented in Sect. 2. Materials and Methods are described in Sect. 3. Architecture of Proposed GANB-based Diabetes Mellitus Prediction System is described in Sect. 4. Dataset, Results, and Discussion are discussed in Sect. 5. Finally, Conclusion are presented in Sect. 6.

2 Related Works

Fuzzy expert system is developed to help doctors to determine the regulation of nephropathy in Type 2 diabetes patients. The study is based on an FES model developed using the guidelines for clinical practice. They performed sixty tests with experts of doctor teams [6].

In data mining classification survey many major types of classification algorithms are consisted with k-nearest neighbor classifier, Naive Bayes, SVM, and IB3. The author provides a general overview of different classification algorithms and their benefits and disadvantages [7].

Author proposed an algorithm using classification technique. They compared and analyzed the performance of various classifiers on the basis of accuracy, execution time, type of dataset, and domain. Data mining technique like decision tree classification is used in their work [8].

In presented classification of diabetes using binary logistic regression, multilayer perceptron, and k-nearest neighbor. Researchers analyzed and performed comparison of all the proposed classification algorithms. In their comparison, author found KNN performs better results than other two method [9].

In [10, 11], authors developed a specialist system that defines the type of diabetes with the likelihood of occurrence as well as percentage rule matches. This was planned with the help of a well-known diabetologist, based on the symptoms identified.

Using the KoGES dataset, a multi-stage model is proposed to predict which people are most likely to have diabetes. A physiological model was suggested to predict the blood glucose level. They have taken data from five patient and processed with physiological features [12–15].

A hybrid model was built to determine whether the patient being diagnosed may or may not develop diabetes disease in 5 years. A tool Weka is used for this research work, and dataset of PIMA Indian diabetes dataset was collected. The proposed hybrid model achieved precision of 92.38% [16].

Another type of hybrid prediction was proposed using Weka tool. They have also used PIMA dataset in their research work. In their research work, first, data are processed to find the F-score value to find the features with high F-score value. Then, clustering and SVM technique are used for classification purpose [17].

Neural networks model is proposed for classification of diabetes. For training and testing, PIMA dataset is collected. Thy have considered 576 samples for training and 192 for testing purpose [18].

In paper [19], the researcher developed predictive model using support vector machine to predict whether or not patient suffering from diabetes. To perform an experiment, they used Pima dataset.

In paper [20], a new method is proposed called FNC technique for diagnosis of diabetes. This work uses fuzzy logic, neural network, and case base reasoning with 16 input attributes of 200 patient data. The tool Matlab and MyCBR plug-in is used to implement case-based reasoning.

In paper [21], decision tree technique is used for type 2 diabetes. Pima Indian diabetes datasets are used. Some preprocessing methods like selecting attributes and handling missing values were used to enhance data quality.

In paper [22], authors suggested a prediction model for diagnosis of diabetes early [23]. They have collected data for prediction from a diabetes clinic about 545 patients. First, neural networks trained and tested patient data with the neurons and found the highest accuracy in a neural network consists seven neurons.

In paper [24], the authors developed decision-support model for expert health care to predict diabetes. This model uses Pima dataset for training the dataset. In this work, decision tree with k-nearest neighbor was presented and find a 90.43% accuracy.

In paper [23], authors proposed an elastic net model for diabetic. This model increases the accuracy of glucose estimation in diabetic patients. Blood sample was not taken in this work. Device is used to collect data for experiment purpose.

3 Materials and Methods

3.1 Synthetic Minority Over Sampling Technique (SMOTE)

In practice, many research have shown that by providing balanced data, better performance on prediction can be obtained. Most machine learning technique perform poor result if data are not balance. The sampling methods are used for imbalance data in machine learning [25–28]. The concept of sampling method is to change the original dataset class value to the equal in label class distribution. In general, dataset contains classes of balance and imbalance. The real-world data in many application like fraud detection and medical diagnosis are imbalanced. This class imbalanced is categorized as supervised learning problem of data science. The SMOTE method is an over sampling approach applied for machine learning; mainly, for imbalance data, it is more effective for medicine dataset [29]. This produces new instances from the minority classes. These new instances are generated according to the features selected from original dataset so that it is similar to the original instances.

3.2 Naïve Bayes

A Naive Bayes is defined as probabilistic classifier technique which is based on Bayes' theorem with assumptions of independence. A Naive Bayes is a class of probabilistic classifier machine learning method used in classification. This is based on following given Bayes theorem [30].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

Using Bayes theorem, given that B has happened, we can find the likelihood of A occurring. Here, evidence is B, and hypothesis is A in Eq. 1. Here, the assumption is that the predictors/functions are independent. This is the presence of one particular attribute does not affect the other. And it is called Naïve. Naïve Bayes is a probability-based classification system that assumes characteristics to be independent, and classification results are unaffected by the influence on each other's characteristics. Naïve Bayes may be used for the classification of health and many other fields. Naive Bayes is a simple, dynamic algorithm that is more or less used for the task of classification. It performs efficiently in contrast to number variables in the case of absolute input variables. Naive Bayes makes significant results when it can be applied for data analysis. Naive Bayes is a strong and aggressive predictor. This can be useful for a very large number of datasets.

The NB classifier is given below:

1. Consider X consists a set of tuples with class labels. Vector with n-dimensional attributes represents each tuple, $X = (x_1, x_2, x_3, ..., x_n)$
2. Consider $C_1, C_2, \ldots, C_m$ classes. Provided a tuple, X is predicted by the classifier that X belongs to the class with the highest posterior probability, conditional on X. The NB classifier will predict X belongs to class C_i if and only if

$P(C_i|X) > P(C_j|X)$ for $1 \leq \mathrm{j} \leq m, j \neq i$.

Hence, need to maximize the $P(C_i|X)$. The C_i class for which $P(C_i|X)$ is maximized. This is called as maximum posteriori hypothesis; using Bayes' theorem, we can express as follows

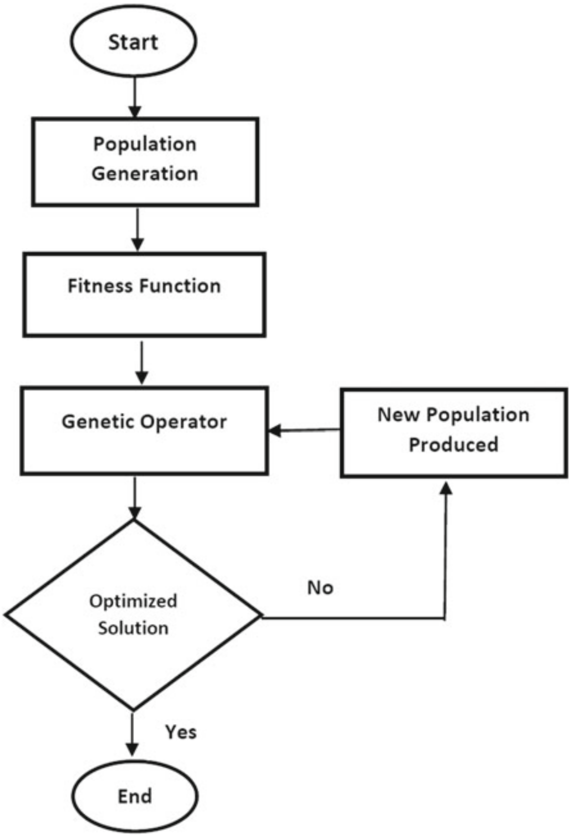
$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}$$

3.3 Genetic Algorithm

GAs is an adaptive heuristic search algorithms which is based on natural selection and evolutionary concepts. GA is an optimization method based on survival of the fittest [31]. GA illustrates the selection mechanism in which the individuals are selected to produce next-generation offspring. The selection process begins with the selection from a population of the best fittest individuals. It generates offspring that inherits parents' traits, which will move on to the next generation. When two individuals have better fitness value, then there is chance of better offspring than parents and higher chance of survival. This cycle is starting to iterate, and at the end, a generation with the fittest individuals will be found. The evolution typically starts with a population of individuals that are randomly created. The fitness value for each individual is measured from population in each generation; individuals are taken from the existing population according to the fitness value and produces a new population. The flowchart is shown in Fig. 1 for GA algorithm. The GA uses three main type of steps to generate the next generation out of the current population at each step [32–34].

1. Selection: In this steps, reproduction process begins with the primary aim of producing the good solutions and remove bad solutions from the populations.

Fig. 1 Flowchart of genetic algorithm



It is achieved by finding and creating duplicate copies of effective solutions (in terms of fitness) within the existing population. Now, remove any poor solutions from the populations. In this process, parents are selected from the current population to produce the next-generation population. Number of methods are used for this process like Roulette Wheel, Boltzmann, and tournament selection method.

2. Crossover: This is recombination process. Genetic operator is used in this process. New solutions are created in the crossover; there are number of crossover method like single-point and two-point are used.
3. Mutation: In mutation probability, some parts of chromosome will be mutated. Mutation is done after completion of crossover when there is no change in offspring. If mutation is made, then one or more parts of a chromosome will be changed.

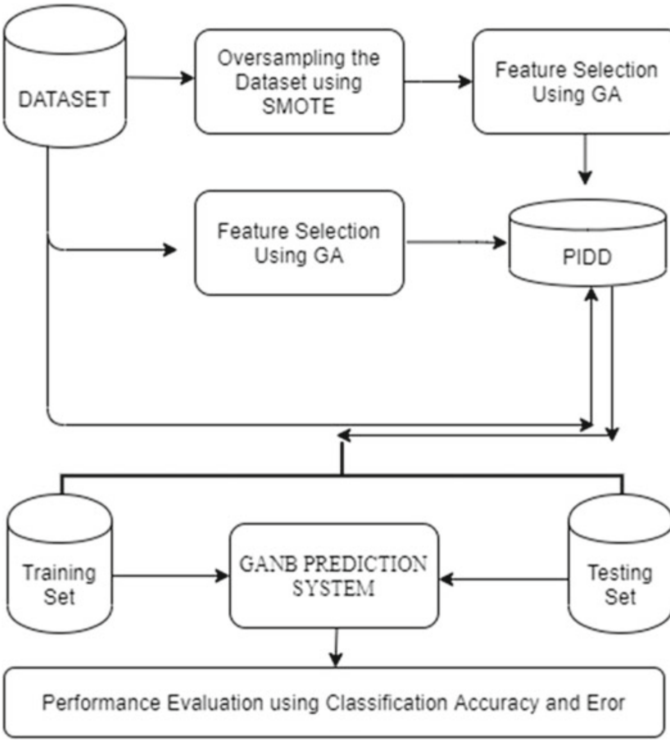


Fig. 2 Architecture of proposed system

4 Architecture of Proposed GANB-Based Diabetes Mellitus Prediction System

See Fig. 2.

5 Dataset, Results, and Discussion

5.1 Dataset and Experiment

In this work, Pima Indians diabetes database (PIDD) dataset is used from UCI repository for experiment. The PIDD dataset consists diabetic and non-diabetic samples, which is class of imbalance data. Dataset description is shown in Table 1.

This experiments are implemented on a PC with Intel(R) Core (TM) i5 8th generation and 8 GB memory, running on Windows 10. For simulation, Weka machine learning library and Java 1.8 are used.

Table 1 Dataset attribute

ID	Name of attribute	Description of attribute
1	Pregnancies	Number of times pregnant
2	Glucose	Plasma glucose concentration 2 h in an oral glucose tolerance test
3	Skin thickness	Triceps skin fold thickness (mm)
4	Blood pressure	Diastolic blood pressure (mm Hg)
5	Insulin	2-h serum insulin (μ U/ml)
6	Diabetes pedigree function	Diabetes pedigree function
7	BMI	Body mass index (weight in kg/(height in m^2))
8	Age	Age (years)
9	Outcome	Class variable (1 or 0) 1: Presence of diabetic 0: Absence of diabetic

5.2 Performance Measure

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

$$\text{Error} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

In this work, performance is measured using Eqs. (2) and (3).

5.3 Results

5.4 Discussion and Comparison with Existing System

Figure 2 shows architecture of proposed GANB-based diabetes prediction system. The class imbalance problem is resolved by using SMOTE; feature selection eliminates the insignificant features using correlation and GA. This reduced dimension of the dataset lowers the training complexity and also solves the issues of over fitting. Naive Bayes is used to train the GANB prediction system. Finally, performance of the proposed model is calculated using classification accuracy and error.

The split test methodology is implemented as a technique for planning and validating the results for the training and test dataset. In split test methodology training–testing set ranges are considered (60–40)%, (65–35)%, (70–30)%, (75–25)%, and (80–20)%. The method is simulated 10 times for each split, and the best outcomes is recorded for each dataset. The four variations of datasets are used, namely PIDD, PIDD_SM, PIDD_GA, and PIDD_SM_GA for simulation. PIDD is the PIMA Indian diabetes dataset; PIDD_SM is the oversampled using SMOTE. In PIDD_GA, features are selected using GA, and in PIDD_SM_GA, over sampled using SMOTE and features are selected using GA. Table 2 shows the result of proposed GANB prediction system, and Table 3 shows the comparison of proposed GANB prediction system with existing system. The proposed method is compared on the basis of accuracy and error. It is worth to mention that the proposed model yields superior results in comparison to the various existing schemes as shown in Table 2. The best outcome is observed on the PIDD_GA variant of dataset, second best result observed on PIDD_SM and PIDD_SM_GA variant of the dataset. Features selected using GA on PIDD are glucose, BMI, DiabetesPedigreeFunction, and age. Whereas, features selected using GA on oversampled dataset using SMOTE are pregnancies, glucose, insulin, BMI, DiabetesPedigreeFunction, and age. Optimal parameters in feature selection with GA are crossover probability 80.6, mutation probability 0.03, population size 20, and maximum generation 50. Optimal parameter configuration in feature selection using GA on oversampled dataset is crossover probability 0.7, mutation probability 0.03, population size 20, and maximum generation 60. The best performance observed on proposed system is accuracy 82.2917% and error 17.7083%; second best result observed is accuracy 81.2672% and error 18.7328%.

Table 2 Result of proposed GANB prediction system

Training and testing set Split size	Key_dataset	Accuracy	Error
60–40 Training–testing set	PIDD	77.5974	22.4026
	PIDD_GA	80.78176	19.21824
	PIDD_SM	80.43478	19.56522
	PIDD_SM_GA	80.48193	19.51807
65–35 Training–testing set	PIDD	78.0669	21.9331
	PIDD_GA	81.4126	18.5874
	PIDD_SM	81.2672	18.7328
	PIDD_SM_GA	81.2672	18.7328
70–30 Training–testing set	PIDD	78.7879	21.2121
	PIDD_GA	81.3853	18.6147
	PIDD_SM	80.0643	19.9357
	PIDD_SM_GA	80.7074	19.2926
75–25 Training–testing set	PIDD	79.6875	20.3125
	PIDD_GA	82.2917	17.7083
	PIDD_SM	79.5367	20.4633
	PIDD_SM_GA	80.6950	19.3050
80–20 Training–testing set	PIDD	78.57143	21.42857
	PIDD_GA	81.16883	18.83117
	PIDD_SM	78.74396	21.25604
	PIDD_SM_GA	79.22705	20.77295

Table 3 Comparison of proposed GANB prediction system with existing system

Key dataset	Accuracy	Error
Sim [35]	75.29	24.71
Sim + F1 [35]	75.84	24.16
Sim + F2	75.97	24.03
FMM [4]	69.28	30.72
FMM-CART [4]	71.35	28.65
FMM-CART-RF [4]	78.39	21.61
Binary-coded GA [5]	74.80	25.20
BP [5]	73.80	26.20
Binary-coded GA [5]	77.60	22.40
GANB (PIDD)	79.6875	20.3125
GANB (PIDD_GA)	82.2917	17.7083
GANB (PIDD_SM)	81.2672	18.7328
GANB (PIDD_SM_GA)	81.2672	18.7328

6 Conclusion

Diabetes is a disorder where blood sugar levels are abnormally high due to the lack of insulin produced by the body to meet its needs. Persons with diabetes are at risk for long-term skin, kidney, lung, brain, feet, and nerve-affected problems. In this research work, we proposed genetic algorithm and Naïve Bayes-based (GANB) diabetes mellitus prediction system. In which, GA is applied for feature selection, Naïve Bayes is used for prediction. The proposed system is trained using PIDD and is preprocessed using SMOTE for solving the issue of class imbalances. GA allows faster training of the prediction model. This reduces a model's complexity, which makes analysis simpler and also increases accuracy. The effectiveness of the system is evaluated using classification accuracy (CA) and error. It is worth to mention that the system giving remarkable accuracy 82.2917% and error 17.7083%. In the future, we will use some optimization techniques like PSO, ACO, and CSO along with the machine learning algorithm.

Acknowledgements We are thankful to Department of Computer Science & Engineering of Katihar Engineering College, Katihar for providing the laboratory facilities to perform this research work. We are heartily thankful to our Principal Mam Prof. Ranjana Kumari and AKU Patna for providing fund under CRS project to accomplish this work.

Funding The financial support provided for this research work is Under Collaborative Research Scheme (CRS) - TEQIP-III Program by AKU Patna.

References

1. Nayak J, Bhat PS, Acharya R, Lim CM, Kagathi M (2008) Automated identification of diabetic retinopathy stages using digital fundus images. J Med Syst 32(2):107–115
2. Alberti KGMM, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med 15(7):539–553
3. Azad CS, Mehta AK, Mahto D, Yadav DK (2020) Support vector machine based eHealth cloud system for diabetes classification. EAI Endors Trans Pervasive Health Technol 6:164627
4. Azad C, Jha VK (2017) Fuzzy min–max neural network and particle swarm optimization based intrusion detection system. Microsyst Technol 23(4):907–918
5. Azad C, Jha VK (2013) Data mining in intrusion detection: a comparative study of methods, types and data sets. Int J Inform Technol Comput Sci (IJITCS) 5(8):75–90
6. Meza-Palacios R, Aguilar-Lasserre AA, Ureña-Bogarín EL, Vázquez-Rodríguez CF, Posada-Gómez R, Trujillo-Mata A (2017) Development of a fuzzy expert system for the nephropathy control assessment in patients with type 2 diabetes mellitus. Expert Syst Appl 72:335–343
7. Archana S, Elangovan K (2014) Survey of classification techniques in data mining. Int J Comput Sci Mob Appl 2(2)
8. Bharathi A, Deepankumar E (2014) Survey on classification techniques in data mining. Int J Recent Innov Trends Comput Commun 2(7)
9. Selvakumar S, Kannan KS, GothaiNachiyar S (2017) Prediction of diabetes diagnosis using classification based data mining techniques. Int J Stat Syst 12(2):183–188
10. Choubey DK, Paul S (2017a) GA_SVM-A classification system for diagnosis of diabetes. Handbook of research on nature inspired soft computing and algorithms. IGI Global, Hershey, pp 359–397
11. Choubey DK, Paul S (2017) GA_RBF NN: a classification system for diabetes. Int J Biomed Eng Technol (IJBET) 23(1):71–93
12. Lee J et al (2011) Development of a predictive model for type 2 diabetes mellitus using genetic and clinical data. Osong Public Health Res Perspect 2(2):75–82
13. Plis K et al (2014) A machine learning approach to predicting blood glucose levels for diabetes management. Assoc Adv Artif Intel
14. Patil BM et al (2010) Hybrid prediction model for type-2 diabetic patients. Expert Syst Appl 37(12):8102–8108
15. Sarojini Ilango B et al (2010) A hybrid prediction model with F-score feature selection for type ii diabetes databases. In: A2CWiC
16. Temurtas H et al (2009) A comparative study on diabetes disease diagnosis using neural networks. Expert Syst Appl 36(4):8610–8615
17. Divya et al (2014) Predictive model for diabetic patients using hybrid twin support vector machine. In: Proceedings of international conference on advances in communication, network, and computing, CNC. Elsevier, Amsterdam
18. Ahmed TM (2016) Developing a predicted model for diabetes type 2 treatment plans by using data mining. J Theor Appl Inf Technol 90(2):181–187
19. Devi MN et al (2016) Developing a modified logistic regression model for diabetes mellitus and identifying the important factors of type II DM. Indian J Sci Technol 9(4)
20. Thirugnanam M et al (2016) Hybrid tool for diagnosis of diabetes. IIOAB J 7(5)
21. AlJarullah AA (2011) Decision tree discovery for the diagnosis of type II diabetes. In: International conference on innovations in information technology, IEEE, New York
22. Meysam J, Mahdi M (2016) Comparison of predictive models for the early diagnosis of diabetes. Healthc Inform Res. 22(2):95–100
23. Zanon M et al (2013) Regularised model identification improves accuracy of multisensor systems for noninvasive continuous glucose monitoring in diabetes management. J Appl Math 2013:1–10

24. Hashi EK et al (2017) An expert clinical decision support system to predict disease using classification techniques. In: International conference on electrical, computer and communication engineering (ECCE), ©2017 IEEE, February 16–18, 2017, Cox's Bazar, Bangladesh
25. Poolsawad N, Kambhampati C, Cleland J (2014) Balancing class for performance of classification with a clinical dataset. In: Proceedings of the world congress on engineering, vol 1
26. Wang J, Xu M, Wang H, Zhang J (2006) Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding. In: 2006 8th international conference on signal processing, vol 3. IEEE
27. García V, Alejo R, Sánchez JS, Sotoca JM, Mollineda RA (2006) Combined effects of class imbalance and class overlap on instance-based classification. In: International conference on intelligent data engineering and automated learning. Springer, Berlin, pp 371–378
28. Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D et al (2008) The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J Magn Reson Imaging 27(4):685–691. https://doi.org/10.1002/jmri.21049 PMID: 18302232
29. Lusa L et al (2015) Joint use of over-and under-sampling techniques and cross-validation for the development and assessment of prediction models. BMC Bioinform 16(1):1
30. Van Der Heijden AA, Ortegon MM, Niessen LW, Nijpels G, Dekker JM (2009) Prediction of coronary heart disease risk in a general, pre diabetic, and diabetic population during 10 years of follow up: accuracy of the Framingham, SCORE, and UKPDS risk functions: the Hoorn Study. Diabet Care 32(11):2094–2098
31. Darwin C (1859) On the origins of species by means of natural selection. Murray, London
32. Yadav DK, Dutta S (2020) Regression test case selection and prioritization for object oriented software. Microsyst Technol 26:1463–1477. https://doi.org/10.1007/s00542-019-04679-7
33. Azad C, Jha VK (2016) A novel fuzzy min-max neural network and genetic algorithm-based intrusion detection system. In: Proceedings of the second international conference on computer and communication technologies 2016, pp 429–439. Springer, New Delhi
34. Yadav DK, Dutta S (2017) Regression test case prioritization technique using genetic algorithm. In: Advances in computational intelligence, pp 133–140. Springer, Singapore
35. Örkcü HH, Bal H (2011) Comparing performances of backpropagation and genetic algorithms in the data classification. Expert Syst Appl 38(4):3703–3709
36. Luukka P (2011) Feature selection using fuzzy entropy measures with similarity classifier. Expert Syst Appl 38(4):4600–4607
37. Seera M, Lim CP (2014) A hybrid intelligent system for medical data classification. Expert Syst Appl 41(5):2239–2249

Multiscale Decomposition of HDR Images Using the Edge-Preserving Filters

A. S. Anand Swamy and N. Shylashree

Abstract Tone mapping is very much required to present the high dynamic range (HDR) images on conventional screens. Most popular and efficient way is to perform decomposition of the high dynamic range image into multi-layers. Large-scale intensity variations are assigned to the base layer, and small-scale intensity variations are assigned to the detail layer. Base layer is compressed retaining detail layers, and both the layers are combined; resulted image is displayed on the conventional display devices. In this process, only, contrast is reduced, whereas the details are preserved. In this paper, various base-detail layers decomposition techniques are compared using the metrics such as psnr, mPSNR, and HDRVDP3. Efficiency of edge-preserving filters to preserve the details while smoothing the contrast is analyzed.

Keywords Edge-preserving filters · HDR decomposition · Tone mapping

1 Introduction

The ratio of highest light intensity value to the lowest light intensity value is called as the dynamic range. It is measured in exposure value or stops. The dynamic range of the natural scenes is very high, but image acquisition devices like digital camera, camcorder, and conventional display devices like CRT or LCD monitors, printers are low dynamic range Medias. Screening devices have a dynamic range of 100:1, natural scene lit by bright sunlight, having deep shadows can have the high dynamic range of 100,000:1. Several LDR images are captured with different exposures using auto-exposure bracketing in a digital camera. Halos and contrast reversals will be

A. S. Anand Swamy · N. Shylashree (✉)
Department of Electronics and Communication Engineering, RV College of Engineering,
Bengaluru 560059, India
e-mail: shylashreen@rvce.edu.in

A. S. Anand Swamy
e-mail: anandswamyas@tjohngroup.com

Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_48

more if the intensity values from different exposures are combined directly. Hence [1], radiometric response function is estimated by properly aligning captured LDR images, and by combining the intensity values from different exposures, the radiance map is estimated. Finally, tone mapping is performed to present HDR images on LDR Medias. Tone mapping decreases the image contrast by preserving the details. Use of global transfer curve is simplest method to compress HDR radiance image into LDR gamut. Here, gamma curve maps the HDR image into displayable gamut. Applying gamma should not be done individual as to each of the channels if so, colors get muted. To overcome this, image is decomposed into two layers based on luminance and chrominance components. Global mapping is applied to luminance layer, and color image is reconstructed. This global approach works fine with images having low range of exposures, and with widely varying exposures, it fails to preserve details. A method similar to dodging and burning can be used for tone mapping. As before, image is decomposed into luminance channel and chrominance channel. The log luminance of the image is found, and it is low-pass filtered which results in the base layer. Subtracting base layer from log luminance yields the detail layer. Base layer is scaled down to the required range of log luminance to reduce the contrast. New log luminance image is produced by adding scaled base layer to the detail layer. Tone-mapped luminance image is obtained by taking exponent of new log luminance image. But, this method produces halo artifacts around the edges with more contrast in detail layer. Therefore, linear filters are not suitable for tone mapping; local edge-preserving filters are used for decomposition. The filters which preserves edges and smoothens the images can be classified into two: (i) local optimization-based filters, includes anisotropic and robust anisotropic diffusion, bilateral filtering and its different versions, guided image filtering, weighted guided image filters, global gradient guided image filters and (ii) global optimization-based filters, includes weighted least square filters, total variation (TV) [2], fast weighted least squares (FWLS) [3], L-norm gradient minimization [4]. Though global optimization-based filters yield excellent quality, their computation cost is high, and they are time-consuming. Local filters have very good efficiency but suffer from halo and gradient reversal effects. This paper is organized as fallows. Section 2—discussion on various local filtering-based edge-preserving smoothing filters. Section 3—study of global optimization-based filters. Section 4—conclusion.

2 Local Optimization-Based Filters

2.1 Anisotropic Diffusion

This is one of the earlier method which reduces the image noise while preserving the edges. It is similar to the technique of forming a scale space, in which an image, based on process of diffusion creates parametrized group of successively blurred images. In this family, each of resulting images is obtained by convolution operation with 2D isotropic Gaussian filter as system response and the image as the input. This method considers each pixel of image as an energy sync interacting with neighbor pixels based upon pixel intensity differences and values of conductance which are obtained from the local edge approximate. [5]. Anisotropic diffusion decomposes the image into multi-layers, where base layer is the smooth layer and detail layer is difference layer. Both the layers are enhanced using enhancement algorithm such as alpha rooting [6] or logarithm transfer shifting [7]. Both the layers are fused together. Energy diffusion equation is used by anisotropic diffusion. Gradients of filtering image itself are used as a guide for the diffusion process, avoiding edge smoothing [8]. Anisotropic diffusion is adiabatic, i.e., it preserves energy [8]. Limitations of anisotropic diffusion are (i) "stair-casing" effect around smooth edges. (ii) It is a slow process due to its discrete diffusion nature. (iii) It is ill posed; due to shock forming process, large changes in output can be caused, even for infinitesimal changes in input [8]. (iv) It over sharpens the image edges [9]. (v) It is non-iterative function which slowly converges. Diffusion process is stopped in between to obtain the piecewise smooth images as it finally converges to a constant image [10]. Qiaosong Chen et al. introduced an algorithm based on affine Gaussian scale space for high dynamic range images [11]. Anisotropic feature is extracted from the HDR images, and they are reformed to be isotropic using fitting and affine transformation. Then, base layer of HDR image is formed by dodging and burning processing. Detail layer of HDR image is formed by two-scale edge-preserving decomposition. Two layers are combined to get the final output image. This method works well for high contortion and stance changes in HDR images.

2.2 Bilateral Filters

Is an efficient filter, which smoothens the image while preserving its discontinuities. It is also capable of decomposing the image into two layers based on different scales. The pixels having similarity in photometric range are also considered close as the pixels occupying nearby spatial locations; this is rationale of bilateral filter, and this is the difference between bilateral and Gaussian filters. Bilateral filter [12, 13] is defined as follows:

$$B[I]_x = \sum G\sigma_s(||x - y||)G\sigma_r\left(I_x - I_y\right)I_y \tag{1}$$

BF[I]x is the normalized weighted average. If pixels y has intensity values different from I_x, influence of such pixels y is decreased by $G\sigma_r$, the range Gaussian. $G\sigma_s$ decreases influence of distant pixels; it is spatial Gaussian. Without parameter $G\sigma r$, bilateral filter becomes Gaussian filter.

where W_x is a normalization factor:

$$W_x = G\sigma_s(||x - y||)G\sigma_r\left(I_x - I_y\right) \tag{2}$$

Amount of filtering for the image is measured by parameters σ_s and σ_r, these parameters filtering. Only, by increasing σ_s, aggressive smoothing cannot be achieved; even, σ_r should be increased. As range Gaussian is flatter, bilateral filter becomes closer to Gaussian blur as σ_r increases that means ability to preserve edges is reduced. Larger features get smoothed when σ_s increases. Weighted average is calculated for each pixel, and pixels value is replaced. The range parameter imposes penalty on adjacent pixels with different intensities and spatial component imposes the penalty on the distant pixels. Combination of spatial and range components ensures pixel is scaled controls that final result is contributed by similar nearby pixels. Bilateral filter decomposes the image into bi-layers. (i) Base Layer: larger-scale components (ii) Detailed layer: smaller-scale components (residual of the bilateral the filter). Other than tone mapping, bilateral filter is also used in tone management, texture and illumination separation, denoising, flash/no flash imaging, and so on. Structures are also preserved in case of videos, here instead of filters are applied to log intensities of HDR image. HDR image is decomposition into bi-layers known as base and detailed layer. Scaled-down base layer is added with the detail layer. Using bilateral filter, multi-scale decomposition can be in performed in several ways. In one of the method spatial filtering, temporal bilateral filter is applied iteratively to the smoothed versions of the input image. To not blur the edges during recursive iterations during the coarsening process, range-Gaussian width is decreased. In other way, during the coarsening process, the spatial Gaussian can be progressively increased to not blur the edges. Smoothing filters like bilateral filters used in image decomposition algorithms treat the detail as low-contrast and extract the local variation in different contrast levels as successive the width of range layers of detail. Fine-scale spatial variation may not be necessarily represented by such layers. But, these filters are best tool for tone mapping as they extract details based on contrast. Drawbacks of bilateral filter are as follows: (i) As at each position, weighted sum as to be estimated over a large neighborhood, it consumes more time for execution. (ii) In the process of preserving the edges, it sharpens the some edges, incurring undesirable effect "stair-case effect." (iii) Predefined pixels neighborhood regions are required; determining them before hand is difficult as filtering performance is limited if a small regions have been selected, may result in cross-region mixing if large neighborhood is smoothing is used. In tone mapping process, the bilateral selected for regions having high textures [14]. (iv) Gradient reversal artifacts are produced along the few edges during detail enhancement process. [13, 15] and HDR compression, the unwanted components n_i from the input image p_i.(v) Involves trade-off between data smoothing and edge preservation. During process of separation of average surface from that of the detail surface, halo artifacts may be resulted at image edges if more smoothing is done. (vi) It is computationally expensive as it is nonlinear filter. Recursive version of bilateral filter was proposed by Yang [16], which defines photometric distance of bilateral filter as

$$pdBF(\boldsymbol{I_x}, \boldsymbol{I_y}) = \sqrt{\sum_{a,a+1\epsilon\phi} ||I_a + 1 - I_a||^2} \tag{3}$$

Here, Φ is the path which is predefined. It is the path connecting the pixels x and y. By adding the all the distances between adjacent pixels along a path Φ, the photometric distances are calculated. Ability to preserve edge becomes difficult, i.e., cross-region mixing is likely to occur though the computational cost is reduced. The Propagation filter [14] has the same goal as that of bilateral filter, and it works on the principle of photometric relationship existing between the image pixels. Cross-region mixing problem is relieved without using any spatial filtering functions. Preservation of image characteristics is better. In fast bilateral filtering [15], two acceleration techniques are used: (i) FFT and fast convolution can be used for computation as here bilateral filter is linearized. (ii) Key operations are decimated. Two strategies are used: (i) sub-sampling in spatial domain and (ii) in intensity domain, piece-wise linear approximation. Bilateral filter is also be further be used to obtain the joint bilateral filter [17], where rather than filtering, another guidance image is used to compute the input weights. This filter performance is better when it is difficult to extract the information on the edges of the images.

2.3 Guided Image Filtering

This is a smoothing filter used for edge preserving which has better behaviors near edges [18], and it does not have gradient reversal artifacts so best suited for detail enhancement and HDR compression. Guidance image is subjected to linear transformation which results in the filtered output, which is less smoothened and more structured than the input image. Regardless of the intensity range and kernel size, this filter exhibits relatively faster algorithm, and it is also non-approximate linear time. This filter is designed by the help of guidance image I, which is modeled with output image q. Let q be the linear transform of I in a window w_k centered at the pixel k.

Drawbacks of guided filter are unwanted smoothing of edges known as “halos” are exhibited near few edges. This is because image cannot be represented well near edges by the linear model used.

$$qi = x_k I_i + y_k \forall i \epsilon \omega_k \tag{4}$$

Here, the coefficients (x_k, y_k) are linear and constant in w_k. If image I has an edge, then q will have an edge; this is ensured by local linear model. Constraints from the filtering are needed to determine the (a_k, b_k). By removing output, q can be obtain.

$$qi = pi - ni \tag{5}$$

2.4 Weighted Guided Image Filter

In the case of WGIF [19], weighting is done considering the edge into the account, and it is added with guided image filter. From both local smoothing filter as well as global smoothing filters, it inherits advantages such as (i) avoids the halo artifacts, (ii) the computational complexity is same as Gaussian filter. To design WGIF, the variances of all intensity values are computed locally in the guidance images, and this result is used to calculate the normalized weighting by normalizing with the local variance of a pixel. Because of this, WGIF can preserve the sharp edges of the HDR image like that of global filters. And also, halo artifacts are reduced. Gradient reversal is avoided in WGIF. An weighting based on edge awareness is calculated first, and it is considered with GIF to come up the with new filter, the WGIF.

2.5 Gradient Domain Guided Image Filtering [GDGIF]

GDGIF is designed to preserve the edges better than the other variants of the guide image filters. In this case, the constraint with edge awareness and of explicit first order is included [20]. Even though weighted guided image filter was working with the principle of the edge awareness, but, to treat edges, there are no explicit constraints in WGIF and even in GIF. As they consider edge-preserving process and image filtering process together, edges cannot be preserved well in some cases. The output image of GDGIF will have gradient similar to that to the input image. The regularization term of GDGIF is different from that of GIF and WGIF as it includes a constraint with an explicit edge awareness.

GDGIF is based on local optimization. First-order regularization term and zeroth-order data fidelity term composes the cost function of GDGIF. Due to this, images are represent more accurately near edges by factors in new local linear model, i.e., preserves edges better than WGIF and GIF. Edge-aware factor in WGIF is single scale, whereas in GDGIF, it is multiscale. This helps it to separate the edges from other details of the image efficiently. The complexity of GDGIF is same of that of GIF and doesn't suffer from gradient reversal artifacts. In Fig. 1, the halo effects introduced by GIF are visible. Though WGIF reduces the effect, but still, they are visible, whereas GDGIF avoids the halo effect.

2.6 Robust-Guided Image Filtering [RGIF]

RGIF [21] consists of data term and a regularization term, to solve the challenges faced by guided image filter [18]. HDR tone mapping is achieved by the decomposition of layers by using RGIF [13, 15, 22]. After nonlinearly mapping the base layer, it is combined with the detailed layer. HDR tone mapping results based on different

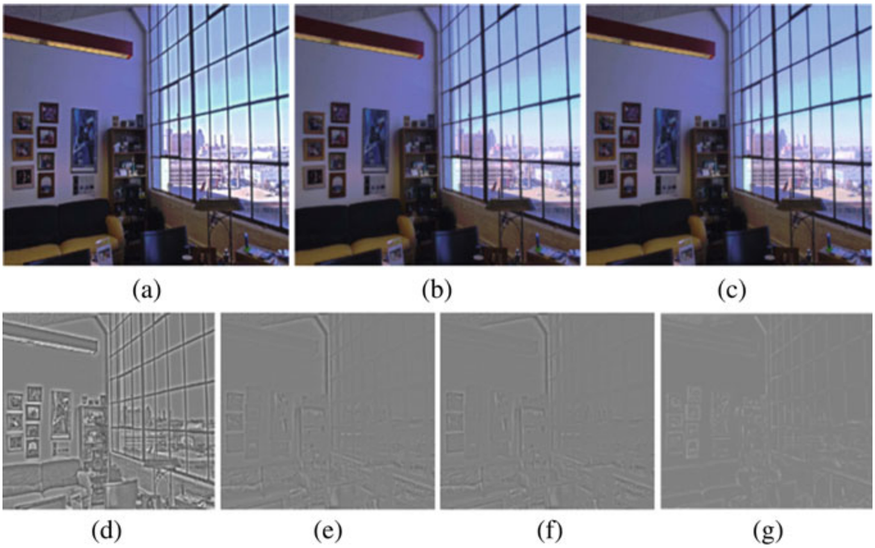


Fig. 1 HDR image "office.hdr" is used in comparison of tone mapping results [20]. **a** GIF output, **b** WGIF output, **c** GDGIF output, **d** detail layer of **a**, **e** detail layer of **b**, **f** detail layer of **c**, **g** subtracted values of **e**, **f** shows that GDGIF is better than WIF

frameworks are shown in Fig. 2. Result of RGIF frame work on HDR images is shown; it can be observed that edges are preserved.

2.7 Anisotropic-Guided Filter

Ochotorena and Yamashita [24] designed the new filter called as anisotropic-guided filter (AnisGF) to preserve the edges better, by making use of the weighted mean to achieve highest possible diffusion. This is designed mainly to take of the detail halos in the output image. Guided filter and its variants perform poor when the both the input and guided images have the inconsistency in their structural features. This is mainly because of the reason that those filters do not use the weighted mean during the final steps. Ochotorena and Yamashita et al., suggested to use the weighted mean to obtain the highest possible diffusion.

2.8 Deep Neural Networks-Based Filters

Zhu et al. came up the solutions such as (i) performance of the algorithms when there is vast span of image parameters using a same setting, (ii) edge-preserving evaluation which is mostly based on the subjective process, (iii) lack of datasets. Here, to set

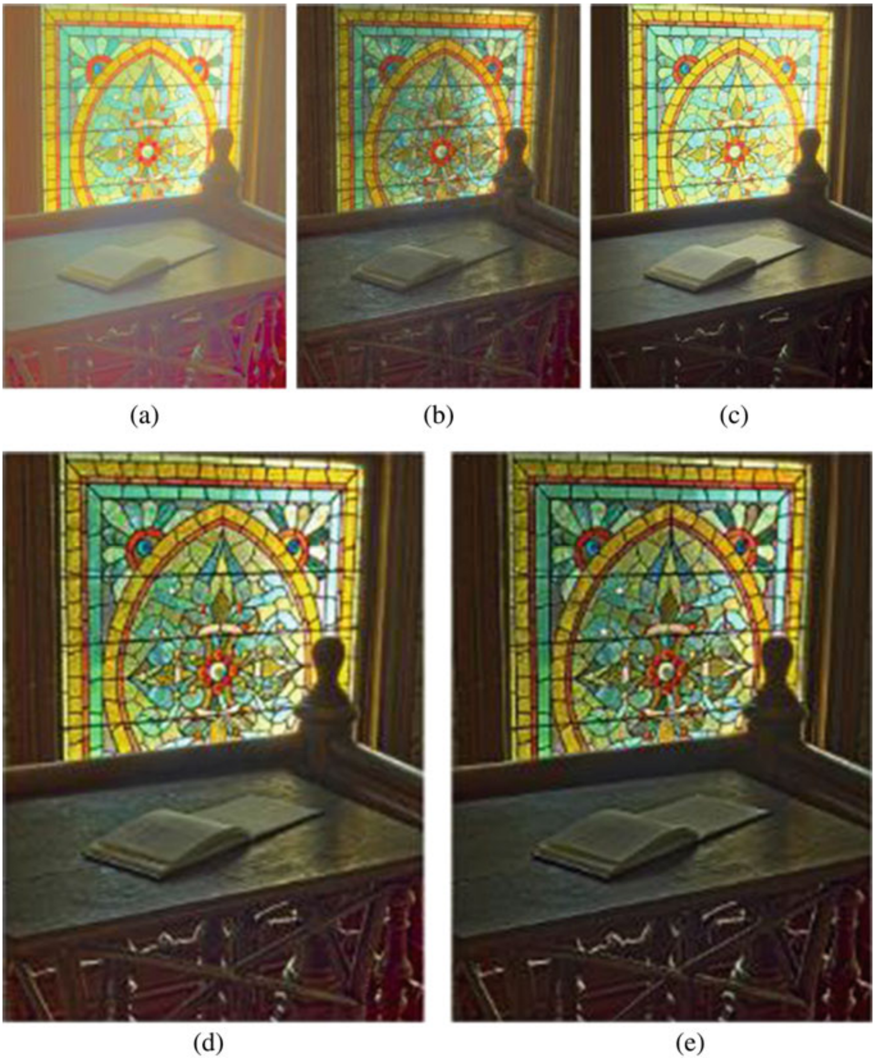


Fig. 2 High dynamic range tone mappers compared: **a** High dynamic range image, **b** L_0 Gradient [4], **c** Tone reproduction [23], **d** Weighted least squares [10], **e** RGIF [21]. In most cases of guided image filtering, texture copy should be avoided. The limitation of RGIF is its ability of suppressing texture copy

up the benchmark, deep neural networks are used. Deep neural networks consist of weights (parameters) in larger number [25]. In other algorithms, even though the weights are used as per the awareness of the edges, they were not trained. Here, the weights are trained, and hence, performance is better even though there exists the different image components in the used datasets. Existing network architecture, deep residual networks (ResNets) are made used here. As ResNet baseline model is trained datasets constructed with criteria such that they can handle the task of avoiding halo artifacts by preserving the dynamic edges. Other part of image will be similar to the output of low-pass filtering whatever be the components of the images. The same tone mapping frame work in [15] is used here by just using the ResNet model in the place of the bilateral filtering model. When compared with other filtering results, such as that of bilateral filtering [15], local edge-preserving filter (LEP) [1], visual adaptation (VAD) [26] tone mappers, ResNet gives better results as shown in Fig. 3 and Table 1.

Fig. 3 Results of different tone mappers. From summit to foot: the bilateral filter, VAD method, LEP method, and ResNet model. Bilateral filter introduces halo artifacts at the edges across the tree in the summit image. Information are lost in case of VAD method. The local edge preserving filter method results are not quit natural look. ResNet model conserves the edge information

Table 1 Tone-mapped image quality index (TMQI) [29] comparisons: Allegiance of structures and naturalness based on image statistics are measured by TMQI

	Bilateral filtering	Visual adaptation	Local edge-preserving filter	ResNet
Tone-mapped image quality index	0.902000	0.904100	0.875000	0.909500

2.9 Domain Transformation

Eduardo S L Gastal and Manuael M Oliveira came up with the new method for edge-preserving filtering [27] based on the edge-preserving filter by Barash et al. [28]. This is based on the isometric between the two dimensional curves manifold in the five-dimensional space and the real line. Here, the geodesic distances are maintained same among the points, meanwhile the input signal is warped adaptively in such a way that in linear time, one dimensional edge-preserving filtering can be performed. This method is faster comparatively as it uses one-dimensional techniques, and also, the memory utilization is less. Computational cost is low, and it can operates on color images at variable scales.

Tone mappers with edge awareness can operate well to preserve the edges avoiding the halo artifacts while converting HDR image to its low dynamic version. Figure 4 provides the comparison of the outputs tone mappers such as (a) recursive filter (b)

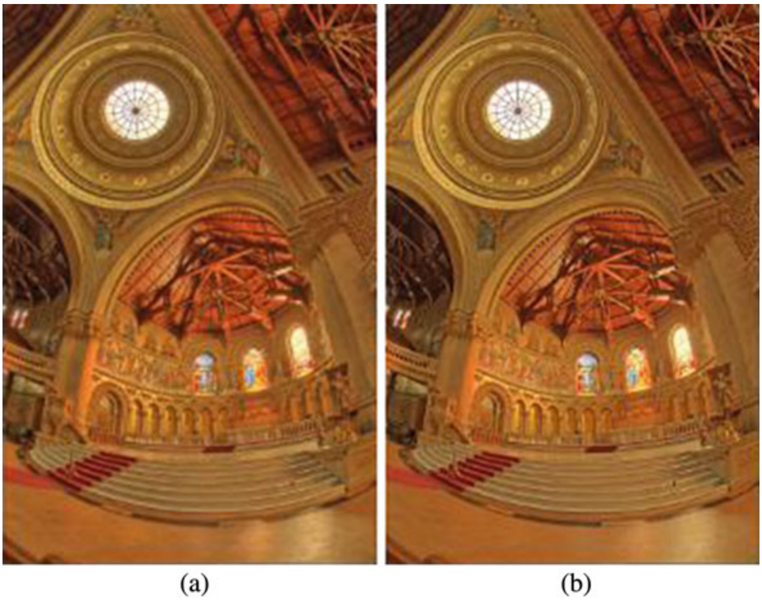


Fig. 4 Results and comparison of different tone mappers: **a** Recursive filter and **b** Weighted least square filter

weighted least square filter. Though the image quality looks same, recursive filter is faster comparatively.

2.9.1 Multi-scale Decomposition-Based Filter

The new method of edge preserving was proposed, where the image is subjected into multiscale decomposition [1]. This filter works on the adaptability principle applied locally on the image. The output image of the filter preserves the edges locally, and also, local mean is maintained everywhere. Here, image is decomposed into three detail layers and one base layer where base layer contains the local mean and edges are conserved in detail layer. This method yields good result in preserving and enhancing local details of the HDR image as per Fig. 5.

3 Global Optimization-Based Filters

Global filter generates filter image by utilization all the pixel information in the image. They are formulated by using an optimization problem which places relationship among the neighboring pixels. As edge-preserving smoothing filters operated locally results in the halo artifacts, the global optimization-based filters are introduced. In case of the globally optimized filters [2, 4, 6, 24], optimization performance criterion will contain two terms, namely regularization term and data term. The smoothness level of the reconstructed image is provided by regularization term, while fidelity of reconstructed image is measured by data term. Global optimization-based filters have high computational cost though they yield very good quality. In case bilateral filter, its spatial and range parameters are fixed. In case of guided image filter, Lagrangian factor is fixed. Whereas, Lagrangian factor varies in case of global optimized filter. For example, in case of weighted least square filter [10], Lagrangian factor is adaptive. This could be major reason for the halo artifacts in bilateral and guided image filters. In the case of adaptive bilateral filter (ABF) [30], which is a training-based approach, range similarity parameter is adaptive. In adaptive filter of [31], range similarity parameter and spatial similarity parameter both are adaptive. But, 3D convolution form is destroyed by adaptation of the parameters.

3.1 Weighted Least Square Filters [WLS]

Bilateral filter-based detail decomposition techniques are less appropriate for the multiscale decomposition of the HDR images as they are not capable to extract detail at variable scales. To perform the multiscale decomposition, details should be extracted in variable scales. WLS is based on the optimization of weighted least squares [10], specifically well suited for (i) multiscale detail extraction (ii)

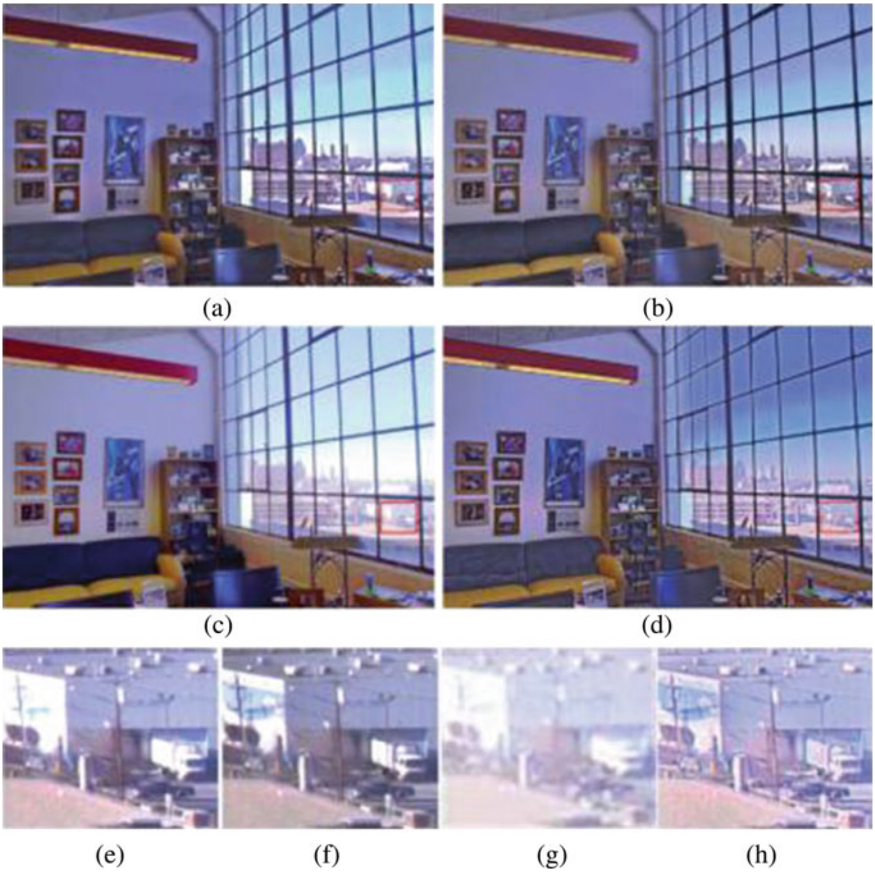


Fig. 5 Reconstruction of HDR images using various algorithms: Images in the red rectangle are enlarged and shown. **a** Output of bilateral filtering [15]. **b** Output of weighted least square filtering. **c** Output of local extrema [22]. **d** Output of local edge-preserving filter [1, 10]. **e** close-up of **a**. **f** close-up of **b**. **g** close-up of **c**. **h** close-up of **d**

for progressive coarsening of images. Edge-preserving smoothing is compromise between smoothing the image everywhere and to retain significant gradients of the image [10]. Disadvantage of WLS is its computational cost.

3.2 L_0 *Gradient Minimization*

This is the global optimization method, helpful for the preserving HDR image edges [4]. This can have control on large count of non-zero gradients in a sparsity control manner to approximate prominent structure. By limiting count of non-zero gradients, edges are strengthened. Smoothing is not performed local; as in the case of local edge-preserving filters, it is done globally. This helps to achieve better smoothing during decompositions of layers in the tone-mapping process and also preserves and strengthens the structures of the images. Limitation of this framework is, in some unavoidable challenging circumstances to remove details, oversharpening occurs.

3.3 Fast Global Image Smoothing

Fast global smoother [3] addresses the limitations of the local optimized edge-preserving filters. This method achieves inhomogeneous spatially while preserving the edges. By the help of the inhomogeneous laplacian matrix, which is specified over the multidimensional spatial domain, the solution for the linear system is obtained (Table 2).

This is fast, compare to the other global optimization filters and halos effects are not found. The advantages of the filter help in the faster and finer tone mapping of HDR image. The metric, Mean Square Error (MSE) calculates the cumulative error square between the modified and input images.

3.4 Scale-Aware Edge-Preserving Image Filtering

Filtering can be obtained by iterative global optimization (IGO) [33] which provides scale awareness and preservation of edges of the HDR images. By the usage of IGO method, with that of the measurement including the scare awareness, it is shown that gradients can be suppressed on lesser-scale characteristics, and variations in the intensities are preserved. But still, this one suffers to preserve the edges in case of the complex-structured images.

3.5 Embedding Bilateral Filter in Least Squares

As discussed, global optimization yields the better results over the local filters but with the disadvantage of the computational cost and usually run slower than the local filters. In this filter [34], the advantages of local and global filters are combined as bilateral filter is embedded least squares. This results in better edge preserving of the HDR image by avoiding halos artifacts and gradient reversals. This filter is found almost ten times faster than weighted least square filters. The tone mapping is done according to the multiscale decomposition. The gradient reversal and halos artifacts problems in [12, 35] are overcome by this method.

Table 2 Comparison table

Ref. No.	Paper title	Outcome
[32]	Computer vision: algorithms and applications	• Provides basics of the HDR imaging such as – Finding the radiometric response function after the alignment of the multi-exposure images – Construction of the radiance map using the multi-exposure pixels – Performing the tone mapping to transform high dynamic range image to low dynamic range image • Discusses various HDR image formats • Discusses the application of HDR images
[2]	Nonlinear total variation-based (TV) noise removal algorithms	• It presents a method for removal of the image noise by applying the constraints using Lagrange multipliers • The method is non-invasive, yielding sharp edges
[3]	Fast global image smoothing based on WLS	• This method addresses the limitations of the locally optimized edge-preserving filters • This achieves inhomogeneous spatially while preserving the edges. By the help of the inhomogeneous Laplacian matrix, which is specified over the multi-dimensional spatial domain, the solution for the linear system is obtained
[4]	Image smoothing via L0 gradient minimization	• Optimization frame work constructed using this method can control as many as non-zero gradients globally • Limitation of this framework is, in some unavoidable challenging circumstances to remove details Over-sharpening occurs
[5]	Edge-preserving image enhancement using anisotropic diffusion	• Anisotropic diffusion considers each pixel of image as an energy sync interacting with neighbor pixels based upon pixel intensity differences • Anisotropic diffusion decomposes the image into bi-layer where base layer is the smooth layer and detail layer is difference layer • Limitations of anisotropic diffusion are (i) “stair-casing” effect around smooth edges. (ii) It is a slow process due to its discrete diffusion nature

(continued)

Table 2 (continued)

Ref. No.	Paper title	Outcome
[6]	Visualization using rational morphology and zonal magnitude-reduction	• Here, the new methods are proposed for detecting the objects and reducing the noise in them image • It can be performed by either by changing the magnitude as per the frequency regions or by utilizing several magnitude changes with different scales
[7]	Transform coefficient histogram-based image enhancement algorithms using contrast entropy	• This paper presents the image enhancement techniques such as logarithmic transform, histogram matching, histogram shifting, and histogram shaping
[8]	LCIS a boundary hierarchy for detail-preserving contrast reduction	• Presents the technique to obtain the conductance values differently to maintain the steady ramp and edges
[9]	Non-linear image filtering with edge and corner enhancement. IEEE transactions on pattern analysis and machine intelligence	• A nonlinear image filtering for noise and reduction and enhancement of edges using anisotropic diffusion is described • The corners and T junctions are as enhanced with the edges
[10]	Edge-preserving decompositions for multiscale tone and detail manipulation	• WLS is based on the weighted least squares optimization • Performs the multiscale detail extraction and good for progressive coarsening of images • Edge-preserving smoothing is compromise between smoothing the image everywhere and to retain significant gradients of the image
[11]	Affine invariant features-based tone mapping algorithm for high dynamic range images	• Anisotropic feature is extracted from the HDR images, and they are reformed to be isotropic using fitting and affine transformation • Multi-layers of HDR image are formed by dodging and burning processing and two-scale edge-preserving decomposition. Two layers are combined to get the final image
[12]	Bilateral filtering for gray and color images	• This paper presents the bilateral filtering for both gray and RGB images. This helps in the decomposition of the HDR image

(continued)

Table 2 (continued)

Ref. No.	Paper title	Outcome
[13]	A gentle introduction to bilateral filtering and its applications	• .Bilateral filter is an efficient filter, which smoothens the image while preserving its discontinuities • It is also capable of decomposing the image into two layers based on different scales • The pixels having similarity in photometric range are also considered close as the pixels occupying nearby spatial locations; this is rationale of bilateral filter, and this is the difference between bilateral and Gaussian filters
[14]	Propagated image filtering	• Propagated image filtering as the same goal as bilateral filter and is dependent on photometric nature noticed among the image pixels • Preservation of image characteristics is better • Predefined pixels neighborhood regions are required; determining them before hand is difficult as filtering performance is limited if a small regions have been selected, may result in cross-region mixing if large neighborhood is selected for regions having high textures
[15]	Fast bilateral filtering for the display of high dynamic range images	• Here, two acceleration techniques are used: (i) Fast Fourier transform and fast convolution can be performed for computation as here bilateral filter is linearized. (ii) Key operations are decimated • Two strategies are used: (i) sub-sampling in spatial domain and (ii) in intensity domain, piece wise linear approximation
[16]	Recursive bilateral filtering	• Recursive bilateral filtering defines photometric distance of bilateral filter • The photometric distances are calculated by adding on the distances between adjacent intensity values over the path Φ • Ability to preserve edge becomes difficult
[17]	Digital photography with flash and no-flash image pairs	• Bilateral filter is also generalized to joint bilateral filter, where rather than filtering, input weights are obtained by another guidance image • This filter performance is better when it is difficult to extract the details of edges

(continued)

Table 2 (continued)

Ref. No.	Paper title	Outcome
[18]	Guided image filtering	• This edge-preserving filter conserves the edges • This overcomes the effect of gradient reversal artifacts (artifacts of unwanted sharpening of edges) • Best suited for enhancing the details of the images and for compressing the HDR images. Linear transformation of the guidance image results in the filtered output, which is less smoothened and more structured than the input image • This can be grouped among the fast locally optimized edge-preserving filters
[19]	Weighted guided image filtering	• The filter incorporates an weights with awareness of edges for the present guided image filter • From both locally optimized and globally optimized smoothing filters, it inherits advantages such as (i) avoids the halo artifacts, (ii) the complexity is same as Gaussian filter
[20]	Gradient domain-guided image filtering	• An explicit constraint with edge awareness and that of first order is included • Filter, thus, conserves edges better than other similar filters
[21]	Robust-guided image filtering	• This filter consists of two terms, namely data term and a regularization term and solves the challenges faced by guided image filter • Re-weighted least squares (IRLSs) are used iteratively as model obtained is highly non-convex
[22]	Edge-preserving multiscale image decomposition based on local extrema	• Provides new technique for the multiscale decomposition of the HDR image by define details are fine-scale spatial oscillations and edges as high variance in spanning values of adjacent local extrema
[24]	Anisotropic-guided filtering	• This filter preserves the edges in the image, by utilizing weighted averaging to achieve maximum diffusion • AnisGF brings anisotropy as an added capability to the existing guided filter • AnisGF is immune to the formation of the artifacts, especially with density dependence it presents some limitations

(continued)

Table 2 (continued)

Ref. No.	Paper title	Outcome
[25]	A benchmark for edge-preserving image smoothing	• Deep neural networks have many weights (parameters). The weights can be fixed after the training process and resulting network can work with various types of inputs with very strong generalization capability • Existing network architecture, deep residual networks (ResNets) are made used
[1]	Local edge-preserving multiscale decomposition for high dynamic range image tone mapping	• Due to its locally adaptive property, it is different from other filters • The output contains local means throughout and conserves local salient edges
[26]	An analysis of visual adaptation and contrast perception for tone mapping	• As per the technique presented, tone mapping is done using both global and local optimization method • Global method is utilized to obtain the visual adaptability, and contrast is locally enhanced
[29]	Objective quality assessment of tone-mapped images	• Algorithm is proposed based on objective quality assessment for tone-mapped images by combination of: (1) a multiscale signal fidelity measure and (2) a naturalness measure
[27]	Domain transform for edge-aware image and video processing	• This is based on the isometric between the two-dimensional curves manifold in the five-dimensional space and the real line • Here, the geodesic distances are maintained same between the points, meanwhile the input signal is warped adaptively in such a way that in linear time, one-dimensional edge-preserving filtering can performed • This method is faster compared to other similar methods as it uses one dimensional techniques, and also, the memory utilization is less
[28]	Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation	• Here, anisotropic diffusion is compared with the bilateral image filter technique • Adaptive smoothing is fixed, and it acts as the medium in between the anisotropic diffusion and bilateral image filter technique

(continued)

Table 2 (continued)

Ref. No.	Paper title	Outcome
[30]	Adaptive bilateral filter for sharpness enhancement and noise removal	• Adaptive bilateral filter rises the edges slope without producing overshoot or undershoot and thereby sharpens the images
[31]	Content adaptive bilateral filtering	• The edges are preserved better than the existing bilateral filter, and also, flat regions are smoothened better
[33]	Scale-aware edge-preserving image filtering	• Filtering can be obtained by iterative global optimization (IGO), which provides scale awareness and preservation of edges of the HDR images • By the usage of IGO method, with that of the measurement including the scare awareness, it is shown that gradients can be suppressed on lesser scale characteristics, and variations in the intensities are preserved • But still, this one suffers to preserve the edges in case of the complex structured images
[34]	Embedding bilateral filter in least squares	• The benefits of locally and globally optimized filters are combined as bilateral filter in embedded least squares • This results in better edge preserving of the HDR image by avoiding halos artifacts and gradient reversals • This filter is found almost ten times faster than weighted least square filters • In tone mapping, this filter can properly eliminate the artifacts, and results are comparable with that of WLS
[35]	Adaptive manifolds for real-time high-dimensional filtering	• Here, powerful method for computing the high-dimensional filtering in real time is presented. • It is faster, and less memory is required for the storage

4 Methodology

The block diagram of the multiscale decomposition of HDR images using the edge-preserving filters is shown in the 6. The high dynamic range image is the input to the edge-preserving filters such as bilateral filter, guided filter, and weighted least square filters. Edge-preserving filters smoothens the HDR image preserving the edges. The output of the edge-preserving filter is taken as the base layer, and the difference between the base layer and input image results in the detail layer, which preserves the edge details. Further, base layer is subjected to the multiscale smoothening using the edge-preserving filters. Here, in our experimentation, edge-preserving filters are used three times for the multiscale smoothening. Further, the smoothen image (final base layer) is added with enhanced detail layer to obtain the output HDR image, there by image is smoothen; hence, memory size is reduced by retaining the edge information. The following metric is employed to compare the output of the edge-preserving filters.

4.1 Peak Signal-to-Noise Ratio (PSNR)

PSNR is a metric in signal processing which measures quality of a signal by calculating ration between original signal and the noise signal. PSNR estimates the image quality by comparing the modified image with the input image. PSNR will have higher value for evenly distributed values rather than sparsely distributed MSE and PSNR are defined as follows:

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - y_{ij})^2$$

m number of rows in cover image
n number of columns in cover image
x_{ij} pixel value from cover image
y_{ij} pixel value from stego image

$$\text{PSNR}(x, y) = \frac{10\log_{10}[\max(\max(x), \max(y))]^2}{x - y^2}$$

4.2 mPSNR

mPSNR is a metric which computes the PSNR for mainly HDR images. It takes following input arguments: input reference image, input distorted image, the minimum exposure for computing mPSNR, and the maximum exposure for computing mPSNR. It provides the following output arguments. mpsnr: the multiple-exposure PSNR value. Higher values mean better quality. eMax: the maximum exposure for computing mPSNR. eMin: the minimum exposure for computing mPSNR.

4.3 HDR-VDP-3

HDR-VDP-3 compute visually significant differences between an image pair with following input arguments: task—the task for which the metric should predict. Test—image to be tested (e.g., with distortions) reference—reference image (e.g., without distortions) color encoding—color representation for both input images. Pixels per degree—visual resolution of the image. The function returns a structure with the following fields: P_{map}—probability of detection per pixel (matrix 0–1) P_{det}—a single-valued probability of detection (scalar 0–1) C_{map}—threshold normalized contrast map so that $C_{\text{max}} = 1$ corresponds to the detection threshold ($P_{\text{det}} = 0.5$). C_{max}—maximum threshold normalized contrast so that $C_{\text{max}} = 1$ corresponds to the detection threshold ($P_{\text{det}} = 0.5$). Q—quality correlate, which is 10 for the best quality and gets lower for lower quality. Q can be negative in case of very large differences. QJOD—quality correlate in the units of just objectionable differences. 1 JOD means that 75 reference image over a test image in a 2-alternative-forced-choice test (Fig. 6).

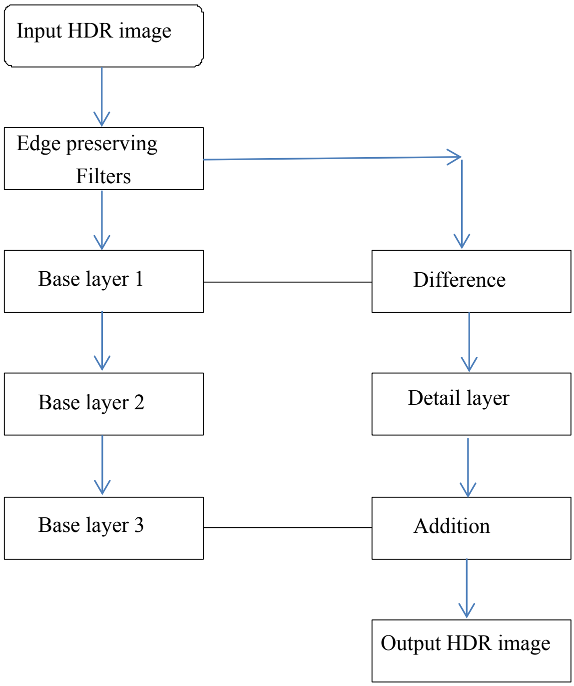


Fig. 6 Block diagram: multiscale decomposition of HDR images using the edge-preserving filters

5 Results and Discussion

Multiscale decomposition of the HDR image was performed on Matlab R2020a using the edge-preserving filters such as bilateral filter, guided filter, and WLS filter. Images were compared with the metrics psnr, mPSNR, and HDRVDP3. The results are as shown below in Tables 3, 4, 5, and 6. Output was evaluated on the two different images “smallOffice.hdr” and “SpheronNiceo9E0.hdr” The table shows the psnr and mPSNR values of the base layer and detail layer compared with the input HDR image. A further table specifies the output arguments of the HDRVDP3, where the quality of the image is given the preferences. The quality of the detail layer is found to be good in the WLS filter.

6 Conclusion

In this paper, filters used for multiscale decomposition of HDR image into base layer and detail layer are performed using the edge-preserving filters. There are two types of filter i) Local filtering-based edge-preserving smoothing filters, includes anisotropic filters and its variants, bilateral filters and its variants, guided filter and its variants. Anisotropic and bilateral filters are slow, and they suffer from staircasing effect. Though guided filter is fast, it suffers from halo effects like all other local filtering-based edge-preserving smoothing filters. AnisGF brings anisotropy as an added capability to the existing guided filter, but still immune to the formation of the artifacts and it presents limitations with density dependence. ii) Global optimization-based filters like WLS are best in edge preserving and smoothing, but drawback is computational cost. WGIF and GDGIF take the advantage of both local filtering-based edge-preserving filters and global optimization-based filters. They preserve the edges better, and the complexity is same as Gaussian filter. Edge-aware factor in WGIF is single scale, whereas in GDGIF, it is multiscale, so it is better in preserving the edges of the HDR image. The embedded BF [34] has the advantages of local and global filters. This results in better edge preserving of the HDR image by avoiding halos artifacts and gradient reversals. This filter is found almost ten times faster than weighted least square filters. So, it can be concluded that filter incorporating the features from both local and global smoothing filters and which performs multiscale decomposition is best suited for HDR image decomposition and tone mapping. Research to be focused is on implementation of an efficient filter with the both local and global optimization features for the multiscale decomposition. The contrast of base layer can be compressed efficiently with a new compression technique, and detail layer can be enhanced further to preserve the edge information. A new frame work can be designed for tone mapping process.

Table 3 Base layer 3 of "smallOffice.hdr"

Filter method	HDR VDP3							PSNR	mPSNR
	P_{map}	S_{Max}	P_{det}	$C_{\max}$	C_{map}	Q	QJOD		
Bilateral filter	[525 × 800 double]	1.69	0.6	2.17	[525 × 800 double]	9.49	9.78	32.67	14.93
Guided filter	[525 × 800 double]	7.8087	0.9715	8.7098	[525 × 800 double]	8.89	9.41	29.60	11.87
WLS filter	[525 × 800 double]	546.0878	0.9999	582.897	[525 × 800 double]	3.60	4.39	−31.95	−4.67

Table 4 Detail layer of "smallOffice.hdr"

Filter method	HDR VDP3							PSNR	mPSNR
	P_{map}	S_{Max}	P_{det}	$C_{\max}$	C_{map}	Q	QJOD		
Bilateral filter	[525 × 800 double]	9661.10	0.99	12,126.00	[525 × 800 double]	3.00	3.71	−2.91	−5.19
Guided filter	[525 × 800 double]	4827.70	1.00	5846.70	[525 × 800 double]	3.05	3.76	−2.89	−5.00
WLS filter	[525 × 800 double]	11,044.00	1.00	12,201.00	[525 × 800 double]	1.98	2.51	−32.25	−5.97

Table 5 Base layer 3 of "SpheronNiceo9E0.hdr"

Filter method	HDR VDP3							PSNR	mPSNR
	P_{map}	S_{Max}	P_{det}	$C_{\max}$	C_{map}	Q	QJOD		
Bilateral filter	[1165 × 2981 double]	0.13	0.08	0.15	[1165 × 2981 double]	9.91	9.97	49.57	28.05
Guided filter	[1165 × 2981 double]	7.72	0.97	8.95	[1165 × 2981 double]	9.19	9.61	41.04	21.55
WLS filter	[1165 × 2981 double]	3681.50	1.00	4848.40	[1165 × 2981 double]	4.14	4.99	−27.01	−4.75

Table 6 Detail layer 3 of "SpheronNiceo9E0.hdr"

Filter method	HDR VDP3							PSNR	mPSNR
	P_{map}	S_{Max}	P_{det}	$C_{\max}$	C_{map}	Q	QJOD		
Bilateral filter	[1165 × 2981 double]	412,840.00	1.00	432,560.00	[1165 × 2981 double]	1.41	1.82	−28.22	−3.71
Guided filter	[1165 × 2981 double]	230,720.00	1.00	255,580.00	[1165 × 2981 double]	1.64	2.10	−28.22	−3.49
WLS filter	[1165 × 2981 double]	7678.10	1.00	8636.30	[1165 × 2981 double]	4.66	5.55	−27.15	−3.97

References

1. Gu B, Li W, Zhu M, Wang M (2012) Local edge-preserving multiscale decomposition for high dynamic range image tone mapping. IEEE Trans Image Process 22(1):70–79
2. Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D 60(1–4):259–268
3. Min D, Choi S, Lu J, Ham B, Sohn K, Do MN (2014) Fast global image smoothing based on weighted least squares. IEEE Trans Image Process 23(12):5638–5653
4. Xu L, Lu C, Xu Y, Jia J (2011) Image smoothing via l0 gradient minimization. In: Proceedings of the 2011 SIGGRAPH Asia conference, pp 1–12
5. Wharton EJ, Panetta KA, Agaian SS (2008) Edge preserving image enhancement using anisotropic diffusion. In: Imageprocessing: algorithms and systems VI, vol. 6812. International Society for Optics and Photonics, p 681218
6. Kogan RG, Agaian SS, Lentz KP (1998) Visualization using rational morphology and zonal magnitude reduction. In: Nonlinear image processing IX, vol 3304. International Society for Optics and Photonics, pp 153–163
7. Agaian SS, Silver B, Panetta KA (2007) Transform coefficient histogram-based image enhancement algorithms using contrast entropy. IEEE Trans Image Process 16(3):741–758
8. Tumblin J, Turk G (1999) Lcis: A boundary hierarchy for detail preserving contrast reduction. In: Proceedings of the 26th annualconference on Computer graphics and interactive techniques, pp 83–90
9. Nitzberg M, Shiota T (1992) Nonlinear image filtering with edge and corner enhancement. IEEE Trans Pattern Anal Mach Intell 8:826–833
10. Farbman Z, Fattal R, Lischinski D, Szeliski R (2008) Edge preserving decompositions for multi-scale tone and detail manipulation. ACM Trans Gr (TOG) 27(3):1–10
11. Chen Q, Li Q, Ding Y, Liu C, Wang J, Deng X (2014) Affine invariant features-based tone mapping algorithm for high dynamic range images. In: 2014 IEEE international conferenceon systems, man, and cybernetics (SMC). IEEE, pp 2373–2378
12. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Sixth international conference on computervision (IEEE Cat. No. 98CH36271). IEEE, pp 839–846
13. Paris S (2007) A gentle introduction to bilateral filtering and its applications. In: ACM SIGGRAPH 2007 courses, p 3
14. Rick Chang J-H, Frank Wang Y-C (2015) Propagated image filtering. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 10–18
15. Durand F, Dorsey J (2002) Fast bilateral filtering for the display of high-dynamic-range images. In: Proceedings of the 29th annual conference on computer graphics and interactive techniques, pp 257–266
16. Yang Q (2012) Recursive bilateral filtering. In: European conference on computer vision. Springer, pp 399–413
17. Petschnigg G, Szeliski R, Agrawala M, Cohen M, Hoppe H, Toyama K (2004) Digital photography with flash and no-flash image pairs. ACM Trans Gr (TOG) 23(3):664–672
18. He K, Sun J, Tang X (2012) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 35(6):1397–1409
19. Li Z, Zheng J, Zhu Z, Yao W, Wu S (2014) Weighted guided image filtering. IEEE Trans Image Process 24(1):120–129
20. Kou F, Chen W, Wen C, Li Z (2015) Gradient domain guided image filtering. IEEE Trans Image Process 24(11):4528–4539
21. Liu W, Chen X, Shen C, Yu J, Wu Q, Yang J (2017) Robust guided image filtering. arXiv preprint arXiv:1703.09379
22. Subr K, Soler C, Durand F (2009) Edge-preserving multiscale image decomposition based on local extrema. ACM Trans Gr (TOG) 28(5):1–9
23. Reinhard E, Stark M, Shirley P, Ferwerda J (2002) Photographic tone reproduction for digital images. In: Proceedings of the 29th annual conference on Computer graphics and interactive techniques, pp 267–276

24. Ochotorena CN, Yamashita Y (2019) Anisotropic guided filtering. IEEE Trans Image Process 29:1397–1412
25. Zhu F, Liang Z, Jia X, Zhang L, Yu Y (2019) A benchmark for edge-preserving image smoothing. IEEE Trans Image Process 28(7):3556–3570
26. Ferradans S, Bertalmio M, Provenzi E, Caselles V (2011) An analysis of visual adaptation and contrast perception for tone mapping. IEEE Trans Pattern Anal Mach Intell 33(10):2002–2012
27. Gastal ES, Oliveira MM (2011) Domain transform for edge-aware image and video processing. In: ACM SIGGRAPH 2011 papers, pp 1–12
28. Barash D (2002) Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. IEEE Trans Pattern Anal Mach Intell 24(6):844–847
29. Yeganeh H, Wang Z (2012) Objective quality assessment of tone-mapped images. IEEE Trans Image Process 22(2):657–667
30. Zhang B, Allebach JP (2008) Adaptive bilateral filter for sharpness enhancement and noise removal. IEEE Trans Image Process 17(5):664–678
31. Li Z, Zheng J, Zhu Z, Yao W, Wu S, Rahardja S (2013) Content adaptive bilateral filtering. In: 2013 IEEE internationalconference on multimedia and expo workshops (ICMEW). IEEE, pp 1–6
32. Szeliski R (2010) Computer vision: algorithms and applications. Springer Science & Business Media
33. Zhou Z, Wang B, Ma J (2017) Scale-aware edge-preserving image filtering via iterative global optimization. IEEE Trans Multimedia 20(6):1392–1405
34. Liu W, Zhang P, Chen X, Shen C, Huang X, Yang J (2018) Embedding bilateral filter in least squares for efficient edge-preserving image smoothing. IEEE Trans Circ Syst Video Technol
35. Gastal ES, Oliveira MM (2012) Adaptive manifolds for real-time high-dimensional filtering. ACM Trans Gr (TOG) 31(4):1–13

Sustainable Wastewater Management Using Algae in Oxidation Ponds

Pallavi Praveen and Tejaswini Singh

Abstract The oxidation pond treatment process is natural because it uses microorganisms such as bacteria and algae. This makes the method of treatment cost-effective in terms of its construction, maintenance, and energy requirements. Oxidation ponds are also productive because they generate effluents that can be used for other applications. Finally, oxidation ponds can be considered a sustainable method for treatment of wastewater.

Keywords Oxidation pond · Wastewater · Waste management and treatment · Algae · Bacteria

1 Introduction

Technologies based on natural systems of waste management and treatment having distinct advantages of being more eco-friendly, with minimum use of mechanical equipment, are being explored in countries like ours for sustainable waste management. Wastewater management has received considerable attention in recent years to move the stringent regulation of the statutory bodies in addition to water scarcity. Proper collection treatment and disposal/utilization of waste water of domestic and industrial origin are essential prerequisites for the protection of public health and improvement in quality of life. Among the various treatment routes available for renovation of wastewater, natural systems form a key component. The main underlying principle of operation of this system is its dependence on natural components for achieving intended purposes.

Oxidation ponds or waste stabilization ponds have been found to be a simple and reliable means of treatment of sewage and certain industrial wastes.

P. Praveen (✉)
Bokaro Steel City College, Bokaro, India
e-mail: pallavipraveen8@gmail.com

T. Singh
Birla Institute of Technology, Mesra, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_49

Some major advantages of oxidation ponds are as follows:

- Low capital requirement (except for land), low operation and maintenance cost. Ability to handle fluctuating organic and hydraulic loads.
- Climatic advantage in warm countries like India.
- Its operation is very simple and without quiring special equipment and speaking skills high BOD is removed.
- Potential use as reservoir for irrigation and fish culture.

An oxidation pond is an open shallow basin excavated in the ground used for secondary treatment (microbial broad production of wastewater. Liquid waste is focused in these ponds and retained for specific periods (15–30 days) to permit desired intervention. The natural combination oxygen using bacteria and oxygen producing algae in such ponds convert organic and inorganic waste into stable secondary effluent that can be released into the receiving waste body.

The aerobic biodegradation process taking place in an oxidation pond can be represented by the following generic equation $OC_xH_y + H_y +$ (microorganisms/nutrients) $\rightarrow H_2O + CO_2, +$ biomass Keeping the potentialities of oxidation pond in wastewater management, the present investigation was undertaken on the ecological role of algae in an oxidation pond situated Dhandobra Bokaro. An attempt has been made to establish a relationship between phytoplankton and waste material. Detailed integral investigations is required two consecutive years.

2 Result and Discussion

The ecological role of algae in the functioning of oxidation ponds has been well recognized. Algae play a dual role of oxygenation as well as stabilization of wastes in oxidation ponds. The determination of the uber and other kinds of the more abundant algae in oxidation ponds can be used as a reliable index of the progress achieved in the oxidation of wastes.

Altogether, 99 taxa belonging to *Chlorophyceae (46), Cyanophyceae (30) Bacillariophyceae (16), and Euglena* (7) were recorded during the entry course of investigation (Table 1). Percentage composition and relative abundance of those classes

Table 1 Course of investigations

Class	Genera	Species	Percentage
Chlorophyceae	25	46	46.5
Cyanophyceae	11	30	30.33
Bacillariophyceae	8	16	16.1
Euglenineane	3	7	7.07

Source Survey results at Dhandobra oxidation pond through Chas and lab facilities show the chlorophyceae were the most dominant, followed by blue-greens and diatoms. Euglenoids were poorly represented.

A total of 16 species were recorded from the mixing site. Out of these, two species belonging to Cyanophyta (*Oscillatoria subbrevis* and *Phormidium luridum*) and one of Euglena (*Phacus longicauda*) were recorded from the mixing site only.

Typical examples of general representatives of the four classes found in stabilization ponds are as follows:

1. Chlorophyceae: *Ankistrodesmus*, *Chlorella,* and *Scenedesmus.*
2. Cyanophyceae: *Spirulina, Oscillatoria,* and *Rivularia.*
3. Bacillariophyceae: *Fragilaria, Eunotia,* and *Navicula.*
4. Euglena: *Phacus* and *Lepocinclis.*

Some algae unusual for oxidation ponds have also been recorded. They are as follows: *Pediastrum, Microspora, Oedogonium, Spirogyra, Gonium, Lyngbya, Fragilaria.*

A drastic reduction in the number and kinds of algae at the mixing site is due to the exposure of algae to sewage pollution stress. Since the pond in such a situation is exposed to altered environmental conditions, the sensitive species disappeared leaving behind the tolerant forms which may be termed as indicators of sewage pollution.

An understanding of the algal fluctuation for an oxidation pond is essential for its proper management. The phytoplankton population of the oxidation pond showed a bimodal pattern of fluctuation with maximum in monsoon and summer. An irregular pattern of fluctuation was followed by green algae, suggesting their wide range of nutrient status showed direct bearing on the productivity of the aquatic body.

The oxidation pond contains a mixed flora and does not exhibit the predominance of any particular species. This suggests that the pond could handle a heavier load. A majority of the common oxidation pond algae were planktonic and did not collect on the surface as mat or bloom.

Primary productivity of a particular water body gives quantitative information about the amount of energy available to support the bioactivity of the system. Gross primary production ranged between 1.13–4.01 gc/m^2/d in the oxidation pond showing peak values from June to September. A clear-cut bimodal pattern of variation was recorded. The results of net production varied from 0.35 to 2.65 gc/m^2/d in the present study. Positive relationship was observed between phytoplankton standing stalk and productivity values.

Proper balance between hydrological factors stimulates better phytoplankton growth. Even a slight change in one of these led to major changes in living biota. Most of the factors in present findings influenced plankton abundance considerably within certain limits.

From the above account, it can be concluded that the oxidation pond under study can be managed scientifically, and that the application of biological principles should make it possible to improve the efficiency of waste water treatment oxidation stabilization pond (Figs. 1 and 2).

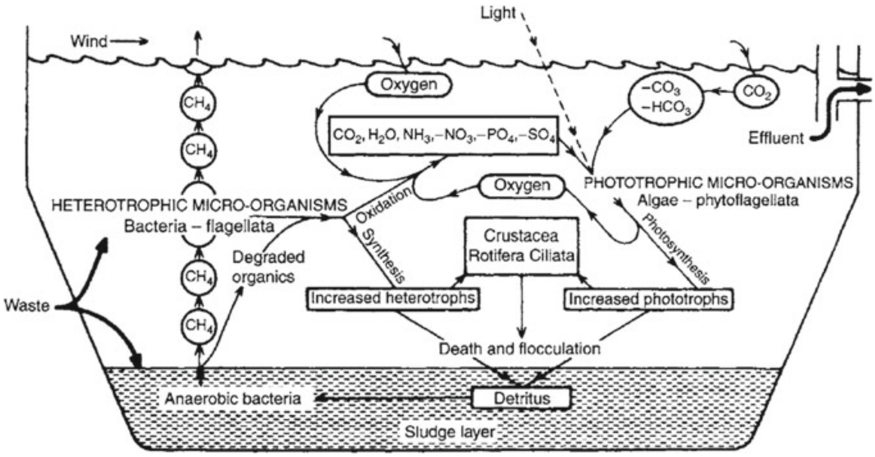


Fig. 1 Example figure for wastewater treatment through stabilization pond (*Source* https://sswm.info/)

Fig. 2 Example figure for stabilization pond

Acknowledgements The author is thankful to Chas Municipality. Chas and for lab facilities provided by Bokaro Steel Plant, Bokaro Steel City.

Bibliography

1. APHA (1985) Standard methods for the examination of water and wastewater
2. BSI (1991) Water—specification IS 10 500 International Standards for drinking water (1984) Issued by World Health Organisation Geneva

3. Manual of Standards of Quality of Drinking Water Supplies (1971) Issued by Indian Council of Medical Research (ICMR). New Delhi
4. Welch PS (1952) Limnology. McGraw Hill, New York
5. https://www.scimagojr.com. Journal IETE Journal of Research—SC Imago
6. www.iete.org—Publications—IETE
7. https://www.tandfonline.com—current
8. IETE Journal of Research, vol 66, No 3. Taylor and Francis Online
9. https://www.springer.com—series
10. IETE Springer Series
11. https://www.springer.com—book
12. Computational urban planning and management for smart cities Stan Geertman Springer
13. https://www.springer.com
14. Sustainable cities. Springer
15. https://link.springer.com—chapter
16. Urban Planning and Smart Cities: Interrelations and Reciprocities

Implementation of Experiential and Project-Based Learning in Mechatronics Course

Ch. Rajendra Prasad, Polaiah Bojja, Sreedhar Kollem, and P. Ramchandar Rao

Abstract With the rapid growth in the technology era, the conventional teaching approach does not suffice for millennial learners. Keeping this in view, making them efficient and globally competitive graduates, the learning environment has to be changed by adopting new teaching pedagogies. To address this, we introduced a multidisciplinary course, i.e., mechatronics in the engineering curriculum. This course is divided into two parts as experiential learning and project-based learning. In the first part, the students are exposed to experiential learning rather than conventional teaching. Then, in the second part, we introduced project-based learning, which helps the students to solve engineering problems. Finally, this course will improve the skills of our next-generation engineers in the fields of innovation, creativity with an entrepreneurial mindset.

Keywords Mechatronics · Experiential learning · Project-based learning (PBL) · Co-design · Multidisciplinary approach

---

Ch. Rajendra Prasad (✉)
Center for Experiential Learning and Department of ECE, S R University, Warangal, Telangana, India
e-mail: chrprasad20@gmail.com

P. Bojja
Department of ECE, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Andhra Pradesh, India
e-mail: polaiah@kluniversity.in

S. Kollem · P. Ramchandar Rao
Department of ECE, S R University, Warangal, Telangana, India
e-mail: ksreedhar829@gmail.com

P. Ramchandar Rao
e-mail: chandupamulaparthi@gmail.com

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_50

1 Introduction

In 1969, a Japanese company senior Engineer (Mr. Tetsuro Mori) first used the word mechatronics [1]. Mechatronics refers to the integration of mechanical and electronic systems. But, mechatronics means it is the combination of multidisciplinary engineering to develop simple to complex systems with ease, economical, and reliable. Figure 1 illustrates the basic streams involved in mechatronics and its broader application areas.

A mechatronic system typically has a mechanical system; its behavior is controlled in a predefined manner using a processor/controller. An array of sensors is connected to the processor/controller using suitable signal conditioning and analog-to-digital converters to monitor the status of the mechanical system.

From this status information, a processor/controller takes the decisions. Decisions are made through code running on the microcontroller. These decisions are given to actuators of mechanical systems through the ADC interface of the microcontroller to make actions like rotate, move, and lift.

The experienced learning in engineering education is worth having, especially within a virtual education upward movement, and that honesty has an impact on creativity and thinking abilities [2, 3].

The learning efficiency and effectiveness are improved year by year over our six years offering of the presented mechatronics pedagogy into mechanical engineering [4]. From the practical point of view, the reform of teaching methodology and content is of great significance to cultivate innovative talents and to improve teaching quality which promotes effective teaching [5, 6]. To realize the 'Make in India' vision, to

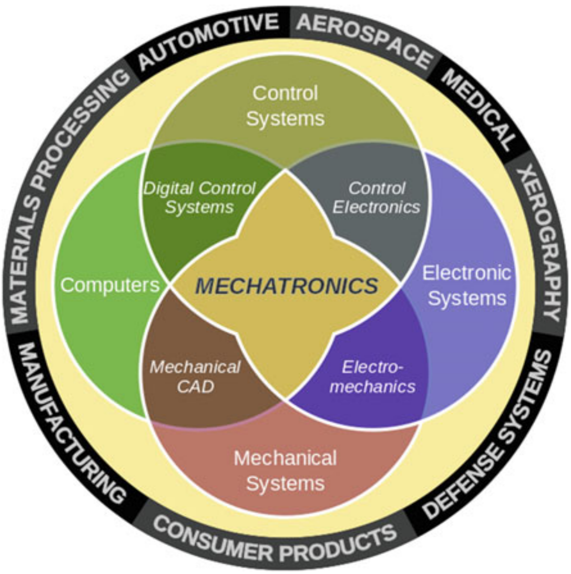


Fig. 1 Basic streams and application domains of mechatronics

classy services and products that are viable in the international market, India requires a large number of skilled and highly qualified graduates [7, 8]. The conventional teaching methods do not meet the needs of a professional activity that currently includes the ability to analyze problems and exploration for solutions by collaborative working, communication skills with professionals from different specialties, and motivation for life-long learning [9, 10].

From the above studies, we identified that conventional teaching is not sufficient for millennial learners, from the following aspects

- Less attention span
- Huge information sources
- Lack of motivation
- Social distance.

The remainder of the paper is as follows. Sect. 2 deals overview of the mechatronics course; Sect. 3 describes the methodology and implementation of the mechatronics course. Section 4 presents the analysis of the proposed methodology and Sect. 5 conclusion of the proposed methodology and overall experience in learning.

2 Course Overview

The course mechatronics includes topics like microcontroller fundamentals, instruments for signal processing, sensors and actuators, and control circuitry. In this course, the theory is combined with immediate hands-on sessions to get effective experiential learning. The subject is based on project-based learning by conducting a few laboratory experiments that introduce the students to the use of electromechanical devices. Finally, the students in the group are required to develop a mechatronic device for automated applications, and by this, the students can understand aspects of multidisciplinary applications. Hence, quality of education can be maintained, and students will have a platform to think upon and showcase their capabilities.

(A) *Basic Information*

This course is introduced for undergraduate students as a multidisciplinary course in II B.Tech. II semester for ECE, EEE students, and III B.Tech. I semester for CSE and ME students. 4 h per week are allocated for this course excluding laboratory hours. Also, two laboratory sessions per week are conducted. The pre-requisite for this course are engineering physics and basic C programming. The course objectives and outcomes are listed below:

(B) *Course Outcomes*

- Describe the key elements of the mechatronic system
- Explain the principle and operation of sensors
- Apply the mechanism of actuators for an application
- Analyze speed and power control of motors

- Examine signal conditioning circuits
- Design a mechatronic system for a specific application.

3 Methodology and Implementation

To overcome the deficiencies of conventional teaching, we modified the teaching–learning process by incorporating experiential learning and PBL in the mechatronics course. This course is divided into two parts as experiential learning and project-based learning. In the first part, the students are exposed to experiential learning rather than conventional teaching. Then, in the second part, we introduced project-based learning, which helps the students to solve engineering problems.

(A) *Experiential Learning*

This is the first part of the proposed methodology, in which we divided the total course syllabus into eight modules. The first four modules are the demonstration, and the remaining four modules are implemented with four active learning pedagogies. These pedagogies include think aloud pair problem solving (TAPPS), student-team-achievement-divisions (STAD), flipped class, and Jigsaw. The implementation of these pedagogies are as follows:

TAPPS [11]: The active learning method is shown in Table 1. The topic selected for this activity is the sensors. Reflections of the TAPPS activity—the students are involved more in this activity with enthusiasm because their involvement and pair selection are of their friends. This activity is only 50% successful. Because maintenance of large class with this activity is difficult and first-time implementation.

STAD [12]: By considering the reflections of the above activity, collaborative learning is introduced to my students on the topic of temperature sensors. There are several activities are presented for collaborative learning. My course is having well-defined objectives; hence, STAD is selected, and it is more appropriate to introduce collaboration to my students. It is best suitable for the mechatronics course. Students learning teams are formed based on the following factors.

Table 1 TAPPS activity

Activity	Topic: sensors		
	Specific question	Details of implementation	Time allotted
TAPPS	What is a sensor? How are sensors classified What is the basic principle of displacement sensors?	The topic will be given. Recorders will be chosen Students pair up and try to explain the topic. Summarize the operation of the sensor In these activities, basic concepts of displacement sensors can be expected	min. list min. share min. explain

- Mixed in performance level (based on performance)
- Sex (2 boys + 2 girls)
- Ethnicity.

Reflections of the STAD activity—the students are involved more in this activity with enthusiasm because their involvement and pair selection are of their friends. This activity only 78.3% success. Hence, maintenance of large classes with this activity is somewhat better than the first activity.

Flipped Class [13]: The topic selected for this activity is electrical actuators.

Step-by-step procedure to implement this activity.

- First, we provided suitable topics for a flipped class in the lecture plan
- Prepare material and recourses required for students
- Announce dates for these activates in my lesson plan
- Before this activity, they can contact during free hours.

Then, students have to explain the topic. To assess the student's understandings, test is conducted on 10 marks. The performance of students on the test is shown in Fig. 2.

Reflections of the Flipped Class activity—most of the students are not prepared for the activity. For the conduction of this activity, student co-operation is very important. Only, fast learners can easily participate, but slow is facing difficulty.

Jigsaw Activity [14]: The topic for the activity is mechanical actuators. The following are the steps and time for the activity (Course project team) (5 min).

- Read allotted topic for you (before the class)
- Members belong to different groups have looked into the identical theme, as well as discuss each other (15 min)
- Back to your group as well as discuss topic in front of group members (45 min)
- Write a test on the topic (20 min).

The student's involvement in the activity is shown in Fig. 3. The test is prepared with Bloom's level 3 questions. The average scores of students are improved with this activity comparing with other activities. The individual student scores after the

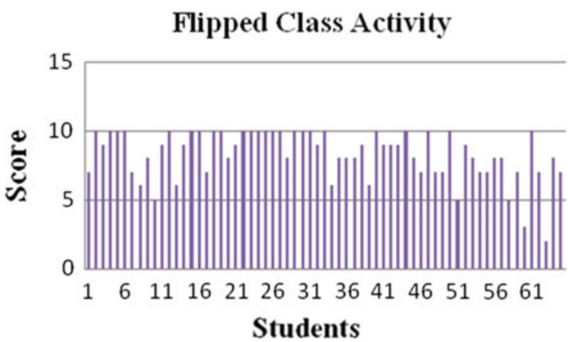


Fig. 2 Students scores of flipped class activity

Fig. 3 Students involvement in Jigsaw activity

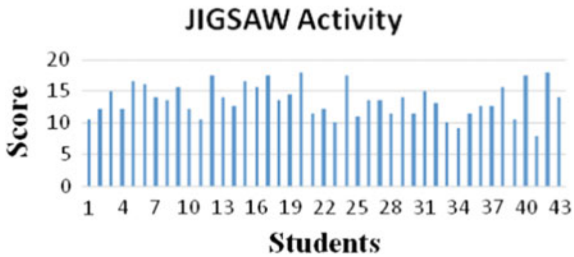


Fig. 4 Individual student scores of the Jigsaw activity

test are shown in Fig. 4. The overall average score of the class is approximately 14 out of 20 marks.

Reflection on the above activity: Classes including more students have been engaged with the help of this activity. The student abilities, namely collaborative and self-learning, have been improved by using this activity. The conclusion of this activity has success of 80% because of topic understanding, student gaining score, and feedback.

(B) *Project-Based Learning*

The second part of this course is provided for PBL which is a learner-centered instruction that involves a dynamic classroom method in which learners attain in-depth knowledge through active exploration of real-world and problems and challenges [15]. Students learn about a subject by working for an extended period to investigate and respond to a complex question, challenge, or problem [16].

After eight weeks of experiential learning, students are grouped into batches. Each batch is allocated with a specific problem statement. Then, each batch has to submit the product or prototype based on their problem statement. For this, we provide the required components for every batch. The remaining eight weeks are provided for the development of their project. By doing this project, students were exposed to project-based and student-centered learning. Finally, they have to exhibit a project

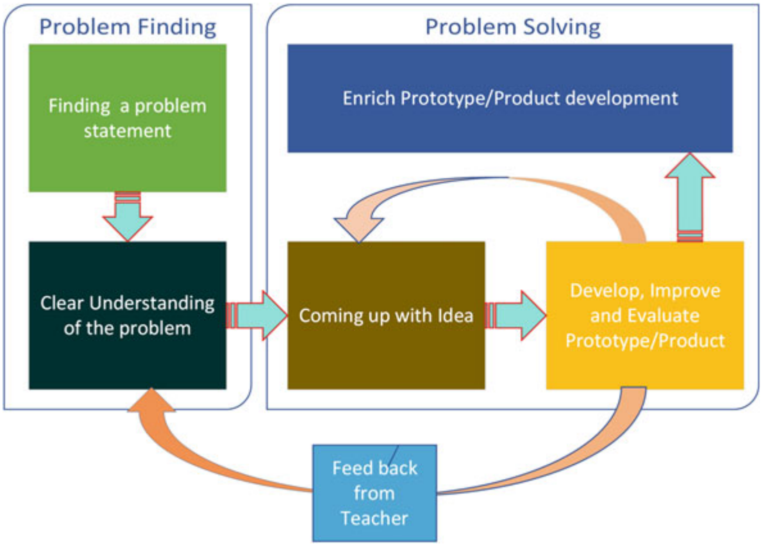


Fig. 5 Implementation procedure of PBL

for evaluation. The complete implementation of PBL is shown in Fig. 5, and the stepwise implementation procedure is as follows:

Step 1: Finding the problem statement by observing societal problems.

Step 2: Clear Understanding of problems by surveys and literature reviews.

Step 3: Coming up with an idea, by agreement with all team members and submit it to the mentor/guide.

Step 4: Development of the product/prototype. This stage is further divided into four stages.

1. Hardware
2. Software
3. Development
4. Evaluate.

In this stage, prototype/product is tested for required specifications, if it is satisfied to continue with step 4, otherwise repeat from step 2 or 3 based on the mentor/guide feedback.

Step 5: Enrich prototype/product: During this stage, students prepare casing for their product/prototype.

The following are a sample of student course projects during the implementation of PBL.

Fig. 6 Remote-controlled robot

Remote-Controlled Robot

This project describes controlling the movements of the robot wirelessly using IR data communication. When we press a button on the remote, the IR transmitter (Television remote) sends IR signal to the TSOP1738 (IR receiver). At the receiver end, photodiode detects light pulses and decodes the information. The decoded information (RC5 format) is sent to the processing device. Based on the received data, the Arduino mega 2560 board performs the necessary actions by sending commands to the robot through a motor driver (L293D), like controlling robot FRONT or BACK or LEFT or RIGHT, or STOP. The designed remote-controlled robots are well suited to collect data from harsh environments, industrial applications, etc. The working model of this project is shown in Fig. 6.

Identify the Headings Gesture-Controlled Robot

The present structure of the project involves the method of controlling the robot using hand gestures as commands using an accelerometer. The control of the locomotion of the robot is presently done by a microcontroller (Arduino mega 2560 board). An accelerometer measures gravitational force. This ADXL 335 gives analog voltage which is read as by the microcontroller. Based on this voltage obtained from the ADXL 335 for the various orientations of the hand, the microcontroller will decide the necessary actions by sending the commands to the robot, like moving the robot LEFT, RIGHT, FRONT, BACK, or STOP. The working model of this project is shown in Fig. 7.

4 Results and Discussion

A. *Analysis of Experiential Learning*

Fig. 7 Gesture-controlled robot

The analysis and success of activities are measured by three parameters such as student performance, observation, and student feedback. A comparison of the above four activities is shown in Table 2, and the mapping of activities with Cos and POs is shown in Table 3.

This comparison is purely based on my class students' performance, feedback given by them, and my observation. Based on my observation, while conducting

Table 2 Comparison of the activities

Activity	Mechatronics conducted activities			
	Performance of students in the activity	Feedback of students on the activity	Observation	Overall success of activity (%)
TAPPS	60	50	40	50
STAD	75	80	80	78.3
Flipped class	45	60	40	48.3
Jigsaw	70	100	90	85

Table 3 Mapping of COs and POs with activities

S. No.	Name of the activity	Topic	CO Mapping	PO Mapping
1	TAPPS	Key Elements of Mechatronics	CO1	PO1
2	STAD	Sensor characteristics and classifications, Selection of sensor	CO2	PO1
3	Flipped Class	Electrical Actuators	CO3	PO2
4	Jigsaw	Mechanical Actuators	CO3	PO3

these activities, the flipped class activity is a failure because they are not prepared for the given topics. But, Jigsaw and the STAD are succeeded in my class.

Continuous Internal Evaluation Assessment: The mechatronics course is evaluated for Continuous Internal Evalution (CIE) of 30 Marks. For CIE, we followed the following rubric for every week.

- Presentation of Idea—5 Marks
- Performance—10 Marks
- Teamwork—5 Marks
- Implementation—10 Marks

This can be averaged based on the first eight weeks. While comparing with previous semester results, the current semester results are improved by 22.7%.

B. Analysis of PBL

As a part of this course, the students learn the basics of automated systems. The students have developed projects based on their knowledge gained during the class as well as in the laboratory. During this semester, 12 projects are prepared by students. The evaluation rubric for the course projects is given in Table 4. The mechatronics course is evaluated for Semester End Examination (SEE) of 70 Marks and the SEE distributed as follows:

- Objectives—15 Marks
- Product/prototype—25 Marks
- Documentation—15 Marks
- Viva-voce—15 Marks

The evaluation of the semester end examination is a project-based competition. To increase the student interest in doing project-based learning, we organized a project demo-day to showcase their projects and share their knowledge. Some of the projects displayed along with posters on demo-day are shown in Fig. 8. By implementing PBL in the mechatronic course, the students have been achieved following aspects such as:

- Work with teams
- Self-learning ability
- Utilization of resources for a given application
- Self-motivation.

5 Conclusion

We have followed two approaches during the mechatronics course, experiential learning and project-based learning. The learning by doing is implemented by experiencing students more on hands-on practice rather than conventional chalk and talk and conducting activities on difficult topics. Project-based learning is implemented

Table 4 The evaluation rubric for the mechatronics course projects

Criteria	Beginning	Developing	Accomplished	Exemplary
Creativity and Innovation [17]	No inquiry and relevant data on the progress of the idea. Followed none of the instructions	The level of inquiry is low and less determination on developing concepts	The basic level of inquiry and collected data is relevant to the design	The proper level of inquiry on a required data, which is sufficient for the design
Cooperation and Knowledge	The teamwork and basic concepts are not up to the mark	The teamwork and basic concepts are satisfactory	The teamwork and basic concepts are fair enough	The teamwork and basic concepts are completely meet the requirements
Design Principles	Student implementation is not up to the mark of the design principles	Student implementation is up to the mark. To meet the design principles, further, improvement is needed	Student implementation fairly meets the design principles	Student implementation completely meets the design principles
Module Design	The design modules are not clear and are difficult to realize	The design modules in modules are clear; interfaces between modules are not clear	The design modules and interfaces between modules are clear. The design is efficient and easy to realize	The design modules and interfaces between modules are clear. The design is efficient and well suited for the design
Testing and Prototyping and [18]	The design does not meet the requirements	The design almost meets the requirements, but still has minor logical errors	The design is meeting the requirements without any logical errors	The design is well suited and producing extraordinary outputs

through a course project. Total 12-course projects are developed with different problem statements. This course helps the students to solve engineering problems and also enlightens the fundamental building blocks of our next-generation engineers like innovation, creativity, and entrepreneurial mindset. Finally, the PBL introduced in the mechatronics course improved student problem-solving skills, team-based learning, and self-learning abilities which develops confidence in my students. But still, we need some extra effort in planning and conducting the activities.

Fig. 8 Projects displayed along with posters on demo-day

Acknowledgements This work is supported by the department of ECE and Center for Experiential Learning, S R University, Warangal, Telangana, India.

References

1. Gesualdo E, Maximiliam L, Oliveira VA (2017) Designing an engineering laboratory course for problem based learning. J Eng Educ Transform
2. Velasquez-Lopez A, Rodriguez-Garcia A (2017) A new mechatronics laboratory for technology integration. Int J Interact Design Manuf (IJIDeM) 11(1):115–124
3. Rai S et al (2017) Mechatronics pedagogy in mechanical engineering capstone design. In: 2017 American Control Conference (ACC), IEEE
4. Li Z et al (2017) Research on the case teaching of graduate mechatronics course in mechanical engineering professional degree. In: DEStech transactions on social science, education and human science EMSE
5. Lord SM, Camacho MM (2007) Effective teaching practices: preliminary analysis of engineering educators. In: 2007 37th annual frontiers in education conference-global engineering: knowledge without borders, opportunities without passports. IEEE
6. Aluvala S, Pothupogu S (2015) A traditional novel approach for skill enhancement of teaching-learning process in engineering education. J Eng Educ Transform 28(4):92–95
7. Pothupogu AG, Reddy S, Sai Priya P (2018) Trenchant pathway to bring innovation through foundations to product design in engineering education. In: IEEE 6th international conference on MOOCs, innovation, and technology in education (MITE), pp 43–47. IEEE
8. Alok G, Anushalini T, Condoor S (2018) Effective approach towards development of idea through foundations to product design. J Eng Educ Transform 31(3):47–52

9. Kumar KSA, Reddy G, Reddy NBS (2018) An integrated Multidisciplinary skill development strategy for effective execution from virtuality to reality in engineering education. In: IEEE 6th international conference on MOOCs, innovation and technology in education (MITE), pp 79–83. IEEE
10. Kolberg E, Reich Y, Levin ILYA (2003) Project-based high school mechatronics course. Int J Eng Educ 19(4):557–562
11. Aydın İ et al (2017) Investigation of preferences from technology faculty mechatronics engineering students. Jr. Educ Instruct Stud World 7(2):22–26
12. Shahri N, Abdul Rahman R, Hussain NH (2017) Enhancing Students' deep approaches to learning among industrial mechatronics engineering technology students. Sains Humanika 9(1):2
13. Anuradha P (2019) The teaching learning process. Int J Adv Sci Technol 28:709–714
14. Dorr M (2017) The effectiveness of project-based learning using digital storytelling technology on improving second-grade students' performance of science standards
15. Pandit CM, Panpat BR (2019) Development of ICT based PBL platform. J Eng Educ Transform 33(1):131–136
16. Hixson NK, Ravitz J, Whisman A (2012) Extended professional development in project-based learning: impacts on 21st century skills teaching and student achievement. West Virginia Department of Education
17. Grover R et al (2014) A competition-based approach for undergraduate mechatronics education using the arduino platform. In: Fourth interdisciplinary engineering design education conference. IEEE
18. Hargrove JB (2002) Curriculum, equipment and student project outcomes for mechatronics education in the core mechanical engineering program at Kettering University. Mechatronics 12(2):343–356

Analysis of Metamaterial-Based Bandstop Filters: A Brief Review

Priyanka Garg and Priyanka Jain

Abstract A comparative analysis of bandstop filters (BSF) with transmission line (TL) employing different metamaterial unit cells is presented in this paper. There are a number of filters proposed in literature that have metamaterial etched on the ground plane; therefore, a comparison-based study is required to be performed to ease the user for selecting the best suitable configuration according to the desired specification. In this paper, six different metamaterial unit cells are taken from the literature and then optimized to perform bandstop operation at 3 GHz. Each structure is simulated and analyzed. The equivalent circuit extraction of the designs is also performed. In the end, a comparison table is presented for all the six filters on the basis of different characteristic parameters of filters such as 3 dB bandwidth, insertion loss, shape factor, quality factor (Q), unit cell size, overall size and group delay.

Keywords Metamaterial · Resonator · Transmission line · Filter · Comparison

1 Introduction

Filters are the essential components required in modern wireless communication systems. They either pass or reject a desired band of frequencies in various RF and microwave applications. With advancement in wireless communication technology, demand of compact sized, low-cost and high-performance filters in increasing. Planar technology was introduced for filter designing in order to fulfill that demand. Also, various tool and techniques have been investigated to overcome the drawbacks of planar technology and design filters with improved performance and minimum complexity.

P. Garg (✉) · P. Jain
Department of Electronics and Communication, Delhi Technological University, New Delhi, India
e-mail: priyankagarg_phd2k16@dtu.ac.in

P. Jain
e-mail: priyankajain@dtu.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_51

In last few decades, metamaterial resonators, especially split-ring resonators (SRRs) [1] and complementary split-ring resonators (CSRRs) [2] have gained lots of interest in the area of filter designing because of their extraordinary electromagnetic properties. SRR leads to negative effective permeability (μ) around the resonance frequency, whereas CSRR being the counterpart of SRR, obtained by etching the negative image of SRR, leads to negative effective permittivity (ε). Various configurations of metamaterial resonator-based filters have been reported in state-of-the-art such as coupled line filters [3, 4] that incorporate microstrip edge-coupled split-ring resonators along the transmission line, defective microstrip line (DML) [5, 6] that have resonators etched out of the microstrip transmission line, defective ground plane (DGS) [7–10] in which microstrip transmission line is on the top of the substrate, and resonator is etched out of the ground plane and many more. DGS-based filters are one of the most simple, efficient and easy to use configurations [11]. Demonstrated that CSRR-based filters attract more attention due to deeper rejection levels and much wider rejection bands as compared to the ones having SRR coupled to the sides of transmission line.

In this paper, a comparison-based analysis of different complementary resonators in designing microstrip transmission line-based bandstop filters is presented. Six metamaterial resonators are taken from the literature; first of all, their dimensions are optimized to provide bandstop operation at 3 GHz, and later, a comparative study is presented on the basis of different filter characteristics like unit cell size, overall size, 3-dB bandwidth, insertion loss, quality factor (Q), group delay and shape factor. A number of metamaterial resonators bandstop filters are proposed in literature; therefore, this paper aims at providing a comparative study with an objective to aid the user point out the best configuration required for a particular specification of bandstop filter as well as evolving the ideas in future by considering the advantages and drawbacks of the present structures. Computer Simulation Technology-Microwave Studio (CST-MW) [12] is used for structure designing, and CST Design Studio [13] is used for circuit design and simulation.

2 Design and Simulation Approach

To design the filter, an FR-4 substrate having dielectric constant (ε_r) = 4.4 and loss tangent ($\tan \delta$) = 0.025 with thickness (h) 1.6 mm is used. On the top of the substrate, a 50-Ω microstrip transmission line is etched and the metamaterial unit cell is etched out of bottom ground plane.

2.1 Metamaterial Unit Cells

All the six configurations of metamaterial unit cells are shown in Fig. 1 that are chosen to investigate the filtering characteristics. Their dimensions are optimized to

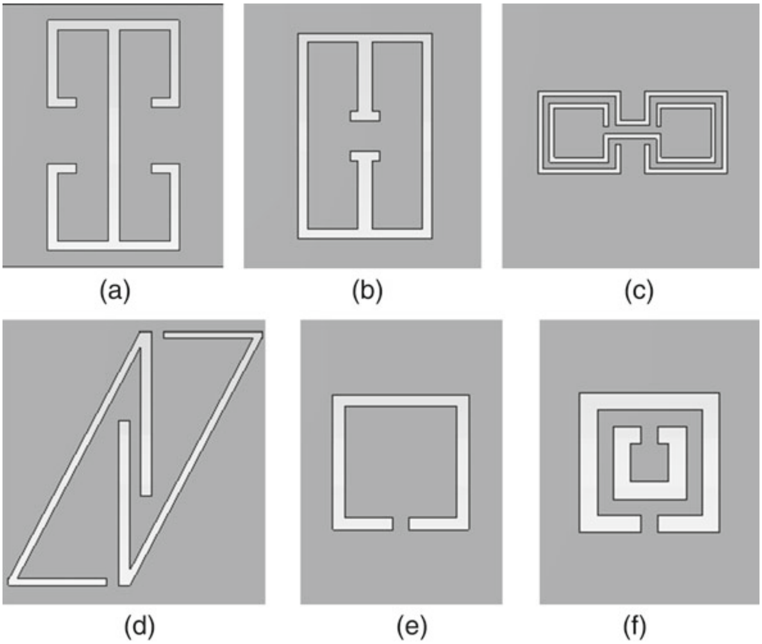


Fig. 1 **a** CELC-I [7], **b** CELC-II [14], **c** OSSRR [8], **d** CSTR [15], **e** SSRR [16], **f** DSRR [11]

operate at 3 GHz and exhibit negative value of effective permeability or permittivity or both near operating frequency.

Two types of complementary electric inductive-capacitive (CELC) resonators are shown in Fig. 1a, b which are defined as mu-negative material. The CELC in Fig. 1a is named as CELC-I and the one in Fig. 1b as CELC-II. An open-slot split-ring resonator (OSSRR), tends to exhibit electric resonance, is shown in Fig. 1c. It is formed by two concentric square slots which are connected to each other and also have openings at the same side. Further, a complementary split triangle resonator (CSTR) is depicted in Fig. 1d which provide double-negative behavior, i.e., negative values of permittivity and permeability simultaneously, therefore providing both electric and magnetic resonances. Single split-ring resonator (SSRR) and double split-ring resonator (DSRR) are the most popular and most explored unit cells which are shown in Fig. 1e, f, respectively. Both of them exhibit negative permittivity near resonance.

2.2 Design and Simulation of Metamaterial Based Filters

For exhibiting filtering characteristics, a metamaterial unit cell is etched out of the ground plane underneath a 50-Ω transmission line. To provide proper impedance matching, the length of transmission line is optimized, and accordingly, the overall structure dimensions are determined.

(1) *CELC-I-based bandstop filter*

The schematic of the bandstop filter utilizing CELC-I unit cell is shown in Fig. 2a, and the results, i.e., impedance curve, reflection coefficient (S_{11}), transmission coefficient (S_{21}), phase in degrees and group delay are indicated in Fig. 2b–e. The S_{11} shows a resonance frequency at 3 GHz along with a weak transmission zero at 5.67 GHz. It can further be optimized on the basis of required specifications. It can be observed from the impedance curve of the structure, shown in Fig. 2b, that it is inductive above the operating frequency and capacitive below it so one can conclude that the circuit equivalent for the filter is a shunt-connected series resonator (Fig. 3). Also, the phase response in Fig. 2d shows a phase changes from positive to negative value at 3 GHz indicating phase lead, which is the characteristics of metamaterial. The filter also provides a group time delay of only −0.548 ns (Fig. 2d).

Furthermore, the values L_1 and C_1 can be calculated by using the conditions given below:

1. Consider f_1 is the resonant frequency, so at f_1, $|S_{11}| = 0$.
2. f_2 is the frequency at the intersection of $|S_{11}|$ and $|S_{21}|$, then $|S_{21}| = 1/\sqrt{2}$.

Now, for a shunt-connected series resonator, the S-parameter matrix of the filter is given as,

$$[S] = \begin{bmatrix} \frac{-YZ_0}{2+YZ_0} & \frac{2}{2+YZ_0} \\ \frac{2}{2+YZ_0} & \frac{-YZ_0}{2+YZ_0} \end{bmatrix}$$

Utilizing the condition 1, we obtain,

$$L_1 = \frac{1}{4\pi^2 f_1^2 C_1} \tag{1}$$

Further, utilizing condition 2, the equation for C_1 obtained in terms of f_1 and f_2 is given below:

$$C_1 = \frac{1}{\pi f_2 Z_0}\left[1 - \left(\frac{f_2}{f_1}\right)^2\right] \tag{2}$$

Figure 2c shows that the value of f_1 and f_2 are 3 GHz and 2.37 GHz, respectively. On using Eqs. (1) and (2), the values of lumped parameters are obtained as $C_1 = 1.01$ pF and $L_1 = 2.79$ nH. CST Design studio is used to design and simulate the obtained equivalent circuit. The simulated S-parameters of the equivalent circuit model extracted using the above procedure is indicated in Fig. 4 which is a close approximation to the S-parameter obtained after design simulation.

The power flow in the BSF is demonstrates in Fig. 5 at 3 and 4 GHz. It can be

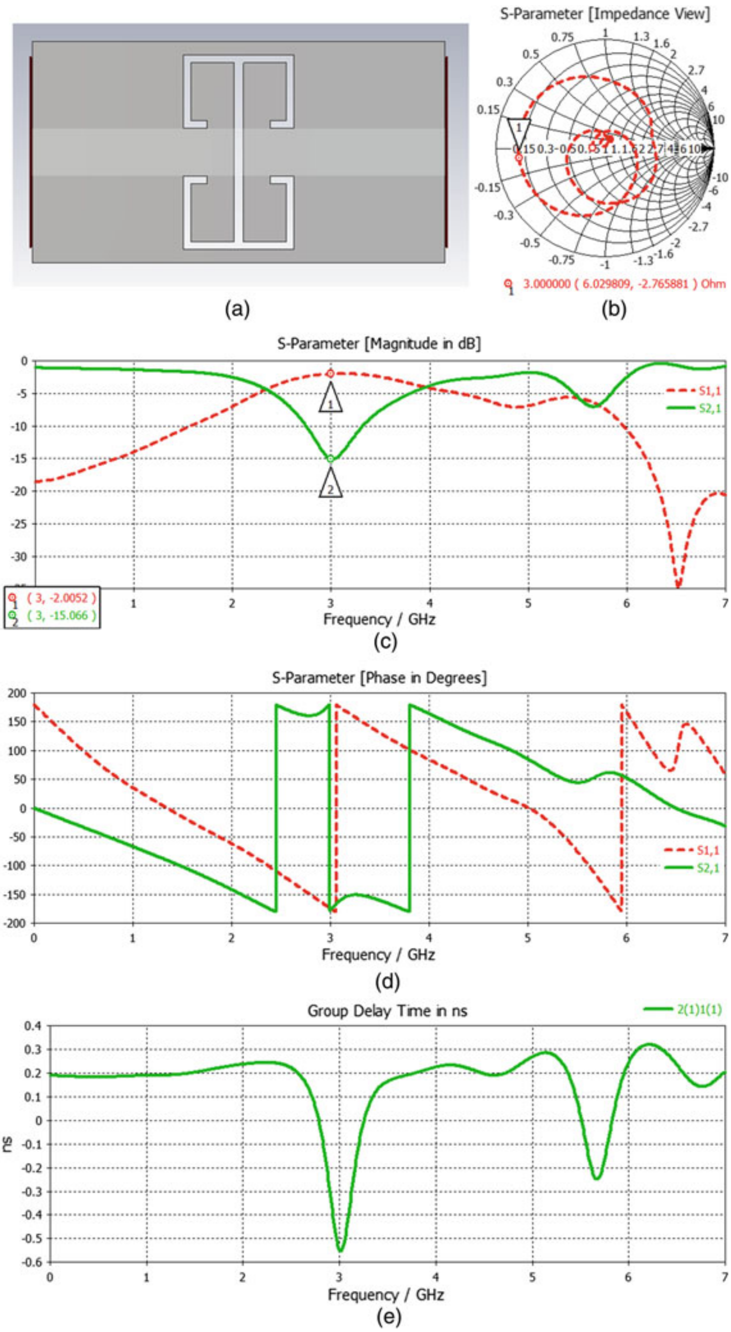


Fig. 2 CELC-I based bandstop filter

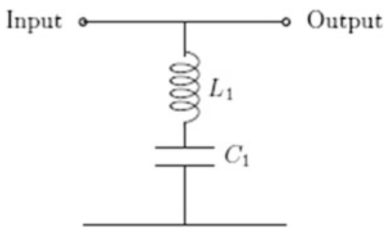


Fig. 3 Equivalent circuit of the CELC-I bandstop filter

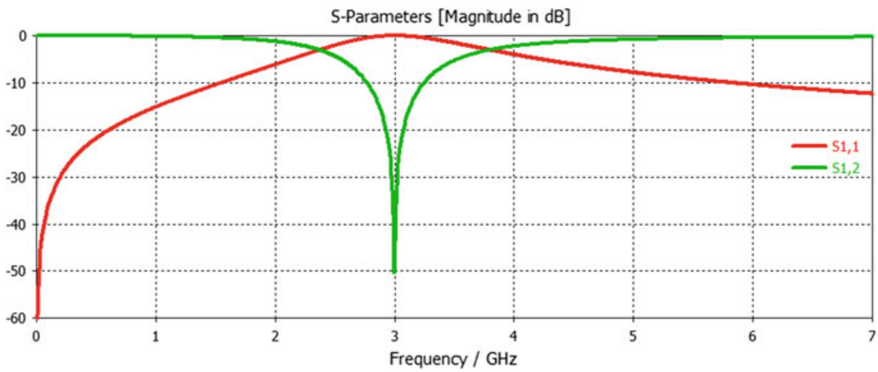


Fig. 4 CELC-I circuit simulation

observed that from port 1 to 2, the power flow is prohibited at 3 GHz, whereas at 4 GHz the power is allowed propagate between the ports. This indicates bandstop characteristics at 3 GHz and passband at 4 GHz.

(2) *CELC-II-based bandstop filter*

The simulated results and schematic of the bandstop filter based on CELC-II resonator are shown in Fig. 6a–e. The filter exhibits a transmission zero at 3 GHz, as indicated by reflection and transmission coefficient curves, shown in Fig. 6c; however, a stop band is also indicated at higher frequencies (greater than 5.5 GHz). Two reflection zeros, one at 3.53 GHz and the other at 4.53 GHz are also provided, showing very good passband characteristics. The equivalent circuit in this case is also shunt-connected series resonator as the impedance curve shows the similar behavior to that of CELC-I. Its parameters can also be extracted by the same procedure already explained earlier. The design is also providing a group delay time of −1.156 ns indicating phase lead.

(3) *OSSRR-based bandstop filter*

The schematic and simulated results of the bandstop filter based on OSSRR are depicted in Fig. 7a–e. It shows that at 3 GHz, there is a transmission zero with a narrow bandwidth of 0.272 GHz. The impedance curve (Fig. 7b) depicts capacitive behavior above and inductive behavior below the operating frequency. Thus, the equivalent

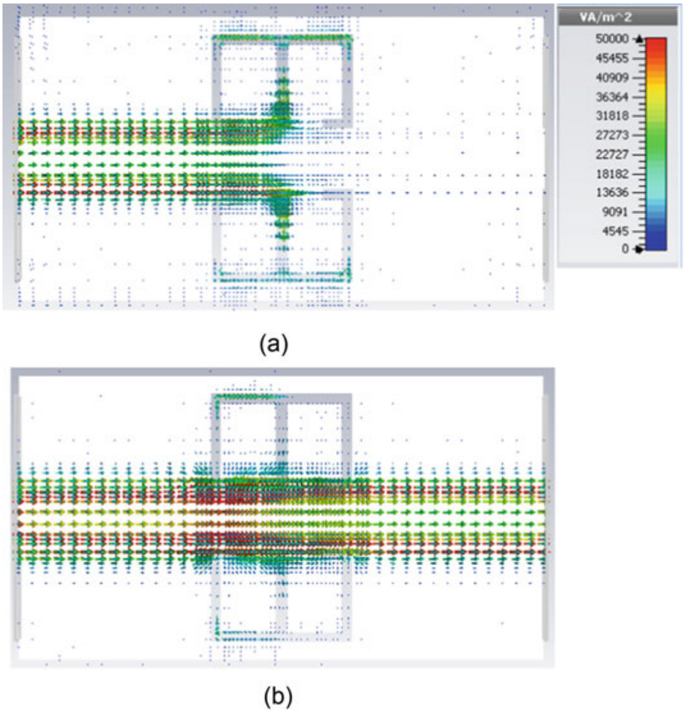


Fig. 5 Flow of power at 3 GHz (**a**) and 4 GHz (**b**)

circuit of OSSRR-based BSF is a series-connected parallel-resonator. However, at the lower frequency 2.12 GHz, there exists a reflection zero, so a capacitor in series with the LC tank circuit is added in order to incorporate it, as depicted in Fig. 8. A phase change from positive to negative at 3 GHz and a group delay time of −9.97 ns are also shown in Fig. 7d, e, respectively.

The same procedure, as discussed earlier, is followed to obtain the values of lumped circuit parameters. The equations for calculating L_1 and C_1 are given as follows:

$$C_1 = \frac{1}{4\pi^2 f_1^2 L_1} \tag{3}$$

$$L_1 = \frac{Z_0}{\pi f_2}\left[1 - \left(\frac{f_2}{f_1}\right)^2\right] \tag{4}$$

The LC tank circuit will behave as an inductor at lower frequencies whose equivalent value is −1.23 nH at 2.12 GHz (lower reflection zero); hence, the capacitor C_2 added in series with −1.23 nH inductor is given by 4.57 pF. The simulated S-parameters of the extracted equivalent circuit model are shown in Fig. 9.

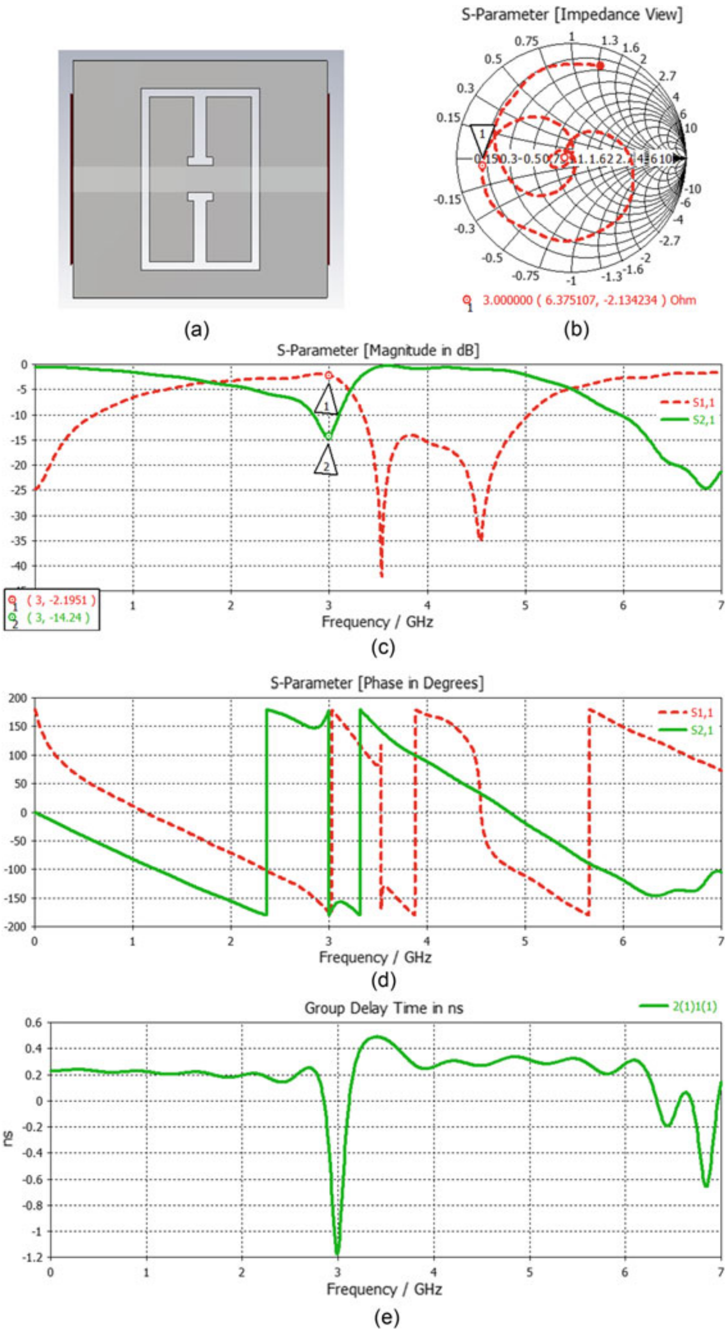


Fig. 6 CELC-II-based BSF

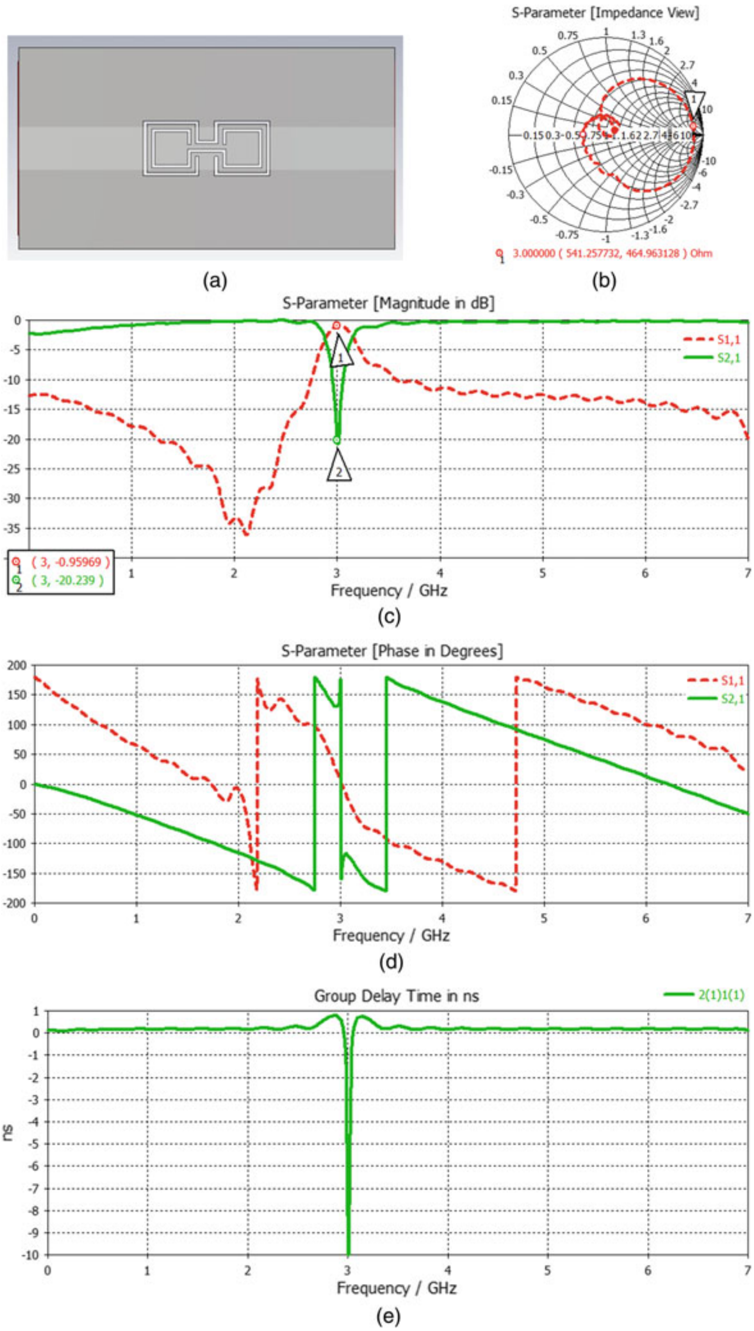


Fig. 7 OSSRR-based BSF

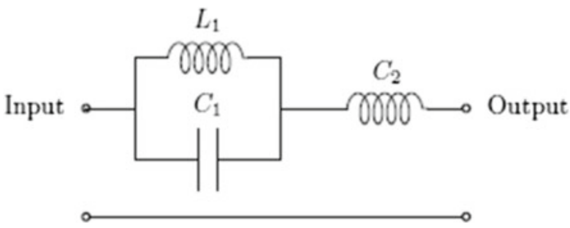


Fig. 8 Equivalent circuit of OSSRR-based BSF

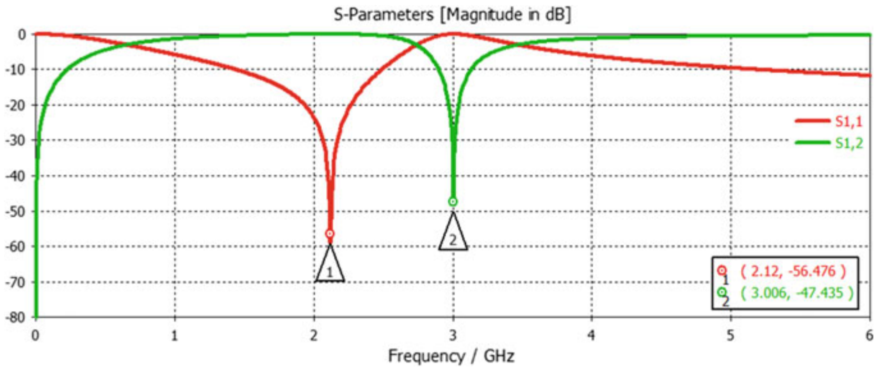


Fig. 9 Circuit simulation of OSSRR BSF

(4) *SSRR-based bandstop filter*

The schematic of the designed BSF based on SSRR is shown in Fig. 10a. The overall dimensions (22 × 20 mm^2) of SSRR-based BSF are compact as compared to aforementioned filters. A series-connected parallel resonator would be the equivalent circuit of this configuration which can be observed from the impedance curve shown in Fig. 10b. Figure 10c shows the *S*-parameters which indicate a transmission zero at 3 GHz and a reflection zero at 2.31 GHz, showing good lower passband response. It provides a wide 3 dB bandwidth. There is a weak series-resonance occurring at 2.31 GHz because of the reflection zero. It can be incorporated in the equivalent circuit if an inductor is added in series with the LC tank circuit, as shown in Fig. 11. The equivalent lumped parameters are calculated using the conditions mentioned above, given as, $C_1 = 1.98$ pF, $L_1 = 1.42$ nH and $L_2 = -3.49$ nH. The simulated *S*-parameters of the extracted equivalent circuit model are shown in Fig. 12 which fairly matches the design simulation results.

(5) *DSRR-based bandstop filter*

The schematic of the designed BSF based on DSRR is shown in Fig. 13a. Figure 13b–e shows that the simulated results are similar to that of a SSRR. The two split rings introduced in DSRR improve the inductive and capacitive coupling in the structure.

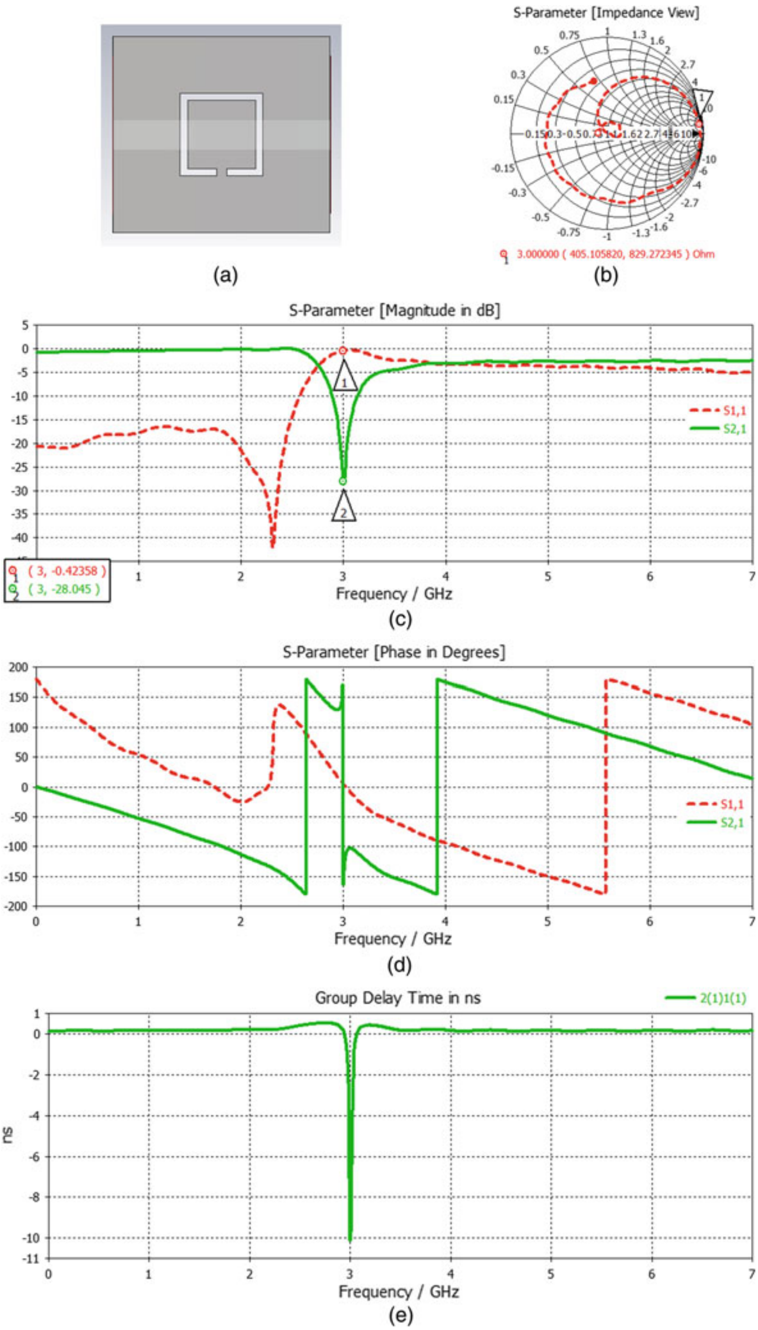


Fig. 10 SSRR-based filter

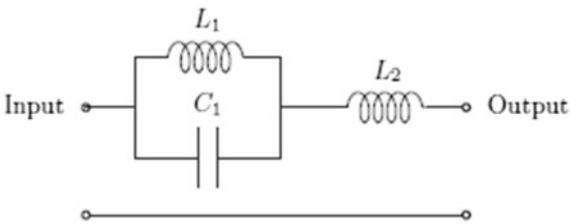


Fig. 11 Circuit equivalent of the SSRR-based BSF

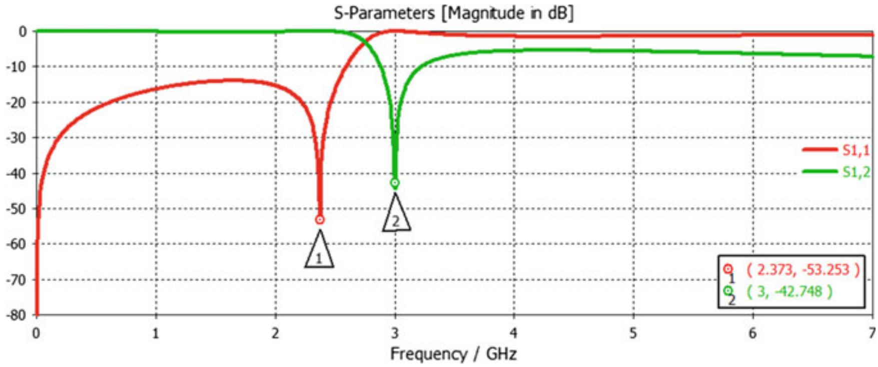


Fig. 12 SSRR BSF circuit simulation

If the mutual coupling is weak, in that case, a DSRR is similar to an SSRR. Also, if the dimensions of two rings are almost same or the gap between two rings is small, the resonant frequency of DSRR is close to single SRR along with larger current density leading to higher magnetic moment [17].

DSRR-based bandstop filter and SSRR-based bandstop filter have similar equivalent circuit, and the lumped parameters extracted using the equations discussed in previous section are given as, $C_1 = 1.75$ pF, $L_1 = 1.608$ nH and $L_2 = -4.28$ nH. The simulated S-parameters of the extracted equivalent circuit model are indicated in Fig. 14 showing close approximation to design simulation results.

(6) *CSTR-based bandstop filter*

Figure 15 illustrates the geometrical configuration of CSTR-based bandstop filter along with its simulated results. Figure 15c shows the S-parameter response of the filter. It shows a stopband resonance at 3 GHz, and an excellent out of band performance can be seen due to the presence of two reflection zeros, one below the stop band at 2.49 GHz and one above it at 3.43 GHz. However, the filter response is disturbed at high frequencies, and the overall dimensions of the filter is also large, i.e., 38×20 mm^2. Figure 15b shows the impedance curve of the BSF which illustrates that the equivalent circuit is a shunt-connected series-resonator. Its parameters

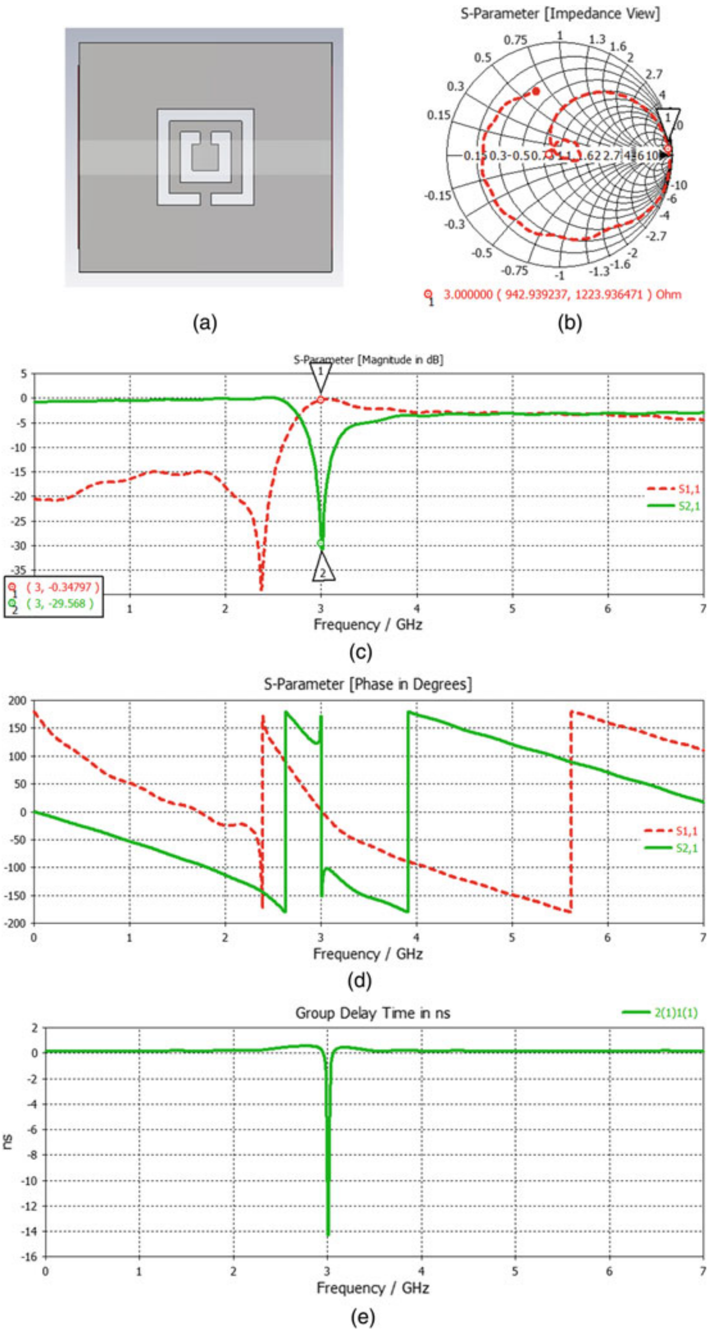


Fig. 13 DSRR-based bandstop filter

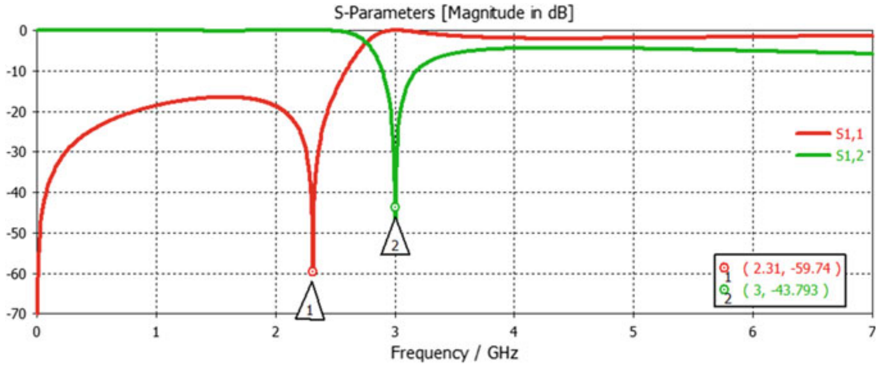


Fig. 14 DSRR BSF circuit simulation

can also be obtained in the similar fashion as discussed above. It also shows negative phase response and a group delay time of −0.088 ns.

3 Comparative Analysis

Table 1 shows a compiled comparative analysis of all the bandstop filters discussed above. Different filter performance parameters are considered while doing the comparison; they are 3-dB bandwidth in GHz (the range of frequencies between the half power points), insertion loss in dB (ratio of signal power delivered to the output terminal and the signal power at input terminal), quality factor (Q) (ratio of energy stored to the energy supplied to the resonator per cycle to keep the amplitude of signal constant at operating frequency (f_r)). It can be obtained as:

$$Q = \frac{f_r}{\text{Bandwidth}}$$

Group delay (τ) in ns (described as change in phase with respect to change in frequency with a negative sign), expressed as:

$$\tau = -\frac{\varphi(\omega)}{\omega}$$

Shape factor quantifies the steepness of the roll-off of a filter. It is the ratio of bandwidth calculated at two different attenuation values. The shape factor obtained in present case is for 14/3 dB attenuation.

In Table 1, it can be observed that the transmission coefficient is minimum for DSRR within a compact geometry and large 3-dB bandwidth but at the expense of shape factor. The shape factor is large for CSTR filter; however, it has large size and low 3-dB bandwidth.

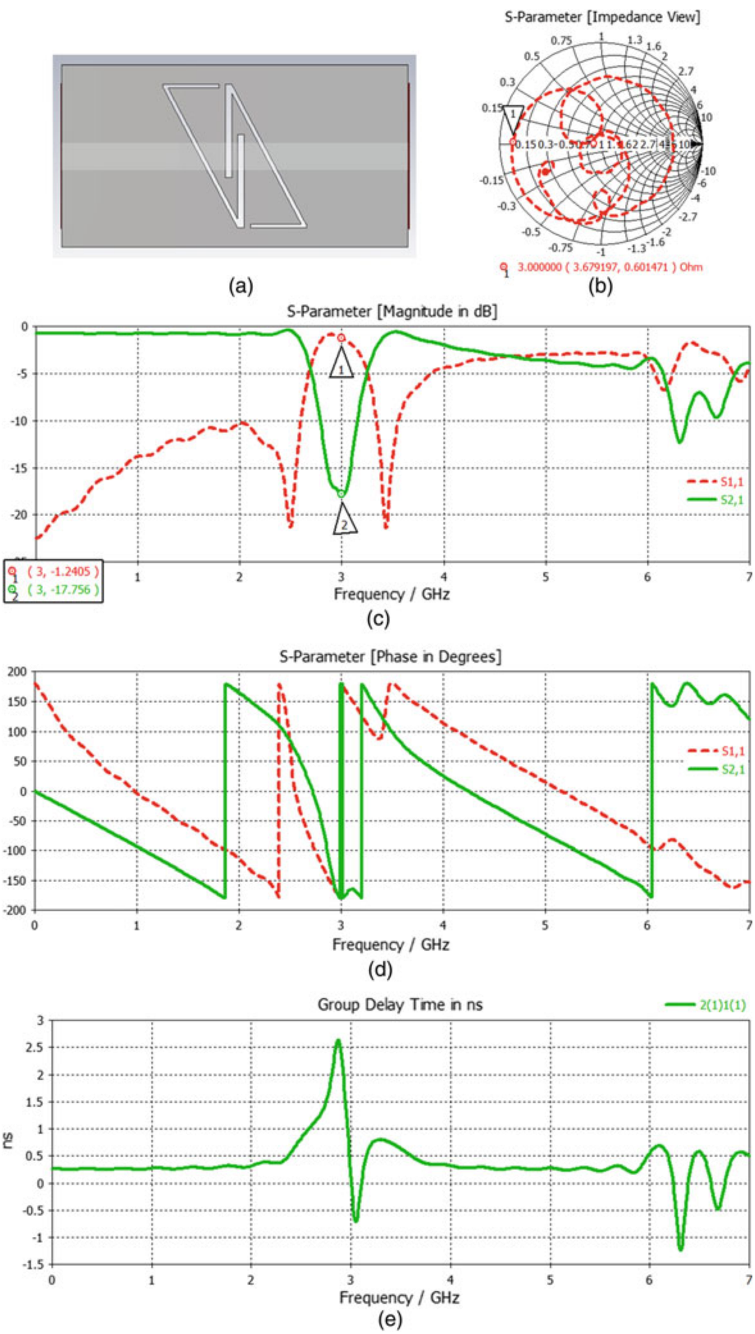


Fig. 15 CSTR-based bandstop filter

Table 1 Performance comparison table

Unit cell	Insertion loss (dB)	3-dB B.W. (GHz)	*Q*-factor	Shape factor	Overall size (mm^2)	Unit cell (mm^2)	Group delay (ns)
CELC-I	−15.06	1.62	1.85	0.125	26 × 14	7 × 12.3	−0.548
CELC-II	−14.24	1.45	2.06	0.045	30 × 28	14 × 21.5	−1.156
OSSRR	−20.24	0.272	11.02	0.238	26 × 14	8.75 × 3.86	−9.97
SSRR	−28.045	1.09	2.75	0.128	22 × 20	8.4 × 8.4	−10.08
DSRR	−29.56	1.68	1.78	0.086	22 × 20	8.4 × 8.4	−14.322
CSTR	−17.75	0.55	5.45	0.467	38 × 20	15.75 × 15.75	−0.088

4 Conclusion

The metamaterial unit cells taken from the literature have been optimized to resonate at 3 GHz, and their performance is investigated as a BSF. The unit cell is etched out of the ground plane of the substrate of the filter with 50-Ω transmission line on the top. The equivalent circuits for the filters are also analyzed using filter theory. The simulated *S*-parameters of the extracted circuits close approximation with the designed bandstop filters. In the end, a comparative analysis table is also presented depicting various performance parameters of the filter such as 3 dB bandwidth, insertion loss, overall size, unit cell size, quality factor (*Q*), shape factor and group delay. The comparative study can helpful for the user to point out the best suitable configuration for a desired application.

References

1. Pendry JB, Holden AJ, Robbins DJ et al (1999) Magnetism from conductors and enhanced nonlinear phenomena. IEEE Trans Microwave Theory Techn 47:2075–2084
2. Falcone F, Lopetegi T, Baena JD, Marques R, Martin F, Sorolla M, Microw IEEE (2004) IEEE Microw Wireless Compon Lett 14:280–282
3. Garcia-Garcia J, Bonache J, Gil I, Martin F, Velazquez-Ahumada MC, Martel J (2006) Miniaturized microstrip and CPW filters using coupled metamaterial resonators. IEEE Trans Microwave Theory Techn 54:2628–2635
4. Jang S-H, Lee J-C (2005) Design of novel cross-coupling elliptic function filters with the miniaturized edge-coupled split ring resonators. Microw Opt Technol Lett 45:495–499
5. Lalj H, Griguer H, Drissi M (2013) Very compact bandstop filters based on miniaturized complementary metamaterial resonators. Wirel Eng Technol 4:101–104
6. Sassi I, Talbi L, Hettak K (2016) Compact bandpass filters based on linked hexagonal-omega resonators. Microw Opt Technol Lett 58:1049–1052
7. Xu H-X, Wang G-M, Pengg Q (2011) Fractal-shaped complementary electric-LC resonator for bandstop filter. Prog Electromagn Res C 23:205–217

8. Karthikeyan SS, Kshetrimayum RS (2010) Composite right/left handed transmission line based on open slot split ring resonator. Microw Opt Technol Lett 52:1729–1731
9. Wang C-J (2016) A low-pass filter based on the slotted ground-plane horn-shaped resonator. IETE Techn Rev 33(2):199–204
10. Wang CJ, Lin CH (2010) Improvement of frequency characteristics for DGS resonators. IETE Techn Rev 27(1):549
11. Öznazı V, Ertürk VB (2008) A comparative investigation of SRR and CSRR-based band-reject filters: simulations, experiments, and discussions. Microw Opt Technol Lett 50:519–523
12. Computer simulation technology microwave studio (CST MWS)
13. Computer simulation technology design studio (CST DS)
14. Bala BD, Rahim MKA, Murad NA (2014) Complementary electric-LC resonator antenna for WLAN applications. Appl Phys A 117:635–639
15. Abessolo MAA, Diallo Y, Jaoujal A, Moussaoui A, Aknin N (2013) Stop-band filter using a new metamaterial complementary split triangle resonators CSTRs. Appl Comput Electromagn Soc J 28:353–358
16. Nornikman H, Ahmad BH, Abd Aziz MZA, Othman AR (2012) Effect of single complimentary split ring resonator structure on microstrip patch antenna design. In: 2012 IEEE symposium on wireless technology and applications (ISWTA), pp 239–244
17. Caloz C, Itoh T (2005) Electromagnetic metamaterials transmission line theory and microwave applications. Wiley-IEEE Press, New York

Home Automation Using Brain–Computer Interface

Utkarsh Raj and Manoj Kumar Mukul

Abstract Brain–Computer Interface (BCI) is a link whichconnects the brain of humans with the outside peripheral devices and machines, and it is a mechanism that allows the users to interact with the outside environment. BCI can be considered an extension of human–computer interaction because BCI is the control of computers by brain waves. The brain waves produced inside the brain can be analyzed with the help of BCI after which simple machine learning algorithms can be used to classify the signals into different classes and hence, map the brain signals into computer commands to operate different external devices. This simple idea is used for home automation to create a basic domotic (Corralejo et al. in A domotic control system using brain-computer interface (BCI). Springer, Berlin, Heidelberg, 2011 [1]) system. This will also be very beneficial for people suffering from neuromuscular disorders like stroke, cerebral palsy and who can't operate their voluntary muscles like those in the hands and legs.

Keywords Brain–computer interface (BCI) · Machine learning algorithms · Domotic system · Neuromuscular disorders

1 Introduction

The brain is one of the most complicated and powerful organs in the human body and along with the spinal cord comprises the Central Nervous System (CNS) which controls the functioning and coordination of the entire human body. Inside the brain is a complex network of billions of neurons firing at each other to produce the desired results giving rise to brainwaves. These brainwaves change as per what we think, feel or do.

U. Raj (✉)
IBM India, Bangalore, Karnataka 560029, India
e-mail: utkarsh77raj@gmail.com

M. K. Mukul
Department of ECE, BIT Mesra, Ranchi, Jharkhand, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_52

For example, when the low-frequency brainwaves are dominant, we feel lazy, sleepy and tired, and on the other hand, the high-frequency brainwaves are produced when we are attentive or alert. Based on the frequency range, brainwaves can be classified into Delta, Alpha, Theta, Beta and Gamma waves [2]. Depending upon the well-established fact that the left part of the body is controlled by the right hemisphere of the brain and vice versa, we can analyze the brainwaves that are generated when we move our left hand and right hand and map these activities to external peripheral devices using BCI. We can achieve home automation by directly controlling simple home appliances using our brains. Simple non-invasive sensors like ones manufactured by Neurosky can be used to collect brainwaves and relay them to the processors to differentiate the signals. Not only this, but we can also use the brainwaves produced when we are in meditative and attentive states which do not require any movement of voluntary muscles and can be used to run and control external devices. This method can be very helpful to those suffering from nervous system disorders [3].

Using the above mechanism, we propose a simple BCI-based home automation system to control devices like fans, lights, switches, etc. For this, we take the motor imagery data set III provided by Graz University to obtain the EEG signals produced when we move our hands and then train an artificial neural network to classify the incoming data correctly and hence control the home appliances.

2 Brainwaves

The synchronized electrical pulses coming from the masses of neurons that are interacting and firing at each other give rise to the brainwaves. Based on their frequency ranges and the reason of origin, the brainwaves are classified as follows:

- Delta Waves—These are high-amplitude brainwaves that are produced during deep sleep stages and the frequency range usually ranges up to 4 Hz [4].
- Theta Waves—These types of brainwaves are common in deep meditation and the frequency range varies between 3 and 8 Hz [5].
- Alpha Waves—These waves are observed during the wakeful state where the subject is in a quiet resting mode and the frequency range varies between 8 and 12 Hz [6].
- Beta Waves—These waves dominate when the person is awake and is totally conscious. These waves are generated during quick activities where one has to engage his/her mind to make decisions, solve problems, etc. The frequency of brainwaves usually lies between 12 and 38 Hz.
- Gamma Waves—These are very high-frequency waves and are the quickest of the brainwaves and this is due to parallel processing of data and information from different parts of the brain. These waves usually lie between 38 and 42 Hz.

3 BCI Block Diagram

Figure 1 shows the basic structure of the BCI system. In our case, we use a non-invasive Neurosky mind wave headgear to acquire the brainwaves [7, 8]. It also performs signal amplification. The raw EEG signal is fed to the processor where the signals are preprocessed and then subjected to feature extraction based on which we train our ANN model to classify the new data. This predicted output is then fed serially to an application interface that interacts with the peripheral devices which in our case was an Arduino Uno microprocessor. This entire system could be totally wireless if we use Bluetooth modules along with the Arduino, hence increasing portability.

A brain–machine interface system is a quite complicated system, it includes 3 basic parts:

1. Means for measuring brainwave signals from the brain
2. A machine learning method that provides us with an algorithm for decoding brain signals
3. Methodology for mapping this decoding to the behavior of action.

3.1 Data Set

We have taken the Motor Imagery data set (Data set III) [9] provided by the Department of Medical Informatics, Institute for Biomedical Engineering, Graz University of Technology. It consists of data from three channels having 140 training and test trials sampled at 128 Hz, each trial is 9 s long.

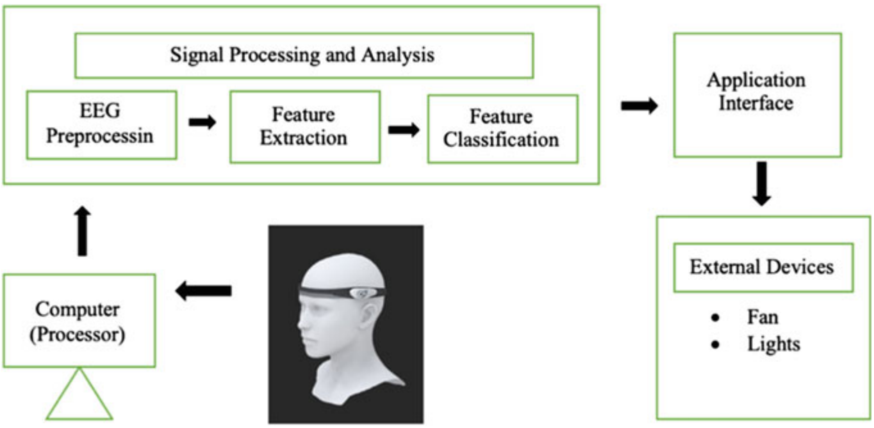


Fig. 1 BCI block diagram

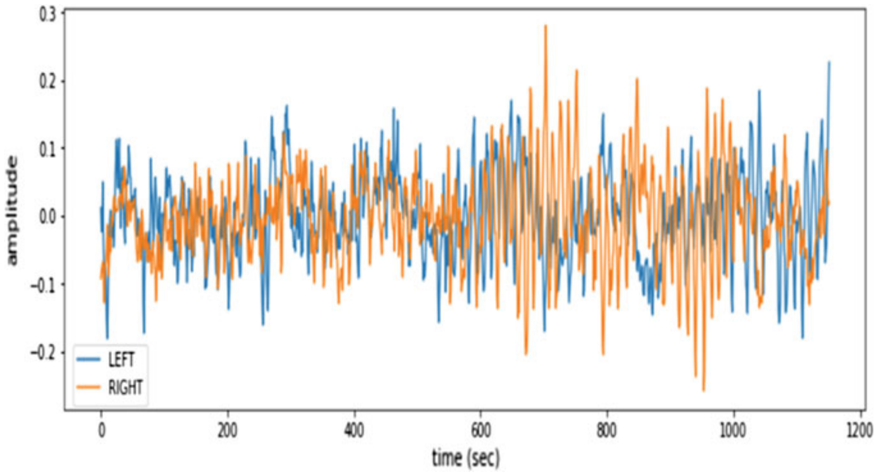


Fig. 2 EEG for left- and right-hand movement (time domain)

This consists of a three-channel data set collected from three points on the brain: left hemisphere, right hemisphere and the central part of the brain corresponding to the brainwaves produced by the movement of the right and the left hands [10].

The recordings correspond to the movement of the left and the right hands, so we take the data from the 2 channels connected to the left and the right hemisphere of the brain controlling the right and the left hands, respectively, and drop the third channel for connection to the central part of the brain.

3.2 Data Analysis

We plotted the data to see the nature of signals from both channels. Figure 2 depicts the time domain plot of one sample data from both channels.

Then we calculated the FFT of the above signal and Fig. 3 shows the plot. We notice that there is a major amplitude difference between the right- and left-hand signals in the delta and alpha range frequency.

We then show the spread of the signals from left and right parts of the brain in Fig. 4.

3.3 Feature Extraction

Since there is a significant variation in the amplitudes of the left- and right-hand movement brainwaves, we calculated the power of each signal in different frequency

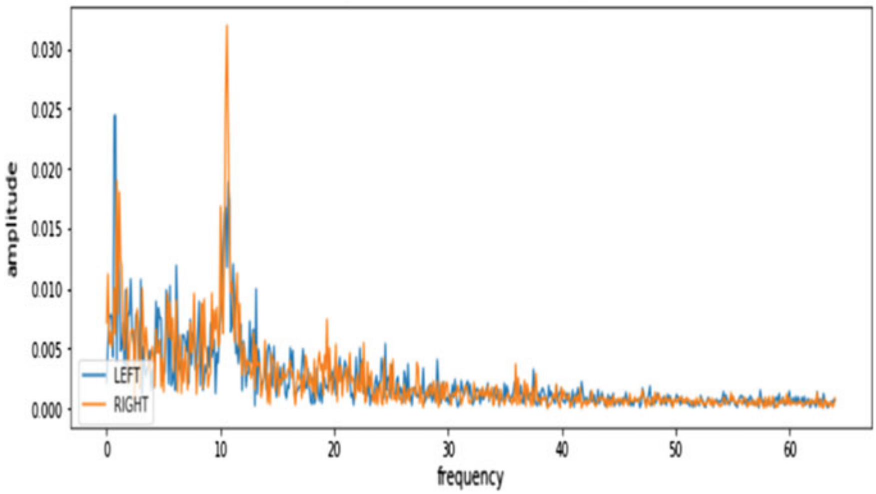


Fig. 3 EEG of the signal in the frequency domain

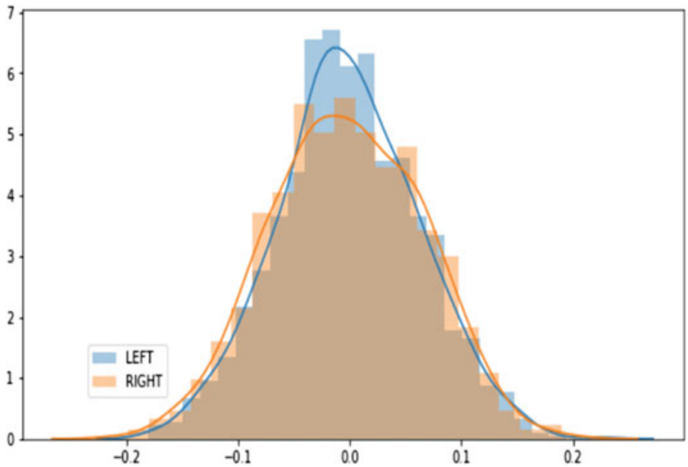


Fig. 4 Spread of the signal about the mean

bands like delta, alpha, theta, beta and gamma. Also, we calculated the mean signal value and standard deviation and hence for each channel we have 7 features. So, each of the 2 classes left and right, is determined by the data coming from both channels, i.e., when we move our left hand, the signals from the right hemisphere of the brain dominate that of the left, but both signals combined characterize the left-hand movement [11]. So, for each hand movement, we have 14 features from 2 channels. Figure 5 depicts the power of the brainwaves for the left and right hands.

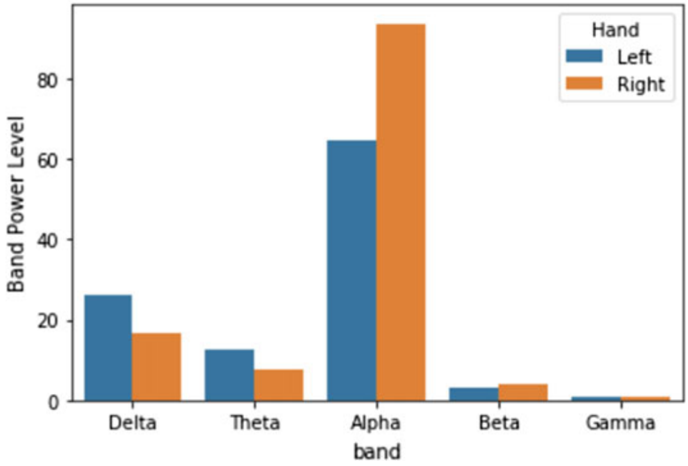


Fig. 5 Band power comparison between left- and right-hand signals

Hence, features to be fed to the classification networks were extracted from the EEG signal [12].

3.4 *Feature Classification*

We used different classifiers for classifying this two-class problem like Decision Trees, Naïve Bayes, SVC, XG boost. We also implemented various component analysis algorithms to reduce the data down to increase accuracy. Finally, we used a neural network for training the model for predicting the results.

4 ANN Model

The ANN model is a 5-layer structure with 3 hidden layers. The number of input nodes in the input layer is equal to the number of features, i.e., 14. The three hidden layers have 8, 12 and 6 nodes, respectively, each having a ReLu activation function. The output layer has 1 node as this is a 2-class classification problem and has a sigmoid activation function.

We used the Adam optimizer and binary cross-entropy as the loss function. We trained the model and ran 500 epochs of a batch size of 5. ANN models can handle data with large features and provide higher accuracy and are very common for EEG classification [13].

4.1 Classification Results

The best results were obtained for the ANN model with the structure as shown in Fig. 6. The accuracies for different classification models are enlisted in Table 1.

We also plotted the accuracy versus error graph and observed that the accuracy remained near unchanged after close to 300 epochs. This is shown in Fig. 7.

We used various kinds of algorithms whose results are displayed in Table 1. We found out that the ANN Model specified in Sect. 4 gives the best results.

5 Implementation

Now that we have processed the raw brainwaves and classified them with a fairly high degree of accuracy, we use the output predicted by the machine learning model to control the external peripheral devices. For this, we feed the output classes (0 and 1 implying left and right, respectively) serially to Arduino which is programmed to control lights and motor (fan) [14]. Figure 8 shows the overall flow of the hardware implementation.

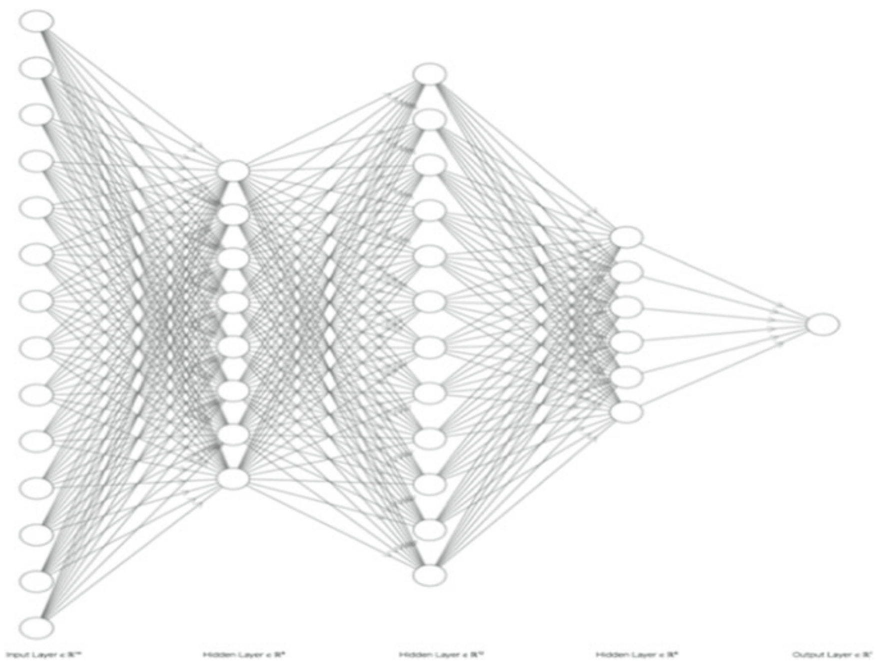


Fig. 6 Implemented ANN model

Table 1 Classification results of various algorithms

	Accuracy (%)
Classification algorithm	
Decision Tree	57.10
Naïve Bayes	52.30
SVC	53.76
XG boost	47.61
Principle component analysis	
XG boost	61.90
SVC	52.30
Linear discriminant analysis	
SVC	66.67
Naïve Bayes	66.68
Decision Tree	54.76
Artificial neural network (ANN)	97.96

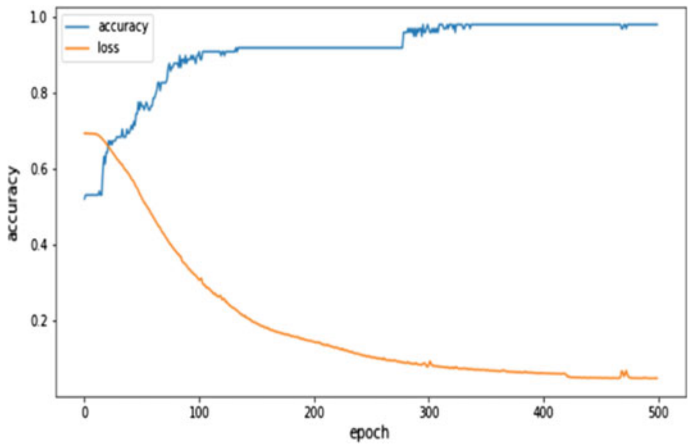


Fig. 7 The plot of model loss and accuracy

The computer and Arduino interact using the UART protocol [15]. Apart from the wired serial data transfer, we can use Bluetooth modules to communicate wirelessly. The Arduino is programmed to operate a 220 V bulb and a stepper motor. The user needs to tell what they want to operate and then control them directly using their brain activity. The light turns on and off as per the arm movement corresponding to on and off. The stepper motor on the other hand is programmed so that it has a minimum and a maximum rpm threshold. Also, the speed increases by a unit of 20 rpm if the user continuously thinks of rotating the motor in that particular direction. Now, if the user thinks of rotating it in another direction, then the motor will start rotating from the minimum set speed again. This way we can quantify the degree of

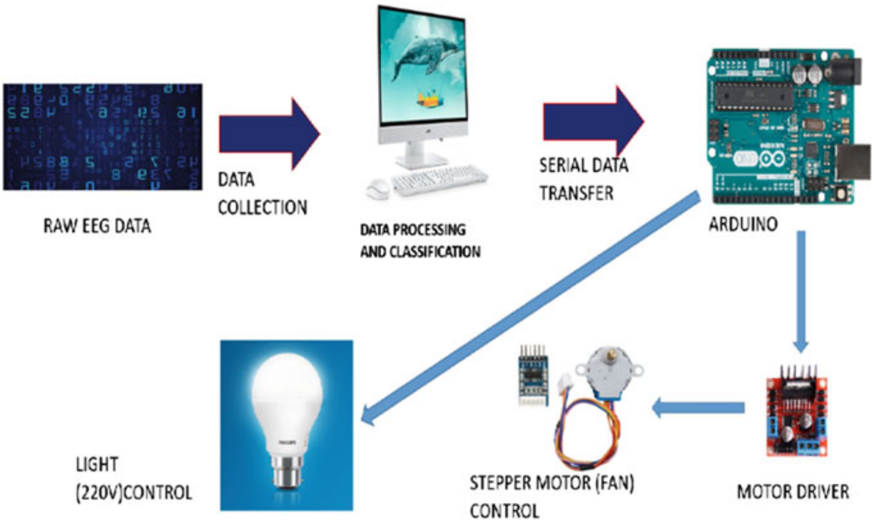


Fig. 8 Proposed system implementation for real-time application

thought, i.e., if a user keeps thinking of turning the motor in one particular direction, then the speed will increase based on how long he keeps thinking of that particular state.

5.1 Comparative Study and Results

Table 2 shows the comparative study of different works with my work. We observe that most models use an SVM classifier and obtain a decent accuracy. They have extracted a large number of features upon which they trained their model. We have

Table 2 Comparative study of different work by other authors

Application type	Processing and feature extraction	Classifier	Accuracy
Subjects identification with imagined speech [3]	Autoregressive coefficients (AR)	SVM	99.76
Subjects identification with VEPs [3]	Autoregressive coefficients (AR)	SVM	98.96
Effect of image stimulus type on neural control of a smart TV [5]	Bandpass filtering	SVM	93.3
BCI arousal detection [16]	Bandpass filtering, asymmetric spatial pattern (ASP)	KNN	82.25
Choice/no choice task [17]	Bandpass filtering, CSP	LDA	80
My work	Power spectral density	ANN	97.96

a small feature set of only 14 features and hence the accuracy with SVM was not that good. Hence, we used ANN which gives us an accuracy of around 97.96%. Also as the number of features increases, the raw signal coming directly from the brain will take a longer time to be model ready (which we can be fed to the classifier) and hence will create a lag in real-time applications. It's more of a trade-off between delay and accuracy: more features might give you higher accuracy but will also increase the delay in real-time applications. Also, most of the work has been done on visually evoked potentials (VEPs) which require a visual stimulus. I used the very basic Power Spectral Density as a feature set that can be very easily calculated. Hence for a feature set, this small, 97.96% can be taken as a good accuracy level. Also, non-invasive sensors [18, 19] won't affect the user and won't have any side effects.

6 Conclusion

Humans have always wanted to wield the power of their minds to do things without moving. This takes things closer to that idea. This BCI was meant to automate a very basic domotic system and was intended to be used by everyone through a certain group of disabled people [20, 21] who will be highly benefitted from this. Right now, we are controlling just simple devices like fans and lights but by using a more complex data set having a greater number of channels associated with different activities, we can automate more complex machines like prosthetics arms and legs as well.

Our aim here is to use the brainwaves to perform such automation with a fairly high degree of accuracy.

References

1. Corralejo R, Hornero R, Álvarez D (2011) A domotic control system using brain-computer interface (BCI). In: Cabestany J, Rojas I, Joya G (eds) Advances in computational intelligence. IWANN 2011. Lecture notes in computer science, vol 6691. Springer, Berlin, Heidelberg
2. Trivedi KR, Thakker RA (2016) Brainwave enabled multifunctional, communication, controlling and speech signal generating system. In: 2016 international conference on electrical, electronics, and optimization techniques (ICEEOT), Chennai. IEEE, pp 4889–4893. https://doi.org/10.1109/ICEEOT.2016.7755650
3. Brigham K, Kumar BV (2010) Subject identification from electroencephalogram (EEG) signals during imagined speech. In: Proceedings of the IEEE fourth international conference on biometrics: theory, applications and systems (BTAS'lO)
4. Malik AS, Amin HU (2017) Designing an EEG experiment. In: Designing EEG experiments for studying the brain. Elsevier Publications. Paperback. ISBN: 9780128111406
5. Kim JJ, Hwang T, Kim M, Oh E, Hwangbo M, Kim M-K, Kim S-P (2013) The effect of stimulus type and distance on neural control of a smart TV. In: 2013 6th international IEEE/EMBS conference on neural engineering (NER). IEEE, pp 1343–1345
6. Kirstein C (2007) Sleeping and dreaming. In: xPharm: the comprehensive pharmacology reference. Elsevier Publications

7. Maskeliunas R, Damasevicius R, Martisius I, Vasiljevas M (2016) Consumer-grade EEG devices: are they usable for control tasks? PeerJ 4:e1746. Published online 2016 Mar 22. https://doi.org/10.7717/peerj.1746
8. On FR, Jailani R, Norhazman H, Zaini NM (2013) Binaural beat effect on brainwaves based on EEG, In: 2013 IEEE 9th international colloquium on signal processing and its applications, Kuala Lumpur, pp 339–343. https://doi.org/10.1109/CSPA.2013.6530068
9. Pfurtscheller G. BCI competition II. Data set III: BCI-experiment data set. Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology, Graz
10. Ackerman S (1992) Discovering the brain. National Academies Press (US), Washington, DC. Available from https://www.ncbi.nlm.nih.gov/books/NBK234151/. https://doi.org/10.17226/1785
11. Kousarrizi MRN, Ghanbari AA, Teshnehlab M, Shorehdeli MA, Gharaviri A (2009) Feature extraction and classification of EEG signals using wavelet transform, SVM and artificial neural networks for brain-computer interfaces. In: 2009 international joint conference on bioinformatics, systems biology and intelligent computing, Shanghai. IEEE, pp 352–355. https://doi.org/10.1109/IJCBS.2009.100
12. Wu T, Yang B, Sun H (2010) EEG classification based on artificial neural network in brain-computer interface. In: Li K, Li X, Ma S, Irwin GW (eds) Life system modeling and intelligent computing. ICSEE 2010, LSMS 2010. Communications in computer and information science, vol 97. Springer, Berlin, Heidelberg
13. Nieva E, Peralta MF, Beltramone D (2013) Home automation by brain-computer interface. https://doi.org/10.4018/978-1-4666-4490-8.ch045
14. Fezari M, Al Dahoud A (2020) Good practice for UART communication using Arduino with applications. https://doi.org/10.13140/RG.2.2.29680.48649
15. Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey B, Kübler A, Perelmouter J, Taub E, Flor H (1999) A spelling device for the paralysed. Nature 398:297–298
16. Huang D, Guan C, Ang KK, Zhang H, Pan Y (2012) Asymmetric spatial pattern for EEG-based emotion detection. In: The 2012 international joint conference on neural networks (IJCNN). IEEE, pp 1–7
17. Jain A, Kim I, Gluckman BJ (2011) Low-cost electroencephalographic acquisition amplifier to serve as a teaching and research tool. In: 2011 annual international conference of the IEEE engineering in medicine and biology society, EMBC. IEEE, pp 1888–1891
18. Allison BZ, Krusienski D (2014) Noninvasive brain-computer interfaces. In: Jaeger D, Jung R (eds) Encyclopedia of computational neuroscience. Springer, New York, NY
19. Jacobsen S, Meiron O, Salomon DY, Kraizler N, Factor H, Jaul E, Tsur EE (2020) Integrated development environment for EEG-driven cognitive-neuropsychological research. IEEE J Transl Eng Health Med 8:2200208. https://doi.org/10.1109/JTEHM.2020.2989768
20. Bhemjibhaih DP, Sanjay GD, Sreejith V, Prakash B (2018) Brain-computer interface based home automation system for paralysed people. In: 2018 IEEE recent advances in intelligent computational systems (RAICS)
21. Sobotnicka E, Sobotnicki A (2018) BCI interface—new opportunities and hopes for the disabled. An overview of available solutions. In: 2018 Baltic URSI symposium (URSI), Poznan. IEEE, pp 91–94. https://doi.org/10.23919/URSI.2018.8406697

Active Learning Pedagogy—Impact on Student Engagement, Scholastic Performance in STEM Courses

Shyamsunder Merugu, R. Archana Reddy, Ramchandar Rao Pamulaparthi, and Tarun Kumar Juluru

Abstract Millennial students have diverse learning styles. Many researchers highlighted that the conventional pedagogical practices were obsolete in Science, Technology, Engineering, and Mathematics (STEM) courses. Therefore, it isn't easy to actively engage students and improve their performance in academics through conventional teaching methods. This paper reports on incorporating active learning activities into electronic courses that give academic improvement and active student engagement that instil creative and critical thinking. We explore the orchestration of dynamic teaching–learning practices in Analog Electronics (AE) and Electromagnetic Theory and Transmission Lines (EMTL) courses. This paper outlines the required methods of student-centric activities. It articulates how these methods are assigned to students and how they perform better in mid-term and end semester examinations. This paper emphasizes on ARCS model, use of ICT in STEM courses, efficient assessment tools, and constructive feedback to students. The results show that these student-centric teaching practices outperform traditional teaching methods and eventually improved academic achievements, evidently shown through CO–PO attainment tables.

Keywords Pedagogical practices · Dynamic learning activities · Power of technology · Constructive feedback

1 Introduction

The students are strongly distinctive these days about ICT tools. The Higher Educational Institutions (HEI) face challenges in fostering the academic performance and effective engagement of millennial students with different learning styles. Much

S. Merugu (✉)
Sumathi Reddy Institute of Technology for Women, Warangal, Telangana, India
e-mail: shyamsunder_m@sritw.org

R. Archana Reddy · R. R. Pamulaparthi · T. K. Juluru
S R University, Warangal, Telangana, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_53

research has proved that the ICT tools are used to help achieve these challenges for educational instructional purposes. They may help the progress of proficiencies in both students and professors. The ICTs offer education several paybacks, like continuous learning, enabling knowledge to be learned and shared [1, 2]. In continuation of learning the course, the students perform well in the formative, summative assessment tests which helps in student learning achievement [3]. The relationship between the student marks and key student learning engagement factors can be attainable by assessment tools [4]. The performance of the individual student team plays a significant role in academic performance. An empirical study was carried on multidimensional students' experience in a group [5]. The quality of e-learning is dependent on how much time the student is engaged. The students are well engaged by the activity based learning through careful designing of the course [6]. The virtual learning environment and assessment-driven learning activities engage the students positively. The style of the instructor humour has a significant impact on effective student engagement, and the kind of humour of the instructor is good-natured and meagre when hostile or sarcastic [7]. When the students have effectively engaged, a student's satisfaction increases, gets motivated to learn and improves performance. Learner-to-Instructor engagement is most important along with Learn-to-Learner, Learner-to-Content in the teaching–learning process [8]. The flipped classroom activity can quickly turn passive learning (in-class lecture) into an active engagement. The students go through the learning material at home and do all the learning activities in the class, in individual and group-based activities on the content under the instructor's supervision. It helps to attain the fruitful learning outcome of the students.

The students will be well engaged in social media, especially WhatsApp. A student group in WhatsApp will be used for the course information, interactivity, longer connectedness, and formal discussion about the posted question [9]. Researchers explained that YouTube is extensively used as a learning resource where the students are engaged with a wide variety of content. It also increases their new findings and interaction with the instructor [10]. However, no scale is predominantly considered to measure how much the users are exclusively envisioned to use YouTube as an educational source. No prototype is available to identify the various impacts of prompting these social intents. YouTube is a content-based online source with a facility for a distinct person to stake their self-made content to other users. Albeit it's treated to be an entertainment medium, it became an appreciable learning educational source. It will be admired as an alternative to web-content-based resources accommodated by course e-blogs and websites [11]. YouTube promotes a lively engagement with an immediate influence on perceived acquisition results in cross-curricular competencies and improves academic performance [12]. Mobile phones, palmtops, laptops, and tablets appear to be enormously used as learning tools because they are convenient [13].

Conduction of various tests on concepts of electronic courses will improve the pedagogical competencies in teaching and learning [14]. The class conceptual survey and homework assignments are used as assessment tools to test pre- and post-conceptual understanding of various essential topics in electromagnetism [15]. The relation between the successful completion of homework and test marks concludes that the higher aptitude skilled students do a more outstanding job than the students with lower aptitude skills. These students can be engaged and helped in completing their homework, but not at deeper learning levels. The effective way of making the students understand electromagnetism is to design such experiments where the students can easily understand the principles, concepts by doing, which can avoid the complexity in learning [16]. An inclusive method to demonstrate electromagnetics at the understudies (UG) level is to arrange the complex, abstract core course as simple as possible with the aid of ICT tools in the teaching atmosphere to assure that the scholars could gain more effectively with hunger and curiosity. The judicious introduction of virtual experiments enhances the students' learning [17].

2 Proposed Methodology-Course Delivery Framework-Pedagogical Design

Before starting the course, the well planned, designed, and balanced course material is given to undergraduate students of Electronics, Sumathi Reddy Institute of Technology for Women, Warangal. The material contains all the information about the course, such as instructor details, POs, PSOs, PEOs, and COs, Bloom's Taxonomy Levelled questionnaire, and diverse learning resources.

The course delivery framework is shown with predominant components (Fig. 1).

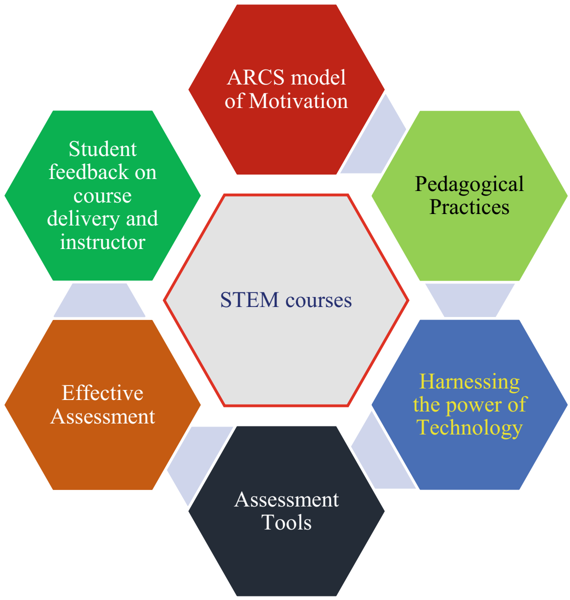


Fig. 1 STEM course delivery framework

Fig. 2 PPT presentation of course

2.1 ARCS Model of Motivation

ARCS motivational model is implemented in course delivery to motivate and engage the students by asking curious questions in their day-to-day life and giving them practical examples. This course helps to find the way we must learn it.

2.2 Pre-survey—Identification of Learning Styles

A survey of 20 questions is conducted to the sample size of 120 students to understand the varied learning sorts of students. To satisfy the training sorts of scholars, various pedagogical practices, ICTs are utilized in delivering the content. Chalk and talk, PPT presentations, video lectures, animations, and dynamic learning activities are conducted. The survey report on the learning styles of the students has been discussed in the discussion section (Figs. 2, 3, 4, 5, 6, and 7).

2.3 Harnessing the Power of Technology

The technology is much used in engaging the students 24 × 7 effectively. A virtual classroom (GOOGLE CLASSROOM) is created. Google forms and simulation software are used for the assessment of the performance of individual students and teams. Virtual Lab links were given and practised for understanding the various principles and concepts. With the help of technology, FLIPPED CLASSROOM activities were conducted, and students are well engaged effectively. The WhatsApp group is created

Fig. 3 Staff self-created YouTube video lectures

Fig. 4 YouTube video on demonstration of TIR

Fig. 5 Think-Pair-Share activity

Fig. 6 Staff interaction during One-Minute activity

Fig. 7 Teaching through ICT tools

for quick information sharing, interaction with the instructor, and discussions among the students. The instructor made self-video lectures on various course topics and made them available on YouTube for students learning engagement (Figs. 8 and 9).

2.4 Assessment Tools

The various assessment tools like quizzes, discussions, and class tests—open and closed books—were considered (Table 1).

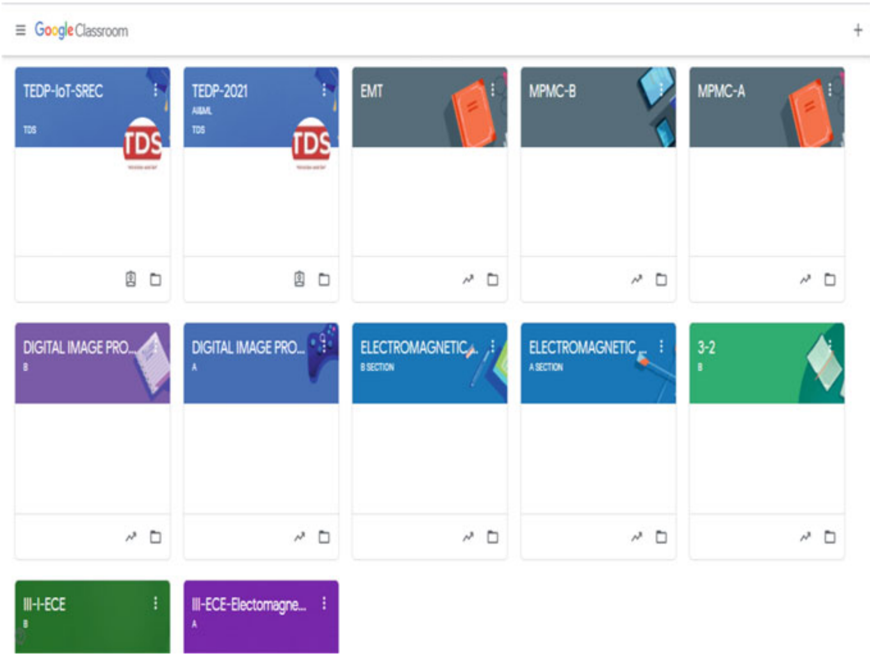


Fig. 8 Google classroom

Fig. 9 WhatsApp group

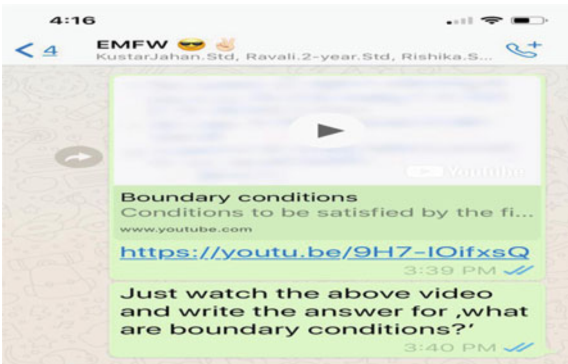


Table 1 Assessment tools used in student engagement

S. No.	Assessment tool	Remarks
1	Quiz	Google form
2	Discussions	WhatsApp
3	Class tests	Closed book
4	Assignments	Open book

2.5 Effective Assessment

The students were given valid reasons for their underperformances in assessment tests. However, the constructive feedback helped them to perform well in College Internal Examinations (CIE), conducted twice. Each time, it is conducted for 25 Marks in which 10 Marks for descriptive, 10 Marks for objective bits, and 5 Marks for assignment. The average of two CIE exams is considered for final internal marks. Including diverse pedagogical practices is to engage students effectively and perform well in semester examinations. A functional assessment is conducted on a sample of 120 students.

2.6 Student Feedback on Course Delivery and Instructor

The success of any dynamic learning activity, quality of resource materials, and student engagement in learning connect with the students' performance index. In addition, the student feedback on various pedagogical practices and the instructor have a prominent role in completing the course from both students and instructor perspectives. The 10-point questionnaire is given to the students for feedback on course delivery and instructor using Survey Monkey standard template (Fig. 10; Table 2).

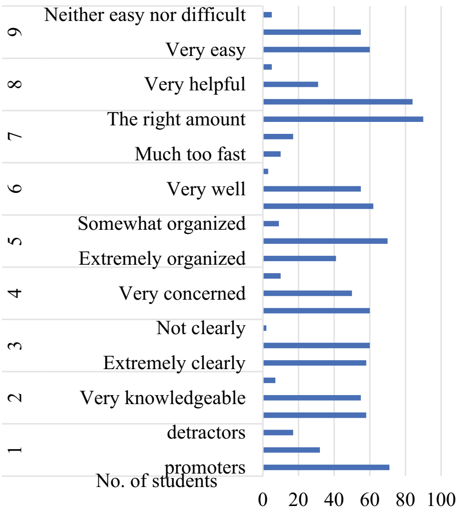


Fig. 10 Horizontal bar charts on course delivery and instructor

Table 2 Feedback on course delivery and instructor

Question	Particular	No. of students
1. How likely is it that you simply would recommend your instructor to a different student?	Promoters	71
	Passives	32
	Detractors	17
2. How far is your instructor knowledgeable?	Awfully knowledgeable	58
	Highly knowledgeable	55
	Fairly knowledgeable	7
3. How is your instructor explaining the course material?	Tremendously clearly	58
	Highly clearly	60
	Not clearly	2
4. How much did your instructor concern about the course material of students' learning?	Enormously concerned	60
	Highly concerned	50
	Fairly concerned	10
5. How well-organized is your instructor for the course?	Tremendously organized	41
	Very well organized	70
	Fairly organized	9
6. The response of your instructor to the questions posed is?	Enormously well	62
	Very well	55
	Not good	3
7. The presentation of the content speed by the instructor is?	Much too fast	10
	Very fast	17
	Right amount	90
8. Are the homework assignments helpful for your understanding of the course material?	Awfully helpful	84
	Very helpful	31
	Fairly helpful	5
9. The availability of your instructor after the class hour is?	Much easy	60
	Easy	55
	Can't say	5
10. The instructor is strong at? Any suggestions to the instructor?	Demonstrations, explanation	–
	Animation videos required to understand very few concepts	–

3 Assessment Analysis and Discussions

(1) *Quiz*: The quiz test is conducted three times with the different levels of easy to difficult from various concepts in the course, and performance is recorded. An individual assessment test is conducted on a scale of 10 points, and a team assessment test is conducted on 20 points. The student team contains three members for team assessment, so there are 40 teams in total (Figs. 11 and 12;

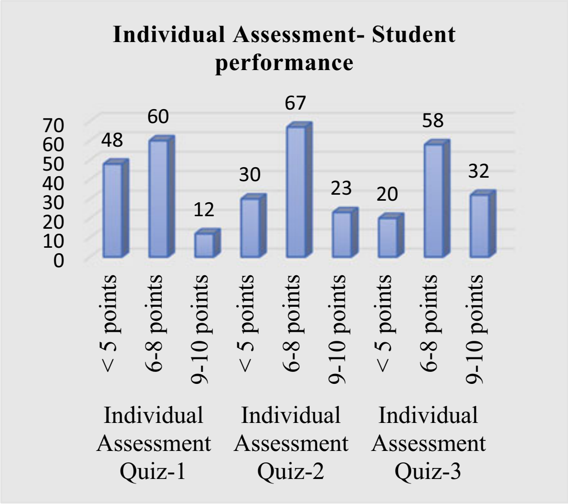


Fig. 11 Individual assessment—student performance

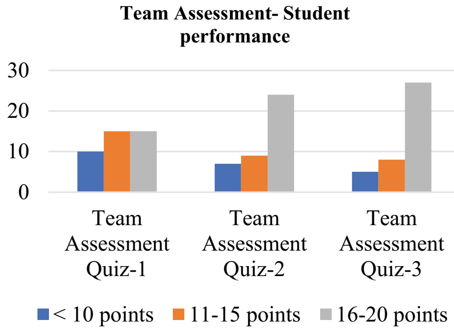


Fig. 12 Team assessment—student performance

Tables 3 and 4).

The individual assessment of the quiz-1 test has a more significant number of underperforming students than the remaining two quiz tests. The analysis is done by taking their feedback on activities and course content. Further improvement is

Table 3 Individual assessment quiz details

Individual assessment quiz 1			Individual assessment quiz 2			Individual assessment quiz 3		
<5 points	6–8 points	9–10 points	<5 points	6–8 points	9–10 points	<5 points	6–8 points	9–10 points
48	60	12	30	67	23	20	58	32

Table 4 Team assessment quiz details

Point scale	Team assessment quiz 1	Team assessment quiz 2	Team assessment quiz 3
<10 points	10	7	5
11–15 points	15	9	8
16–20 points	15	24	27

observed in students' performances. The students were first time exposed to online tests, and they are from rural backgrounds in our case. With the Team assessment of quiz 1, there is a lack of students' coordination within a team of students. Later, the students realized how individual student cooperation leads to team success among all the groups. It is identified that the student teams performed well in subsequent team assessment quiz tests, which are tangible things.

(2) *Group Discussions*: The selected topics were given to students to discuss and share their views. Before conducting the group discussions, the students could search and get data from any reliable learning resources. The entire debate is conducted under the supervision of the instructor. The best performer is appreciated with small prizes to encourage them further.

(3) *Class Tests*: The descriptive class tests were conducted thrice for 10 Marks for each time. It contained four questions in which two questions are to be answered. Each question carries 5 Marks (Fig. 13; Table 5).

The performance of the students in the class tests was observed to improve gradually. Therefore, these class tests are conducted to enhance students' performance in CIE and SEE examinations.

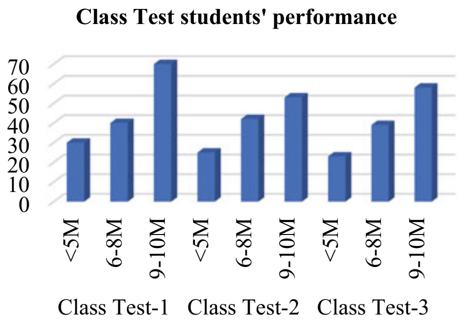


Fig. 13 Students' performance in class tests

Table 5 Class tests marks

Item	Class test-1			Class test-2			Class test-3		
Scores (avg)	<5 M	6–8 M	9–10 M	<5 M	6–8 M	9–10 M	<5 M	6–8 M	9–10 M
No. of students	30	40	70	25	42	53	23	39	58

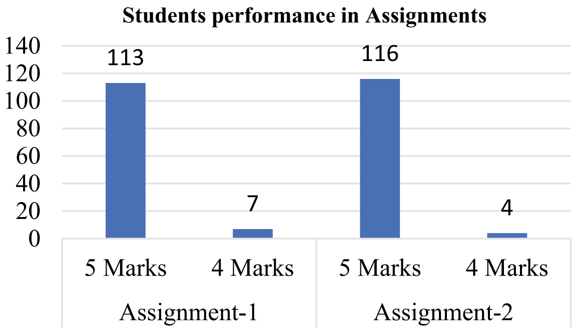


Fig. 14 Students' performance in assignments

Table 6 Assignment marks

Item	Assignment-1		Assignment-2	
Score out of 5 M	5 M	4 M	5 M	4 M
No. of students	113	7	116	4

(4) *Assignments*: The assignments are given as open book activities. The given well-defined RUBRICS was stressed on Accuracy, Timeliness, Completion, and Neatness. The works are evaluated and given constructive feedback as per the RUBRICS. So, there is no scope for showing student partiality, which may not lead to student demotivation. Two assignments were given separately, which may be considered for CIE evaluation. Except for a few students out of 120 got 5 Marks out of 5 Marks (Fig. 14; Table 6).

The class tests, assignments, and individual and team assessment tests are conducted to improve the students' performance in CIE and SEE (Table 7).

For example, in the final examination conducted by JNTU, Hyderabad students scored a 90% above pass result including 15 outstanding performances in Electromagnetic Theory and Transmission Lines course. The exceptional performances are those students who got above 90 Marks out of 100 Marks (Fig. 15).

Table 7 CIE marks

Item	CIE-1 (No. of students)			CIE-2 (No. of students)		
Scores	<5 M	6–8 M	9–10 M	<5 M	6–8 M	9–10 M
Descriptive (10 M)	8	22	90	9	25	86
Objective (10 M)	0	19	101	3	12	105

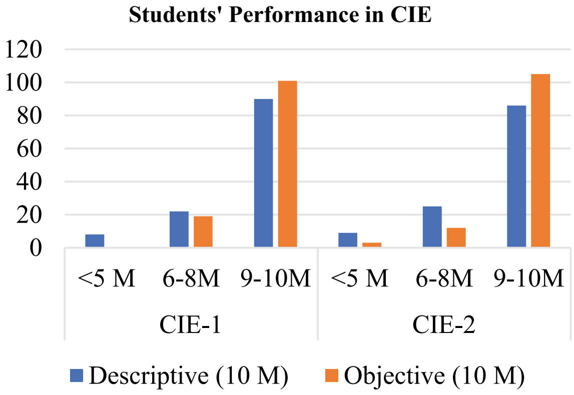


Fig. 15 Students' performance in CIE

4 Attainment of COs–POs of STEM Courses

Tables 8, 9, and 10 give the CO–PO mapping, attainment level of AE course without/with ALA (Active Learning Activities), whereas Tables 11, 12, and 13 give about EMTL course.

Table 8 CO–PO mapping of AE course

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	3	2	3	–	2	–	2
CO2	1	3	3	–	–	–	2
CO3	2	–	3	–	3	–	–
CO4	2	2	–	–	2	–	2
Avg	2	1.75	2.25	–	1.75	–	1.5

Table 9 CO–PO attainment of AE course without ALA

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	1.51	0.85	1.81	–	0.92	–	0.75
CO2	0.52	1.42	1.79	–	–	–	0.63
CO3	0.87	–	1.82	–	1.21	–	–
CO4	0.99	1.02	–	–	0.78	–	0.81

Table 10 CO–PO attainment of AE course with ALA

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	2.75	1.5	2.66	–	1.87	–	1.5
CO2	0.93	2.67	2.78	–	–	–	1.25
CO3	1.90	–	2.85	–	2.65	–	–
CO4	1.79	1.82	–	–	1.53	–	1.75

Table 11 CO–PO mapping of EMTL course

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	3	3	2	–	3	–	2
CO2	2	2	3	2	1	–	2
CO3	2	3	–	–	2	–	2
CO4	3	3	2	3	2	–	2
Avg	2.5	2.75	1.33	2.5	2	–	2

Table 12 CO–PO attainment of EMTL course without ALA

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	1.78	1.38	0.72	–	1.67	–	0.56
CO2	0.86	0.95	1.85	1.01	0.62	–	0.28
CO3	0.83	1.92	–	–	0.78	–	0.42
CO4	1.62	1.45	1.22	1.35	0.83	–	0.93

Table 13 CO–PO attainment of EMTL course with ALA

COs	PO1	PO2	PO3	PO4	PO5	PO6–11	PO12
CO1	2.65	2.72	1.76	–	2.75	–	1.81
CO2	1.73	1.91	2.59	1.56	0.89	–	1.22
CO3	1.81	2.72	–	–	1.82	–	1.31
CO4	2.69	2.58	1.92	2.66	1.76	–	1.11

CO—Course outcome; PO—Program outcome; ATA—Active learning activities

5 Conclusions

This paper cited and analysed the various pedagogical practices used to improve the students' learning engagement and academic performance in Continuous Internal Evaluation (CIE) and Semester-End Examinations (SEE) when practised with STEM courses. For Electromagnetic Theory and Transmission Lines course, Tables 3 and 5 infer the individual assessment of quiz-1 and class tests. From the results, it is evident that when the students were exposed to such practices their academic performance gradually increased. Also, it was observed that the number of students underperforming is reduced. From Table 4, it is identified that the student teams performed well in subsequent team assessment quiz tests, which are tangible things. Hence, at the individual-level and team-level quiz test, the team-level performance has the edge due to the implementation of the pedagogical practice. Table 6 renders that 113 out of 117 students performed well in open book assignment-1 and assignment-2, respectively, with the help of well-structured RUBRICS. In Table 7, the students' performance in CIE-1 and CIE-2 are mentioned. The number of students who performed well increased to a maximum of 105 out of 120 students who scored more than 9 Marks. From the results announced by JNTU, Hyderabad, the result was 90%, and there are 15 outstanding performances in Electromagnetic Theory and Transmission Lines, whereas, in AE, it is 95% with 25 outstanding performances. Tables 8 and 11 show the CO–PO mapping of EMTL and AE courses. Tables 9 and 12 infer that the attainment of outcomes is at a low level. It is evident from Tables 10 and 13 that the inclusion of active learning activities helped scholastic performance through active engagement during the classwork. The discussion assessment tool is much helpful to share knowledge and clarify doubts instantly.

The FLIPPED CLASSROOM, a vigorous learning activity, has a significant problem: the students' learning engagement is highly un-monitored. This problem is to be dealt with carefully. However, the recorded data of course delivery is beneficial for future modifications of the teaching–learning process to achieve greater results. The ARCS model of motivation, dynamic activities, practical assessment, and assessment tools usage proved that it is possible to engage students effectively and improve their academic performance to achieve the course outcomes.

References

1. Hernandez-de-Menendez M, Díaz CE, Morales-Menendez R (2019) Technologies for the future of learning: state of the art. Int J Interact Des Manuf (IJIDeM). https://doi.org/10.1007/s12008-019-00640-0
2. Petrović J, Pale P, Jeren B (2017) Online formative assessments in a digital signal processing course: effects of feedback type and content difficulty on students learning achievements. Educ Inf Technol 22(6):3047–3061. https://doi.org/10.1007/s10639-016-9571-0
3. Strang KD (2016) Beyond engagement analytics: which online mixed-data factors predict student learning outcomes? Educ Inf Technol 22(3):917–937. https://doi.org/10.1007/s10639-016-9464-2
4. Lerís D, Fidalgo Á, Echaluce MLS (2014) A comprehensive training model of the teamwork competence. Int J Learn Intellect Cap 11(1):1. https://doi.org/10.1504/ijlic.2014.059216
5. Nienaber K, Abrams G, Segrist D (2019) The funny thing is, instructor humor style affects likelihood of student engagement. J Scholarsh Teach Learn 19(5). https://doi.org/10.14434/josotl.v19i5.24296
6. Martin F, Bolliger DU (2018) Engagement matters: student perceptions on the importance of engagement strategies in the online learning environment. Online Learn 22(1). https://doi.org/10.24059/olj.v22i1.1092

7. Holmes N (2017) Engaging with assessment: increasing student engagement through continuous assessment. Active Learn High Educ 19(1):23–34. https://doi.org/10.1177/1469787417723230
8. Hakim L, Sinaga P, Setiawan A (2019) The characteristics of development teachers' pedagogy competencies of electromagnetics test. J Phys Conf Ser 1280:052021. https://doi.org/10.1088/1742-6596/1280/5/052021
9. Zhou S, Yeung Y-Y, Wang Y, Wang X, Xiao H (2014) Integrating effective pedagogies in science education with a design of alternative experiments on electromagnetics. EURASIA J Math Sci Technol Educ 10(1). https://doi.org/10.12973/eurasia.2014.1011a
10. Mukhopadhyay SC (2006) Teaching electromagnetics at the undergraduate level: a comprehensive approach. Eur J Phys 27(4):727–742. https://doi.org/10.1088/0143-0807/27/4/005
11. Mady MA, Baadel S (2020) Technology-enabled learning (TEL): YouTube as a ubiquitous learning aid. J Inf Knowl Manag 2040007. https://doi.org/10.1142/s0219649220400079
12. Chintalapati N, Daruri VSK (2017) Examining the use of YouTube as a learning resource in higher education: scale development and validation of TAM model. Telemat Inform 34(6):853–860. https://doi.org/10.1016/j.tele.2016.08.008
13. Orús C, Barlés MJ, Belanche D, Casaló L, Fraj E, Gurrea R (2016) The effects of learner-generated videos for YouTube on learning outcomes and satisfaction. Comput Educ 95:254–269. https://doi.org/10.1016/j.compedu.2016.01.007
14. Fidalgo-Blanco Á, Sein-Echaluce ML, García-Peñalvo FJ (2017) APFT. In: Proceedings of the 5th international conference on technological ecosystems for enhancing multiculturality—TEEM 2017. https://doi.org/10.1145/3144826.3145433
15. Aşıksoy G, Özdamlı F. Flipped classroom adapted to the ARCS model of motivation and applied to a physics course. EURASIA J Math Sci Technol
16. Mhlana S, Krauss K (2017) The impact of information and communication technology in education: gender issues. In: 2017 international conference on infocom technologies and unmanned systems (trends and future directions) (ICTUS), Dec 2017. https://doi.org/10.1109/ictus.2017.8286090
17. Prahl K (2016) Best practices for the think-pair-share active-learning technique. Am Biol Teach 79(1):3–8. https://doi.org/10.1525/abt.2017.79.1.3

Analysis on Image Compression for Multimedia Communication Using Hybrid of DWT and DCT

Aamir Junaid Ahmad, Syed Danish Hassan, Rahul Priyadarshi, and Vijay Nath

Abstract Image compression is a technique to compress information that converts the original image with limited bits and reconstructs it with some distortion. This compression leads to native disc memory saving; conjointly, it helps in speedy multimedia system files transfer. This analysis presents a projected technique for the compression of multimedia system pictures victimization hybrid compression technique (discrete trigonometric function remodel and separate moving ridge transform). DWT and DCT Area unit applied initial on individual RGB parts. Once applying this, image is amount for every distinctive amount. Therefore, on verifying the distinctive code for every distinctive image for his or her cryptography, the standard of the new image compression technique is evaluated victimization the height signal to noise magnitude relation (PSNR) through simulations victimization MATLAB. The simulation results show that the projected JPEG codec delivers increased performance over the standard JPEG code in terms of PSNR.

Keywords DCT · DWT

A. J. Ahmad (✉)
Maulana Azad College of Engineering and Technology, Aryabhatta Knowledge University, Patna, Bihar, India
e-mail: aamirjunaid5@gmail.com

S. D. Hassan
Electronics and Communication Engineering Discipline Science, Engineering and Technology School, Khulna University, Khulna 9028, Bangladesh

R. Priyadarshi
Department of Electronics and Communication Engineering, National Institute of Technology Patna, Patna, India

Present Address:
V. Nath
Department of Electronics and Communication Engineering, Birla Institute of Technology Mesra, Ranchi, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_54

1 Introduction

Compression may be a technique that compresses the information to be transmitted. The first image is encoded with few bits and reconstructed with some distortion. The aim of compression is to scale back the redundancy of the first image by taking advantage of human's complicated sensory system. The most drawbacks of DCT area are unit overview of false contouring special effects and interference artifacts' at advanced compression, while DWT is demand of enormous process properties. As such, demand arises with efficiency store or transmits information and increase transmission rates with smaller information measure consumption. So as to satisfy these wants, totally different strategies of still compression, each loss and lossless, are developed. In lossless compression, there is no loss of data. It reduces bits by finding and eliminating redundancy. In lossless compression, each single little bit of information that was originally within the same once the file is uncompressed. All of the knowledge is totally remodeled. The foremost common image formats that use lossless compression is GIF (graphical interchange format) and BMP (bitmap file). Lossy compression technique reduces the scale {of data of knowledge of info} by finding non-essential information and destroys it. Lossy compression reduces a file by for good eliminating bound info, particularly redundant info. Once the file is uncompressed, solely a vicinity of the first info remains. During this case, the ensuing image is predicted to be like the original image associate example of a picture format that uses this compression technique is JPEG (joint photographic consultants group) [1, 2].

2 Motivation

Because it is that of data and technology, an oversized variety of multimedia system information are transmitted through Web and alternative wireless networks. However, of this multimedia system, information like image and video needs massive information measure and needs brooding angina storage space; as a result, it is fascinating to characterize the info in the information with significantly less bits by the mean of information compression methods. At an equivalent time, we want such quite information compression technique that has to be able to reconstruct {the information |the info |the information} terribly on the point of original data. The DCT and DWT Area unit the most unremarkably used procedures. The DCT has great energy compaction property and needs less process resources. On the opposite hand, DWT may be a multi-resolution remodel, and moveable compression will be simply attained. The most drawbacks of DCT area unit overview of false contouring special effects and interference artifacts' at advanced compression, and that of DWT is demand of enormous process properties. So, the idea of reconnoitering the blessings of each procedure actuated the U.S.A. to examine grouping of DWT and DCT procedures. Such grouping of two algorithmic programs is denoted as 'hybrid' algorithm.

3 Literature Review

Sharma et al. in their paper, "Image Compression victimization hybrid of DWT, DCT and Huffman Coding," given a hybrid compression technique (DWT) DCT and Huffman coding to win higher compression rates. Their effect showed that the cryptography performance is considerably improved with their projected method [3]. Kaur et al. in their analysis paper, "DWT based mostly HYBRID compression FOR swish web TRANSFERS," have applied compression supported DWT and compared it with the existing system. Finally, they noticed that the projected compression technique is healthier in terms of PSNR, MSE, CR etc. for Swish Web transfers [3]. Thilepa and Thanikachalam mention the image process technique victimization MATLAB that establish the fault presents on the materials. For this image process technique, initial image is taken. Noise filtering bar chart and thresholding techniques area unit are applied on the image, and therefore, the output is obtained. Initial input color cloth fault image to the MATLAB in image process system is completed, and then, color image to gray image conversion is done. Noise removal and filtering from the image are completed. Binary image conversion from the noise removed output is completed, and then, bar chart output is obtained. Thresholding technique is applied then [4].

Mohammed and Hussein expected a hybrid schemes for operational compression that has gained huge quality among researchers. To achieve great compression rates, a mix of DWT and DCT is used as a hybrid compression technique at totally dissimilar threshold worth on DWT rotten sub-bands. When DCT remodel is applied on hectoliter, and luteinizing hormone sub-bands it conserved the standard of the decompressed medical image. The sub-bands area unit are the amount basis on threshold values. Finally, the entropy cryptography is applied to amount sub-band. Finally, the entropy cryptography is applied to amount sub-band. The cryptography presentation enclosed compression rate, compression magnitude relation, and image quality.

Simulation outcomes showed that the cryptography performance is considerably upgrade by the hybrid DWT and DCT method [5]. Corinthios et al. [6] projected better hybrid technique. Hybrid technique is a mixture of DWT and DCT. The comparative outcomes for DCT, DWT, and hybrid DWT–DCT compression techniques area unit are assumed. The quality is shown by victimization PSNR. DWT 2 threshold technique offers higher feature parallel to DCT and one threshold DWT. DWT 2 threshold is additionally referred to as "improve DWT." Materials and strategies distinct trigonometric function remodel are

- Original image is split into blocks of eight $\times$ eight.
- Constituent standards of a black and white image vary from 0 to 255; however, DCT is meant to figure on constituent values, starting from -128 to 127. Therefore, every block is changed to figure within the vary.
- DCT is applied to each block by multiplying the improved block with DCT matrix on the left and transpose of DCT matrix on its right.
- Every block is then compressed by quantization.
- Entropy encryption is then wont to quantize matrix.

- Compressed image is restored through reverse method.
- Inverse DCT is employed for decompression [7].
- The supply image is digitized into a sign that may be a string of numbers.
- The signal is rotten into a sequence victimization hear ripple coefficients.
- Threshold is employed to switch the ripple coefficients.
- Quantization is employed to convert sequence.
- Entropy encryption is applied to convert [8].
- Hybrid DWT–DCT algorithmic rule image of or any resolute provided by thirty two is first divided into blocks of $N \times N$.
- Every block is rotten victimization 2-D: Consequent stage wherever the high frequency constants (HL, LH, and HH) are rejected. Passed LL elements are any another 2-D DWT.
- 8-point DCT is applied to the DWT Coefficients image is reconstructed by following the inverse procedure.

4 Conclusion

Finally, we have a tendency to conclude that, the planned JPEG Codec based mostly on hybrid DWT (discrete ripple transform) and DCT (discrete trigonometric transform) function provides higher performance over standard JPEG with higher peak signal to noise magnitude relation [9, 10]. For DCT, 22 decibel on the average and for DWT a mean price of −32 dB are achieved. On the opposite hand, the planned DWT and DCT technique gave a mean of 36 decibel, that is way over DCT or DWT applied one by one. Thus, increased quality of the reconstructed image is obtained on victimization of the planned JPEG Codec.

5 Results and Discussion

The PSNR value swings between −36 to −29 dB and 32 to 41 dB are shown in Fig. 1.

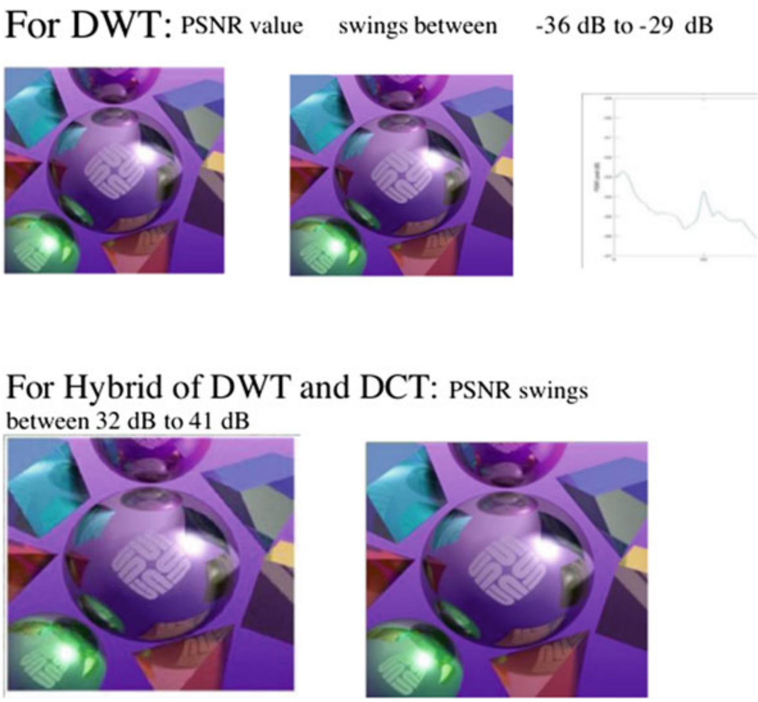


Fig. 1 For DCT: PSNR value swings between −36 and −29 dB, and PSNR swings between 32 and 41 dB

References

1. http://en.wikipedia.org/Image_compression
2. Ames G (2002) Image compression, 07 Dec 2002
3. https://scholar.google.com/scholar?q=recent+research papers on dct dwt quantization&hl=en&as_sdt=0& as_vis=1&oi=scholart#d=gs_qabs&u=%23p%3DmgggvgLjfb gJ. Date: 19/03/2019, 11:00 AM
4. Thilepa R, Thanikachalam M (2010) Analysis of compression algorithmic rule victimisation DCT. J Eng Anal Appl (IJERA) 2(1):515–552
5. Mohammed AA, Hussein JA (2010) Hybrid remodel secret writing theme for medical image application. In: 2010 IEEE international conference on signal process and knowledge technology (ISSPIT), Dec 2010, pp 237–240
6. Corinthios, Benchikh S, Michael (2011) A hybrid compression technique supported DWT and DCT transforms. IEEE, Canada, pp 1049–1065
7. https://www.techradar.com/news/computing/all-youneed-to-know-about-jpeg-compression-586268/2. Date: 29/03/19, 11:43 PM
8. http://en.wikipedia.org/wiki/Discrete_wavelet_transform
9. http://www.3.interscience.wiley.com
10. http://www.vectorsite.net/ttdcmptwo.html

11. http://ijsetr.org/wpcontent/uploads/2015/01/IJSETVOL-4-ISSUE-1-192-198.pdf. Date: 19/03/2019, 12:34 PM. https://www.academia.edu/1291979/Comparative_Analysis_between_DCT_and_DWT_Techniques_of_Imagee_Compression?auto=download. Date: 29/03/19, 10:43
12. https://www.maketecheasier.com/how-image-compression-works/. Date: 29/03/19, 8:43 PM

Face Detection Using Convolutional Neural Network

Abhishek Pandey, Deepak Prasad, K. Kushwanth Reddy, K. Venkatesh, Ajay Chand, and Vijay Nath

Abstract In this paper, a face detection system based on neural network has been reported. It is a novel hybrid face-recognition approach which is based on a convolutional neural network architecture, and it is designed to robustly detect the highly variable face patterns. Automatic recognitions of the human faces are a bit concern in the development and application of face pattern recognition, where in this paper, human faces are identified in cluttered scenes which are generally based on neural nets. Face detection is complex and a challenging task since human faces may appear different in different orientations and in various head poses. And even become complex due to various changing parameters like light intensity, facial expressions, face shadows, etc.

Keywords Face detection · Convolutional neural networks (CNN) · Convolution operation · Pooling types · Activation function

1 Introduction

Face recognition is a challenging task to perform on a machine. Over the last five to ten years, it has been one of the most popular areas of research in computer vision [1]. Due to its wide area of application, it is considered as one of the most promising areas of computer vision. Face detection algorithms are formulated using template-matching approaches that are focused on detecting properties of images such as intensity, face color, and edges [2]. The main challenge in training face

A. Pandey (✉)
Faculty of Engineering and Applied Science, Usha Martin University, Ranchi, JH, India
e-mail: abhishek@umu.ac.in

D. Prasad
Department of Electronics and Communication Engineering, Sarala Birla University, Ranchi, JH, India

K. Kushwanth Reddy · K. Venkatesh · A. Chand · V. Nath
Department of Electronics and Communication Engineering, Birla Institute of Technology Mesra, Ranchi, JH, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_55

detection algorithms is getting a large database that contains thousands of images [3]. The purpose of this research is to construct a face detection model such a way that it can accurately match or predict the subject's face (image) from our database. Input data (image) to recognition system can be either from the camera footage or from the database itself [4]. The essential parts for human facial recognition consist of eyes, skin color, nose shapes, and mouth shapes. After that, the input data is uploaded to recognition system, the image data is preprocessed first, where the facial data images are of in random shapes. Those are made uniform for easy processing, and then, the colored facial images are converted into grayscale images for easy computation; normalization is done to reduce color complexities [5]. Then, the image data is transformed using convolutional neural network (CNN), which is a class of deep learning neural networks and match with the database referencing the similarity measurement of both images. CNN removes most of variation in lighting and contrast and can reduce intrapersonal variation.

The remainder of the paper is organized as follows. Convolutional neural network (CNN) is discussed in Sect. 2. In Sect. 3, demonstration of the components and its convolutional neural network has been done. The obtained results are shown in next section. Finally, the paper ends with the conclusion.

2 Convolutional Neural Network

Neural networks also called as artificial neural networks (ANN) are computational processing systems which are inspired from human nervous system. Neural networks are similar to human brains and can be thought in terms of neural activation and strength of connection between each pair of neurons. It is also called as feedforward neural network [6]. The basic architecture of neural network is shown in Fig. 1. Neural networks are mainly composed of high number of interconnected neurons which work collectively to optimize the output.

The input is in the form of multidimensional vector which is loaded into input layer. The input will be sent to the hidden layers and the number of hidden layers can

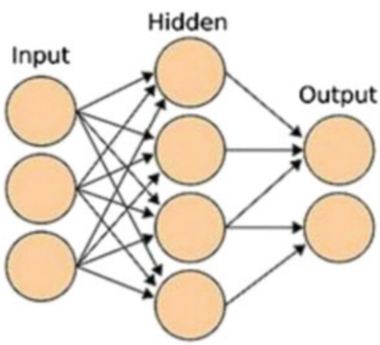


Fig. 1 A simple three-layered neural network comprised of input, hidden, and output layer

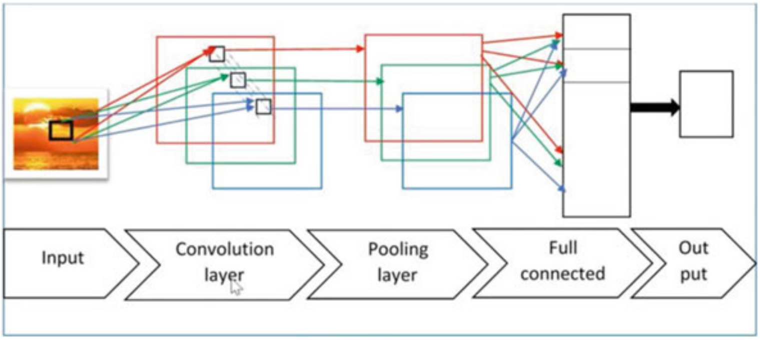


Fig. 2 A simple CNN architecture consisting of different layers

be varied according to the requirement. The hidden layers take decision from previous layers if there are more than one hidden layer. Having multiple hidden layers stacked upon each other is commonly known as deep learning. Convoluted neural networks are similar to artificial neural networks. Instead of being composed of neurons, CNNs are learning systems that are self-aware. They are used for extracting patterns and images from data [4].

3 CNN Architecture

The three important layers which convolution neural network comprised of are convolution layer, pooling layers, and fully connected layer. When all these layers are stacked up, convolutional neural network architecture is formed as shown in Fig. 2. The input layer takes the 2D image as input. Pixels are the smallest indivisible segments, and an image is comprised of these pixels. The strength of pixel is called as pixel intensity. This network sees image as a matrix of numbers where each value represents pixel intensity. The output of convolutional layer is sent to the pooling layer which will reduce the parameters to be trained by down sampling which is further reducing dimensionality of the input. The fully connected layer processes the output of pooling layer [7]. This layer is similar to that of ANN where flattening happens, and then, activation function is applied and sent to further layers.

4 Convolutional Layer

This is the first layer of the CNN network. Convolution operation happens in this layer. Convolution operation is one of the fundamental blocks of CNN. The convolution layer consists of one or more kernels with different weights that are used to extract feature from the image. Convolution is a linear operation which involves

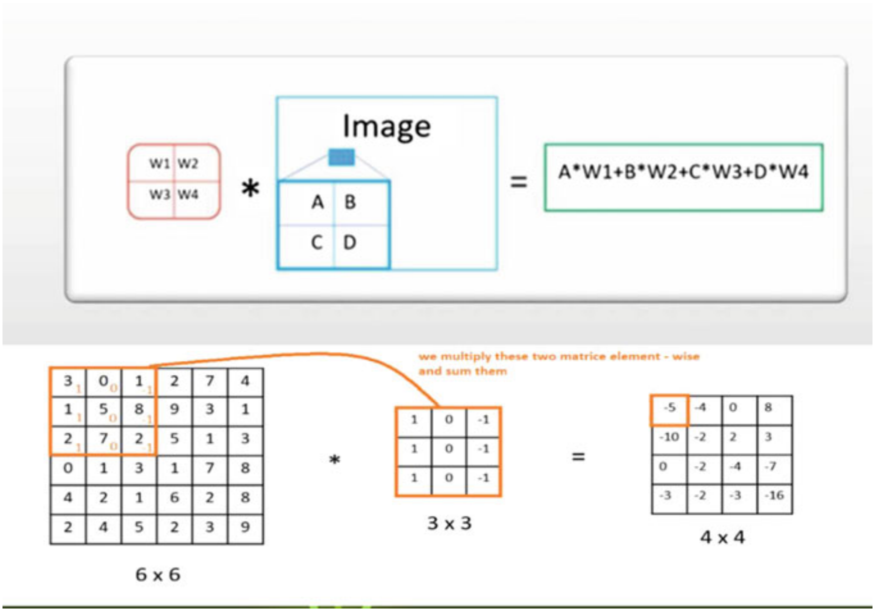


Fig. 3 Convolution operation with an example

the multiplication of a set of weights with the input which is similar to traditional neural network. The multiplication is performed between an array of input data and a two-dimensional array of weights which is generally called filter or kernel [8]. As the filter is applied multiple times to the input array, the result is a two-dimensional array of output values that represents the filtering of the input. The output array is called feature map which summarizes the presence of detected features in the input image (Fig. 3).

5 Pooling Layer

Pooling layer is another building block of CNN architecture. The pooling layer processes the output of convolutional layer. The pooling layer helps in discarding some unimportant information by reducing the dimensionality which further reduces the number of parameters and computation in the network. There are two types of pooling. They are max pooling which returns the maximum value from the portion of the image covered by pooling kernel, and the other one is average pooling which returns the average of values covered by pooling kernel. The most common approach is max pooling (Fig. 4).

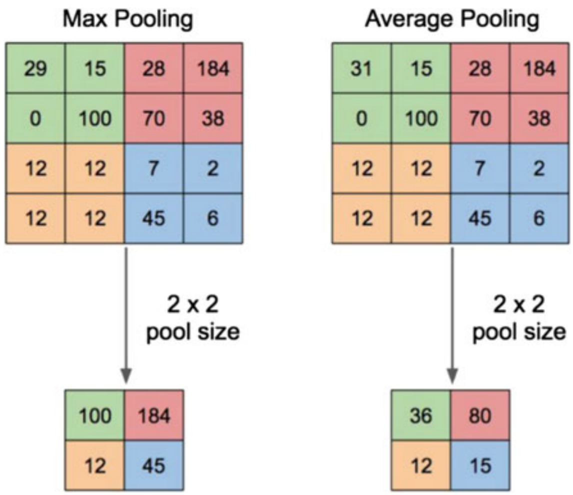


Fig. 4 Working of max pooling and average pooling

6 Code

See Fig. 5.

7 Result

See Fig. 6.

8 Conclusion and Future Research

The proposed algorithm can detect faces with an accuracy of 80–90% in a set of 400 test images, with 5000 images consisting of images with face, and images without face used for training the classifier. There are many directions for future work of our project. The main limitation of our proposed current system is that it can only detect faces in images only. Our future work will include detection of faces in video which will have huge benefit in security surveillance. Personalized emoji creation similar to human faces.

```python
import cv2
from PIL import Image
import numpy as np
import os
from tensorflow import keras
from tensorflow.keras.preprocessing.image import ImageDataGenerator
import matplotlib as plt
```

```python
base_dir = r"C:\Users\user\Documents\anaconda scripts\project 2.0\Face_Database"
train_dir=os.path.join(base_dir,"training_set")
validation_dir=os.path.join(base_dir,"cross_val_set")
names = os.listdir(train_dir)
print(names)
```

```python
train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=40,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest')
validation_datagen = ImageDataGenerator(
    rescale=1./255,
    fill_mode='nearest')
train_generator = train_datagen.flow_from_directory(
                  train_dir,
                  target_size=(150,120),
                  batch_size=20,
                  color_mode="rgb",
                  class_mode='categorical')
validation_generator=validation_datagen.flow_from_directory(
                  validation_dir,
                  target_size=(150,120),
                  batch_size=16,
                  color_mode="rgb",
                  class_mode='categorical')
```

```python
model = keras.Sequential([
    keras.layers.Conv2D(16,(3,3),activation='relu',input_shape=(150,120,3)),
    keras.layers.MaxPooling2D(2,2),
    keras.layers.Conv2D(32,(3,3),activation='relu'),
    keras.layers.MaxPooling2D(2,2),
    keras.layers.Conv2D(64,(3,3),activation='relu'),
    keras.layers.MaxPooling2D(2,2),
    keras.layers.Conv2D(64,(3,3),activation='relu'),
    keras.layers.MaxPooling2D(2,2),
    keras.layers.Flatten(),
    keras.layers.Dense(units=128, activation='relu'),
    keras.layers.Dense(units=128, activation='relu'),
    keras.layers.Dense(units=4, activation='softmax')
])
model.compile(optimizer='adam', loss='categorical_crossentropy',metrics=['accuracy'])
```

Fig. 5 Code for proposed model

```
history=model.fit_generator(
    train_generator,
    steps_per_epoch=50,
    epochs=10,
    validation_data=validation_generator,
    validation_steps=6,
    verbose=2)
```

```
Epoch 1/10
50/50 - 16s - loss: 1.3222 - accuracy: 0.3810 - val_loss: 0.8952 - val_accuracy: 0.5000
Epoch 2/10
50/50 - 15s - loss: 0.8991 - accuracy: 0.6150 - val_loss: 0.4191 - val_accuracy: 0.8646
Epoch 3/10
50/50 - 15s - loss: 0.7373 - accuracy: 0.6990 - val_loss: 0.1619 - val_accuracy: 0.9583
Epoch 4/10
50/50 - 15s - loss: 0.4120 - accuracy: 0.8270 - val_loss: 0.3442 - val_accuracy: 0.8229
Epoch 5/10
50/50 - 15s - loss: 0.3028 - accuracy: 0.8790 - val_loss: 0.0310 - val_accuracy: 1.0000
Epoch 6/10
50/50 - 15s - loss: 0.2382 - accuracy: 0.9120 - val_loss: 0.0117 - val_accuracy: 1.0000
Epoch 7/10
50/50 - 15s - loss: 0.1335 - accuracy: 0.9500 - val_loss: 0.0097 - val_accuracy: 1.0000
Epoch 8/10
50/50 - 16s - loss: 0.0763 - accuracy: 0.9700 - val_loss: 0.0070 - val_accuracy: 1.0000
Epoch 9/10
50/50 - 15s - loss: 0.2092 - accuracy: 0.9220 - val_loss: 0.0054 - val_accuracy: 1.0000
Epoch 10/10
50/50 - 16s - loss: 0.0535 - accuracy: 0.9850 - val_loss: 9.3041e-04 - val_accuracy: 1.0000
```

Fig. 6 Simulation results

References

1. Wiskott L, Fellous J, Kruger N, der Malsburg CV (1997) Face recognition by elastic bunch graph matching. IEEE Trans Pattern Anal Mach Intell 19(7):775–779
2. Lajevardi SM, Wu HR (2012) Facial expression recognition in perceptual color space. IEEE Trans Image Process 21(8):3721–3733
3. Yang MH, Abuja N, Kriegman D (2000) Face detection using mixtures of linear subspaces. In: Proceedings fourth IEEE international conference on automatic face and gesture recognition 2000, pp 70–76
4. Yang G, Huang TS (1994) Human face detection in a complex background. Pattern Recogn 27(1):53–63
5. Sung KK, Poggio T (1998) Example-based learning for view-based human face detection. IEEE Trans Pattern Anal Mach Intell 20(1):39–51
6. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86
7. Yang MH, Kriegman DJ, Ahuja N (2002) Detecting faces in images: a survey. IEEE Trans Pattern Anal Mach Intell 24(1):34–58
8. Wang YQ (2014) An analysis of the Viola-Jones face detection algorithm. Image Process On Line 4:128–148

Utilization of Smartphone-Based Wireless Sensors in Agricultural Science: A State of Art

Rahul Priyadarshi, Pragya Bhardwaj, Prabal Gupta, and Vijay Nath

Abstract Smartphones have become an vital contraption in agribusiness when you consider that their movability arrange with developing, the price of the tool is relatively on hand, and their getting ready power lets in a grouping of all the way down to earth applications to be made. Similarly, mobile phones are nowadays supplied with diverse styles of proper sensors that cause them to a promising mechanical meeting to assist grouped growing duties. This paper proficiently analyses PDA requests noted in research composing that utilization of mobile labored in sensors to present agrarian sport plans. The hidden 1500 articles identified through informational index request had been screened reliant upon limit policies and sometime later investigated inside and out in full substance, reaching 22 articles related to this evaluation. The packages are organized as established by way of their rustic capacities. Those articles analyzed portray 12 developing programs, 6 homestead the board programs, 3 facts structure programs, and 4 increment association programs. GPS and cameras are maximum prestigious sensors applied inside evaluated papers. This indicates a risk for destiny applications to apply distinct sensors like accelerometer to give improved plant blueprints.

Keywords Land utilization · Cultivation · Intelligent farming · Agricultural crops

---

Vijay Nath: Member IEEE

R. Priyadarshi (✉)
National Institute of Technology, Patna, Patna, Bihar, India
e-mail: rahul.glorious91@gmail.com

P. Bhardwaj · P. Gupta
Department of Chemical Engineering, B.I.T. Mesra, Ranchi, JH 835215, India

V. Nath
Department of ECE, B.I.T. Mesra, Ranchi, JH 835215, India
e-mail: vijaynath@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_56

1 Introduction

Inside the context of improvement of farmers, "land usage" marks a primary aid that desires right reforms for the development of financial system of Indian farmers. One thing that improves the cellular phones' capacity to support clients to accomplish specific assignments is various underlying sensors (e.g., situating sensors, movement sensors, and cameras mouthpieces). Several companies have embraced cell telephones to paintings with their work, for example, scientific services [1–3] and preparation [4, 5]. This paper surveys using cell phone sensors in a full-size vicinity, agribusiness. Reasonably priced mobile telephones supplied with one-of-a-kind sensors are beginning new freedoms for the USA ranchers who lately had restricted admittance to up-to-date rural statistics and assist from farming professionals also, management enlargement laborers. In the period in between, ranchers in large scope ranches, who as of now embody help from different records improvements, would now be capable of use mobile phone-based sensors to build performance and paintings with one-of-a-kind assignments in the course of the cultivating cycle. Making certain meals protection has been a worldwide difficulty at some stage in humankind has set of reports, and the worldwide meals emergency has underscored the significance of expanding each measure and nature of meals creation [6]. Agribusiness, which is the upstream food creating area, hence, needs new and present day techniques to guarantee food security. Conventional farming was typically done inside a personal home or a town and collective cultivating aptitude; also, information were passed down to their people in the future. Human was the primary eyewitness of arena situations and also arrangement supplier as issues emerge. Nevertheless, this conventional strategy is presently not reasonable because cultivating yields rely generally upon the regular encompassing circumstances (e.g., climate and water) and an unnatural weather change issues (causing regular dry spells and floods) and yield illness flare-ups are troublesome of cultivating efficiency. Conventional bunch of inside information and human perceptions alone are not adequate to fight questions. In this way, new strategies for cultivating have been imagined, for instance, exactness agribusiness. One key thought of novel performance in agribusiness is utilization of innovation to gauge or screen field and harvest circumstances all together for ranchers to settle on educated choices in different parts regarding their cultivating cycle. Sensors for consistent and programmed estimations of various field esteems are one of the primary achievement factors for exactness farming [7].

2 Smartphone Sensors and Applications

This element studies standard sensors of progressing phones. Sensors are a contraption that sports an actual sum and renovates to a signal that may be scrutinized and used by distinctive expedients [8–12]. A representation of standard sensors is thermometer, that changes over how blistering or bloodless ether, is into more than a few that is called temperature. There are various styles of sensors that are accessible currently, and numerous trendy PDAs are fortified with various sensors. The sensors added in telephones may be masterminded into three characterizations: improvement sensors, regular sensors, and position sensors [13]. The important elegance, development sensors, offers pace increment electricity and rotational force exams. The second magnificence, natural sensors, gives tests of the overall conditions. This compasses from encompassing air temperature from a thermometer and squeezing aspect from a measure to brilliance from a photometer. The second one fee magnificence, role sensors, offers exams of the tool's actual role. Such sensors consolidate magnetometers, GPS, except, heading sensors [14]. Desk 1 portrays the regular PDA sensors. With the ultimate objective of this overview, we live away from correspondence channels and show channels from the sensor list. Past due developments in acceptance of PDAs have called recognize researchers to sightsee the use of mobile sensors in their exertion. Various sensors had been practical adequately in specific districts. Normal positioning machine (GPS) licenses applications to get the cutting-edge area of maneuver. Beside direct groups of the distance in programs (e.g., maps), GPS records have additionally been used to instigate transference methods [15]. Accelerometers, that quantity pressure of pace increment whether or not performed through the phone's turn of occasions or gravity in three hatchets, were used as sensors in fall place [16], pastime affirmation [17], using event location [18], and so forth hidden cameras (both back and front cameras) of progressing cell phones have amended highly to the quantity conclusions over non-stop years. The digicam goals are right now similar with committed handheld cameras. Therefore, Wi-Fi packages would now have the choice to be a solitary tool, which takes as per records the photographs or bills from worked in cameras, then uses its estimation capacity to accomplish laptop apparition computations, and yields accommodating statistics from records pics. Instances of such claims fuse indoor arranging [19], beat extraction the usage of video imaging [20], and heartbeat charge appraisal [21].

3 Methods

3.1 Information Sources of Systematic Review

We have taken on Prisma [22–25], that cause framework for broadcasting actual evaluations, for our investigation. A composing quest preserving Prisma rule for traineeships from SCOPUS database become done on February 2, 2015. Our objective changed to examine for articles remembering sensors for cell phones and their usage in cultivating. Usage of the watchword "sensor(s)" inside the pursuit guidelines yielded articles associated with such sensors but not principally PDA-based sensors, and we agreed many effects with an awesome quantity of articles to brush aside. Along these lines, we settled on a choice not to incorporate the watchword "sensor(s)" as the consideration presented an excessive number of bogus positive articles to be later physically taken out. Furthermore, a utilization of different cell phone parts as sensors in exploration can be very unobtrusive. For instance, analysts regularly do not unequivocally bring up that the utilization of cameras to gather picture information is "detecting" pictures utilizing cameras on cell phones. Intermittently, they are clarified as using cell phones to play out certain horticultural assignments. Hence, we chose to incorporate "smartphone(s)" in our inquiry models all things being equal. Since cell phones in the writing might be alluded to by countless terms, for example, "portable phone(s)" or "versatile device(s)," we concocted a bunch of words that may depict cell phones. The arrangement of catchphrases are accompanying "portable application," "portable application," "cell phone," "cell phone," and "versatile telephone." With our objective of exploring use of smartphone-based sensors in agribusiness, past set of watchwords developed quest measures related to the watchword "horticulture." Specifically, the expression (("cell phone" or "portable gadget" or "cell phone" or "portable application" or "versatile application") or "agribusiness") was gone into SCOPUS Web Index. The inquiry terms zeroed in on the terms related to cell phones and agribusiness. The indexed lists were not restricted by verbal or any time requirements. Outcome set from this pursuit was then deliberately decreased to as it was article including sensors in cell phones and their utilization in farming by rejection measures [26–29].

3.2 Consideration and Exclusion Criteria

In this orderly audit, we are keen on courses that extant utilization just as turn of events and assessment of cell phone-based sensors in farming. Hence, we looked SCOPUS information base utilizing the catchphrases "horticulture" and numerous varieties of "smartphone(s)." Inquiry delivered a rundown of articles with their utilization of installed cell phone sensors, as planned. The hunt likewise brought about articles that highlight cell phones yet not their sensors. For example, cell phones are utilized uniquely as a correspondence implies in agribusiness work, when ranchers use voice call or SMS to convey crop costs, or when network-enabled cell phones are exclusively utilized as gadgets for getting to rural related data. We chose to bar those articles, as cell phones are not concocted as sensors yet just as specialized gadgets. Additionally, we prohibited articles that portray general frameworks that were not explicitly intended for rural use. In addition, extent of our survey is plant cultivating; we in this manner eliminated articles related just too animal cultivating. Besides, we audit use of cell phone-based sensors in horticulture to increment ranchers' familiarity with various claims that can be utilized in their cultivating cycle deprived of buying new devoted gadgets. We, in this way, chosen to bar articles in which cell phones are utilized uniquely as a distant incurable to screen and control other outer sensors, for instance, outer temperature and dampness sensors as in [30, 31].

Much of the time, it is yet not satisfactory from an occupied text survey if any underlying sensors are utilized in claims. Nevertheless, we chose to remember convinced traineeships for which albeit the creators did not obviously indicate the sensors utilized, we can surmise from the functionalities of the applications, for example, recovering area information by means of GPS as in [32, 33]. We likewise prohibited traineeships inscribed in dialects except English and those of which we were unable to get to full text.

4 Results

This segment portrays the cell phone applications from our overview. The applications were ordered into four classes: cultivating, ranch the board, data framework, and, expansion administrations.

4.1 Farming Applications

Cultivating alludes to a progression of agrarian cycles, which include different everyday exercises on the field, for instance, planting, weeding, preparing, and settling on related horticultural choices. Cultivating exercises are zeroed in on the most proficient method to develop plants, kill weeds/bothers, distinguish furthermore, right plant infections, apply composts, and gauge development/yield of harvests. Late advances in cell phone application improvement and an expanding accessibility of cell phones take into account a portion of these rural weights to be lifted and directed. For instance, ranchers might ascertain appropriate measures of manures for crop fields after investigating shade of crop leaves with some assistance from cell phone applications. This part surveys cell phone applications under cultivating classification and how they might help ranchers to accomplish cultivating errand.

4.2 Disease Detection and Diagnosis

Cell phone claims under this section are committed to infection discovery/determination in ranches while using sensors on cell phones. Some authors portrayed a versatile apparition framework that supports plant sickness ID measure [34]. The framework worked by catching pictures of plant leaves being explored for sicknesses, then, at that point, preprocessing those pictures, and sending the prepared pictures to far off research facilities. The picture preprocessing step was essential for saving communication price of referring unhealthy leaf pictures to shrub pathologists in far off research facilities. Leaf pictures were fragmented by a grouping calculation into three regions: foundation non-diseased piece of the leaf, and infected portion(s) of the leaf. Leaf pictures were then edited just to the area of the biggest ailing patch on the leaf and sent over any accessible organization to lab specialists for additional sickness ID.

4.3 Fertilizer Calculator

Applying manure is a significant cultivating action with a possibility to enormously influence ranch efficiency. Choices on which ranchers should make synthetic substances to apply and their yield explicit proper amounts. Some authors inspected a couple of portable device-based optical applications for agribusiness [35]; one of them was a cell phone-based shading assessor committed for rice leaves' chlorophyll assessment. Application and its fresher delivered adaptation assessed the shading level of rice leaves and suggested mandatory measures of nitrogen manure for smearing to the rice field [36].

4.4 Water Study

Water quality influences cultivating and horticulture in neighboring districts. A venture in Scotland created a cell phone claim committed to urge clients to submit data of water situations, that is, water level, water lucidity, deterrent in waterway, green growth cover, temperature, non-native plants in water, and going with photos of the river. Different clients in space could then give opinion on water circumstances information.

5 Conclusion

This paper presented the top-tier investigation of the exploration writing in how wireless sensors have been used in cultivating, without the prerequisites for external sensors. Outright of 22 articles portraying 25 mobile phone-based agrarian claims were seen to be inside our certification. Cameras and GPS were the most largely used sensors. Out of 25 applications, cameras and GPS were used in 16 and 14 applications, independently. Several usages each used various sensors (enhancers, accelerometer, and whirligig).

References

1. Mosa ASM, Yoo I, Sheets L (2012) A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 12(1). Article 67

2. Cheung WS, Hew KF (2009) A review of research methodologies used in studies on mobile handheld devices in K-12 and higher education settings. Australas J Educ Technol 25(2):153–183
3. Priyadarshi R, Gupta B, Anurag A (2020) Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues. J Supercomput 1–41
4. Milrad M, Spikol D (2007) Anytime, anywhere learning supported by smart phones: experiences and results from the MUSIS project. Educ Technol Soc 10(4):62–70
5. Mesas-Carrascosa FJ, Castillejo-Gonzalez IL, de la Orden MS, García-Ferrer A (2012) Real-time mobile phone application to support land policy. Comput Electron Agric 85:109–111
6. Kwon S, Kim H, Park KS (2012) Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society (EMBC'12), San Diego, CA, Aug–Sept 2012. IEEE, pp 2174–2177
7. Priyadarshi R, Singh MP, Bhardwaj A, Sharma P (2017) Amount of fading analysis for composite fading channel using Holtzman approximation. In: 2017 fourth international conference on image information processing (ICIIP), Dec 2017. IEEE, pp 1–5
8. Yang F, Li S (2008) Development of information support system for the application of new maize variety based on smartphone. In: IFIP international federation for information processing, vol 259, pp 817–824
9. Priyadarshi R, Rawat P, Nath V (2019) Energy dependent cluster formation in heterogeneous wireless sensor network. Microsyst Technol 25(6):2313–2321
10. de Silva L, Goonetillake J, Wikramanayake G, Ginige A (2013) Farmer response towards the initial agriculture information dissemination mobile prototype. In: Computational science and its applications—ICCSA 2013. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 7971. Springer, pp 264–278
11. Priyadarshi R, Soni SK, Nath V (2018) Energy efficient cluster head formation in wireless sensor network. Microsyst Technol 24(12):4775–4784
12. Schmidt R, Möhring M, Härting R-C, Reichstein C, Neumaier P, Jozinović P (2015) Industry 4.0—potentials for creating smart products: empirical research results. In: International conference on business information systems. Springer, Cham, pp 16–27
13. Priyadarshi R, Singh L, Singh A (2018) A novel HEED protocol for wireless sensor networks. In: 2018 5th international conference on signal processing and integrated networks (SPIN), Feb 2018. IEEE, pp 296–300
14. Priyadarshi R, Nath V (2019) A novel diamond–hexagon search algorithm for motion estimation. Microsyst Technol 25(12):4587–4591
15. Priyadarshi R, Soni SK, Sharma P (2019) An enhanced GEAR protocol for wireless sensor networks. In: Nanoelectronics, circuits and communication systems. Springer, Singapore, pp 289–297
16. Priyadarshi R, Rawat P, Nath V, Acharya B, Shylashree N (2020) Three level heterogeneous clustering protocol for wireless sensor network. Microsyst Technol 26:3855–3864
17. Priyadarshi R, Singh L, Singh A, Thakur A (2018) SEEN: stable energy efficient network for wireless sensor network. In: 2018 5th international conference on signal processing and integrated networks (SPIN), Feb 2018. IEEE, pp 338–342
18. Priyadarshi R, Gupta B (2021) Area coverage optimization in three-dimensional wireless sensor network. Wireless Pers Commun 117(2):843–865
19. Priyadarshi R, Gupta B, Anurag A (2020) Wireless sensor networks deployment: a result oriented analysis. Wireless Pers Commun 113(2):843–866
20. Intaravanne Y, Sumriddetchkajorn S, Nukeaw J (2012) Cell phone-based two-dimensional spectral analysis for banana ripeness estimation. Sens Actuators B Chem 168:390–394
21. Priyadarshi R, Kumar RR (2021) An energy-efficient LEACH routing protocol for wireless sensor networks. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 423–430

22. Anurag A, Priyadarshi R, Goel A, Gupta B (2020) 2-D coverage optimization in WSN using a novel variant of particle swarm optimisation. In: 2020 7th international conference on signal processing and integrated networks (SPIN), Feb 2020. IEEE, pp 663–668
23. Priyadarshi R, Gupta B (2020) Coverage area enhancement in wireless sensor network. Microsyst Technol 26(5):1417–1426
24. Werner M, Kessel M, Marouane C (2011) Indoor positioning using smartphone camera. In: Proceedings of the international conference on indoor positioning and indoor navigation (IPIN'11), Sept 2011, vol 6, p 1
25. Rawat P, Chauhan S, Priyadarshi R (2020) Energy-efficient clusterhead selection scheme in heterogeneous wireless sensor network. J Circuits Syst Comput 29(13):2050204
26. Rafoss T, Sælid K, Sletten A, Gyland LF, Engravslia L (2010) Open geospatial technology standards and their potential in plant pest risk management—GPS-enabled mobile phones utilising open geospatial technology standards web feature service transactions support the fighting of fire blight in Norway. Comput Electron Agric 74(2):336–340
27. Priyadarshi R, Yadav S, Bilyan D (2019) Performance analysis of adapted selection based protocol over LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 247–256
28. Priyadarshi R, Yadav S, Bilyan D (2019) Performance and comparison analysis of MIEEP routing protocol over adapted LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 237–245
29. Priyadarshi R, Soni SK, Bhadu R, Nath V (2018) Performance analysis of diamond search algorithm over full search algorithm. Microsyst Technol 24(6):2529–2537
30. Soni SK, Kumar S, Priyadarshi R (2018) Energy-aware clustering in wireless sensor networks. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 453–461
31. Kumar S, Soni SK, Priyadarshi R (2018) Performance analysis of novel energy aware routing in wireless sensor network. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 503–511
32. Policy CF (2010) Department of Industrial Policy and Promotion, Ministry of Commerce and Industry, Government of India
33. Rawat P, Chauhan S, Priyadarshi R (2021) A novel heterogeneous clustering protocol for lifetime maximization of wireless sensor network. Wireless Pers Commun 117(2):825–841
34. Sateesh VA, Kumar A, Priyadarshi R, Nath V (2021) A novel deployment scheme to enhance the coverage in wireless sensor network. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 985–993
35. Singh L, Kumar A, Priyadarshi R (2018) Performance and comparison analysis of image processing based forest fire detection. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 473–479
36. Gómez-Robledo L, López-Ruiz N, Melgosa M, Palma AJ, Capitán-Vallvey LF, Sánchez-Marañón M (2013) Using the mobile phone as Munsell soil-colour sensor: an experiment under controlled illumination conditions. Comput Electron Agric 99:200–208

Development of Smart Village for Better Lifestyle of Farmers by Crop and Health Monitoring System

Abhishek Pandey, Deepak Kumar, Rahul Priyadarshi, and Vijay Nath

Abstract This experiment manuscript demonstrates the smart way to monitor farm conditions using IoT. To analyse the conditions prevailing over the field we should have to gathered and club the data. With the help of this data, we can reach out to them and help them. The idea demonstrated here is simple, cheap and effective. Useful in running several tests on the soil before starting plantations. We are using quality sensors which will measure almost all the factors that are needed to be considered when we are testing the land. It will be helpful in measuring the salinity level of the sand, water content, temperature, humidity and air quality of the surrounding environment, which will be a great help. This task will be performed in Raspberry Pi. The components include a single-board computer with wireless LAN and Bluetooth connectivity. Seed sunlight sensor, soil moisture sensor, serve their purpose as their names are. Not only hardware but we also need an IDE platform, Arduino, to make all the components mentioned above work together.

Keywords Raspberry Pi4 · Arduino · Seed sunlight sensor · Soil moisture sensor · Arduino IDE

A. Pandey (✉)
Faculty of Engineering and Applied Science, Usha Martin University, Ranchi, Jharkhand 835103, India
e-mail: abhishek@umu.ac.in

D. Kumar · V. Nath
Department of Electronics and Communication Engineering, B.I.T. Mesra, Ranchi, Jharkhand 835215, India
e-mail: vijaynath@bitmesra.ac.in

R. Priyadarshi
Department of Electronics and Communication Engineering, National Institute of Technology, Patna, Patna, Bihar, India

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_57

1 Introduction

Arduino is a development board used for making projects and few end products. You can consider it as a microcontroller transformed for a modern world. Its customization is possible, but it already comes in many variants [1].

How it is useful:

1. Students can utilize microcontrollers to make projects to learn how they work.
2. Many electronics aficionados can build robots or any other useful product that can be used in the home or office.
3. It is an open-source electronics, which means you can download the design and use it to create your own kits and products. Many Arduino shields, collaborative circuits, and development boards are now available [2].
4. It may be used to teach programming to electronics and computer science engineers in order to create a practical design.
5. Any useful gadget can be made by completely understanding the hardware–software interaction.
6. To comprehend any Internet of things project's concepts.

It becomes a little less equipped after a certain level. This indicates that any industry or company will not use Arduino. However, the ideas may cross paths. As a result, you can study it and subsequently progress to circuit boards with better specifications and purposes.

We will take the help of Sigfox for between circuits and display. Sigfox provides low-powered wireless network connectivity, better than Wi-Fi, Bluetooth, or Zigbee. Sigfox is a novel wireless technology that fits within the LPWAN umbrella (low-power wide area network). Sigfox enables the expansion of IoT (Internet of things) and IR4.0 by providing very low-power bidirectional wireless communication for small devices (Industrial Revolution 4.0) [3].

The suitable action can be taken against this problem by first analysing the situation. The data, which needs to be collected, are

1. The intensity of light.
2. Soil moisture.
3. Other factors through sensors.

Israel is using high quality sensors, drones, high-resolution cameras, and GPS for monitoring farms. We lack such technology, but we can start with small setups. Transportation and marketing comes into picture when the crops are healthy.

Our work is to connect Raspberry Pi to Arduino and Arduino to the sensors for observing condition on the farm and then sending the data through Sigfox to farmers through email, SMS, app, or website and preparing an e-card for crop safety [4].

The paper is constructed in following manner: Raspberry Pi working information is provided in Sect. 2. Raspberry Pi 3 Model B working is explained in Sect. 3. Arduino UNO and its communication through are explained in Sect. 4. Last section is conclusion followed by references.

2 Raspberry Pi Working

The Raspberry Pi is a PI series development board. It can be thought of as a single-board computer that runs on the LINUX operating system. The board not only offers a lot of functionality, but it also has a fast processor, making it ideal for complex applications. The PI board was created with hobbyists and engineers in mind who are interested in LINUX systems and IoT (Internet of things) [5].

After Arduino, the Raspberry Pi platform is the most popular. PI is recommended when designing complex applications, despite the fact that it has fewer total uses. In addition, the Raspberry Pi is an open-source platform that provides a wealth of associated data, allowing you to tailor the system to your own needs [6].

Raspberry Pi 3 is preferred over other microcontrollers and development boards in the following scenarios:

1. When the system's processing power is enormous, because most Arduino boards' clock speeds are less than 100 MHz, they can only perform functions that are within their capabilities. They are unable to handle high-end programmes for applications such as weather stations, cloud servers, gaming consoles, and so on. Raspberry Pi can accomplish all of these complex functions because of its 1.2 GHz clock speed and 1 GB RAM.
2. Wherever there is a requirement for wireless communication, the Raspberry Pi 3 features wireless LAN and Bluetooth capabilities, allowing you to create a Wi-Fi Hotspot for Internet access. This functionality is ideal for the Internet of things.
3. The Raspberry Pi has a dedicated connector for connecting a touch LCD display, which eliminates the need for a monitor.
4. The Raspberry Pi also contains a dedicated camera connection, allowing users to connect a camera to the PI board without difficulty.
5. PWM outputs are also available on the Raspberry Pi for application use.

Many other features, such as HD streaming, help to promote the use of Raspberry Pi.

Here, we used Raspberry to connect the DF Robot Turbidity sensor. In addition, Raspberry Pi is further connected to Walabot creator. The Pi is attached to the Sigfox module for wireless connectivity to the Walabot SDK installed on your PC. Then, that SDK shows the image of the object taken by Walabot by radio frequency.

3 Arduino UNO Design and Working

The ATmega328P microprocessor is used in the Arduino UNO microcontroller board (datasheet). There are 14 digital input/output pins, 6 analogue inputs, a 16 MHz ceramic resonator (CSTCE16M0V53-R0), a USB connector, a power jack, an ICSP header, and a reset button on this board. It includes everything you'll need to get started with the microcontroller, including a USB cable to connect it to a computer and an AC-to-DC adapter or battery to power it. You can play about with your UNO without worrying about making a mistake; if something goes wrong, you can replace the chip for a few dollars and start over [7, 8].

The Getting Started with Arduino UNO page has all you need to know about configuring your board, using the Arduino Software (IDE), and getting started with code and electronics.

The Raspberry Pi is a small computer, and Arduino is a microcontroller board. As a result, Arduino is merely a component of the Raspberry Pi [9, 10]. The Raspberry Pi excels in software applications, but the Arduino simplifies hardware tasks. We can connect soil moisture sensor, seed sunlight sensor, and other sensors with Arduino to acquire data. This data is sent through Sigfox module connected with Arduino to the serial monitor and then to the other platforms [11].

- Of course some coding is needed to configure, but it is a one-time process.
- Open the serial monitor and check if your sensors are printing data [12] (Figs. 1 and 2).

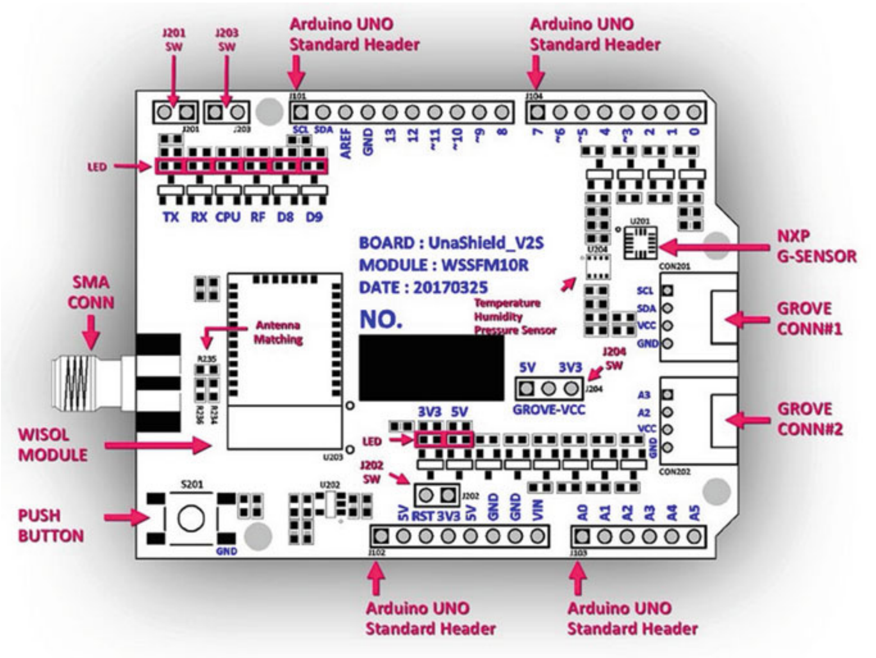


Fig. 1 Design of an Arduino UNO

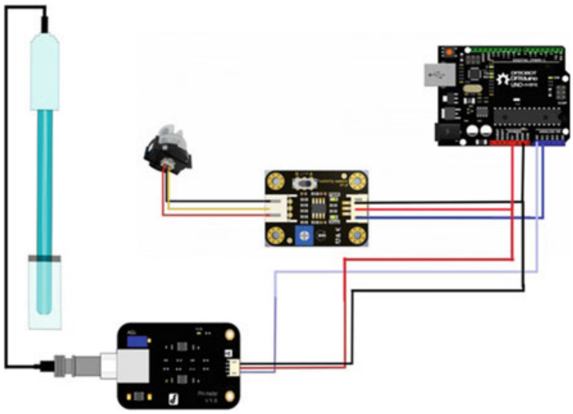


Fig. 2 Shows connection of some sensors with Arduino UNO

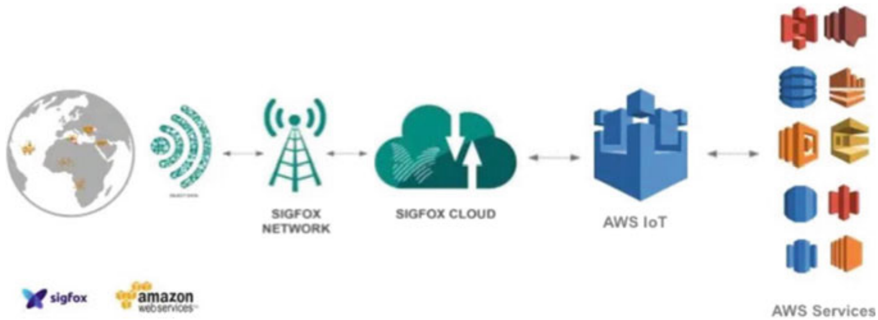


Fig. 3 Data transfer network using Sigfox and AWS IoT

4 Using Sigfox for Data Sharing

After collecting the above data, we can send it to farmers and scientists. Arduino UNO is a microcontroller board based on the ATmega328P (datasheet). We are sending the data of sensors through Sigfox. Sigfox is eco-friendly and leaves no carbon footprints. Using machine learning can help the scientists recognize the problems more specifically (Fig. 3).

5 Conclusion

The project is proposed to demonstrate how we can keep check on temperature, humidity, plantation technique, soil moisture, and other factors. If any of the above factors change, we will be notified by email or SMS. The project can be made more interesting by adding some more sensors and some ML and AI works. Switching this project to a large scale can sorely help us in increasing crop productivity. Arduino and Raspberry bring new gadgets that can be implemented in our life for learning IoT. Technology is advanced and will continue to escalate. Technology has made everything large to small and small to smaller. Therefore, it will be our loss if we do not know them, learn about them, and apply them in our life.

References

1. Shim K-S, Yoon S-H, Lee S-K, Kim M-S (2017) SigBox: automatic signature generation method for fine-grained traffic identification. J Inf Sci Eng 33(2):537–569
2. Hassan QF (ed) (2018) Internet of things A to Z: technologies and applications. Wiley
3. Barrett SF (2013) Arduino microcontroller processing for everyone! Synth Lect Digit Circuits Syst 8(4):1–513
4. Tsay J, Dabbish L, Herbsleb J (2014) Influence of social and technical factors for evaluating contribution in GitHub. In: Proceedings of the 36th international conference on software engineering, pp 356–366
5. Pearson RSGPH (2002) Microprocessor architecture programming and applications with the 8085. In: Atilim University library, 2002, vol 1. Bukupedia
6. Priyadarshi R, Gupta B, Anurag A (2020) Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues. J Supercomput 1–41
7. Khan MA, Salah K (2018) IoT security: review, blockchain solutions, and open challenges. Future Gener Comput Syst 82:395–411
8. Upton E, Halfacree G (2014) Raspberry Pi user guide. Wiley
9. Severance C (2013) Eben Upton: Raspberry Pi. Computer 46(10):14–16
10. Priyadarshi R, Gupta B, Anurag A (2020) Wireless sensor networks deployment: a result oriented analysis. Wireless Pers Commun 113(2):843–866
11. Mekki K, Bajic E, Chaxel F, Meyer F (2018) Overview of cellular LPWAN technologies for IoT deployment: Sigfox, LoRaWAN, and NB-IoT. In: 2018 IEEE international conference on pervasive computing and communications workshops (percom workshops), Mar 2018. IEEE, pp 197–202
12. Lavric A, Petrariu AI, Popa V (2019) Long range Sigfox communication protocol scalability analysis under large-scale, high-density conditions. IEEE Access 7:35816–35825

A Novel Approach for Sink Route in Wireless Sensor Network

Rahul Priyadarshi, Harsh Rana, Aman Srivastava, and Vijay Nath

Abstract In a wireless sensor network, sensor hubs have obliged power. Thus, the battery force of the sensor hubs should be monitored. It will blast the life expectancy of the remote sensor organization. A sensor hub identifies the abnormal event and sends the records to sink hub through multi jumping. Sensor hub inside the sink hub will eat up more prominent battery strength. As a result of those reasons, sensor hub fastly empties out battery power and declines the life expectancy of them. Repositioning of sink is an incredible procedure for increasing the WSN lifetime. Strength cognizant, a few sink repositioning (MSR) plot, is utilized appropriately here. SAMSR technique has significant parts, for example, energy mindful transmission range change and various sinks repositioning. Transmission range change component upgrades the life expectancy of the sensor hubs to sensor organization. The sink repositioning strategy comprises of two significant parts. The primary one is utilized to choose whether the sink hub met the repositioning circumstance or not. The second is utilized to choose the sink migration way.

Keywords Wireless · Sensor · Networks · Sinks · Battery · Repositioning

1 Introduction

Wireless sensor community (WSN) [1, 2] is a fixed allocation of self-sustaining sensors that is used to feel the environmental situations consisting of temperature, pressure, humidity, rainfall, and lots of others and to cooperatively transmit their facts through the network. The improvement of wireless sensor networks became recommended by way of army applications including battlefield surveillance. When WSN is utilized in sensing environment, sensor nodes can be liable of detecting an

R. Priyadarshi (✉)
Department of ECE, National Institute of Technology, Patna, Bihar, India
e-mail: rahul.glorious91@gmail.com

H. Rana · A. Srivastava · V. Nath
VLSI Design Group, Department of ECE, B.I.T. Mesra, Ranchi, JH 835215, India
e-mail: vijaynath@bitmesra.ac.in

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_58

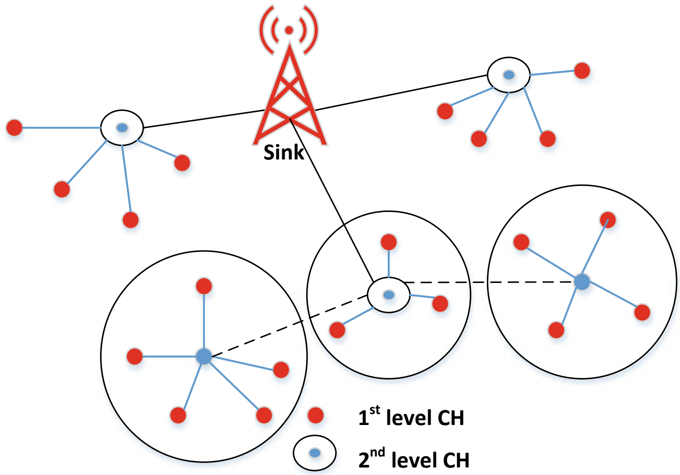


Fig. 1 Working scheme of a WSN

extraordinary scenario or for gathering the sensed statistics of the surroundings. In case of distinguishing an extraordinary occasion in a sensor node or being set to detect the statistics, it gives a report to a unique node which is referred to as a sink node. Then, sink node records the administrator via net [3–7].

As seen in Fig. 1, the sensor node detects a strange occasion and it sends a caution notification to the sink node to inform the manager through a pre-decided path. The routing path can be either static or dynamic, depending on given routing protocol.

Every sensor node in the wireless sensor network has several components along with radio transceiver with antenna, a microcontroller, a virtual circuit, and battery [5, 8–10]. Programs of WSN are several, together with climate monitoring battlefield surveillance healthcare tracking, forest fireplace detection, and so on. Strength is the restrained useful resource of WSN nodes, and it comes to a decision the lifetimes of WSNs. Sensors are unable to recharge or substitute, while batteries run out of power [11–14]. The battery tired out sensor nodes can also additionally motive several troubles together with communication holes and coverage troubles. Many WSN researches are involved in designing productive techniques in order to defend battery drain. WSNs may be categorized into two classes. They are static and dynamic WSNs. When a static WSN is employed in sensing surroundings, every sensor node points at a set function to carry out sensing and sensed data relaying functions until a sensor node drain out their battery power. For the class of dynamic WSNs, sensor nodes or the base stations are able to move. One of the energy-saving techniques is sensor relocation. In this technique, mobile sensors with high-residual energies alter their places from a place with a high-energy sensor node region to a low-strength sensor node location. Although this approach can growth community lifetime, the transpose of sensor nodes may also reduce strength of the battery. Sink relocation is another technique for extending lifetime of a network where the other sensor nodes stay at rest to preserve strength of the battery [15–20].

The node close to the sink will be draining out its battery strength, and thereby, the entire WSN will die. This node is known as a "hotspot". Within the case of the sink being capable of transferring earlier than the latest spot that node "*a*" drains out of its battery power, the sink can attain a few other sensor node "*b*" to reduce the situation of heavy energy consumption of node "*a*" [3, 21–23].

As in the occasion, the proper part, the sink relocates its feature from a close by node "*a*" to the node "*b*". In this way, the characteristic of the hotspot can be exchanged from one sensor node to different sensor node, which improves the network lifetime. A sink repositioning scheme gives facts to the sink when and in which to transport [24–26].

2 Related Work

On this phase, we describe a few lifetime enhancement strategies of existing structures. Energy-aware routing protocols [27] give fundamental idea of energy-saving routing protocols to increase the survivability of networks which can also use most useful paths. There are so many routing paths among source and destinations nodes. Messages are dispatched from the supply to destination through the paths that are randomly decided on. Energy-conscious routing protocol is likewise a reactive protocol. Energy-aware routing protocol uses best one path at all times. The protocol has three phases. They may be setup segment, statistics verbal exchange, and route upkeep [28].

Energy-efficient routing protocol is constructed by [29] to decrease the power usage on the nodes which maximizes the network lifetime. Transmission energy control technique is used to adjust the battery strength ranges at node. The main motive of this set of rules is used to make the strength-efficient layout protocol that uses the idea of transmission variety of the nodes. Higher transmission range makes use of the much less wide variety of nodes that had to attain the destination nodes, whereas low transmission range desires extra quantity of forwarding nodes which reduce energy usage. In path discovery phase, every node communicates with the neighboring nodes. As quickly because the direction is located, every individual node manages their transmission range, so that exceptional possible energy is utilized for packet transmission.

In a broadcast transmission, a supply node sends message to excursion spot node via intermediate node. Every node is ready with omnidirectional antenna. Each node has confined preliminary battery electricity. The battery energy is reduced after each transmission, counting on the transmission distance. Broadcast incremental strength (BIS) algorithm is used to build a spanning tree with conventional minimal electricity. Electricity-aware multiple paths set of regulations became described through, and it distributes the burden in addition among all of the nodes. In unmarried route set of guidelines, there may be one direction available from deliver to sink. Because of unmarried direction, there is constantly very heavy website traffic at the path, and moreover, its lifetime is brief. In a couple of paths, each node is established to the wide variety of paths. All nodes are on foot with two modes at the side of relay and experience mode. Sink node will broadcast a commercial to all nodes. Sink node chooses random nodes that are referred to as cluster node (cn). All cluster nodes advertise their popularity to its cluster humans. A sensor node wants to deliver records manner; it will deliver to its on hand cn to sink.

Distributed and localized set of regulations for sinks movements is changed. Here, sinks take relocation based at the community situations, surroundings conditions, and sensor deployment. Every sink serves as a cluster head. Each sensor node keeps statistics approximately the nearest sink and forwards messages to its nearer sink. A sink node makes a ramification to action its new region which relies at the power ranges. In order to find out the extreme region to transport, firstly a sink must examine all the locations in its very own cluster. The near locations are taken into consideration first for sink movement. Related backbone set of rules is used to check the latest location to fulfill the requirement of sink connectivity. This is known as a localized set of guidelines.

The mobility of the base station which is not always most effective improves the life of the networks and also lets in to balance the burden. Routing method is used to discover the shortest path for transmission of statistics. The base station is the most effective movements at the round path with radius. Nodes inside the circle use the shortest routing direction to ship the facts. Nodes which might be present within the outdoor of the round trajectory not use the shortest direction which helps to stability the load. Sink repositioning algorithms are described in [30–33]. The trouble of multi-hop relaying and a couple of sinks may be prevented with the resource of sink repositioning approach. Repositioning the sink at some stage in the ordinary network operation can be very hard. The simple problems are when the sink ought to flow into, in which it needs to be moved. Preferred strength transmission of the sensors for the previous and subsequent sink positions is evaluated. Battery strength transmission is checked, and if it is far more than a threshold cost means, the sink is probably moved to a next function, otherwise it remains at the previous region [34–37].

3 Proposed Method

SAMSR scheme is used to enhance the life of the wireless sensor network. SAMSR is based on the following steps.

The SAMSR approach involved inside the following steps:

1. Calculate the residual battery power (RBE) of the sensor nodes.
2. Regulate the transmission range primarily based at the battery electricity of nodes.

3. Find out the routing path among the sensors and sink nodes and the usage of maximum potential path (MCP). Above three steps are used to determine the circumstance of the sink movement.

The moving destinations of sink nodes are determined based on the complete neighbor set and weight value. The proposed SAMSR consists of three additives including strength consumption version, strength-efficient load balancing routing protocol, and energy-aware sink repositioning.

3.1 Energy Consumption Model for WSN

Let $E_{\mathrm{Tx}}(k, d)$ denote the full electricity needed for sensor node for transmitting a message of length "k" bits to the nodes behind the sensor node at a distance "d". $E_{\mathrm{Rx}}(k)$ denotes the entire power required for a sensor node to collect a "k" bit period message from the neighboring sensor node. The power consumed for message transmission $E_{\mathrm{Tx}}(k, d)$ can be divided into two components in which the first element is energy ate up within the transmitter or the receiver for k bit message that is equal to $E_{\mathrm{elec}} \times k$, where E_{elec} shows the electricity ate up for transmission or receiving circuitry consistent with bit. The other element which shows power consumption of an amplifier element is clearly identical to amp $\times\, d^2$, where amp identifies the strength sufficient for the transmitter amplification. The complete energy transmission and receiving are shown in below equations. Range adjustment might be depending upon the residual power of a sensor node. Right here, sensor nodes are classified into three kinds based upon their battery power, and they are able to alter their transmission variety. For this cause, let R be preliminary transmission variety, B be preliminary battery energy, and t be adjusted transmission range. Transmission variety may be adjusted in predicated upon the residual battery power. Transmission range can be decided in three ways. They are

Case 1: If the battery of the node belongs to $0 \leq r(u) < B/3$ means, $t = R/4$.
Case 2: If the battery of the node belongs to $B/3 \leq r(u) < B/2$, $t = R/2$.
Case 3: If the battery of the node belongs to $B/2 \leq r(u) \leq B$, $t = R$.

3.2 Electricity-Efficient Route-Balancing Routing Protocol

Routing protocols in a WSN can normally be labeled two into sorts. They may be (i) static routing and (ii) dynamic routing. For dynamic routing protocol, sensor nodes send information via extraordinary paths in each transmission in line with the present day residual battery strength. Due to this reason, the dynamic routing can be stable to the burden on every sensor node, so it will grow the community lifetime. Dynamic routing protocol is known as maximum capability route (MCP) routing. MCP routing protocol is utilized properly right here for sink repositioning technique. A current battery electricity of sensor nodes may be described by way of the use of capacity graph $g = (v, e)$. The MCP specially includes two procedural steps: (i) figuring out the maximum potential route for each sensor node and (ii) appearing routing and residual power replace. The MCP will repeat the above-mentioned steps.

3.3 Energy-Aware Multiple Sink Repositioning

EAMSR method incorporates the approaches of energy-aware transmission range adjustment and sinks repositioning. Depending on residual battery energy (RBE), each node adjusts its own transmission variety. If the RBE becomes low, transmission range might be changed to small for energy saving. In this proposed method, the transmission.

(1) *Energy-Aware Transmission Range*

A sensor node with best a bit residual electricity can exchange the transmitting range to be low which stores its electricity. This adjustable transmission variety mechanism enlarges the life of the sensor node and eventually the network.

(2) *Base Station Repositioning Mechanism*

This mechanism has two additives. The first issue is used to determine whether a repositioning circumstance is encountered or not. The second main thing is to identify which characteristic of the sink is to be relocated. For the repositioning, the sink will eventually gather the battery energy of every sensor node. After sensing the records series, the sink will make avail of one of the routing protocols to compute the most ability direction Pus for each maximum capacity direction Pus, and the most capability charge with appreciate to Pus is C(Pus). Allow the gathering of the sensor pals of a sensor "S" be "N". Then, the reporting state of affairs could be met at the same time as one of the described conditions occur: (1) even as one of the most functionality values C(Pus) with respect to the sensor neighbor u in N drops underneath $B/2$ or (2) the common battery energy of the neighboring set drops underneath $B/2$. That is, while both of the conditions occur, this means that the residual energies of the close by sink consisting of sensor nodes emerge as small or the electricity of a few routing path will fail under a given threshold.

The sink repositioning technique can be carried out to relocate the sink to a new role, which will increase the community lifetime.

4 Results and Performance Analysis

Energy-aware multiple sink repositioning is simulated by using MATLAB.

Figure 2 gives the performance results of energy-aware multiple sink repositioning

Fig. 2 Network lifetime

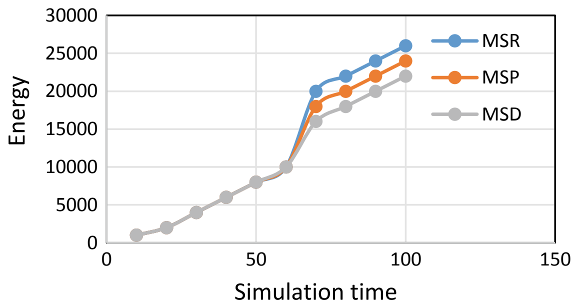
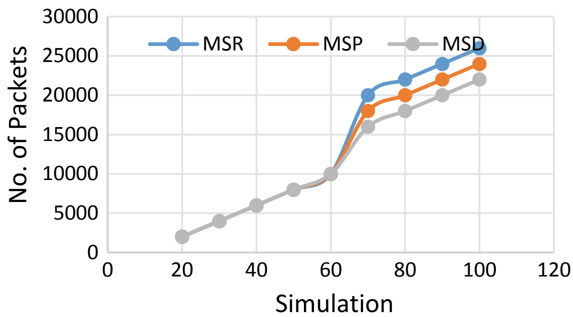


Fig. 3 Packet delivery



and mobile collector WSN lifetime enhancement methods. When compared to mobile collector method, MSR increases the lifetime of the WSN.

Figure 3 shows the performance results for packet delivery in energy-aware multiple sink repositioning and WSN lifetime enhancement methods. In MSP method, sensor nodes in WSN deliver more data packets than mobile collector method.

5 Conclusion

Energy is the limited source of wireless sensor network nodes (WSNs). One of the main problems in the WSN is to increase the lifespan of the sensor nodes. Strength-aware multiple sink repositioning is one way for enhancing the network lifetime. It avoids the stay of the sink nodes at fixed locations. The sink nodes are static means which causes the communication holes in WSN. Multiple sink repositioning is done with the aid of the MSR approach. MSR is primarily based on the concept of modifying sensor node's transmission ranges according to their residual energies, which enhances the lifetime of the WSN. When compared to mobile collector method, multiple sink repositioning with energy-aware routing protocol minimizes the energy consumption of sensor nodes in the WSN.

References

1. Jain N, Agrawal DP (2005) Current trends in wireless sensor network design. Int J Distrib Sens Netw 1(1):101–122
2. Shah RC, Rabaey J (2002) Energy aware routing for low energy ad hoc sensor networks. In: Proceedings of IEEE wireless communications and networking conference, Mar 2002, vol 1, pp 350–355
3. Priyadarshi R, Kumar RR (2021) An energy-efficient LEACH routing protocol for wireless sensor networks. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 423–430
4. Srikrishnan A, Pandidurai R (2017) RF energy harvesting with data collection approach incorporating in WSN. Int Innov Res J Eng Technol 62–66
5. Priyadarshi R, Gupta B, Anurag A (2020) Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues. J Supercomput 1–41
6. Shrivastava P, Pokle SB (2012) Survey on sink repositioning techniques in wireless sensor networks. Int J Comput Appl 51(4)
7. Luo J, Hubaux JP (2005) Joint mobility and routing for lifetime elongation in wireless sensor networks. In: Proceedings of IEEE INFOCOM, Mar 2005, pp 1735–1746
8. Jayalath C, Stephen J, Eugster P (2014) From the cloud to the atmosphere: running MapReduce across data centers. IEEE Trans Comput 63(1)
9. Priyadarshi R, Singh MP, Bhardwaj A, Sharma P (2017) Amount of fading analysis for composite fading channel using Holtzman approximation. In: 2017 fourth international conference on image information processing (ICIIP), Dec 2017. IEEE, pp 1–5
10. Abdu AI, Salamah M (2003) An adaptive energy aware transmission scheme for wireless sensor networks. Int J Wireless Commun Mob Comput 1(1):14–20
11. Priyadarshi R, Rawat P, Nath V (2019) Energy dependent cluster formation in heterogeneous wireless sensor network. Microsyst Technol 25(6):2313–2321
12. Priyadarshi R, Singh L, Singh A (2018) A novel HEED protocol for wireless sensor networks. In: 2018 5th international conference on signal processing and integrated networks (SPIN), Feb 2018. IEEE, pp 296–300
13. Priyadarshi R, Nath V (2019) A novel diamond–hexagon search algorithm for motion estimation. Microsyst Technol 25(12):4587–4591
14. Priyadarshi R, Soni SK, Sharma P (2019) An enhanced GEAR protocol for wireless sensor networks. In: Nanoelectronics, circuits and communication systems. Springer, Singapore, pp 289–297
15. Karkvandi HR, Pecht E, Yadid-Pecht O (2011) Effective lifetime aware routing in wireless sensor networks. IEEE Sens J 11(12):3359–3367
16. Priyadarshi R, Rawat P, Nath V, Acharya B, Shylashree N (2020) Three level heterogeneous clustering protocol for wireless sensor network. Microsyst Technol 26:3855–3864
17. Priyadarshi R, Singh L, Singh A, Thakur A (2018) SEEN: stable energy efficient network for wireless sensor network. In: 2018 5th international conference on signal processing and integrated networks (SPIN), Feb 2018. IEEE, pp 338–342
18. Panigrahi CR, Sarkar JL, Pati B, Das H (2015) S2S: a novel approach for source to sink node communication in wireless sensor networks. In: International conference on mining intelligence and knowledge exploration. Springer, Cham, pp 406–414
19. Priyadarshi R, Gupta B (2021) Area coverage optimization in three-dimensional wireless sensor network. Wireless Pers Commun 117(2):843–865
20. Tahiri A, Egea-López E, Vales-Alonso J, García-Haro J, Essaaidi M (2009) A novel approach for optimal wireless sensor network deployment. In: Symposium on progress in information & communication technology
21. Priyadarshi R, Gupta B, Anurag A (2020) Wireless sensor networks deployment: a result oriented analysis. Wireless Pers Commun 113(2):843–866
22. Marta M, Cardei M (2009) Improved sensor network lifetime with multiple mobile sinks. Pervasive Mob Comput 5(5):542–555

23. Anurag A, Priyadarshi R, Goel A, Gupta B (2020) 2-D coverage optimization in WSN using a novel variant of particle swarm optimisation. In: 2020 7th international conference on signal processing and integrated networks (SPIN), Feb 2020. IEEE, pp 663–668
24. Priyadarshi R, Gupta B (2020) Coverage area enhancement in wireless sensor network. Microsyst Technol 26(5):1417–1426
25. Rawat P, Chauhan S, Priyadarshi R (2020) Energy-efficient clusterhead selection scheme in heterogeneous wireless sensor network. J Circuits Syst Comput 29(13):2050204
26. Priyadarshi R, Yadav S, Bilyan D (2019) Performance analysis of adapted selection based protocol over LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 247–256
27. Kaur S, Grewal V (2020) A novel approach for particle swarm optimization-based clustering with dual sink mobility in wireless sensor network. Int J Commun Syst 33(16):e4553
28. Priyadarshi R, Yadav S, Bilyan D (2019) Performance and comparison analysis of MIEEP routing protocol over adapted LEACH protocol. In: Smart computational strategies: theoretical and practical aspects. Springer, Singapore, pp 237–245
29. Sarkar JL, Panigrahi CR, Pati B, Das H (2016) A novel approach for real-time data management in wireless sensor networks. In: Proceedings of 3rd international conference on advanced computing, networking and informatics. Springer, New Delhi, pp 599–607
30. Priyadarshi R, Soni SK, Bhadu R, Nath V (2018) Performance analysis of diamond search algorithm over full search algorithm. Microsyst Technol 24(6):2529–2537
31. Soni SK, Kumar S, Priyadarshi R (2018) Energy-aware clustering in wireless sensor networks. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 453–461
32. Kumar S, Soni SK, Priyadarshi R (2018) Performance analysis of novel energy aware routing in wireless sensor network. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 503–511
33. Wang C-F, Shih J-D, Pan B-H, Wu T-Y (2014) A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks. IEEE Sens J
34. Rawat P, Chauhan S, Priyadarshi R (2021) A novel heterogeneous clustering protocol for lifetime maximization of wireless sensor network. Wireless Pers Commun 117(2):825–841
35. Sateesh VA, Kumar A, Priyadarshi R, Nath V (2021) A novel deployment scheme to enhance the coverage in wireless sensor network. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems. Springer, Singapore, pp 985–993
36. Singh L, Kumar A, Priyadarshi R (2018) Performance and comparison analysis of image processing based forest fire detection. In: International conference on nanoelectronics, circuits and communication systems, Nov 2018. Springer, Singapore, pp 473–479
37. Jain B, Brar G, Malhotra J, Rani S (2017) A novel approach for smart cities in convergence to wireless sensor networks. Sustain Cities Soc 35:440–448

Automated Cheque Processing Through Data Verification and Siamese Networks

Anil Muthigi, Ashutosh Kumar, Gaurav Bhagchandani, Kiran Muthigi, and Vijay Nath

Abstract Bank cheques are still majorly used for financial transactions all over the world. A large number of cheques are processed manually on a daily basis, thereby requiring a lot of time, money and human effort. In such a manual verification, information like date, signature, amount present on the cheque has to be physically verified. This paper aims at finding a solution for processing cheques which increases the efficiency of this process, while minimizing human intervention. The service was hosted locally on a webpage, we first accepted the cheque image from the user, passed it to OpenCV which returned the various parts of the cheque which were then passed to Google Vision API to convert them into text, but MICR code was passed to tesseract OCR. After the successful extraction, the details were verified from the information in the SQLite database and signature was verified using the model trained in Jupyter Notebook.

Keywords OpenCV · MICR · Siamese networks · Tesseract OCR

1 Introduction

Due to security and trust issues, paper cheques are still estimated to play a big role in financial transactions worldwide. For clearance, the cheque is first converted into its digital form and then passed onto the cheque clearing unit for further processing of the cheque, which involves visual verification of all the details and digital transfer of cheque details between banks for amount confirmation and for validating the amount transfer.

A. Muthigi · A. Kumar · G. Bhagchandani · V. Nath
Department of Computer Science, B.I.T. Mesra, Ranchi, JH 835215, India
e-mail: vijaynath@bitmesra.ac.in

K. Muthigi (✉)
Department of Computer Science, Columbia University, New York, NY 10027, USA
e-mail: km3712@columbia.edu

V. Nath and J. K. Mandal (eds.), *Microelectronics, Communication Systems, Machine Learning and Internet of Things*, Lecture Notes in Electrical Engineering 887,
https://doi.org/10.1007/978-981-19-1906-0_59

2 Extracting Cheque Details

Different bank cheques have different size, shape and different relative positions of the fields in a cheque. A different searching region bounded by coordinates can be determined for every field in the cheque which can then be extracted from their regions after being grayscaled.

2.1 *Noise Reduction*

Different cheques have different background styles which thereby bring certain unwanted errors ("noise") in the extraction of text from the different fields in the cheque. To avoid extracting unnecessary details we can convert the extracted text into greyscale format and use other techniques like erosion–dilation to enhance the quality of extracted information.

2.2 *Extracting Fields from Their Regions*

As demonstrated in Fig. 1, the cheque image can be broken down into different bounding boxes based on their relative positions in the cheque to extract various fields from the cheque image. This process is called image slicing.

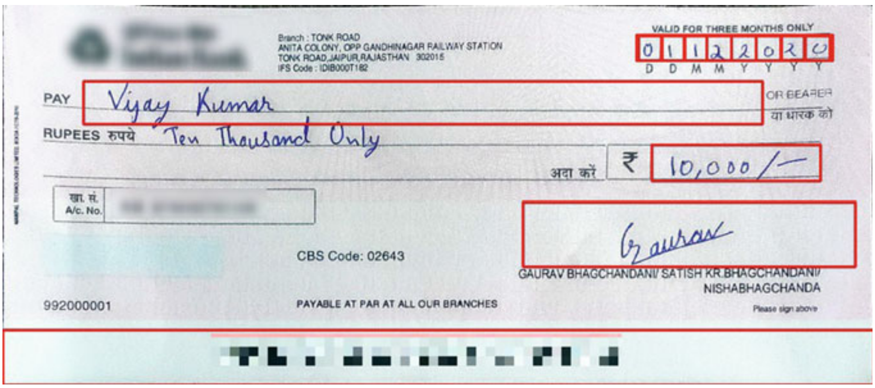


Fig. 1 Extracting cheque details from an Indian bank cheque

2.3 Extracting MICR Code

For Indian banks, there is a common pattern of the location of magnetic ink character recognition (MICR) code in the bank cheque. The MICR code is always found to be present in the lower 10% of the cheque image. The MICR code is written using a special font which can only be extracted from an optical character recognition model specifically trained to read that font. Example of one such optical character recognition model is the tesseract OCR. The extracted MICR code is divided into four parts. The first part of the code is the cheque number. The second part of the code is divided into three parts—the first part denotes the city name; the second part denotes the bank code to distinguish between different banks in India; and the third part is the branch code which distinguishes different branches of the same bank. The third part of the code is the Reserve Bank of India (RBI) code and the fourth part denotes the transaction number. Thus, we can see that extracting the MICR code will actually help us uniquely identify the payer's bank details.

3 Data Verification

After extracting the data, we firstly need to verify the extracted cheque details for cheque validity. The payee's name field on a cheque denotes the name of the person to whom the money is to be paid through that cheque. The name written in the payee-name field on the cheque is to be extracted and verified with that given in the database.

3.1 Verifying Cheque Details

In India, payee name is located above the amount field as shown in Fig. 1. For a cheque to be processed without any errors, the payee's name should match with the given account number, the account number on the cheque should match with that of the payer, the signature of the payer should match with that given in the database, and the payer should have enough amount in his/her balance in order to complete the transaction. In India, cheques cannot be processed after 3 months of the date written on the cheque, hence verifying the date of when the cheque was issued also has its importance.

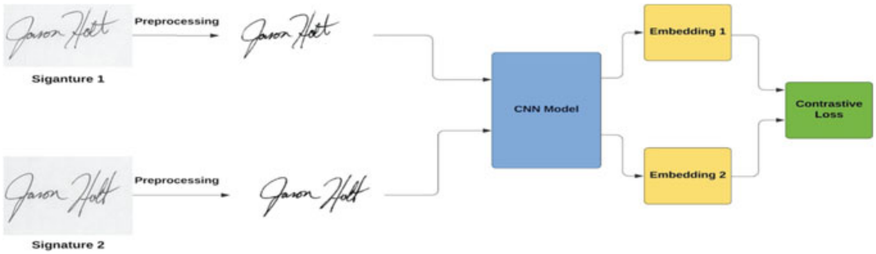


Fig. 2 Comparing the similarity of two signatures

3.2 Signature Verification

This is not only the most important step in the cheque processing industry but also the most sensitive out of all the cheque details. The thing that makes it so sensitive is that there is no exact measure to quantify if a signature is genuine or not. If the given signature seems to be similar to the one given in the database it is considered to be genuine. Hence, it is obvious that there will be certain human bias involved in what is considered to be genuine and what is not, thus making this a very non-uniform process.

3.2.1 Siamese Networks

As shown in Fig. 2, Siamese network (also called as twin network) is a special type of convolutional neural network (CNN) model in which two or more inputs are encoded into vector embeddings and their distance is computed. The contrastive loss function takes in the output of this network and treats them as vectors pointing in a multidimensional vector space.

The loss function computes its loss in a way so as to minimize the distance between similar (positive) samples and maximize the distance between dissimilar (negative) samples.

$$L\left(W, Y, \vec{X}_1, \vec{X}_2\right) = (1 - Y)\frac{1}{2}(D_{\mathrm{W}})^2 + (Y)\frac{1}{2}\{\max(0, m - D_{\mathrm{W}})\}^2$$

3.2.2 Contrastive Loss

The mathematical expression of the contrastive loss function is as follows:

Contrastive loss is a distance-based loss function (which can be either cosine distance or Euclidian distance). It tries to ensure that semantically similar examples are embedded close together. The original sample has a margin (m) around its vector space. This loss function is mathematically defined to push the negative samples outside of the neighbourhood by a margin while keeping positive samples within the neighbourhood.

4 Comparison with Current State-of-the-Art Models on CEDAR Dataset

State-of-the-art models	#Signatures	Accuracy	FAR	FRR
Word shape	55	78.50	19.50	22.45
Graph matching	55	92.10	8.20	7.70
Zernike moments	55	83.60	16.30	16.60
Surroundedness features	55	91.67	8.33	8.33
Signet model (convolutional Siamese networks)	55	86.5	13.70	13.15

5 Conclusion

When we are dealing with issues as sensitive as money, experimentations and risks are not very much appreciated. An automated cheque processing system can therefore be introduced into the process only if it offers reliability. With advanced technologies and further research, accuracy of the model can be further enhanced and the cheque clearance process can be truly automated and put to industrial use. There is no denying the fact that achieving human-level accuracy and understanding is yet to be achieved, but it is definitely a step in the right direction.

Bibliography

1. Bromley J, Bentz JW, Bottou L, Guyon I, LeCun Y, Moore C, Säckinger E, Shah R (1993) Signature verification using a Siamese time delay neural network. Int J Pattern Recogn Artif Intell 7(04):669–688
2. Khosla P, Teterwak P, Wang C et al (2020) Supervised contrastive learning
3. Gonzales RC, Woods RE (2002) Digital image processing, 2nd edn. Prentice Hall
4. Liu W, Wen Y, Yu Z, Yang M (2016) Large-margin softmax loss for convolutional neural networks
5. Tcheslavski GV (2010) Morphological image processing: grayscale morphology. ELEN 4304/5365 DIP, Spring
6. Lee C-Y, Gallagher PW, Tu Z (2016) Generalizing pooling functions in convolutional neural networks: mixed, gated, and tree. In: AISTATS
7. Krizhevsky A (2009) Learning multiple layers of features from tiny images. Technical report

8. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: CVPR
9. Zeiler MD, Fergus R (2013) Stochastic pooling for regularization of deep convolutional neural networks. arXiv preprint arXiv:1301.3557
10. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: Computer vision—ECCV 2014. Springer, pp 818–833
11. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
12. Ji S, Xu W, Yang M, Yu K (2013) 3D convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell 35(1):221–231
13. Tieleman T, Hinton G (2012) Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA Neural Netw Mach Learn 4(2):26–31
14. Ciresan DC, Meier U, Gambardella LM, Schmidhuber J (2011) Convolutional neural network committees for handwritten character classification. In: 2011 international conference on document analysis and recognition (ICDAR). IEEE, pp 1135–1139
15. Jayadevan R, Kolhe SR, Patil PM et al (2012) Automatic processing of handwritten bank cheque images: a survey. IJDAR 15:267–296. https://doi.org/10.1007/s10032-011-0170-8
16. Parcham E, Ilbeygi M, Amini M (2021) CBCapsNet: a novel writer-independent offline signature verification model using a CNN-based architecture and capsule neural networks. Expert Syst Appl
17. Thakur A, Prakash A, Mishra AK, Goldar A, Sonkar A (2020) Facial recognition with OpenCV. In: Smys S, Tavares J, Balas V, Iliyasu A (eds) Computational vision and bio-inspired computing. ICCVBIC 2019. Advances in intelligent systems and computing, vol 1108. Springer, Cham. https://doi.org/10.1007/978-3-030-37218-7_24
18. Pramkeaw P (2016) The study analysis knee angle of color set detection using image processing technique. In: 2016 12th international conference on signal-image technology & internet-based systems (SITIS), Naples, pp 657–660. https://doi.org/10.1109/sitis.2016.109
19. Dhanawade A, Drode A, Johnson G, Rao A, Upadhya S (2020) OpenCV based information extraction from cheques. In: 2020 fourth international conference on computing methodologies and communication (ICCMC), Erode, pp 93–97. https://doi.org/10.1109/iccmc48092.2020.iccmc-00018
20. Kumar V, Kaware P, Singh P, Sonkusare R, Kumar S (2020) Extraction of information from bill receipts using optical character recognition. In: 2020 international conference on smart electronics and communication (ICOSEC), pp 72–77. https://doi.org/10.1109/ICOSEC49089.2020.9215246
21. Palekar RR, Parab SU, Parikh DP, Kamble VN (2017) Real time license plate detection using OpenCV and tesseract. In: 2017 international conference on communication and signal processing (ICCSP), pp 2111–2115. https://doi.org/10.1109/ICCSP.2017.8286778
22. Sidhwa H, Kulshrestha S, Malhotra S, Virmani S (2018) Text extraction from bills and invoices. In: 2018 international conference on advances in computing, communication control and networking (ICACCCN), pp 564–568. https://doi.org/10.1109/ICACCCN.2018.8748309