Lecture Notes in Electrical Engineering 509

Sri Niwas Singh · Fushuan Wen
Monika Jain
Editors

Advances in System Optimization and Control

Select Proceedings of ICAEDC 2017

Springer

Lecture Notes in Electrical Engineering

Volume 509

** **Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Springerlink** **

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

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering
- Engineering

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

For general information about this series, comments or suggestions, please use the contact address under "service for this series".

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

Asia:

China, *Jessie Guo, Assistant Editor* (jessie.guo@springer.com) (Engineering)

India, *Swati Meherishi, Senior Editor* (swati.meherishi@springer.com) (Engineering)

Japan, *Takeyuki Yonezawa, Editorial Director* (takeyuki.yonezawa@springer.com)
(Physical Sciences & Engineering)

South Korea, *Smith (Ahram) Chae, Associate Editor* (smith.chae@springer.com)
(Physical Sciences & Engineering)

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

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

Europe:

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

(Applied Sciences and Engineering; Bio-Inspired Robotics, Medical Robotics, Bioengineering; Computational Methods & Models in Science, Medicine and Technology; Soft Computing; Philosophy of Modern Science and Technologies; Mechanical Engineering; Ocean and Naval Engineering; Water Management & Technology)
(christoph.baumann@springer.com)
(Heat and Mass Transfer, Signal Processing and Telecommunications, and Solid and Fluid Mechanics, and Engineering Materials)

North America:

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

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

Sri Niwas Singh · Fushuan Wen
Monika Jain
Editors

Advances in System Optimization and Control

Select Proceedings of ICAEDC 2017

Springer

Editors
Sri Niwas Singh
Madan Mohan Malaviya
University of Technology
Gorakhpur, Uttar Pradesh
India

Fushuan Wen
Department of Electrical
Engineering
Zhejiang University
Hangzhou, Zhejiang
China

Monika Jain
Department of Electrical
and Electronics Engineering
I.T.S. Engineering College
Greater Noida, Uttar Pradesh
India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-0664-8 ISBN 978-981-13-0665-5 (eBook)
https://doi.org/10.1007/978-981-13-0665-5

Library of Congress Control Number: 2018942156

Printed on acid-free paper

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

Preface

This volume contains twenty-four papers from International Conference on Advancement in Energy, Drive & Control (ICAEDC-2017), which was devoted to the gamut of emerging engineering issues from theoretical aspects to application-dependent studies and the validation for specific application. This was envisioned and founded to represent the growing needs of research in the areas of electrical, computer and electronics engineering and technology. This proceeding is intended as a forum for practitioners and researchers to share their ideas in the fields of sustainable energy, intelligent control, signal processing and communication.

The format consists of plenary sessions, tutorials and workshops on niche area. Around 300 people have participated in ICAEDC-2017, 186 papers have been received, and approximately selected 90 papers have been presented by authors.

The conference was inaugurated by honourable Chief Guest Mr. Dinesh Jain, President-Legal & Corporate Affairs, Uflex Ltd, India. Followed by first plenary session from distinguished Speaker Prof. Abul Hassan Siddiqi, Ex. Pro Vice-Chancellor, AMU & Visiting Consultant, ICTP entitled "Mathematical models & Methods of oil Industry". In his session, some well-known mathematical models in the form of partial differential equation representing real-world system and their wavelet-based simulation were presented. The plenary session on "Building a Cyber Resilient Energy Sector" was delivered by Dr. Anuj Goel, SM-IEEE, Co-Founder and CEO, Cyware Labs, USA. Further, Keynote Speaker Prof. Prerna Gaur, SM-IEEE, NSIT, New Delhi, has delivered a session on the "Application of doubly fed induction generator & permanent magnet synchronous generator for variable speed wind turbines". Next session was on the "Applications of Mathematical Modeling in Energy Drives," which is considered as a standard tool in energy planning and management, delivered by Prof. V. N. Jha, Prince Sattam Bin Abdulaziz University, Kingdom of Saudi Arabia. The last keynote session was by Prof. Seema Arora, Waljat College of Applied Sciences, Muscat, Oman, on "Broadband over Power Lines & Smart Grids," which defines the technology how a high degree of integration of existing electrical infrastructure with the communication infrastructure allows power and bidirectional information flow efficiently.

The conference was also featured with four tutorial sessions. The first tutorial was delivered by Mr. Antonio VILEI, Software Platforms and Cloud, STMicroelectronics, Lecce, Italy, on "The Internet of Things revolution: Smart Cities and Smart Industry". In parallel to this, second tutorial was on the "Impacts of Artificial Intelligence & Machine Learning on Workforce and Society, a Macro Economic Perspective" by Mr. Anshuman Tripathi, Head Consulting, Americas, USA. Mr. Anil Kumar Goel, Ex. DGM-NTPC, has delivered the third tutorial on "Power Scenario in India and Present Challenges". Prof. M.A Ansari, SM-IEEE, Gautam Buddha University Ltd, has delivered the fourth tutorial session on "Renewable Energy Scenario and Future Scope in India".

The workshop was another feature of this conference. It was scheduled in successive sessions on the " Trends in Microgrid Management" by Mr. Himadri Endow, Ex. Head Smart Grid, Alstom Grid, India, South East Asia, followed by another workshop on "Advancement and skill Challenges in Instrumentation and Automation" by Mrs. R. Mahalakshmi, DGM-Yokogawa India Limited. Mr. Piyush Chandra Ojha, VP CABCON, India, has delivered the third workshop on "Microgrids—The Changing Paradigm for Grid Control". The fourth workshop on "IOT in building and Home Automation" was delivered by Mr. Nitin Jain, Director and Chief Operating Officer, Aviconn Solutions Pvt. Ltd. Applications of IoT in buildings and home automation such as climate control HVAC (heating, ventilation, AC) automation and control safety and alarms detection of smoke, fire, water leak, gas leak, air quality, earthquake, device fault detection, energy savings lighting and appliance control home lighting, building/peripheral lighting and appliance control have been highlighted.

The ICAEDC-2017 has received the partial financial support from Science and Engineering Research Board, SERB, Department of Science & Technology—Govt. of India. Conference has also received its technical support from IEEE U.P section.

Different chapters of these volumes are outlined into the following major categories:

- Economical Assessment of Electric Utilities
- Sustainable Energy and Performance Analysis
- Power Quality Enhancement
- Power Line Communication
- Intelligent Control Systems:
- Reliability Measures and Optimization
- Antenna Design and Processing
- Wireless, Private Security and Routing Techniques
- Biomedical and Biometric Applications
- Artificial Intelligence and Optimization
- Signal and Image Processing

Economical Assessment of Electric Utilities

Since the operating characteristics of a modern power system are becoming increasingly complicated, it is imperative to assess the daily power systems' dispatching quality. Hence, Pengcheng Cui et al. have presented a post-evaluation index system for daily power system dispatching, based on the average effect, buckets effect and abnormalities effect, from the perspectives of system security, economics, energy saving, environmental friendliness and impartiality. Stated case studies show that the proposed method can not only identify the weak links, but also reflect the overall level accurately and is feasible and effective for applications in practical power systems. Kalyan Chaterjee et al. have presented a technique for calculating comparative cost efficiencies of Indian State Owned Electric Utilities (SOEUs), which have been primarily accountable for the distribution of electricity in India. The performance of twenty-eight state utilities was evaluated using the nonparametric approach of data envelopment analysis (DEA).

Sustainable Energy and Performance Analysis

As renewable energy has got a forefront focus for power generation due to the gradual depletion of fossil fuel and increased concerns about global warming, challenges associated with wind power plant have also become a major concern for researchers. As wind speed, being stochastic in nature, is making energy extracted from it non-dispatchable, Singh et al. have proposed a method to regulate it by coupling the wind energy system (WES) with flywheel, an energy storage device which requires less maintenance, quick acting, mechanically robust, high energy efficient, and least environment pollutant. It quickly reacts to variation in electrical output of wind energy system, thereby delivering smoother power to grid. Active power regulation due to 500 kW, 100 kgm^2 flywheel energy system integrated with 1.5 MW type 3 wind turbine generator is considered for variable wind condition in MATLAB/Simulink. Currently, solar energy is identified as a sustainable renewable energy source which is catching the eyes of big power industries and entrepreneurs who are constantly doing research for enhancing the efficiency and cost of energy production. Towards those aspects, authors have proposed a performance analysis of PV cell using one-diode and two-diode models. Another author has presented a review work of various maximum power point tracking (MPPT) techniques which experience partial shading conditions (PSC), an unavoidable complication that significantly reduces the efficiency of the overall system. The exhaustive comparison of various MPPTs taken by the researchers which tracks the global peak (GP) of a photo-voltaic (PV) array under PSC is compared, and the best available option is proposed for researchers. MPPT techniques such as perturb and observe

(P&O), improved particle swarm optimization (IPSO) and grey wolf optimization (GWO) are the recent techniques proposed by the researchers. Various DC–DC converter topologies with PV system are also compared for the purpose of maximum power point tracking.

Power Quality Enhancement

The penetration of renewable energy is increasing worldwide. The advancement in renewable energy is exciting, but at the same time it is creating significant technical challenges to power industry. The term "Power Quality" is a broad concept. It is associated with electrical transmission, distribution and utilization systems. Due to power quality problems, industries have to invest a large amount for mitigation of voltage sags, distortions, harmonics and short-term interruptions/disturbances, etc. Hence, authors have proposed to cover various possible sources and compensation methods of reactive power in power system under different contingencies using FACTS devices and advanced power converters.

Power Line Communication

Distribution load management is an important concern for the utilities to exploit the existing infrastructure of the power system network. There is a need to improve the reliability of power line networks and provide standards and systems for utilizing power networks as a medium of communication. There are areas where the digital communication may be made through low-voltage distribution network. This helps the utilities to reach remote locations. Mapping of the consumer through control centre is achieved through interfacing of coupling circuits that leads in distribution load management. Arora and Jain present designing of broadband coupling circuit that satisfies specific signal transmission, appropriate bandwidth and limited number of components. The paper also discusses the significant parameter variations during capacitive couplings and inductive couplings for passive and active network topologies.

Intelligent Control Systems

To avoid a catastrophic failure in any process plant, precise control of all associated drives is critical for real-time application. Singh et al. have proposed implementation of ANFIS controller-based algorithm in precise control techniques

particularly suitable to the textile, aerospace and the automobile sectors. In this chapter, a method has been proposed by implementing artificial intelligence (AI)-based software algorithm which tracks the position of the motor from signals received from the resolver and thereby calculates speed from the tracked position. Further, they have presented a comparison between AI-based controller and PI controller showing how it eliminates the delay and noise effects introduced by hardware circuit and can be readily implemented in real time with minimum complexity and cost. In subsequent chapter, for the complex higher-order system analysis, several order reduction techniques have been proposed. In another paper, a mixed method is proposed which combines the improved pade approximations and the eigen spectrum analysis for reducing the higher-order system. In this method, system stiffness and pole centroid of both original- and reduced-order system remain same. The denominator of higher-order system (HOS) is derived using the eigen spectrum analysis, and the numerator is derived by using improved pade approximation. The later method of order reduction utilizes both time moments and Markov parameters. The stability and quality of the reduced-order system are compared with the existing methods of order reduction. To understand the proposed method, paper includes some numerical examples of single-input single-output (SISO) systems.

Reliability Measures and Optimization

As complexity increases the system unavailability, reliability analysis plays an important role in its assessment to ensure the performance and availability of any specific systems at the point of requirement. Most preferred techniques of reliability optimization include the incorporation of redundant element either at the component or at the system level, and generally, the redundant components are of identical types. But the occurrence of common cause failure (CCF) defeats the benefit of redundancy optimization. To validate the same, authors have presented five exact approaches to evaluate the reliability of complex network and also extended the work to evaluate and compare the reliability of the same network under CCF. To incorporate the CCFs in reliability evaluation, conditional probability approach is used.

Antenna Design and Processing

Articles here exemplify the analysis and exploration of communication systems. They provide invaluable insights for the improvement of bandwidth and efficiency of types of antennas used for communication systems, the wide band antenna for

millimetre-wave application, where single antenna is used rather than more antennas in a device, to reduce the size and cost of device. An investigation into the design of Giuseppe Peano fractal patch antenna and reactive impedance surface (RIS) is discussed for the improvement of bandwidth and improvement in antenna radiation. A review of metamaterial structures is also presented for the enhancement of various antenna properties. Objectives of these articles are to get the best configuration required for the desired specification and also to help developing the future ideas by taking into account the advantages of the available structures.

Wireless, Private Security and Routing Techniques

In this section, a chapter on spectrum sensing technique for cognitive radio network and enhancement of trunking efficiency of mobile switching centre (MSC) is presented. Further, privacy issues of user identification by the use of encryption for pseudonymization using MAC and HMAC have also been illustrated.

Biomedical and Biometric Applications

Authors have focused on the aspects related to EMD, which is applied to EEG recordings from epileptic patients. The classification was done by using these selected features by artificial neural network (ANN). Next, articles illustrate signal processing in the field of biometrics. In next chapter, low-power EEG-based BCI or brain–computer interface for the physically challenged section of the society has been presented.

Artificial Intelligence and Optimization

In the very first article, author has demonstrated the importance of integrated novel image processing techniques—which shows the potential of such intelligent approaches. Hence, to illustrate the application, an intelligent approach based on generalized neural network (GNN) is proposed and applied for the short-term solar PV power forecasting. In the next chapter, Volterra filtering, for feed-forward active noise controls (ANCs), suitable for the nonlinear controller has been illustrated. As removing noise from images is an integral part of the image processing field, noises can occur in images during acquisition on transmission. The presence of noise can hinder the proper utilization of these images for various applications such as medical imaging, satellite imaging.

Signal Processing and Advanced Digital Design

The papers selected to be included in this category contribute to the understanding of relevant trends of current research on signal and image processing technologies. With context to that, a focus on recursive algorithm for the computation of discrete cosine transform (DCT) and inverse discrete cosine transform (IDCT) has been proposed; algorithm is both time and hardware efficient. In the next article, wavelet shrinkage function (WSF) used for fault detection and diagnosis in a transmission line has been presented. In the subsequent chapter, field programmable gate array (FPGA)-based architecture for implementation of DST has been proposed. To establish the communication between the blocks operating in different clock domains is a challenging task due to synchronization failure. Therefore, in the next article, mixed-clock first-in-first-out (FIFO) as an interface between different clock domains is briefed.

We immensely thank all the participants and authors for their contributions to ICAEDC-2017. The review committee has done an excellent job in reviewing the articles and approving the high-quality research articles to be published in this conference proceedings. The editors are thankful to various committees for their dedication in making this a very successful conference. We are grateful to Springer for making possible the publication of these proceedings of ICAEDC 2017. We would like to express our heartfelt thanks to College Administration for their constant encouragement and facilitating with all logistics support. A special thanks goes to the core committee members Ms. Kalpana Hazarika and Mr. Nitin Kathuria, assistant professors, EEE Department, I.T.S Engineering College, Greater Noida. Also, a deep sense of appreciation is extended to entire ICAEDC organizing committee for this successful endeavour.

Sincere gratitude is extended to Honorary Advisors of ICAEDC-2017: Senior advisors Prof. Abul Hassan Siddiqi, Ex. Pro Vice-Chancellor, AMU & Visiting Consultant, ICTP, and Prof. Krishna B. Misra, Founder and Past Editor-in-Chief, International Journal of Performability Engineering, Editor, Book Series on Performability Engineering . We also pleased to thank Dr. Manoj Kumar, Senior Group Manager, Leading and Managing System R&D and Applications Group at STMicroelectronics, India.

We sincerely hope that this volume will inspire researchers and scholars for innovation.

Gorakhpur, India — Sri Niwas Singh
Hangzhou, China — Fushuan Wen
Greater Noida, India — Monika Jain
Editors of ICAEDC-2017

Contents

About the Editors

Sri Niwas Singh is currently the vice-chancellor of Madan Mohan Malviya University of Technology, Gorakhpur, India, and a professor at the Department of Electrical Engineering, Indian Institute of Technology Kanpur, India. He obtained his M.Tech. and Ph.D. in electrical engineering from the Indian Institute of Technology Kanpur, in 1989 and 1995, respectively. Prof. Singh was an assistant engineer at UP State Electricity Board, from 1988 to 1996, and was an assistant professor at the Roorkee University (now IIT Roorkee) from 1996 to 2000 and at the Asian Institute of Technology, Bangkok, Thailand, from 2001 to 2002. Prof. Singh has received numerous awards, including Young Engineer Award from the Indian National Academy of Engineering (INAE), Khosla Research Award of IIT Roorkee and Young Engineer Award of CBIP New Delhi, India. He is also a recipient of the German Humboldt Fellowship and the Danish Otto Monsted Fellowship. In 2013, Prof. Singh became the first Asian to receive the IEEE Educational Activity Board Meritorious Achievement Award in Continuing Education. He was also the recipient of the INAE Outstanding Teacher Award in 2016 and IEEE R10 region (Asia-Pacific) Outstanding Volunteer Award in 2016. His research interests include power system restructuring, FACTS, power system optimization and control, wind power and security analysis. Prof. Singh has published more than 440 papers in international/national journals/conferences and has written two books.

Fushuan Wen is a professor at the School of Electrical Engineering at Zhejiang University, China. He completed his B.E. and M.E. at Tianjin University, Tianjin, China, in 1985 and 1988, respectively, and went on to do his Ph.D. at Zhejiang University, Hangzhou, China, in 1991, all in electrical engineering. He joined the faculty of Zhejiang University in 1991 and has been a full professor and the director of the Institute of Power Economics and Information since 1997 and the director of Zhejiang University-Insigma Joint Research Center for Smart Grids since 2010. He has undertaken various teaching, research and visiting appointments at the National University of Singapore (NSTB Postdoctoral Fellowship, Research Fellowship), Hong Kong Polytechnic University (Research Fellowship, Visiting Assistant

Professorship), University of Hong Kong (Research Assistant Professorship), South China University of Technology (University Distinguished Professorship), University of New South Wales (ARC Senior Fellowship), Queensland University of Technology (CSIRO Visiting Fellowship), University of Technology Brunei (Professorship in Power Systems), Technical University of Denmark (Otto Monsted Guest Professorship in Power Systems), Nanyang Technological University (Visiting Fellowship). His research interests include power economics and electricity markets; power system investment; planning and operation optimization; smart grids and electric vehicles; power system alarm processing; fault diagnosis; and system restoration. Prof. Wen is an editor of IEEE Transactions on Power Systems and IEEE Power Engineering Letters and associate editor of IET Generation, Transmission and Distribution, the Journal of Energy Engineering (ASCE), and the Journal of Modern Power Systems and Clean Energy (Springer). He is also on the editorial boards of more than ten journals.

Monika Jain is a professor and head of the Department of Electrical & Electronics Engineering at ITS Engineering College, Greater Noida. Prof. Jain is a senior administrator and passionate educationist with more than 20 years of experience in various prestigious institutions. She completed her M.Sc. in Instrumentation at IIT Roorkee and subsequently an M.Tech. from DAVV Indore and Ph.D. from BITS, Pilani. She has been the recipient of many awards and travel grants and has convened various high-profile conferences and seminars sponsored by AICTE, DST, IEEE and other funding agencies/sponsors. She has received several financial grants from the Government of India for hosting these events. Her major contributions include the establishment of a "Project Lab on Virtual Instrumentation" funded by AICTE under the "MODROBS" scheme. Prof. Jain has published more than 40 research papers in respected international journals and conferences, and her Ph.D. work was published in the IEEE Transactions on Broadcasting. Her research interests include MPEG systems, set-top box domains, embedded systems, IoTs, PLC, SCADA and automation. She serves as a reviewer for various leading conferences and journals. She is a member of IEEE, life member of IETE and member of the Indian Society of Industrial and Applied Mathematics (ISIAM).

Part I
Computer Architecture

FTSM: Design and Reliability Measures

Shilpa Gupta and G. L. Pahuja

Abstract Multistage interconnection networks (MINs) provide an effective way to communicate between processors and memory modules in supercomputer systems. Although shuffle-exchange network (SEN) is viewed as a potential network to manage as MIN for its regular structure and most modest size of switching element (SE), the fault tolerance is an issue in SEN because it is a unique path network. Many networks have been introduced in the past to increase the fault tolerance in the SEN such as SEN−, SEN+, SEN+2, SHSEN. It has been observed that all these networks provide maximum of one fault tolerance at input/output stage. In this paper, a new network has been proposed as FTSM (fault-tolerant SEN-Minus), which addresses the fault tolerance issue with increased number of totally disjoint paths, thus improving the reliability of the entire network. To assure that reliability has been improved and increased, a thorough comparison of 2-terminal reliability has been done with all other SENs.

Keywords Multistage interconnection network (MIN) · Fault-tolerant SEN-Minus (FTSM) · Shuffle-exchange network (SEN) · Switching element (SE) Reliability · Fault tolerance

Abbreviations

IN	Interconnection network
MIN	Multistage interconnection network
SEN	Shuffle-exchange network
SEN+	Shuffle-exchange network with one extra stage
SEN+2	Shuffle-exchange network with two extra stages
SEN-Minus	Shuffle-exchange network with one less stage
SHSEN	Symmetric homogeneous shuffle-exchange network

S. Gupta (✉) · G. L. Pahuja
Department of Electrical Engineering, NIT Kurukshetra, Kurukshetra, Haryana, India
e-mail: shilpa1_goyal@rediffmail.com

G. L. Pahuja
e-mail: pahuja.gl@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_1

FTSM	Fault-tolerant SEN-Minus
SE	Switching element
S-D	Source–destination
MUX	Multiplexer
DMUX	Demultiplexer
N/W	Network
RBD	Reliability block diagram
TR	Terminal reliability
$PM2^i$	Plus-Minus 2^i
N	Number of input/output in the network
R_{SE}	Reliability of switching element in the network
i	Number of stage

1 Introduction

In this computing era, MIN is becoming most favourable choice, as they provide efficient communication between processors and memory modules in parallel as well as distributed systems [1–15]. MIN has been acknowledged worldwide as most economical transmission media which provides programmable data paths within the command unit of multiprocessing systems. Performance and reliability of these fast multiprocessing systems totally rely on the performance of MIN employed in that system [1–4]. The reliability of a system/network may be defined as the capability of a system/network to carry out its desired operation successfully in any given condition with any SE failure in the network [4, 5]. So MIN employed in these systems should offer multiple paths to obtain desired operation in case of SE failures. Lot of research has been done in recent times to increase the number of paths of MIN so that reliability and fault tolerance capability of the network can be improved. Many new topologies have been suggested to provide multiple paths between each source–destination (S-D) node pair. In these topologies, either the size or the number of SE has been increased [5–10, 12, 13, 15]. In some other topologies, MUX/DEMUX has been used at input/output stage to increase fault tolerance and reliability [11–14]. MIN with the regular topology called SEN has been modified to offer multiple paths by increasing, either the number or the size of SE, or by using MUX/DEMUX at input/output in its basic topology. As SEN is a unique path network, it has a topological equivalence with many other networks of this class like Omega network, cube network, baseline network [8, 15]. All modifications done on SEN can likewise be applied to these networks too. SEN is a square MIN with $N \times N$ and configuration, where N is the number of inputs as well as outputs. It uses 2×2 SE in its basic topology which is recognized as a smallest SE in MIN. To increase redundancy in unique path SEN, one or two extra stages (SEN+ and SEN+2 respectively) have been suggested to add so as to increase the

number of paths between each S-D node pair [6, 7]. In symmetric homogeneous SEN (SHSEN) for a same network size, the hardware has been doubled to improve redundancy, which increases the hardware cost to a great deal [9]. In SEN-Minus network, an attempt has been made to reduce hardware cost by using MUX and DEMUX at input and output stages, respectively [12, 13]. In this network, two totally disjoint paths are offered between each S-D node pair, and it can tolerate one fault at each stage, including input and output stages. But if more than one fault occurs at any stage between same S-D node pair, then the entire network will get failed. To palliate this problem, a new topology FTSM (fault-tolerant SEN-Minus) has been introduced here in this paper which provides four totally disjoint paths and in total eight paths between each S-D node pair. This newly proposed network provides tolerance against minimum of three faults at each stage including input and output stages, which automatically increase the reliability. In this paper, reliability of newly proposed network has been evaluated and compared with existing networks. The results show that FTSM has highest reliability among all other existing variants of SEN network. Highlights of work performed in this paper are presented here as follows:

1.1 Highlights

SEN is most used interconnection network (IN), for its lowest cost, uncomplicated design topology and smallest SE. Hence, this network is viewed as a probable candidate to increase the reliability and efficiency of supercomputer systems.

SEN-Minus provides high reliability with two totally disjoint paths between each S-D node pair at lowest cost. It can tolerate one fault at each stage, but if more than one fault occurs at any stage then network will fail.

With a considerable increment in the cost of SEN-Minus, the reliability and fault tolerance can be improved and this became the motivation to pursue the research in the same. In this newly proposed FTSM network, bigger size of MUX and DEMUX have been used which provide four paths between each S-D node pair so as to make this network highly reliable and fault-tolerant (three faults can be tolerated at each stage including input as well as output stage). FTSM mitigates the problem of being lesser fault-tolerant arising in SEN-Minus.

Organization of this paper is as follows: In Sect. 2, literature review of SEN has been presented. In Sect. 3, topology of newly proposed FTSM has been explained. In Sect. 4, terminal reliability and fault tolerance evaluation of FTSM have been computed and compared with existing topologies, whereas in Sect. 5 conclusion and future scope have been given followed by the references.

2 Literature Review

SEN is a unique path MIN having regular size of $N \times N$, where 'N' is the number of inputs as well as outputs [3–15]. It comprises of '$N/2$' number of SE of size 2×2 at each stage, with $\log_2 N$ number of stages and switching complexity of $N/2$ ($\log_2 N$). The topological block diagram of SEN MIN is shown in Fig. 1a. It can be seen from the block diagram that there is one path between each S-D node pair [6–8, 10, 12, 13, 15] in SEN. In SEN+ there is one additional stage than SEN MIN and possesses two paths between each S-D node pair. It can tolerate one fault at intermediate stages but no fault can be tolerated at input as well as output stage. Topological block diagram of SEN+ MIN is shown in Fig. 1b. In SEN+2, there are two additional stages than SEN MIN and four paths between each S-D node pair. This network can tolerate multiple faults at intermediate stages [6–8, 10, 12, 13, 15] but again no fault can be tolerated at input as well as output stage. The topological block diagram of SEN+2 MIN is shown in Fig. 1c. In another SEN variant, i.e. SHSEN (symmetric homogeneous SEN) MIN, four paths are available between each S-D node pair. It comprises N inputs and N outputs with $((\log_2 N) + 1)$ stages having 'N' SE at each stage [9]. The topological block diagram of SHSEN MIN is shown in Fig. 1d.

There is one more network, i.e. SEN-Minus, which belongs to the category of SEN MIN. In this network, there is one less stage than SEN with $(\log_2 N - 1)$ number of stages and '$N/2$' SE at each stage of configuration 2×2. It uses 'N' MUX and 'N' DEMUX at input and output stages, respectively, [8, 12, 13] as they give lower cost and high reliability as compared to SE used in the network. SEN-Minus network provides two totally disjoint paths between each S-D node pair. The topological block diagram of SEN-Minus is shown in Fig. 1e.

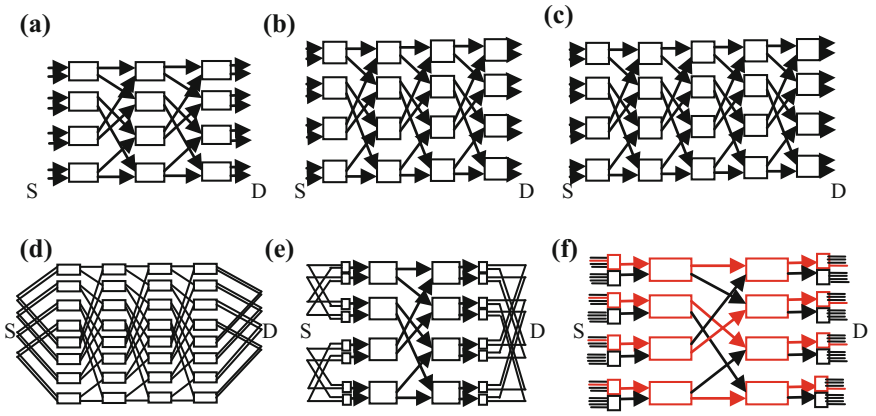

Fig. 1 Topologies for size 8×8 **a** SEN network, **b** SEN+ network, **c** SEN+2 network, **d** SHSEN network, **e** SEN-Minus network and **f** new FTSM network

To mitigate all the problem arising in all topologies discussed above, a new topology, i.e. FTSM (fault-tolerant SEN-Minus), has been proposed here in this paper. FTSM network uses 4 × 1 MUX and 1 × 4 DEMUX at input and output stages, respectively, to increase the number of totally disjoint paths between each S-D node pair to make network more reliable and fault-tolerant. The topologies' block diagram of FTSM is depicted in Fig. 1f.

3 Proposed FTSM Network

FTSM network proposed in this paper contains N input and N output with ($\log_2 N-1$) stages and $N/2$ SE at each stage. It comprises 'N' 4 × 1 MUX and 'N' 1 × 4 DEMUX at input and output stages, respectively, whereas in SEN-Minus network, 2 × 1 MUX and 1 × 2 DEMUX were used [12, 13]. Due to this configuration, FTSM contains four totally disjoint paths and in total eight paths for each S-D node pair. Due to the presence of multiple disjoint paths, FTSM becomes highly fault-tolerant. There are two MUX/DEMUX attached to each SE at input/output stage, out of which one is attached to all even inputs/outputs (such as '0', '2', '4' and '6') and another is attached to all odd inputs/outputs (such as '1', '3', '5' and '7'). There exist four paths between any S-D node pair (shown in red in Fig. 1f). Terminal reliability evaluation of this network has been computed in the preceding section and compared with existing networks to demonstrate supremacy of this network.

4 Terminal Reliability and Fault Tolerance

Reliability of a system can be defined as the ability of a system to perform as per the requirements in normal as well as critical circumstances [4, 6, 7] and [12, 13], whereas terminal reliability of a system can be defined as the probability of existing at least one fault-free path between a S-D node pair [7, 9, 11, 15]. It is an important parameter for performance evaluation of MIN. For analysis of terminal reliability, the Reliability Block Diagram (RBD) of FTSM network is shown in Fig. 2 and RBD of all other networks can be found in ref. [8–10]. Reliabilities of *2 × 2* SE, *4 × 1* MUX and *1 × 4* DEMUX are taken as *'r'* (*in each case, total cross-point used are 4*).

Terminal reliability (R_t) for new FTSM network can be expressed as:

$$R_t(\text{FTSM}) = 1 - \left[1 - \left\{1 - \left(1 - r^2\right)^2\right\}\left\{1 - \left(1 - r^2 r^{\log_2 N-3}\right)^2\right\}\right]^2 \qquad (1)$$

Terminal reliability of SEN, SEN+, SEN+2, SHSEN and SEN-Minus can be expressed by the Eqs. (2)–(6).

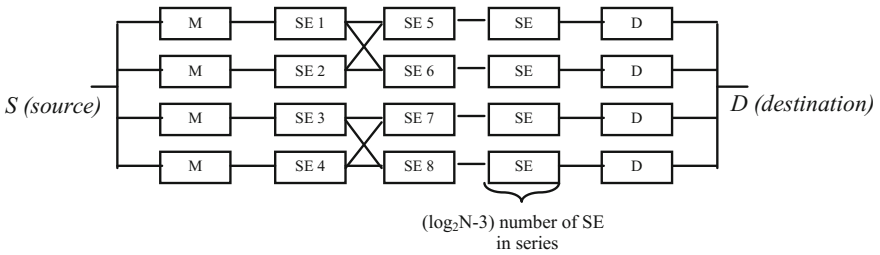

Fig. 2 RBD of FTSM (fault-tolerant SEN-Minus) network for TR analysis

$$R_t(\text{SEN}) = r^{\log_2 N} \tag{2}$$

$$R_t(\text{SEN}+) = r^2\left[1-\left(1-r^{(\log_2 N)-1}\right)^2\right] \tag{3}$$

$$R_t(\text{SEN}+2) = r^2\left[1-\left\{1-r^2\left(1-\left(1-r^{(\log_2 N)-2}\right)^2\right)\right\}^2\right] \tag{4}$$

$$R_t(\text{SHSEN}) = \left[1-(1-r)^2\right]\left[1-\left(1-r^{\log N}\right)^2\right] \tag{5}$$

$$R_t(\text{SEN}-) = \left[1-\left(1-r^{\log_2 N}\right)^2\right] \tag{6}$$

Values computed for terminal reliability evaluation of all above networks for network size $N = 8$ are given in Table 1 and for different values of N varying from $N = 3$–10 are given in Table 2 for SE reliability $R_{SE} = 0.95$. As shown in Tables 1 and 2, proposed FTSM shows the highest reliability among all other SEN variants for all values of N.

Table 1 Comparison of TR of all networks with FTSM for 8×8 N/W size

R(SE)	R_t (FTSM)	R_t (SEN−)	R_t (SEN)	R_t (SHSEN)	R_t (SEN+)	R_t (SEN+2)
0.99	0.9999993729	0.9991	0.9703	0.9990	0.9797	0.9797
0.97	0.9999513714	0.9924	0.9127	0.9915	0.9376	0.9375
0.96	0.9998498066	0.9867	0.8847	0.9851	0.9159	0.9157
0.95	0.9996419529	0.9797	0.8574	0.9772	0.8939	0.8935
0.92	0.997825709	0.9510	0.7787	0.9449	0.8264	0.8250
0.90	0.9949736451	0.9266	0.7290	0.9173	0.7808	0.7782

Table 2 Comparison of TR of all networks with FTSM for $R_{SE} = 0.95$

$\log_2 N$	R_t (FTSM)	R_t (SEN−)	R_t (SEN)	R_t (SHSEN)	R_t (SEN+)	R_t (SEN+2)
3	0.9996419529	0.9797	0.8574	0.9772	0.8939	0.8935
4	0.9991205948	0.9656	0.8145	0.9632	0.8841	0.8923
6	0.9963765822	0.9298	0.7351	0.9275	0.8563	0.8876
8	0.9900715586	0.8867	0.6634	0.8845	0.8204	0.8792
10	0.9789975857	0.8390	0.5987	0.8369	0.7791	0.8665

4.1 Fault Tolerant

SEN communicates data between input devices and output devices through a number of stages [1–3]. Fault tolerance of SEN is to provide service even in the presence of faulty components. A fault can be either permanent or transient; here in this case, faults are assumed to be permanent [4–6].

Fault tolerance of a given network is defined with respect to a chosen fault-tolerant model of that network, which characterizes all faults assumed to occur in that network. A network is single fault-tolerant if it can function as per the desired functionality despite of single fault. More generally, if a network can tolerate any set of i faults, then that network is i-fault tolerant. Fault tolerance at each stage of all networks under study is shown in Table 3 for network size $N = 8$. As can be seen in Table 3, new FTSM network shows highest fault tolerance among all

Table 3 Fault tolerance of FTSM and SEN networks

Network	Number of fault tolerated					
	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
FTSM	3 faults	If SE1 is operational SE 7,8,5/6 If SE2 is operational SE 7,8,5/6 If SE3 is operational SE 5,6,7/8 If SE4 is operational SE 5,6,7/8	–	–	–	–
SEN-Minus	1 fault	One SE at stage 1 can fail	–	–	–	–
SEN	No fault	No fault	No fault	–	–	–
SEN+	No fault	One SE at stage 1 can fail	One SE	No fault	–	–
SEN+2	No fault	One SE at stage 1 can fail	Any 2 SE	Any 1 SE	No fault	–
SHSEN	1 fault	One path with all series element can fail	One SE in series fail	One SE in series fail	One SE in series fail	1 fault

other SEN networks and can be preferred over all SEN variants in application where SE faults are more likely to happen.

5 Conclusion and Future Scope

SEN is most commonly used for fast computing in multiprocessors for their low cost and uncomplicated design. But at the same time, it needs attention as it is a unique path MIN; hence, the reliability of these SEN MIN has to be addressed. In this paper, new topology of SEN MIN, i.e. FTSM, has been proposed, which shows highest reliability among all studied networks. It has an advantage of providing multiple and totally disjoint paths between each S-D node pair. Hence, fault tolerance of the proposed network is highest, as it can tolerate maximum of three path failures for a given S-D node pair and can perform well. It also possesses smallest path between each S-D node pair and hence, reduces the latency produced by the larger path lengths. With all these advantages, FTSM proves itself as most capable candidate for hassle-free broadband communication in parallel and distributed systems. For future guidelines, same work can be extended to other class of networks such as Indra network, $PM2^i$ network (Plus-Minus 2^i), Gamma network, and reliability of these networks can be evaluated and compared with their counterparts.

References

1. V. Cherkassy, M. Malek, in *Reliability and Fault Diagnosis Analysis of Fault-Tolerant Multistage Interconnection Networks*. Proceeding 14th International Conference Fault Tolerant Computing (1984), pp. 246–251
2. J.H. Patel, Processor-memory interconnections for multiprocessors. IEEE Trans. Comput. **C-30**, 301–310 (1981)
3. G. Adams, D. Agrawal, H. Siegel, A survey and comparison of fault-tolerant multistage interconnection networks. Computer, 14–27 (1987)
4. J.T. Blake, K.S. Trivedi, Multistage interconnection network reliability. IEEE Trans. Comput. **38**(11), 1600–1604 (1989)
5. N.S. Fard, I. Gunawan, Terminal reliability improvement of shuffle-exchange network systems. Int. J. Reliab. Qual. Saf. Eng. **12**(1), 51–60 (2005)
6. I. Gunawan, Reliability analysis of shuffle-exchange network systems. Reliab. Eng. Syst. Saf. **93**(2), 271–276 (2008)
7. F. Bistouni, M. Jahanshahi, Analyzing the reliability of shuffle exchange network using reliability block diagrams. J. Reliab. Eng. Syst. Saf. **132**, 97–106 (2017)
8. N.A.M. Yunus, M. Othman, in *Reliability Performance of Shuffle Exchange Omega Network*. IEEE International Symposium on Telecommunication Technologies (ISTT) (2012)
9. Mohsen Jahanshahi, Fathollah Bistouni, A new approach to improve reliability of the multistage interconnection networks. Comput. Electr. Eng. **40**, 348–374 (2014)
10. N.A.M. Yunus, M. Othman, Shuffle exchange network in multistage interconnection network; A review and challanges. Int. J. Comput. Electr. Eng. **3**(5), 724 (2011)

11. S. Gupta, G.L. Pahuja, Terminal reliability assessment for a new gamma minus multistage interconnection networks. Procedia Comput. Sci. **70**, 476–482 (2015)
12. S. Gupta, G.L Pahuja, *Terminal Reliability Assessment and Comparison of New SEN-MIN: A Critical Component in Power System* (ICEPE-2015 IEEE) (2015)
13. S. Gupta, G.L Pahuja, *A New SEN Minus: Design and Reliability Measures*, vol. 23, No. 4 (World Scientific Publishing Company, 2016), pp. 16500121–16500129
14. S. Gupta, G.L Pahuja, in *Evaluation and Comparison of Performability of Gamma Network and its Variants*. IEEE Conference on ICCIC-2016 (2016)
15. N.A.M. Yunus, M. Othman, Reliability evaluation for shuffle exchange interconnection network. Procedia Comput. Sci. **59**, 162–170 (2015)

FPGA-Based Architecture for Implementation of Discrete Sine Transform

Anamika Jain, Neeta Pandey and Priyanka Jain

Abstract This paper presents architecture for discrete sine transform (DST) algorithm using VHDL description and its implementation on field programmable gate array (FPGA). Several algorithms have been proposed to implement the DST in its recursive structure form and focus on minimizing the number of additions and multiplications. The hardware designs of the algorithms are largely ignored. The implementation of DST algorithm on the FPGA is motivated by the fact that large memory FPGAs are now available providing a platform for processing real-time algorithms on application-specific hardware with substantially higher performance.

Keywords Recursive structure · Data compression · DST · FPGA

1 Introduction

In electronics and computer fields, DSP and digital design are two important and closely related areas. Digital design devoted to the issues in Boolean algebra, logic optimization, and digital implementation. DSP is related to the theory and applications of Fourier transform (FT). This family includes the discrete sine transform

A. Jain (✉)
Department of Electronics & Communication Engineering,
Bhagwan Parshuram Institute of Technology, PSP-4 Sector-17,
Rohini, Affiliated to Guru Gobimd Singh Indraprasta University, Delhi, India
e-mail: annamikajain@yahoo.co.in

N. Pandey · P. Jain
Electronics & Communication Engineering Department,
Delhi Technological University, Shahbad Daulatpur, Bawana Road,
Delhi 42, India
e-mail: neetapandey@dce.ac.in

P. Jain
e-mail: priyajain2000@rediffmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_2

(DST), discrete cosine transform (DCT), and the discrete Fourier transform (DFT). These transforms find applications in a variety of areas such as speech processing, data transmission on telephone channels, image processing, instrumentation, biomedical engineering, seismology, oil exploration, detection of nuclear explosion, and in the processing of signals received from the outer space [1, 2]. The DFT is a complex transform (due to the multiplication by $e^{-j2\pi nk/N}$) and, therefore, stipulates that both the image magnitude and the phase information be encoded. DFT also exhibits good decorrelation and energy compaction characteristics. The implicit periodicity of DFT gives rise to boundary discontinuities that result in significant high-frequency content. After quantization, the Gibbs phenomenon causes the boundary point to take on erroneous values [3]. DCT and DST use sinusoidal functions involving only real arithmetic operations. Energy compaction and optimum performance of DCT close to Karhunen–Loeve transform (KLT) which makes DCT most widely used transform in speech and image processing for data compression [4]. Jain [5] has shown that for a first-order Markov sequence with certain boundary conditions, the Karhunen–Loeve transform, reduces to the DST especially for signals with low correlation coefficient ($\rho < 0.6$). DST is widely used in image reconstruction [6], image coding [7], and interpolation [8].

Various algorithms have been introduced for the computation of the DST [9–13] and to minimize the number of addition and multiplications. There are many research papers that propose hardware implementation of DCT [14–16] but hardware design of the DST algorithm, however, is largely ignored. The implementation can be done either by using field programmable arrays (FPGAs) or application-specific integrated circuits (ASICs) [17]. The advantages of the FPGA approach for DST implementation include higher sampling rates than are available from traditional DSP chips, rapid prototyping and lower costs than its ASIC counterpart for moderate volume applications. This paper presents FPGA-based implementation of DST algorithms proposed in Ref. [18]. These algorithms employ Chebyshev polynomial of the first kind as it requires lesser computation overhead as compared to the use of Chebyshev polynomial of second and third kind. The choice of selected algorithm [18] is motivated by the fact that DST is performed on segments of a continuously incoming data stream and no complex multiplication is needed for processing, require less memory and fewer recursive cycles are needed for computation which results in smaller round-off error in the computations. The corresponding recursive structures are regular, modular, and free of global interconnections and easy to implement on FPGA.

The paper is organized as follows. Section 2 briefly describes a recursive DST algorithm [18] and its recursive architecture followed by architecture modeling and synthesis using VHDL in Sect. 3. In Sect. 4, synthesis results from our VHDL models have been reported. Finally, Sect. 5 summarizes the conclusion from this paper.

2 DST Algorithm

The DST (type II) of a sequence $x(n)$ with length N^* is defined by

$$X_s(k) = a_k \sum_{n=1}^{N} x(n) \sin \frac{(2n-1)k\pi}{2N}, \quad k = 1, 2, \ldots, N, \tag{1a}$$

where

$$a_k = \sqrt{\frac{2}{N}} \varepsilon_k \tag{1b}$$

$$\varepsilon_k = \begin{cases} \frac{1}{\sqrt{2}}, & k = N \\ 1, & k = 1, 2, 3, \ldots, N-1 \end{cases} \tag{1c}$$

$$\theta_k = \frac{k\pi}{N} \tag{1d}$$

Recursive structures for the computation of DST-II for k even/odd are shown in Figs. 1 and 2, respectively, and the corresponding values of $G[k]$ and $H[k]$ are reproduced below for completeness.

$$X[k] = a_k (-1)^{\left(\frac{k}{2}-1\right)} G_i[k], \tag{2a}$$

where

$$G_i[k] = \sum_{n=0}^{i-1} w_k[i-n] \sin\left(n + \frac{1}{2}\right)\theta_k, \quad k \text{ even} \tag{2b}$$

$$G_i[k] = \sin\frac{\theta_k}{2}\{w_k[i] + w_k[i-1]\} + 2\cos\theta_k G_{i-1}[k] - G_{i-2}[k], \quad k \text{ even} \tag{2c}$$

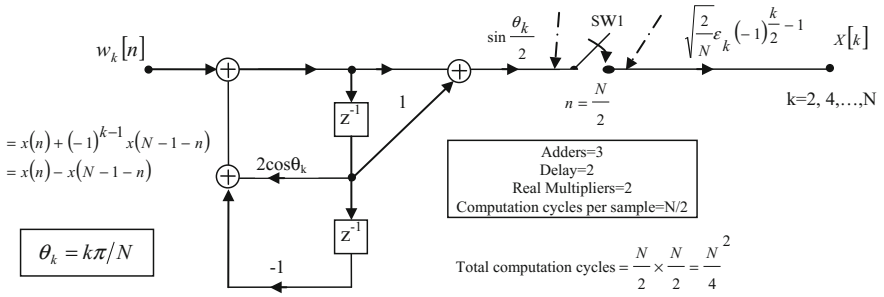

Fig. 1 Computation of DST kernels $X[k]$ for even values of k

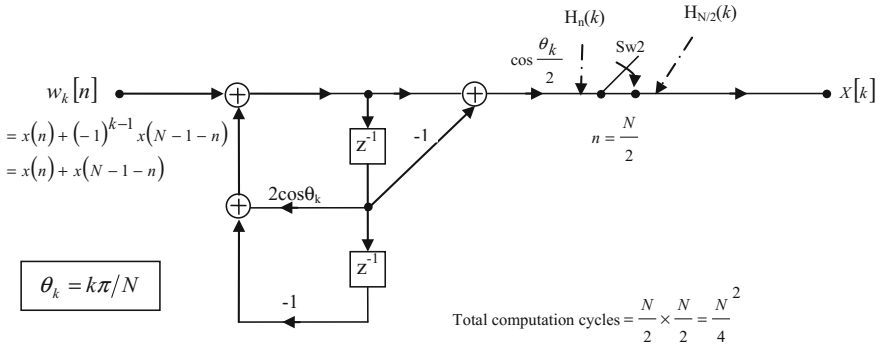

Fig. 2 Computation of DST kernels $X[k]$ for odd values of k

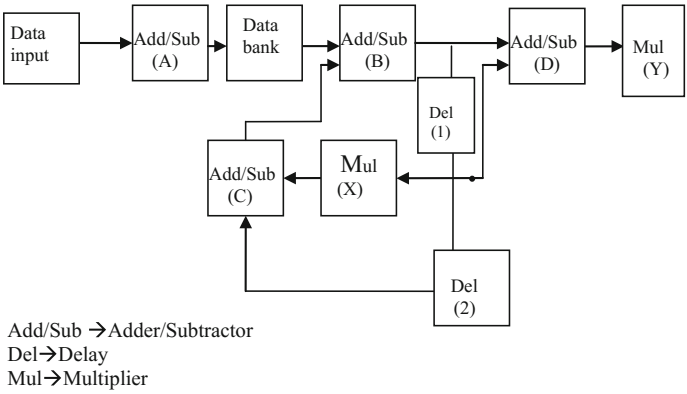

Fig. 3 Generalized functional block diagram of DST

and

$$X[k] = \sqrt{\frac{2}{N}}\varepsilon_k(-1)^{\left(\frac{k-1}{2}\right)}H_{N/2}[k], \quad k = 1, 3, \ldots, N-1 \tag{3a}$$

where

$$H_i[k] = \sum_{n=0}^{i-1} w_k[i-n]\cos\left(n+\frac{1}{2}\right)\theta_k, \quad k = 1, 3, \ldots, N-1 \tag{3b}$$

$$H_i[k] = \cos\frac{\theta_k}{2}\{w_k[i] - w_k[i-1]\} + 2\cos\theta_k H_{i-1}[k] - H_{i-2}[k], \quad k \text{ odd} \tag{3c}$$

Detailed mathematical analysis is provided in the [18]. This work focuses on the architecture modeling of DST and its synthesis using VHDL.

The generalized block diagram for hardware realization of the algorithm is shown in Fig. 3.

3 Architecture for the Proposed Algorithm

Architecture of the proposed algorithm is shown in Fig. 4. This paper uses VHDL as a design entity, and their Synthesis by Xilinx (Virtex7) Synthesis Tool. The schematic depicts the various entities defined in accordance with the recursive structure that computes the discrete sine transform using separate entities for odd and even values of kernels (k). At each rising edge of the clock, 8-bit data is entered serially into the element's block. After 'N' clock pulses, i.e., after all the data values, i.e., $x(n)$ have been entered in the data bank, the start signal becomes high. As soon as the start signal is high, the process of computation of discrete sine transform begins. The data is then fed into the w(k)_gen blocks, which generates the $w_k[n]$ terms according to the proposed algorithm for both odd and even values of 'k,'

where

$$w_k[n] = \left[x[n] + (-1)^{k-1} x(N-n+1)\right], \quad k = 1, 2, \ldots, N$$

Its syntax is:

```
architecture Behavioral of gen_wk_odd is

signal sum : std_logic_vector(7 downto 0): = (others=>'0');

    -sum is represented by ADD/SUB (A) block in Fig. 3

signal n : std_logic_vector(5 downto 0): =(others=>'0');

process(clk, en, rst, start)
  begin
      if(rising_edge(clk)) then
             if(rst = '1') OR start='0' then
                    sum <= (others=>'0');
                    n <=(others=>'0');
                    N_2 <= '0';
                    DV  <='0';
             elsif(en = '1') AND start = '1' then
                      sum <= std_logic_vector(signed(data(n))+signed(-
data(N-n-1)));
                      DV<= '1';
```

```
                    if (n = N/2)
                        then N_2 <= '1';
                          n <= '0';
                    else
                        n <=n+1;
                    end if;
               else
                    DV<='0';
                    N_2<='0';
                    sum<=(others=>'-');
                end if;
           end if;
   end process;

w_k <= sum;                                    -sum is represented by data
bank
block in Fig. 3.

 end Behavioral;
```

Depending on the value of k, the remaining computation takes place in separate DSP blocks. Cosine and sine terms required for computation of DST have been generated cos_ bank_block. This could be implemented using CORDIC algorithm or a lookup table. In this paper, lookup table has been developed (cos_ bank_block) for extracting cosine and sine values to ensure simplicity of design. This, however, has restricted the values of 'N' for which DST can be computed. For ensuring a small-sized lookup table, assumption is made on 'N' can take even values ranging from 2 to 32. The design can also be implemented for values of N greater than 32 simply by increasing the size of lookup table. The cosine and sine values obtained from the cos_bank_block are 24 bits wide and multiplied with a factor of (222). The dsp_odd and dsp_even blocks are the multiplier accumulator (MAC) blocks that compute $G(k)$ and $H(k)$, respectively.

$$G_i[k] = \sin\frac{\theta_k}{2}\{w_k[i] + w_k[i-1]\} + 2\cos\theta_k G_{i-1}[k] - G_{i-2}[k], \quad k\ even$$

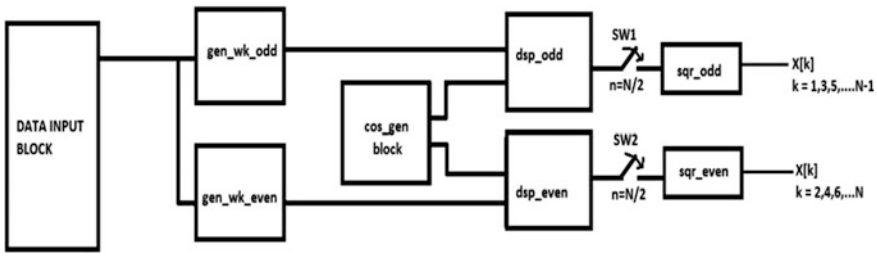

Fig. 4 Proposed architecture for DST operation

and

$$H_i[k] = \cos\frac{\theta_k}{2}\{w_k[i] - w_k[i-1]\} + 2\cos\theta_k H_{i-1}[k] - H_{i-2}[k], \quad k\,odd$$

The syntax for this operation is:

```
-sum1,sum2,sum3,mul1,mul2 are variables used inside process block and
are
mapping
- the functions ADD/SUB (B), (C), (D) and MUL(X), (Y) respectively in Fig. 3
- z1,z2 are delay elements indicating the mapping functions    Del(1) and
Del(2)
- respectively in Fig. 3

- inside the process block :

z2 <= z1;
mul1 : = std_logic_vector((signed(z1)*signed(cos_t1)));

             - mul1 is resized and then used in further computation

sum2: = std_logic_vector((signed(mul1)-signed(z2)));

             - sum1 is assigned value of sum2 after multiplying by a
suitable factor.

sum3 : = std_logic_vector((signed(sum1)-signed(z1)));

mul2 : = std_logic_vector((signed(sum3)*signed(cos_t2)));

z1 <= sum1;

DV <='1';
value <= mul2;
```

The dsp units are the main computing blocks of the design implemented in VHDL. It can be clearly seen that dsp block is implemented using two 36 × 24 multipliers and three 36-bit adders/subtractors apart from one 1-bit register (data valid register), one 16-bit register (register storing output value) and two 36-bit registers (delay elements). Here, sqr_even block multiplies the G(k) value by

Table 1 Resources that operators use on FPGA

Device utilization summary (estimated values)			
Logic utilization	Used	Available	Utilization (%)
Number of slice registers	299	437,600	0
Number of slice LUTs	448	218,800	0
Number of fully used LUT-FF pairs	130	617	21
Number of bonded IOBs	329	300	109
Number of BUFG/BUFGCTRLs	1	32	3
Number of DSP4SEls	18	840	29

$$\sqrt{\frac{2}{N}}\varepsilon_k(-1)^{\frac{k}{2}-1}$$

As for k even, one value of $X[k]$ takes $N/2$ computational cycles, and sqr_even block will operate after $N/2$ clock cycles of the dsp block. Similarly, sqr_odd block multiplies $H(k)$ value by

$$\sqrt{\frac{2}{N}}(-1)^{(k-1)/2}$$

and operates after N/2 clock cycles of the dsp block operation. Thus, by parallel processing even or odd values of $X[k]$ are computed in $N^2/4$ computational cycles.

4 Synthesis Results

The proposed DST algorithm architecture of Fig. 4 is implemented using VHDL and synthesized on Xilinx Virtex7 (vh290thcg1155-2) FPGA. Input is a 12-bit data matrix. The utilization summary is given in Table 1. This architecture is operating at a clock frequency of 61.104 MHz. Total time required for computation of 12 samples is 0.8 us and total of 309,312 kB memory are used in this architecture. Further improvement in the speed could be obtained by appending the next input data simultaneously when DSP block is performing the operation on previous data. There is another possibility of increasing the speed by introducing parallel processing. This design is easily extendable to other matrix sizes.

Values of $x(n)$ used:

$$x[n] = [\,12 \quad 3 \quad 14 \quad 12 \quad 4 \quad 35 \quad 60 \quad 7 \quad 8 \quad 12 \quad 13 \quad 15\,] \quad n = 1,\ 2,\ \ldots,\ 12$$

Values: results for $X[k]$

S. No.	Virtex7
1.	59.4700927734375
2.	−4.3626708984375
3.	−14.3636474609375
4.	1.224609375
5.	32.299072265625
6.	−13.57275390625
7.	−17.65478515625
8.	8.4957275390625
9.	21.8388671875
10.	−6.6588134765625
11.	5.00048828125
12.	7.8046875

5 Conclusion

FPGA-based architecture for discrete sine transform algorithm is proposed and implemented using VHDL on Xilinx Virtex7. The proposed architecture is realized for $N = 12$ and can be easily extended for any value of N simply by changing the lookup table size. Operating frequency of the proposed architecture on the selected FPGA is found to be 61.104 MHz.

References

1. J.G. Proakis, D.G. Manolakis, *Digital Signal Processing Principles, Algorithms, and Applications*, 2nd edn. (Prentice Hall, Englewood Cliffs, New Jersey, 1992)
2. L.R. Rabiner, B. Gold, *Theory and Application of Digital Signal Processing* (Prentice Hall, Englewood Cliffs, New Jersey, 1975)
3. R.C. Gonzalez, P. Wintz, *Digital Image Processing* (Addison-Wesley, Reading, MA, 1977)
4. N. Ahamed, T. Natarajan, K.R. Rao, Discrete cosine transform. IEEE Trans. Comput. **COM-23**, 90–93 (1974)
5. A.K. Jain, A fast Karhunen-Loeve transform for a class of random processes. IEEE Trans. Commun. **COM-24**, 1023–1029 (1976)
6. S. Cheng, Applications of the sine transform method in time of flight positron emission image reconstruction algorithms. IEEE Trans. Biomed. Eng. **BME-32**, 185–192 (1985)
7. K. Rose, A. Heiman, I. Dinstein, DCT/DST alternate-transform image coding. IEEE Trans. Commun. **38**(1), 94–101 (1990)
8. Z. Wang, L. Wang, Interpolation using the fast discrete sine transform. Sig. Process. **26**, 131–137 (1992)

9. P. Yip, K.R. Rao, A fast computation algorithm for the discrete sine transform. IEEE Trans. Commun. **COM-28**(2), 304–307 (1980)
10. A. Gupta, K.R. Rao, An efficient FFT algorithm based on the discrete sine transform. IEEE Trans. Sig. Process. **39**(1), 486–490 (1991)
11. A. Gupta and K.R. Rao, A fast recursive algorithm for the discrete sine transform. IEEE Trans. Acoust. Speech Sig. Process. **ASSP-38**(3), 553–57 (1990)
12. F.Y. Huang, P.Z. Lee, in *Design of Algorithms for the Discrete Sine Transform*. Proceedings of International Computer Symposium, Taichung, Taiwan, December 1992, pp. 733–42
13. P.Z. Lee, F.Y. Huang, Restructured recursive DCT and DST algorithms. IEEE Trans. Sig. Process. **43**(7), 1600–1609 (1994)
14. S. An, C. Wang, Recursive algorithm, architectures and FPGA implementation of the two-dimensional discrete cosine transform. IET Image Process. **2**(6), 286–294 (2008)
15. E.D. Kusuma, T.S. Widodo, in *FPGA Implementation of Pipelined 2D-DCT and Quantization Architecture for JPEG Image Compression*. International Symposium on Information Technology (ITSim), vol. 1 (Kuala Lumpur, 15–17, June 2010), pp. 1–6
16. M. Jridi, A. Alfalou, in *A Low-Power, High-Speed DCT Architecture for Image Compression: Principle and Implementation*. 18th IEEE/IFIP, VLSI System on Chip Conference (VLSI-SoC), 27–29 September 2010, pp. 304–309
17. Z. Guo, B. Buyukkurt, W. Najjar, K. Vissers, in *Optimized Generation of Data-Path from C Codes for FPGAs. Proceedings of the Conference on Design, Automation and Test in Europe*, vol. 1 (March 2005), pp. 112–117
18. Jain Priyanka, B. Kumar, S.B. Jain, Discrete sine transform and its inverse realization through recursive algorithms. Int. J. Circ. Theor. Appl., Wiley Interscience **36**, 441–449 (2007)

Design of High-Performance Mixed-Clock FIFO

Shilpi Maurya, Satyendra Sharma and Jitendra Saroj

Abstract Modern digital design techniques are based on globally asynchronous locally synchronous (GALS) architecture which promises to combine the advantages of both synchronous and asynchronous design techniques. System on chip (SOC) based on GALS architecture is divided into many synchronous blocks where each synchronous blocks operate at its own clock. To establish, the communication between the blocks operating in different clock domains is challenging task due to synchronization failure. Therefore, it is desirable to design mixed-clock FIFO (First In First Out) as an interface between different clock domains. In this paper, we have designed an 8 × 32 mixed-clock FIFO which is able to interface two synchronous systems with independent clock frequencies. Our design is simulated using Xilinx 14.7 and modelsim14.4a. Dynamic power of mixed-clock FIFO is 5 mw when write clock and read clock frequencies are 100 MHz which is a most desirable power for an enhanced design. Delay of the design is 4.886 ns.

Keywords GALS · SOC · Mixed clock · FIFO

1 Introduction

Modern design techniques integrate benefits of both synchronous and asynchronous design techniques to meet the demand of current applications. Fully asynchronous design technique lacks global timing reference; hence, it has shown to be good solution to the problems related to clocks. However, several factors prevent a major

S. Maurya (✉) · S. Sharma · J. Saroj
Department of ECE, Noida Institute of Engineering and Technology, Greater Noida, India
e-mail: shilpi.et2007@gmail.com

S. Sharma
e-mail: satyendracommn@gmail.com

J. Saroj
e-mail: jitendra.saroj007@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_3

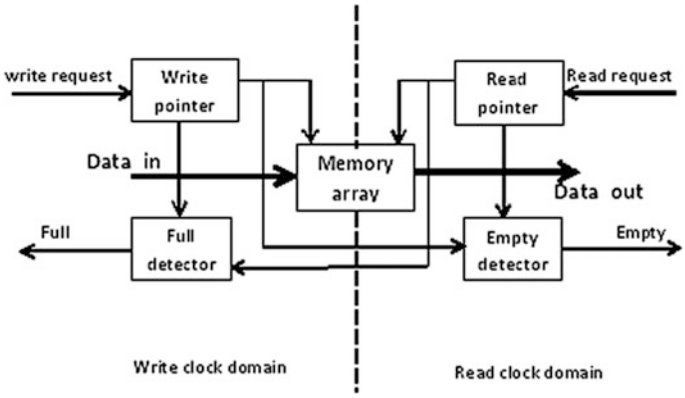

Fig. 1 Block diagram of FIFO

transition to fully asynchronous design style, including the lack of mature design tools and the unwillingness of the industry to incur the cost and risk of moving away from a synchronous design style that has been acknowledged as successful in the past. GALS-based architecture is an intermediate solution that combines the benefits of both design styles [1, 2]. A GALS system consists of locally clocked synchronous blocks. Communication between different synchronous modules is handled asynchronously. We have presented mixed-clock FIFO or dual clock FIFO which is inserted between sender block and receiver block when both the sender and the receiver are synchronous, potentially under arbitrary uncorrelated clock. To implement the design, token ring-based architecture is used. Like other design styles of FIFO, our mixed-clock FIFO (Fig. 1) consists of write pointer, read pointer, memory array, detector, and empty detectors. Write pointer and full detector operate in write clock domain while read pointer and empty detector operate in read clock domain. Memory array can operate in any one out of two domains. Write pointer points to the next location where the data is to be written in. Read pointer points to the location where the data is to be read from. Both write and read events take place at the rising edge of their respective clocks.

2 Related Works

Fully asynchronous FIFOs generally come into view in the literature but these design styles do not utilize clocks, and therefore, they cannot be employed in cases of synchronizing data between clock domains. One common technique for designing an asynchronous FIFO is to use gray code pointers that are got synchronized to the opposite clock domain before generating synchronous FIFO full or empty status signals [3]. But these FIFOs were able to implement only $\log_2 N$ depth. Other dual clock FIFO architectures use independent registers for memory element

[4]. The pointers of mixed-clock FIFO proposed in [5, 6, 7] benefit from token ring-based structure while the FIFOs presented in [8] use Johnson Coding for read and write pointers. A unique idea of haltable clock has been used for full and empty detection in [9]. In our mixed-clock FIFO, we have used register-based memory and token ring-based pointers. Synchronization issue is resolved by using modified two flip-flop synchronizer as has been done in [6].

3 FIFO Architecture

Figure 2 shows the memory structure used for FIFO. FIFO memory is generally an instantiated dual port random access memory (RAM) but other memory styles can also be adapted to function as memory array [10]. In our design, we have used register-based memory as data buffer for FIFO. The data buffer can be accessed by both read and write clock domain.

Mixed-clock FIFO architecture is composed of two token ring-based pointers each containing a token. In the sender interface, the token ring pointer is driven by the **w_clk**, synchronous to incoming data; it generates the write pointer while in receiver interface, the token ring pointer is driven by the **r_clk**; it generates the read pointer (Fig. 3). Token ring structure is succession of single-bit registers connected like cyclic shift register with single token. Token is represented by logic state 1 of register. The position of the token (logic state '1') determines the position of

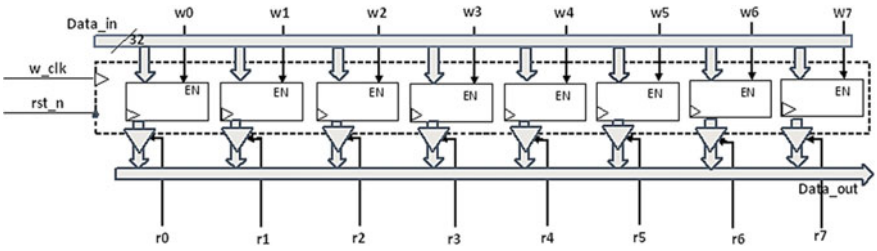

Fig. 2 Register-based memory

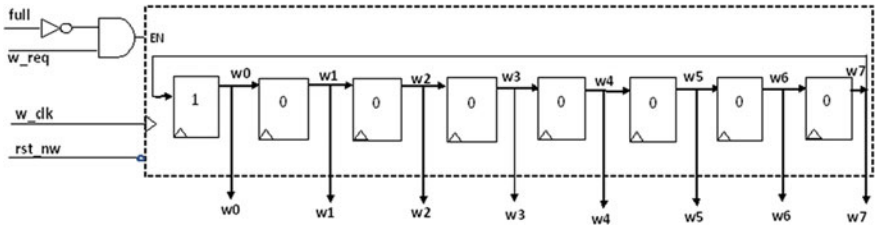

Fig. 3 Write pointer based on token ring

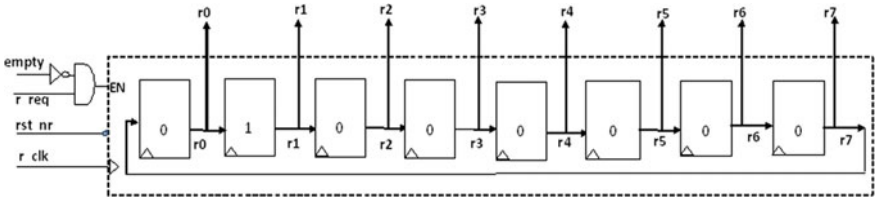

Fig. 4 Read pointer based on token ring

pointers, hence the location in the FIFO memory; therefore, length of write or read pointers is equal to the FIFO depth [10] (Fig. 4).

3.1 Full and Empty Flags

As shown in Fig. 5, full and empty detectors compute the full and empty status of the FIFO using read and write pointer outputs. These detectors perform an asynchronous comparison between the FIFO write and read pointers that are generated in clock domain asynchronous to each other so few synchronizer flip-flops are used. The condition for full flag is when write pointer points to the previous position of the read pointer; i.e., the FIFO contains ($N-1$) elements where N is the depth of the FIFO. FIFO is considered empty when write pointer points to the same position of the read pointer; i.e., the FIFO contains zero element. Empty flag is generated in similar way as full flag is generated. The flag logic in the FIFO also inhibits reading from an empty FIFO and writing to a full FIFO. When reading an empty FIFO, the output will always show that last valid data read from the FIFO. Writes to a full FIFO are discarded.

The full detector decides the value of the full signal depending on the content of write and read pointer. It requires 2-input AND gates (N nos) and N-input OR gate (1 no), where N is the FIFO depth. The detector computes the logic AND operation between the write and read pointer and then collects the result along with an OR gate, thus obtaining **full_pre** signal. After that, modified two flip-flop synchronizer is used to give full flag [6]. Assertion of the **full_pre** signal is synchronous to the

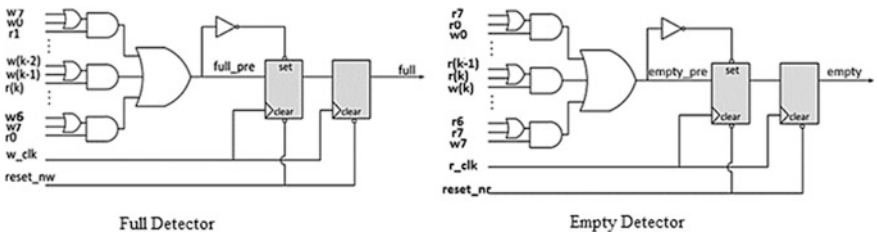

Fig. 5 Full and empty detection

w_clk-domain, since **full_pre** can only be asserted when the write pointer is incremented, but de-assertion of the **full_pre** signal happens when the read pointer is increased, which is an asynchronous event to **r_clk**. On the contrary, assertion of the **empty_pre** signal is synchronous to the **r_clk domain**, since **empty_pre** can only be asserted when the read pointer is incremented, but de-assertion of **empty_pre** happens when the write pointer increments, which is an asynchronous event to **w_clk**. As a consequence, the **full_pre** and **empty_pre** signals, coming out of an asynchronous comparison of read and write pointers, need to be synchronized by means of modified two flip-flop synchronizers.

4 Performance Analysis

Subsequently, after construction of all the key constituents, i.e., data buffer, read pointer, write pointer, full detector, empty detector, the mixed-clock FIFO is designed at gate level using Verilog. Performance analysis is carried out for FIFO with depth 8 and data width 32 bits. The Field Programmable Gate Array (FPGA) area can be estimated by considering the number of Lookup Tables (LUTs) after synthesis illustrated in Table 1. The estimated delay of the design is 4.886 ns. Dynamic power analysis done with the help of Xilinx is illustrated in Figs. 6 and 7. Dynamic power of the design is 5 mw when both write and read clock frequencies are 100 MHz.

Table 2 shows the frequency estimation of 32-bit mixed-clock FIFO as function of FIFO depth. Maximum frequency of operation is greater than 6000 MHz.

Figure 8, 9, 10, and 11 show a gate-level simulation results for 8 × 32 FIFO for different sender/receiver frequency and phase relation scenarios. It can be observed that data is written into the FIFO whenever write request (w_req) is high and FIFO is not full. A data is read out whenever read request (r_req) is high and FIFO is not empty; otherwise, previous data is available at the output port. When receiver's

Table 1 Device utilization summary

Logic utilization	Used	Available	Percentage utilization (%)
No. of slice flip-flop	323	1920	16
No. of four input LUTs	219	1920	11
Number of occupied slices	286	960	29
Number of slices containing only related logic	286	244	100
Number of slices containing unrelated logic	0	244	0
Total number of four input LUTs	219	1920	11
Number of bonded IOBs (Input/output block)	70	83	83
Number of BUFG MUXs	2	24	8
Avg. fan-out of non-clock nets	3.05		

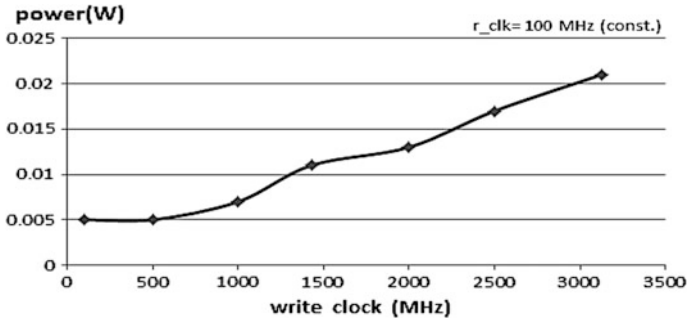

Fig. 6 Dynamic power and write clock frequency relation

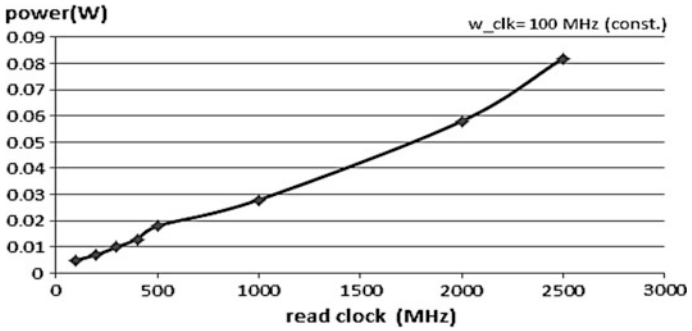

Fig. 7 Dynamic power and read clock frequency relation

Table 2 Frequency and FIFO depth

FIFO depth	Max. write frequency	Max. read frequency
8 × 32	Greater than 6000 MHz	Greater than 6000 MHz
7 × 32	Greater than 6000 MHz	Greater than 6000 MHz

clock frequency is three times the sender's clock and vice versa, FIFO operates normally with correct full and empty flags as shown in Figs. 8 and 9. Similar operation is shown when both frequencies are totally different in phase and frequency (Fig. 10). Figure 11 shows the erroneous value of *full* and *empty* flags at circled position of the waveform as FIFO cannot be full and empty at the same time.

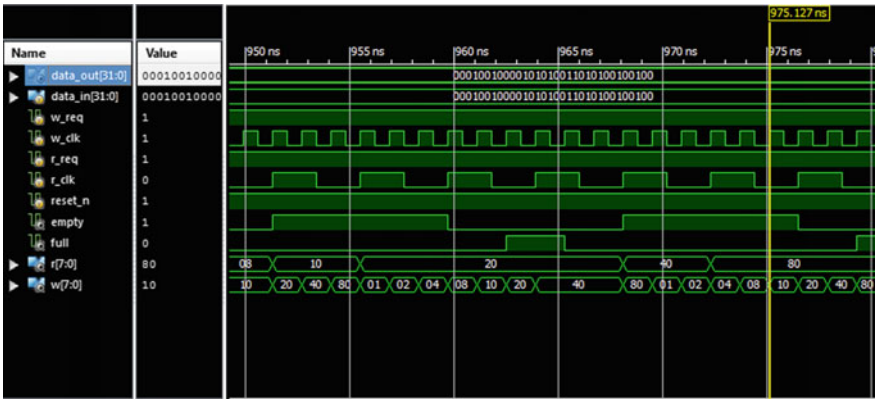

Fig. 8 FIFO operation when write and read frequency has a relation 3:1

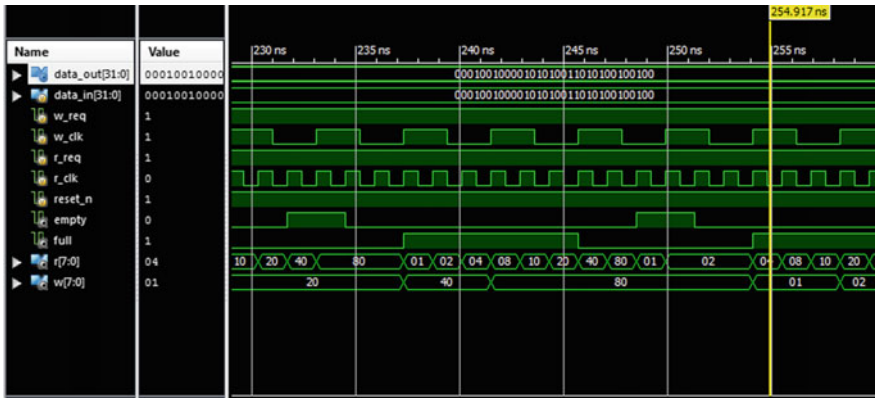

Fig. 9 FIFO operation when write and read frequency has a relation 1:3

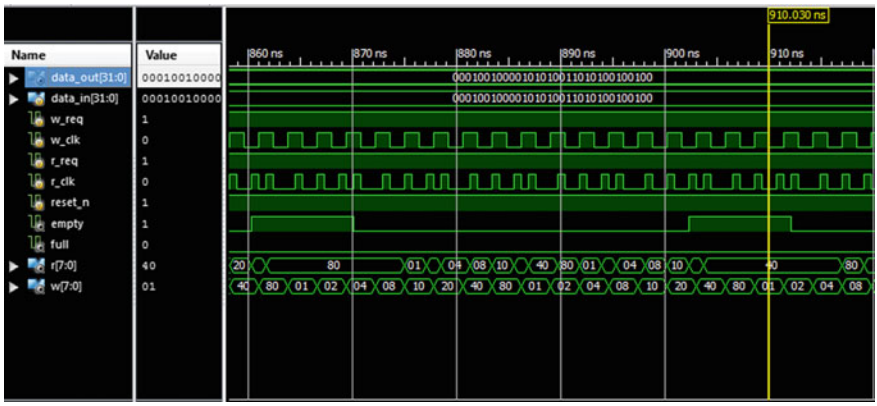

Fig. 10 FIFO operation when write and read frequency has arbitrary relation

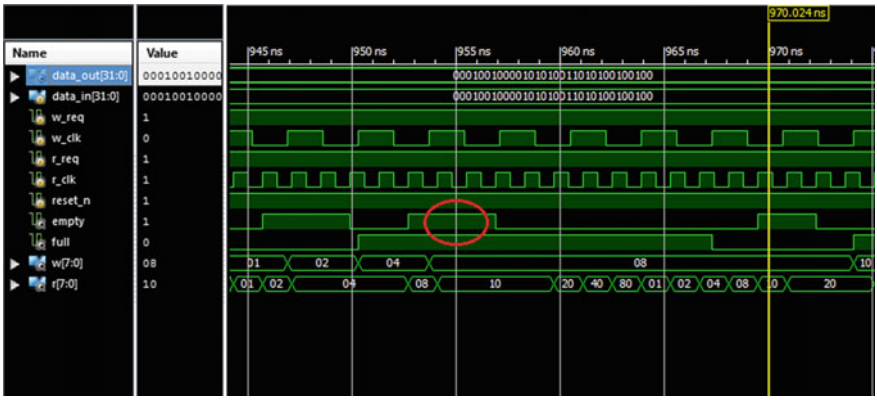

Fig. 11 Erroneous full and empty flags

5 Conclusions

A mixed-clock FIFO that can transfer data between unrelated clocks is presented in this paper. Dynamic power consumption of the design is 5 mw at frequency of 100 MHz. Most of the circuit components (LUTs) are utilized in the range of 8–16%. This circuit also reduces the delay period of 4.886 ns which is feasible for an enhanced design. Max. read and max. write frequencies of operation are greater than 6000 MHz. Erroneous full and empty flag detection is subject to future research work.

References

1. L. Benini, G. De Micheli, Networks on chips: a new SoC paradigm. Computer **35**(1),70–78 (2002)
2. P. Teehan, M. Greenstreet, G. Lemieux, A survey and taxonomy of GALS design styles. IEEE Des Test Comput 418–428 (2007)
3. C.E. Cummings, *Simulation and Synthesis Techniques for Asynchronous FIFO Design with Asynchronous Pointer Comparisons* (San Jose, CA, 2002)
4. T. Chelcea, S.M. Nowick, Robust interfaces for mixed-timing systems. IEEE Trans. Very Large Scale Integr. Syst. **12**, 857–873 (2004)
5. I. Miro Panades, A. Greiner, in *Bi-Synchronous FIFO for Synchronous Circuit Communication Well Suited for Network-on-Chip in GALS Architecture*. Proceedings of the ACM/IEEE International Symposium on Networks-on-Chip (Princeton, 2007), pp. 83–89
6. T.T. Nguyen, X.T. Tran, in *A Novel Asynchronous First-In-First-Out Adapting to Multi-synchronous Network-on-Chips*. 2014 International Conference on Advanced Technologies for Communications (ATC2014) (2014), pp. 365–370
7. T. Ono, M. Greenstreet, in *A Modular Synchronizing FIFO for NoCs*. Proceedings of the 3rd ACM/IEEE International Symposium on Networks-on-Chip (2009), pp. 224–233

8. A.-M. Rahmani, P. Liljeberg, J. Plosila, H. Tenhunen, in *An Efficient VFI-Based NoC Architecture Using Johnson-Encoded Reconfigurable FIFOs*. Proceedings of the IEEE International Norchip Conference (2010), pp. 1–5
9. R.W. Apperson, Z. Yu, M.J. Meeuwsen, T. Mohsenin, B.M. Baas, A scalable dual clock FIFO for data transfers between arbitrary and haltable clock domains. IEEE Trans. Very Large Scale Integr. Syst. **15**, 1125–1134 (2007)
10. A.M. Rahmani, P. Liljeberg, J. Plosila, H. Tenhunen, Design and implementation of reconfigurable FIFOs for voltage/frequency Island-based networks-on-chip. Microprocess. Microsyst. **37**(4–5) (2013)

Mixed Approach of Order Reduction for Single-Input Single-Output (SISO) Systems

R. V. S. Sengar, Kalyan Chatterjee and Jay Singh

Abstract In this paper, a mixed method is proposed which combines the improved Pade approximations and the eigen spectrum analysis for reducing the higher-order system. In this method, system stiffness and pole centroid of both original and reduced-order system remain same. The denominator of higher-order system (HOS) is derived using the eigen spectrum analysis, and the numerator is derived by using improved Pade approximation. The latter method of order reduction utilizes both time moments and Markov parameters. The stability and quality of the reduced-order system are compared with the existing methods of order reduction. To understand the proposed method, paper includes some numerical examples of single-input single-output (SISO) systems.

Keywords Improved Pade approximation · Eigen spectrum analysis
Integral square error · Order reduction · Stability

1 Introduction

Many of the physical systems when realized using mathematical and theoretical considerations, large-order models are obtained which include higher-order differential equations that are difficult to analyze. Therefore, researchers have developed several methods to reduce the higher-order model into required form keeping the original characteristics of the system. Hence, reduction of higher-order differential

R. V. S. Sengar (✉) · K. Chatterjee
EE Department, Indian Institute of Technology Dhanbad, Dhanbad, Jharkhand, India
e-mail: ramveerchiro@gmail.com

K. Chatterjee
e-mail: kalyanbit@gmail.com

J. Singh
EEE Department, GL Bajaj Institute of Technology and Management,
Greater Noida, UP, India
e-mail: jaysinghism@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_4

equation is called model order reduction (MOR). For order reduction of a linear continuous system in the time domain and frequency domain, there are several techniques available in the literature [1–4]. In addition to this, several methods have also been opted by combining the features of two different techniques [5–7]. Mixed methods have advantages and disadvantages both when applied on a particular system. There are no methods available in the literature, which is always suitable for order reduction of any physical system. The current effort is developing an order reduction method in which both the pole centroid and system stiffness of the original and reduced-order systems are kept exactly the same. As per the analysis, the proposed method is applicable on systems having real poles only. Method is illustrated through three numerical examples. The performance of the method is tested on the MATLAB platform.

2 Description of the Method

The proposed order reduction is a mixed transforming system, where the reduced-order denominator polynomial is calculated by eigen spectrum analysis and numerator coefficients are determined by improved Pade approximations. To find the reduced-order model from the original HOS, the method is described in the following sections. Section 2.1 describes the process to obtain the denominator polynomial, whereas Sect. 2.2 describes numerator polynomial.

2.1 Denominator Polynomial for Reduced-Order Model

Let the 'n'-order original SISO system is written as

$$G_n(s) = \frac{N_n(s)}{D_n(s)} = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_{n-1} s^{n-1}}{b_0 + b_1 s + b_2 s^2 + \cdots + b_n s^n} \tag{1}$$

The poles of the HOS are: $-p_1 < -p_2 < \cdots < -p_n$.
Also, 'r'-order reduced SISO system is written as

$$G_r(s) = \frac{N_r(s)}{D_r(s)} = \frac{c_o + c_1 s + c_2 s^2 + \cdots + c_{r-1} s^{r-1}}{d_o + d_1 s + d_2 s^2 + \cdots + d_r s^r} \tag{2}$$

The poles of the lower-order system (LOS) are: $-p_1' < p_2' < \cdots < -p_r'$.

Step 1: If the poles of $-p_i (i = 1, n)$ are located at $-(\mathrm{Re}\, p_i + \mathrm{Im}\, p_i), (i = 1, m)$ within eigen spectrum zone (ESZ), then the ESZ is formed by the two lines passing through the nearest $\mathrm{Re}\, p_1$ and the farthest $\mathrm{Re}\, p_m$ real poles as shown in Fig. 1.

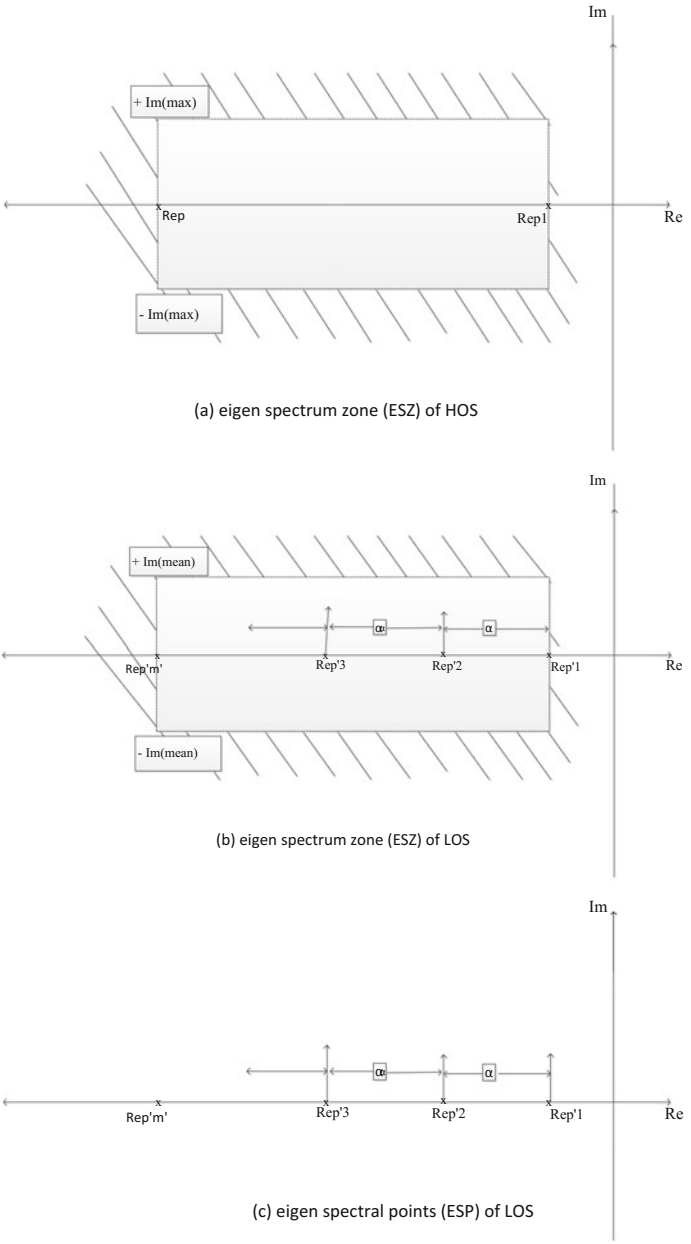

Fig. 1 ESZ and points of the system

Step 2: Now the pole centroid (mean of real parts of the poles) and stiffness (the ratio of the nearest to the farthest pole of a system in terms of real poles only) of the system are evaluated as:

$$p_k \triangleq \frac{\sum_{i=1}^{m} \mathrm{Re}\, p_i}{m} \tag{3}$$

$$p_s \triangleq \frac{\mathrm{Re}\, p_1}{\mathrm{Re}\, p_m} \tag{4}$$

where p_k and p_s are pole centroid and stiffness, respectively.
Step 3: To find eigen spectral points of LOS, assume p'_k and p'_s are pole centroid and stiffness, such that $p'_k = p_k$ and $p'_s = p_s$. With these assumptions, following situation may arise:

$$p'_s = \frac{-\mathrm{Re}\, p'_1}{\mathrm{Re}\, p'_{m'}} = p_s \tag{5}$$

$$p'_k = \frac{\mathrm{Re}\, p'_1 + \mathrm{Re}\, p'_2 + \cdots + \mathrm{Re}\, p'_{m'}}{m'} = p_k \tag{6}$$

where, $p'_i (i = 1, r)$ are the poles of LOS, situated at $[-(\mathrm{Re}\, p'_i \pm \mathrm{Im}\, p'_i), (i = 1, m')]$.

Let;

$$\frac{\mathrm{Re}\, p'_{m'} - \mathrm{Re}\, p'_1}{m' - 1} = \alpha \tag{7}$$

Hence,

$$\mathrm{Re}\, p'_1 + \alpha = \mathrm{Re}\, p'_2; \mathrm{Re}\, p'_2 + \alpha = \mathrm{Re}\, p'_3; \ldots; \mathrm{Re}\, p'_{m'-1} + \alpha = \mathrm{Re}\, p'_{m'}$$

then

$$p_k = \frac{\mathrm{Re}\, p'_1 + \mathrm{Re}\, p'_{m'} + (\mathrm{Re}\, p'_1 + \alpha) + (\mathrm{Re}\, p'_2 + \alpha) + \cdots + (\mathrm{Re}\, p'_{m'-2} + \alpha)}{m'}$$

i.e.,

$$\mathrm{Re}\, p'_1 + \mathrm{Re}\, p'_{m'} + (\mathrm{Re}\, p'_1 + \alpha) + (\mathrm{Re}\, p'_1 + 2\alpha) + \cdots + (\mathrm{Re}\, p'_1 + (m' - 2)\alpha) = p_k m'$$

or

$$\beta = \mathrm{Re}\, p'_1 (m - 1) + \mathrm{Re}\, p'_{m'} + Q\alpha \tag{8}$$

where $\beta = p_k m'$ and $QM = \alpha + 2\alpha + \cdots + (m' - 2)\alpha$

Now, put $\mathrm{Re}\, p'_1 = p_s \mathrm{Re}\, p'_{m'}$

$$\mathrm{Re}\, p'_{m'}(1 - p_s) + \alpha(1 - m') = 0 \tag{9}$$

$$\mathrm{Re}\, p'_{m'}[p_s(m' - 1) + 1] + Q\alpha = \beta \tag{10}$$

$$\begin{bmatrix} p_s(m' - 1) + 1 & Q \\ 1 - p_s & 1 - m' \end{bmatrix} \begin{bmatrix} \mathrm{Re}\, p'_{m'} \\ \alpha \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix} \tag{11}$$

By solving Eq. (11), $\mathrm{Re}\, p'_{m'}$ and α can be calculated and reduced-order denominator $D_r(s)$ can be rewritten as:

$$D_r(s) = (s + p'_1)(s + p'_2)\ldots(s + p'_r) = d_o + d_1 s + d_2 s^2 + \cdots + d_r s^r \tag{12}$$

2.2 Numerator Polynomial for Reduced-Order Model

Let numerator coefficients for the reduced-order model

$$N_r(s) = c_o + c_1 s + c_2 s^2 + \cdots + c_{r-1} s^{r-1}$$

Series expansion of high-order system is

$$G(s) = \sum_{i=0}^{\infty} M_i s^{-i-1} (\text{about } s = \infty) \tag{13}$$

$$= -\sum_{i=0}^{\infty} T_i s^i (\text{about } s = 0) \tag{14}$$

where T_i and M_i are the ith-time moment and Markov parameter of $G(s)$, respectively. The rth-order reduced model may be written as:

$$G_r(s) = \frac{N_r(s)}{D_r(s)} = \frac{\sum_{i=0}^{r-1} c_i s^i}{\sum_{i=0}^{r} d_i s^i} \tag{15}$$

The coefficients of the reduced-order numerator polynomial $N_r(s)$ can be obtained from the following set of equations [8].

$$\left.\begin{aligned}
&c_o = d_o T_o \\
&c_1 = d_o T_1 + d_1 T_o \\
&c_2 = d_o T_2 + d_1 T_1 + d_2 T_o \\
&\cdots \\
&\cdots \\
&c_{\alpha-1} = d_o T_{\alpha-1} + d_1 T_{\alpha-2} + \cdots + d_{\alpha-2} T_1 + d_{\alpha-1} T_o \\
&c_{r-\beta} = d_r M_{\beta-1} + d_{r-1} M_{\beta-2} + \cdots + d_{r-\beta+2} M_1 + d_{r-\beta+1} M_o \\
&c_{r-\beta+1} = d_r M_{\beta-2} + d_{r-1} M_{\beta-3} + \cdots + d_{r-\beta+3} M_1 + d_{r-\beta+2} M_o \\
&\cdots \\
&\cdots \\
&c_{k-2} = d_r M_1 + d_{r-1} M_o \\
&c_{r-1} = d_r M_o
\end{aligned}\right\} \tag{16}$$

The coefficients $c_j; j = 0, 1, 2\ldots, (r-1)$ of the numerator can be calculated by solving the above 'r' linear equations. Finally, the numerator $N_r(s)$ is obtained as

$$N_r(s) = c_o + c_1 s + c_2 s^2 + \cdots + c_{r-1} s^{r-1} \tag{17}$$

Hence, by using the proposed method of order reduction, denominator and numerator polynomials of the reduced-order model can be obtained as explained in Sects. 2.1 and 2.2, respectively.

3 Numerical Examples

The proposed method is explained by two numerical examples chosen from the literature. The method has been compared with the existing methods through performance indices, i.e., an Integral of the Absolute Magnitude of Error (IAE) and an Integral Square of Error (ISE) in between the transient portion of reduced and HOS.

$$\text{IAE} = \int_0^\infty |y(t) - y_r(t)| \mathrm{d}t \tag{18}$$

$$\text{ISE} = \int_0^\infty [y(t) - y_r(t)]^2 \mathrm{d}t \tag{19}$$

where $y_r(t)$ and $y(t)$ are the unit step responses of the reduced model and original systems, respectively.

Example 1 Consider a 4th-order system taken from Mittal [9]

$$G_4(s) = \frac{N_n(s)}{D_n(s)} = \frac{s^3 + 7s^2 + 24s + 24}{s^4 + 10s^3 + 35s^2 + 50s + 24}$$

Poles of the above transfer function are: −1, −2, −3 & −4.

Steps for reducing the 4th-order system into second order are as given below:

Step 1: Fixing of ESZ of HOS:
Since all poles are real, hence it will be a line joining the nearest and farthest pole.
Step 2: Quantification of pole centroid and stiffness of

$$\text{HOS} : p_k \triangleq \frac{-\sum_{i=1}^{m} \text{Re}\, p_i}{m} = \frac{10}{4} = 2.5$$
$$p_s \triangleq \frac{-\text{Re}\, p_1}{\text{Re}\, p_m} = \frac{1}{4} = 0.25$$

Step 3: Determination of eigen spectral points of LOS:
From Eq. (11)

$$\begin{bmatrix} 1.25 & 0 \\ 0.75 & -1 \end{bmatrix} \begin{bmatrix} \text{Re}\, p'_{m'} \\ \alpha \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$\text{Re}\, p'_{m'} = 4\,\alpha = 3$$
$$p'_1 = 1,\ p'_2 = 4$$

Therefore, reduced-order denominator is given by

$$D_r(s) = s^2 + 5s + 4$$

From Sect. 2.2, few time moments and Markov parameters are calculated as

Time moments: $T_o = 1, T_1 = -1.0833, T_2 = 1.0903, T_3 = -1.0666$
Markov parameters: $M_o = 1, M_1 = -3, M_2 = 19, M_3 = -111$

Since,

$$\sigma + \rho = r$$

where

σ Number of time moments
ρ Number of Markov parameters
r Order of reduced model.

We calculate numerator coefficients as follows

$$c_o = d_o T_o = 4 \times 1 = 4$$
$$c_1 = (d_o T_1) + (d_1 T_o) = (4 \times -1.0833) + (5 \times 1) = 0.668$$

Therefore, numerator polynomial is obtained as $N_3(s) = 0.668s + 4$, taking $\rho = 2$, $\rho = 0$ and $r = 2$.

Using the proposed method, the 2nd-order reduced model can be written as

$$G_{12}(s) = \frac{0.668s + 4}{s^2 + 5s + 4}$$

Performance error indices are calculated between the transient portion of 2nd-order reduced model and original model (ISE = 0.0002664 and IAE = 0.02618).

The step response of original and reduced-order models is shown in Fig. 2. The comparative performance indices for reduced-order systems are given in Table 1.

Example 2 Consider an 8th-order system taken from the literature [14].

$$G_8(s) = \frac{N_n(s)}{D_n(s)} = \frac{18s^7 + 514s^6 + 5982s^5 + 36,380s^4 + 122,664s^3 + 222,088s^2 + 185,760s + 40,320}{s^8 + 36s^7 + 546s^6 + 4536s^5 + 22,449s^4 + 67,284s^3 + 118,124s^2 + 109,584s + 40,320}$$

Poles of the above transfer function are: (−1, −2, −3, −4, −5, −6, −7, −8).

Eighth-order original system is converted into 2nd-order reduced model as follows

Step 1: Fixing of ESZ of HOS:
Since all poles are real, it will be a line joining the nearest and farthest poles.
Step 2: Quantification of pole centroid and stiffness of

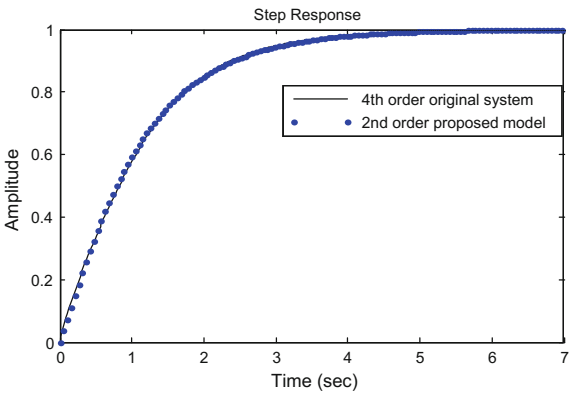

Fig. 2 Step responses of original and reduced-order model for Example 1

Table 1 Comparison of the proposed method for Example 1

Reduction methods	Reduced models	ISE	IAE
Proposed method	$R_2(s) = \frac{0.668s+4}{s^2+5s+4}$	0.0002664	0.02618
Pal [8]	$R_2(s) = \frac{16.008s+24}{30s^2+42s+24}$	0.011	0.2533
Gutman [10]	$R_2(s) = \frac{96s+288}{70s^2+300s+288}$	0.045593	0.3749
Krishnamurty [11]	$R_2(s) = \frac{20.5714s+24}{30s^2+42s+24}$	0.095891	2.4090
Chidambara [12]	$R_2(s) = \frac{-1s+2}{s^2+3s+2}$	0.220	0.9164
Prasad and Pal [13]	$R_2(s) = \frac{s+34.2465}{s^2+239.8082s+34.2465}$	1.53247	1.000

$$\text{HOS} : p_k \triangleq \frac{-\sum_{i=1}^{m} \text{Re}\, p_i}{m} = \frac{36}{8} = 4.5$$
$$p_s \triangleq \frac{-\text{Re}\, p_1}{\text{Re}\, p_m} = \frac{1}{4} = 0.125$$

Step 3: Determination of eigen spectral points of LOS:
Equation (11) can be formed as under:

$$\begin{bmatrix} 1.25 & 0 \\ 0.875 & -1 \end{bmatrix} \begin{bmatrix} \text{Re}\, p'_{m'} \\ \alpha \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

$$\text{Re}\, p'_{m'} = 8, \quad \alpha = 7$$
$$p'_1 = 1, p'_2 = 8$$

Therefore, reduced-order denominator is given by

$$D_R(s) = s^2 + 9s + 8$$

Time moments: $T_o = 1, T_1 = 1.8904, T_2 = -2.5592, T_3 = 2.7924$
Markov parameters: $M_o = 18, M_1 = -133.9520, M_2 = 975.8720, M_3 = -7276.40$

Since,

$$\sigma + \rho = r$$

We calculate numerator coefficients as follows

$$c_o = d_o T_o = 8 \times 1 = 8$$
$$c_1 = (d_o T_1) + (d_1 T_o) = (8 \times 1.8904) + (9 \times 1) = 24.1232$$

Therefore, numerator polynomial is obtained as $N_3(s) = 24.1232s + 8$, taking $\sigma = 2$, $\rho = 0$ and $r = 2$.

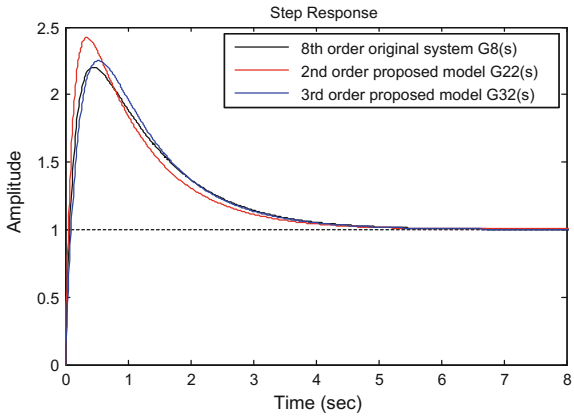

Fig. 3 Step responses of original and reduced-order model for Example 2

Table 2 Comparison of the proposed method for Example 2

Reduction methods	Reduced models	ISE	IAE
Proposed method	$G_{22}(s) = \frac{24.1232s+8}{s^2+9s+8}$	0.0482	0.3006
	$G_{32}(s) = \frac{13.0532s^2+116.5544s+36}{s^3+13.5s^2+48.5s+36}$	0.01499	0.1454
Mukherjee et al. [4]	$R_2(s) = \frac{11.3909s+4.4357}{s^2+4.2122s+4.4347}$	0.05689	0.4572
Mittal et al. [9]	$R_2(s) = \frac{7.0908s+1.9906}{s^2+3s+2}$	0.2689	0.8054
Prasad and Pal [13]	$R_2(s) = \frac{17.98561s+500}{s^2+13.24571s+500}$	1.4584	1.000
Krishnamurthy [11]	$R_2(s) = \frac{155658.6152s+40320}{65520s^2+75600s+40320}$	1.6533	2.4090
Hutton and Friedlad [15]	$R_2(s) = \frac{1.98955s+0.43184}{s^2+41.17368s+0.43184}$	1.9171	10.0702

Using the proposed method, the 2nd-order reduced model is synthesized as

$$G_{22}(s) = \frac{24.1232s+8}{s^2+9s+8}$$

Similarly, the 3rd-order reduced model is synthesized as

$$G_{32}(s) = \frac{13.0532s^2+116.5544s+36}{s^3+13.5s^2+48.5s+36}$$

The step response of original and reduced-order models is shown in Fig. 3. Also, error index, i.e., ISE and IAE are calculated between the transient response of reduced model and original model, which is given in Table 2.

4 Conclusions

In this paper, an order reduction technique by combining the approach of improved Pade approximation and eigen spectrum analysis has been presented. In this mixed method, the reduced-order denominator polynomial is calculated by using eigen spectrum analysis and numerator is determined by improved Pade approximation. The proposed method has been explained with two numerical examples having real poles only. The obtained reduced model response is quite good, but in few cases the reduced model develops tendency to be non-minimum phase. Though it is not investigated accurately, the reason may be forced equalization of system stiffness. The method is simple, uneven and takes slight computational time. This approach can also be used for reducing the order of multi-input multi-output (MIMO) systems. Examples 1 and 2 are explained using proposed method, and comparison between the proposed and other existing order reduction techniques is also given in Tables 1 and 2, respectively. From these analyses, it is clear that the projected method is comparable in quality with other existing techniques of model order reduction. The method preserves model stability and avoids any error in between the initial or final values.

References

1. V. Singh, D. Chandra, H. Kar, Improved Routh Pade Approximants, a computer aided approach. IEEE Trans. Autom. Control **49**(2), 292–296 (2004)
2. C.B. Vishwakarma, Order reduction using modified pole clustering and Pade approximations. World Acad. Sci. Eng. Technol. **56**, 787–791 (2011)
3. O.M.K. Alsmadi, Z.S. Abo-Hammour, Substructure preservation model order reduction with power system model investigation. Wulfenia J. **22**(3), 44–55 (2015)
4. S. Mukherjee, R.C. Mittal, Model order reduction using response matching technique. J. Franklin Inst. **342**, 503–519 (2005)
5. S. Mukherjee, Order reduction of linear system using Eigen spectrum analysis. J. Inst. Eng. India IE (I) J. EL **77**(August), 76–79 (1996)
6. J. Singh, K. Chattterjee, C.B. Vishwakarma, MIMO System using Eigen Algorithm and Improved Pade approximations. SOP Trans. Appl. Math. **1**(1), 60–70 (2014)
7. J. Singh, C.B. Vishwakarma, K. Chattterjee, Biased reduction method by combining improved modified pole clustering and improved Pade approximations. Appl. Math. Model. **40**(2016), 1418–1426 (2015)
8. J. Pal, Improved Pade approximants using stability equation methods. Electron. Lett. **19**(11), 426–427 (1983)
9. A.K. Mittal, R. Prasad, S.P. Sharma, Reduction of linear dynamic systems using an error minimization technique, J. Inst. Eng. India IE (I) J. EL **84**, 201–206 (2004)
10. P. Gutman, C.F. Mannerfelt, P. Molander, Contribution to the model reduction problem. IEEE Trans. Autom. Control **AC-27**(2), 454–455 (1982)
11. V. Krishnamurthy, V. Seshadri, Model reduction using the Routh stability criterion. IEEE Trans. Autom. Control **23**(4), 729–731 (1978)
12. M.R. Chidambara, On a method for simplifying linear dynamic dystem. IEEE Trans. Autom. Control **AC-12**, 119–120 (1967)

13. R. Prasad, J. Pal, Stable reduction of linear systems by continued fractions. J. Inst. Eng. India IE (I) **72**, 113–116 (1991)
14. Y. Shamash, Linear System Reduction using Pade Approximation to allow retention of dominant modes. Int. J. Control **21**(2), 257–272 (1975)
15. M.F. Hutton, B. Friedland, Routh approximations for reducing order of linear time invariant systems. IEEE Trans. Autom. Control **20**(3), 329–337 (1975)

Part II
Communication Systems

Designing Coupling Circuits for Communication of High-Frequency Signals Over Power Lines

Seema Arora, Mini Shaji Thomas and Monika Jain

Abstract There are areas where the digital communication may be made through low-voltage distribution network. This helps the utilities to reach at remote locations. Mapping of the consumer through control center is achieved through interfacing of coupling circuits that leads to distribution load management. This paper presents designing of broadband coupling circuit that satisfies specific signal transmission, appropriate bandwidth, and limited number of components. The paper also discusses the significant parameter variations during capacitive couplings and inductive couplings for passive and active network topologies.

Keywords Active filter · Band pass filters · Bandwidth · Coupling circuits Digital communication · Distribution automation · Line filters · Passive filters Voltage gain

1 Introduction

Distribution load management is an important concern for the utilities to exploit the existing infrastructure of the power system network [1]. There is a need to improve the reliability of power line networks [2, 3] and provide standards and systems for utilizing power networks as a medium of communication. The load management of power operations with predefined logic is achieved [4, 5]. The technology may be beneficial for remote and rural areas [6, 7]. There are limitations to transmit data

S. Arora (✉)
Waljat College of Applied Sciences, Muscat, Oman
e-mail: seema.arora@waljat.net

M. S. Thomas
NIT Tiruchirappalli, Tiruchirappalli, India

M. Jain
Department of Electrical and Electronics Engineering,
I.T.S. Engineering College, Greater Noida, India
e-mail: monikajain.bits@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_5

over distribution systems [8, 9], and transmission characteristics of signals over power lines are also based upon topological connections of the power networks [10]. The connection to the power lines is made by using either inductive couplings or capacitive couplings [11]. These coupling circuits require large number of elements to obtain a good bandwidth. The paper presents different topologies of coupling circuits with reduced number of components. The investigation is being carried out for evolving and designing of couplers with cascaded stages involving active and passive band pass filters in the range of 1–22 MHz and even up to 50 MHz.

2 Coupling Circuits

Coupling circuits and signal coupling are the two important aspects in communication using power lines as a medium. If the coupling circuits are not designed well, the communication devices, modems, or other measuring instruments would be at risk by AC mains supply.

The most important function of a coupling circuit is as a filter. It has to block and filter the AC mains voltage and allow the high-frequency communication signals to pass through. Most of the coupling circuits designed are either high pass filters or band pass filters [11]. The coupling circuits need to superimpose the communication signal onto the network. The circuits should aim to provide galvanic isolation and should be capable of withstanding and adapting to varying impedance between the modem and input impedance of power line. There exist mainly two ways of realization which can be either Differential Mode coupling or Common Mode coupling. The Differential Mode coupling is realizable and implementable in low-voltage power line networks where Live (L) constitutes as the first terminal and Neutral (N) constitutes as the second terminal. In Common Mode coupling, both Live (L) and Neutral (N) constitute as the first terminal and Ground (G) constitutes as the second terminal. Theoretically, this is not achievable but practically, it can be achieved because of the high inductance existing between the coupling points and short circuit point. Inductive coupling results in losses up to several decibels. Capacitive coupling, if realized for high pass filtering, is easy and with a compact design. However, practical coupling circuits often apply a combination of both techniques. Multisim is an industry-standard, simulation environment. It is the solution in teaching practices to build expertise through practical application in designing, prototyping, and testing electrical and electronic circuits and observe the waveforms based on the designed circuits. The Multisim design approach helps researchers to save costs and efforts involved in prototype iterations, and therefore, printed circuit board (PCB) designs can be optimized.

Figure 1 represents schematic diagram of a coupler circuit where the filters are connected between Live and Neutral wires and serves to present high impedance to the carrier band frequency of relatively higher frequency range than AC mains frequency. Generally, T-filters are preferred over other configurations. Some

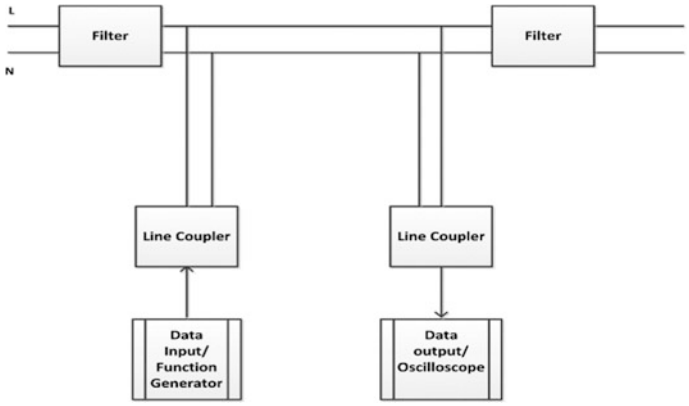

Fig. 1 Schematic representation of a coupler circuit

coupling circuits are discussed which are good enough for narrow bandwidths in kHz frequency range. The advantage offered by them is that their implementation is inexpensive.

2.1 Capacitive Coupling Circuit

A capacitive coupling circuit as given in Fig. 2 is bidirectional in the frequency range up to 0.5 MHz. The power system frequency is damped by the LC-filter at the grid terminals with the 230 V/50 Hz voltages reduced to a 1.411 pV. The zener diodes block voltage spikes and possible resonances in the circuit. The circuit employed for simulation in Fig. 2 has the limitation that there is no galvanic separation with the power circuit. Therefore, in case of a short circuit, a few

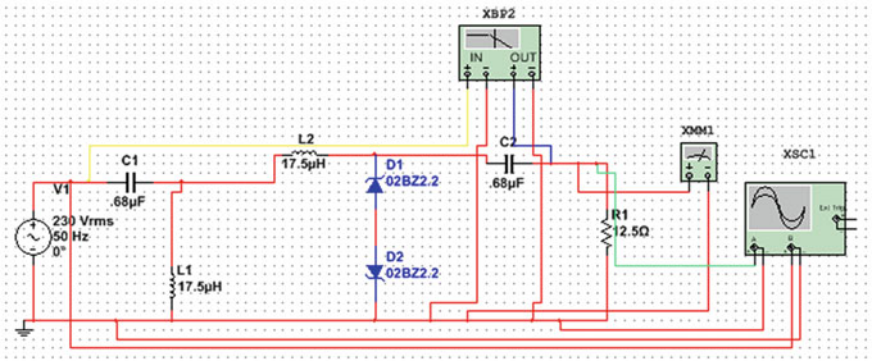

Fig. 2 Capacitive coupling network

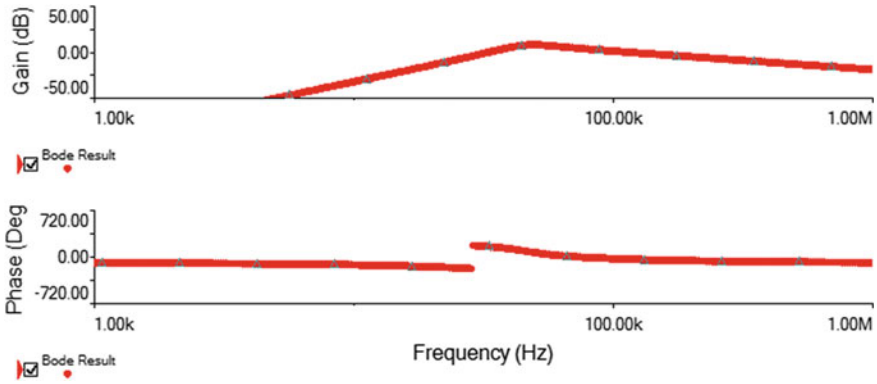

Fig. 3 Bode plot for voltage transfer function and phase transfer function

hundreds of milliamperes current may pass through the circuit. The capacitive coupling, however, remains a cost-effective alternative for communication modems. The bode plots (Fig. 3) developed using Multisim 12.0 represent a summarized result of attenuation and phase shift for a logarithmic range of frequencies, and these results can be used for the design of modulation schematics.

2.2 Capacitor-Coupling Transformer Coupling

A capacitor-coupling transformer circuit as given in Fig. 4 employs high-voltage capacitors to filter out 230 V/50 Hz grid terminals voltage which is reduced to a 153.445 mv. The coupling transformer provides galvanic isolation.

The simulated results are represented in the form of bode plots and oscilloscope waveforms (Figs. 4 and 5).

3 Wideband Band Pass Filter Designing

Band pass filters need to be designed which will allow a particular band of frequencies from ω_1 to ω_2, and attenuate signals with frequencies outside this range. From a typical bode plot for a second-order filter, it can be observed that as Q-factor increases, the bandwidth of the filter becomes smaller, and $H(j\omega)$ gets sharply peaked around ω_0, the center frequency (Fig. 6).

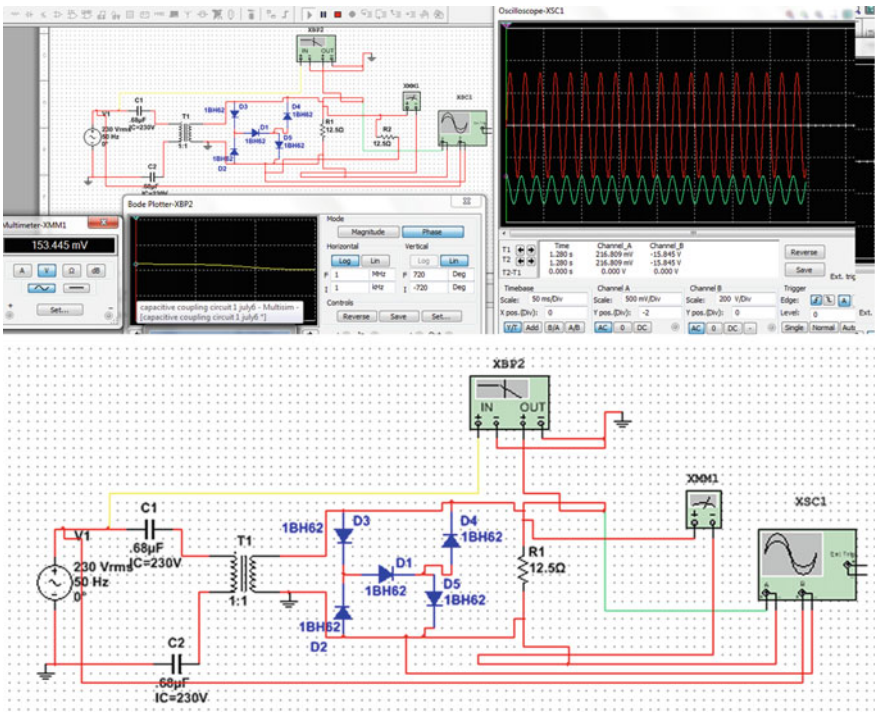

Fig. 4 Capacitor-coupling transformer coupling

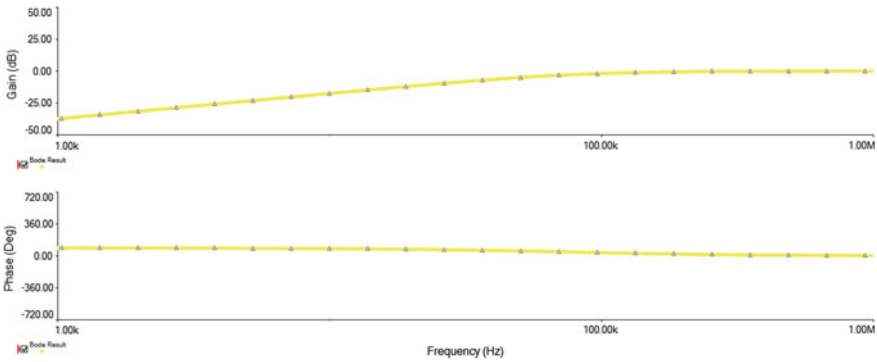

Fig. 5 Bode plot for voltage transfer function and phase transfer function

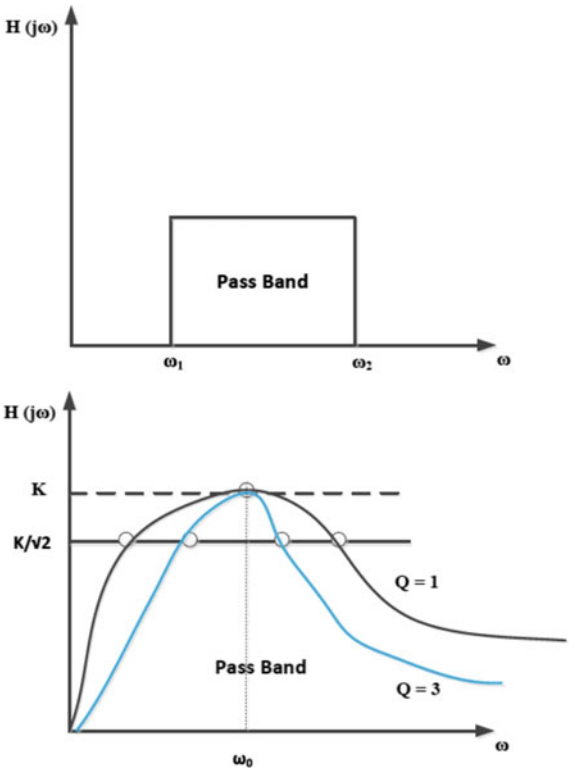

Fig. 6 Band pass filter characteristics

3.1 *Numerical Calculations for a Typical LC Coupling Circuit*

With the values for capacitors selected to be $C_1 = C_2 = 47$ nF, $C_3 = 1$ µF, $C_4 = 10$ nF, and inductances with values of $L_1 = L_2 = 47$ µH and a value of 230 V, 50 Hz mains voltage, the series branch impedance (Z_1) comprising of C_1 and L_1 is (−j67,725.49) ohms and the shunt branch comprising of C_2 and L_2 results in a parallel branch impedance (Z_2) of j0.014765 ohms. Thus, the theoretical value of voltage at output terminals expressed by the Eq. (1) is 50×10^{-6} V. The attenuation as calculated by the Eq. (2) is 133.25 dB (Fig. 7).

The simulations were performed for the circuit configuration of Fig. 8 (at 50 Hz) and Fig. 10 (at 200 kHz), and the results are depicted in Figs. 9 and 11, respectively. It is observed that simulated results are closely linked to theoretical calculated values.

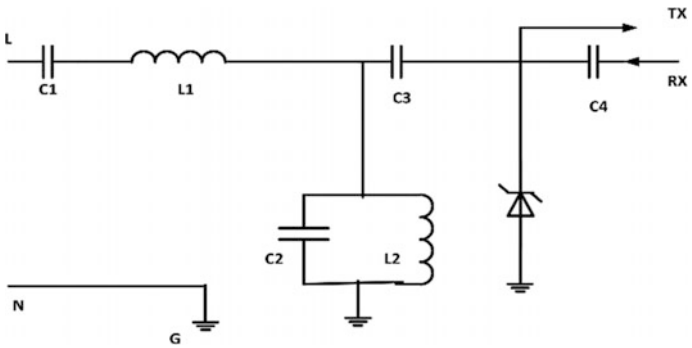

Fig. 7 LC coupling circuit

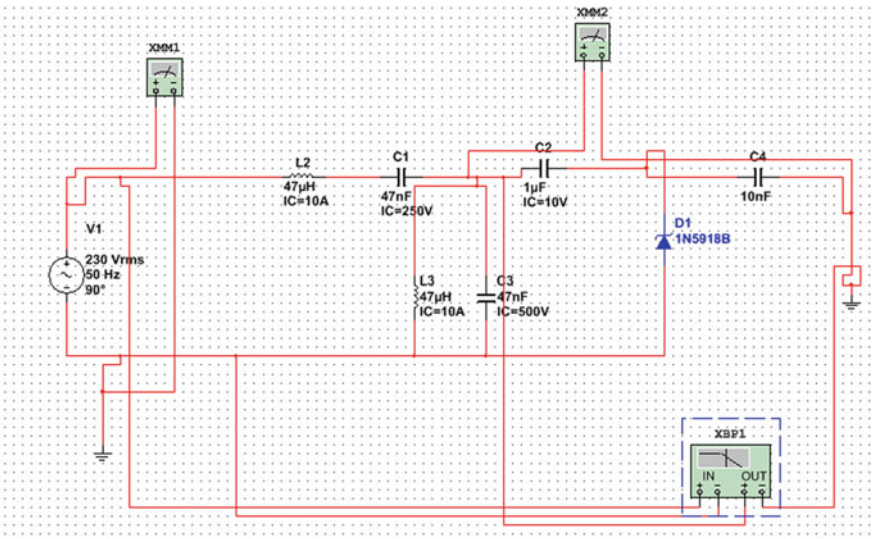

Fig. 8 T-Network

$$V_{oc} = V_{in} \frac{Z_2}{Z_1 + Z_2} \tag{1}$$

$$\text{Attenuation} = 20 \log_{10} \frac{V_{in}}{V_{oc}} \tag{2}$$

Thus, $\text{Attenuation} = 20 \log_{10} \frac{230}{50 \times 10^{-6}} = 133.25$ dB.

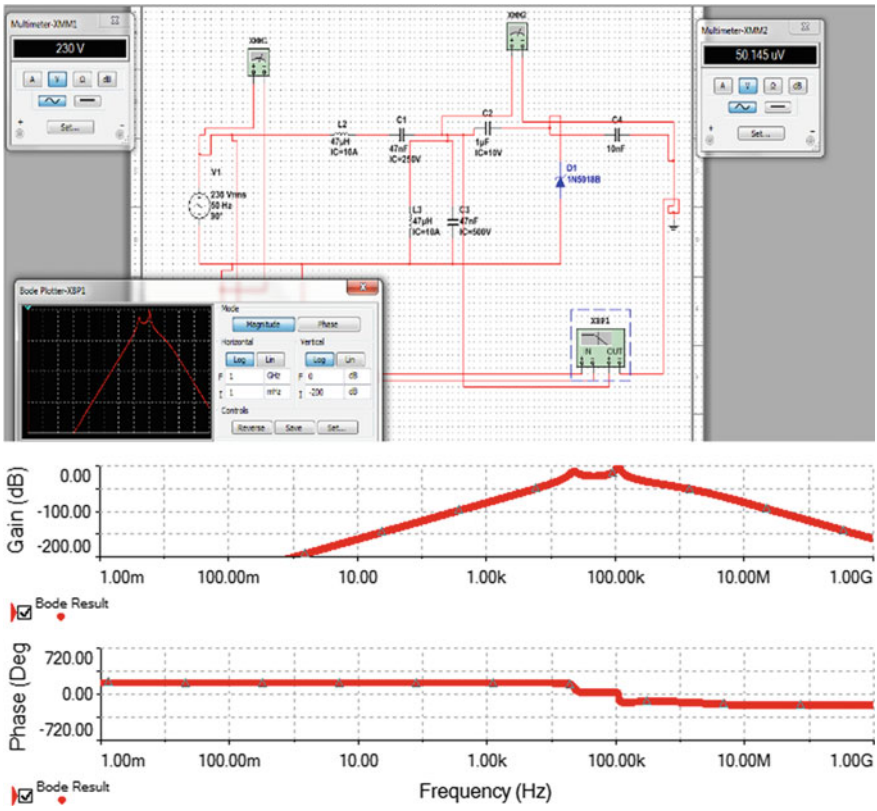

Fig. 9 Results of simulation using Multisim

3.2 Attenuation in DB for Varying Load Impedances

To observe the results for attenuation for differing load impedances and behavior of the network under these conditions, the simulations were performed for the circuit configuration of Fig. 13, and the results are shown in Table 1. The results of simulation (Fig. 14) reveal that with increased resistance of the load (with no inductive reactance), attenuation has decreased. However, keeping load (resistive) value constant at 12.5 ohms and increasing the inductive reactance of load, the signal gets drastically attenuated. The measurements were observed at a fixed typical frequency of 9.807 MHz from Bode Plotter (Fig. 12).

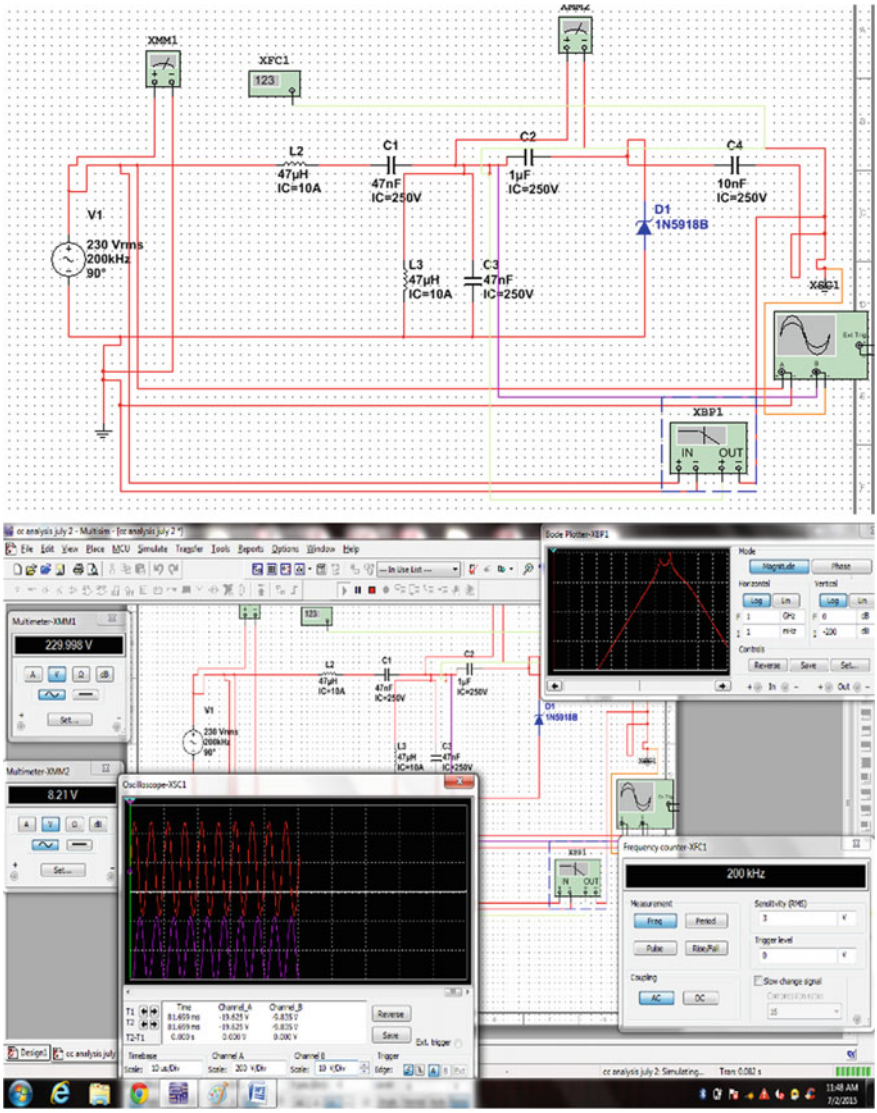

Fig. 10 T-Network and simulation results at 200 kHz

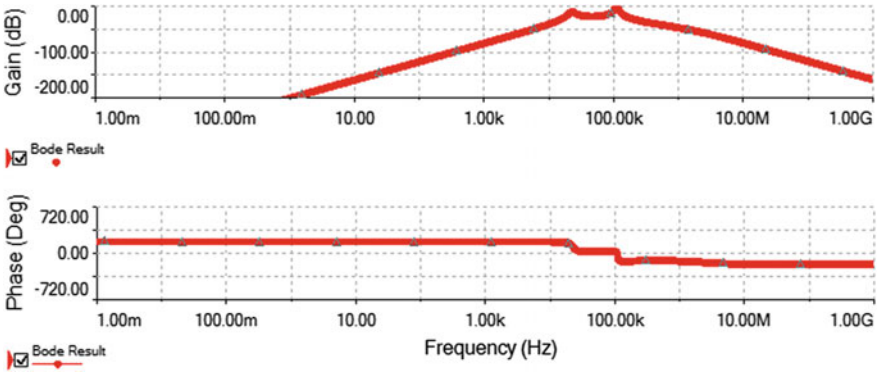

Fig. 11 Results of simulation using Multisim (200 kHz)

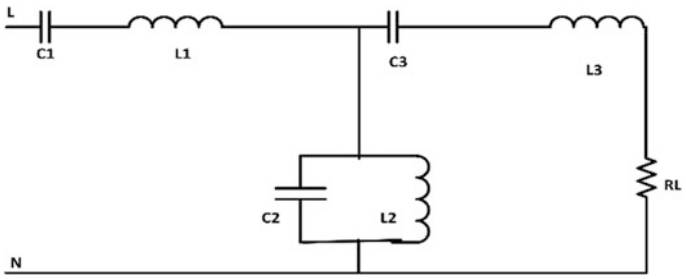

Fig. 12 LC coupling circuit with loaded conditions

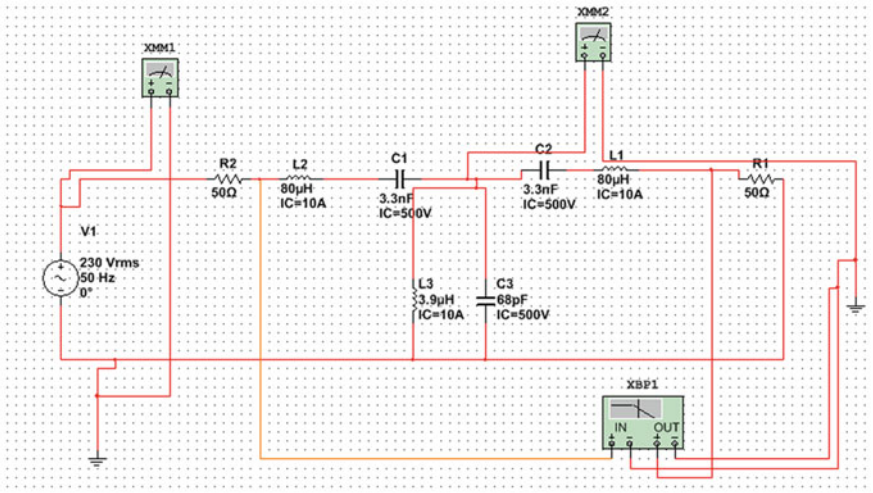

Fig. 13 LC coupling circuit for simulation

Table 1 Bode plot results for attenuation in dB for varying load impedances

Load impedance	Signal attenuation (dB)
50 +j0	−49.373
100 +j0	−43.353
200 +j0	−37.335
12.5 +j $2\pi f \times 18 \times 10^{-6}$	−49.39
12.5 +j $2\pi f \times 30 \times 10^{-6}$	−20.244

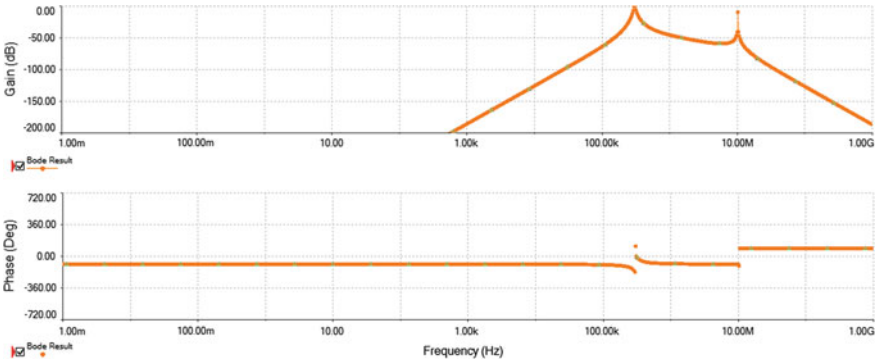

Fig. 14 Bode plot for voltage transfer function and phase transfer function (with 50 +j0 ohms load) for LC coupling circuit

3.3 *Passive Topology Chebyshev Band Pass Filter*

A passive topology, Chebyshev band pass filter (Fig. 15), with the design specifications (Table 2) was built and simulated. The expected typical bode plot for voltage transfer function for a passive band pass filter is shown in Fig. 16, and the bode plots actually obtained along with simulated waveforms are depicted in Fig. 17. Also, the same topology was conceived using Table 3 with still higher values for High End Stop frequency (50 MHz), and the results are revealed in Fig. 18.

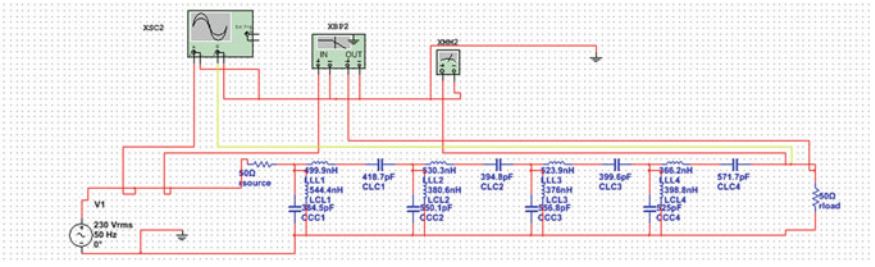

Fig. 15 Passive topology, Chebyshev band pass filter (for 1–22 MHz)

Table 2 Design specifications and criterion used

Design parameter	Specified value
Low End Stop Frequency	1 MHz
Low End Pass Frequency	2 MHz
High End Pass Frequency	20 MHz
High End Stop Frequency	22 MHz
Pass Band Gain	−1 dB
Stop Band Gain	−25 dB
Pass Band ripple	0.01

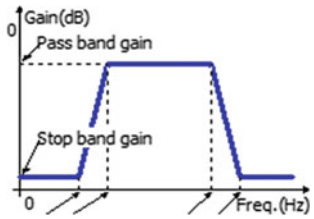

Fig. 16 Typical bode plot for voltage transfer function for a passive band pass filter

3.4 Active Topology Chebyshev Band Pass Filter

Chebyshev band pass filter with an active topology and having design specifications of Tables 2 and 4 were built, and simulated results are depicted in Figs. 20 and 21, respectively. In comparison to the passive circuit, Pass Band ripple of 0.1 was selected as a design value in each case (Fig. 19).

Based on the specifications and number of components to be used in case of active and passive topologies for the coupling circuits to be used for a higher frequency range of 1–50 and 1–45 MHz, respectively, cost comparison was also carried out and it was observed that active design is more economical (by five times) in comparison to the other design worked out for passive topology.

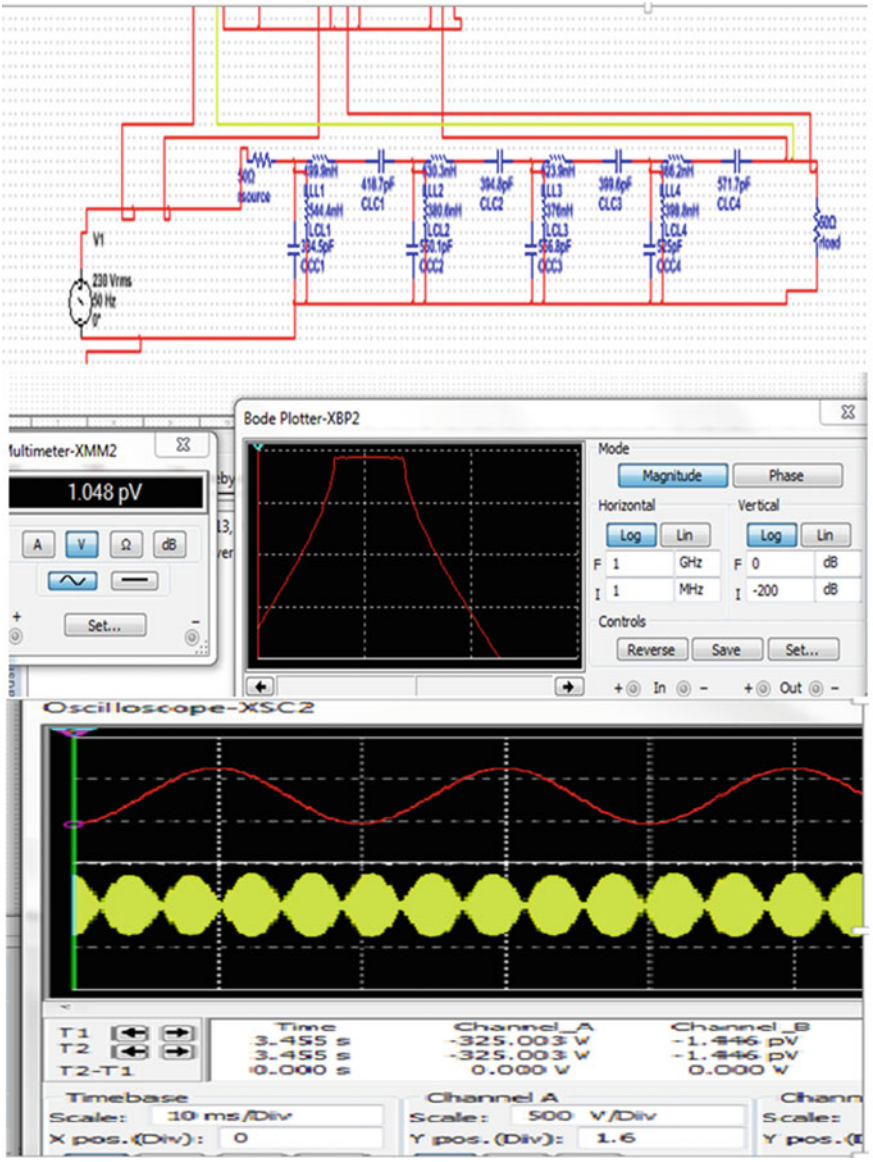

Fig. 17 Actual bode plot for a passive band pass filter (for 1–22 MHz) and results of simulation

Table 3 Design specifications and criterion used

Design parameter	Specified value
Low End Stop Frequency	1 MHz
Low End Pass Frequency	2 MHz
High End Pass Frequency	48 MHz
High End Stop Frequency	50 MHz
Pass Band Gain	−1 dB
Stop Band Gain	−25 dB
Pass Band ripple	0.01

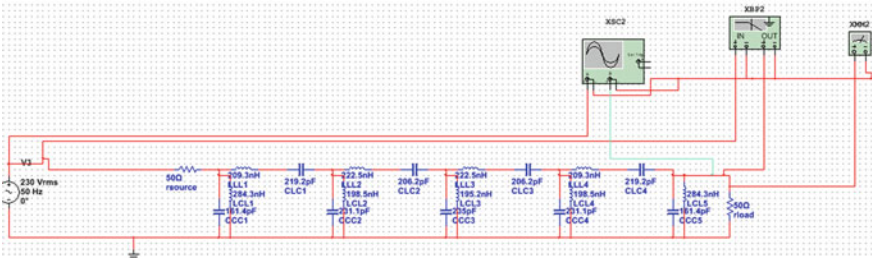

Fig. 18 Passive topology, Chebyshev band pass filter for 1–50 MHz

4 Conclusion

In this work, band pass coupling circuit for broadband communication over power lines is presented. The proposed coupling circuits meet system requirements such as specific signal transmission, appropriate bandwidth, and reduced number of components. The obtained results reveal that the attenuation will decrease as the load resistance increases. However, the signal gets drastically attenuated with increased inductive reactance, provided that the load resistance remains constant. The results also reveal the possibility of reduced cost active circuits, capable of blocking the AC mains and still enabling the designer to use a very accurate, and precise set of components' values to be used for band pass filters. The study demonstrates that the designed systems can be used efficiently and accurately in comparison to previously designed coupling circuits by others.

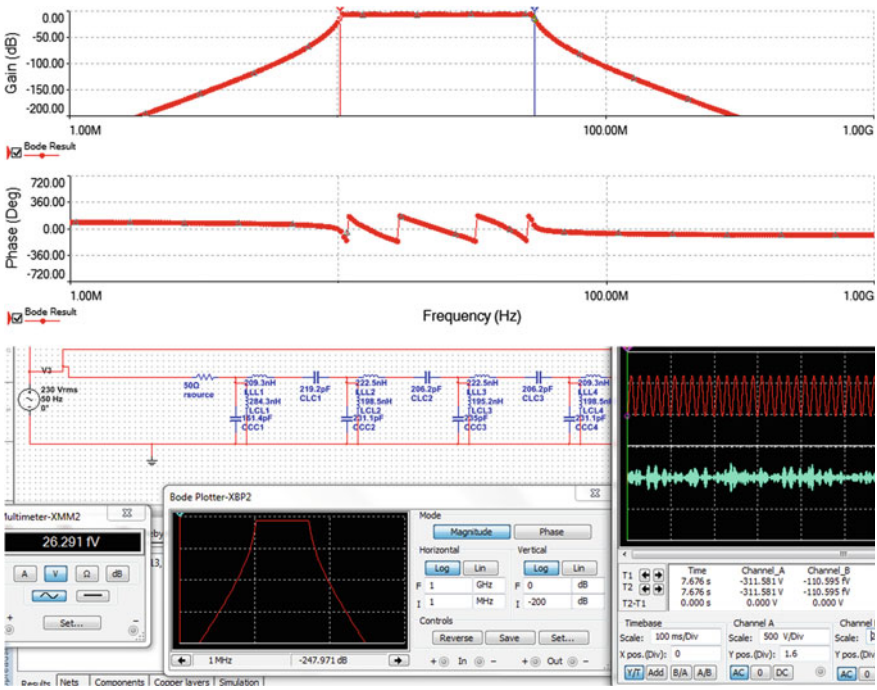

Fig. 19 Actual bode plots for voltage transfer function and phase transfer function (50 +j0 ohms load and 1–50 MHz)

Table 4 Design specifications and criterion used

Design parameter	Specified value
Low End Stop Frequency	1 MHz
Low End Pass Frequency	2 MHz
High End Pass Frequency	40 MHz
High End Stop Frequency	45 MHz
Pass Band Gain	−1 dB
Stop Band Gain	−25 dB
Pass Band ripple	0.1

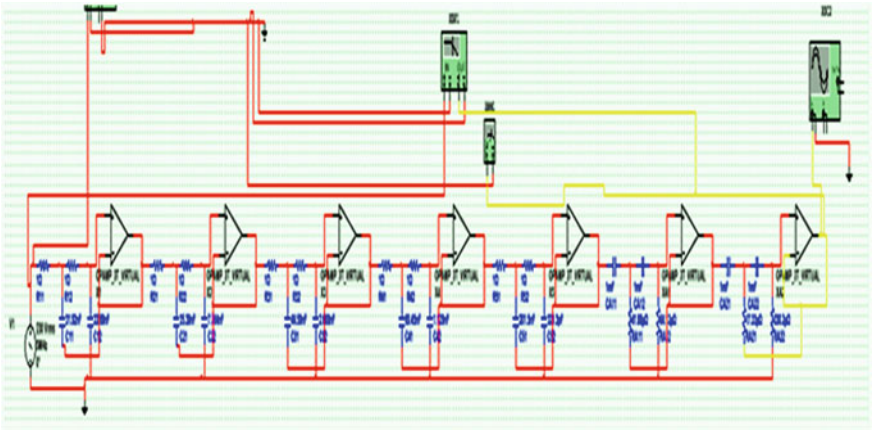

Fig. 20 Active topology, Chebyshev band pass filter (50 +j0 ohms load and for 1–22 MHz)

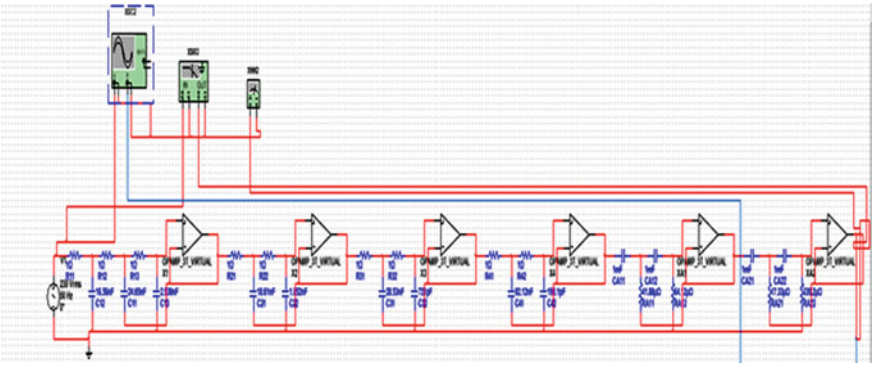

Fig. 21 Active topology, Chebyshev band pass filter (50 +j0 ohms load and for 1–45 MHz)

References

1. H. Philipps, Performance measurements of power line channels at high frequencies, in *Proceedings of the 1998 International Symposium on Power Line Communications and its Applications (ISPLCA'98)*, Apr 1998, pp. 229–237
2. K. Moon, Y. Lee, Y. Son, C. Kim, Universal home network middleware guaranteeing seamless interoperability among the heterogeneous home network middleware. IEEE Trans. Consum. Electron. **49**, 546–553 (2003)
3. C. Lien, Y. Bai, H. Chen, C. Hung, Home appliance energy monitoring and controlling based on power line communication, in *IEEE ICCE* (2009)
4. B. Bolsens, J. Van den Keybus, E. Lemaire, J. Driesen, R. Belmans, in *The Power Line as a Telecommunication Channel for Metering Applications: A Case Study*. Accepted for Distributech-Europe '01, Berlin, Germany, 15 p

5. Y.-S. Son, Home energy management system based on power line communication, in *2100 Digest of Technical Papers International Conference on Consumer Electronics*, 01/2010
6. A. Ipakchi, F. Albuyeh, Grid of the future. IEEE Power Energy Mag. **7**(2), 52–62 (2009)
7. T. Garrity, Getting smart. IEEE Power Energy Mag. **6**(2), 38–45 (2008)
8. S. Galli, A. Scaglione, Z. Wang, Power line communications and the smart grid, in *IEEE International Conference on Smart Grid Communications* (Smart Grid Comm), Gaithersburg, MD, 4–6 Oct 2010
9. D. Nordell, Communication systems for distribution automation, in *IEEE Transmission and Distribution Conference and Exposition*, Bogota, Colombia, 13–15 Apr 2008
10. P. Rengaraju, C.H. Lung, A. Srinivasan, Communication requirements and analysis of distribution networks using WiMAX technology for smart grids, in *International Wireless Communications and Mobile Computing Conference*, 2012
11. M.P. Sibanda, P.A.J. Van Rensburg, H.C. Ferreira, Passive, transformer less coupling circuitry for narrow-band power line communications, in *IEEE International Symposium on Power Line Communications and its Applications*, 2009, pp. 125–130

Metamaterial-based Patch Antennas—Review

Priyanka Garg and Priyanka Jain

Abstract Metamaterials have gained a huge interest in research area due to their extraordinary electromagnetic properties. Metamaterial can be used to overcome disadvantages of patch antenna such as size reduction with acceptable amount of bandwidth and power. This paper gives a brief introduction and classification of basic metamaterial structures and their usage in antenna engineering. The applications of these structures in enhancement of various antenna properties are also discussed. The paper also describes their advantages over conventional antenna structures. Major objective is to analyze the best configuration required while designing the antenna that suits the desired specification and helps in developing the future ideas by utilizing the advantages of the available structures.

Keywords Metamaterial · Patch antenna · Split ring resonator (SRR) Interdigital capacitor

1 Introduction

In the last few decades, a new research area dealing with the study of metamaterials has evolved. Most of the isotropic materials existing in nature have positive values of permittivity (ε) and permeability (μ) greater than unity, hence considered as double positive materials (DPS). These recently developed structures are called metamaterials. These materials are typically engineered with novel or artificial structures, made by unit cells of dimensions much smaller than the wavelength. These unit cells are periodically arranged and tend to exhibit electromagnetic properties that are difficult or unusual to be found in nature. Such properties

P. Garg (✉) · P. Jain
Department of Electronics and Communication Engineering,
Delhi Technological University, Bawana Road, Delhi, India
e-mail: garg.priyanka16@yahoo.com

P. Jain
e-mail: priyajain2000@rediffmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_6

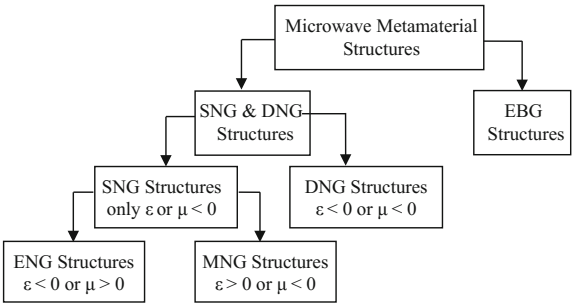

Fig. 1 Basic classification of metamaterial structures

comprise a simultaneous negative permeability and permittivity that render into a negative refractive index; these structures are referred to as double negative (DNG) or left-handed materials (LHM). By using LHM, the properties of a structure can be modified in a better way as compared to the conventional materials. Metamaterials can be used for application in various microwave devices such as antennas, matching networks, and sensors.

This paper is focused on the use of metamaterial structures in patch antennas. Patch antennas are being widely used in wireless links and other microwave applications. Use of metamaterial helps in improving various properties and to overcome the drawbacks of patch antennas, such as miniaturizing the substrates, reducing radiator size, increasing radiation efficiency, controlling antenna's bandwidth, and providing multi-band operation.

First, it is necessary to define the basic classification of metamaterials [1] which has been demonstrated by Buriak et al. using Fig. 1. According to Engheta et al., all metamaterials are broadly classified into two categories [2]. First category refers to DNG (double negative) and SNG (single negative, either ε-negative or μ-negative) structures, in which inclusions and inter-inclusion distances are much smaller than a wavelength. The second category refers to EBG (electromagnetic band gap)-structures, also known as photonic crystals, which involve distances that are nearly half a wavelength or more. These are accounted by the Bragg's reflection and other periodic media concepts.

2 Epsilon Negative Metamaterials ($\varepsilon < 0$)

Epsilon negative (ENG) metamaterials are the materials having permittivity less than zero and permeability more than zero ($\varepsilon < 0$, $\mu > 0$). Many plasma displays possess these characteristics for certain frequency ranges. The first known structure presented by Pendry et al. [3] that gave ε-negative behavior is the one having infinitely long parallel thin metal wires, connected in square matrix form and embedded in a dielectric medium Fig. 2a. Few years later, Falcone et al. [4] presented complementary split ring resonators (CSRRs), shown in Fig. 2b, which was

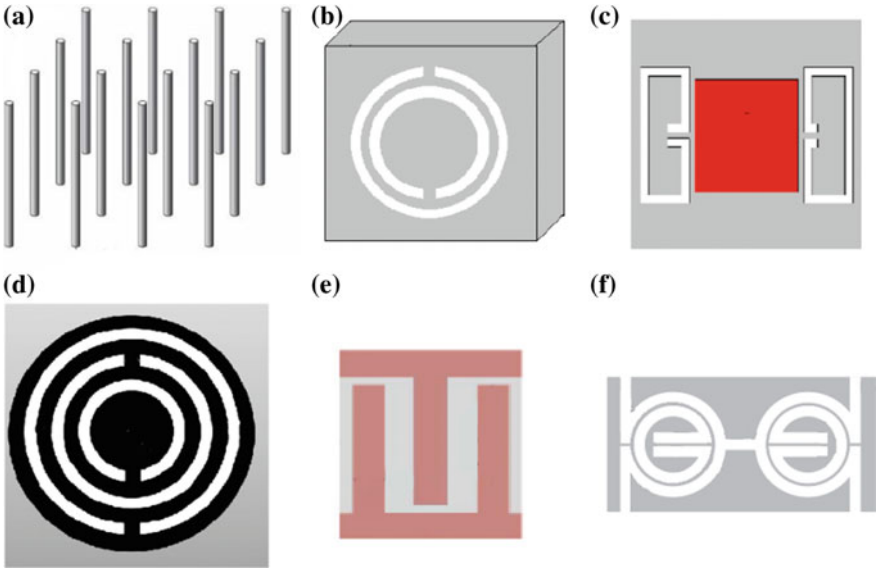

Fig. 2 Structures offering ε-negative behavior

proven to exhibit the negative permittivity. Now, CSRR structures are being widely used for several applications such as mutual coupling reduction, beam steering, and size reduction.

Figure 2c shows a double CSRR-loaded antenna, proposed by Cao et al., which was used for both positive and negative beam steerings [5]. To minimize antenna size and to improve bandwidth, a CSRR was presented by Ouedraogo et al. as shown in Fig. 2d, where SRR is placed horizontally between the bottom ground plane and the top patch [6]. Further, 55% patch size reduction was achieved by Ha et al. in [7] by using CSRR slot on the bottom ground plane along with an inter-digital capacitor on the top side, as shown in Fig. 2e. Later, a slot combined complementary split ring resonators (SC-SRR), shown in Fig. 2f, offering ε-negative behavior is also introduced by Qamar et al. [8]. This structure minimizes mutual coupling between two elements of antenna array. The SC-SRR structure is placed in middle of the two patches to reduce mutual coupling with small separation distance as compared to techniques employing EBG for coupling suppression. EBG structures require complicated fabrication process or large antenna separation distance, as observed by Rajo-Iglesias et al. [9].

Results discussed above are summarized in Table 1.

Table 1 Comparative analysis of various ε-negative structures

Reference	Methodology	Structure	Application and results	Operating frequency
[8]	Placing metamaterial superstrate between the array elements	Slot-combined complementary split ring resonators (SC-CSRR)	Mutual coupling suppression of 27 dB	3.6 GHz
[5]	Parameter extraction for the CSRR structure is investigated	CSRR-loaded ground structure	−51° to 48° beam steering is possible	1.96 GHz
[6]	Study effect on bandwidth due to area reduction	CSRR between patch and ground	Easy miniaturization at low cost due to simple geometry	2.45 GHz
[7]	Resonant frequency reduces on increasing series capacitance, thus reducing the electrical size	Interdigital capacitor and CSRR Slot	55% reduction in patch size. Efficiency = 96%, gain = 3.85 dBi	3.67–3.93 GHz

3 Mu-Negative Metamaterials ($\mu < 0$)

Mu-negative (MNG) material has permittivity greater than zero and permeability less than zero ($\varepsilon > 0$, $\mu < 0$). Some gyrotropic materials present these characteristic in certain frequency ranges. The first μ-negative metamaterial introduced was split ring resonator (SRR). Due to the presence of artificial magnetic dipole moments provided by ring resonator, SRR structure possesses a magnetic response even though it does not include magnetic conducting materials. Various SRR structures were presented in [10–15] and were used for different applications to enhance antenna properties.

A circular patch antenna loaded with μ-negative metamaterial, shown in Fig. 3, was presented by Bilotti et al. which support the expected resonant mode, despite the small size of the antenna [16]. Few years later, an elliptical patch with

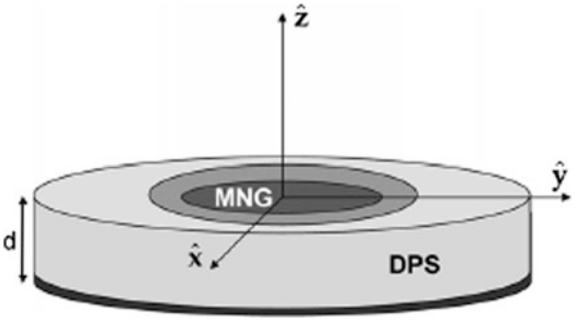

Fig. 3 Metamaterial-loaded circular patch antenna

metamaterial loading was introduced by Chen et al. for antenna miniaturization [17]. This structure provides better gain and aperture efficiency compared to circular patch, without compromising the radiated power.

Conventional planar SRRs have a drawback of narrow resonance band and require large dimension at low frequency. To resolve this problem, a compact nested tridmensional SRR was presented in high-permittivity low temperature co-fired ceramic (LTCC) medium by Liu et al. [11]. Figure 4a shows the tridimensional SRRs unit cell used for enhancement of gain. It provides an increase in gain of antenna to 6.93 dBi from 5.37 dBi compared to conventional antenna. An alternative approach was introduced by Bilotti et al. in [16] by employing composite multi-band resonators where bandwidth of MNG could be enhanced, with easier fabrication approach. The paper also gave a comparative analysis of different SRRs and was found that co-directional SRR (Fig. 4b) is the best choice to reduce mutual coupling while considering different inner-ring rotation angles. Another planar antenna was introduced by Farzami et al., which presents a row of nine SRRs (basic unit cell shown in Fig. 4c) kind of structure, embedded on the lower section of antenna [10]. This provides about one-third size reduction compared to conventional antenna with the same impedance bandwidth characteristics.

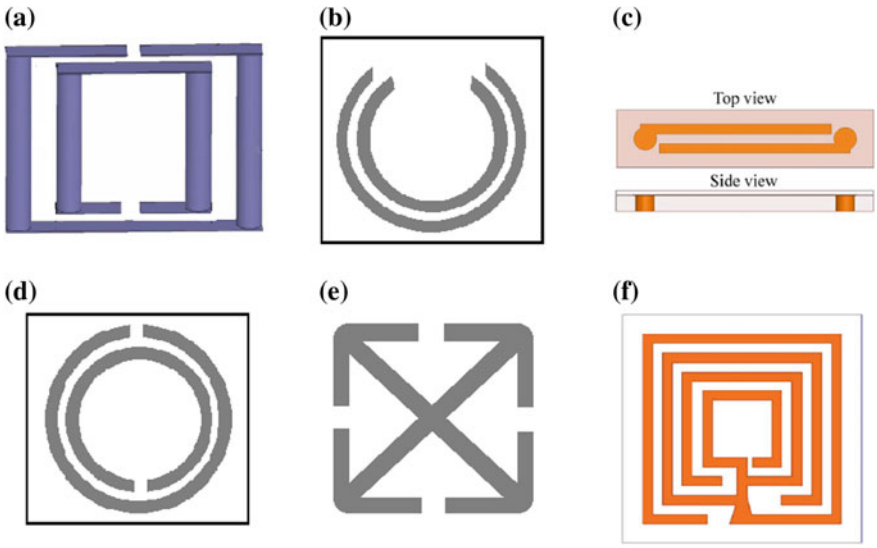

Fig. 4 Various available SRR structures

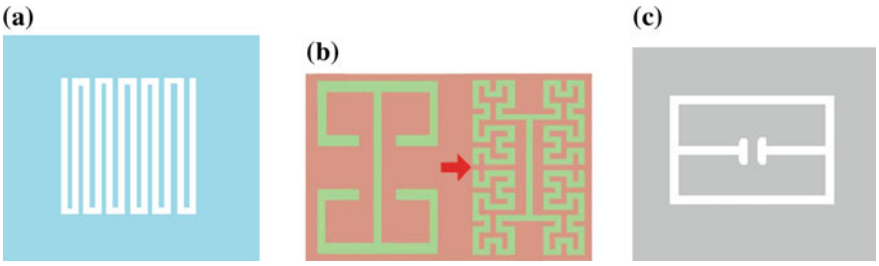

Fig. 5 Different CELC structures

SRR metamaterial also has filtering characteristics which can be used to reduce surface waves. Gupta et al. in [13] presented an SRR; as shown in Fig. 4d for band reject filter application as the number of SRRs increases, the rejection level in the stop band of the filter also increases. Another application of planar SRR structures is in THz frequencies [14]. Koutsoupidou et al. observed that since using simple dielectric substrate leads to shock waves at air–substrate interface, SRR metamaterials provide an effective way to direct the radiations from surface to free space. By using SRRs-CS (cross-shaped) substrate, as shown in Fig. 4e, the gain enhancement of 6.24 dB as compared to 5.15 dB for the single SRR substrate was obtained by Koutsoupidou et al. [14]. SRRs can also help to exhibit circular polarization property. A modified square split ring resonator (MSSRR) to show the circular polarization was presented by Lou et al. [15]. MSSRR unit cell (Fig. 4f) was composed of thin strips in which each of them produced negative refraction index.

Other than SRRs, the complementary electric inductive capacitive (CELC) structure, which is a dual counterpart of the electric inductive–capacitive resonator (ELC), has been known to deliver a resonant magnetic response, i.e., negative magnetic permeability in the proximity of resonance as demonstrated by Thomas et al. [18]. Tang et al. [19] presented the complementary meander line (CML), as shown in Fig. 5a, as a typical CELC structure. The structure has very low cross-polarization and quite high radiation efficiency. Another CELC structure was demonstrated by Xu et al. [20], which is a fractal-based Hilbert-shaped complementary electric inductive–capacitive resonator (H-CELC), as shown in Fig. 5b. It helps in reducing mutual coupling more than 9.7 dB. Also, a slot-loaded ELC on the ground plane (Fig. 5c) was presented by Bala et al. [21] to offer a peak realized gain of 2.63 dB and efficiency of 86%.

Table 2 gives a comparative analysis of various MNG structures discussed so far.

Table 2 Comparative analysis of various μ-negative structures

Reference	Methodology	Structure used	Application and results	Operating frequency (GHz)
[10]	A row of nine SRRs is embedded on lower section of antenna	SRRs (split ring resonators)	one-third size reduction maintaining the same impedance bandwidth	2.4 GHz
[11]	Tridimensional SRRs are embedded in LTCC substrate	Tridimensional SRRs	Antenna gain increases to 6.93 dBi from 5.37 dBi	5.2 GHz
[12]	Comparison of different SRR structures	Co-directional split ring resonator	Low mutual-coupling considering different inner-ring rotation angles	5.18 and 5.68 GHz
[13]	Increasing the number of SRRs increases the rejection level of the filter in stop band	Split ring resonator (SRR)	Reducing outer radius of SRR shifts central filtering frequency of band stop filter upward	–
[14]	Study of four different types of SRRs-based metamaterial	Cross-shaped split ring resonators (SRRs-CS)	Gain enhancement from 5.15 to 6.24 dB	–
[15]	Each thin strip of MSRR produces negative refractive index	Modified square split ring resonator (MSSRR)	Circular polarization is obtained in three bands	2.18–2.5, 3.21–3.76, and 4.1–7.89 GHz
[16]	Comparative analysis	Circular patch antennas loaded by MNG material	Radiation efficiencies: at 470 MHz = 0.67 at 2.44 GHz = 0.92, gains: at 470 MHz = 3.4 dBi at 2.44 GHz = 6.5 dBi	470 MHz–2.44 GHz
[17]	Closed form solution followed by full wave simulation	Elliptical patch antennas partially loaded with MNG	Elliptical patch provides better efficiency and gain compared to the circular geometry	0.5 GHz
[19]	Use of an ultralow-profile antenna	Complementary meander line (CML) structure	High radiation efficiency, low cross-polarization	2.68515 GHz
[20]	Hilbert-shaped complementary electric inductive–capacitive resonator (H-CELC)	Hilbert space-filling curves	Mutual coupling reduction of more than 9.7 dB	3.5 and 4.42 GHz

(continued)

Table 2 (continued)

Reference	Methodology	Structure used	Application and results	Operating frequency (GHz)
[21]	Use of slot-loaded ELC on the ground plane and excited by a microstrip line	Slot-loaded electric-LC resonator	Peak realized gain and efficiency of 2.63 dB and 86% are obtained	2.39–2.48 GHz

4 Double-Negative Metamaterials ($\varepsilon < 0$, $\mu < 0$)

Double-negative (DNG) indexed materials have both permittivity and permeability less than zero ($\varepsilon < 0$, $\mu < 0$). These DNG classes of materials are artificially tailored structures. These materials are not available in nature. The first experimental DNG structure of thin wires and SRRs was introduced by UCSD team (University of California, San Diego) [22], as shown in Fig. 6a.

Mutual coupling between antennas can be reduced by using DNG structures in an antenna array. The use of conventional metamaterial structures, i.e., thin wire and SRR, involves bulky formations or multilayer structures. Such structures are difficult to miniaturize and are less cost-efficient, so a wire-loaded spiral resonator (SR), shown in Fig. 6b which was introduced by Alqadami et al. [23], reduces mutual coupling in an efficient manner. Using this, the mutual coupling better than −28 dB was achieved. Another disadvantage of SRR-based metamaterial is that the structural rotation influences the properties of these unit cells. This drawback was overcome by cross-type structures presented by Wang et al. [24]. Further, Jerusalem Cross (JC) structures represented as a replacement of SRR by Katko [25]. Idea of JCs was further implemented through 3×3 array of criss-cross structure to enhance gain (achieved gain = 4.61 dB) and bandwidth (improved by 86.66% compared to conventional antenna) of antenna, as shown in Fig. 6c, which was presented by Inamdar et al. [26].

DNG structures are used for antenna miniaturization as well. Mahmoud presented miniaturization of antenna [27] by partially loading metamaterial ring into an annular ring patch resonator, as shown in Fig. 7a. It was demonstrated that the

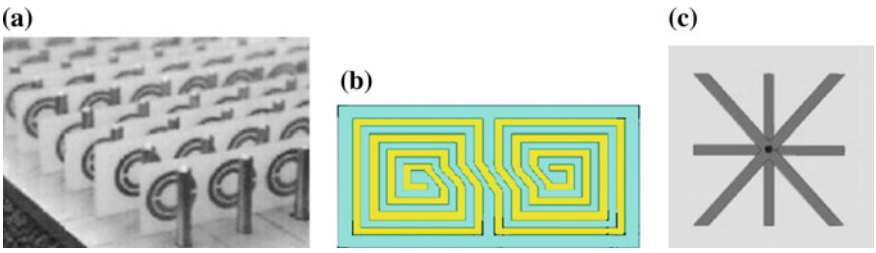

Fig. 6 SRR structures for reduction of mutual coupling

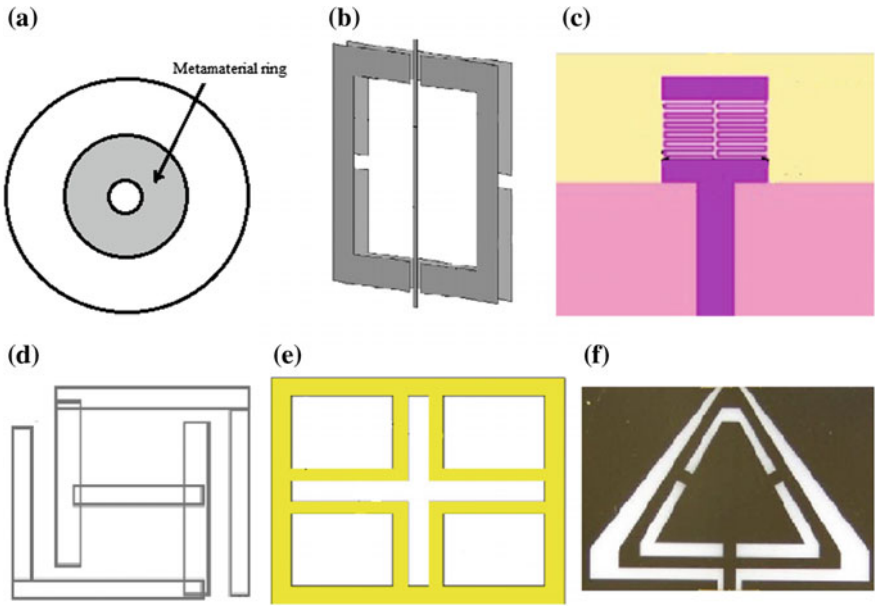

Fig. 7 Variants of DNG structures

resonant frequency can be made as small as desired by adjusting the width of metamaterial ring. To construct a DNG metamaterial, modified split ring resonators (MSRRs) and metal strips were used by Wu et al. [28], as shown in Fig. 7b. Physical dimension reduction from 0.5λ to 0.17λ and the bandwidth expansion to 17.56% were achieved through DNG filling.

Various other structures are also available for bandwidth enhancement [29–31]. The metamaterial unit cell made of meander lines in Fig. 7c showing both permittivity and permeability negative simultaneously was presented by Islam et al. [29], which provided compactness and improved bandwidth of 600 MHz with 23.81% fractional bandwidth. A superstrate of double H-shaped metamaterial structure was presented by Kaur et al. Figure 7d was also used for bandwidth enhancement [30] giving efficiency improvement from 57.3 to 80.2% and bandwidth improvement from 70 to 220 MHz. A fishnet-based metamaterial unit cell proposed by Sirmaci et al. [31], shown in Fig. 6e, is another DNG unit cell and was used for bandwidth improvement (8.2% bandwidth enhancement).

Most of the structures used either square or circular geometry. Use of triangular split ring resonator (Fig. 7f) is quite rare which was presented by Lucibello et al. [32], where the cantilever MEMS switches are kept along the arms of the triangle to make splits. When the switch is in ON state, it offers DNG behavior. Triangular structures are more compact compared to the circular and square geometries. These structures can be easily coupled side-by-side and have axial symmetry.

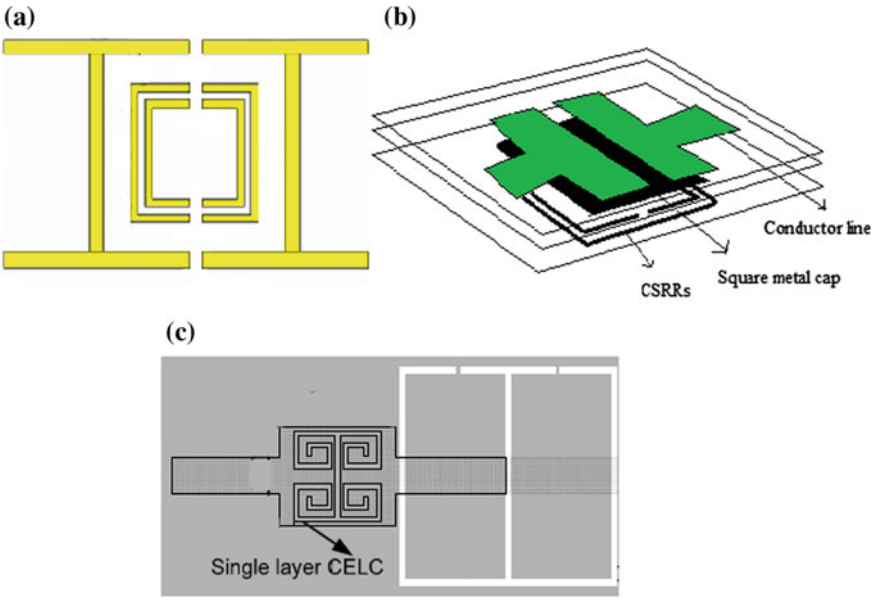

Fig. 8 DNG structures developed by combining SNG structures

Various DNG structures were developed by combining two structures having SNG behavior independently [33–35]. One MSRR between two pairs of capacitance-loaded strips (CLSs), shown in Fig. 8a, was demonstrated by Majid et al. [33]. Wherein, MSRR exhibits response like magnetic material and produces negative permeability, while the CLS exhibits response like strong dielectric and produces the negative permittivity. Similarly, Xu et al. [34] demonstrated the effect of combining the use of capacitive gap on the top substrate and CSRR on the bottom ground plane Fig. 8b. The double-layered CSRRs provide negative permittivity; similarly, in the longitudinal direction, the capacitive gap exhibits the negative permeability, producing a DNG structure. The structure provides continuous frequency-scanning capabilities from backward −29° to forward 72°. Liang et al. [35] developed a novel DNG structure by combining CSSRR, offering negative permittivity, and CELC (complementary electric inductive capacitive) cells offering negative permeability Fig. 8c. It provided band-pass filter characteristics with good upper band skirt performance.

Table 3 indicates the summary of above discussion.

Table 3 Comparative analysis of various double-negative structures

Reference	Methodology	Structure used	Application and results	Operating frequency
[23]	Unit cell arranged in 2 × 3 array inthe ground plane	Spiral resonator	Mutual coupling reduction (more than 28 dB)	515 GHz
[26]	3 × 3 array of criss-cross structure inspired from the famous Jerusalem Cross	Criss-cross structure	Bandwidth obtained = 560 MHz (improved by 86.66%), achieved gain = 4.61 dB	4.92–5.48 GHz
[27]	Annular ring patch resonator loaded by a metamaterial ring	Metamaterial ring	Adjusting the width of metamaterial ring shifts resonant frequency	–
[28]	DNG unit cells are stacked and embedded into a host substrate	Modified split ring resonators (MSRRs) and metal strips	Physical dimensions reduction from 0.5λ to 0.17λ and the bandwidth is enhanced to 17.56%	8.45–11.05 GHz
[29]	Double-negative TL metamaterials show negative permittivity and permeability simultaneously	Double-negative meander lines structure	Antenna provides 600 MHz bandwidth and gain of 3.72 dBi	2.51 GHz
[30]	A superstrate of double H-shaped metamaterial has been applied to H-shaped metamaterial-embedded RMPA	Double H-shaped metamaterial	Efficiency improvement from 57.3 to 80.2%, bandwidth improvement from 70 to 220 MHz	5.2 GHz
[31]	Use of three layers: rectangular microstrip patch antenna (RMPA), substrate, and fishnet-based MTM	Fishnet-based metamaterial unit cell	Gain is found to be 3.57 dB and bandwidth is 8.2%	1.05 and 1.1 THz
[32]	Use of MEMS for narrow-band microwave switching	Triangular split ring resonator	Triangular structure offers compactness and axial symmetry	16.38 GHz
[33]	Parametric studies	MSRR and the capacitance loaded strip (CLS)	Size increment of the LHM will shift the resonance frequency to a lower region	–
[34]	Using extra metal cap increases left-handed capacitor by 36%	CSRRs, capacitive gaps, and metal caps	Continuous frequency-scanning capabilities of the antenna from backward −29° to forward 72°	4.39–6.5 GHz

(continued)

Table 3 (continued)

Reference	Methodology	Structure used	Application and results	Operating frequency
[35]	Use of three CELC cells giving wideband stop band performance	Cascaded complementary single split ring resonator (CCSSRR) with a CELC	Total suppression is lower than 14 dB	2.28–2.70 GHz

5 Electromagnetic Bandgap Structures

Electromagnetic bandgap (EBG) structures are periodic arrangements of dielectric or metallic inclusions or sometimes combination of both. These structures are also called photonic bandgap structures (PBG). The periodicity can be in one, two, or three dimensions. These periodic arrangements of structures give rise to stop band, in which frequencies are strongly reflected, and passband, in which frequencies are propagated without attenuation, along with the direction of periodicity in electromagnetic transmission spectrum. This property of EBG structure displays the forbidden behavior of electromagnetic waves in certain frequency region and propagation in another region. These forbidden frequency regions are referred to as forbidden bands or more commonly, as band gaps. According to the geometrical configurations, these structures can be categorized into three groups: (i) one-dimensional transmission lines, (ii) two-dimensional planar surfaces, and (iii) three-dimensional volumetric structures. Two-dimensional planar EBG surfaces are further classified into two categories: uniplanar EBG surfaces and mushroom-like EBG surfaces. The EBG structures are utilized in microwave, millimeter wave, and infrared devices such as antennas, planar reflectors, filters, resonators, integrated circuits, and waveguides.

Various EBG structures used in patch antennas for different applications are discussed and reported in literature [36–40]. A 2-D mushroom such as electromagnetic bandgap structure (Fig. 9a) was introduced by Xu et al. for high directivity, resulting in a gain of 21.6 dB, increased by 12.6 dB compared to conventional antenna [36]. Another mushroom-type EBG structure (Fig. 9b) was presented by Liang et al. [37] and its performance as a tunable antenna was investigated by loading diode switches in between wires and the ground plane. An ON state gain of 2.2 dBi and OFF state gain of 4.4 dBi was observed. Further, to improve the radiation efficiency, slots were added on the ground plane of a conventional hexagonal mushroom such as electromagnetic band gap structure by Pyo et al. [38], as shown in Fig. 9c. By increasing the slot length, efficiency improvement from 24.1 to 38.8% was observed.

EBG structures are also used for multi-band applications [39, 40]. A modification in mushroom-type electromagnetic bandgap (EBG) structure, shown in Fig. 9d, was presented by Cao et al. [39]. For the analysis of these

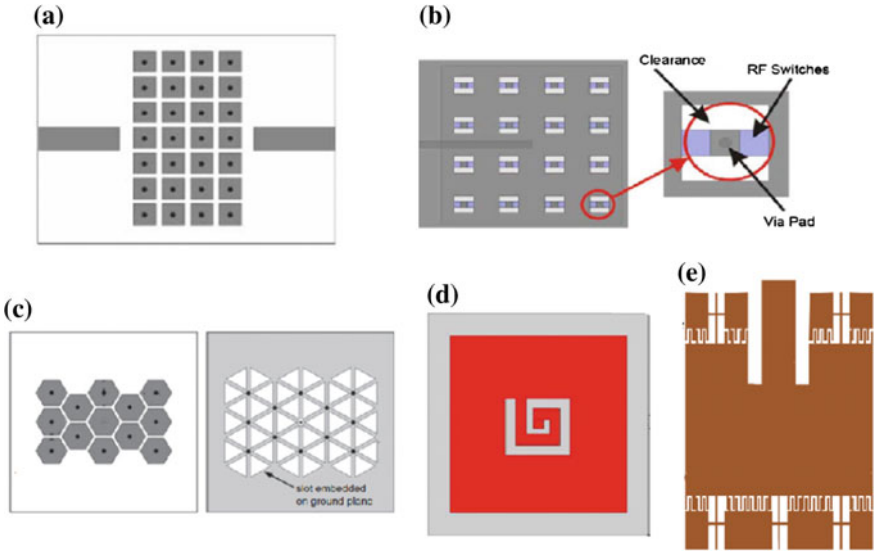

Fig. 9 EBG structures

Table 4 Comparative analysis of various EBG structures

Reference	Methodology	Structure used	Application and results	Operating frequency
[36]	Use of metamaterial superstrate and an EBG substrate	Mushroom-like electromagnetic bandgap (EBG) structure	Gain is 21.6 dB, increased by 12.6 dB compared to conventional antenna	14.6 GHz
[37]	Loading diode switches in between vias and the ground plane	Mushroom-type Electromagnetic bandgap (EBG) structure	On state: gain—2.2 dBi, efficiency—38.7%, Off state: gain—4.4 dBi, efficiency—55%	On state: 1.98 GHz, Off state: 2.45 GHz
[38]	Increase slot length for efficiency improvement	Periodic EBG cells on the top plane and triangular slots on the ground plane	Efficiency improvement from 24.1 to 38.8%	3.605, 4.715, 5.280 GHz
[39]	CRLH-TL theory and multiconductor transmission line (MTL) theory	Modified mushroom-type EBG structure	Dual polarization exists	1.34, 1.91 and 2.38 GHz
[40]	Microstrip-fed patch with MTM–EBG sections placed on the radiating edges	Use of MTM–EBG structure	One operating frequency in each of the band gap and passband regions	2.4 and 5.0 GHz

structures, composite right-/left-handed transmission line (CRLH-TL) and multi-conductor transmission line (MTL) theory were used. It results in a dual-mode and multi-frequency antenna. Another dual-band patch antenna was presented by Smyth et al. [40] that employed metamaterial–EBG integrated into its radiating edge Fig. 9e in which the lower frequency was determined by the combination of cavity and dispersive nature of EBG, whereas high frequency was determined by the size of cavity without EBG.

Table 4 gives a comparative analysis of various EBG structures discussed so far.

6 Conclusion

The study of various articles is summarized in this paper. Use of metamaterial in antenna provides advantages of size reduction, bandwidth enhancement, mutual coupling reduction, gain, and directivity improvement and provides self-tuning capability. In-depth literature has been analyzed to classify them based on the basic metamaterial structure. An application-based study is carried out specifying the suitable structure for an application. Through the study of this paper, user can point out the best configuration required while designing the antenna that suits the desired specification.

Glossary

CELC Complementary electric inductive–capacitive

CLSs Capacitance-loaded strips

CML Complementary meander line

CRLH-TL Composite right-/left-handed transmission line

CSRRs Complementary split ring resonators

DNG Double negative

DPS Double positive materials

EBG Electromagnetic bandgap

ELC Electric inductive–capacitive resonator

ENG Epsilon negative

H-CELC Hilbert-shaped complementary electric inductive–capacitive resonator

LHM Left-handed materials

LTCC Low temperature co-fired ceramic

MNG Mu-negative

MSRRs Modified split ring resonators

MSSRR Modified square split ring resonator

MTL Multi-conductor transmission line

PBG Photonic band gap structures

SC-SRR Combined complementary split ring resonators

SNG Single negative

SR Spiral resonators

SRR Split ring resonator

SRRs-CS Split ring resonators—cross-shaped

UCSD University of California, San Diego

References

1. I.A. Buriak, V.O. Zhurba, G.S. Vorobjov, V.R. Kulizhko, O.K. Kononov, O. Rybalko, "Metamaterials: theory, classification and application strategies (Review). J. Nano Electron. Phys. (2016)
2. N. Engheta, R.W. Ziolkowski, *Metamaterials: Physics and Engineering Explorations* (IEEE-Wiley, New York, 2006)
3. J.B. Pendry, A.J. Holden, D.J. Robbins, W.J. Stewart, Magnetism from conductors and enhanced nonlinear phenomena. IEEE Trans. Microw. Theory Tech. **47**(11), 2075–2084 (1999)
4. F. Falcone, T. Lopetegi, J.D. Baena, R. Marques, F. Martin, M. Sorolla, Effective negative epsilon stopband microstrip lines based on complementary split ring resonators. IEEE Microw. Wirel. Compon. Lett. **14**(14), 280–282 (2004)
5. W. Cao, Y. Xiang, B. Zhang, A. Liu, T. Yu, D. Guo, A low-cost compact patch antenna with beam steering based on CSRR-loaded ground. IEEE Antennas Wirel. Propag. Lett. **10**, 1520–1523 (2011)
6. R.O. Ouedraogo, E.J. Rothwell, A.R. Diaz, K. Fuchi, A. Temme, Miniaturization of patch antennas using a metamaterial-inspired technique. IEEE Trans. Antennas Propag. **60**(5), 2175–2182 (2012)
7. J. Ha, K. Kwon, Y. Lee, J. Choi, Hybrid mode wideband patch antenna loaded with a planar metamaterial unit cell. IEEE Trans. Antennas Propag. **60**(2), 1143–1147 (2012)
8. Z. Qamar, U. Naeem, S.A. Khan, M. Chongcheawchamnan, M.F. Shafique, Mutual coupling reduction for high-performance densely packed patch antenna arrays on finite substrate. IEEE Trans. Antennas Propag. **64**(5), 1653–1660 (2016)
9. E. Rajo-Iglesias, O. Quevedo-Teruel, L. Inclan-Sanchez, Mutual coupling reduction in patch antenna arrays by using a planar EBG structure and a multilayer dielectric substrate. IEEE Trans. Antennas Propag. **56**(6), 1648–1655 (2008)
10. F. Farzami, K. Forooraghi, M. Norooziarab, Miniaturization of a microstrip antenna using a compact and thin magneto-dielectric substrate. IEEE Antennas Wirel. Propag. Lett. **10**, 1540–1542 (2011)

11. Z. Liu, P. Wang, Z. Zeng, Enhancement of the gain for microstrip antennas using negative permeability metamaterial on low temperature co-fired ceramic (LTCC) substrate. IEEE Antennas Wirel. Propag. Lett. **12**, 429–432 (2013)
12. M.-C. Tang, S. Xiao, T. Deng, Y. Wang, Y. Bai, C. Liu, Y. Shang, J. Xiong, B. Wang, Design of a broadband μ-negative planar material with low frequency dispersion. Appl. Phys. A **106**, 821–828 (2012)
13. M. Gupta, J. Saxena, Microstrip filter designing by SRR metamaterial. Wirel. Pers. Commun. **71**, 3011–3022 (2013)
14. M. Koutsoupidou, I.S. Karanasiou, N. Uzunoglu, Substrate constructed by an array of split ring resonators for a THz planar antenna. J. Comput. Electron. **13**, 593–598 (2014)
15. R.K.M. Lou, M. Naser-Moghadasi, R.A. Sadeghzadeh, Compact multi-band circularly polarized CPW fed antenna based on metamaterial resonator. Wirel. Pers. Commun. (2016)
16. F. Bilotti, A. Alú, L. Vegni, Design of miniaturized metamaterial patch antennas with μ-negative loading. IEEE Trans. Antennas Propag. **56**(6), 1640–1647 (2008)
17. P.Y. Chen, A. Alu, Sub-wavelength elliptical patch antenna loaded with μ-negative metamaterials. IEEE Trans. Antennas Propag. **58**(9), 2909–2919 (2010)
18. H.H. Thomas, J. Gollub, S. Sajuyigbe, D.R. Smith, S.A. Cummer, Characterization of complementary electric field coupled resonate surface. Appl. Phys. Lett. **93** (2008)
19. M.-C. Tang, S. Xiao, Y.-Y. Bai, T. Deng, C. Liu, Y. Shang, C. Wei, B.-Z. Wang, Design of hybrid patch/slot antenna operating in induced TM120 mode. IEEE Trans. Antennas Propag. **60**(5), 2157–2165 (2012)
20. H.-X. Xu, G.-M. Wang, M.-Q. Qi, Hilbert-shaped magnetic waveguided metamaterials for electromagnetic coupling reduction of microstrip antenna array. IEEE Trans. Magn. **49**(4), 1526–1529 (2013)
21. B.D. Bala, M. Kamal, A. Rahim, N.A. Murad, Complementary electric-LC resonator antenna for WLAN applications. Appl. Phys. A **117**, 635–639 (2014)
22. R.A. Shelby, D.R. Smith, S. Schultz, Experimental verification of a negative index of refraction. Science **292**, 77–79 (2001)
23. A.S.M. Alqadami, M.F. Jamlos, P.J. Soh, S.K.A. Rahim, A. Narbudowicz, Left-handed compact MIMO antenna array based on wire spiral resonator for 5-GHz wireless applications. Appl. Phys. A **123** (2017)
24. J. Wang, S. Qu, H. Ma, S. Xia, Y. Yang, L. Lu, X. Wu, Z. Xu, Q. Wang, Experimental verification of anisotropic three-dimensional left-handed metamaterial composed of Jerusalem Crosses. PIERS Online **6**(1) (2010)
25. A.R. Katko, Artificial negative permeability based on a fractal Jerusalem Cross. Undergraduate Honors Thesis (The Ohio State University, 2009)
26. K. Inamdar, Y.P. Kosta, S. Patnaik, Criss-Cross metamaterial-substrate microstrip antenna with enhanced gain and bandwidth. Radioelectron. Commun. Syst. **58**(2), 69–74 (2015)
27. S.F. Mahmoud, A new miniaturized annular ring patch resonator partially loaded by a metamaterial ring with negative permeability and permittivity. IEEE Antennas Wirel. Propag. Lett. **3**, 19–22 (2004)
28. M.F. Wu, F.Y. Meng, Q. Wu, J. Wu, J.C. Lee, An approach for small omnidirectional microstrip antennas based on the backward waves of double negative metamaterials. Appl. Phys. A **87**, 193–198 (2007)
29. M.M. Islam, M.R.I. Faruque, M.T. Islam, M.F. Mansor, Compact and broadband antenna using double-negative transmission line metamaterial. Appl. Phys. A **123** (2017)
30. P. Kaur, S.K. Aggarwal, A. De, Performance enhancement of rectangular microstrip patch antenna using double H shaped metamaterial. Radioelectron. Commun. Syst. **59**(11), 496–501 (2016)
31. Y.D. Sirmaci, C.K. Akin, C. Sabah, Fishnet based metamaterial loaded THz patch antenna. Opt. Quantum Electron. **48** (2016)
32. A. Lucibello, E. Proietti, R. Marcelli, G.M. Sardi, G. Bartolucci, Narrow-band filtering by means of triangular metamaterial resonators based on RF MEMS cantilevers in CPW configuration. Microsyst. Technol. (2016)

33. H.A. Majid, M.K.A. Rahim, Parametric studies on left-handed metamaterial consist of modified split-ring resonator and capacitance loaded strip. Appl. Phys. A **103**, 607–610 (2011)
34. H.-X. Xu, G.-M. Wang, M.-Q. Qi, A leaky-wave antenna using double-layered metamaterial transmission line. Appl. Phys. A **111**, 549–555 (2013)
35. J.-G. Liang, H.-X. Xu, Harmonic suppressed bandpass filter using composite right/left handed transmission line. J. Zhejiang Univ.-Sci. C (Comput. Electron.) 552–228 (2012)
36. H. Xu, Z. Zhao, Y. Lv, C. Du, X. Luo, Metamaterial superstrate and electromagnetic band-gap substrate for high directive antenna. Int. J. Infrared Millimeter Waves **29**, 493–498 (2008)
37. J. Liang, H.Y.D. Yang, Microstrip patch antennas on tunable electromagnetic band-gap substrates. IEEE Trans. Antennas Propag. **57**(6), 1612–1617 (2009)
38. S. Pyo, J.-W. Baik, S.-H. Cho, Y.-S. Kim, Metamaterial-based antenna with triangular slotted ground for efficiency improvement. Electron. Lett. **45**(3) (2009)
39. W. Cao, B. Zhang, A. Liu, T. Yu, D. Guo, X. Pan, Multi-frequency and dual-mode patch antenna based on electromagnetic band-gap (EBG) structure. IEEE Trans. Antennas Propag. **60**(12), 6007–6012 (2012)
40. B.P. Smyth, S. Barth, A.K. Iyer, Dual-band microstrip patch antenna using integrated uniplanar metamaterial-based EBGs. IEEE Trans. Antennas Propag. **64**(12), 5046–5053 (2016)
41. X. Cheng, D.E. Senior, C. Kim, Y.-K. Yoon, A compact omnidirectional self-packaged patch antenna with complementary split-ring resonator loading for wireless endoscope applications. IEEE Antennas Wirel. Propag. Lett. **10**, 1532–1535 (2011)

Realization of Recursive Algorithm for One-Dimensional Discrete Cosine Transform and Its Inverse

Pragati Dahiya and Priyanka Jain

Abstract In this paper, recursive algorithm for the computation of discrete cosine transform (DCT) and inverse discrete cosine transform (IDCT) has been proposed. This algorithm has been implemented using recursive structure, i.e., infinite impulse response (IIR) filter. Further, any desired output kernel can be computed independently which is an added attraction of the proposed design. Besides being recursive, this algorithm requires less multiplier and adders than other DCT algorithms. The suggested scheme is time efficient as it provides output in $N/2$ computational cycles, where N is the length of the input sequence. Therefore, the proposed algorithm is both time and hardware efficient. Further, the recursive property of proposed algorithm results in a regular and modular structure which results in easier VLSI implementation.

Keywords Discrete Cosine Transform (DCT) · Recursive structures
Infinite impulse response filter (IIR)

1 Introduction

Real and orthogonal discrete cosine transform (DCT) [1] and discrete sine transform (DST) [2] are members of sinusoidal unitary transforms [3] having numerous applications in field of digital signal processing and image processing especially in data compression, digital filtering, feature extractions, and recognition. Modified form of DCT and DST are used in MPEG-1 and MPEG-2 (moving picture expert group) for audio coding. DCT and DST are used in standard JPEG for image compression/decompression [4]. When transform of highly correlated signal in time

P. Dahiya (✉) · P. Jain
Department of Electronics and Communication Engineering,
Delhi Technological University, Bawana Road, Delhi, India
e-mail: pragatidahiya@gmail.com

P. Jain
e-mail: priyajain2000@rediffmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_7

domain like speech signal is computed, it gets decorrelated in frequency domain such that most of the transform coefficients are near to zero value and maximum information gets stored in few transform coefficients which can be encoded with less number of bits as compared to original signal. DCT-II has shown performance identical to Karhunen–Loeve transform (KLT) which is an optimal transform in terms of decorrelation and energy compaction based on value of the correlation coefficient [5]. For highly correlated signal, DCT operates similar to KLT, whereas for low correlated signal DST operates similar to KLT [6]. The basis vectors of DCT and DST are independent of input signal statistics; hence, development of fast algorithm for these transform is an easy task as compared to KLT whose basis vector gets altered with the change in input. Different authors have used different methods to simplify the DCT computation which can be broadly classified as (1) indirect approach [7] in which other transforms like FFT, Walsh–Hadamard transforms are used to compute DCT of a signal, (2) direct factorization of transform matrix using sparse matrix factorization method [8], (3) recursive approach [9–12] in which lower-order transform coefficient values are used to compute higher-order transform coefficient values. This paper is aimed to reduce the multiplication and additions required for DCT and inverse discrete cosine transform (IDCT) computation. The proposed algorithm provides improved results in terms of number of mathematical operation and time required to compute the transform coefficients as compare to [13–18], and also, the proposed recursive structures show saving in terms of hardware.

2 Recursive Algorithm for Computation of DCT-II

DCT-II for input sequence $x(n)$ for $n = 0, 1, \ldots, (N-1)$ where N has even value is given by following equation:

$$X(k) = A_k \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\left[\frac{(2n+1)k\pi}{2N}\right], \quad k = 0, 1, \ldots, N-1 \tag{1a}$$

where

$$A_k = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & k = 1, 2 \ldots, N-1 \end{cases} \tag{1b}$$

Omitting the scalar multiplier $A_k \sqrt{\frac{2}{N}}$ from (1a) as it is constant so not included in derivation procedure.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cos[(2n+1)k\pi/2\mathrm{N}], \quad k = 0, 1, \ldots, N-1 \tag{2}$$

2.1 Input Folding Recursive (IFR) Transform Algorithm

Equation (2) can be expressed as (3a, 3b) by folding input sequence:

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} w_k(n) \cos[(2n+1)k\pi/2\mathrm{N}], \quad k = 0, 1, \ldots, N-1 \tag{3a}$$

where

$$w_k(n) = \left\{ x(n) + (-1)^k x(N-1-n) \right\} \tag{3b}$$

Replacing n by $(N/2 - n)$ in Eq. (3a), we obtain

$$X(k) = \sum_{n=1}^{N/2} w_k\left(\frac{N}{2} - n\right) [\cos\left(\frac{k\pi}{2}\right) \cos[(2n-1)k\pi/2N] + \sin\left(\frac{k\pi}{2}\right) \sin\left[\frac{(2n-1)k\pi}{2N}\right], \quad k = 0, 1, \ldots, N-1 \tag{4}$$

In next subsection, Kernels for even and odd values of k are developed as follows.

2.1.1 IFR Kernel for k Even

When k is even, Eq. (4) can be written as

$$X_1(k) = (-1)^{k/2} \sum_{n=1}^{N/2} w_k\left(\frac{N}{2} - n\right) \cos[(2n-1)k\pi/2N], \quad k = 0, 2, \ldots, N/2 \tag{5}$$

Let

$$P_n^k = \frac{\cos[(n-1/2)\varnothing_k]}{\cos(\varnothing_k/2)}, \quad \varnothing_k = \frac{k\pi}{\mathrm{N}}, \quad k \text{ even} \tag{6a}$$

$$P_{n+1}^k = \frac{\cos[(n+1/2)\varnothing_k]}{\cos(\varnothing_k/2)}, \quad k \text{ even} \tag{6b}$$

$$X_1(k) = (-1)^{k/2}\cos\left(\frac{\varnothing_k}{2}\right)\sum_{n=1}^{N/2} w_k\left(\tfrac{N}{2}-n\right)P_n^k, \quad k \text{ even} \tag{7}$$

Equation (6b) can be rewritten using the following trigonometric identity

$$\cos(t\varnothing_k) = 2\cos[(t-1)\varnothing_k]\cos\ \varnothing_k - \cos[(t-2)\varnothing_k] \tag{8}$$

$$\sin(t\,\varnothing_k) = 2\sin[(t-1)\varnothing_k]\cos\ \varnothing_k - \sin[(t-2)\varnothing_k] \tag{9}$$

$$P_{n+1}^k = \frac{2\cos[(n-1/2)\varnothing_k]\cos\ \varnothing_k - \cos[(n-3/2)\varnothing_k]}{\cos(\varnothing_k/2)}, \quad k \text{ even} \tag{10}$$

$$P_{n+1}^k = 2P_n^k\cos\ \varnothing_k - P_{n-1}^k, \quad k \text{ even} \tag{11}$$

Assume

$$S_{\frac{N}{2}}^k = \sum_{n=1}^{\frac{N}{2}} w_k\left(\tfrac{N}{2}-n\right)P_n^k, \quad k \text{ even} \tag{12}$$

Above equation can be rewritten by substituting value $P_1^k = 1$

$$S_{N/2}^k = w_k\left(\tfrac{N}{2}-1\right) + \sum_{n=1}^{\frac{N}{2}-1} w_k\left(\tfrac{N}{2}-n-1\right)P_{n+1}^k, \quad k \text{ even} \tag{13}$$

Using recursive relation Eq. (11), above equation can be expressed as:

$$S_{N/2}^k = w_k\left(\tfrac{N}{2}-1\right) + \sum_{n=1}^{\frac{N}{2}-1} w_k\left(\tfrac{N}{2}-n-1\right)\left[2P_n^k\cos\varnothing_k - P_{n-1}^k\right], \quad k \text{ even} \tag{14}$$

Or

$$S_{N/2}^k = w_k\left(\frac{N}{2}-1\right) - w_k\left(\frac{N}{2}-2\right) + 2\cos\varnothing_k\sum_{n=1}^{\frac{N}{2}-1} w_k\left(\frac{N}{2}-n-1\right)P_n^k - \sum_{n=2}^{\frac{N}{2}-1} w_k\left(\frac{N}{2}-n-1\right)P_{n-1}^k, \quad k \text{ even} \tag{15}$$

Using above equation, we get following recursive relation:

$$S_{N/2}^k = w_k\left(\frac{N}{2}-1\right) - w_k\left(\frac{N}{2}-2\right) + 2\cos\ \varnothing_k S_{\frac{N}{2}-1}^k - S_{\frac{N}{2}-2}^k, \quad k \text{ even} \tag{16}$$

We get Eq. (17) used to compute the even coefficient of DCT-II by substituting (12) in (7)

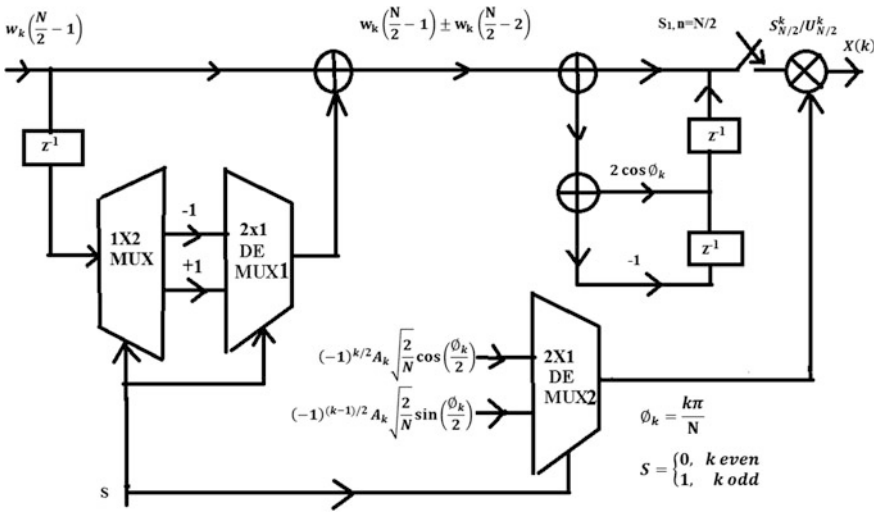

Fig. 1 Filter structure for calculating N-point DCT-II

$$X_1(k) = (-1)^{k/2} A_k \sqrt{\frac{2}{N}} \cos\left(\frac{\varnothing_k}{2}\right) S^k_{\frac{N}{2}}, \quad k = 0, 2, \ldots, N/2 \tag{17}$$

$X_1(k)$ can be realized using infinite impulse response (IIR)-filters structure as shown in Fig. 1. Thus, $S^k_{N/2}$ is available at $n = \frac{N}{2}$. For a specific value $k = k_1$, say, the computation of DCT coefficient $X(k_1)$ takes $\frac{N}{2}$ clock cycles. In other word, the full $X(k)$ for even kernel is computed in $(N/2)$. $(N/2)$ clock cycles.

2.1.2 IFR Kernel, *k* Odd

Let

$$Q^k_n = \frac{\sin\left[\left(n-\frac{1}{2}\right)\varnothing_k\right]}{\sin\left(\frac{\varnothing_k}{2}\right)}, \quad k \text{ odd} \tag{18}$$

$$Q^k_{n+1} = \frac{2\sin\left[\left(n+\frac{1}{2}\right)\varnothing_k\right]}{\sin\left(\frac{\varnothing_k}{2}\right)}, \quad k \text{ odd} \tag{19}$$

When k is odd, Eq. (4) can be written as

$$X_2(k) = (-1)^{(k-1)/2} \sin(\varnothing_k/2) \sum_{n=1}^{N/2} w_k\left(\tfrac{N}{2} - n\right) Q^k_n, \quad k = 1, 3, \ldots, N \tag{20}$$

By algebraic manipulation and by using Eq. (9), Eq. (19) can be expressed in following recursive form:

$$Q_{n+1}^{k} = 2Q_{n}^{k}\cos\varnothing_{k} - Q_{n-1}^{k}, \quad k \text{ odd} \tag{21}$$

Assume

$$U_{N/2}^{k} = \sum_{n=1}^{N/2} w_{k}\left(\tfrac{N}{2} - n\right)Q_{n}^{k}, \quad k \text{ odd} \tag{22}$$

Above equation can be rewritten by substituting value of $Q_{1}^{k} = 1$ [obtained from Eq. (18)]

$$U_{N/2}^{k} = w_{k}\left(\tfrac{N}{2} - 1\right) + \sum_{n=2}^{N/2} w_{k}\left(\tfrac{N}{2} - n\right)Q_{n}^{k}, \quad k \text{ odd} \tag{23}$$

$$U_{N/2}^{k} = w_{k}\left(\tfrac{N}{2} - 1\right) + \sum_{n=1}^{\frac{N}{2}} w_{k}\left(\tfrac{N}{2} - n - 1\right)Q_{n+1}^{k}, \quad k \text{ odd} \tag{24}$$

Using recursive relation obtained from Eq. (21), above equation can be expressed as:

$$U_{N/2}^{k} = w_{k}\left(\tfrac{N}{2} - 1\right) + \sum_{n=1}^{\frac{N}{2}-1} w_{k}\left(\tfrac{N}{2} - n - 1\right)\left[2Q_{n}^{k}\cos\varnothing_{k} - Q_{n-1}^{k}\right], \quad k \text{ odd} \tag{25}$$

Or

$$U_{N/2}^{k} = w_{k}\left(\tfrac{N}{2} - 1\right) + w_{k}\left(\tfrac{N}{2} - 2\right) + 2\cos\varnothing_{k}U_{\frac{N}{2}-1}^{k} - U_{\frac{N}{2}-2}^{k}, \quad k \text{ odd} \tag{26}$$

Using Eqs. (20), (22), and (26), we get odd coefficient of DCT-II

$$X_{2}(k) = (-1)^{(k-1)/2}A_{k}\sqrt{\tfrac{2}{N}}\sin\left(\tfrac{\varnothing_{k}}{2}\right)U_{\frac{N}{2}}^{k}, \quad k \text{ odd} \tag{27}$$

Using above equation, recursive filter structure is realized as shown in Fig. 1. When k is odd, computation time for $X(k)$ is again the same (i.e., $\left(\tfrac{N}{2}\right).\left(\tfrac{N}{2}\right)$) as for even case. Equations (16) and (26) show that the equations are same in the feedback path so the feedback structures are same for both the cases. These equations are different in feed-forward path; i.e., it is implemented by a control unit (MUX, DEMUX1 and DEMUX2) and select line as represented in Fig. 1. This control unit is used to select the inputs for the recursive structure to get any desired output coefficient. The selection is done through a mode selector switch, S, mode 0 provides even coefficients of the output, and mode 1 gives the odd coefficients of

k. Therefore, the proposed structure can compute any desired output coefficient independently.

2.2 Recursive Output Folding Algorithms (OFR) for Inverse DCT

IDCT-II of sequence $X(k)$for $k = 0, 1, 2, \ldots, (N-1)$ where N has even value is given by following equation:

$$x(n) = A_k \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} X(k) \cos[(2n+1)k\pi/2\mathrm{N}], \quad n = 0, 1, \ldots, N-1 \tag{28}$$

Following Eqs. (29) and (30) are obtained by folding the data from above equation in which value of $x(n) + x(N-1-n)$ is obtained for even values of k and of $x(n) - x(N-1-n)$ is obtained for odd values of k.

$$x(n) + x(N-1-n) = \sum_{k=0}^{N-1} X(k)\left\{1 + (-1)^k\right\} \cos\left[\frac{(2n+1)k\pi}{2\mathrm{N}}\right], \quad n = 0, 1, \ldots, \tfrac{N}{2} - 1 \tag{29}$$

$$x(n) - x(N-1-n) = \sum_{k=0}^{N-1} X(k)\left\{1 - (-1)^k\right\} \cos\left[\frac{(2n+1)k\pi}{2\mathrm{N}}\right], \quad n = 0, 1, \ldots, \tfrac{N}{2} - 1 \tag{30}$$

We now proceed to obtain the folded output (LHS) of Eqs. (29) and (30) for the even and odd values of k.

2.2.1 OFR Kernel for the Even Index Sequence $X(k)$

When k is even, Eq. (29) may be written as

$$x(n) + x(N-1-n) = 2 \sum_{k=0}^{\frac{N}{2}-1} X(2k) \cos\left[\frac{(2n+1)k\pi}{\mathrm{N}}\right], \quad n = 0, 1, 2, \ldots, \frac{N}{2} - 1 \tag{31}$$

Replacing k by $(N/2 - k)$

$$x(n) + x(N-1-n) = 2(-1)^n \sum_{k=1}^{\frac{N}{2}} X(N-2k)\sin(k\beta_n), \quad n = 0, 1, \ldots, \frac{N}{2} - 1 \tag{32a}$$

where

$$\beta_n = \frac{(2n+1)\pi}{N} \tag{32b}$$

Equation (32a) can also be written as

$$x(n) + x(N-1-n) = 2(-1)^n \sin(\beta_n) \sum_{k=1}^{\frac{N}{2}} X(N-2k) D_k^n, \quad n = 0, 1, 2, \ldots, \frac{N}{2} - 1 \tag{33a}$$

where

$$D_k^n = \frac{\sin(k\beta_n)}{\sin(\beta_n)}, \quad n = 0, 1, 2, \ldots, \tfrac{N}{2} - 1 \tag{33b}$$

Equation (33b) can be mathematically manipulated using Eq. (9) as

$$D_{k+1}^n = 2cos\,\beta_n D_k^n - D_{k-1}^n, \quad n = 0, 1, 2, \ldots, \tfrac{N}{2} - 1 \tag{34}$$

Assume

$$L_{N/2}^n = \sum_{k=1}^{\frac{N}{2}} X(N-2k) D_k^n, \quad n = 0, 1, 2, \ldots, \tfrac{N}{2} - 1 \tag{35}$$

Above equation can be rewritten by substituting value of $D_1^n = 1$ obtained from Eq. (33b)

$$L_{N/2}^n = X(N-2) + \sum_{k=1}^{\frac{N}{2}} X(N-2k-2) D_{k+1}^n, \quad n = 0, 1, 2, \ldots, \tfrac{N}{2} - 1 \tag{36}$$

Or

$$L_{N/2}^n = X(N-2) + 2\cos\,\beta_n L_{\frac{N}{2}-1}^n - L_{\frac{N}{2}-2}^n, \quad n = 0, 1, 2, \ldots, \tfrac{N}{2} - 1 \tag{37}$$

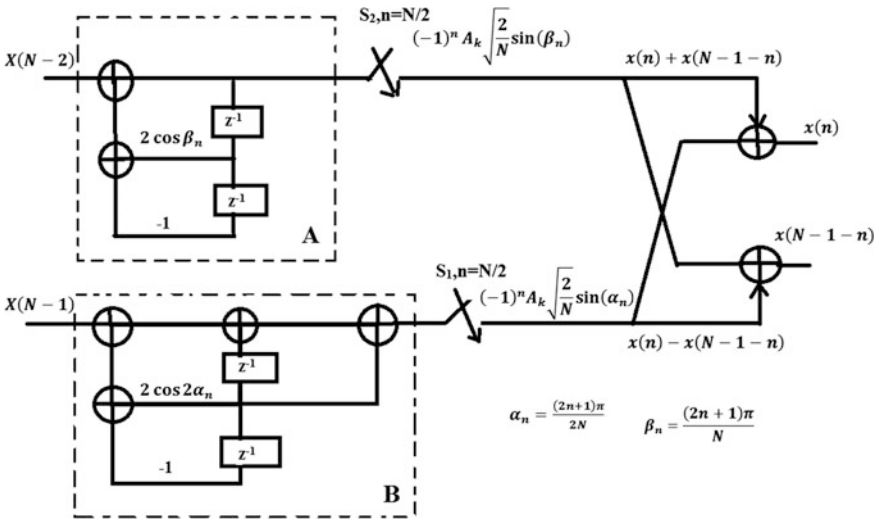

Fig. 2 Recursive filter structure for calculating *N*-point IDCT-II

Even coefficient of IDCT-II can be obtained by substituting (34) in (32a)

$$x(n)+x(N-1-n)=2(-1)^n A_k\sqrt{\frac{2}{N}}\sin(\beta_n)L_{\frac{N}{2}}^n, \quad n=0,1,2,\ldots,\frac{N}{2}-1 \tag{38}$$

Filter structure is used to realize Eq. (38) as shown in Fig. 2 marked A. Thus, thereafter we can determine the value of $x(n)+x(N-1-n)$ by using even-indexed $X(k)$.

2.2.2 OFR Kernel for the Odd Index Sequence $X(k)$

When k is odd, Eq. (30) may be written as

$$x(n)-x(N-1-n)=2\sum_{k=0}^{\frac{N}{2}-1}X(2k+1)\cos\left[\frac{(2n+1)(2k+1)\pi}{2N}\right], \quad n=0,1,2,\ldots,\tfrac{N}{2}-1 \tag{39}$$

Replacing k by $(N/2-k)$, we get following equation

$$x(n) - x(N-1-n) = 2(-1)^n \sum_{k=1}^{N/2} X(N-2k+1)\sin((2k-1)\alpha_n), \quad n = 0, 1, 2, \ldots \tfrac{N}{2} - 1 \tag{40a}$$

where

$$\alpha_n = \frac{(2n+1)\pi}{2N} \tag{40b}$$

$$x(n) - x(N-1-n) = 2(-1)^n \sin(\alpha_n) \sum_{k=1}^{N/2} X(N-2k+1) F_k^n, \quad n = 0, 1, 2, \ldots \frac{N}{2} - 1 \tag{41a}$$

where

$$F_k^n = \frac{\sin((2k-1)\alpha_n)}{\sin(\alpha_n)}, \quad n = 0, 1, 2, \ldots \tfrac{N}{2} - 1 \tag{41b}$$

It may be seen after applying the trigonometric identity

$$F_{k+1}^n = 2\cos 2\alpha_n F_k^n - F_{k-1}^n, \quad n = 0, 1, 2, \ldots \tfrac{N}{2} - 1 \tag{42}$$

Assume

$$M_{N/2}^n = \sum_{k=1}^{N/2} X(N-2k+1) F_k^n, \quad n = 0, 1, 2, \ldots \tfrac{N}{2} - 1 \tag{43}$$

Above equation can be rewritten by substituting value of $F_1^n = 1$ obtained from (41b)

$$M_{N/2}^n = X(N-1) + \sum_{k=1}^{\frac{N}{2}-1} X(N-2k-1) F_{k+1}^n, \quad n = 0, 1, 2, \ldots \tfrac{N}{2} - 1 \tag{44}$$

Using recursive relation Eq. (42), above equation can be expressed as:

$$M_{\frac{N}{2}}^k = X(N-1) + \sum_{k=1}^{\frac{N}{2}-1} X(N-2k-1)\left[2F_k^n \cos 2\alpha_n - F_{k-1}^n\right], \quad n = 0, 1, 2, \ldots \frac{N}{2} - 1 \tag{45}$$

or

$$M_{\frac{N}{2}}^{k} = X(N-1) + X(N-3) + 2\cos 2\alpha_n M_{\frac{N}{2}-1}^{n} - M_{\frac{N}{2}-2}^{n},$$
$$n = 0, 1, 2, \ldots \frac{N}{2} - 1 \tag{46}$$

We get Eq. (47) which is used to compute the odd coefficient of IDCT-II by substituting (43) in (40a)

$$x(n) - x(N-1-n) = 2(-1)^n A_k \sqrt{\frac{2}{N}} \sin(\alpha_n) M_{\frac{N}{2}}^{k},$$
$$n = 0, 1, 2, \ldots \frac{N}{2} - 1 \tag{47}$$

Filter structure is used to realize Eq. (47) as shown in Fig. 2 marked B. Thus, thereafter we can determine the value of $x(n) + x(N-1-n)$ by using odd-indexed $X(k)$

By adding and subtracting Eqs. (38) and (47), we get $x(n)$ and $x(N-1-n)$. Clearly, having computed the values of $x(n) + x(N-1-n)$ and $x(n) - x(N-1-n)$ for $n = 0, 1, 2, \ldots \frac{N}{2} - 1$, we obtain all the values of $x(n) n = 0, 1, \ldots, N$. Refer to Fig. 2 for its realization.

3 Performance Comparison and Discussion

A comparison of proposed algorithm with available recursive algorithms [13–18] is done on the basis of data throughput per transformation (DTPT) and computation speed and hardware complexity. Following observations are made on the basis of Table 1:

- Though the number of computation cycles is equal to N (maximum) among available algorithms [15] and [18], these use minimum hardware for realization.
- Number of computation cycles in proposed algorithm for IDCT is same as those in [13, 16] but hardware requirement is far less.

Table 1 Hardware complexity of different DCT/IDCT algorithms

	No. of multipliers	No. of adders	No. of latches	DTPT	Number of computation cycles
[13]	4/4	6/7	4/4	2/2	(*N*/2)/(*N*/2)
[15]	2/2	3/2	3/2	1/1	(*N*)/(*N*)
[16]	6/7	11/14	4/6	2/4	(*N*/4)/(*N*/2)
[18]	3/2	3/2	3/2	1/1	(*N*)/(*N*)
Proposed	2/4	3/7	3/4	1/2	(*N*/2)/(*N*/2)

- The proposed DCT algorithm requires two multipliers, three adders, and three latches which is similar to [15], but the former is far better in terms of number of computation cycles.
- The data throughput per transformation (DTPT) of proposed IDCT algorithm is twice of that [15, 18], while DTPT of the suggested algorithm for DCT is less as compared to [16].
- The speed of proposed algorithm may be increased by replicating the proposed hardware. The hardware complexity with this enhancement is far less than those given in [16].
- Computational speed of the proposed algorithm is better as compare to [15, 18]; however, this translates into hardware penalty.

Another comparison is made on the basis of the number of addition and multiplication operation per transformation among different algorithms for DCT computation enlisted in Table 2. It is noted that proposed algorithm requires $N/2$ multiplication and $(2N - 4)$ addition per transformation which is minimum among available DCT algorithms. Table 3 summarizes the findings related to requirement of multiplication/addition per transformation for IDCT with previously reported recursive algorithms [13, 15]. The number of multiplications requirement in the proposed algorithm is N and addition $(5/2N - 6)$, which is same as that [13] and less than that of [15]. Through the proposed structure, two coefficients are obtained simultaneously, whereas in [15] only one output coefficient is obtained. Therefore,

Table 2 Comparison of number of multiplications/additions required for DCT algorithms

N	[13] $(N)/(2N - 4)$	[14]		[15] $(N)/(3N - 4)$	Proposed $(N/2)/(2N - 4)$
		Even output $(N/2)/(N - 1)$	Odd output $(3N/2)/(2N - 1)$		
4	4/4	2/3	6/7	4/8	2/4
8	8/12	4/7	12/15	8/20	4/12
16	16/28	8/15	24/31	16/44	8/28
32	32/60	16/31	48/63	32/92	16/60
64	64/124	32/63	96/127	64/188	32/124

Table 3 Comparison of number of multiplications/additions required for IDCT algorithms

N	[13] $(N)/\left(\frac{5}{2}N - 6\right)$	[15] $(N)/(2N - 3)$	Proposed $(N)/\left(\frac{5}{2}N - 6\right)$
4	4/4	4/5	4/4
8	8/14	8/13	8/14
16	16/34	16/29	16/34
32	32/74	32/61	32/74
64	64/154	64/125	64/154
DTPT	2	1	2

saving in time/high throughput can be achieved by proposed algorithm for realization of DCT/IDCT.

4 Conclusion

In this paper, computationally efficient recursive algorithms for DCT and IDCT have been proposed. The proposed algorithms are implemented using recursive structures. The necessary formulas required for the recursive structures have been derived using mathematical manipulations. The number of addition, multiplications and computation cycles required in the recursive algorithms for DCT/IDCT is less in comparison with some existing structures. Therefore, saving in time, hardware can be achieved by the proposed algorithm for DCT/IDCT in its realization. The realized structures can be easily implemented in VLSI due to their regular, modular, and parallel structures. Moreover, any desired output kernel is computed independently. No complex multiplication is an added advantage of the design.

Acknowledgements The authors wish to acknowledge research fellowship from University Grants of commission, Government of India.

References

1. N. Ahmed, T. Natrajan, K.R. Rao, Discrete cosine transform. IEEE Trans. Comput. **C-23**(1), 90–93 (Jan 1974)
2. P. Jain, A. Jain, Regressive structures for computation of DST-II and its inverse. Int. Sch. Res. Netw. ISRN Electron., pp. 1–4 (2012)
3. A.K. Jain, A sinusoidal family of unitary transforms. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-1**(4), 356–365 (Oct 1979)
4. C.H. Chen, B.D. Liu, J.F. Yang, Recursive architectures for realizing modified discrete cosine transform and its inverse. IEEE Trans. Circuit Syst. II Anal. Digit. Signal Process. **50**(1) (Jan 2003)
5. Z. Wang, B. Hunt, The discrete cosine transform—a new version, in *IEEE International Conference on ICASSP'83 Acoustics, Speech, and Signal Processing*, vol. 8, pp. 1256–1259 (Apr 1983)
6. P. Jain, B. Kumar, S.B. Jain, A general design for one dimensional discrete sine transform. Analog Integr. Circuits Signal Process **61**(2), 211–214 (Nov 2009)
7. D. Hein, N. Ahmed, On a real-time Walsh-Hadamard cosine transform image processor. IEEE Trans. Electromagn. Compat. **1**(EMC-20), 453–457 (Aug 1978)
8. Z. Wang, Fast algorithms for the discrete W transform and for the discrete Fourier transform. IEEE Trans. Acoust. Speech, Signal Process. **ASSP-32**, 803–816 (Aug 1984)
9. J.-F. Yang, C.-P. Fan, Recursive discrete cosine transforms with selectable fixed-coefficient filters. IEEE Trans. Circuits Syst. II: Analog Digit. Signal Process. **46**(2), 211–216 (Feb 1999)
10. Z. Cvetkovic, M.V. Popovic, New fast recursive algorithms for the computation of discrete cosine and sine transforms. IEEE Trans. Signal Process. **40**(8), 2083–2086 (Aug 1992)
11. P. Lee, F.Y. Huang, Reconstructed recursive DCT and DST algorithms. IEEE Trans. Signal Process. **42**(7), 1600–1609 (1994)

12. L.-P. Chau, W.-C. Siu, Recursive algorithm for the discrete cosine transform with general lengths. Electron. Lett. **30**(3) (Feb 1994)
13. J.L. Wang, C.B. Wu, B.D. Liu, J.F. Yang, Implementation of the Discrete Cosine Transform and its inverse by recursive structures, in *SiPS 99. 1999 IEEE Workshop on Signal Processing Systems, 1999*, pp. 120–130 (Oct 1999)
14. M.N. Murthy, Radix-2 algorithms for implementation of type-II discrete cosine transform and discrete sine transform. Int. J. Eng. Res. Appl. **3**(3), 602–608 (May–June 2013)
15. M.N. Murthy, Recursive algorithms and systolic architectures for realization of type-II discrete cosine transform and inverse discrete cosine transform. Int. J. Eng. Res. Appl. **4**(4), 24–32 (Apr 2014)
16. C.-H. Chen, J.-F. Yang, J.-L. Wang, Efficient recursive structures for forward and inverse discrete cosine transform. IEEE Trans. Signal Process. **52**(8), 2665–2669 (Sept 2004)
17. P. Lee, Restructured recursive DCT and DST algorithms. IEEE Trans. Signal Process. **42**(7), 1600–1608 (July 1994)
18. Z. Wang, G.A. Jullien, W.C. Miller, Implementation of the discrete cosine transform and its inverse by digital filtering, in *Proceedings of the 37th Midwest Symposium on Circuits and Systems*, vol. 23, pp. 821–824 (5 Aug 1994)

Wideband Patch Antenna for Millimeter-Wave Application

Shilpa Srivastava, Navneet Sharma and Pradyot Kala

Abstract This paper presents a wideband antenna for millimeter-wave application. In modern era, all wireless devices are multipurpose and versatile, so this requires a wideband antenna to perform different tasks. A single antenna is used rather than more antennas in a device, to reduce the size and cost of the device. Proposed antenna has a wide bandwidth of operating frequency bandwidth at 31.47–38.9 GHz, respectively. Gain about 7 dBi with voltage standing wave ratio (VSWR) is less than 2 at operating frequency 35.8 GHz. Design is simulated by high-frequency structural simulator (HFSS).

Keywords Patch · Wideband · HFSS · Millimeter wave

1 Introduction

Radars and satellites are the examples of the microwave applications. A radar must have a wide band to detect the objects, and a satellite must have a great bandwidth to perform different communication aspects. To fulfill these requirements, a wideband antenna is required. In present scenario, every wireless device has a different operating frequency to perform. For example, if we talk about the mobile phone, it has Bluetooth, Wi-Fi, and cellular technologies needed to operate on different frequency bands. The solution to this scenario is either to use multiple antennas in a device or use a wideband antenna to operate in a wide frequency band

S. Srivastava (✉) · N. Sharma
Department of ECE, ABES Engineering College, Ghaziabad, India
e-mail: shilpa.srivastava@abes.ac.in

N. Sharma
e-mail: navneet.sharma@abes.ac.in

P. Kala
Department of ECE, Galgotia Engineering of College, Greater Noida, India
e-mail: pradyot_bly@yahoo.co.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_8

range with multiple resonant frequencies. A wideband antenna with smaller in size makes the device cheaper and compatible.

A design of wideband antenna with a patch makes fabrication simple and economical [1]. A wide range of resonant frequencies makes it more useful in wireless communication within microwave range. Since the Federal Communications Commission (FCC) has been allocated the 7-GHz (57–64-GHz) band as unlicensed band, the research of many innovative technologies at 60 GHz band has led to many exciting commercial applications, including interfacing of high-definition multimedia, streaming of high-definition video, mobile-distributed computing, and high-speed Internet. Millimeter-wave and quasi-millimeter-wave frequencies are used for investigations on high-speed wireless communications. Research on wideband antenna has been started many years ago, and since then, a lot of enhancement and improvement is done. A key-shaped fractal antenna [2] for wideband applications is designed, while in [3] a wideband L-shaped microstrip antenna for millimeter band and working for 60 GHz band application has been presented. A proximity-coupled multiband microstrip antenna is presented that operates on two bands in [4]. The other methods are probe-fed patch antenna [5] where the substrate thickness is increased. Furthermore, the enhancement in bandwidth of patch antenna can be done with stack patch [6]. A wide impedance bandwidth is achieved using dual layer patch antenna with L-strip-fed technique [7]. Table 1 shows the comparison of performance characteristics between ultra-wideband (UWB) antennas and wideband antennas. A summarized comparison of the parameters of the different types of wideband antennas is shown in Table 1.

In this paper for wideband, one single-patch antenna has been proposed for multiple functions operating on wideband frequency range which reduces the size of the antenna. The paper is organized as follows. Introduction is described in Sect. 1, and Sect. 2 describes proposed antenna design. In Sect. 3, simulated result validates the accuracy of the analysis and demonstrates the improved performance of the proposed antenna. Section 4 concludes the paper.

Table 1 Antenna parameter for comparison for different types of wideband antennas

Antenna type	Gain (dBi)	Size	Bandwidth (GHz)
Frequency-notched UWB microstrip slot antenna with tuning stub	upto 5	48 × 41 mm	2.66–10.76
CPW-fed planar with frequency band notch function	upto 5	22 × 31 mm	1.97–12.73
Square planar monopole antenna using notching technique	4−6	30 × 30 mm with ground plane	1.65–20
Rectangular planar monopole antenna using beveling technique	5−7	60 × 39 mm with ground plane	31.47–38.9
Rectangular slot antenna	upto 7	52 × 38 mm	31.47–38.9

2 Proposed Antenna

2.1 Design Consideration

The microstrip patch antenna has narrow bandwidth, but in the proposed antenna, the substrate material, size, and height of the antenna are calculated so that it can operate over a wide range of frequencies. The antenna operating bandwidth increases with the slots cut in patch as electric length changes. Resonance frequency increases as the electrical length increases. The inductive impedances are zero, and impedance of transmission line is in match with antenna impedances, so maximum current flows through the antenna. The dimensions of the proposed wideband patch antenna and the length (S_{L1}) and width (S_{W1}) of the slots are given in Table 2 (Fig. 1).

Table 2 Proposed Antenna Dimensions

Parameters	Proposed antenna dimensions (mm)
S_{L}	38
S_{X}	52
S_{H}	2.8
S_{L1}	7.8
S_{W1}	10.4

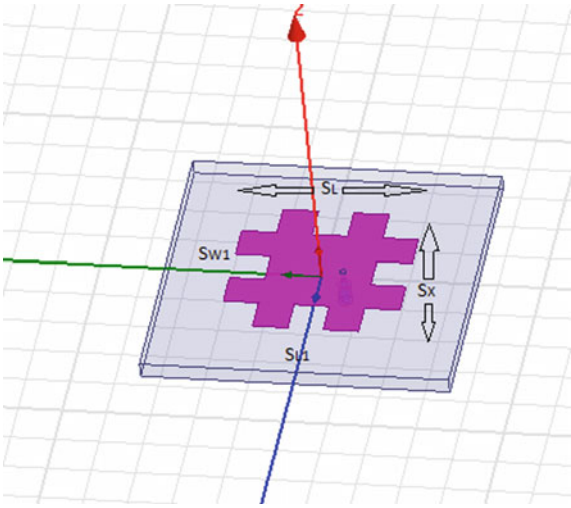

Fig. 1 Simulated design of proposed antenna

2.2 Equations

The length (L) of the patch is about $\lambda_g/2$ (where λ_g is an effective wavelength), and substrate height (h) is of order of $\lambda_g/20$. The width of the patch is w. The wavelength of the dielectric medium is given by Eq. (1); the effective relative permittivity of the substrate (S_H) of the proposed antenna is calculated by Eq. (2). The length (S_L) is calculated using Eqs. (3), (4), and (5); resonant frequency as given by Eq. (6) and width(S_x) of the rectangular patch is calculated using Eq. (7)

$$\lambda_g = \frac{\lambda_o}{\sqrt{\varepsilon\,\mathrm{eff}}} \tag{1}$$

where

$$\varepsilon\,\mathrm{reff} = \frac{\varepsilon r + 1}{2} + \frac{\varepsilon r + 1}{2}\left[1 + 12\frac{h}{w}\right]^{-\frac{1}{2}} \tag{2}$$

where ε_r is the relative permittivity and ε_{reff} is the effective relative permittivity of the substrate, λ_o is the resonant wavelength, and ΔL is the due to fringing effect.

$$\mathrm{Leff} = \frac{c}{2fr\sqrt{\varepsilon\,\mathrm{reff}}} \tag{3}$$

$$L = \mathrm{Leff} - 2\Delta L \tag{4}$$

$$\Delta L = \frac{(\varepsilon\,\mathrm{reff} + 3)\left(\frac{w}{h} + 0.264\right)}{(\varepsilon\,\mathrm{reff} - 8.258)\left(\frac{w}{h} + 0.8\right)} \tag{5}$$

For rectangular patch antenna, the resonance frequency (f_o) is given by

$$f_o = \frac{c}{2\sqrt{\varepsilon\,\mathrm{reff}}}\left[\left(\frac{m}{L}\right)^2 + \left(\frac{n}{W}\right)^2\right]^{\frac{1}{2}} \tag{6}$$

where m and n are modes, and for efficient radiation, the width of the patch is (w), and c is the velocity of em wave, respectively.

$$w = \frac{c}{2f_o\sqrt{(\varepsilon r + 1)/2}} \tag{7}$$

The slots are inserted in the patch to increase the electrical length, and it can operate over a wide range of frequencies making it ultra-wideband (UWB) in the millimeter range.

3 Simulated Results and Discussion

The proposed antenna is simulated using high-frequency simulator structure (HFSS). The Rogers RO4232(tm) material is used as a substrate with relative permittivity of 3.2. It is observed in simulation results that the proposed antenna is operating at three bands successfully. Figure 2 shows the return loss of the antenna with resonance frequency bands below than −10 dB. The VSWR, gain, and radiation pattern are shown in Figs. 3, 4 and 5, respectively.

3.1 Return Loss

The most important parameter of any antenna is return loss, which shows about operating frequency of antenna. Its value must be greater than 10 dB because this is a standard value of any antenna. The values of return loss for proposed antenna are −21.9, −21.87, and −25.49 dB at 35.8, 31.3, and 35.72 GHz, respectively. The conventional antenna is working at single frequency band, and this is a narrowband, but the proposed antenna is working at different frequency bands, and the last band is a wide band as shown in Fig. 2.

3.2 Voltage Standing Wave Ratio (VSWR)

The value of VSWR is an important parameter for designing of any antenna. The value of this parameter should be less than 2. It is clear from Fig. 3 that for entire frequency range, the VSWR is less than 1.5.

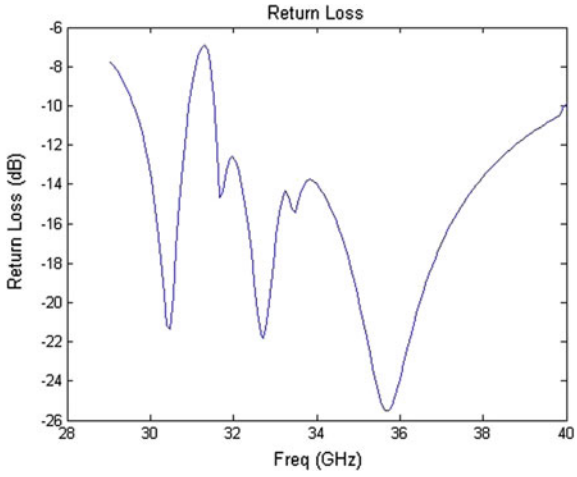

Fig. 2 Return loss of the proposed antenna

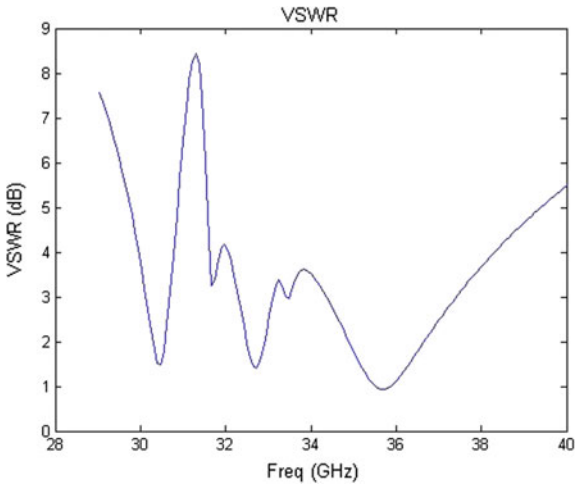

Fig. 3 VSWR of the proposed antenna

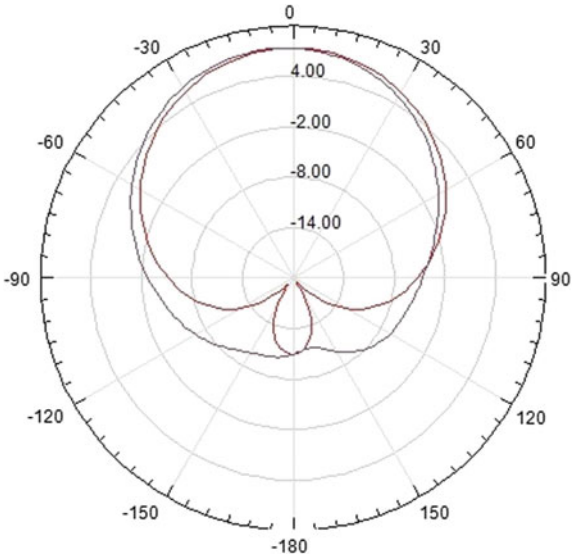

Fig. 4 Gain of the proposed antenna

3.3 Gain

Most important parameter of antenna is gain. The term antenna gain describes how much power is transmitted in the direction of peak radiation to that of an isotropic source. Antenna gain is more commonly quoted than directivity in an antenna's specification sheet because it takes into account the actual losses that occur. Gain of this antenna is obtained 7 dBi as shown in Fig. 4.

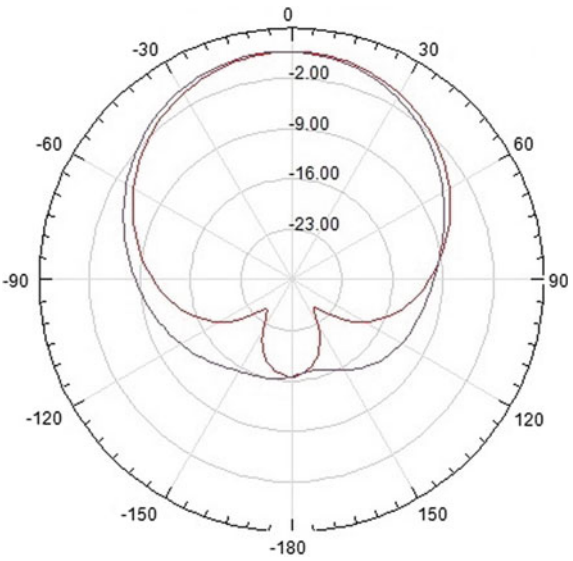

Fig. 5 Radiation pattern of the proposed antenna

3.4 Radiation Pattern

A radiation pattern defines the variation of the power radiated by an antenna as a function of the direction away from the antenna. This power variation as a function of the arrival angle is observed in the antenna's far field. Radiation pattern of the proposed antenna is shown in Fig. 5 which shows the direction of maximum radiation in single direction with minimum radiation in other directions.

4 Conclusion

The proposed microstrip patch antenna is successfully simulated in HFSS. Results show that it operates over a wide range of frequencies such that it is useful for ultra-wideband operations (UWBs) in the millimeter-wave range of 35.8, 31.3, and 35.72 GHz. Due to small in size and simplified structure, it will prove beneficial for high-end devices, thereby reducing the overall cost of the device.

References

1. D.M. Pozar, Microstrip antennas. Proc. IEEE **80**(1), 79–91 (1992)
2. D. Kumar, M. Sharma, S. Bansal, in *Novel Design of Key-Shaped Fractal Antenna for UWB Applications*. IET, International Conference on Computational Intelligence and Communication Networks, vol. 6 (2014)

3. M. Li, K.-M. Luk, Low-cost wideband microstrip antenna array for 60-GHz applications. IEEE Trans. Antennas Propag. **62**(6) 2014
4. P.S. Bakariya, S. Dwari, M. Sarkar, M.K. Mandal, Proximity-coupled multiband microstrip antenna for wireless applications. IEEE Antennas Wirel. Propag. Lett. **14** 2015
5. K.-L. Wong, C.-L. Tang, J.-Y. Chiou, Broad-band probe-fed patch antenna with a W-shaped ground plane. IEEE Trans. Antenna Propag. **50**(6) 2002
6. M.T. Islam, M.N. Shakib, N. Misran, B. Yatim, Ultrawideband EH shaped stack patch antenna for wireless communications 978-1-4244-4565-3/09/$25.00 ©2009 IEEE
7. Z. Wang*, S. Fang, S. Fu, Wideband dual-layer patch antenna fed by a modified L-strip. J. MicroWaves Optoelectron. Electromagn. Appl. **9**(2) 2010

Enhancement of Trunking Efficiency and Analysis of Cell Capacity in CDMA

Navneet Sharma, Himani Garg and Shilpa Srivastava

Abstract In this work, the presence of soft and softer handoffs on code-division multiple access (CDMA) is explored to evaluate system capacity for unsectorized and sectorized hexagonal cells, and analysis of capacity has been done with the application of soft handoff as well as softer handoff, thereby increasing the macrodiverse area of the cell. Proposed model has also enhanced the trunking efficiency of mobile switching center (MSC), by generating a breathing cluster (rather than a breathing cell) and reporting to MSC only when a cluster capacity is reached, thus increasing the trunking efficiency by the factor of the cluster size

Keywords Handoffs · Code-division multiple access · Soft handoff
Softer handoff

Abbreviations

BS	Base stations
BER	Bit error rate
CDF	Cumulative distribution function
CDMA	Code-division multiple access
MS	Mobile station
MSC	Mobile switching center
PSD	Power spectral density
QoS	Quality of service
SIR	Signal to interference ratio
VAF	Voice activity factor

N. Sharma (✉) · H. Garg · S. Srivastava
Department of ECE, ABES Engineering College, Ghaziabad, India
e-mail: navneet.sharma@abes.ac.in

H. Garg
e-mail: himani.garg@abes.ac.in

S. Srivastava
e-mail: shilpa.srivastava@abes.ac.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_9

1 Introduction

The coverage area in a cellular network is divided into cells and every cell has one base station; therefore, the numbers of base stations (BSs) are fixed. A mobile station (MS) communicates with the BSs close to it (Fig. 1). There are two main issues in the cellular communication at the physical and medium access layers and they are multiple access and interference management. The former issue addresses how the overall resource (time, frequency, and space) of the system is shared by all the users in the same cell (intra-cell) and the latter addresses the interference caused by simultaneous signal transmissions in different cells (inter-cell).

Soft handoff enhances the capacity and coverage of a CDMA system. Its effect on the capacity-coverage trade-off curve will now be considered here. Partitioning the cell enables the comparison with the capacity-coverage trade-off with no handoff [1]. When the MS is communicating only with the BS, designing of R_{in} is such that the outage probability at a distance R_{in} from the cell is P_{out} (required outage probability). So, the inner circle is the region of no soft handoff. R_{cov} represents the coverage of the pilot and R_{out} is such that it allows soft handoff with other BS as shown in Fig. 1. The analysis carried out is the same as in no handoff case but with A_{out} and B_{out} replaced with their equivalent expressions for soft handoff. A_{out} will be the event that all BSs connected to the user under consideration do not have a feasible solution and B_{out} will be the event that all BSs require power from the user greater than the maximum possible. Seamless connectivity to the MS is an important issue to be considered at the network layer, as it moves from one cell to the another (and thus communication is switched from one BS to another, an operation popularly known as handoff) [2, 3]. The analysis is based on the soft handoff process, in which an MS is connected to more than one BS or more than one sectors and softer handoff is where more than one sectoral antenna of similar or another cell is involved.

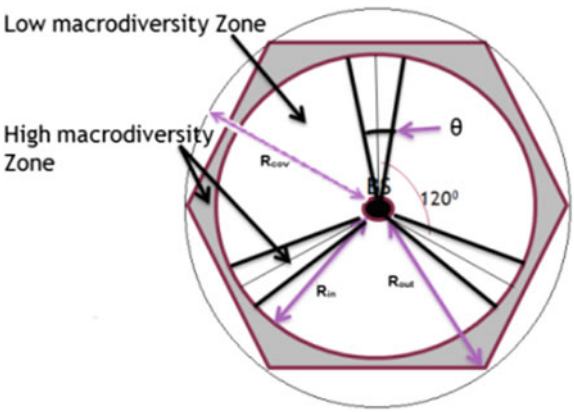

Fig. 1 Cell ADII and macrodiverse regions

2 Capacity Issues

The communication resources can be represented by virtue of capacity, and quantitative expressions of capacity can be implemented by considering:

(a) number of users that can be included in a communication system as an indication of the capacity of the system.
(b) number of calls that can be handled by a communication system as an indication of the capacity of the system.

The concept of breathing cell [4] where the cell resources are consumed for every entry of the user, thus, exists the concept of breathing cluster, where the resources are shared by each user (MS). On reaching the maximum capacity of the cell (by determining the value of Quality of Service (QoS)), a trunking signal is sent to MSC for notification of capacity outreach.

3 Methodology

The following methodology has been followed for the generation of a system model:

Step 1. Health of each MS is determined by checking its bit error rate (BER). BER determines the QoS to the user.
Step 2. Entry of a MS due to either call initialization or handoff process is taken as an additional input to the system. It is to be mentioned that handoff process is on higher priority than call initialization in the event of capacity reach.
Step 3. The user is placed randomly in the cell; user may lie in a cell where it has a strong control of the single-host BS or at a macrodiverse area where more than one BS or sector handles the call.
Step 4. Entry of a user to the cluster increases the BER, which is compensated with the QoS of the system, thus creating a breathing cluster (contrary to breathing cell) [4].
Step 5. By exhausting the predetermined QoS, the capacity is achieved which can be reported to MSC.

4 Softer and Soft Handoff in Sectorized Cells

The system capacity effects on both forward and reverse links are examined.

4.1 Capacity Loss (Forward Link)

The overlapping zone of the antenna pattern at the sector boundaries is used for softer handoff. Softer handoff overlapping coverage areas are shown in Fig. 2.

These overlapping areas between the sector boundaries can be modeled as segments of a cell with angle θ_0 as shown in red in Fig. 1. The categorization into softer, soft, and soft or softer handoff subzones in the sectorized cells can be done, as in Fig. 3. The overlapping area of soft as well as softer handoff subzones is termed as the area of soft and softer handoff subzones. Any MS could be handled by either softer or soft handoff depending on its location. Only those MS which are located in the subzone of softer handoff would be using softer handoffs. The area can be defined as in Eq. (1):

$$\psi_{\text{sh}} = \frac{1}{2}\theta_0 R_{\text{h}}^2 \tag{1}$$

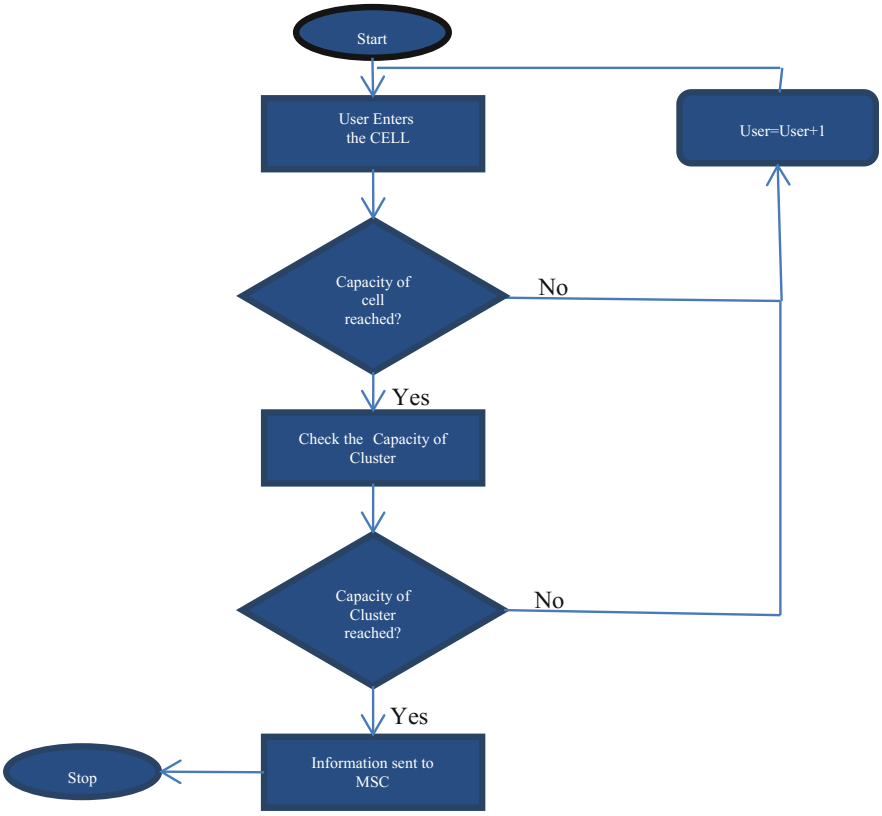

Fig. 2 Flow diagram for a user entry to cell and cluster

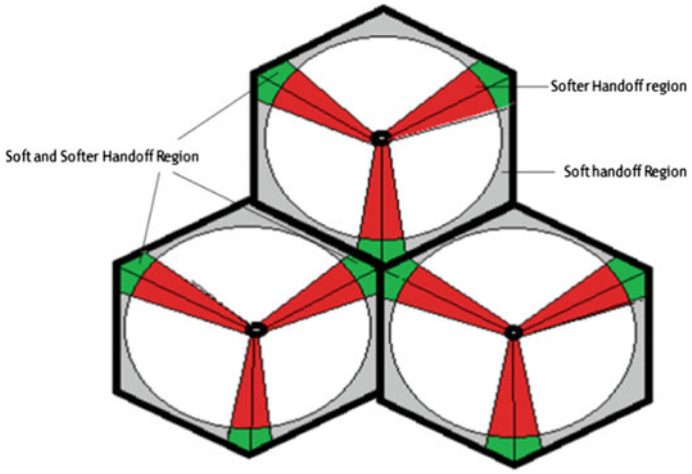

Fig. 3 Soft and softer handoff regions

where

θ_0 is the overlapping area at the cell boundary.
R_h is the starting distance of soft handoff region.

Assuming uniformly distributed users with ρ as MS density in the zone of softer handoff. In the softer handoff zone, two BSs are used to handle the MS, and the cell capacity because of the MS in these zones is $1/2\rho\psi_{sh}$. The loss of capacity due to softer handoffs can be accounted from Eq. (2) as:

$$F_{sh} = \frac{\frac{1}{2}\psi_{sh}}{\frac{\sqrt{3}}{2}R^2} = \frac{\theta_0}{4\sqrt{3}}\left(\frac{R_h}{R}\right)^2 \tag{2}$$

where

R is the distance between BS and edge of the hexagon.

Fractional capacity loss (F_T) in a sector due to soft and softer handoffs can be formulated as in Eq. (3):

$$F_T = \frac{1}{2}\left[1 - \frac{1}{\sqrt{3}}\left(\frac{2\pi}{3} - \theta_0\right)\left(\frac{R_h}{R}\right)^2\right] \tag{3}$$

Soft and softer handoffs give the fraction of the capacity loss with θ_0 as the function of the overlapping angle. The handoff threshold, taken as 3 dB higher than the lowest detectable signal level for $\theta_0 = 5^\circ$. In comparison with the capacity loss because of soft handoff in unsectorized cell, the loss of capacity in soft and softer handoffs is 9% less [5]. Additional 1.8% capacity loss is due to softer handoff only [5].

4.2 Capacity Gain (Forward Link)

The forward link capacity is first calculated without macrodiversity and then with macrodiversity. The number of active MS at the boundaries of the cell limits the capacity in the forward link, accounting for the intercellular interference higher than for the MS inside the cell. The areas having the maximum level of interference are considered for capacity evaluation.

4.2.1 Without Macrodiversity

In addition to the interfering sectors in the adjacent areas, the antenna overlapping areas is also interfering sector for MS at the corner of the hexagon. Considering the first BS and the zeroth BS at the same site of the cell, forming first and the zeroth sector, respectively, the shadowing conditions and the path loss from the BS to the MS in softer handoff zone are same. Further, it can be assumed for simplicity that the power received from the zeroth sector BS and the first sector BS is same as shown in Eq. (4), i.e.,

$$(r_0/r_1)^{\propto} 10^{\lambda_1-\lambda_0} = 1 \tag{4}$$

where

r_0 is the distance of BS from the zeroth cell
r_1 is the distance of BS with the consecutive cell

without macrodiversity the power spectral density (PSD) ratio (E_b/I_0) for mobiles in this overlapping area [1], for a BER less than 10^{-3}, path loss γ_{req} is approximately 7 dB. Specifying γ_{req}, the capacity for forward link can be evaluated as [1] with minimal number of users N as shown in Eq. (5)

$$N \geq \frac{1 + \frac{G_P}{\mu}\left(\frac{1}{\gamma_{req}} - \frac{\eta}{S}\right)}{1.5 + \frac{1}{2}\sum_{j=1}^{J-1}\left(\frac{r_0}{r_j}\right)^{\propto} E\left[\phi_j\left(\zeta_j, \frac{r_0}{r_j}\right)10^{\zeta_j/10}\right]} - 1 \tag{5}$$

where

G_p is the propagation gain
γ_{req} is the estimated path loss
$\frac{r_0}{r_j}$ is the ratio of the distance between the BS and zeroth BS
E is the electric field
ϕ_j is the angle which given BS makes with zeroth BS
$\zeta_j = \lambda_j - \lambda_0$ is a normal random variable having a standard deviation $\sqrt{2}\sigma$ as λ_0 and λ_j are two independent fading variables.

Considering the situation of an active MS at center of the sector, the calculation for the number of minimum users per sector can be done in accordance with the interfering sectors [5]. If the MS is located near boundary, there are two or more interfering sectors in non-softer handoff zone [6]. Macrodiversity that is provided by soft handoff handles the additional interference due to overlapping antenna sector pattern. Worst case of signal to interference ratio (SIR) occurs at the MS location of soft or softer handoff subzones, maximizing at the corners. For voice activity factor (VAF) of 1/2 and 3/8, 22–30 users per sector could be supported in the forward link system for analysis. The forward link capacity in the sectorized cells, without soft and softer handoff subzones, is limited by minimum number of users.

4.2.2 With Macrodiversity

To the MS located at the overlapping of the BS sectorized antennas, we get the worst SIR, and the evaluation of gain of capacity due to macrodiversity in the areas of overlapping has to be done. It is because of maximum ratio combination of the two signals from two neighboring sectors BS are employed by the MS in soft and softer handoff regions, similar procedure as the one employed earlier can be used to calculate the combined E_b/I_0 [5].

5 Results

The soft and softer handoff regions are shown separately in Fig. 1. The presence of both soft and softer handoffs increases the macrodiverse area of the cell and creating several macrodiverse regions in the cluster. Secondly, it had been shown in the capacity analysis where simulation model for three different practical scenarios had been created, keeping position of the users and BER random and other parameters deterministic. The cluster capacity has been evaluated in all the scenarios such as *rural* location, *semi-urban* location, and *urban* location. The simulations are done using MATLAB.

5.1 Capacity Analysis for Rural Region

One of the unique characteristics of CDMA system is that it is possible to trade off between capacity and its coverage [7] and is called the cell-breathing. The above-mentioned parameters are used for creating a model for a rural region where the following parameters are taken to simulate a rural environment where a number of users are less and the cell radius is kept high, keeping this in mind the lower noise interference path loss factor is kept low. VAF is kept 3/8 which is kept a standard for the whole analysis.

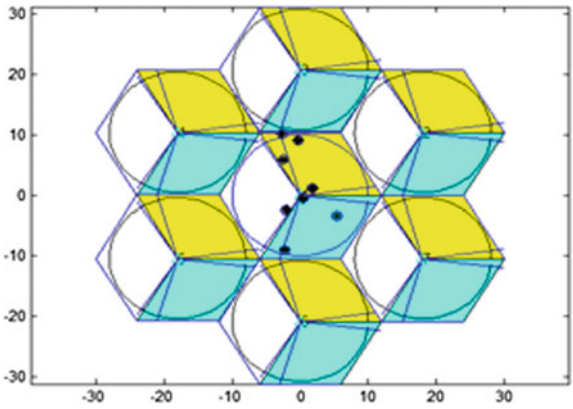

Fig. 4 A cluster showing regions of macrodiversity (rural)

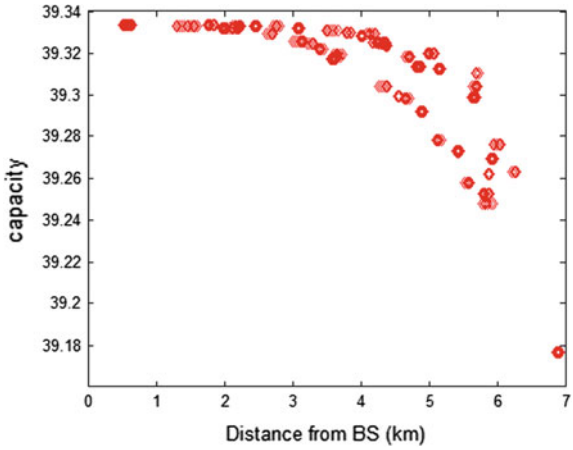

Fig. 5 A graph of capacity without soft or softer handoff within the cell (rural)

The parameters taken are as follows:

- Cell radius: 12 km.
- Number of mobile stations (MS): 08
- Path loss factor: 02
- VAF: 3/8

The resultant outputs are shown in Fig. 4.

Figure 4 shows 8 mobile stations randomly distributed over the area in the rural region where MS at high microdiverse area is subjected to lesser BER and hence provides greater capacity as shown in the cumulative distribution function (CDF) in Fig. 7.

As shown in Fig. 5, the analysis is done without soft and softer handoffs. As shown, the capacity achieved without macrodiverse effect is lesser as compared to capacity achieved at macrodiverse area; a graph of capacity is shown for the rural region, where the macrodiverse area of a particular cell is explored. A graph showing capacity with soft as well as softer handoffs within the cell (Rural)

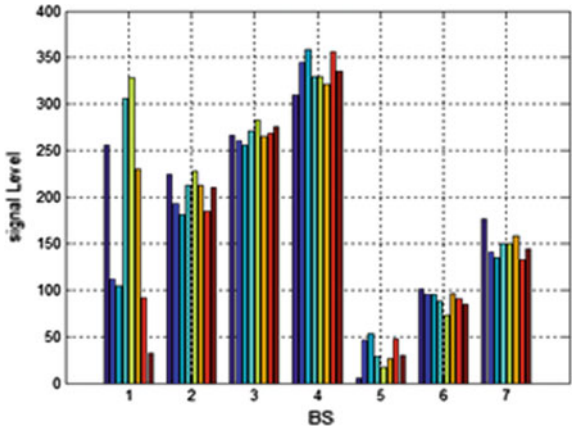

Fig. 6 Signal levels of individual MS at BS depending on its position (rural)

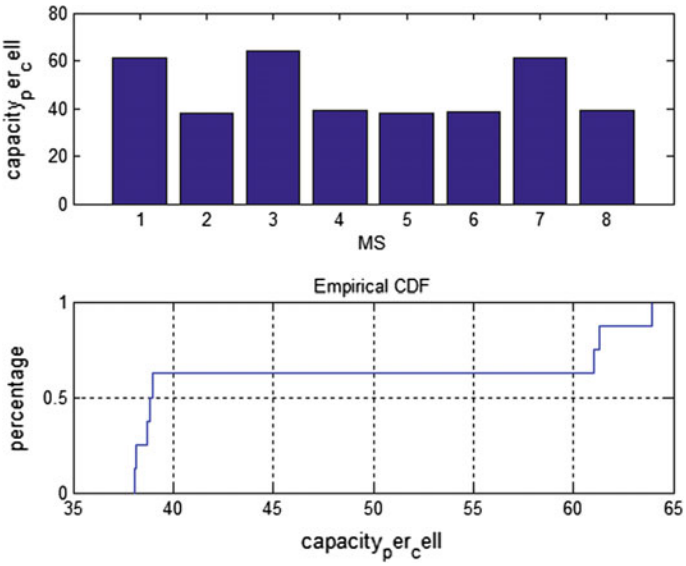

Fig. 7 Cumulative distribution function on capacity (rural)

increases the cell capacity, though the capacity is limited due to co-channel interference.

Figure 6 shows the signal level for individual MS as seen from a base station. The signal levels are measured by generating the BER.

Figure 7 shows the cumulative distribution function of the capacity; the capacity of the cluster, on application of soft as well as softer handoffs, is achieved at around

64 in the analysis. This is the point at which MSC has to be informed about the capacity outreach.

5.2 *Capacity Analyses for Semi-urban Region*

The above-mentioned parameters are used for creating a model of a semi-urban region; the following parameters are taken to simulate a semi-urban environment where number of users, cell radius, and path loss factor are kept moderate. VAF is kept 3/8.

The parameters taken are as follows:

- Cell radius: 7 km.
- Number of mobile stations (MS): 50
- Path loss factor: 05
- VAF: 3/8

The resultant outputs are shown in Figs. 8, 9, 10, 11 and 12.

Figures 8, 9, 10, 11 and 12 show the analysis for semi-urban region. Figure 9 shows the graph of capacity without soft or softer handoff, and it can be seen that the capacity is reached earlier because we are not using soft or softer techniques. Figure 10 shows the soft and softer handoff applications within the cell. The capacity is limited because of co-channel interference, which is increased due to the reduced cell size and increased number of users.

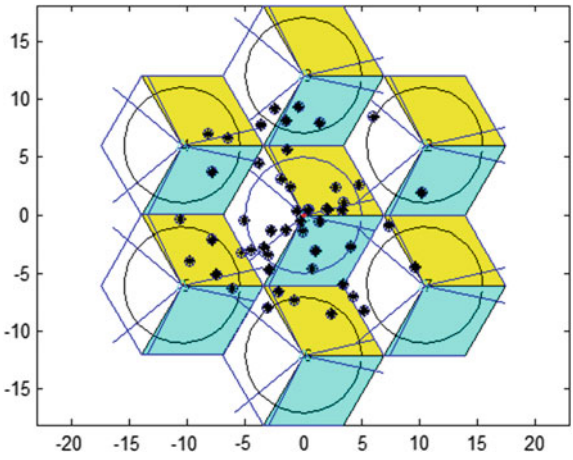

Fig. 8 A cluster showing regions of macrodiversity (semi-urban)

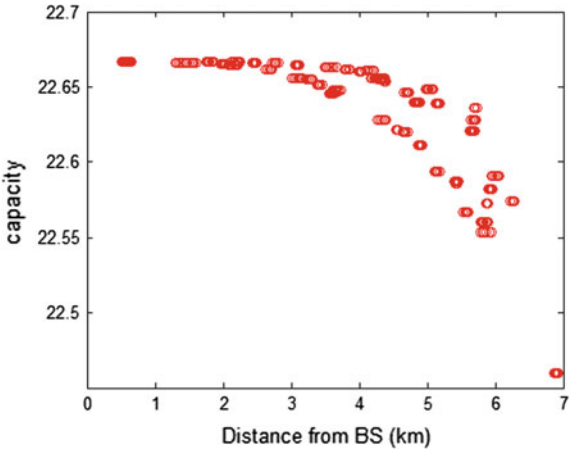

Fig. 9 A graph of capacity without soft or softer handoff within the cell (semi-urban)

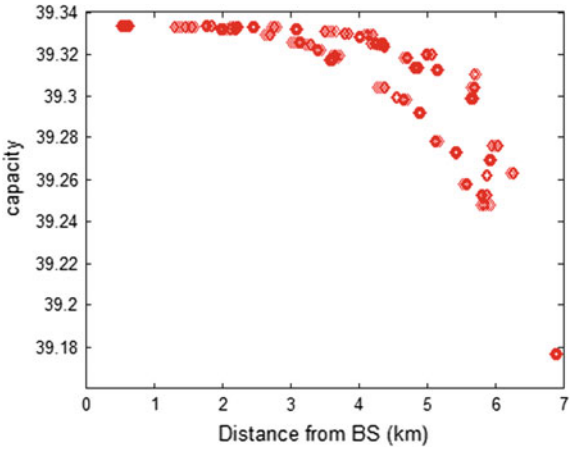

Fig. 10 A graph showing capacity with soft as well as softer handoff within the cell (semi-urban)

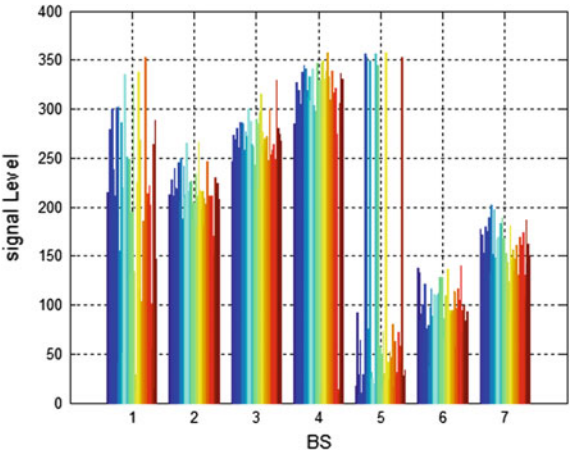

Fig. 11 Signal levels of individual MS at BS depending on its position (semi-urban)

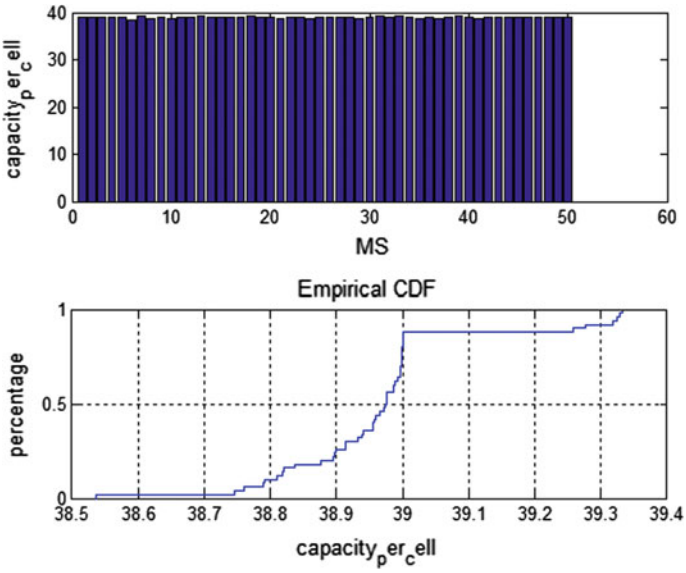

Fig. 12 Cumulative distribution function on capacity (semi-urban)

5.3 Capacity Analysis for Urban Region

The above-mentioned parameters are used for creating a model of an urban region where the following parameters are taken to simulate an urban environment, where number of users are higher (densely populated), and the cell radius is kept low, whereas path loss factor is kept high, keeping in mind the high noise interference. VAF is 3/8.

The parameters taken are as follows:

- Cell radius: 3 km.
- Number of mobile stations (MS): 100
- Path loss factor: 08
- VAF: 3/8

The resultant outputs are shown in Figs. 13, 14, 15, 16 and 17.

Figures 13, 14, 15, 16, and 17 show the analysis for the urban region. Figure 13 shows the random distribution of MS over the region, where cell radius is reduced to 3 km. Figure shows the graph of capacity without soft or softer handoff; it can be seen that the capacity is reached earlier because we are not using soft or softer techniques. Figure 15 shows the occurrence of soft and softer handoffs within the cell, though there is a capacity enhancement as shown in Fig. 17. The capacity is limited because of co-channel interference, which is increased due to the reduced cell size and increased number of users. Figure 16 shows the signal level of individual MS as seen at the BS.

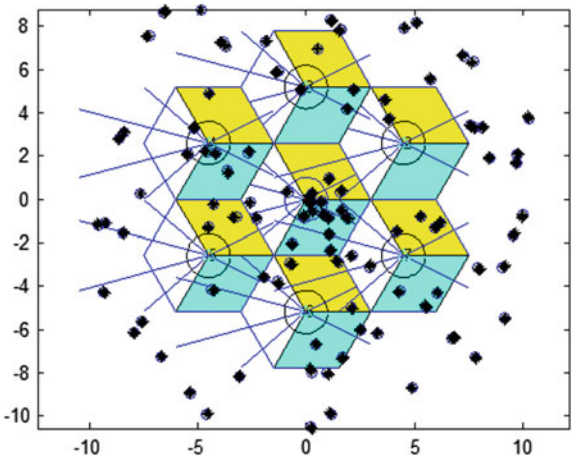

Fig. 13 A cluster showing regions of macrodiversity (urban)

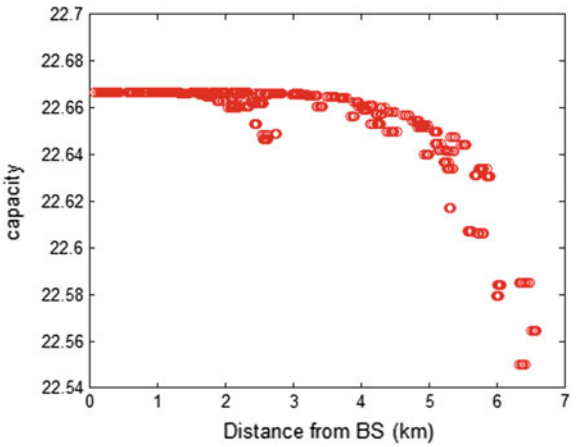

Fig. 14 A graph of capacity without soft or softer handoff within a cell (urban)

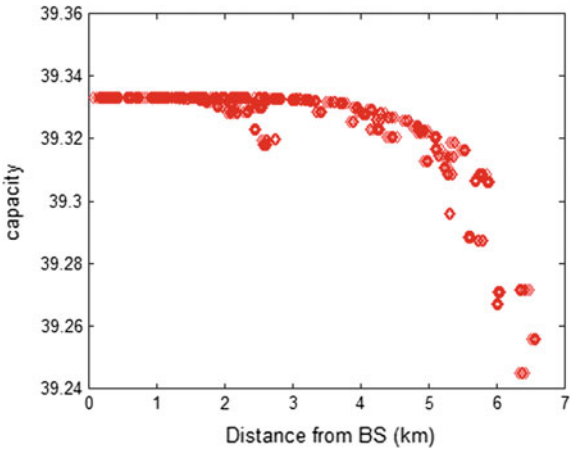

Fig. 15 A graph showing capacity with soft as well as softer handoff within a cell (urban)

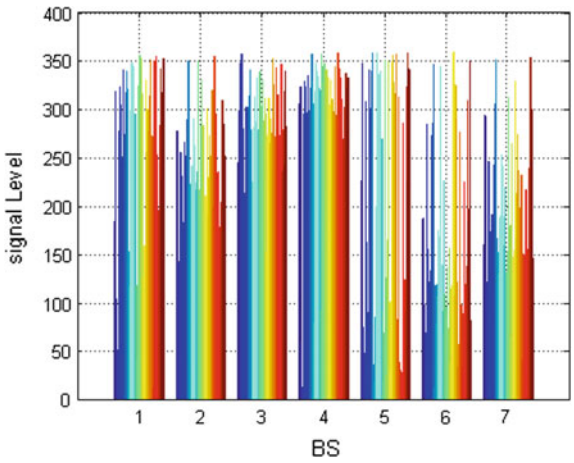

Fig. 16 Signal levels of individual MS at BS depending on its position (urban)

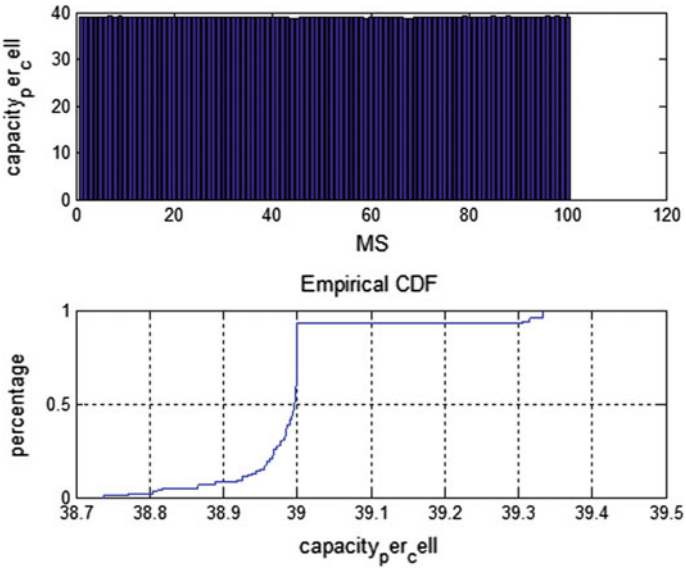

Fig. 17 Cumulative distribution function on capacity (urban)

6 Conclusion

The information is conveyed to MSC for the capacity reach out at the point where the CDF graph shows 100%. This decreases the amount of information sent to MSC and thus increasing the trunking efficiency. The analysis for the capacity is done

and has been observed that, when soft and softer handoffs are implemented, there is enhancement in the capacity.

References

1. A.J. Viterbi, A.M. Viterbi, K.S. Gilhousen, E. Zehavi, Soft handoff extends CDMA cell coverage and increases reverse link capacity. IEEE J. Select. Areas Commun. **12**(8), 1281–1287 (1994)
2. T.S. Rappaport, *Wireless Communications* (Prentice-Hall Inc., 1996)
3. A. Salmasi, K.S. Gilhousen, in *On the System Design Aspects of Code Division Multiple Access (CDMA) Applied to Digital Cellular and Personal Communications Networks*. IEEE 41st VTS Conference (St. Louis, MO, May 1991), pp. 57–62
4. J. Yao, D.T.C. Wong, Y.H. Chew, Capacity balancing between the reverse and forward links in multiservice CDMA cellular networks with cross layer design. IEEE Trans. Veh. Technol. **55** (4), 1397–1411 (2006)
5. C.C. Lee, R. Steele, Effect of soft and softer handoffs on CDMA system capacity. IEEE Trans. Veh. Technol. **47**(3), 830–841 (1998)
6. Q. Wu, Q. Liang, in *Increasing Capacity of Multi-cell Cooperative Cellular Networks with Nested Deployment*. IEEE International Conference in Communications (2015)
7. A. Kurniawan, I. Zakia, E. Wartika, A.G. Austin, in *Coverage and Capacity Trade-off Using Admission and Power Controlled CDMA System for Wireless Internet Services in Rural Environments, IEEE Conference*, 25–26 November 2015

RIS-Based Multiband Microstrip Patch Antenna Using Square and Giuseppe Peano Fractals

Vishal Upmanu, Ajay Kumar Yadav, Nitin Kathuria and Pradyot Kala

Abstract An investigation into the design of Giuseppe Peano fractal patch antenna and reactive impedance surface (RIS) is presented in this study. We investigate the properties and possibilities of application of RIS-based Giuseppe Peano fractal geometry for the antenna miniaturization, bandwidth increment, gain improvement, resonance frequency decrement and return loss decrement of microstrip patch antennas and compare its performance with those of the simple rectangular antenna. RIS has been realized using two-dimensional metallic patches printed on a metal-grounded substrate. The fractal antenna not only has a large effective length, but the contours of its shape can generate a capacitance or inductance that can help to match the antenna to the circuit. RIS is equal to the parallel combination of inductor and capacitor. RIS increases the equivalent value of inductor and capacitor; hence, the resonance frequency decrease, bandwidth increase and antenna radiation performance improved significantly.

Keywords Fractal antenna · Reactive impedance surface (RIS) Microstrip · Miniaturization · Patch

V. Upmanu (✉) · A. K. Yadav
Mewar University, Chittorgarh, Rajasthan, India
e-mail: vupmanu@gmail.com

A. K. Yadav
e-mail: ajay.ajaikr@gmail.com

N. Kathuria
Department of ECE, I.T.S Engineering College, Greater Noida, India
e-mail: nkathuria.niet@gmail.com

P. Kala
Department of ECE, Galgotia Engineering of College, Greater Noida, India

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_10

1 Introduction [1, 2]

Fractal geometry with electromagnetic meta-surfaces is a field of intense research with remarkable progress. There are various types of meta-surfaces available such as reactive impedance surface (RIS). It is composed of two-dimensional periodic metallic patches on a grounded substrate. The RIS is used to further decrease the resonance frequency and to improve the antenna radiation performance. The edge coupling of the square patch provides a shunt capacitor, while the short-circuited dielectric-loaded transmission line can be modelled as a shunt inductor. The circuit diagram of RIS is equivalent to the parallel combination of inductor and capacitor.

When it works as an inductive surface, it is able to store the magnetic energy because inductor stores the kinetic energy of moving electrons in the form of a magnetic field. They behave quite differently than resistors (which simply dissipate energy in the form of heat) in a circuit. Energy storage in an inductor is a function of the amount of current through it. When it works as a capacitive surface, separated by a dielectric,will be able to store the electric energy and increases the capacitive value for the patch type resonance. In order to enhance the 3-dB bandwidth of single layer, an RIS method is employed. The perfectly electric conductor PEC ground plane of RIS surface is inductive in nature with thickness, and RIS patches give the capacitive effect which forms a parallel combination, thus forming the LC parallel circuit. RIS is like an LC resonant circuit. So by varying the thickness parameter of the RIS substrate and optimizing other elements of the array, radiation properties of the antenna can be improved. The resonant frequency of the surface depends on the value of the capacitive elements, the distance between the capacitive layer and the metallic surface and the permittivity of the dielectric layer.

Fractal geometries are self-similar configurations. Mandelbrot first introduced the fractal geometry in 1975 [3], in which each subsection has the characteristics of the whole structure in a smaller scale. This is the basic property of self-similarity. Generally, the fractal geometries in antennas reduce their physical sizes and produce multiband response in their radiation characteristics. Fractal structures have a repetitive geometry; they generate long paths in a limited volume, so resonance frequency decreases. This self-similarity property of fractal geometry is used to get multiband operations, and their space-filling property is used for antenna miniaturization [4, 5]. Fractal shapes are used in radiating systems and microwave.

In this paper, Giuseppe Peano fractal geometry with RIS for the microstrip patch antenna is presented, and miniaturization with high bandwidth, high gain, high radiation performance, low resonance frequency, low return loss, etc., is obtained. In Sect. 2, fractal and why Giuseppe Peano fractal shape has taken is introduced. In Sect. 3, the RIS and its properties are detailed. In Sect. 4, Giuseppe Peano fractal microstrip multiband patch antenna with RIS is defined. In Sect. 5, performance of this fractal with RIS is compared to the other simple rectangular patch antennas with RIS and without RIS.

2 Giuseppe Peano Fractal Shape

Fractal is a repetitive or self-similar geometry which is obtained by cutting slots on the patches. Fractals have space-filling curve which is used for the antenna miniaturization, and this self-similar property of fractal is used to achieve multiband operation. Giuseppe Peano fractal geometry has better efficiency as compared to earlier fractal antenna designs such as Koch antenna, Tee-Type antenna, Sierpiski antenna, and it is quite stable and flat for the specified frequency band. Figure 1 gives a comparative analysis of radiation efficiency versus frequency band for Giuseppe Peano fractal geometry with other fractal geometries.

Giuseppe Peano fractal geometry for the microstrip antenna produces comparatively a larger 10 dB return loss (S11) bandwidth with lower number of iterations and also achieves better miniaturization than that of others, as can be observed in Fig. 2.

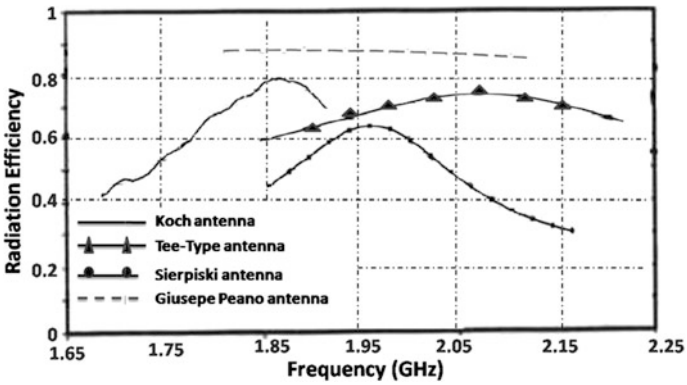

Fig. 1 Radiation efficiency of various fractal antennas

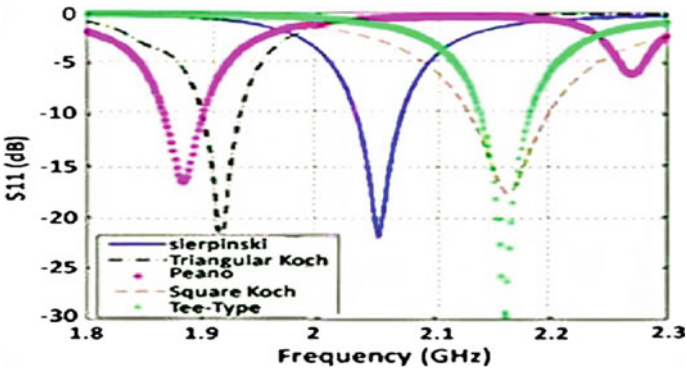

Fig. 2 Return loss (S11) versus frequency for various fractal geometries

3 Reactive Impedance Surface (RIS)

The RIS is a meta-surface composed by two-dimensional periodic metallic patches printed on a grounded substrate [6]. The edge coupling of the square patch provides a shunt capacitor, while the short-circuited dielectric-loaded transmission line can be modelled as a shunt inductor. The unit cell of this RIS structure is first designed between the perfectly electric and magnetic conductor, i.e. perfectly electric conductor (PEC) and perfectly magnetic conductor (PMC) boundary limits with an aim to achieve the best compromise between optimal bandwidth and miniaturization factor.

Hence, the equivalent circuit diagram of RIS is equal to the parallel combination of inductor and capacitor as shown in Fig. 3.

When RIS works as an inductive surface, it is able to store the magnetic energy and increase the inductance value of the patch-type resonance, and when it works as a capacitive surface, it is able to store the electric energy and increase the capacitive value of the patch-type resonance. Hence because of RIS, the equivalent value of inductor and capacitor increases and the resonance frequency slightly decreases. So the RIS is employed to decrease the resonance frequency and improve the antenna radiation performance. RIS minimizes the interaction between the antenna and its substrate, which simplifies matching and allows wideband operation. The impedance can be obtained as follows:

$$\begin{aligned} Z_{\text{ris}} &= \frac{X_L X_C}{X_L - X_C} \\ X_L &= jZ_d \tan kh_2 = jZ_d \tan k_0\sqrt{\varepsilon_r}h_2 \\ X_C &= \frac{1}{j\omega C_{\text{ris}}} \end{aligned} \tag{1}$$

where

Z_{ris} Reactive impedance of RIS
X_L Inductive impedance of RIS

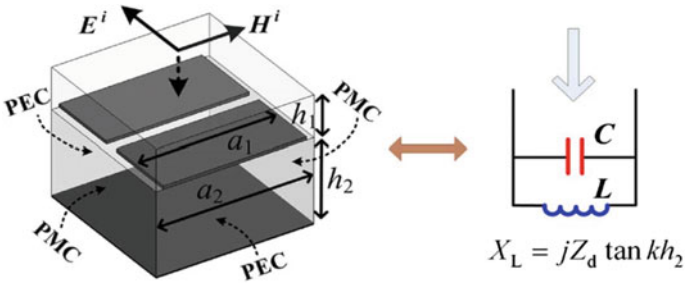

Fig. 3 Equivalent circuit of the RIS bounded with PEC and PMC walls

X_C Capacitive impedance of RIS
Z_d Impedance of substrate
ε_r Effective permeability
h_2 Substrate height of RIS.

The variation of the patch size and slot width mainly changes the capacitor value, while the substrate thickness and dielectric constant mainly affect the inductance value, all of which can be used to control the resonance frequency. Either an inductive RIS (below the PMC surface frequency) or a capacitive RIS (above the PMC surface frequency) can be obtained depending on the geometry and the operating frequency. The inductive RIS is capable to provide a wider matching bandwidth; therefore, it is more suitable for antenna application. Miniaturization is achieved by combining the reactive characteristic of the RIS with the capacitive and inductive behaviour of the antenna input impedance below its natural resonance in order to tune the resonance to a lower frequency. The resonant frequency is determined by

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \tag{2}$$

where L and C are the inductance and capacitance of the RIS equivalent circuit.

4 Proposed Antenna Design

The proposed antenna is "RIS-based Giuseppe Peano fractal microstrip multiband patch antenna" which is shown in Fig. 4. Design specifications of proposed antenna are shown in Table 1.

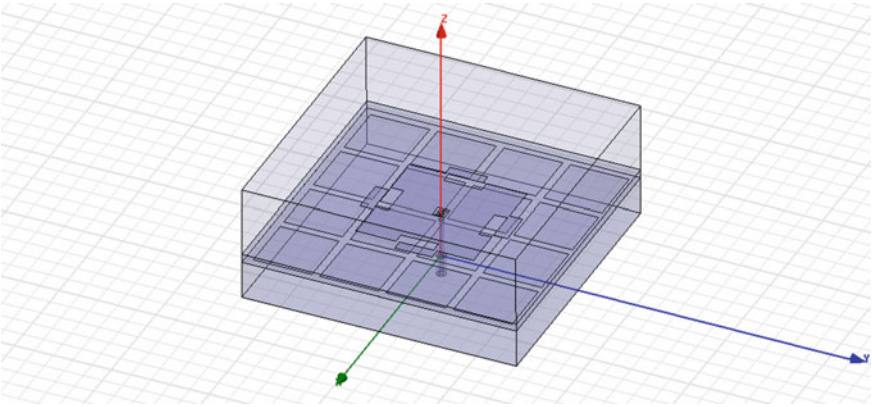

Fig. 4 Reported RIS-based Giuseppe Peano fractal microstrip multiband patch antenna

Table 1 Design specifications of RIS-based Giuseppe Peano fractal microstrip multiband patch antenna

	Length (mm)	Width (mm)	Height (mm)	Spacing (mm)	Material ($€_r$)
Ground	70	70	0	–	–
Lower substrate	70	70	10	–	Silicon nitrate, 7
Upper substrate	70	70	12	–	Roger RO 3003 (tm), 3
RIS	15	15	10	2	–
Patch	30	30	12	–	Copper
Cover	70	70	30	–	Air, 1

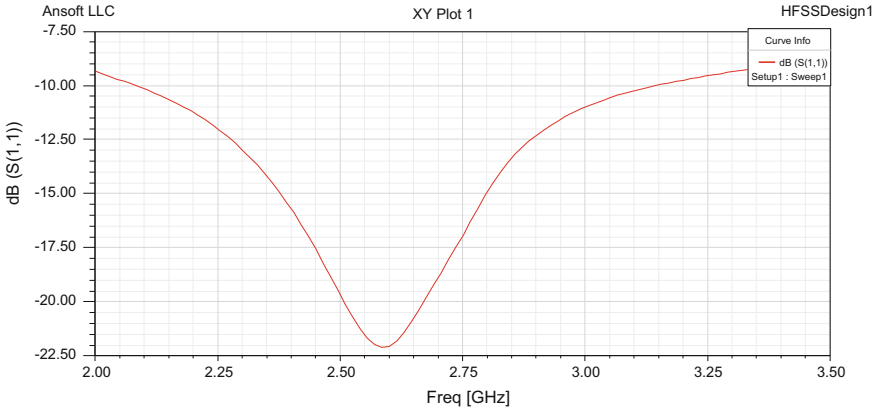

Fig. 5 Return loss for RIS-based Giuseppe Peano fractal microstrip multiband patch antenna of HFSS simulation

We can increase the bandwidth, gain and decrease the return loss of the Giuseppe Peano fractal antenna by using RIS. RIS is placed on the grounded substrate shown in Fig. 4. The curve between S11 and frequency is obtained by using high-frequency structural simulator (HFSS) simulation's software which works on the principle of finite element method (FEM), and measurement data are drawn in Fig. 5. This structure resonates at the frequency 2.6 GHz. It has a bandwidth of 1600 MHz and return loss = −22 dB at 2.6 GHz. Radiation patterns in Fig. 6 give the directional property of the proposed antenna. Figure 7 gives the voltage standing wave ratio (VSWR) this fractal antenna versus frequency The antenna fed is through a 50 Ω coaxial cable. The inductive effect of the probe makes the impedance matching difficult. To compensate this effect, a square slot is cut around the feed point, where the inner and outer squares have sides 2 and 3 mm, respectively. The capacitance of this square slot compensates for the inductance of the probe.

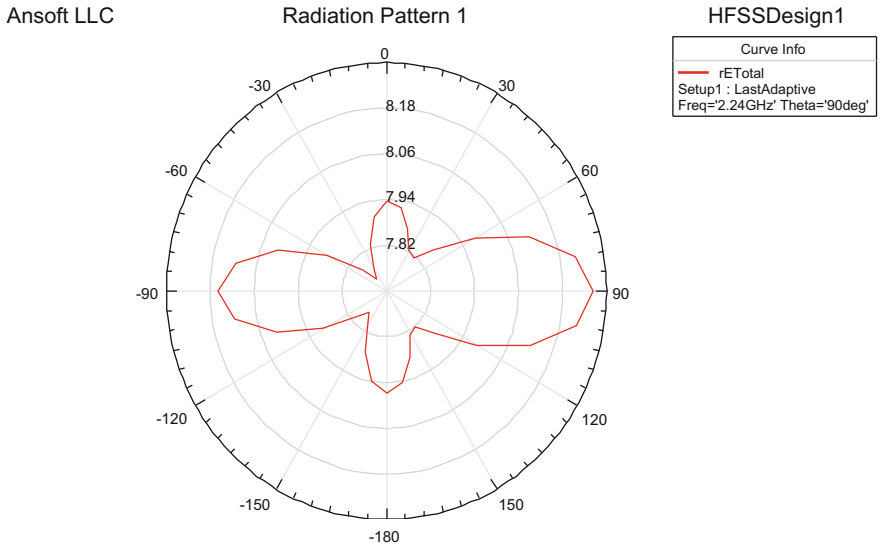

Fig. 6 Radiation pattern of proposed antenna by the result of HFSS

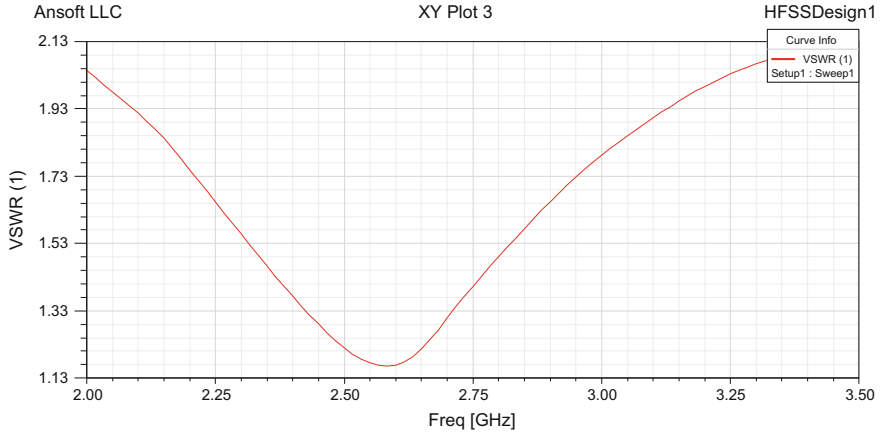

Fig. 7 VSWR for RIS-based Giuseppe Peano fractal microstrip multiband patch antenna

5 Comparison of Proposed Antenna with Previous Designs

The proposed antenna is RIS-based Giuseppe Peano fractal microstrip multiband patch antenna. Simulation results for the design are compared with Giuseppe Peano fractal microstrip multiband patch antenna without RIS, and simple square patch

with and without RIS has been compared for the designed specifications with the simulated radiation parameters observed.

5.1 Giuseppe Peano Fractal Microstrip Multiband Patch Antenna Without RIS

The design as shown in Fig. 8 is proposed by Oraizi and Hedayati [7]. The presented design was simulated, and results through HFSS were measured. The antenna resonates at the centre frequency 2.9 GHz with the bandwidth of 706 MHz and return loss of −15.2 dB. The design specifications are shown in Table 2 with the return loss as shown in Fig. 9.

5.2 Simple Square Patch Without RIS and with RIS

Simple square patch antennas with RIS and without RIS were simulated using HFSS design tool. Dimensions of the square patch are 30 mm × 30 mm at 12 mm height for both antenna structures (with RIS and without RIS). The structures are shown in Fig. 10.

For the antenna without RIS, bandwidth is 848 MHz and return loss of −18 dB at centre frequency of 3.25 GHz as shown in Fig. 11. With RIS, structure resonates at the frequency 2.56 GHz and has a bandwidth of 1000 MHz with return loss = −20 dB as shown in Fig. 12.

A parametric comparison of bandwidth, return loss and resonating frequency of the proposed antenna with existing designs is presented in Table 3.

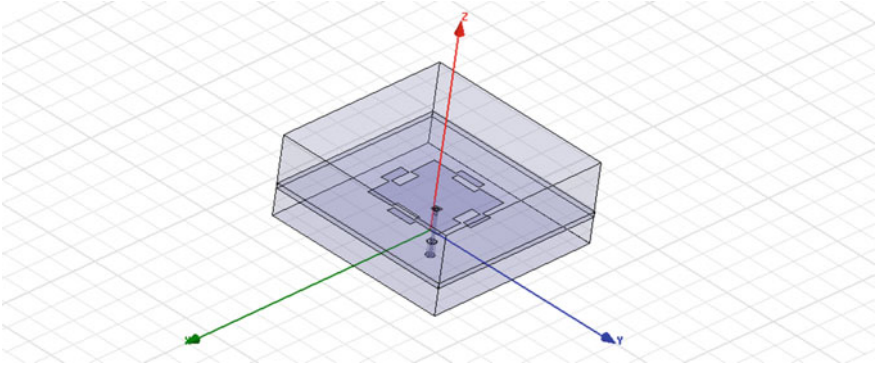

Fig. 8 Giuseppe Peano fractal microstrip multiband patch antenna without RIS

Table 2 Design specifications of Giuseppe Peano fractal microstrip multiband patch antenna without RIS

	Length (mm)	Width (mm)	Height (mm)	Material (ϵ_r)
Ground	70	70	0	–
Lower substrate	70	70	10	Air, 1
Upper substrate	70	70	12	FR4, 4.4
Patch	30	30	12	Copper
Cover	70	70	30	Air, 1

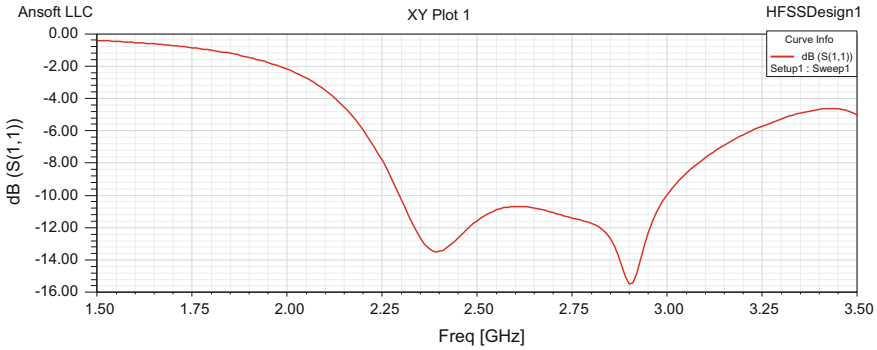

Fig. 9 Return loss for Giuseppe Peano fractal microstrip multiband patch antenna without RIS

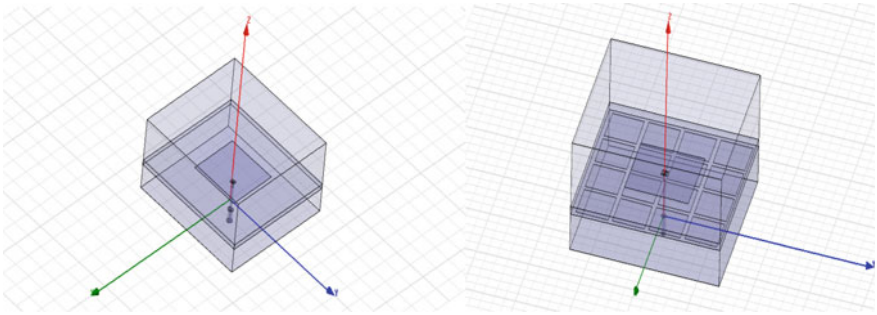

Fig. 10 Simple square patch antenna without RIS and with RIS

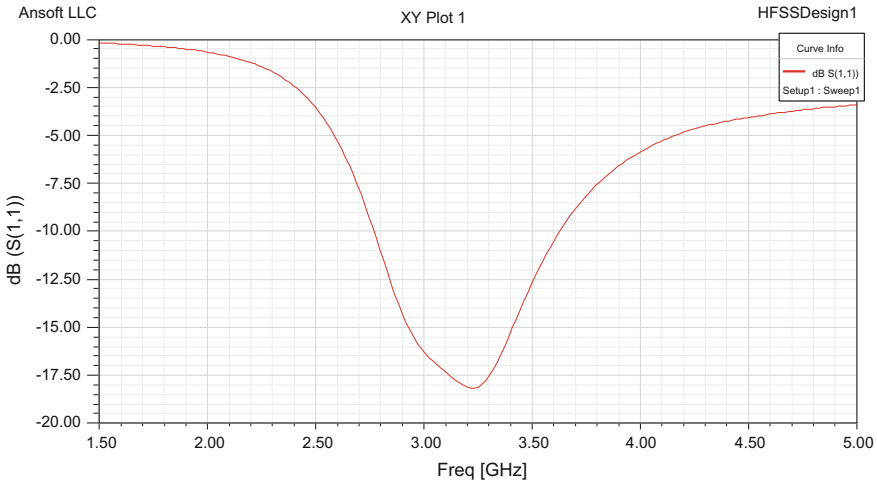

Fig. 11 Return loss of simple square patch without RIS

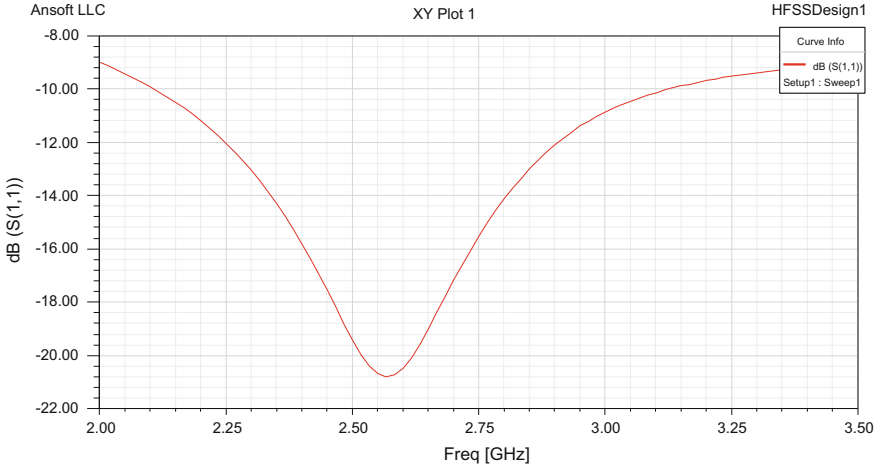

Fig. 12 Return loss of simple square patch with RIS

Table 3 Variations of bandwidth, return loss and resonance frequency for four antennas

Antenna type	Bandwidth (MHz)	Return loss (dB)	Resonating frequency (GHz)
G.P. Fractal antenna (without RIS)	706	−15.2	2.9
G.P. Fractal antenna (with RIS) (Proposed)	1600	−22	2.6
Simple rectangular (without RIS)	848	−18	3.25
Simple rectangular (with RIS)	100	−20	2.56

6 Conclusion

Proposed RIS-based Giuseppe Peano fractal microstrip multiband patch antenna has been compared with the similar model without RIS and simple square patch with and without RIS. The use of the meta-surface enables the antenna miniaturization. Implementation of RIS slightly decreases the resonance frequency whereas improves the antenna radiation performance. It also increases −10 dB bandwidth and reduces the reflection coefficient. As the proposed antenna operates with high bandwidth and less return loss (S11) for the same resonating frequency range makes it a good candidate for WiMAX applications

References

1. C.A Balanis, *Antenna Theory Analysis and Design*. (Wiley, New York, 1997)
2. I.J. Bahal, P. Bhartia, *Microstrip Antennas* (Artech House, Dedham, MA, 1980)
3. D.H. Werner, R. Mittra, *Frontiers in Electromagnetics* (IEEE Press, Piscataway, NJ, 2000), pp. 48–81
4. D.H. Werner, S. Ganguly, An overview of fractal antenna engineering research. IEEE Antennas Propag. Mag. **45**(1), 23–29 (2003)
5. J.P. Gianvittori, Y. Rahmat-Samii, Fractal antenna: a novel antenna miniaturization technique, and applications. IEEE Antennas Propag. Mag. **44**(1), 20–36 (2002)
6. H. Mosallaei, K. Sarabandi, Antenna miniaturization and bandwidth enhancement using a reactive impedance substrate. IEEE Trans. Antennas Propag. **52**(9), 2403–2414 (2004)
7. H. Oraiz, S. Hedayati, Miniaturization of multiband antenna by the novel application of the giuseppe piano fractal geometries. IEEE Trans. Antennas Propag. **60**(8) (2012)

Spectrum Sensing Techniques for a Cognitive Radio Network

Bhupesh Aneja, Kanchan Sharma and Amita Rana

Abstract This paper carries out a detailed review of the various techniques that have emerged in spectrum sensing in the recent years in cognitive radio (CR). It discusses the current innovations in techniques and algorithms, and within this context, the operation of an intelligent wireless communication system using CR technology has been reviewed. The important techniques used in past are also tabulated. Various technical papers and articles on wireless communication and CR systems have been analyzed. The paper takes an inquisitive approach to the proposals and prototypes of an effective spectrum sensing technique obtained using various algorithms, which can be used to eliminate the spectrum scarcity and overcrowdings. After reviewed various techniques, we work on data rate improvement and suggest a method of improving the quality by maintaining the BER at an average level by adaptive modulation.

Keywords Cognitive radio (CR) · Spectrum sensing (SS) · Wireless network Primary user (PU) · Secondary user (SU)

1 Introduction

Wireless communication is developing rapidly due to increase in demand for service quality and data rates. Emergence of new communication schemes has helped fulfill these demands, but it has also led to scarcity of available spectrum. A recent study of spectrum consumption has shown that spectrum which actually licensed is

B. Aneja
Department of ICE, JSSATE, Noida, India
e-mail: er_bhupesh@jssaten.ac.in

K. Sharma · A. Rana (✉)
Department of ECE, IGDTUW, Delhi, India
e-mail: amita.rinky@gmail.com

K. Sharma
e-mail: joinkanchansharma@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_11

not used in wide geographic and temporal areas. A promising solution to this problem is the development of smart wireless system that is capable enough to detect its surroundings and change its method of operation accordingly like cognitive radio.

In the frequency spectrum system, certain users have been authorized to occupy the frequency band; such high priority users are known as licensed user or primary users. However, the Secondary users or Cognitive users are authorized to use the idle spectrum without interrupting the Primary user. The CR enables the SU to share the authorized spectrum resources with the PU, when they are idle.

The key activity of a cognitive radio is to monitor dynamically the spectrum availability over the spectrum band. The radio comprises of users which are secondary users. These are the users which have not been licensed that band in the spectrum. They continuously look for changes in the spectrum bands of primary users (licensed users) for transmission when they are not in use. This is collectively known as spectrum management. This task can be accomplished by the secondary user with the process of spectrum sensing. Spectrum sensing deals with searching for spectrum holes. The concept of spectrum hole is shown in Fig. 1, [1].

The main work of CR is spectrum sensing, i.e., detection of spectrum holes. A spectrum hole is a band of frequencies which has been primarily assigned to a PU but is not being used by that user at a particular time.

SS enables the SU to sense the existence of PU for an interference-free spectrum access. Sect. 2 explains the various types of spectrum sensing method. Sect. 3 consists of an overview of contributions of various authors. Section 4 presents the proposed model for adaptive modulation for BER improvement of transmission.

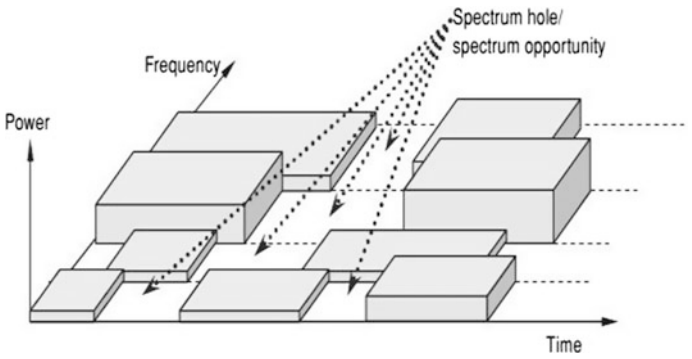

Fig. 1 Concept of spectrum hole

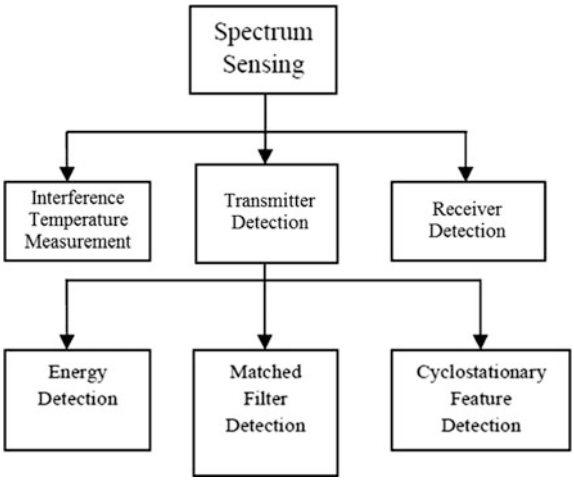

Fig. 2 Classification of spectrum sensing techniques

2 Spectrum Sensing

SS is the important step which is to be used in order to give SU opportunities for transmission. The developed algorithm should be efficient enough to identify the changes in the moving holes for the whole spectrum, at the same time maintaining the quality of service without interfering with other users.

SS algorithm depends on parameters like signal-to-noise ratio (SNR), number of samples and noise uncertainty and their effect on probability of detection and probability of false alarm. Figure 2 shows the classification of various SS techniques.

2.1 Interference Temperature Measurement

This approach involves CR working with ultra-wideband (UWB) technology where a state of coexistence of licensed users and cognitive users occurs. Cognitive user signals are transmitted with low power such that they do not interrupt with the licensed users.

2.2 Receiver Detection

In this method, licensed receiver's local oscillator leakage power is used to sense spectrum characteristics and interference.

2.3 Transmitter Detection

In this method, primary user is detected on the basis of signal received by secondary user. Different methods for this approach are as follows:

Matched Filter Detection It is a filter that is used to increase the SNR of the input signal to an optimal level. Figure 3 shows the basic diagram for a matched filter. The drawback of this method is that it requires some prior knowledge of the primary user, which is almost impossible to achieve. However in this method comparitively less detection time is required as it has only (1/SNR) samples to satisfy the given constraint for detection of probability. It can also distinguish between the users sharing the same channel and can also work in low SNR.

Expression for matched filter detection is given by Eq. (1)

$$Y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] \tag{1}$$

where 'x' is unknown signal, 'h' is impulse response of the matched filter and 'Y' is the output signal. 'x' and 'h' is convoluted to maximize the SNR.

Energy detector In this method, a band-limited signal of bandwidth W is sampled and its energy is calculated in order to compare it to the reference threshold. If the energy is above the threshold, the primary user is present, or else absent. Figure 4 represents the working of an energy detector with the help of a block diagram. This method has low computational cost, and it does not need any former information, and it does not work in low SNR values also cannot differentiate between the users using the same channel.

Fig. 3 Block diagram of matched filter

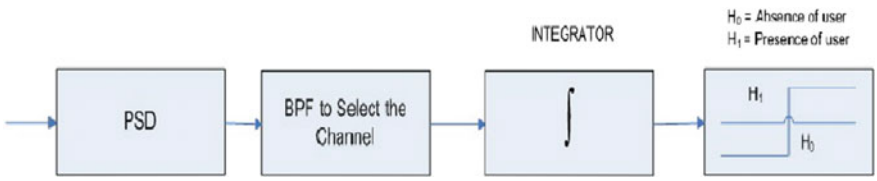

Fig. 4 Block diagram of energy detection

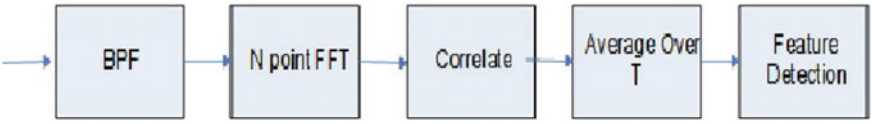

Fig. 5 Block diagram of cyclostationary feature detection

Expression for energy detection analysis is given by Eqs. (2) and (3):

$$y(k) = n(k); H0 \tag{2}$$

$$y(k) = h * s(k) + n(k); H1 \tag{3}$$

where $y(k)$ represents the sample to be analyzed at every moment k, $n(k)$ is the noise of variance, and $s(k)$ represents the sampled signal. $H0$ and $H1$ indicate the nonexistence and existence of user, respectively.

The decision rule at the detector can be stated by Eqs. (4) and (5):

$$H0 : y(k) = n(k) \ldots PU \text{ Absent} \tag{4}$$

$$H1 : y(k) = x(k) + n(k) \ldots PU \text{ Present} \tag{5}$$

Cyclostationary Feature Detection In this method, the primary user is detected by the use of periodicity of the received primary signal. Figure 5 diagrammatically represents cyclostationary feature detection. This technique has a high computational cost, but it is robust to perform in the low SNR region and can also interpret some exclusive features of detected signal. Due to complex computation of this method, it is unable to detect and search real-time voids in the wireless spectrum.

3 Overview of Contributions

Table 1 lists the various contributions of the authors that has been considered and reviewed throughout the study of the topic.

4 Our Contribution

4.1 Proposed Model for Adaptive Modulation for BER

The objective of adaptive modulation here in our work is to suggest a method to improve the data transmission by SU in the CR network. The work focuses on improving the data rate and also the quality of data transmission. The model is

Table 1 List of contribution by various authors

S. No	Paper title	Authors	Outcomes
1.	Ref. [2]	Wei Z., Ben L.K.	Half-voting rule was used to reduce the total error probability. A method for calculating optimal threshold was also proposed
2.	Ref. [3]	Ling-ling Z., Cheng-kai T., Jian-guo H.	The study is done on this fact that noise variation leads to detect error. In order to reduce detection error, detection threshold was calculated using noise uncertainty
3.	Ref. [4]	Joshi M., Borde S.D.	The association of energy detection with SNR, noise uncertainty, and the number of samples is studied. Graphically, it was summarized that as SNR increases the performance of energy detector scheme increases
4.	Ref. [5]	Verma P., Singh B.	A scheme in which the fusion center averages the observed energy values so that it can be compared with a threshold value proposed
5.	Ref. [6]	Waleed E.	A minimizing sensing time algorithm based on fuzzy logic was proposed that gives acceptable results at higher SNR. This proposal was found to be time-consuming but gives more accurate results
6.	Ref. [7]	Dusadee T., Dimitrie C.	The use of hidden Markov model (HMM) accurately estimates the energy detector performance in terms of false alarm and detection probability
7.	Ref. [8]	Iglesias E.R.	A machine learning algorithm which learns from the past and present values was implemented in order to achieve a real CR environment
8.	Ref. [9]	Ying-Chang L., Zeng Y., Peh Edward C.Y., Hoang A.T.	The paper uses medium access control (MAC) frame structure design to describe the sensing throughput tradeoff problem by considering the interests of both the users

(continued)

Table 1 (continued)

S. No	Paper title	Authors	Outcomes
9.	Ref. [10]	Sutton P.D., Nolan K.E., Doyle L.E.	The use of cyclostationary signatures is done for signal detection and network identification by using Dynamic Spectrum Access. The paper focuses on OFDM based systems
10.	Ref. [11]	Tevfik Y., Huseyin A.	The paper makes use of SS and spectrum opportunity concepts in different dimensions of spectrum space
11.	Ref. [12]	MacKenzie A.B., Athanas P., Raman S.	The study of cognitive systems by means of game theory and purpose of cognitive techniques was considered to decrease the problems in dynamic SS
12.	Ref. [13]	Shobana S., Saravanan R., Muthaiah R.	The paper establishes that sensing accuracy is the most significant factors of SS for CR network
13.	Ref. [14]	Wang B., Yongle W., Liu K. J. Ray	A classification of different state-of-the-art game theories and their contributions on different CR networks were made
14.	Ref. [15]	Cabric D., Mishra S.M., Brodersen R.W.	The paper establishes that cyclostationary feature is advantageous because it can differentiate between signals in the low SNR region
15.	Ref. [16]	Qing Z., Sadler B.M.	The paper provides an outline of key technical and regulatory issues in opportunistic sensing access (OSA)
16.	Ref. [17]	Ghasemi A., Sousa E.S.	The paper quantifies the essential limits of SS in fading environments. It also investigates the capacity of fading networks when power is received at a third-party receiver
17.	Ref. [18]	Ganesan G., Li Y.	The benefits of cooperation are shown by increasing the agility of CR networks. Non-cooperative and totally cooperative schemes are employed on different degrees of cooperation
18.	Ref. [19]	Stevenson C.R., Hu W., Caldwell W.	The IEEE 802.22 operates on a non-interfering basis in the VHF/UHF TV broadcast bands which lead to efficient consumption of the spectrum

(continued)

Table 1 (continued)

S. No	Paper title	Authors	Outcomes
19.	Ref. [20]	Ghasemi A., Sousa E.S.	Performance limitations and various other problems related to SS design are investigated which can be overcome by combining local signal processing, user-level cooperation, and system-level coordination among CRs
20.	Ref. [21]	Farag H.M., Ehab M.	This paper proposes an efficient energy detection algorithm for CR SS. This algorithm predicts the LU's current activity with the help of the stored succeeding energy measurements
21.	Ref. [22]	Li G., Fang J.	On increasing the values of SNR, detection performance can be improved significantly for a certain threshold value. It is also a function of filter bandwidth and detection time

based on fuzzy inference engine. This engine works on If-Then rules and takes decision on the basis of these rules. This model can be used to suggest a way in which the data transmission can be improved. During the transmission, some of the bits received are erroneous. The ratio of erroneous bit to the total number of bits sent is given by a ratio called the bit-error rate (BER). This bit-error rate for any transmission scheme should be under the considerable limits. It is a probabilistic measure of the error rate and hence lies between 0 and 1. The system proposed here senses the channel fluctuation, previous rate, and the current BER to decide whether which modulation scheme has to be chosen for transmission to reduce the errors and improve the transmission.

The philosophy of this model is that we try to adaptively select the digital modulation scheme by which the user should transmit data so as to leave the channel as soon as possible and also reduce the error. This is the improvement in transmission which is intended in this work. Four types of digital modulation schemes have been taken, namely BPSK, 4-QAM, 16-QAM, and 64-QAM. Each modulation scheme has a different expression of BER and due to which they have a different BER v/s SNR curves for the channel as shown in Fig. 6.

First input parameter is the current SNR of the channel as shown in Fig. 7. We will assume that the cognitive radio network determines this channel property. This particular property keeps on changing, and we have made our transmission system of the cognitive radio capable of adjusting to this fluctuation. The SNR value ranges from 0 to 30 dB. This input shown in Fig. 8 has four MBFs. The next input parameter is BER which the current bit-error rate generated by the user in its present modulation scheme. This input is given and fuzzified under 2 MBFS. BER value

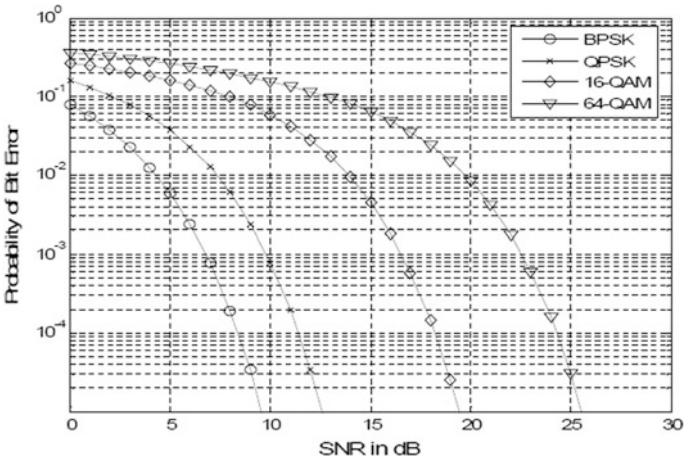

Fig. 6 BER versus SNR curves for various modulation schemes

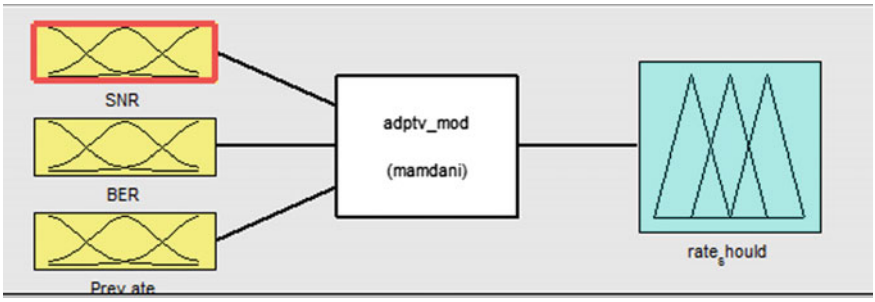

Fig. 7 Model of adaptive modulation

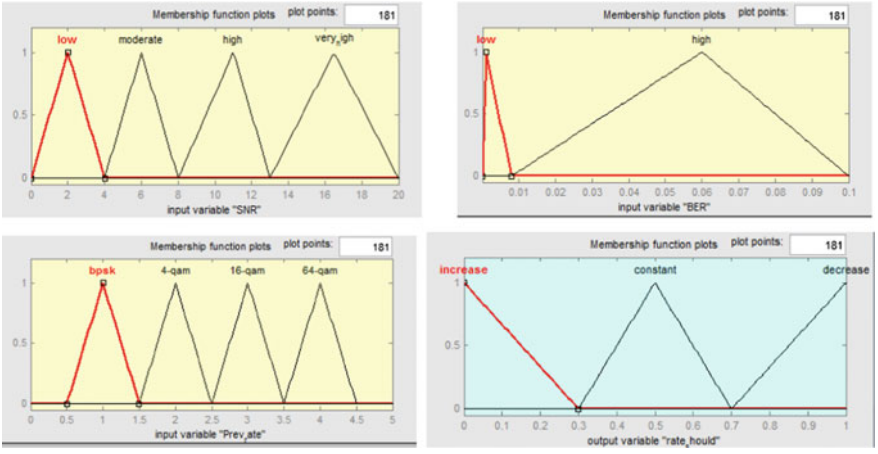

Fig. 8 Membership function of SNR, BER, previous rate and output

Table 2 If-Then rules for the adaptive modulation model

S.No.	SNR	BER	Prev_rate	Output
1.	L	L	BPSK	Rate should be constant
2.	L	L	4-QAM	Rate should be constant
3.	L	H	4-QAM	Rate should be decrease
4.	L	H	16-QAM	Rate should be decrease
5.	L	H	64-QAM	Rate should be decrease
6.	M	L	4-QAM	Rate should be constant
7.	M	L	16-QAM	Rate should be constant
8.	M	H	4-QAM	Rate should be increase
9.	M	H	16-QAM	Rate should be decrease
10.	M	H	64-QAM	Rate should be decrease
11.	H	L	16-QAM	Rate should be constant
12.	H	L	4-QAM	Rate should be constant
13.	H	L	64-QAM	Rate should be constant
14.	H	H	4-QAM	Rate should be increase
15.	H	H	BPSK	Rate should be increase
16.	H	H	16-QAM	Rate should be increase
17.	H	L	64-QAM	Rate should be constant
18.	H	L	16-QAM	Rate should be constant
19.	H	H	64-QAM	Rate should be decrease
20.	H	H	16-QAM	Rate should be increase
21.	H	H	4-QAM	Rate should be increase
22.	H	H	BPSK	Rate should be increase
23.	H	H	BPSK	Rate should be increase
24.	H	H	4-QAM	Rate should be increase

L low, *M* moderate, *H* high

greater than 0.008 is considered to be high. The third input parameter is the current modulation scheme or data rate. A rule set for appropriate decision making is fed to the fuzzy inference engine. Table 2 shows 24 If-Then rules have been made on the observations. The output for the system is a value which determines whether the current rate has to be increased, decreased, or should remain constant. This tends to improve the transmission quality in terms of BER.

5 Conclusion

Various papers have been analyzed, and the spectrum sensing algorithms for cognitive radio networks have been studied effectively. A comparative evaluation of the papers and their proposed technologies has been tabulated, and the outcomes of each paper have been drawn out. We work on data rate improvement and suggest

Table 3 Output generated by adaptive modulation system for various input sets

S.No	SNR	BER	Current rate	Output
1.	5	0.05	3 (16-QAM)	0.8859 (Decreases rate)
2.	10	0.08	3(16-QAM)	0.1141 (Increase rate)
3.	2	0.0064	1 (BPSK)	0.5 (Keep rate constant)

a method for improving the quality by maintaining the BER at an average level by adaptive modulation. This model works as a feedback system by continuously sensing the change of SNR of the channel and suggesting whether we should change the modulation order and data rate. This helps the system to maintain the BER to an average level. The limit of BER is this system is 0.008. After this value, the BER is considered to be high and the modulation order is either increased or decreased so as to maintain the BER under this value. When BER is below, this value it is said to be low and the modulation order has to remain constant. Table 3 shows the few results. The increase or decrease of the order of modulation has to be taken in steps. This is the reason the output is not specifying the actual modulation order to bring the BER within limits. It hence tells the user to decrease or increase it step by step.

Hence, we see that on increasing or decreasing the rate, we can keep the BER under the specified limit and hence improve the channel performance.

References

1. Y.H. Wang, S.L. Liao, J.L. Chang, in *A Fuzzy—Based Dynamic Channel Allocation Scheme in Cognitive Radio Networks*. 8th International Conference on Ubi-Media Computing (UMEDIA), pp. 049–054, 2015
2. K. Letaief, W. Zhang, Cooperative communications for CR networks. Proc. IEEE **97**(5), 878–893 (2009)
3. Z. Ling-ling, H. Jian-guo, T. Cheng-kai, in *Novel Energy Detection Scheme in CR*. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1–4, 14–16 Sept 2011
4. M. Joshi, S.D. Borde, in *Comprehensive Analysis of Various Energy Detection Parameters in SS for CR Systems*. International Conference on Advances in Communication and Computing Technologies (ICACACT), pp. 1–4, 10–11 Aug 2014
5. P. Verma, B. Singh, in *Simulation Study of Double Threshold Energy Detection Method for CRs*. 2nd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 232–236, 19–20 Feb 2015
6. W. Ejaz, *SS in CR Networks*. (National University of Science and Technology, Pakistan)
7. D. Treeumnuk, D.C. Popescu, in *Using Hidden Markov Models to Enable Performance Awareness and Noise Variance Estimation for Energy Detection in CR*. 46th Annual Conference on Information Sciences and Systems (CISS), pp. 1–5, 21–23 Mar 2012

8. E.R. Iglesias, Cognitive-radio and antenna functionality. IEEE Antenna Propog. Mag. **56**(1) (2014)
9. Y.C. Liang, Y. Zeng, E.C.Y. Peh, A.T. Hoang, Sensing-throughput tradeoff for CR networks. IEEE Trans. Wireless Commun. **7**(4), 1326–1337 (2008)
10. P.D. Sutton, K.E. Nolan, L.E. Doyle, Cyclostationary signatures in practical CR applications. IEEE J. Sel. Areas Commun. **26**(1), 13–24 (2008)
11. T. Yucek, H. Arslan, A survey of SS algorithms for CR applications. IEEE Commun. Surv. Tutorials **11**(1), 116–130 (2009)
12. A.B. MacKenzie, J.H. Reed, P. Athanas, C.W. Bostian, R.M. Buehrer, L.A. DaSilva, S.W. Ellingson, Y.T. Hou, M. Hsiao, J.M. Park, C. Patterson, S. Raman, C. da Silva, CR and Networking research at virginia Tech. Proc. IEEE **97**(4), 660–688 (2009)
13. S. Shobana, R. Saravanan, R. Muthaiah, Matched Filter based SS on CR for OFDM WLANs. IJET **5**(1), 142 (2013)
14. B.Wang, W. Yongle, K.J.R. Liu, Game theory for CR networks: an Overview. Comput. Networks **54**(14), 2537–2561 (2010)
15. D. Cabric, S.M. Mishra, R.W. Brodersen, in *Implementation issues in SS for CRs*. Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, vol. 1, pp. 772–776, 7–10 Nov 2004
16. Q. Zhao, B.M. Sadler, A survey of dynamic spectrum access. Sig. Process. Mag. IEEE **24**(3), 79–89 (2007)
17. A. Ghasemi, E.S. Sousa, Fundamental limits of spectrum-sharing in fading environments. IEEE Trans. Wireless Commun. **6**(2), 649–658 (2007)
18. G. Ganesan, Y. Li, Cooperative SS in CR, Part I: two user networks. IEEE Trans. Wireless Commun. **6**(6), 2204–2213 (2007)
19. C. Stevenson, G. Chouinard, L. Zhongding, H. Wendong, S.J. Shellhammer, W. Caldwell, IEEE 802.22: The first CR wireless regional area network standard. Commun. Mag. IEEE **47** (1), 130–138 (2009)
20. A. Ghasemi, E.S. Sousa, SS in CR networks: requirements, challenges and design trade-offs. Commun. Mag. IEEE **46**(4), 32–39 (2008)
21. H.M. Farag, M. Ehab, in *An Efficient Dynamic Thresholds Energy Detection Technique for CR SS*. 10th International Computer Engineering Conference (ICENCO), pp. 139–144, 29–30 Dec 2014
22. G. Li, J. Fang, H. Tan, J. Li, in *The Impact of Time-Bandwidth Product on the Energy Detection in the CR*. 3rd IEEE International Conference on IC-BNMT, pp. 634–638, 26–28 Oct 2010

Part III
System Optimization

Implementation of ANFIS Controller-Based Algorithm Measuring Speed to Eliminate RDC Hardware in Resolver-Based PMSM

Prerna Gaur, Bhim Singh, A. P. Mittal and Seema Arora

Abstract Resolver-to-digital converter (RDC) converts resolver analog signals to analogous digital values representing the angular shaft position. The disadvantages of the circuit are limited noise rejection, poor resolution, and higher cost. An AI-based algorithm is developed to process the signals received from resolver sensor to extract position and speed and is compared with the existing PI-based algorithm. This algorithm developed in MATLAB minimizes the error between actual rotor angle and computed angle using a controller in a feedback loop. The problems caused by hardware can be eliminated by implementing the algorithm in Simulink. As real-time applications of such algorithms need to be very precise, the error is further reduced by replacing the PI controller with an adaptive neuro-fuzzy inference system (ANFIS)-based AI controller. The real-time implementation of the algorithm is analyzed by using dSPACE 1104 as the hardware interface. The developed control technique is beneficial for systems used in precise control techniques, and it finds applications, particularly in the textile, aerospace, and the automobile sectors.

Keywords dSPACE · AI · ANFIS · Resolver-to-digital converter
Real-time interface · DS1104

P. Gaur (✉) · A. P. Mittal
Instrumentation and Control Engineering Division, NSIT, Sec-3, Dwarka 110078, New Delhi, India
e-mail: prernagaur@yahoo.com

A. P. Mittal
e-mail: mittalap@gmail.com

B. Singh
Department of Electrical Engineering, IIT Delhi, Hauz Khas 110016, New Delhi, India
e-mail: bsingh@ee.iitd.ac.in

S. Arora
Waljat College of Applied Sciences, Muscat, Oman
e-mail: aroraseema1@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_12

1 Introduction

PMSM has found wide applications in high-performance applications such as industrial robots, satellite, aerospace actuators, and machine tools, due to their high power density and high torque-to-inertia ratio. The vector control PMSM drive is popular in the area where precision in speed control is highly desirable. The highly accurate position sensor is required in the vector control algorithm. Resolver is used as precision sensor that can measure the initial position at standstill. This feature is important to gain the maximum starting torque in such drive systems [1, 2].

Resolver signals are not degraded through long cables; hence, it may be used in severe atmospheric conditions such as humidity and high temperature. The resolver has one rotating coil as the primary winding and two stationary coils as the secondary winding. The primary coil is supplied by carrier wave of high-frequency voltage in order of kHz. Two secondary coils are placed in stator at 90° apart from each other. Once the rotor is rotated, the induced voltages would be produced in secondary coil. The amplitudes of these sinusoidal voltages are modulated with rotor position [3]. The demodulation must be designed to extract the rotor position. This demodulation process could be implemented on R/D converter IC [4–11].

Resolver-to-digital converter has many disadvantages such as relatively poor digital resolution, particularly at high speeds, and limited tracking error tolerance [5]. The phase relationship between the modulating source and the modulated feedback to provide the sine and cosine values is critical and often requires special cables. Also, such systems have limited noise rejection [1, 2, 5]. The implementation of resolver-to-digital converter hardware circuit increases the overall cost of the drive system. The rotor angle is computed by means of the feedback loop control which is presented in [2]. The error signal is introduced by the difference of the cross product of modulated signals and measured excitation signal. Then, this error would be controlled by PI controller.

The objective is to implement artificial intelligence (AI)-based software algorithm which tracks position of the motor from signals received from the resolver and thereby calculates speed from the tracked position. The AI-based controller as compared to PI controller further eliminates the delay and noise effects introduced by hardware circuit and can be readily implemented in real time with minimum complexity and cost [1, 2, 11–16].

The algorithm used implements PI- and AI-based controllers separately to present a comparative study to establish supremacy of AI-based controller. AI logic is a rule-based approach that does not require the knowledge of mathematical model of the system for designing the controller. Adaptive neuro-fuzzy inference system (ANFIS)-based speed controller realizes a good dynamic behavior of the motor, a perfect speed tracking with no overshoot, and a good rejection of impact load disturbance [13–15]. The final stage involves interfacing the computer with the hardware through dSPACE, to implement control techniques [10]. The analog signals from the motor are sent to dSPACE through the ADC channels provided on the DS 1104 R&D dSPACE controller board.

2 Mathematical Modeling of PMSM and RDC

A. *Mathematical Modeling of PMSM*

Stator voltages in *d*–*q* reference frame of the PMSM in the rotor reference frame are as [11]

$$v_q = Ri_q + p\lambda_q + \omega_s\lambda_d \tag{1}$$

$$v_d = Ri_d + p\lambda_d - \omega_s\lambda_q \tag{2}$$

The flux linkages are as

$$\lambda_q = L_q i_q \tag{3}$$

$$\lambda_d = L_d i_d + \lambda_{af} \tag{4}$$

where v_d and v_q are the *d*-, *q*-axis voltages, i_q and i_q are the *d*-, *q*-axis stator currents, L_d and L_q are the *d*-, *q*-axis inductances, λ_d and λ_q are the *d*-, *q*-axis stator flux linkages, while R and ω_s are the stator resistances per phase and inverter fundamental frequency, respectively. λ_f is the flux linkage due to the rotor magnets linking the stator. The electromagnetic torque developed is as

$$T_e = 3p\{\lambda_f i_q + (L_d - L_q) i_d i_q\}/2, \tag{5}$$

Figure 1 shows the Simulink model of PMSM to be used with dSPACE. Under subsystem block of dSPACE-based model, there is PWM generator block of PMSM drive system. The input to the dSPACE1104SL_PWM blocks is error between measured currents and reference currents $i_a - i_{ar}, i_b - i_{br}, i_c - i_{cr}$, where i_a and i_b are the outputs of ADC channels of dSPACE 1104 and i_c is calculated using $i_a + i_b + i_c = 0$. i_{ar}, i_{br} and i_{cr} are three-phase reference currents.

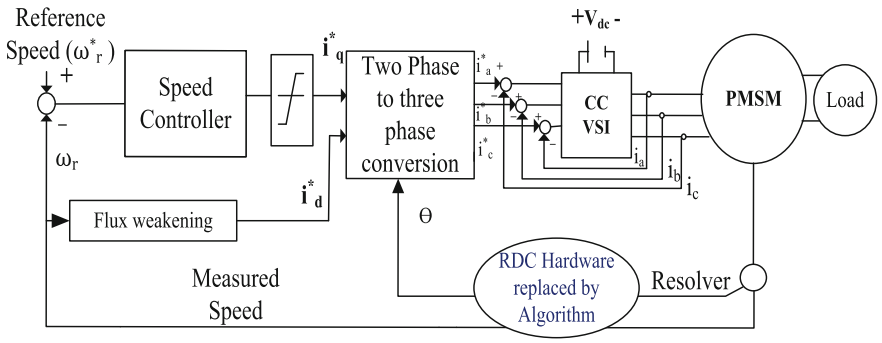

Fig. 1 Line diagram of PMSM with algorithm in place of RDC

B. *Mathematical Modeling of Resolver-to-Digital Conversion*

Resolver is a mechanical position sensor measuring the instantaneous angular position of the attached rotating shaft. Resolvers are typically built like small motors with a rotor attached to the shaft whose position is to be measured, and a stator which produces the output signals. Resolvers are preferred over other sensors such as the encoders. Encoders are fairly complex and contain some delicate parts [2]. This makes them less tolerant of mechanical abuse and restricts their allowable temperature. The relevant equations of resolver rotor winding voltage (U_0) and stator windings (U_1 and U_2) are summarized as follows:

$$U_0(t) = \hat{U}_0 . \sin \omega_{\text{ref}} t \tag{6}$$

$$U_1(\phi, t) = \hat{U}_0 . k . \sin \phi . \sin \omega_{\text{ref}} t \tag{7}$$

$$U_2(\phi, t) = \hat{U}_0 . k . \cos \phi . \sin \omega_{\text{ref}} t \tag{8}$$

where k is the turn ratio of the resolver, $\hat{U}_0$ is the peak value, $\omega_{\text{ref}} t$ is the frequency in rad/sec of the excitation signal, and ϕ is the rotor position in rad. Resolver signals are normally detected through IC Chips. The equations of algorithm are given below for demodulation of modulated SSB signals obtained from rotor windings of the resolver. The algorithm attempts to minimize the error between the actual rotor angle ϕ and the computed angle θ, using a feedback loop, and eliminates the requirements of hardware [1, 2].

C. *Mathematical Algorithm of Resolver-to-Digital Converter*

Figure 2 shows the Simulink model of resolver algorithm.

$$\begin{aligned} \text{err} = {} & (\hat{U}_0 . \sin \omega_{\text{ref}} t . \cos \theta)(\hat{U}_0 . k \sin \phi . \sin \omega_{\text{ref}} t) \\ & - (U_0 . \sin \omega_{\text{ref}} t . \sin \theta)(U_0 . k \cos \phi . \sin \omega_{\text{ref}} t) \end{aligned} \tag{9}$$

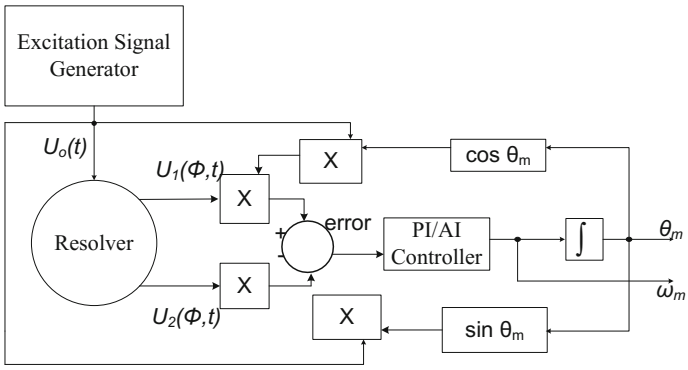

Fig. 2 Line diagram of mathematical algorithm [1]

$$\text{err} = \hat{U}_0(t).(\hat{U}_0.k.\sin\omega_{\text{ref}}t)[\sin\phi\cos\theta - \cos\phi\sin\theta] \tag{10}$$

$$\text{err} = A.[\sin(\phi - \theta)] \tag{11}$$

$$A = \hat{U}_0(t).(\hat{U}_0.k.\sin\omega_{\text{ref}}t) \tag{12}$$

The error is controlled to zero by PI or proposed AI-based controller. The integrator is used to increase the resolution of computed angle. Once the error decreases to zero, the computed angle θ becomes equal to actual rotor angle ϕ. In view of overcoming all the above-mentioned problems, a mathematical algorithm is developed in MATLAB/Simulink to eliminate the use of this hardware. The computed angle now is calculated by means of feedback loop control. The error signal is produced by difference of cross product of modulated signals and measured excitation signals [1, 2]. This error is controlled by PI- or AI-based controller. Thus, complete demodulation is done in the Simulink model itself.

Once the control loop is accomplished (i.e., error = 0), then the computed angle (θ), which is limited to the 0–2π rad, is equal to the actual rotor angle (ϕ).

3 Implementation of Algorithm Using PI Controller

The sine and cosine waves utilized in the model in Fig. 3 have been generated by pre-defined functions in MATLAB. The SSB1 and SSB2 signals are generated by modulation of sine and cosine waves over the carrier signal having a frequency of 2 kHz. The input to the PI/AI controller is the error between actual and calculated rotor angle, and the outputs of the controller are speed and position. The controller removes oscillation produced in the response. The controller has to be tuned to achieve accurate results.

4 Implementation of Algorithm Using AI Controller

The adaptive neuro-fuzzy inference system (ANFIS) is a FIS implemented in the framework of an adaptive fuzzy neural network. It combines the explicit knowledge representation of a FIS with the learning power of ANNs. The basic structure of the type of fuzzy inference system is a model that maps input characteristics to input membership functions, input membership function to rules, rules to a set of output characteristics, output characteristics to output membership functions, and the output membership function to a single-valued output or a decision associated with the output [12–17].

As shown in Fig. 3, the PI controller is replaced by an AI (ANFIS-based) controller. Figure 4 shows the ANFIS controller. The input through port 1 is that of

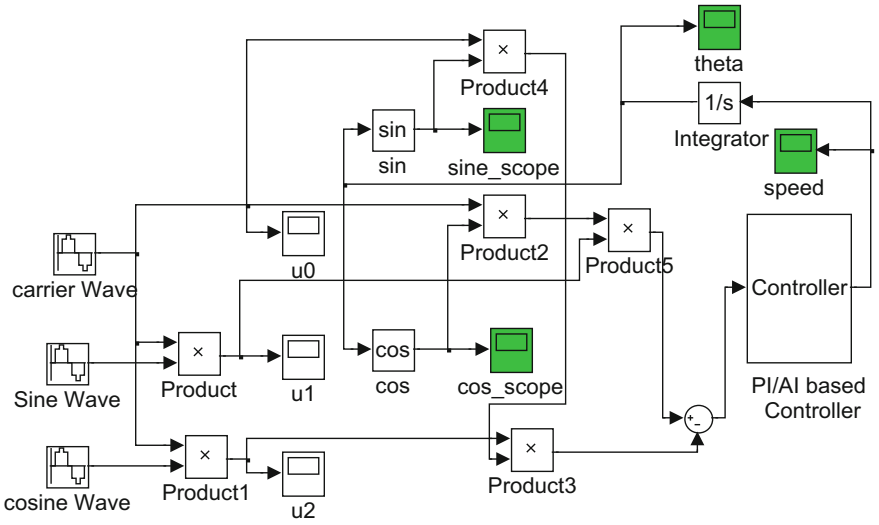

Fig. 3 Simulink model implementing the algorithm with a PI controller

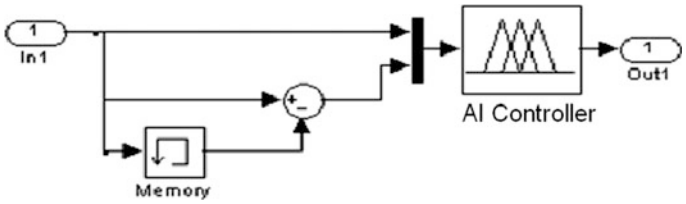

Fig. 4 ANFIS controller

the error between the actual and calculated rotor angle. Two variables are fed into the neuro-fuzzy controller, i.e., error and the rate of change of error. The rate of change of error is calculated using a memory block to store last value of the error which is then compared to the new error value.

The AI controller requires a FIS structure fismat1 which is generated by writing a MATLAB code. Using a given input/output data set, the ANFIS function constructs a fuzzy inference system (FIS) whose membership function parameters are tuned using backpropagation algorithm [12, 17]. The number of membership functions used for each input is seven. The type of membership function is dsigmoidal.

5 Implementation of Algorithm Using dSPACE

The proposed mathematical algorithm for tracking of position using AI controller is experimentally implemented using dSPACE 1104. The DSP has been supplemented by a set of onboard peripherals used in digital control systems, such as A/D, D/A converters, and incremental encoder interfaces. The dSPACE 1104 is also equipped with a TI TMS320C240 16-bit DSP processor. DSP acts as a slave processor and provides the necessary digital input/output (I/O) ports and powerful timer functions such as input capture, output capture, and pulse width modulation (PWM) waveform generation. Due to high processor speed in dSPACE, the sampling of high-frequency signals is accurate.

As shown in Figs. 5 and 6, the mathematical algorithm using AI controller is interfaced with the hardware. The ADC blocks used are provided in the RTI blockset library. Externally, the resolver signals and the carrier are connected to these channels. The signals received in the computer are to be multiplied by a certain gain to remove the attenuation introduced by the ADC channels at the time of acquisition. A block of zero-order hold is introduced for proper sampling of the analog signals. Figure 6 shows the implementation of algorithm using AI controller in real time using dSPACE.

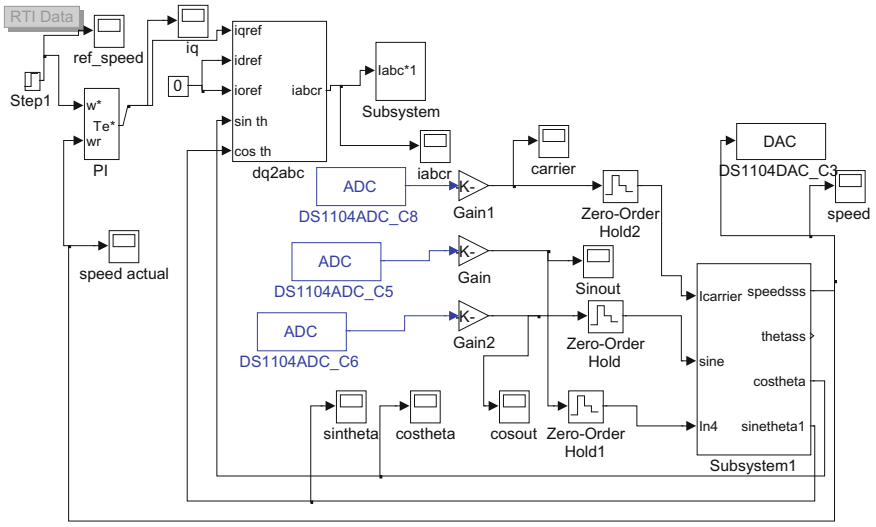

Fig. 5 Simulink model of PMSM for implementing the algorithm using dSPACE

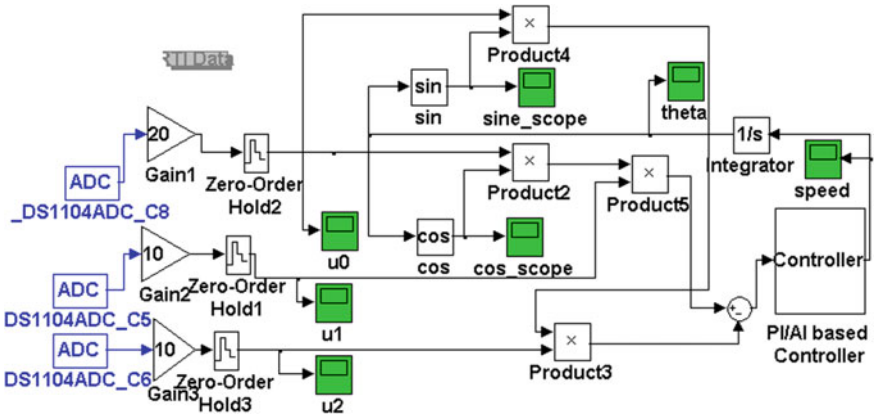

Fig. 6 Simulink model of the algorithm implementation using dSPACE and a PI/AI controller

6 Results and Discussion

Figure 7 shows the demodulated sine and cosine waves as output results from RDC using IC LM1496. The noise in the signals received is clearly seen in figure. Figure 8 shows the results obtained by simulating the MATLAB-/Simulink-based model of the algorithm to replace RDC. From the SSB signals, the demodulated sine and cosine signals and speed of 70 rad/sec are detected using PI and then AI controller as shown in Fig. 8a, b, respectively.

The algorithm to replace RDC is run with PI and then AI-based controllers with the PMSM drive model, and the hardware results using dSPACE to detect the resolver signals in terms of demodulated sine and cosine waves with the speed as output are obtained from very low speed to a larger speed. Figure 9 shows the detected speed at 28 and 57 rad/sec with demodulated signals at 4.5 and 9 Hz, respectively, with PI controller.

Figure 10 shows the detected speed at 104 and 140 rad/sec with demodulated signals at 16 and 22 Hz, respectively, with PI controller.

Figure 11 shows the detected speed at 28 and 52 rad/sec with demodulated signals at 4.5 and 8.3 Hz, respectively, with AI controller.

Figure 12 shows the detected speed at 105 and 126 rad/sec with demodulated signals at 16.7 and 20 Hz, respectively, with AI controller

The detected speed using AI controller is smooth as compared to as obtained using PI controller, and oscillations are considerably reduced at every speed range.

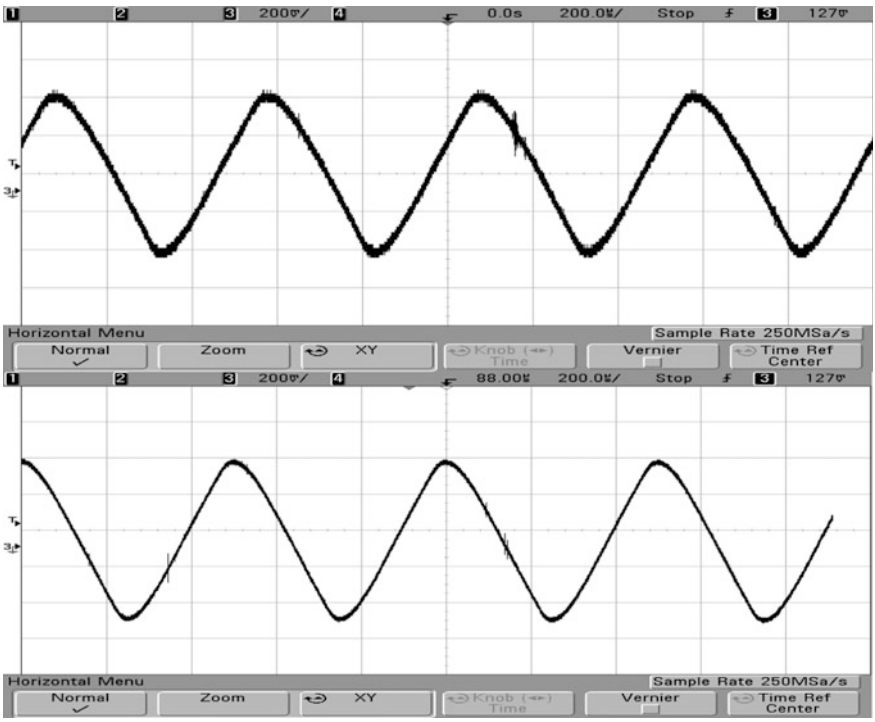

Fig. 7 Demodulated sine and cosine waves with RDC

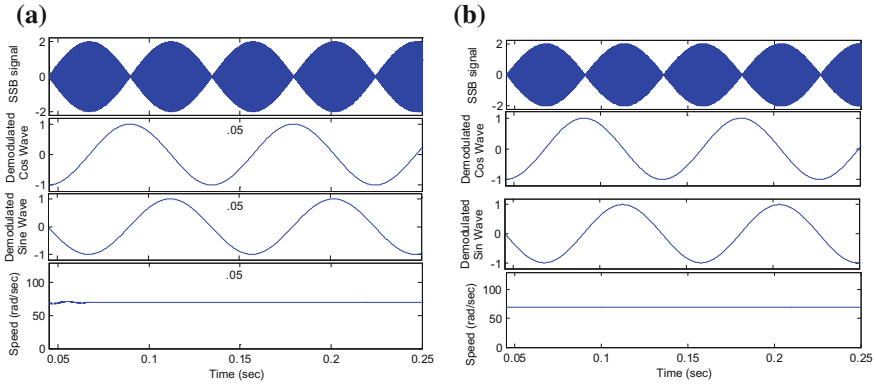

Fig. 8 MATLAB-/Simulink-based SSB signal, demodulated cosine and sine waves for 70 rad/sec of speed

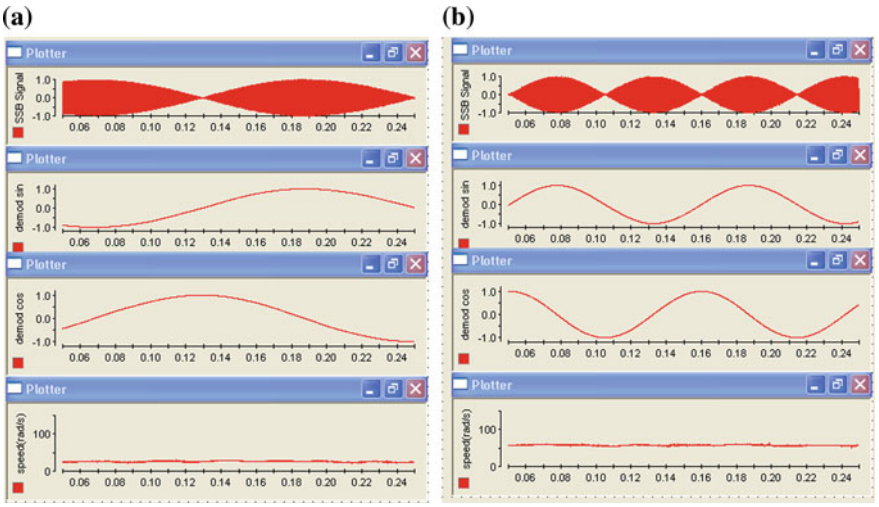

Fig. 9 dSPACE-based output signals at frequency 4.5 Hz and at speed 28 rad/sec and at frequency 9 Hz and at speed of 57 rad/sec with PI controller

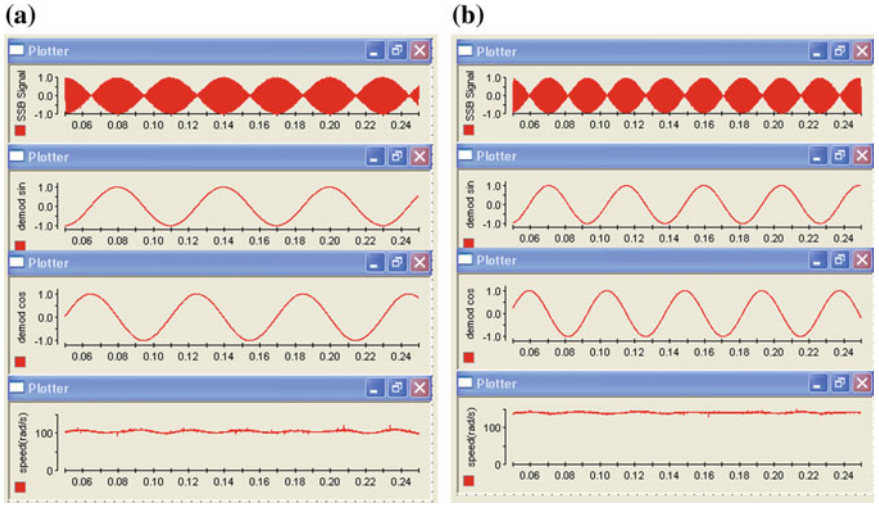

Fig. 10 dSPACE-based output signals at frequency 16 Hz and at speed 104 rad/sec and at frequency 22 Hz and at speed of 140 rad/sec with PI controller

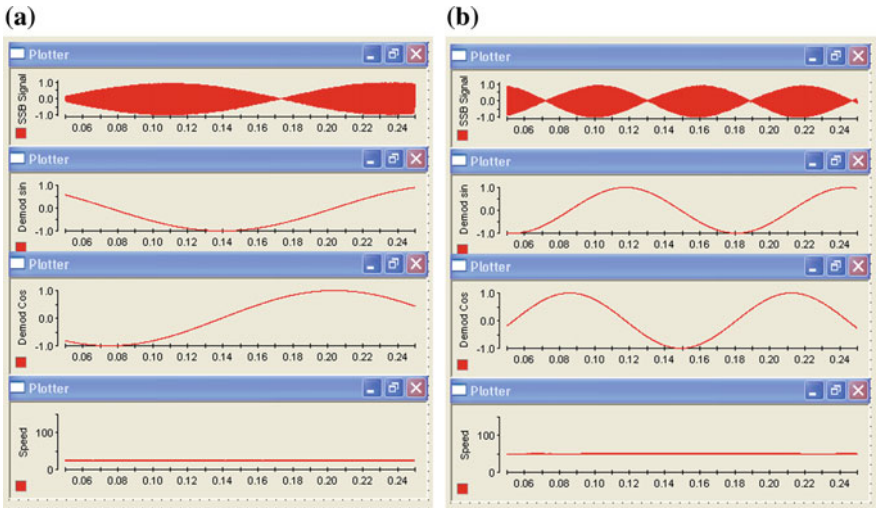

Fig. 11 dSPACE-based output signals at frequency 4.5 Hz and at speed 28 rad/sec and at frequency 8.3 Hz and at speed of 52 rad/sec with AI controller

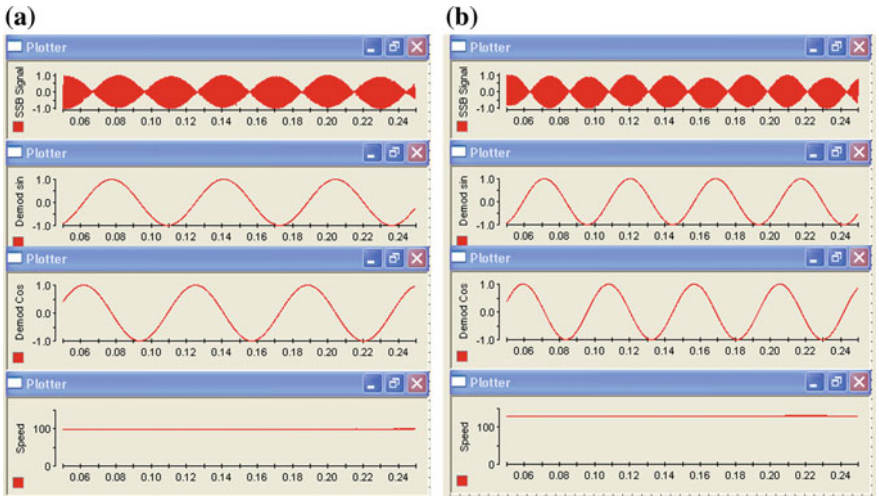

Fig. 12 dSPACE-based output signals at frequency 16.7 Hz with speed 105 rad/sec and at frequency 20 Hz with the speed of 126 rad/sec with AI controller

7 Conclusion

Successful simulation and implementation of the proposed AI control-based algorithm to replace hardware requirements to detect speed in real time are achieved using dSPACE. The motor is run at variable speed, and the speed is tracked with minimum error in the large range. As dSPACE incorporates a superior processor, it is capable of acquiring data at very high speed which leads to efficient tracking of position and speed.

A comprehensive study of AI controller as compared to PI controller based on results obtained is performed. The PI controller implemented in the algorithm is replaced by an AI controller. With the experience gained by working on AI network, it is observed that the rotor position is tracked accurately and the computed speed is achieved accurately.

The proposed AI-based software algorithm used for resolver-to-digital conversion can force the speed error to minimize. Then, the correct rotor speed computation is guaranteed. The implementation of this algorithm filters out noise and reduces the complexity and cost due to hardware circuit. The delay introduced due to filters in hardware circuit is also removed.

The implementation of RDC software algorithm using AI controller reduces noise in speed response and introduces stability in the system.

References

1. A. Kusagur, S.F. Kokad, B.V.S. Ram, Modeling, design & simulation of an adaptive neuro-fuzzy inference system (ANFIS) for speed control of induction motor. Int. J. Comput. Appl. **6**(10), 29–44 (2010)
2. W. Kaewjinda, M. Konghirun, Vector control drive of permanent magnet synchronous motors using resolver sensor. ECTI Trans. Electr. Eng. **5**, 560–583 (2007)
3. A.K. Yadav, P. Gaur, in *AI Based Adaptive Control and Design of Autopilot System for Nonlinear UAV*. Proceedings Sadhana—Academy Proceedings in Engineering Science Sadhana, vol. 39, no. 4 (Springer, 2014), pp. 765–783
4. D.C. Hanselman, Resolver signal requirements for high accuracy resolver-to-digital conversion. IEEE Trans. Ind. Electron. **37**, 556–561 (1989)
5. B. Murray, Hare, A. Hirao, in *Resolver Position Sensing System with Integrated Fault Detection for Automotive Applications*. Conference Record 2002 IEEE Proceedings in Sensors, vol. 2, pp. 864–869
6. A.O. Di Tommaso, R. Miceli, in *A New High Accuracy Software Based Resolver-to-Digital Converter*. Conference Record 2003 IEEE International Conference of the Industrial Electronics Society, vol. 3, pp. 2435–2440
7. C. Attaianese, G. Tomasso, D. De Bonis, in *A Low Cost Resolver to Digital Converter*. Conference Record 2001 IEEE International Conference on Electric Machines and Drives, pp. 917–921
8. J. Wan, X. Li, H. Guo, in *The Analysis and Design of High-Speed Brushless Resolver Plus R/D Converter Shaft-Angle Measurement System*. Conference Record 2001 IEEE International Conference on Electrical Machines and Systems, vol. 1, pp. 289–292

9. P. Gaur, S. Bhardwaj, N. Jain, N. Garg, A. Prashant, A.P. Mittal, B. Singh, in *A Novel Method for Extraction of Speed from Resolver Output Using Neural Network in Vector Control of PMSM*. Proceedings of IEEE International Conference Delhi, IICPE 2010, January 2011
10. H. Li, Z. Qin, W. Xuhui, X. Zhijie, in *The Position Information Acquisition of Rotor Flux in a PMSM Control System Based on DSP*. Proceedings of ICEMS, Conference on Electrical Machines and Systems, 2003, pp. 538–540
11. B.K. Bose, *Power Electronics and Variable Frequency Drives* (IEEE Press, Standard Publishers Distributors, New Delhi, 2000)
12. P. Gaur, A.K. Singh, in *Adaptive Control for Non-linear Systems Using Artificial Neural Network and its Application Applied on Inverted Pendulum.* Proceedings of IEEE International Conference Delhi, IICPE 2010, January 2011
13. J. Cui, G. Mu, Y. Fu, in *Design of PMSM Control System Based on Fuzzy Logic*. Conference Record October 2008 IEEE International Conference on Electrical Machines and Systems, pp. 1037–1041
14. F. Betin, D. Pinchon, G.A. Capolino, Fuzzy logic applied to speed control of a motor drive. IEEE Trans. Ind. Electron. **47**, 610–622 (2000)
15. A.K. Yadav, P. Gaur, Comparative analysis of modern control and AI based control for maintaining constant ambient temperature. World Rev. Sci. Technol. Sustain. Dev. **10**(1/2/3) (2013)
16. H.R. De Azevedo, K.P. Wong, in *A Fuzzy Logic Controller for Permanent Magnet Synchronous Machine*. Proceedings of IEEE Power Conversion Conference, Yokohama 1993, 19–21 April 1993, pp. 672–677
17. P. Vas, *Artificial-Intelligance Based Electrical Machines and Drives* (Oxford University Press, New York, 1999)

SISO Method Using Modified Pole Clustering and Simulated Annealing Algorithm

Jay Singh, Kalyan Chatterjee and C. B. Vishwakarma

Abstract A mixed method by using the modified pole clustering technique and simulated annealing is proposed to reduce higher-order mathematical model into a smaller one. The denominator and numerator polynomials are obtained by using a modified pole clustering technique and simulated annealing algorithm, respectively. The proposed-biased method generates k number of reduced models from higher-order systems. The compatibility of the method has been checked via time responses of the original higher-order system and the reduced-order system, respectively. Also, the proposed method has been compared with few known model-order reduction techniques through performance indices.

Keywords Pole clustering · Simulated annealing · Transfer function
Inverse distance measure (IDM)

1 Introduction

The mathematical representation of the any large-scale physical system (high-order system) is the easiest way to analyze, design, and understand deeply. Also, sometimes the high-order systems have many complexities and those may not be eliminated in a short time due to the large size of the system. Thus, we can say that

J. Singh (✉)
EEE Department, GL Bajaj Institute of Technology
and Management, Greater Noida, UP, India
e-mail: jaysinghism@gmail.com

K. Chatterjee
EE Department, Indian Institute of Technology
Dhanbad, Dhanbad, Jharkhand, India
e-mail: kalyanbit@yahoo.co.in

C. B. Vishwakarma
EE Department, SOE, GBU, Greater Noida, UP, India
e-mail: cbvishwa@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_13

analysis, controlling, evaluation, and designing of high-order system parameters increase the complexity. Therefore, for the complex higher-order system analysis, several order reduction techniques have been proposed in the literature. A reduction technique converts the higher-order mathematical model into the required small-order model, which maintains the original characteristics of the system also. Also, rise time, overshoot, settling time, and transient period are kept in mind during the reduction of original models, i.e., model-order reduction that may be in the frequency domain [1–7] or time domain [8–12].

A few decades ago, the pole clustering technique had been proposed by Sinha and Pal [5], where system poles are grouped in required numbers, i.e., orders of the reduced model. Further, it has been modified by Vishwakarma and Prasad [6] to get the most suitable reduced model which maintains the original response patterns of the system also. Modified pole clustering method is based on pole clustering techniques. It has a few iterations to calculate most dominant pole for the reduced-order model. Here, the number of iterations depends upon the order of the reduced model. Therefore, it generates most dominant pole for reduced model.

Another technique is the optimization tool, i.e., simulated annealing, used to calculate the zeros of reduced-order model. The simulation of annealing is a systematic approach that reduces a function of big number of variables to the numerical process of equilibration (annealing) of the corresponding artificial multiatomic system which was first articulated by Armen G. Khachaturyan.

In both the techniques [5, 6], an IDM criterion has been used to initiate the reduction process. Further, numerous reduction techniques have been proposed by using optimization techniques [13–16]. Keeping the several non-optimization and optimization techniques, it has been concluded that the mixed method is also a powerful reduction process, which generates efficient reduced-order model also. Hence, in this paper, a mixed reduction method has been proposed by using optimization techniques, i.e., simulated annealing and modified pole clustering method [6] to get the powerful reduced-order model. The proposed technique is simple, easy understandable, and computer oriented. The obtained reduced models will be stable if the original system is stable.

2 Statement of the Problem

Let the single input single output (SISO) transfer function is

$$G_n(s) = \frac{N(s)}{D(s)} = \frac{\sum_{j=0}^{n-1} a_j s^j}{\sum_{j=0}^{n} b_j s^j} \tag{1}$$

where a_j and b_j are known scalar constants.

Let the kth $(k = 1, 2, 3, \ldots, n-1)$-order reduced model from the original system (1) is

$$R_k(s) = \frac{N_k(s)}{D_k(s)} = \frac{\sum_{j=0}^{k-1} c_j s^j}{\sum_{j=0}^{k} d_j s^j} \tag{2}$$

where c_j and d_j are unknown scalar constants.

The objective of this paper is to get the kth-order reduced model (2) from model (1), such that it holds the original features of the system in reduced-order model (2).

3 Description of the Method

Reduction process is elaborated in two steps as follows:

3.1 Determination of the Denominator Polynomial

There are some rules to get a denominator polynomial [6] for the reduced-order models.

A. Make a separate group of poles for real and complex, respectively.
B. Any group poles of a system should not contain a pole of other side of the s-plane.
C. In the reduced model, jw axis poles have to be retained in jw axis only.
D. Most dominant pole, i.e., pole at origin, should be retained.

The reduction technique is given as follows:

Let there are r real poles in ith cluster, i.e., $p_1, p_2, \ldots p_r$, where $|p_1| < |p_2| \cdots < |p_r|$.

(i) Let 'm' pair of complex conjugate poles in the jth cluster be

$$[(\alpha_1 \pm j\beta_1), (\alpha_2 \pm j\beta_2), \ldots (\alpha_m \pm j\beta_m)],$$

where

$$|\alpha_1| < |\alpha_2| \cdots < |\alpha_r|$$

By using the same algorithm as below, a modified pole cluster center is calculated as follows:

$$c_j = \left[\sum_{i=0}^{r} \left(\frac{-1}{|p_i|} \right) \div r \right]^{-1}$$

The calculated pole cluster may be written as follows:

$$\phi = A_{ei} \pm jb_{ei}$$

where $\phi = A_{ei} + jb_{ei}$ and $\phi = A_{ei} - jb_{ei}$.

Therefore, the complete iterative process may be written as follows:

Step-I. Let if 'r' real poles in a selected cluster as follows:

$$|p_1| < |p_2| \cdots < |p_r|.$$

Step-II. Now set $j = 1$ in the equation of the modified cluster center.
Step-III. The pole cluster center is calculated as follows:

$$c_j = \left[\sum_{i=0}^{r} \left(\frac{-1}{|p_i|} \right) \div r \right]^{-1}$$

Step-IV. Further set $j = j + 1$ in the equation of the modified cluster center.
Step-V. Now find a modified cluster center from the given equation

$$c_j = \left[\left(\frac{-1}{|p_1|} + \frac{-1}{|c_{j=1}|} \right) \div 2 \right]^{-1}$$

Step-VI. Now check, whether $r = (j + 1)$ or not, if not, go to Step-IV.
Step-VII. Finally center of the kth cluster as $p_{ak} = c_j$.

Further, in kth-order denominator polynomial, one of the cases may occur.

Case-1: If all pole clusters are real

$$D_k(s) = (s - p_{e1})(s - p_{e2}) \cdots (s - p_{ek}) \tag{3}$$

where $p_{e1,}$ $p_{e2,}$ and p_{ek} are first, second, and kth cluster center.
Case-2: If all pole clusters are complex conjugates

$$D_k(s) = (s - \phi_{e1})(s - \phi_{e2}) \cdots (s - \phi_{ek/2}) \tag{4}$$

Case-3: If some pole cluster centers are real and some are complex conjugate.

$$D_k(s) = (s - p_{e1})(s - p_{e2}) \cdots (s - p_{e(k-2)})(s - \phi_{e1})(s - \phi_{e2}) \tag{5}$$

Finally, the denominator polynomial $D_k(s)$ is obtained as follows:

$$D_k(s) = d_o + d_1 s + d_2 s^2 + \cdots d_k s^k \tag{6}$$

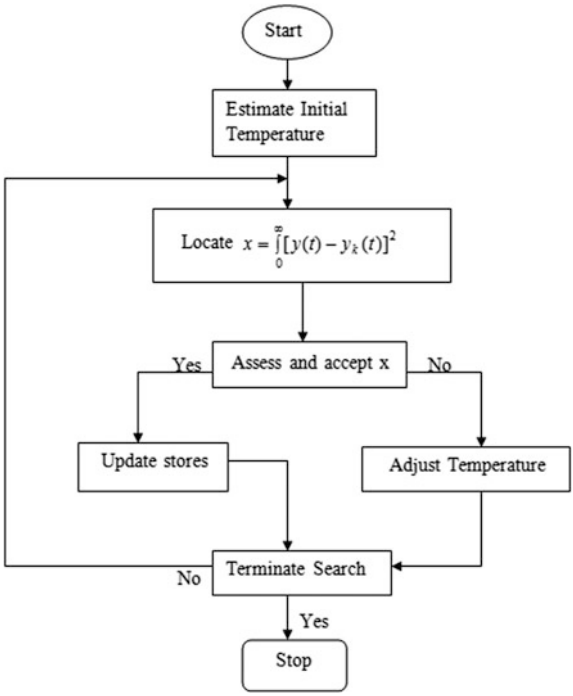

Fig. 1 Flowchart of simulated annealing

The numerator polynomial is obtained by using reduced-order denominator polynomial, and original system function via simulated annealing flowchart process as shown in Fig. 1. It is a powerful optimization search, where numerator coefficients are generated.

3.2 Determination of the Numerator Polynomial via Simulated Annealing

3.2.1 Error Function Minimization Using Simulated Annealing (SA)

SA is a process to solve constrained and unconstrained optimization problems. The process simulates the physical process of warming a material, then slowly sinking the temperature to drop the defects, hence minimizing the system energy. During every iteration of the SA algorithm, a new point is arbitrarily generated. The space of the new point from the existing point is based on a probability dissemination with a scale related to the temperature. An algorithm admits all new points with definite probability. By accepting the points that increase the objective, the process avoids being trapped in the local minima in the initial iterations and is capable to explore worldwide for better solutions.

Therefore, to get the perfect reduced-order model, an error index must be minimized between reduced model (2) and original high-order system (1). Hence, to get unknown numerator coefficients $c_0, c_1, \ldots c_{k-1}$, an error minimization technique, i.e., simulated annealing, has been applied as shown in Fig. 1.

An error function, i.e., integral square error (ISE), is written as follows:

$$\text{ISE} = \int_0^\infty (y(t) - y_k(t))^2 \mathrm{d}t \tag{7}$$

where $y(t)$ is the original high-order unit step response and $y_k(t)$ is the reduced-order unit step response. Numerator coefficients are generated through the simulated annealing process as shown in Fig. 1 and can be written as follows:

$$N_k(s) = c_0 + c_1 s + c_2 s^2 + \cdots + c_{k-1} s^{k-1} \tag{8}$$

4 Numerical Example

To elaborate the proposed-biased method, a mathematical transfer function model has been taken from the literature. A second-order reduced model has been calculated in details to understand the proposed method. To measure the error, ISE is calculated between the transient quantities of the high-order original transfer function and reduced model transfer function using MATLAB.

Example 1 Consider an eighth-order system [5]. $G(S) = \frac{N(s)}{D(s)}$

$$N(s) = 18s^7 + 514s^6 + 5982s^5 + 36,380s^4 + 122,664s^3 + 222,088s^2 + 185,760s + 40,320$$

$$D(s) = s^8 + 36s^7 + 546s^6 + 4536s^5 + 22,449s^4 + 67,284s^3 + 118,124s^2 + 109,584s + 40,320$$

The poles are as follows: $(-1, -2, -3, -4, -5, -6, -7, -8)$.

Here, two pole clusters have been taken for the generation of second-order reduced model, i.e., cluster 1 $(-1, -2, -3, -4)$ and cluster 2 $(-5, -6, -7, -8)$.

Using the proposed reduction technique, an original high-order transfer function is reduced into second-order reduced model.

$$R_2(s) = \frac{N_2(s)}{D_2(s)} = \frac{15.11s + 5.2220}{s^2 + 6.0964s + 5.2220}$$

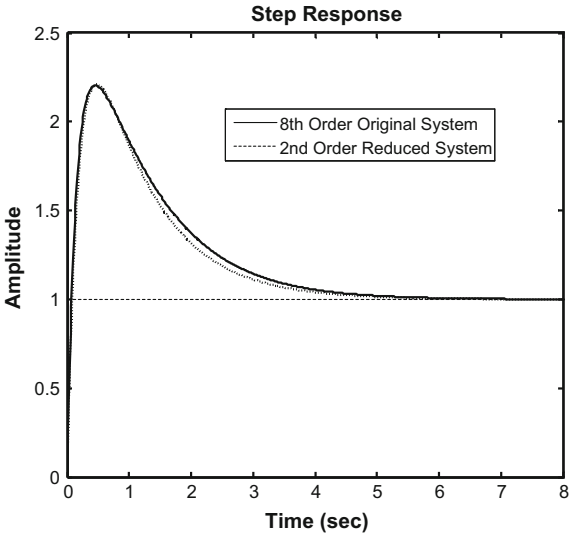

Fig. 2 Comparison of the step response between original and reduced system

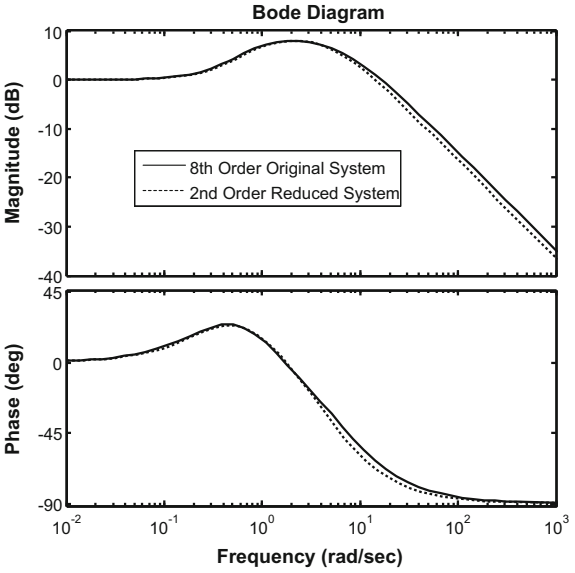

Fig. 3 Comparison of the frequency response between original and reduced system

Also, for the reduction of the original system into third-order reduced model, three pole clusters should be selected as $(-1, -2), (-3, -4, -5)$ and $(-6, -7, -8)$; it may be any other combination in a systematic manner.

A reduced-order and higher-order unit step response is shown in Figs. 2 and 3, respectively; also error index is shown in Table 1. From Figs. 1, 2, and Table 1, it is clear that the reduced-order model may be suitable for the analysis, designing, and evaluation, etc.

Table 1 Comparisons of reduced models

Methods	Reduced models $R_2(s)$	ISE
Proposed method	$\frac{15.11s+5.2220}{s^2+6.0964s+5.2220}$	0.0073
Vishwakarma and Prasad [6]	$\frac{16.51145s+5.45971}{s^2+6.19642s+5.45971}$	0.0140
Parmar et al. [17]	$\frac{24.11429s+8}{s^2+9s+8}$	0.0480
Mukherjee et al. [18]	$\frac{11.3909s+4.4357}{s^2+4.2122s+4.4347}$	0.05689
Mittal et al. [19]	$\frac{7.0908s+1.9906}{s^2+3s+2}$	0.2689
Prasad and Pal [20]	$\frac{17.98561s+500}{s^2+13.24571s+500}$	1.4584

5 Conclusions

In this paper, authors have proposed a model-reduction method based on optimization technique and modified pole clustering technique, where the denominator polynomial is generated by modified pole clustering method, whereas numerator coefficients are obtained from simulated annealing optimization technique. To show the quality of the proposed reduced model, two types of responses, i.e., time and frequency response, have been shown graphically in Figs. 2 and 3, respectively. Also, an error index has been shown in Table 1 for the more viability of the proposed method. From these graphical and tabular comparisons, it is concluded that the method consisting of modified pole clustering and simulated annealing is simple and capable to achieve transient response as well as steady-state response to get the original characteristics of the system.

References

1. J. Singh, K. Chattterjee, C.B. Vishwakarma, MIMO system using eigen algorithm and improved Pade approximations. SOP Trans. Appl. Math. **1**(1), 60–70 (2014)
2. J. Singh, K. Chattterjee, C.B. Vishwakarma, System reduction by eigen permutation algorithm and improved Pade approximations. Int. J. Math. Comput. Sci. Eng. **8**(1), 1–5 (2014)
3. J. Singh, C.B. Vishwakarma, K. Chattterjee, Biased reduction method by combining improved modified pole clustering and improved Pade approximations. Appl. Math. Model. Elsevier **40**, 1418–1426 (2016)
4. J. Pal, Improved Pade approximants using stability equation methods. IEEE Electron. Lett. **19** (11), 426–427 (1983)
5. A.K. Sinha, J. Pal, Simulation based reduced order modeling using a clustering technique. Comput. Electr. Eng. **16**(3), 159–169 (1990)

6. C.B. Vishwakarma, R. Prasad, MIMO system reduction using modified pole clustering and genetic algorithm. Model. Simul. Eng. **2009**, 1–6 (2009). Hindawi Publishing Corporation
7. A.K. Sinha, J. Pal, Simulation based reduced order modelling using a clustering technique. Comput. Electr. Eng. **16**(3), 159–169 (1990)
8. C.B. Vishwakarma, R. Prasad, Time domain model order reduction using Hankel matrix approach. J. Franklin Inst. Elsevier **351**, 3445–3456 (2014)
9. N.K. Sinha, G.T. Bereznai, Optimal approximation of high order systems by low order models. Int. J. Control **14**, 951–959 (1971)
10. S.A. Marshall, An approximation method for reducing the order of a large system. Control Eng. **10**, 642–648 (1966)
11. C.P. Therapos, Internally balanced minimal realization of discrete SISO systems. IEEE Trans. Autom. Control **30**(3), 297–299 (1985)
12. S.K. Nagar, S.K. Singh, An algorithmic approach for system decomposition and balanced realized model reduction. J. Franklin Inst. **341**, 615–630 (2004)
13. A.K. Mittal, R. Prasad, S.P. Sharma, in *Reduction of Multivariable Systems Using Stability Equation Method and Error Minimization Technique*. Proceedings of the 27th National Systems Conference (NSC-2003), Indian Institute of Technology, Kharagpur, India, 17–19 December 2003, pp. 34–38
14. S. Mukherjee, R.N. Mishra, Reduced order modelling of linear multivariable systems using an error minimization technique. J. Franklin Inst. **325**(2), 235–245 (1988)
15. S. Mukherjee, R.N. Mishra, Order reduction of linear systems using an error minimization technique. J. Franklin Inst. **323**(1), 23–32 (1987)
16. A.K. Mittal, R. Prasad, S.P. Sharma, Reduction of linear dynamic systems using an error minimization technique. J. Inst. Eng. India IE(I) J. EL **84**, 201–206 (2004)
17. G. Parmar, S. Mukherjee, R. Prasad, Division algorithm and eigen spectrum analysis. Appl. Math. Model. Elsevier **31**, 2542–2552 (2007)
18. S. Mukherjee, Satakshi, R.C. Mittal, Model order reduction using response matching technique. J. Franklin Inst. **342**, 503–519 (2005)
19. A.K. Mittal, R. Prasad, S.P. Sharma, Reduction of linear dynamic systems using an error minimization technique. J. Inst. Eng. India IE (I) J. EL **84**, 201–206 (2004)
20. R. Prasad, J. Pal, Stable reduction of linear systems by continued fractions. J. Inst. Eng. India IE (I) J. EL **72**, 113–116 (1991)

Evaluation and Comparison of Effects of Common Cause Failures on Redundancy Optimization of a Complex Network

Kalpana Hazarika and G. L. Pahuja

Abstract Reliability evaluation of a complex network is mostly done by exact and approximate approaches. Redundancy allocation (with s-identical and s-independent components) is one of the commonly used techniques for reliability optimization. But the occurrence of common cause failure (CCF) defeats the benefit of redundancy optimization. To validate the same, we have presented five exact approaches to evaluate the reliability of complex network. Also extended the work to evaluate and compare the reliability of the same network under CCF. To incorporate the CCFs in reliability evaluation, conditional probability approach is used. The qualitative and quantitative analysis of the evaluation process is discussed by implementing on a complex network with five subsystems. It includes the process of reliability optimization, process to incorporate CCF, quantitative demonstration of the concept with five evaluation methods, observations and conclusions.

Keywords Redundancy optimization · CCF · Exact methods · Complex network

Notations

$R_s(x_i)$	Reliability of the system consisting x_i components
$R_{IC}(t)$	Reliability of the system under CCFs at the instant t
r_i	Reliability of subsystem i
$R_{ICi}(t)$	Reliability of subsystem i under CCFs at the instant t
q_i	Unreliability of subsystem i
F_{ICi}	Unreliability of the subsystem i under CCFs at the instant t
p_i	Reliability of single component of individual subsystem
λ_{ij}	Failure rate of subsystem i involving j no of components

K. Hazarika (✉)
Department of Electrical and Electronics, I.T.S Engineering College, Greater Noida, UP, India
e-mail: hazarika.kalpana@gmail.com

G. L. Pahuja
Department of Electrical Engineering, NIT Kurukshetra, Kurukshetra, HR, India
e-mail: pahuja.gl@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_14

1 Introduction [1]

Most preferred techniques of reliability optimization include the incorporation of redundant element either at component or system level. Generally, the redundant components are of identical types. Hence, when operated under same technical and physical conditions, the probability of common cause failure increases. This defeats the purpose of redundancy optimization [2]. This paper introduces a method of reliability optimization with identical component and extended the evaluation of reliability optimization incorporating CCFs to observe the effects. Based on the number of components involved in a particular fault process, different failure rates are assumed [2].

2 Process for Reliability Optimization Using Redundancy Optimization [3, 4]

By incorporating redundant element either at component or at system level, reliability of any systems or networks can be optimized. To optimize the system's reliability under cost constraints, one of the well-known heuristic methods [5] is adopted here. The technique [2] applied to optimize the network is demonstrated as under. The technique applied to optimize the network by surfeit allocation at component level is briefed as under:

(a) Based on cost constraints, evaluate the path selection and component selection factor [5].
(b) Until its limiting constraint is met, increase the designated component of the designated path by 1.
(c) By apportioning the redundant element in the network, final optimized structure is determined.
(d) Calculate the reliability of the upgraded network.

3 Technique to Incorporate the CCF for Reliability Evaluation [6–8]

3.1 Assumptions

(a) Each subsystem component is having two states and composed of s-identical components.
(b) Components are exposed to different failure processes. Every failure process of subsystem involves a specific number of components.

(c) All subsystems have their exclusive group of identical components. Hence, different failure rates are considered for all.

3.2 Technique to Integrate CCF

(a) At a particular time instant t, the possibility of a specified element to be in working mode is projected based on the possibility of k-specific element out of n elements to be in well condition [9] and s-independence of the Poisson processes.
(b) Identify all probable failure processes of subsystems and respective failure rates.
(c) Estimate the reliability of optimized subsystems using proposed concept.
(d) Formulate the reliability idiom of complete structure considering CCFs.
(e) Considering CCFs, evaluate the reliability of optimized network.

4 Example

Following complex network [5] as shown in Fig. 1 may be assumed as the network configuration of any networking systems with five variables. Its reliability data are:

x_i	1	2	3	4	5
P_i	0.7	0.85	0.75	0.80	0.9
C_i	2	2	2	3	1

(Reliability and cost of single element)

Based on different failure processes involved, failure rates for every subsystem of above network are assumed as shown in Table 1.

A. Reliability idiom of the network "Fig. 1" using **"Shi"** is

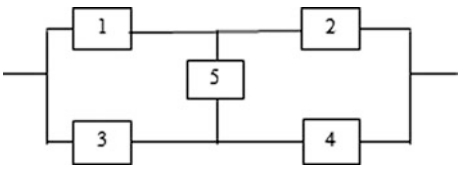

Fig. 1 Complex bridge network

Table 1 Chart of failure rates

Subsystem 1		Subsystem 2		Subsystem 3		Subsystem 4		Subsystem 5	
λ_{11}	0.035	λ_{21}	0.016	λ_{31}	0.028	λ_{41}	0.022	λ_{51}	0.0105
λ_{12}	0.0175	λ_{22}	0.008	λ_{32}	0.014	λ_{42}	0.011	λ_{52}	0.00526
λ_{13}	0.0087	λ_{23}	0.004	λ_{33}	0.007	λ_{43}	0.0055	λ_{53}	0.00263

$$R_S(x) = R_1(x_1)R_2(x_2)Q_1(x_1)Q_5(x_5) + Q_1(x_1)R_3(x_3)R_4(x_4)Q_5(x_5) + [R_1(x_1)R_3(x_3) + R_3(x_3)R_5(x_5) + R_5(x_5)R_1(x_1) - 2 * R_1(x_1)R_3(x_3)R_5(x_5)] * [R_2(x_2) + R_4(x_4) - R_2(x_2)R_4(x_4)]$$

The network reliability of optimized structure {x^* = 3, 2, 2, 1, 1} integrating CCFs is

$$R_{IC}(t) = R_{IC1}R_{IC2}F_{IC3}F_{IC5} + F_{IC1}R_{IC3}R_{IC4}F_{IC5} + [R_{IC1}R_{IC3} + R_{IC3}R_{IC5} + R_{IC5}R_{IC1} - 2 * R_{IC1}R_{IC3}R_{IC5}] * [R_{IC2} + R_{IC4} - R_{IC2}R_{IC4}]$$

B. Reliability idiom of the network "Fig. 1" using **Inclusion and Exclusion law** is

$$R_s(x) = r_1r_2 + r_3r_4 + r_1r_4r_5 + r_2r_3r_5 - r_1r_2r_3r_4 - r_1r_2r_4r_5 - r_1r_3r_4r_5 - r_1r_2r_3r_5 - r_2r_3r_4r_5 + 2r_1r_2r_3r_4r_5$$

The network reliability of optimized structure {x^* = 3, 2, 2, 1, 1} integrating CCFs is

$$R_{IC}(t) = R_{IC1}R_{IC2} + R_{IC3}R_{IC4} + R_{IC1}R_{IC4}R_{IC5} + R_{IC2}R_{IC3}R_{IC5} - R_{IC1}R_{IC2}R_{IC3}R_{IC4} - R_{IC1}R_{IC2}R_{IC4}R_{IC5} - R_{IC1}R_{IC3}R_{IC4}R_{IC5} - R_{IC1}R_{IC2}R_{IC3}R_{IC5} - R_{IC2}R_{IC3}R_{IC4}R_{IC5} + 2R_{IC1}R_{IC2}R_{IC3}R_{IC4}R_{IC5}$$

C. Reliability idiom of the network "Fig. 1" using **Sum of Disjoint Product (SDP)** is

$$R_s(x) = r_1r_2 + q_1r_3r_4 + r_1q_2r_3r_4 + q_2q_3r_1r_4r_5 + q_1q_4r_1r_3r_5$$

The network reliability of optimized structure {x^* = 3, 2, 2, 1, 1} integrating CCFs is

$$R_{IC}(t) = R_{IC1}R_{IC2} + F_{IC1}R_{IC3}R_{IC4} + R_{IC1}F_{IC2}R_{IC3}R_{IC4} + F_{IC2}F_{IC3}R_{IC1}R_{IC4}R_{IC5} + F_{IC1}F_{IC4}R_{IC1}R_{IC3}R_{IC5}$$

D. Reliability idiom of the network "Fig. 1" using **Bay's Decomposition** is

$$R_s(x) = [(1 - q_1q_3)(1 - q_2q_4)] * r_5 + [(1 - (1 - r_1r_2)(1 - r_3r_4))] * q_5$$

The network reliability of optimized structure {x^* = 3, 2, 2, 1, 1} integrating CCFs is

$$R_{IC}(t) = [(1 - F_{IC1}F_{IC3})(1 - F_{IC2}F_{IC4})] * R_{IC5} + [-(1 - R_{IC1}R_{IC2})(1 - R_{IC3}R_{IC4})] * F_{IC5}$$

E. Reliability idiom of the network "Fig. 1" using **Exhaustive Search Method** is

$$\begin{aligned} R_S(x) = {} & q_3q_4q_5r_1r_2 + q_1q_2q_5x_3x_4 + q_4q_5x_1x_2x_3 + q_3q_5x_1x_2x_4 + q_2q_5x_1x_3x_4 \\ & + q_1q_5x_2x_3x_4 + q_1q_4x_2x_3x_5 + q_2q_3x_1x_4x_5 + q_3q_4x_1x_2x_5 + q_1q_2x_3x_4x_5 \\ & + q_1x_2x_3x_4x_5 + q_2x_1x_3x_4x_5 + q_3x_1x_2x_4x_5 + q_4x_1x_2x_3x_5 \\ & + q_5x_1x_2x_3x_4 + x_1x_2x_3x_4x_5 \end{aligned}$$

The network reliability of optimized structure {x^* = 3, 2, 2, 1, 1} integrating CCFs is

$$\begin{aligned} R_{IC}(t) = {} & F_{IC3}F_{IC4}F_{IC5}R_{IC1}R_{IC2} + F_{IC1}F_{IC2}F_{IC5}R_{IC3}R_{IC4} + F_{IC4}F_{IC5}R_{IC1}R_{IC2}R_{IC3} \\ & + F_{IC3}F_{IC5}R_{IC1}R_{IC2}R_{IC4} + F_{IC2}F_{IC5}R_{IC1}R_{IC3}R_{IC4} + F_{IC1}F_{IC5}R_{IC2}R_{IC3}R_{IC4} \\ & + F_{IC1}F_{IC4}R_{IC2}R_{IC3}R_{IC5} + F_{IC2}F_{IC3}R_{IC1}R_{IC4}R_{IC5} + F_{IC3}F_{IC4}R_{IC1}R_{IC2}R_{IC5} \\ & + F_{IC1}F_{IC2}R_{IC3}R_{IC4}R_{IC5} + F_{IC1}R_{IC2}R_{IC3}R_{IC4}R_{IC5} + F_{IC2}R_{IC1}R_{IC3}R_{IC4}R_{IC5} \\ & + F_{IC3}R_{IC1}R_{IC2}R_{IC4}R_{IC5} + F_{IC4}R_{IC1}R_{IC2}R_{IC3}R_{IC5} + F_{IC5}R_{IC1}R_{IC2}R_{IC3}R_{IC4} \\ & + R_{IC1}R_{IC2}R_{IC3}R_{IC4}R_{IC5} \end{aligned}$$

Results of five exact approaches with and without considering CCFs are as shown in Table 2 to compare the results.

Table 2 Comparative reliability charts

S. No.	Evaluation techniques	Reliability of the system without CCFs	Reliability of the system with CCFs for t = 10 h	Reliability of the system with CCFs for t = 100 h
A	Shi	0.993215	0.969826381	0.0049684516
B	Inclusion and Exclusion	0.993215	0.930623	0.00501881
C	Sum of Disjoint Product	0.993195	0.924850181	0.003541498592
D	Bay's Decomposition	0.9932157	0.930623084	0.00494437066
E	Exhaustive Search Method	0.993916	0.930623	0.00493220937

5 Observation and Conclusion

(a) In quantitative analysis it is observed that, the reliability of the same network when evaluated with different exact approaches without considering CCF, in most of the approaches, the results are similar till six decimal places. But in techniques "C" and "E" as stated in Table 2, is varying from third decimal places.

(b) When reliability of this network is analysed under CCFs at the time instant $t = 10$ h, in all approaches except "A" and "C", results are found similar till six decimal places. However, when time instant t is extended to 100 h to observe the effects on the results, remarkable variation is found in all. The highest difference is observed in SDP evaluation technique amongst all other techniques. This is because the impact of CCF on subsystem 1 (it is a system consists of three s-identical components connected in parallel as per optimized network configuration) is found highest in its quantitative analysis. And in the reliability idiom of the complex network based on SDP evaluation technique, the impact of reliability and unreliability of subsystem 1 is associated in every term of the expression. Hence, it may be concluded that the result is highly influenced by its performance as it highly prone to CCFs.

References

1. K.K. Aggarwal, Redundancy optimization in general system. IEEE Trans. Reliab. **R-25**(5) (1976)
2. K. Hazarika, G.L. Pahuja, Effect of common cause failures on redundancy optimization. SCOPUS Indexed. Int. J. Appl. Mech. Mater. **592–594**, 2491–2495 (2014). ISSN-1662-7482
3. K. Gopal, K.K. Aggarwal, J.S. Gupta, An improved algorithm for reliability optimization. IEEE Trans. Reliab. **R-27**(5) (1978)
4. T. Kohda, K. Inoue, A reliability optimization method for complex systems with the criteria of local optimality. IEEE Trans. Reliab. **R-31**(1) (1982)
5. S. Dighua, A new heuristic algorithm for constrained redundancy optimization in complex system. IEEE Trans. Reliab. **R-36**(5) (1987)
6. J. Yaun, S.-T. Wong, On inclusion & exclusion and sum of mixed products formulas to calculate system reliability with common-cause failures. Reliab. Eng. Syst. Safety **27**, 219–230 (1990)
7. J.M. Nahman, Minimal path & cuts of networks exposed to common-cause failures. IEEE Trans. Reliab. **R-41**(1) (1992)
8. J. Yaun, Pivotal decomposition to find availability & failure frequency of systems with common cause failures. IEEE Trans. Reliab. **R-36**(1) (1987)
9. K.C. Chae, G.M. Clark, System reliability in the presence of common-cause failures. IEEE Trans. Reliab. **R-35**(1) (1986)

Privacy by Using an Asymmetry Encryption for Pseudonymization in Mix-Zones over Road Networks

M. S. Needa Nazeeema, Ishrath Fathima and Rubina Shahin Zuberi

Abstract In this paper, we resolve the privacy issues, giving solution of the user's identification revealing problem. To prevent this leading information, we defined the Mix-zones and roadside unit (RSU), for securing the moving objects (not only vehicles but also humans who are registered to LBS) real identity by the malicious eavesdroppers who are illegally tracking the users. We also distinguish the symmetry and asymmetry encryptions processes and the method of applying encryption for pseudonymization using MAC and HMAC.

Keywords Location privacy · Mix-zones (MZ) · Roadside unit (RSU) Encryption scenario

1 Introduction

Growth of self-routing handsets and the services such as Google Maps, which is also a routing service, not only in handsets but also in Internet has done important job for social life which are used for reliable location-based ideas and functioning. It seems as good that minimising costs of new mobile handsets, pushes Information Society to the mobiles devices. This created opportunity for growth and acceptance of location-based services (LBS), which uses mobile devices, wireless networks, and positioning technologies to bring distinctive and vital value to users.

Location-based services are very smart services and play a vital role in every part of life. However, present location-based services are growing much slower than

M. S. N. Nazeeema (✉) · I. Fathima · R. S. Zuberi
Electronics and Communication Engineering, I.T.S. Engineering College, Greater Noida, India
e-mail: needanazeema@gmail.com

I. Fathima
e-mail: zoya.ishrath@gmail.com

R. S. Zuberi
e-mail: rubinazuberi.ece@its.edu.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*, Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_15

what is expected. Privacy is the main concern as user will get easily traced up here. The capability of constantly examining the position of an individual in real time is a critical privacy concern. Mobile users can be continuously tracked without knowing him/her, which may cause to unsafe consequences. The effects of privacy problem of prospective and skilled customer's acceptance of location-based services have made beyond approachable.

It simply provides a customer to use services or to turn off the services. An alternative solution for this is to provide privacy which is done in this paper through encrypted HMAC. Let us consider a region with k number of users. This is well explained in Aggarwal et al. in 2005. In this, a technique called k-anonymity is used which is cited as in [1–3].

Another substitute and contributory approach to spatial cloaking depends on protection of location privacy is to break the flow of location exposure by developing model, known as Mix-zones [2, 4, 5]. Mix-zones anonymize identity of client by limiting the locality where users can be located. Mix-zones are areas at roadside junctions where the moment of the user not be traced by any applications. This gives sure-fire solution by enforcing that an entered set of users, pseudonyms are replaced and leave a Mix-zone, in such a way that the mapping is done in between their old and new pseudonyms not revealed [4].

In this, encryption process is done by using MAC. HMAC can be used in conjunction with any iterated cryptographic hash function [6]. MD5 and SHA-1 are examples of such hash functions. HMAC also uses a secret key for calculation and verification of the message authentication values.

2 Secure Road Network Junctions

Primarily, the Mix-zones are boundaries declared on the road networks' junctions. These are used to explain the network managements and communicating aspects of the mobile and location-based service (LBS) providers and also ready to permit the enhanced security aspects within these covering zones. Basically, it is discussed in two dimensions—first is pseudonymization of the MO and second is due to the confusion made by the anonymisation of the traffic in the region. The LBS provider provides pseudonyms only after the MO's registration. When the MO object crosses the Mix-zone boundaries, message packets with new aspects are embedded. The above discussed method provided the pseudonyms as the MO authentication and process to verify the availability of signal for an indivdual within the Mix-zone.

The emphasis is applied upon the encryption with hash message authentication code (HMAC) and pseudonyms which are assigned to the MOs. They play major role in the success of the privacy enhancement over proposed Mix-zones.

3 Assumptions for Defining System Architecture

By considering and defining, there are some important things that helps us to understand:

(a) Client/Moving Object (MO)
(b) Trusted LBS server
(c) Privacy agents
(d) Subscription

(a) Client/moving object (MO): This is a handset/device which is used by the user with enabled data connectivity, wireless access (Wi-Fi), Global Positioning System (GPS), and privacy constituents which will work in parallel with the networking service provider and trusted LBS server. The device/handset treated as a MO/client which may be included simple GPS, Internet connectivity handsets, and smart car with GPS installed.
(b) Trusted LBS server: The responsibility of trusted LBS server is to make the interrelation between the MO/Client's application program and the various. This server responsible for the communication varies with some services such as messaging, friend finder, store finder, restaurant finder in the form of message packets between these service providers and MO.
In addition to this, it is also reliable for the identification provision of the MO to the service providers. Thus, the trusted LBS server plays vital thing for privacy enhancing points in the of LBS system architecture.
(c) Privacy agents: The privacy agents known to the encryptions (encoders of the cryptographic codes) and decryptions (decoders of the cryptographic codes) existing on the MO including the trusted servers and the networking service providers. They also mold the coded identities (pseudonyms) that are decoded on the trusted server according to requirements.
(d) Subscription service: This will be going to permit the user to manage their pseudonyms based on Internet/data connection interface, aliases, and privacy access rules. Thus, pseudonyms' initialization, their synchronization, and coordination are established.

4 Authentication of the Users

In this, firstly the users are subscribed, secondly registered, and finally assigning pseudonym in random manner when the MO crosses the Mix-zones. Before continuing, let us go over the Mix-zones. Simply, the Mix-zones are constructed by identifying the traffic flow intersection that to circle shape or rectangle shape taking its intersection on the center. The Mix-zones' boundaries are originated and these are mainly based on total number of roads joining the junctions of the traffic signals and traffic signal flow. The Mix-zone boundaries are doorway for performing the

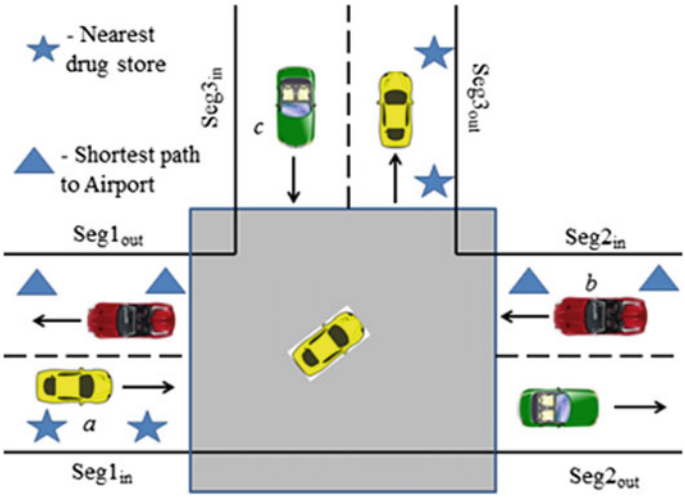

Fig. 1 Rectangular Mix-zones

process of authentication. The message packets of MO/client are changed in this authentication process.

Normally, packets will step into the Mix-zone and these MO packets will be replaced with messaging packet until it again combat with another authentication process or Mix-zone. The overall module of the authentication process is described by a Fig. 1.

5 Roadside Unit (RSU)

The multiple Mix-zones at road network are simply the roadside unit (RSU) which is shown in Fig. 2 with pseudonymization and encryption.

It can be drawn by assuming Mix-zone geometry, the user population records, and spatial and velocity constrict on the of MO motion paradigm.

The time of Internet connection and within the mobile connectivity is also accessed by the MO. The area in which the Wi-Fi, data communication links seize to exist should be considered as Mix-zones, constructed in various ways.

The Mix-zone is a virtually bounded area where the Wi-Fi communication links seize to exist. These boundaries, however, can be constructed in several ways. We are considering in circular shape as above case (Fig. 2).

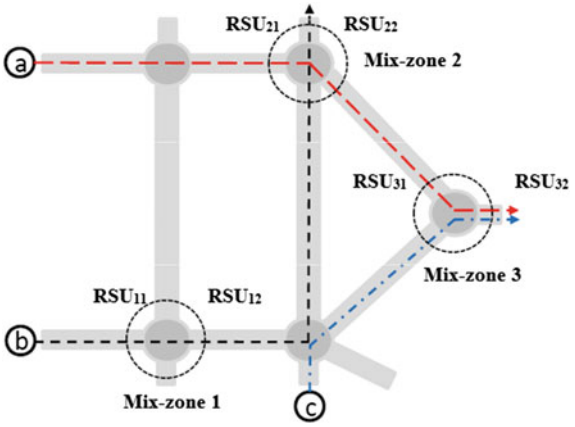

Fig. 2 Overview of the proposed secure road network containing Mix-zones

6 Encryption Scenario Considered

The process of converting a plaintext message into cipher text, again decoded back to initial message, is referred as encryption. In this algorithm, a secret key is used to encrypt data as well as also to decrypt the data. Here, the key length and type which should be used depend on encryption algorithm, and the level of security needs to be eliminated. Only the single symmetric key is used in this entire process.

By using this key, the message can be encrypted by the sender and the message can be decrypted by the recipient; there are again security problems in the circumstances.

So, the drawback of symmetry encryption is overcome by using asymmetric encryption. Two different keys are used: namely, one is a public key by which the sender can encrypt the message, and the second is a private key by which a recipient can decrypt the message. In this, we have used the asymmetry encryption algorithm as shown in Fig. 3.

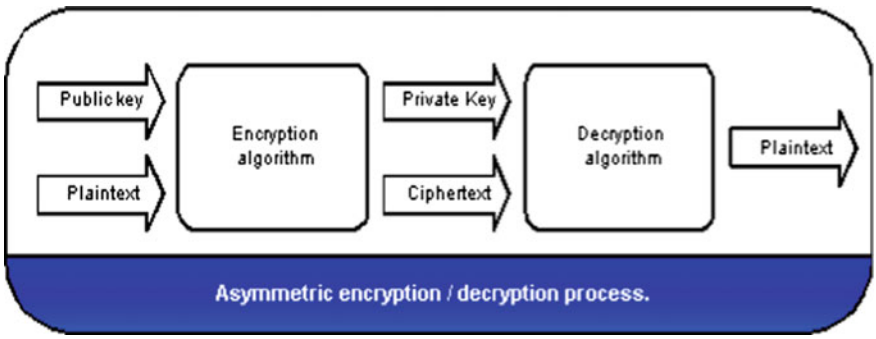

Fig. 3 Asymmetry encryption and decryption

7 Encryption and Pseudonym Generation Using MAC

Here, it provides way to understanding the integrity information must be sent through authenticate sender rather than unknown or false sender. It is very important thing in computing and communications. This process of checking the integrity of message using a secret key is known as 'message authentication codes' (MAC). In this mechanism, two parties use the MAC with a secret key so as to crosscheck the transmitted information of two parties.

In this paper, we are performing a MAC mechanism in conjunction with cryptographic hash functions.

Hence, encryption and pseudonymization generation process are described as follows:

The pseudonym calculation is performed to enhance the privacy which is existing on the MO. Consider the cryptographic function with '$C_{\mathrm{r.}}$'

The calculation of pseudonym has to be performed by the privacy module residing on the MO. We can define a cryptographic function C_{r} as,

$$C_{\mathrm{r}} = \mathrm{enc}\left(M_{\mathrm{s}}, K_{\mathrm{y}}\right) \tag{1}$$

Here, M_{s} is the message, enc is the encryption function, and K_{y} is the key.

Here, (H_{e}) entropy is randomness or uncertainty. If moving objects/clients are deterministic, entropy is zero. So,

$$H_{\mathrm{e}}\left(C_{\mathrm{r}} | M_{\mathrm{s}}, K_{\mathrm{y}}\right) = 0 \tag{2}$$

Given a decryption function 'dec,' the message is received as,

$$M_{\mathrm{s}} = \mathrm{dec}\left(\mathrm{enc}\left(M_{\mathrm{s}}, K_{\mathrm{y}}\right)\right) \tag{3}$$

Here also, the entropy becomes zero,

$$H_{\mathrm{e}}\left(M_{\mathrm{s}} | C_{\mathrm{r}}, K_{\mathrm{y}}\right) = 0 \tag{4}$$

There could be many keys for the pair (M_{s}, C_{r}); hence, it becomes possible to make $H_{\mathrm{e}}\left(K_{\mathrm{y}} | M_{\mathrm{s}}, C_{\mathrm{r}}\right) = 0$ only when there is a single key. In this work, we have assumed the same.

If we consider information theory's generally used notion for mutual information I_{m},

$$I_{\mathrm{m}}\left(C_{\mathrm{r}} | K_{\mathrm{y}}\right) \geq 0 \tag{5}$$

Equation (4) proved true here if and only if there is some guess amount of information C_{r} given relative to K_{y}, which gets satisfied here. Also,

$$I_{\mathrm{m}}\left(C_{\mathrm{r}}|K_{\mathrm{y}}\right) = 0 \tag{6}$$

where C_{r} and K_{y} must be going to independent which does not stand true. But in symmetry case here because there is information conveyed by C_{r} which relative to K_{y} is almost the similar as conveyed by K_{y} relative to C_{r}. Hence,

$$I_{\mathrm{m}}\left(C_{\mathrm{r}}|K_{\mathrm{y}}\right) = I_{\mathrm{m}}\left(K_{\mathrm{y}}|C_{\mathrm{r}}\right) \tag{7}$$

From Eqs. (5), (6), and (7), it is proved for an authentic information to get conveyed

$$I_{\mathrm{m}}\left(C_{\mathrm{r}}|K_{\mathrm{y}}\right) \gg 0 \tag{8}$$

While for perfect secrecy establishment,

$$I_{\mathrm{m}}(C_{\mathrm{r}}|M_{\mathrm{s}}) = 0 \tag{9}$$

Equations (8) and (9) are thus useful for this work. The MO/client selects a personal master key which will be going to use further to calculate pseudonym chain. Here, the pseudonyms are referring to hash message authentication codes. For the GPS-enabled MOs, the predefined encrypted calculated codes will be used as pseudonyms.

Each pseudonym is a keyed-hash message authentication code (HMAC). A hash function has an arbitrary length of bits as input and output produces as 128 bits. For example, HMAC-MD5 produces output of 128 bits and HMAC-SHA-1 output is 160 bits. Two inputs are given to HMAC: One is secret key 'K_{y}' and second one is message 'M_{s}.'

The sender and receiver must share a secret key K_{y}. Here, the size of hashing block is 64 bytes, and zeros are filled at remaining empty blocks. This key and 64-byte ipad both are now XOR-ed. This calculated result then appended by a message M_{s}; further, it is applied to HMAC. Similarly, the key is XOR-ed with opad of 64 bytes.

This is now appended with the previous result of appended 'M_{s}' which results in a 128 bit string. This result is now applied to the hash function 'H_{s}' gives us the HMAC string. Substituting the process in the equation,

We define two fixed and different strings ipad and opad as follows (the 'i' and 'o' are mnemonics for inner and outer):

$$\begin{aligned}\text{ipad} &= \text{the byte 0x36 repeated B times}\\ \text{opad} &= \text{the byte 0x5C repeated B times.}\end{aligned}$$

H_{s} (K_{y} XOR opad, H_{s} (K_{y} XOR ipad, Xi)). Here, Xi is the input string has to be converted to the HMAC string after this entire operation. If the key is longer than 64 bytes, then it is hashed and made shorter. Longer keys do not give us more

security. Take care of choosing key size must be in coordination with the hashing block to denote procedural simplicity so doing not complicate the decoding process.

8 Subscription Process

As described in the beginning of the block diagram, trusted server contains more than one pseudonym. Hence, the moving objects must be subscribed to the location-based services. Next, the number of users will be subscribed so it forms a list-containing user's aliases. But when the subscription process takes place, a tracker can track the user since the subscription process is done by the network providers and must be encrypt. Here, we used the cryptographic encryption technique named as HMAC (algorithm). Almost encryption in digital signature is similar to HMAC encryptions. By comparing both HMAC and digital signatures, HMAC is more secure, simple, and easy to perform. The entire subscription process and encryption technique is shown in below Fig. 4. Taking the care of synchronous option neither disabled nor turn off, the subscription process will be treated as pending and status showing Pending Subscription (.) by the trusted server.

The messages will get queued up to every user have been subscribed. So, every Pending Subscription (.) contains a self-identifying pseudonym. Finally, the subscription process is carried out by the trusted server and sends an acknowledgement to the waiting subscribed users. Now, the user will go to accept the request by sending acceptSubscriptionRequest (.). The message declineSubscriptionRequest (.) is used to decline the request. If user accepts the request, then it is stored, and if the

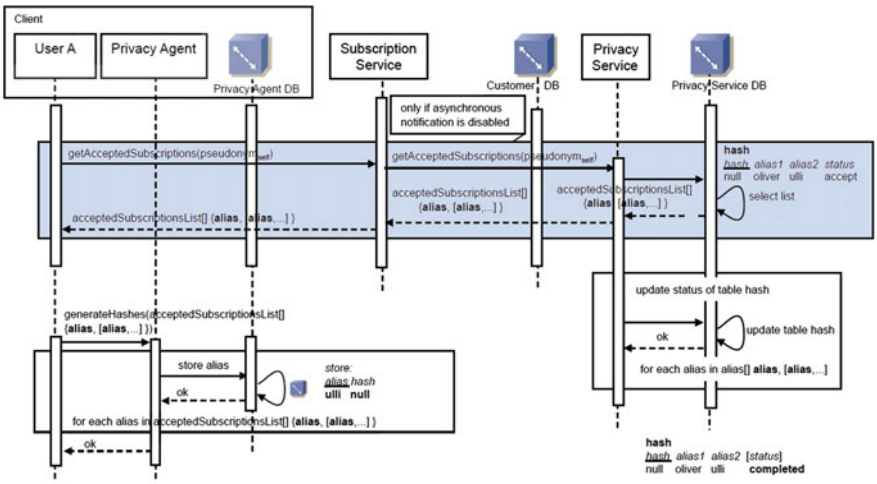

Fig. 4.8 Flow diagram for showing the completion of subscription process

Fig. 4 Flow diagram of showing the complete subscription process

particular user denies the request, then the previous pending entry of subscription is deleted. In addition to this, there is a timestamp to avoid further communication repeating denied requests of users. For accepting the requests, the alias names added into a single list for a user as message acceptSubscriptionRequest (.) sent to the privacy service. The user's status from 'pending' to 'accepted' changes by the privacy services. Then, alias name of the user is updated back. Get Pending Subscriptions (.) having the self-identifying pseudonym for the user. The privacy service translates their pseudonyms for the user want to subscribe, and 'null' will show for the denied users. Privacy services also perform alias name for initializing hash chain. Now, the trusted server sends the first request to network service provider as 'call get location.' Corresponding to the user location, lookup is initiated and immediately returns the result.

It initiates a lookup of the locations of the corresponding users and returns the result immediately. For example, a1 = HMAC (alias, trusted Server) (message 1) is being sent by the trusted server to the location/network service provider as part of the get Location (.). Each change of the subscription state induces a notification. Pending requests could be processed periodically or when the application started.

9 Conclusion

Our model's main goal is to provide the security for the users. Here, we are performing the encrypted HMAC which is carried out in the Mix-zones for the LBS users. Hence, it provides almost impossible thing for track the person's identity and location. The paper, here to make an effort for improving privacy with enhanced technique for LBS and also used the dynamic Mix-zones used here as activating and deactivating according to the traffic flow analysis. We also derived all the possible equations related to HMAC encryption with entropy and mutual information.

References

1. L. Sweeney, K-anonymity: a model for protecting privacy. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **10**(5), 557–570 (2002)
2. J. Freudiger, R. Shokri, J.-P. Hubaux, On the optimal placement of mix zones, in *PETS* (2009)
3. G. Aggarwal, T. Feder, K. Kenthapadi, R. Motwani, R. Panigrahy, D. Thomas, Approximation algorithms for k-anonymity, in *The Preceding to the 10th International Conference on Database Theory* (2005)
4. Z. Xu, H. Zhang, X. Yu, in *Multiple Mix-Zones Deployment for Continuous Location Privacy Protection* (2016)
5. B. Bamba, L. Liu, P. Pesti, T. Wang, Supporting anonymous location queries in mobile environments with privacygrid, in *WWW* (2008)
6. J. Freudiger, M. Raya, Mix-zones for location privacy in vehicular networks

Part IV
Signal Processing

EMD Analysis of EEG Signals for Seizure Detection

Mohd Hamza Naim Shaikh, Omar Farooq and Garima Chandel

Abstract The Electroencephalogram (EEG) is the electrical signals which contain the information related to activities within the brain. Neurological disorders such as epilepsy can be diagnosed effectively by analyzing EEG signals. In the present work, the empirical mode decomposition (EMD) is applied to EEG recordings for the automated detection of seizures in epileptic patients. For this purpose, intrinsic mode functions (IMFs) from the EMD are processed to extract the features from normal and seizure EEG signals. The extracted features are ranked to select the useful features for classification. The classification was done by using these selected features by Artificial Neural Network (ANN). The EEG dataset used in the present study is the well-known publicly available Bonn University EEG dataset. Three different classification problems were done by using the sets of this data. The classification accuracy achieved for these three cases were 96.1, 96.4, and 99.3%.

Keywords EMD · IMF · Seizure · EEG · ANN

1 Introduction

Epilepsy is the second most common neurological disease [1]. Its clinical symptoms may involve the loss of awareness, drop attack of facial muscles and eye movements, aggressive outbursts, etc. Seizures in epileptic patients are due to syn-

M. H. N. Shaikh (✉)
Department of Electronics and Communication Engineering, IIIT, Delhi, India
e-mail: shamzanaim@gmail.com

O. Farooq
Department of Electronics Engineering, A.M.U., Aligarh, India
e-mail: omar.farooq@amu.ac.in

G. Chandel
Department of Electronics and Communication Engineering,
I.T.S. Engineering College, Greater Noida, India
e-mail: garimachandelsn@its.edu.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_16

chronous discharge by the groups of neurons, creating a large amplitude signal and leading to uncontrollable oscillations. It is important to diagnose seizures as it affects drastically the daily life of patients and their family members. One of the conventional methods for seizure detection is by visually interpreting the EEG signals. But the major limitations of this method are the need of trained neurologists (24×7 basis), time-consuming, and probability of human errors due to multi-channel EEG.

Hence, in recent years, computer-aided automatic system is developed for automatic seizure detection to warn the patient and their caregivers by raising alarm.

Therefore, patients could avoid activities for some time (such as driving a vehicle, using stairs, horse riding) that may cause harm and quickly administer Anti-Epileptic Drugs (AEDs).

Various automatic seizure detection techniques have been introduced using EEG records in previous years [1–5]. A multidimensional probability evolution-based technique was proposed by McSharry et al. [1] which gave less false alarms in comparison to using variances. Although the group found less difference between the two methods and interpreted, the seizure detection using linear statistics such as variance is not effective. Mohseni et al. [2] detected seizures by variance threshold from windows for classification of normal and seizure EEG activities. Adeli et al. [3] suggested that chaotic features are effective in classifying different types of EEG signals including seizure activities. For automatic seizure detection, they extracted largest Lyapunov exponent (LLE) and Correlation Dimension (CD) from wavelet sub-bands. Guler and Ubeyli [4] used Lyapunov spectra for seizure detection in epileptic patients and employed multiclass support vector machines (SVMs) for classification. Schneider et al. [5] used fractal dimension-based feature and ANN for seizure detection. Tzallas et al. [6] extracted features using various time–frequency techniques such as smoothed pseudo-Wigner–Ville. High classification accuracy was achieved by implementing ANN.

Recently, the EMD technique has drawn the attention of researchers for the analysis of nonlinear signals. The major advantage of EMD analysis is that it does not require any assumption regarding stationarity and linearity of the signals [7]. Since the EEG signals are nonstationary in behavior, many algorithms have been developed for seizure detection using EMD technique. The group [8] used IMF-based features for effective classification between ictal and non-ictal EEG signals, that is, seizure and non-seizure EEG activities. The energy of an IMF and minimum distance duration is used as discriminating features for seizure detection.

In the present work, an algorithm is proposed by using online available EEG dataset for detection of normal and seizure EEG. The EMD-based features on different IMFs are extracted from EEG signals of Bonn University dataset. To select the relevant features and reduce the feature vector dimension, ranking was done. Finally, the selected features are fed to ANN for classification of different EEG activities. Three different classification problems were considered in the study by using the five types of EEG recordings from the dataset. These three cases of classification are explained in Sect. 2.

2 Materials

For the present study, Bonn University database [3] is used. The dataset consists of five subsets A, B, C, D, and E. The recordings in each subset consist of 100 segments of single-channel recording. Each segment is of 23.6 s duration. The sampling frequency is taken as 173.61 Hz in this dataset. The recording location of signals has been randomized with montage being the common average reference. Sets A and B consist of surface EEG recording from five healthy subjects with closed and open eyes. Sets C and D contain recordings of seizure-free interval of five epileptic patients. And Set E contains seizure activity from the same patients, selected from all recording sites showing ictal activity. Here, we have considered the following three cases for classification of features: Case I is for the classification between non-seizure from all four sets and seizure EEG, i.e., Sets A, B, C, D versus Set E. Case II is for the classification between all healthy EEG record and seizure EEG, i.e., Set AB versus Set E. Case III is for the classification between healthy record and seizure EEG, i.e., Set A versus Set E.

3 Methods

The framework of the proposed method for seizure detection is shown in Fig. 1. Firstly, the EEG signal is fed to the preprocessing stage to remove the artifacts. Preprocessing involve high-pass filtering at 0.1 Hz followed by decomposing the signal into intrinsic modes. Then features are extracted and are ranked using the statistical *t*-test. The reduced feature set is then fed to a two-layer feedforward neural network.

Algorithm 1

Residue = s (t)
I1 (t) = Residue
i = 1, k = 1
while Residue = 0 or Residue is not monotone

while Ii has non-negligible local mean

U (t) = spline through local maxima of Ii
L (t) = spline through local minima of Ii
Av (t) = 1 2 (U (t) + L (T))

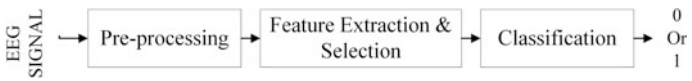

Fig. 1 Framework of the proposed automatic seizure detector. 0 output represents non-seizure class and 1 output represents seizure class

Ii (t) = Ii (t) − Av (t)
i = i + 1

end
IMFk (t) = Ii (t)
Residue = Residue − IMFk
k = k + 1

end

3.1 Empirical Mode Decomposition (EMD)

EMD is a recently developed method for the analysis of nonstationary signals [7, 8] developed by N. Huang. This method decomposes a signal into a sum of components, each with slowly varying amplitude and phase. EMD being an adaptive method, the decomposition is specific to the signal being analyzed. In theory, the output components should be separate phenomena based on different timescales. The components of the EMD are called an IMFs and these are derived using the basis obtained from the data itself.

The IMFs satisfy two criteria, in order to resemble a generalized Fourier decomposition: The number of zero crossing and the number of extremas (minima and maxima) should differ at most by one and an IMF must have a zero "local mean."

The procedure of extracting IMFs is as follow:

- First, subtract the mean of upper and lower envelope from the original signal, and then check the conditions for IMF.
- For next IMF, subtract the previous IMF from the original signal; the difference is the residue for the outer loop of the process.
- The construction of an IMF is now repeated for the residual signal.
- This EMD procedure is repeated until the residue is either constant or monotone. This final residue should reveal any trend that exists in the data.

Schematically, a pseudocode for this procedure for a signal $s(t)$ is shown in algorithm 1.

The procedure results in the decomposition

$$s(t) = \sum_i \text{IMF}_i + \text{residue} \tag{1}$$

Figure 2 shows the sample plot of the various IMFs and the original EEG record. These are from a healthy subject. They are just for the sake of example to visualize

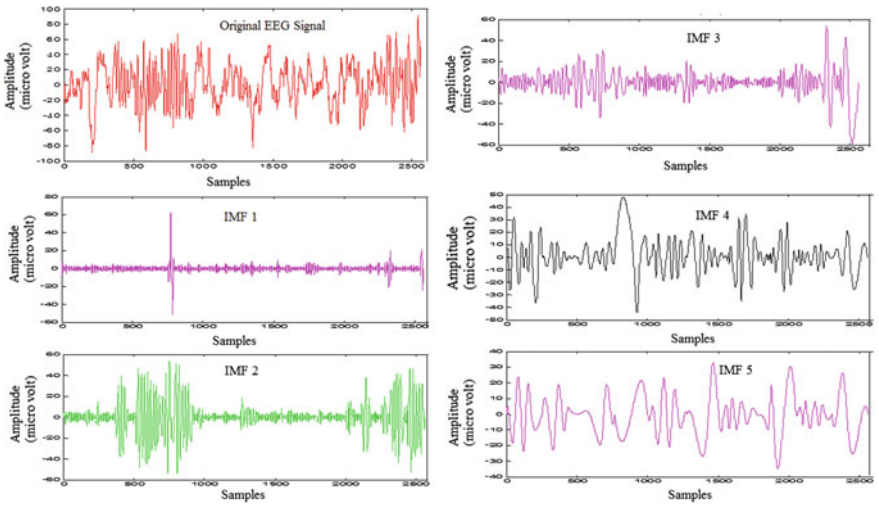

Fig. 2 EEG signal from Set A with five IMFs

the decomposition. It is clear from Fig. 2 that the oscillations are slowing down as we proceed from first IMF gradually.

3.2 Feature Extraction and Selection

The features were extracted on epoch of 200 samples to capture the characteristics of different EEG activities from all the subsets. These features are Energy, Entropy, Mean, Variance, Standard Deviation, Skewness, Kurtosis, Geometric Mean, Harmonic Mean, Waveform Length, Interquartile Range (IQR). These are calculated for all the five IMFs. Thus in total for each epoch, we have fifty-five features of each sample point. Further, these features are ranked according to their *t*-test scores to select useful features and to discard irrelevant and redundant features for classification between different EEG activities.

3.3 Classification

The selected features were fed to ANN for classification between different EEG activities. ANN is an assembly of neurons interconnected. It can be viewed as weighted directed graphs. The nodes here are artificial neurons and the directed edges (with weights) are connections [5]. Multilayer feedforward networks can be designed to form arbitrarily complex decision boundaries, and thus, they can represent any typical function. Back-propagation learning algorithm to evaluate the

weights has made multilayer feedforward network as one of the most popular of among the networks [6].

The performance measure of binary classifier namely accuracy, sensitivity, and specificity are observed. In classification, we have taken two class viz. seizure and non-seizure. Sensitivity is the effectiveness of the detecting seizure by the algorithm, whereas specificity represents that non-seizure is recognized as non-seizure by the algorithm and low specificity leads to more false alarms. Accuracy shows the overall performance of the algorithm to detect correctly both the classes.

4 Results and Discussion

In this work, Bonn University database was used. The epochs of 200 samples were extracted by using non-overlapping rectangular window for each of the segments from all subsets.

The feature set consists of 11 features of 5 IMF, so we have in total 5 * 11 = 55 features per epochs. These 55 features are then ranked accordingly and first 15 features are taken for classification purpose. Hence, dimension of feature vector was reduced from 55 to 15 for each epoch. Figure 3 represents the top 15 features for all the three classification cases, and Table 1 describes the names of top-ranked features. The feature ranking shows that Waveform Length, Interquartile Range (IQR), and Standard Deviation are three best features for this dataset.

In classification stage, the selected feature set from non-seizure and seizure classes is divided into three parts, i.e., training, validation, and testing. We have taken 70, 15% of the total features from each class for training and validation, respectively. The remaining 15% of features are used for testing the network. The results for all the three cases under study are shown in Table 2. This table shows the efficacy of the present algorithm to discriminate between non-seizure and seizure EEG activities. The Case III achieved the best results, and it is the case of classification between non-seizure EEG from normal subjects and seizure EEG from

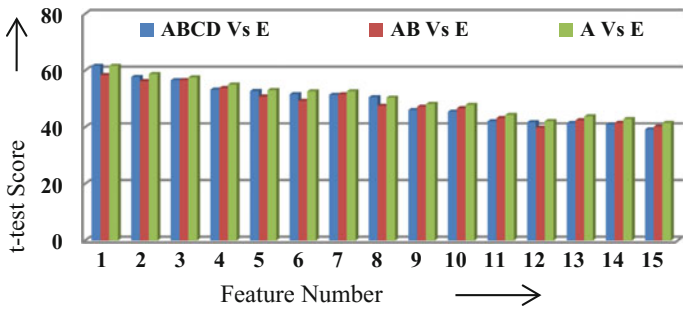

Fig. 3 Top 15 features selected for all classification problems under study

Table 1 Selected features

Feature No.	Feature name/IMF No.	Feature No.	Feature name/IMF No.	Feature No.	Feature name/ IMF No.
1	Waveform length/2	6	Std. deviation/1	11	IQR/4
2	Std. deviation/2	7	IQR/3	12	IQR/1
3	Waveform length/3	8	Waveform length/1	13	Std. deviation/5
4	Std. deviation/3	9	Waveform length/4	14	Waveform length/5
5	IQR/2	10	Std. deviation/4	15	IQR/5

Table 2 Classifier performance for Bonn University database showing the three cases

Case	Spe (%)	Sen (%)	Acc (%)
Case I—Set ABCD versus Set E	96.5	93.6	96.1
Case II—Set AB versus Set E	96.7	95.7	96.4
Case III—Set A versus Set E	99	99.7	99.3

Table 3 Comparison of proposed method with other available methods using same database

Papers	Set	Method used	Acc (%)	Sens (%)	Spe (%)
Sharmila and Geethanjali [9]	ABCD/E	Wavelet-based features and NB	95.85	89.92	97.31
This work	ABCD/E	EMD and ANN	96.1	93.6	96.5
Lee et al. [10]	A/E	Phase-space reconstruction and NN	98.17	96.3	–
This work	A/E	EMD and ANN	99.3	99.7	99

epileptic subjects. This represents that the algorithm could be used to detect the epileptic patients efficiently.

Table 3 illustrates the comparison of the presented method with the existing methods. For comparison purpose, the work from other researchers by using same database and same case of classification was considered. The results showed better performances for both the cases. Sharmila and Geethanjali [9] used wavelet-based features and Naïve Bayes classifier for Case I. In their methods specificity was better but with less sensitivity. Our method detected seizure data more effectively with sensitivity of 93.6%. Also for Case I, the proposed EMD-based method performed well over the method based on phase-space reconstruction [10].

5 Conclusion

The conclusion of the present work is that for the analysis of EEG signals EMD can be used for analysis of EEG signals to detect seizures effectively. In this study, ANN was employed for classification of non-seizure and seizure EEG. In future, the proposed method may be used for analysis of EEG signals from other datasets also.

References

1. P.E. McSharry, T. He, L.A. Smith, L. Tarassenko, Linear and non-linear methods for automatic seizure detection in scalp electroencephalogram recordings. J. Med. Biol. Eng. Comput. **40**, 447–461 (2002)
2. H.R. Mohseni, A. Maghsoudi, M.B. Shamsollahi, Seizure detection in EEG signals: a comparison of different approaches. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 6724–6727 (2006)
3. H. Adeli, S.G. Dastidar, N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy. IEEE Trans. Biomed. Eng. **54**(2), 205–211 (2007)
4. I. Guler, E.D. Ubeyli, Multiclass support vector machines for EEG signals classification. IEEE Trans. Inf. Technol. Biomed. **11**(2), 117–126 (2007)
5. M. Schneider, P.N. Mustaro, C.A.M. Lima, Automatic recognition of epileptic seizure in EEG via support vector machine and dimension fractal. Proc. Int. Joint Conf. Neural Networks 2841–2845 (2009)
6. A.T. Tzallas, M.G. Tsipouras, D.I. Fotiadis, Automatic seizure detection based on time-frequency analysis and artificial neural networks. Comput. Intell. Neurosci. **2007**, 80510-1–80510-13 (2007)
7. R. Djemili, H. Bourouba, M.A. Korba, Application of empirical mode decomposition and artificial neural network for the classification of normal and epileptic EEG signals. Biocybern. Biomed. Eng. **36**(1), 285–291 (2016)
8. L. Orosco, E. Laciar, A.G. Correa, A. Torres, J.P. Graffigna, An epileptic seizures detection algorithm based on the empirical mode decomposition of EEG. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2651–2654 (2009)
9. A. Sharmila, P. Geethanjali, DWT based epileptic seizure detection from EEG signals using Naïve Bayes/k-NN classifiers. IEEE Access **4**, 7716–7727 (2016)
10. S.H. Lee, J.S. Lim, J.K. Kim, J. Yang, Y. Lee, Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance. Comput. Methods Programs Biomed. **116**(1), 10–25 (2014)

Novel Biometric Modalities

Jaspreet Kour, Ankita Katiyar and Gauri Katiyar

Abstract Biometrics is a study of physiological and behavioral characteristics of a person for the purpose of verification. There is an increase in number of biometric traits involved in biometric authentication system depending on the application and the demand in the market. Each modality has its own advantages and disadvantages, and no method is able to satisfy the security needs in completeness. Therefore, there is a continuous effort to look for novel and emerging modalities that are able to provide high security, robust, and also cost-effective. This paper presents some of the novel biometric modalities.

Keywords Security · Biometrics · Modalities · Multimodal

1 Introduction

Every human being is unique, and this uniqueness leads to biometrics. It is a science which verifies or authenticates a person based on physiological or behavioral characteristics. Biometrics is reliable and more capable than traditional knowledge-based and token-based techniques. Biometrics is secure and reliable measure of security. Physiological traits, also termed as passive characteristics, include fingerprint, palm print, iris, knuckle print to name a few, whereas behavioral traits also called as active characteristics include signature, gait, keystroke dynamics. The multimodal biometric system utilizes more than one biometric

J. Kour
EIE Department, GCET, Gr. Noida, UP, India
e-mail: tojaspreet@gmail.com

A. Katiyar
CSE Department, VIT, Vellore, Tamil Nadu, India
e-mail: ankita.katiyar2013@vit.ac.in

G. Katiyar (✉)
EEE Department, ITS Engineering College, Greater Noida, UP, India
e-mail: gaurikatiyarrl@its.edu.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_17

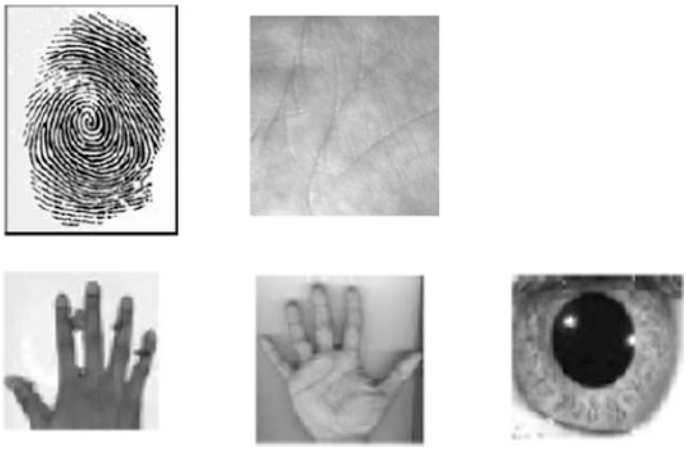

Fig. 1 Conventional biometric modalities

modality. Unimodal biometric system based on a single trait suffers from various limitations like noise, intra-class variation, spoof attack, distinctiveness. Multiple data can be collected from multiple sensors, multiple traits, multiple instances, multi algorithms, or multiple samples. The advantage of multimodal biometrics is that it improves performance and robustness of the system (Fig. 1).

In addition to mature and wider used traits, there are some emerging modalities in various stages of development and implementation. Each modality has its pros and cons as no trait is impeccable. The biometric industry is evolving to move beyond conventional and traditional physiological attributes to new technology relying on physiological and behavioral modalities.

2 Novel Modalities

2.1 Gait

Recognition using gait is a novel biometric that depends on the walk of an individual and stems from the fact that gait of a person is unique. Gait data is obtained from the video of walking subject. Gait as a biometric has various advantages: non-contact biometric; obtained from distance; noninvasive. The challenges associated with this trait are light conditions, surface condition, viewing angle, carrying object or behavioral changes in the person due to physiological challenges or illness leading to variations in gait [1] (Fig. 2).

Fig. 2 Gait

2.2 Eye Movements

Promising results have been obtained for biometric verification using eye movements. Eye-tracking device that consists of two video cameras catches the movement of pupil of each eye. Eye movements which can be used as biometric can be produced in different ways. One of them is that the subject follows a horizontally jumping light dot on screen with his gaze for 60 s (Fig. 3).

Fig. 3 Eye movement [2]

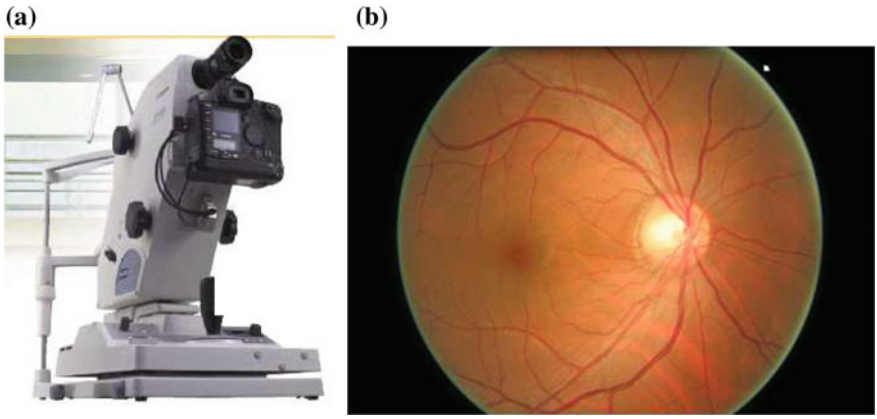

Fig. 4 **a** Retinal camera, **b** color retinal image [3]

2.3 Retinal Vascular Pattern

Retina vascular pattern has huge potential to be one of the most secure biometric traits. Noninvasive color retinal cameras capture retinal images. The map of branch, crossover point, and bifurcation of blood vessels in retina are the characteristics which contribute to the uniqueness of retinal images. The unique invariant features are constructed from these three characteristics. This novel biometric has the advantage that it is a stable physiological characteristic which does not change with age; it is unlikely to be affected by environmental conditions and is highly distinctive. The drawbacks it suffers are that it may be uncomfortable to have retinal scan for some people, some diseases may damage retina, retinal scammers are expensive, and data collection is tedious process (Fig. 4).

2.4 Palm Finger Joints

Hand is used as conventional biometric where fingerprint, palm print, knuckle print are the areas used as trait [4]. A new hand-based biometric extracted from the hand image is palm finger joint surface which is one-third of knuckle print having metacarpo-phalangeal joint only (Fig. 5).

2.5 Heart Sound

The natural physiological signal produced by sudden closure of atrioventricular and semilunar valves termed as heart sound is a new biometric modality. It has the

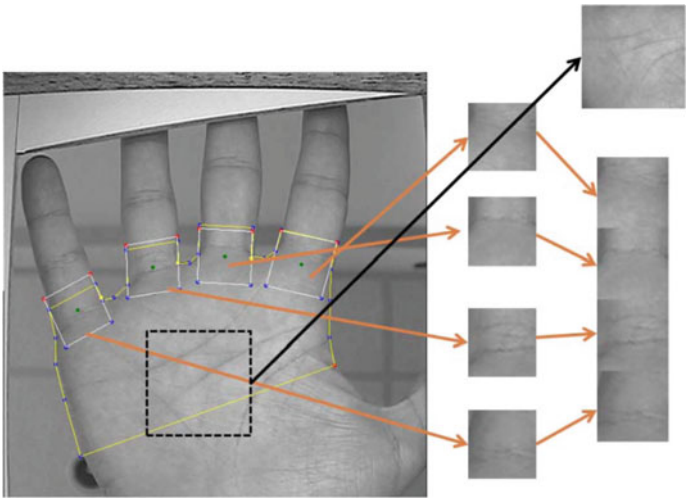

Fig. 5 Palm finger joint surface biometrics

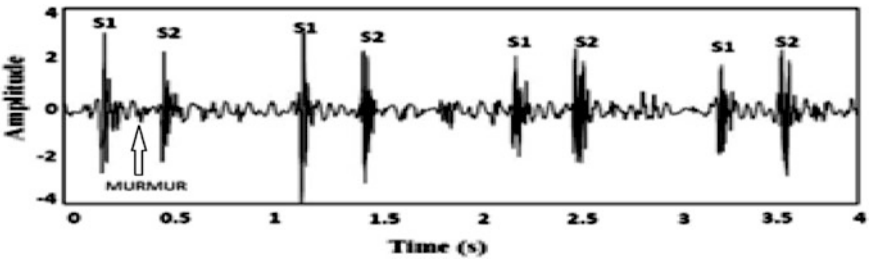

Fig. 6 Time-domain heart sound signal

advantage that it cannot be reproduced by any artificial means. Heart sound consists of three parts: first heart sound (*S*1), second heart sound (*S*2), and murmur. The peaks and valleys of heart sound carry important information. A digital stethoscope is used to collect heart sound and can be stored in WAV format. Heart cycle is a quasi-periodic signal with period of 0.7–0.8. The noise in heart signal is removed before processing it for extracting suitable features [5] (Fig. 6).

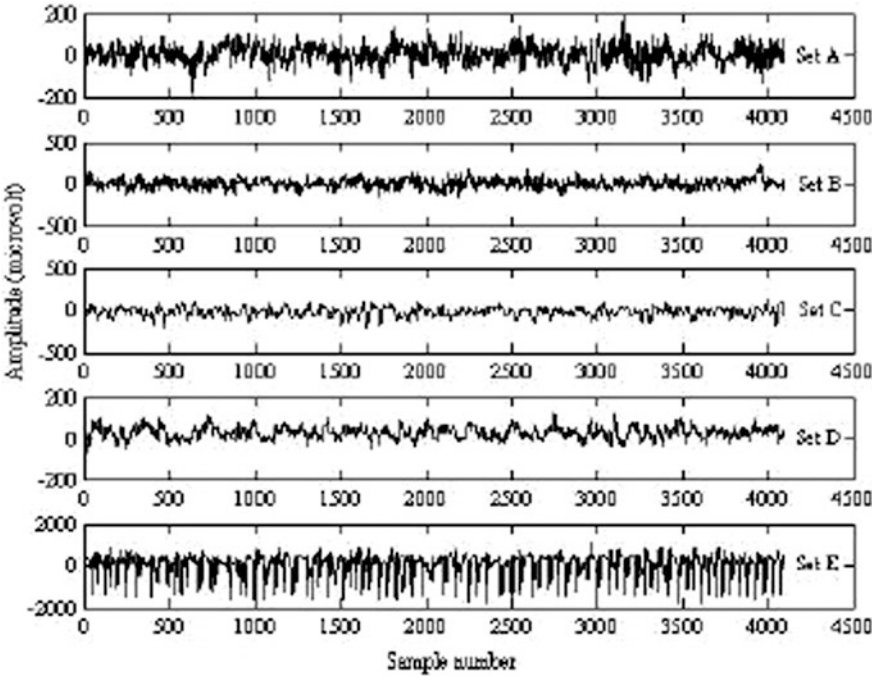

Fig. 7 Overt EEG signal

2.6 *Overt Mental Stimuli Brain Signal*

One of the emerging biometric identifiers is electroencephalogram (EEG) which is basically the graphical representation of electrical activities of brain. These signals are obtained by placing the electrodes on the scalp which record the brain signals. There are frequency bands in EEG signal which have specific frequency range. Alpha wave, theta wave, and beta wave are the bands, and each band is characterized by certain dominant mental activities [6]. These bands can be used individually, or combination of bands is used for person authentication of overt stimuli brain signal. The brain pattern of an individual has genetic information which is unique for an individual thus serving as secure biometric trait. Also, the procedure of collecting data is painless and noninvasive. It is non-vulnerable to spoof attack and is applicable to alive human beings (overt) (Fig. 7).

2.7 *Ocular Surface Vasculature*

Eye has been used as biometric modality since long, where iris and retina were the areas used as trait unique to an individual. It is followed by new physiological

Fig. 8 Ocular surface vasculature [7]

Fig. 9 Palm vein

modality termed as ocular surface vasculature. Our eye has blood vessels on the white portion which are visible and carry unique pattern rich in texture information. OSV features have been recently used for biometric authentication as they are universal, distinct, collectable, and non-circumventable which are the characteristics required for a biometric authentication.OSV region is extracted from the eye image, and features are extracted and given to a classifier (Fig. 8).

2.8 Palm Vein

The palm vein as biometrics is based on blood vascular structure. The multispectral palm vein database is obtained by capturing palm image under blue, green, and red lights. The blood vessels are classified into two: artery and vein [8]. The edges of vein can be extracted by canny edge detector which acts as features for the recognition system (Fig. 9).

2.9 Dental Radiograph

The condition of teeth, roots, jaw, and facial bone composition is used in dental biometrics. The radiographs are acquired and used in dental biometrics. Generally, this modality is used in forensics. This method is highly intrusive, and X-ray

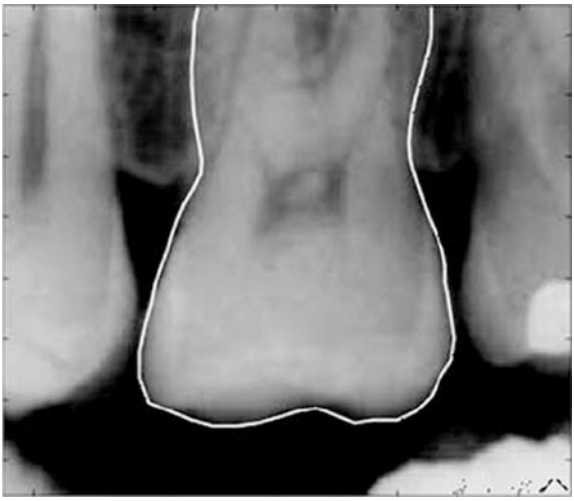

Fig. 10 Dental radiograph [9]

radiation may pose health hazards to individuals. Some of the subjects may have missing teeth which results in mismatching of the features extracted (Fig. 10).

3 Multimodal Biometrics

A biometric system using only one human characteristic for verification/ authentication is termed as unimodal biometric system. Such a system is simple, cheap, requires less storage space, etc., but suffers from various disadvantages: intra-class variations, noise in the input data, non-universality, spoofing, etc. In order to overcome the disadvantages of unimodal systems, we have multimodal biometric system in which more than one biometric modality is used; these systems increase the performance of its system and robustness. It is hard to spoof multi-modal systems. The data from different modalities can be combined using various fusion techniques. Fusion in multimodal biometric systems can be done at sensor level, score level, etc. Multimodal biometric systems are certainly superior to unimodal systems in terms of performance and have lower error rates.

4 Conclusion

The physiological and behavioral characteristics are being explored to be used as biometrics. Some advanced human features are described to be unique and experimented by the researchers. There are problems encountered with new traits such as expensive equipments, time-consuming enrollment process, high

intrusiveness, testing of new modalities on large database. Multimodal biometrics comes as a solution to the problems with unimodal biometrics as it increases the performance of the system and makes it more secure.

References

1. S.M. Darwish, Design of adaptive biometric gait recognition algorithm with free walking directions. IET Biom. **6**(2), 53–60 (2016)
2. Y. Zhang, M. Juhola, On biometrics with eye movements. IEEE J. Biomed. Health Inf., pp. 99 (2016)
3. A. Bhuiyan, A. Hussain, A. Mian, T.Y. Wong, K. Ramamohanarao, Y. Kanagasingam, Biometric authentication system using retinal vessel pattern and geometric hashing. IET Biom. **6**(2), 79–88 (2016)
4. O. Makul, M. Ekinci, Palm and palm finger joint surfaces based multibiometric approach, in *Proceedings—International Conference on Telecommunications and Signal Processing*, pp. 665–668 (2016)
5. G. Gautam, K. Kumar, Biometric system from heart sound using wavelet based feature set, in *Proceedings of International Conference on Communication and Signal Processing IEEE*, pp. 551–555 (2013)
6. W.M. Rahman, M. Gavrilova, Overt mental stimuli of brain signal for person identification, in *Proceedings of International Conference on Cyberworlds*, pp. 197–203 (2016)
7. P.S. Tankasala, P. Doynov, S. Chrihalmeanu, R. Derakhshani, Ocular surface vasculature recognition using curvelet transform. IET Biom. **6**(2), 97–107 (2016)
8. S.D. Raut, V.T. Humbe, A novel approach for palm vein feature extraction using Gabor and canny edge detector, in *Proceedings of IEEE International Conference on Computational Intelligence and Computing Research*, pp. 1–4 (2015)
9. M. Hofer, A.N. Maranara, Dental biometrics: human identification based on dental work information, in *IEEE Proceedings of the XX Brazalian Symposium on Computer Graphics and Image Processing*, pp. 288–286 (2007)

Handwritten Character Recognition—An Analysis

Usha Tiwari, Monika Jain and Shabana Mehfuz

Abstract Handwriting recognition has always been an active and challenging research area with various applications in daily life and also in area of pattern recognition. Some of the applications are reading tool for blind people, reading handwritten bank cheques, by converting it into properly structured text which can be easily detected by the designed algorithms. This paper gives brief analysis of various handwritten recognition techniques, such as Optical Character Recognition (OCR), Artificial Neural Network (ANN), Intelligent Character Recognition (ICR), and Intelligent Word Recognition (IWR). Accuracy rates of all these methods were compared, and the comparative analysis demonstrates that OCR method is the best among them for the recognition of English handwritten characters. Merits and demerits are also discussed for the methods of recognition. Basic steps involved in the handwritten recognition process are also briefed.

Keywords Optical Character Recognition (OCR) · Neural network topologies Back propagation algorithm · Feature extraction process

1 Introduction

With the growing computational power, demand for character recognition methodologies has increased for various applications. It is a difficult task to develop a practical system of handwritten character recognition with high accuracy of

U. Tiwari (✉) · S. Mehfuz
Jamia Millia Islamia, New Delhi, India
e-mail: ushapant@rediffmail.com

S. Mehfuz
e-mail: smehfuz@jmi.ac.in

M. Jain
Department of Electrical and Electronics Engineering,
I.T.S. Engineering College, Greater Noida, India
e-mail: monikajain.bits@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_18

recognition. In the existing systems, the accuracy of recognizing the text depends immensely on the quality of the input document. The process of scanning and recognizing static images of the characters is generally called as Offline Character Recognition and also called as Optical Character Recognition (OCR) [1]. To increase the accuracy of system, different classifiers use uppercase and lowercase English alphabets. Handwriting recognition can be divided into two types: online and offline. As discussed in [2], online method is based on the pen trajectory data while offline method relies on the pixel data only. Online method has an advantage that for spatially overlapped characters, segmentation is comparatively easier than offline method. First, the handwritten or printed text is converted into the machine-readable form with the help of OCR system. Basic steps of character recognition are:

1.1 Preprocessing

Preprocessing is the first and the major step of OCR software. At this stage, certain operations are performed on the scanned image, i.e., de-skew, converting an image from color to black and white, cleans up non-glyph boxes and lines, identifies columns, paragraphs, captions as different blocks, and normalization.

1.2 Segmentation

The process of assigning a label for every available pixel in a particular image, so that same set of pixel possessing some common visual characteristics shares same type of label, is called segmentation. This process is generally used for finding the boundaries in a particular image like curves and lines in image. Edge detection is used in this as a method of segmentation.

1.3 Feature Extraction

For acquiring the main traits of input symbols, feature extraction process is used. This process leaves some unimportant traits and extracts the main features. For achieving high recognition performance, the key factor is selection of appropriate feature extracting method [3].

1.4 Classification and Recognition

Classification is the process of choosing character and then assigning it to proper character class to which it belongs. Classification is done by the decision-making unit which uses the features previously extracted by earlier processes [4].

1.5 Postprocessing

The main function of this block is to print the text which is already recognized in previous stages. This is done by finding the ASCII value of input samples available [5] (Fig. 1).

2 Literature Review

Earlier, creating reading devices and expanding telegraphy for the blind people were done by OCR method [6]. Emanuel Goldberg, a renowned scientist in 1914, designed a new device which converts the character (by reading) into some standard codes (telegraphs). During that time, scientist put hard in process of developing a portable machine or scanner-type device which if moved over the written material generated some unique sounds or tones, which help in recognizing the character in handwritten material. But it failed to read non-optical characters for which different researches are done in past. M. Sheppard in 1951 [7] developed Intelligent Character Recognition (ICR), which is one of the specific recognition systems that can be used for variety of fonts and different kind of handwriting samples and helps in increasing the recognition rate and accuracy of designed system.

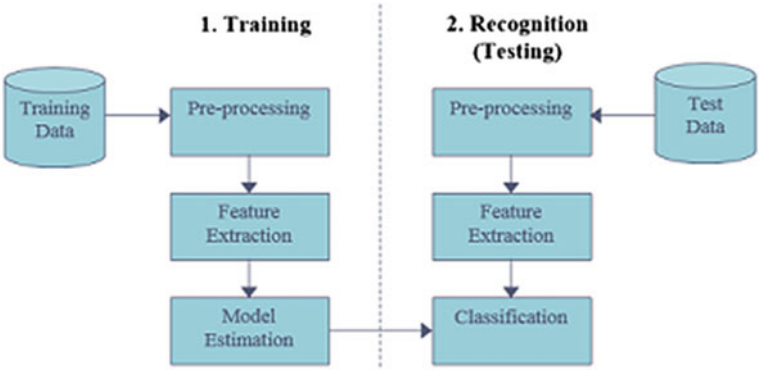

Fig. 1 OCR block diagram

Most ICR systems are incorporated with a self-learning mechanism called as Artificial Neural Network, ANN, which basically updates the database for every handwritten samples received. It improves the system by detecting handwritten samples in place of printed samples, which is the main function of OCR [8]. It can achieve 97%+ accuracy rates; sometime the accuracy is not good in reading the handwritten data in structured forms. Intelligent Word Recognition (IWR) is very effective system not only for printed handwritten material but also for various types of handwriting forms like cursive writing and others [3].

Some of the recently used techniques are listed below:

2.1 *Artificial Neural Networks*

Various topologies such as Radial Basis Function, Back Propagation, and Nearest Neighbor Network of Neural Network are used for the process of classification. ANN basically learns from the various image training sets available [9]. After training, it is used for the process of character identification. Then for the new image entered, ANN basically finds the similarity with the training images and then conclude the result on the basis of similar factors. The neural network has three layers: an input layer consisting of 100 nodes (for the 10 by 10 letter input), a hidden layer which consists of 50 nodes, and an output layer with 26 nodes (each letter) as shown in Fig. 2.

2.2 *Radial Basis Function*

It is a type of function which depends on a point which is basically distance from particular location. For getting good results, it requires 1800 neurons on validation data sets [10].

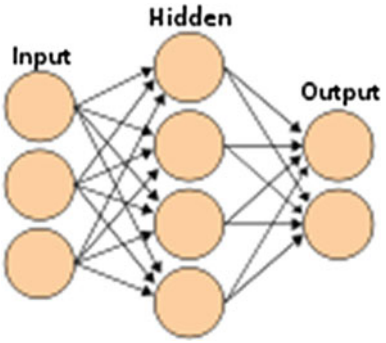

Fig. 2 Artificial Neural Network

2.3 Statistical Methods

Character recognition fields are mainly divided in two types: nonparametric and parametric [11]. In nonparametric recognition, the pattern is identified by the usage of cluster, which has a center at particular short distance from the pattern over the set of clusters. Any prior information about the data is not required. In parametric recognition, some priori information is available, and by that a parametric model for each and every character can be obtained.

2.4 Structural Methods

In the area of structural recognizing, a method called syntactic is considered one of the most prevalent approaches. Concept of grammar can be used to find the similarity content in structural components. The character composition is given on the basis that each individual class of variable is having its own grammar definition. The strings or trees representation can be used to represent grammar.

2.5 Feature Extraction Techniques

For specifically handwritten character, Diagonal-Based Technique is used for the feature extraction [4]. Gupta et al. [10] used the potential feature like Horizontal/Vertical strokes and end points for recognition and reported an accuracy of about 90.50% for handwritten Kannada numerals. Zoning, Template matching, *N*-tuples, Moments, Crossing, and Distances are some of the widely used techniques.

Above section gives the brief description of various handwritten detection techniques used in past followed by some recent techniques used for detection of handwritten data of various types of languages other than English also. Accuracy of methods and other parameters is also compared.

3 Conclusion

This paper gives an analysis of various recognition techniques for handwritten text images. For character recognition, the use of ANNs has also been described which results in high noise tolerances at the output as compared to other methods. Comparative analysis shows that excellent results are obtained by using these systems. Most important step of OCR is feature extraction. A set of features should be properly chosen to yield good neural network classification rates. While analyzing all methods, it is found that among the other methods, OCR method is one of

the best methods specifically for English text and cannot be replaced with other methods. In this by using the distance among the connected component, probability of word boundary is estimated. This is done by combining segmentation and recognition distance. The past work discussed will provide the better understanding of the method and will help the new researchers who are keen to work in this field for the betterment of the new technology.

References

1. V.L. Sahu, B. Kubde, Offline handwritten character recognition techniques using neural network: a review. Int. J. Sci. Res. (IJSR) **2**(1), (Jan 2013)
2. R. Seiler, M. Schenkel, E. Eggimann, Off-line cursive handwriting recognition compared with on-line recognition, in *IEEE Proceedings of ICPR '96*, 1015-4651/96 (1996)
3. J. Pradeep, E. Srinivasan, S. Himavathi, Diagonal based feature extraction for handwritten alphabets recognition system using neural network. Int. J. Comput. Sci. Inf. Technol. (IJCSIT) **3**(1) (Feb 2011)
4. N. Arica, F. Yarman-Vural, An overview of character recognition focused on off-line handwriting. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **31**(2), 216–233 (2001)
5. K. Prasad, D.C. Nigam, A. Lakhotiya, D. Umre, Character recognition using matlab's neural network toolbox. Int. J. u- and e- Serv. Sci. Technol. 6(1) (Feb 2013)
6. N. Arica, F.T. Yarman-Vural, An overview of character recognition focused on off-line handwriting. IEEE Trans. Syst. Man Cybern.—Part C: Appl. Rev. **31**(2) (May 2001)
7. B. Verma, M. Blumenstein, S. Kulkarni, Recent achievements in offline handwriting recognition systems
8. R. Plamondon, S.N. Srihari, On-line and off-line handwriting recognition: a comprehensive survey. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1) (Jan 2000)
9. J. Pradeep, E. Srinivasan, S. Himavathi, Neural network based recognition system integrating feature extraction and classification for English handwritten. Int. J. Eng. **25**(2), 99–106 (May 2012)
10. A. Gupta, M. Srivastava, C. Mahanta, Offline handwritten character recognition using neural network, in *International Conference on Computer Applications and Industrial Electronics (ICCAIE-2011)*
11. T. Ahmad, A. Jameel, B. Ahmad, Pattern recognition using statistical and neural techniques. 978-1-61284-941-6/11/$26.00<S>20 11 IEEE

Design of Low-Power EEG-Based Brain–Computer Interface

Piyush Yadav, Mayank Sehgal, Prateek Sharma and Komal Kashish

Abstract The idea of communication through thoughts developed into reality with the growth of computer and communication technologies and the understanding of how a human brain works. One of the main benefactors of this technology and the prime reasons for development of BCI or brain–computer interface is the physically challenged section of the society which will be able to carry on with their regular activities with ease in the near future with the help of ongoing research works in BCI which are primarily focused in the area of neuroprosthetics to restore the damaged hearing, movement and sight in the patients. In this paper, a low-power brain–computer interface design is suggested. The main motive of this paper was to check whether a cost-efficient BCI system can be implemented using basic components and discuss the possible applications where such low-range devices can be of use.

Keywords Brain–computer interface · EEG · Brain signals · Neuroprosthetics Electrodes

1 Introduction

Brain–computer interface, or BCI, a term coined by Professor Jacques Vidal, a computer scientist in 1973 [1], refers to an artificial system which can establish a communication medium between brain and an external device, i.e. a computer. BCI gives the user a medium apart from the normal output mediums of the brain such as muscles. BCI recognizes a specific set of brain signal patterns which follow successive five stages: signal acquisition, pre-processing, feature extraction, classification and control interface [2]. Technologies such as cognitive neuroscience and brain imaging have allowed us to create a direct interface with our brains in the recent times [3, 4]. A prime area of work for BCI was with the electrophysiological

P. Yadav · M. Sehgal · P. Sharma · K. Kashish (✉)
G.L. Bajaj Institute of Technology and Management, Greater Noida, India
e-mail: kashishkomal@ymail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_19

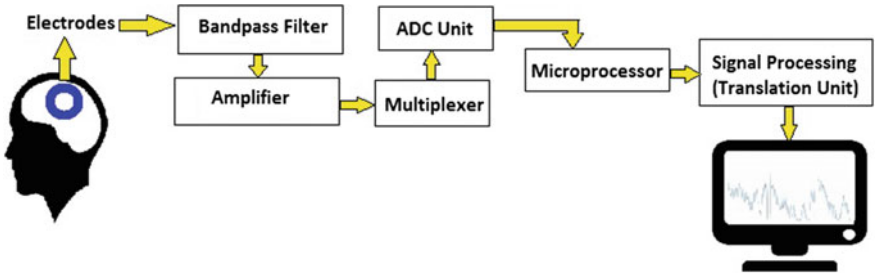

Fig. 1 BCI model

signals that might be found over, underneath and within the brain of patients suffering from neurodegenerative diseases such as amyotrophic lateral sclerosis and stroke patients [5] who have partially or completely lost control of their voluntary muscular activity but still have a functioning cognitive system. At present, over 400 research groups worldwide are working in improving the signal acquisitions, processing and interpreting algorithms [6].

In rudimentary terms, a BCI model works using electrodes which fetch the brain signals. Usually, these acquired signals are of both low amplitude (0.1–0.5 V) as well as low frequency (1–4 Hz). To interpret these signals, amplification and processing units are required next in process to amplify the signals and weed out the unwanted noise. Basically, any BCI system aims to detect the brain signals and then quantify these signals to interpret the intention of the users in real time. Figure 1 is a simple model for a basic BCI system where the signals from a brain are fetched using electrodes, amplified and then processed with the help of a microcontroller, and finally, the signals are quantized and interfaced with any application.

2 Origin and Evolution of BCI

The concept of BCI began with Hans Berger's discovery of electrical signals in human brain followed by the first ever record of electroencephalography (EEG) in 1929 [7]. Berger's first model used invasive technique of signal acquisition by using silver wires that were inserted under the scalp. In the 1950s, Jose Delgado's technique of using electrodes in the brains of animals and humans became famous. A radio receiver, called stimoceiver, was used to stimulate the test subjects. Delgado was successful in producing behaviours such as calming and aggression in the patients. This test and its result could be regarded as one of the earliest proofs that a human brain can be externally controlled [8].

One of the most commercial applications of BCI was the development of bionic ear or cochlear implant in 1976 [9] which was designed on the principle of stimulating the cochlear nerve by sending a signal to a group of electrodes placed in the

cochlea of hearing-impaired patients. Over 400,000 cochlear implants have been surgically implanted all over the world till now.

Another such application in the area of medical devices was the BCI system to aid the blind developed by Dr William Dobelle's Institute in 1978 which was successfully tested on a blind man named, Jerry, who was able to walk and read to some extent with the help of a bulky system consisting of an artificial visual system which collected the surrounding information using a camera and an ultrasonic system. Electrodes placed on the surface of Jerry's visual cortex stimulated him to visualize specks of light, shades of grey and corners of nearby objects [10]. In recent years, one of the most notable contributions of BCI has been enabling a man named, Mathew Nagle, to regain his lost functionality that occurred due to paralysis in 2002.

3 Classification of BCI Systems

The classification of BCI systems is done on the basis of the position of the sensors placed for signal acquisition. Basically, BCI systems are classified into three types: invasive, where the sensors are placed inside the cortex to measure the brain activity, semi-invasive, where the sensors are inside the skull but not in direct contact of grey matter and non-invasive, where the sensors are placed on the scalp of the patient.

3.1 Invasive Technique

Invasive technique devices [11] are implanted directly inside the surface of the brain, targeting the grey matter inside. Invasive BCIs have proved to be effective for helping patients with visual impairment and paralysis. Since these systems interact directly with the grey matter, they are required to be implanted surgically and are prone to be rejected by the body but also give the highest signal quality among all the BCI categories.

3.2 Semi-invasive Technique

Partial or semi-invasive BCI systems are implanted inside the skull but not inside the grey matter of the brain. They have lower chances of getting rejected by the patient's body than invasive techniques but can still produce high-quality signals. Electrocorticography (ECoG) [12] is an example of semi-invasive BCI technique which measures the brain activity with the help of electrodes placed above the

cortex and beneath the dura matter of the test patient's brain, wrapped in a plastic outing.

3.3 Non-invasive Technique

Non-invasive BCI systems are not implanted either inside the skull or grey matter of the patient. Hence, the signal quality suffers in comparison to invasive and semi-invasive techniques. However, these techniques are both faster and safer for the patients involved. Non-invasive devices use medical sensor or imaging units which are capable of reading the brain signals.

Electroencephalogram (EEG) and magnetoencephalography (MEG) signals are often used in most non-invasive techniques. Synchronous activity of cortical neurons inside the brain is responsible for the generation of EEG signals. EEG signals are often used in diagnosing epilepsy [13], sleep disorders, monitoring the comatose state, etc. Magnetoencephalography(MEG) [14] is neuroimaging technique which is used for brain mapping based on the magnetic system produced from the neural activity of the brain. MEG has become an important technique used in presurgical evaluation of patients with epilepsy and brain tumours [15].

4 Methodology and Outline of the System

Figure 2 shows the basic flow of processes in our system which are dependent on each other.

Figure 3 gives an outline of our system, a compilation of brain signal acquisition unit, an amplification unit, signal processing and an output unit.

The first step in our BCI model was acquiring brain signals which was done with the use of electrodes placed on the forehead region of the volunteer applied with a conduction gel [16]. The gel helps in increasing the conduction between the electrode and the forehead. For grounding purpose, another electrode is placed at the foot of the volunteer.

The electrode placed at the forehead area is connected to an amplification and filter board implemented using TL-074, a low noise quad Op-Amp IC with high-performance capacity in the low-frequency spectrum, while the electrode connected to the feet is grounded. The amplification and filter board is a two-stage amplifier circuit as shown in Fig. 4, which filters out the noise, smoothens and

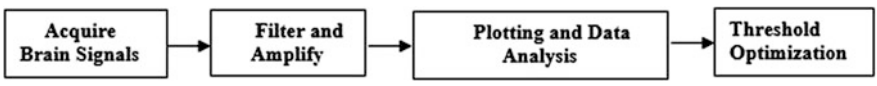

Fig. 2 Flow chart of methodology

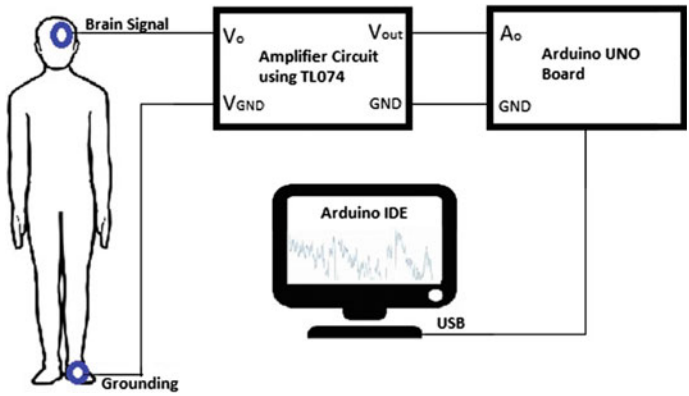

Fig. 3 Block diagram of our BCI system

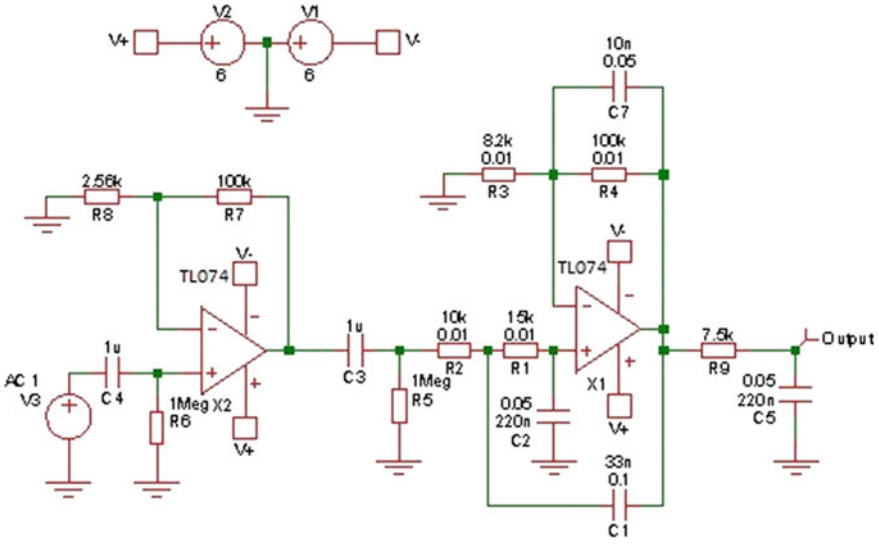

Fig. 4 Amplification circuit

amplifies the incoming raw signal. The gain provided to the signals was up to the order of 10^3 Table 1.

Once testing was done as seen in the values in Table 1, we needed to implement the entire amplification unit on a board which was done by designing a PCB. We chose ExpressPCB, an open-source PCB designing tool, to design the circuit in Fig. 5.

Figure 6 is the actual board with our amplification unit designed and functioning on it.

Table 1 Standardization of amplifier board

Input voltage (V_i) μV	Output voltage (V_o) mV	Gain
5.1	1.73	10^3
5.5	1.79	10^3
4.7	1.69	10^3
6.2	1.83	10^3
5.9	1.80	10^3

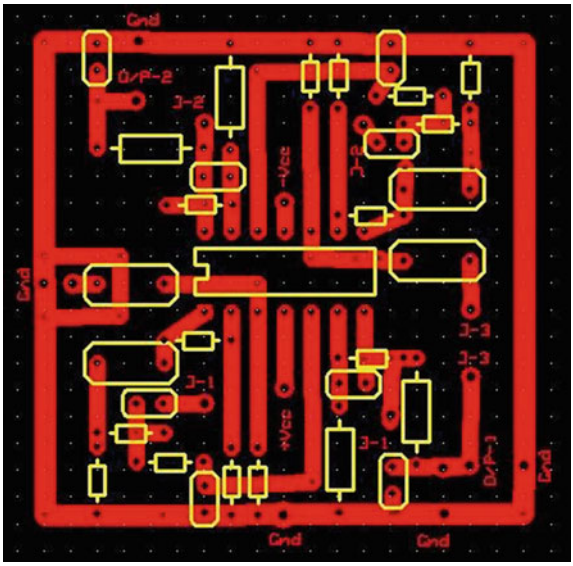

Fig. 5 PCB layout of the amplification board

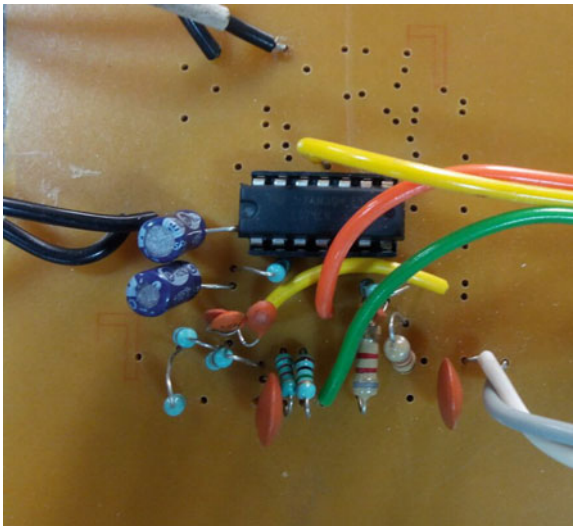

Fig. 6 PCB of amplifier circuit

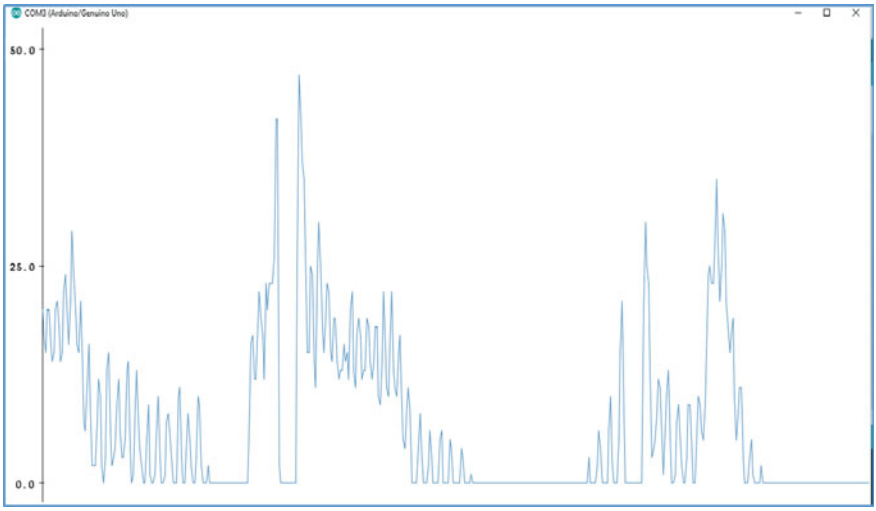

Fig. 7 Brain signal supplied to Arduino UNO

The output of amplification unit is fed to the analog pin of Arduino UNO which operates on Atmega 328p chipset [17, 18]. Arduino UNO has an inbuilt ADC (analog to digital) module which eases the process of analysing the incoming analog value from the amplifier stage.

Figure 7 shows the graph of the analog values supplied to Arduino UNO by the amplifier. The graph is obtained on Arduino IDE and plotted between amplitude on y-axis and time on x-axis. The lower-range peaks correspond to normal tasks such as seeing or hearing and the higher peaks occur when the test subject performs activities such as movement, blinking.

The final function performed by Arduino UNO is threshold optimization which means defining certain amplitude- or intensity-oriented operation on the captured signals. For this, we took an example of the Arduino UNO inbuilt LED blinking corresponding with the blinking of the volunteer. Multiple records of eye-blinking instances of the person were taken, and a conclusion was drawn that in case of an eye blink, the value jumped to 0.15 V. Thus, a threshold to identify an eye blink can be set to 0.15 V.

Using the above method, a threshold was set and the incoming brain signal was optimized accordingly. In Fig. 8, the non-optimized plot of Fig. 7 where the peaks corresponding to either low or high level of activity is optimized in Fig. 8 using threshold method.

Using the steps above, we repeated the same process multiple times to understand the voltage range of other activities such as hand movement, breathing, walking. Once a number of thresholds were decided upon, we asked our volunteer to perform these activities in a random order and checked to see whether our system was able to tell the difference between the activities from the optimized signals and

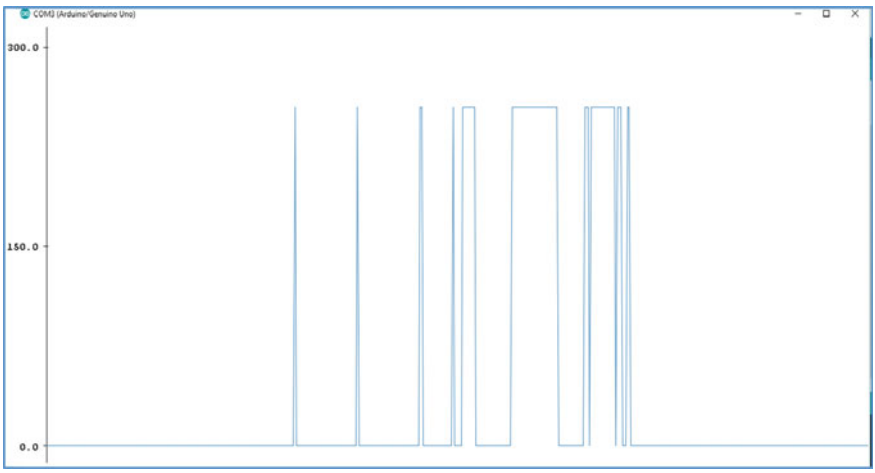

Fig. 8 Optimized graph

our system was successfully able to identify the different movements from each other. Our implementation of low-power and low-cost brain cost interface system was successful to a degree. In future, motors could be used along with the designed BCI system as a representation of limbs where they would function depending upon the user's brain signals whether they want to move that certain limb. This could further be used on physically challenged individuals where the prosthetics could be operated through the collected brain signals.

5 Conclusion

In our work, we were able to successfully fetch, filter, amplify, digitize and optimize the brain signals using a simple, low-cost and low-power system. The operation of the project and plotting of the brain signals was achieved in real time. Two main hurdles of BCI systems that we were able to overcome with our project were reducing the cost of the project and minimizing the noise interference in the obtained signals.

Feasibility of such low-cost devices with considerable accuracy can help in establishing an active BCI developer's community which in turn will promote new researches and developments in real-life applications [19] of BCI which still face a number of challenges. Although the performance of our proposed system is still inferior to the ones being developed on commercial level, with further research and analysis higher accuracy and precision for such low-cost and power efficient devices can be obtained in the future.

References

1. J.J. Vidal, Toward direct brain-computer communication. Ann. Rev. Biophys. Bioeng. **2**(1), 157–180 (1973)
2. M.B. Khalid, N.I. Rao, I. Rizwan-i-Haque, S. Munir, F. Tahir, Towards a brain computer interface using wavelet transform with averaged and time segmented adapted wavelets, in *Proceedings of the 2nd International Conference on Computer, Control and Communication (IC4'09)* (Karachi, Sindh, Pakistan. February 2009), pp. 1–4
3. G. Kucukyildiz et al, Design and implementation of a multi sensor based brain computer interface for a robotic wheelchair. J. Intell. Robot. Syst. 1–17 (2017)
4. S. Rihana, T. Azar, E. Bitar, Portable EEG recording system for BCI application, in *2016 3rd Middle East Conference on Biomedical Engineering (MECBME)*. (IEEE, 2016)
5. A. Remsik et al, A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke. Expert Rev. Med. Devices **13**(5), 445–454 (2016)
6. T.M. Vaughan, J.R. Wolpaw, The third international meeting on brain-computer interface technology: making a difference. IEEE Trans. Neural Syst. Rehabil. Eng. **14**(2), 126–127 (2006)
7. H. Berger, Ueber das Elektrenkephalogramm des Menschen. Arch. Psychiat. **94**, 16–60 (1931)
8. J.M.R. Delgado, W.W. Roberts, N. Miller, ELearning motivated by electrical stimulation of the brain. Am. J. Physiol. **179**, 587–593 (1954)
9. A.A. Eshraghi et al, The cochlear implant: historical aspects and future prospects. Anat. Rec. **295**(11), 1967–1980 (2012)
10. I. Arafat, *Brain-computer interface: past, present & future* (International Islamic University Chittagong (IIUC), Chittagong, Bangladesh, 2013)
11. T.N. Lal et al, Methods towards invasive human brain computer interfaces, in *NIPS* (2004)
12. N.E. Crone, A. Sinai, A. Korzeniewska, High-frequency gamma oscillations and human brain mapping with electrocorticography. Prog. Brain Res. **159**, 275–295 (2006)
13. M.S. Berger et al, Low-grade gliomas associated with intractable epilepsy: seizure outcome utilizing electrocorticography during tumor resection. J. Neurosurg **79**(1), 62–69 (1993)
14. M. Hämäläinen et al, Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. **65**(2), 413 (1993)
15. S.M. Stufflebeam, N. Tanaka, S.P. Ahlfors, Clinical applications of magnetoencephalography. Hum. Brain Mapp. **30**(6), 1813–1823 (2009)
16. C.-T. Lin et al, Novel dry polymer foam electrodes for long-term EEG measurement. IEEE Trans. Biomed. Eng. **58**(5), 1200–1207 (2011)
17. M. Schmidt, in *Arduino*. Pragmatic Bookshelf (2011)
18. Y.A. Badamasi, The working principle of an Arduino. Electronics, Computer and Computation (ICECCO), 2014 11th International Conference on. IEEE (2014)
19. M.M. Moore, Real-world applications for brain-computer interface technology. IEEE Trans. Neural Syst. Rehabil. Eng. **11**, 162–165 (2003)

Denoising of Images Using Neural Network: A Review

Ankita Katiyar and Gauri Katiyar

Abstract Removing noise from images is an integral part of the image processing field. Noises can occur in images during acquisition on transmission. The presence of noise can hinder the proper utilization of these images for various applications such as medical imaging, satellite imaging. Use of soft computing methods, especially neural networks, is studied in this paper. Soft computing techniques are preferred for their ability to deal with inexact data and model accurately. Neural network is an approach that uses layers of neurons to arrive at a result. It is modeled after the human brain cell structure. The works done in the field of image denoising using neural networks over the past decade are presented here.

Keywords Denoise · Neural network · Image processing · Soft computing

1 Introduction

Images almost always contain noises, which are unwanted information in terms of colors, brightness, and pixel intensity, in the way of proper utilization of the image [1]. Image processing requires a number of steps, and denoising image is one of the first tasks to be undertaken. Noise degrades the quality of the image and must be removed to use the image further for various purposes such as object identification, edge detection, feature extraction. The noise that affects image can either originate during image acquisition or it could be the result of a random disturbance in the signal. There are different types of noises that affect images, namely Gaussian, salt and pepper, anisotropic, quantization noise, shot noise. The categorization can also be done on the basis of how the noise corrupts the data. The noise can be additive or

A. Katiyar
CSE Department, VIT, Vellore, Tamil Nadu, India
e-mail: ankita.katiyar2013@vit.ac.in

G. Katiyar (✉)
Department of Electrical Engineering, ITS Engineering College, Greater Noida, UP, India
e-mail: gaurikatiyarrl@its.edu.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_20

multiplicative in nature [2]. Additive noise in the unwanted signal gets added to the image signal, whereas multiplicative noise occurs when the unwanted signal gets multiplied to the original signal. The noisy image can be represented by the following equation:

$$f(x, y) = h(x, y) * \eta_m + \eta_a \tag{1}$$

where *h(x,y)* represents the noise-free image. η_m and η_a are multiplicative and additive noise components, respectively. The filters traditionally used for the removal of additive noises usually blur the details such as edges of the image, whereas filters available for preservation of edges end up amplifying the noises.

Removing of noises from the image is one of the most important steps in the field of image processing. Denoised images are then used for a myriad of purposes such as pattern recognition, satellite imaging, recognition of shapes. The noise model that affects the image decides the approach taken toward the removal of noises. The filtering of images can be divided up into two categories: spatial domain and transform-based filtering. In spatial filtering, the filters are directly applied to the image whereas sometimes the images are transformed into frequency domain and then processed. There can be linear or nonlinear filtering in the spatial domain. Transform domain filtering can be categorized on the choice of basis function which can be data adaptive or nonadaptive.

The flow of the paper is as follows. Section 2 explains the basics of neural network. Section 3 describes the neural network-based denoising techniques, and the paper is concluded in Sect. 4.

2 Basics of Neural Network

Soft computing is a way of taking of computationally hard tasks and arriving at inexact answers for which perfect algorithms do not exist. Soft computing approaches are being widely used and employed into problem solving for their ability to deal with incomplete or approximate data. Artificial neural network (ANN) is a soft computing approach to solve computational problems which is based on how human brain solves tasks [3]. It consists of a number of neurons which are based on the biological neurons. These neurons are interconnected and are arranged in layers. There is one output and an input layer along with intermittent hidden layers. The layers have weighted connections between them which are usually modified using some learning rules. These layers have usually had an associated threshold value which decides whether the neuron fires or not. Neural networks are highly effective in various tasks such as regression analysis, data mining, robotics.

3 Neural Network-Based Denoising Techniques

Ultrasound images have characteristic impulse noises called speckles. It reduces image resolution and contrast. In the work done by Hyunkyung Park and Toshihiro Nishimura [4], they use the cellular neural network to remove the noises for effective usage of ultrasound images. A cellular neural network is a 2D arrangement of cells which are analogical nonlinear processors. In their experiment, a three-layered neural network with infinite impulse response filter is used for speckle reduction and boundary enhancement. Sigmoidal function is used as the activation function in the network. The network is trained using sample images and then tested using ultrasound images. The experiment is said to perform well in terms of speckle reduction and enhancement of boundaries.

Speckle noises also appear in synthetic aperture radar (SAR) images. Pham Duc Long and Pham Thuong Cat [5] proposed a cellular neural network to remove noises in real time. A hybrid parallel/serial computing implementation of speckle reducing anisotropic diffusion (SRAD) on cellular neural network universal machine is used to denoise the images. The computational time is reported to be shorter in comparison to being done on digital computers.

In works performed by S. M. E. Sahraeian, F. Marvasti, and N. Sadati [6], the aim is to remove noises from images based on wavelet transform. This experiment uses a two-step setup, first of which is a new class of nonlinear thresholding functions as activation functions in a new thresholding neural network system (TNN). A cycle-spinning-based technique is then used with the help of the TNN to remove noises from the images. The experiment is reported to provide good results and even perform better than some established noise removal methods in terms of peak signal-to-noise ratio (PSNR) and visual quality.

A method proposed by Sudhansu Kumar Mishra, Ganpati Panda, Sukadev Meher, and Ajit Kumar Sahoo is presented in [7]. The functional link NN (FLNN) is a single-layered network in which the input patterns are enhanced with nonlinear function expansion. The novel method proposed in this method requires much less computational time than multi-layer perceptron (MLP) and is highly effective against Gaussian noise.

Work by Chen Junhong and Zhang Qinyu [8] uses a back-propogation neural network to remove white noise from the images. It also uses a data fusion system based on self-organizing mapping neural networks. This combination of neural networks is said to perform well visually and also eliminate the problem of selecting training data as it uses the same image for training as well as denoising.

A denoising approach proposed by Haijiao Wen and Jie Wen [9] uses pulse-coupled neural networks (PCNNs) which are reported to be effective against salt-and-pepper noise. The experiment first uses a high-resolution image and introduces noise to it. Then PCNN is used to remove noise within five neighborhood and three neighborhood. Then an adaptive median filter is used to reconstruct the modified image. The method is not as effective against Gaussian noise, but with salt-and-pepper noise, it is reported to perform well.

A multilevel discrete wavelet transform (DWT)-based method has been proposed by Torali Saikia and Kandarpa Kumar Sarma [10]. In this method, a combination of DWT and feed-forward neural network (FFNN) is used to remove multiplicative noise from the images. The proposed algorithm is tested with four different images, and the results are evaluated in comparison with a few traditional methods of noise removal such as median filtering, bilateral filtering. The results indicate that based on the visual quality and also the PSNR, the denoised images using this algorithm show less noise and also retain details through the process.

A new approach for denoising images based on neural network is proposed by Satyasis Mishra and Ranjeeta Bisoi [11]. This approach uses least mean square (LMS) to update the weights using accelerated particle swarm optimization (PSO) which is said to be faster in terms of computational time than PSO or genetic algorithm (GA). A PSO is a relatively newer development, and it uses learning parameters to have faster convergence. This approach can provide good results without knowing the mathematical model of the noise, i.e., additive or multiplicative.

In the work done by Emir Turajlic [12], denoising of images of lungs is undertaken to successfully remove noises but also preserve diagnostically important data. The capacity of the neural network framework in removal of the noise from the image is evaluated using a database of computed tomography images of lungs using various performance metrics, such as PSNR and mean squared error. This project tries to evaluate how block size, network architecture, and training methods affect the removal of noise from medical images. Kohonen's self-organizing maps are used for segmentation of feature space, and the use of multiple, finely tuned multi-layer perceptrons is made so as to achieve better results. The experiment is conducted on 20 CT images from the End-Use Load and Consumer Assessment Program (ELCAP) public database which are 512×512 in size. The results obtained in the experiment demonstrate that the scaled conjugate gradient is a more effective choice for the network training rather than the commonly used back-propagation (gradient descent) algorithm.

4 Conclusions

We can conclude from the above-mentioned approaches adopted for denoising of images that usage of neural networks in the field is providing better results than the traditional filtering and statistical methods used before. Integration of soft computing with image processing is preferable due to its ability to model real life better and the ability to deal with approximate data with more flexibility. The works done in the last decade is summarized here, and it can be observed that it is highly important to keep developing new techniques to remove noise from the images. With this review, we hope to provide a better understanding of the work already done and help researchers looking to work in this field.

References

1. V. Rohit, J. Ali, A comparative study of various types of image noise and efficient noise removal techniques. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(10), 617–622 (2013)
2. A. Maity, A. Pattanaik, S. Sagnika, S. Pani, A comparative study on approaches to speckle noise reduction in images, in *Proceedings—1st International Conference on Computational Intelligence and Networks, CINE 2015*, no. 1 (2015) pp. 148–155
3. M.S. Uddin, K.K. Halder, M. Tahtali, A.J. Lambert, M.R. Pickering, Speckle reduction and deblurring of ultrasound images using artificial neural network. Pict. Coding Symp. (PCS) **2015**, 1–4 (2015)
4. H. Park, T. Nishimura, Reduced speckle noise on medical ultrasound images using cellular neural network, in *Conference Proceedings of IEEE Engineering in Medicine and Biology Society*, vol. 2007 (2007), pp. 2138–2141
5. D.L. Pham, T.C. Pham, Image ramp edge enhancing and denoising by complex nonlinear diffusion using Cellular Neural Network, in *7th Asian Control Conference 2009. ASCC 2009* (2009) pp. 1509–1514
6. S.M.E. Sahraeian, F. Marvasti, N. Sadati, Wavelet image denoising based on improved thresholding neural network and cycle spinning, in *2007 IEEE International Conference on Acoustics Speech and Signal Processing—ICASSP'07*, vol. 7 (2007) pp. I-585-I-588
7. S.K. Mishra, G. Panda, S. Meher, A.K. Sahoo, Exponential functional link artificial neural networks for denoising of image corrupted by gaussian noise, in *Proceedings—International Conference on Advanced Computer Control. ICACC 2009* (2009), pp. 355–359
8. J.C.J. Chen, Q.Z.Q. Zhang, Image denoising based on combined neural networks filter, in *2009 International Conference on Information Engineering and Computer Science* (2009) pp. 1–4
9. H. Wen, J. Wen, image denoising and restoration using pulse coupled neural networks, in *6th International Congress on Image and Signal Processing*, no. Cisp (2013), pp. 282–287
10. T. Saikia, K.K. Sarma, Multilevel-DWT based image de-noising using feed forward artificial neural network, in *Advanced Communication and Computing* (2014), pp. 93–102
11. S. Mishra, Image denoising using neural network based accelerated particle swarm optimization, *Power, Communication Information Technology Conference (PCITC), 2015 IEEE* (2015), pp. 808–813
12. E. Turajlic, E. Engineering, S. Bosnia, Application of neural networks to denoising of CT images of lungs, in *2016 XI International Symposium on Telecommunication* (2016), pp. 49–54

Adaptive Volterra Filters for Active Control of Nonlinear Noise Processes

Amrita Rai, Kalpana Hazarika and Monika Jain

Abstract Active noise controls (ANCs) are a challenging application, where there are many considerations such as a secondary path modeling and compensation, fast-changing environment and noise characteristics, interactions between acoustic and electrical domains, placement of microphones and loudspeakers as well as undesired acoustic feedback from the canceling loudspeaker to the reference microphone. ANC algorithms need to identify the secondary path transfer functions between the loudspeaker-to-microphone error sensors and then use the information to guide the direction of control filter coefficient updating. The conventional ANC systems are inefficient in the nonlinear or deterministic noise process, where noise comes from a dynamic system rather than a stochastic, white, or tonal noise process. This paper presents a Volterra filtering for feed-forward ANC which is suitable for the nonlinear controller.

Keywords Active noise controls (ANCs) · Volterra filtering · LMS (least mean square) algorithm · Nonlinear system

1 Introduction

Active noise control (ANC) systems have been developed with an increasing demand for current research and industrial application [1, 2]. Basic techniques used in ANC are based on the superposition principle of the primary noise source and the

A. Rai (✉)
Department of Electronics and Communication Engineering,
GLBITM, Greater Noida, India
e-mail: amrita.tu@gmail.com

K. Hazarika · M. Jain
Department of Electrical and Electronics Engineering,
I.T.S Engineering College, Greater Noida, India
e-mail: hazarika.kalpana@gmail.com

M. Jain
e-mail: monikajain.bits@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_21

secondary noise source. The acoustic outputs of both the sources are of same amplitude and opposite phases so they cancel each other. A single-channel acoustic ANC has been employed for the system to reduce computational complexity and stability problems [1, 3] as illustrated in Fig. 1. An ANC system consists of a reference sensor (reference basic microphone), a secondary source (secondary output loudspeaker), an error sensor (error microphone), and an adaptive controller. The reference signal is used as input to an adaptive filter to generate the anti-noising signal that drives the secondary (canceling) source in the ANC system. The secondary noise produced by the source has equal amplitude but is 180° out of phase with the primary noise and thus is able to cancel it. The error microphone measures the signal which is known as error signal or residual noise, which is used for updating the coefficients of the adaptive filter in the ANC system. From the literature survey, we can conclude that the linear active noise controls system could suffer control performance degradation and may encounter the following problems in some situations. Firstly, the noise process is more chaotic rather than a stochastic or tonal noise. Secondly, when primary path may exhibit nonlinearl distortion then the secondary path designate transfer function between the secondary source (secondary speaker) and an error microphone to have non-minimum phase difference [4].

Therefore, many researchers investigated and applied the nonlinear ANC system using various adaptive algorithms like filtered-X least mean square (LMS), recursive least squares (RLS), filtered-s LMS, and neural networks for tackling various nonlinear effects [1, 3]. In adaptive second-order Volterra filters [5], ANC is being addressed with an assumption that the primary path exhibits the nonlinear behavior and also indicates potential application in active noise and vibration control through the simulation results. The result shows that proposed algorithm can improve control performance over the conventional linear LMS algorithm. The following two situations where the improvements can be observed are the signal generated by a reference microphone as a nonlinear and predictable noise, whereas the secondary path transfer functions of an ANC system has non-minimum phase and nonlinear behavior exhibited by the primary path. In this paper, an adaptive Volterra filter with a dynamic variable step-size least mean square (DVSS-LMS) algorithm, yielding lower misalignment performance than conventional linear approaches, has

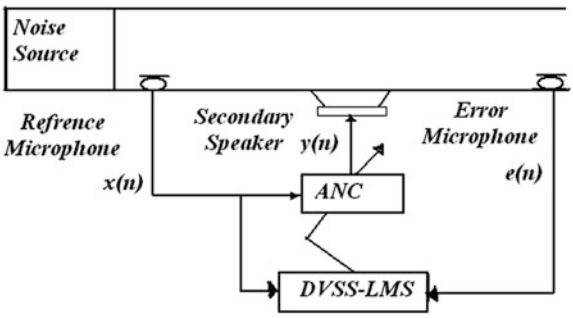

Fig. 1 Schematic diagram of single-channel active noise controls

been proposed. Section 2 explains the Volterra filter for active noise controls. In Sect. 3, an analysis of nonlinear ANC has been described using Volterra filter with non-minimum secondary path transfer function phase. Finally, concluding remarks are presented in Sect. 4.

2 A Volterra Nonlinear ANC System

In engineering applications, nonlinear controllers for ANC using neural networks were anticipated, where the actuators are used in the system to demonstrate nonlinear characters [6]. Alternatively, training algorithm can be used, which is based on an extended back-propagation proposal fabricated for multilayer perceptron neural networks [4]. The following adaptive nonlinear filters are used for above applications: (1) order statistic filters, (2) holomorphic filters, (3) morphological filters, and (4) polynomial filters. A nonlinear adaptive Volterra filter with the corresponding various LMS algorithms was independently developed for solving saturation in ANC systems. The key problem associated with a nonlinear Volterra filter is that its computational complexity increases exponentially as the order of the nonlinearity increases. To reduce the computational burden, recent research work [7] has explored partial update strategies for the third- and second-order Volterra filter X least mean square (VFXLMS) algorithm for the case of multichannel noise control.

Figure 2 demonstrates an adaptive filter with the basic single-channel broadband ANC system to estimate an unknown plant. The primary path transfer function $P(z)$ comprises of the acoustic response from the reference sensor to the error microphone, where the noise attenuation has to be realized for the canceled primary noise signal $d(n)$, and $S(z)$ is the secondary transfer function between adaptive filter and error microphone used to minimize residual error signal $e(n)$. The adaptive algorithm will be continuously tracking and estimating secondary path transfer function $\overline{S}(z)$ using VFXLMS algorithm [1]. The most important difference between Fig. 2 and the conventional system identification scheme is the use of an acoustic summing junction instead of the subtraction of electrical signals. However, for

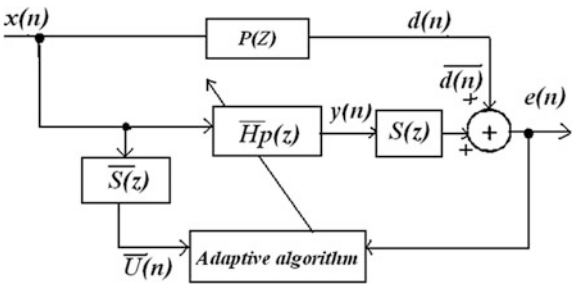

Fig. 2 Block diagram of ANC system using adaptive algorithm

inequality, it will represent the summing junction by a subtraction; it is arbitrary because it can be implemented with a sign change of the secondary signal [8].

Since the output of the Volterra filter is linear with respect to the Volterra kernels, the existing linear theory can be directly applied to Volterra filtering for nonlinear ANC systems [3, 7]. As illustrated in Fig. 2, the secondary signal $y(n)$ is generated by filtering the reference signal $x(n)$ by the Volterra filter $h_p(n)$ expressed as:

$$y(n) = h_p^T(n)x(n) + e(n) \tag{1}$$

where n is the time index, $e(n)$ is the residual noise measured by the error microphone, and $\boldsymbol{T}$ denotes the transpose of the coefficient vector.

$$h_p(n) = \begin{bmatrix} h_0, h_1(0;n), h_1(1;n), \ldots h_1(M-1;n), h_2(0,0;n), h_2(0,1;n), \ldots h_2(0,M-1;n) \\ , h_2(1,1;n), \ldots h_2(M-1,M-1;n), h_3(0,0,0;n), \ldots h_3(M-1,M-1,M-1;n), \end{bmatrix}^T \tag{2}$$

$\boldsymbol{M}$ is length of the filter.

$$x(n) = \left[x_1^T(n)x_2^T(n)\ldots x_3^T(M-1;n)\right]^T \tag{3}$$

$x(n)$ is the reference input signal received by the reference microphone. The reference signal $x(n)$ passes through the Volterra filter $h_p(n)$ and secondary path $S(z)$. The adaptive Volterra filter minimizes the difference between the desired signal $d(n)$ and the estimated signal $\overline{d}(n)$; thus, we can obtain a residual error signal $e(n)$ received by the error microphone as shown in Fig. 1. Also, the Volterra filter coefficient can be utilized to update filter input signal $\overline{U}(n)$ (estimated by secondary

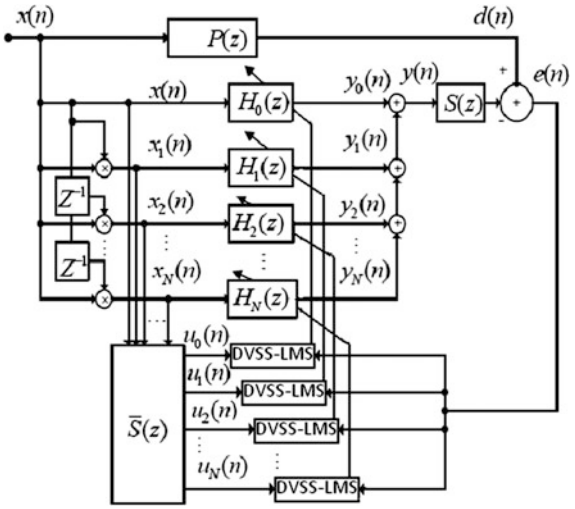

Fig. 3 Third- and second-order Volterra filtered DVSS-LMS algorithm for a feed-forward ANC system

path $S(z)$) and error signal $e(n)$. The adaptive filter minimizes the instantaneous squared error $e(n)^2$, using the proposed algorithm as expressed in the preceding section (Fig. 3).

3 Secondary Propagation Path Transfer Function with Non-Minimum Phase

The secondary propagation path is the path of anti-noise which takes from the output loudspeaker to the error microphone within the quiet zone. Estimation of the secondary path is performed offline prior to the implementation of ANC algorithm, but in practical cases the secondary path is usually time-varying in nature. To examine the feasibility of the developed DVSS-LMS algorithm used in a situation where the secondary path transfer function has a non-minimum phase, while the reference noise is a nonlinear noise process as initiate with the Z-domain derivation. The objective is to illustrate that if a nonlinear reference noise process is to be controlled with a non-minimum-phase secondary path estimate, the DVSS-LMS algorithm will achieve better performance than using the conventional LMS algorithm. To illustrate above error function $e(n)$ (from Fig. 2) in the Z domain as

$$\begin{aligned} E(z) &= P(z)X(z) + \sum_{p=1}^{P} Y_p(z)S(z) \\ &= S(z)\left[P(z)S^{-1}(z)X(z) + \sum_{p=1}^{P} Y_p(z)\right] \end{aligned} \tag{4}$$

The implementation of Eq. (4) can be modeled and derived in the block diagram Fig. 4. When $S(z)$ is non-minimum phase, the $S^{-1}(z)$ may be approximated as $\overline{S}(z)z^q$ [3], where $\overline{S}(z)$ is a causal approximation to the inverse of secondary path $S(z)$ and q is a positive integer delay value. Using the approximate value of $S^{-1}(z)$ in (4),

$$E(z) = S(z)\left[\overline{S}(z)z^q P(z)X(z) + Y(z)\right] \tag{5}$$

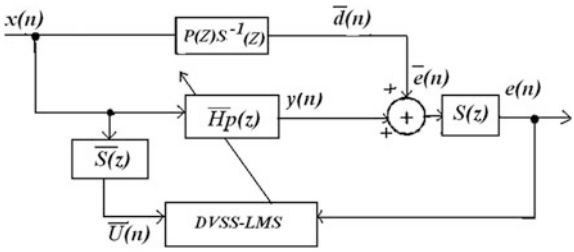

Fig. 4 Modified block diagram of ANC with DVSS-LMS (derived by Fig. 2)

We assume that $\overline{S}(z)P(z) \approx \sum_{i=0}^{N-1} h(i)z^{-i}$, where $h(i)$ is set of filter coefficient. From Fig. 4, substituting $\overline{D}(z) = S^{-1}(z)P(z)X(z)$ in (4) the residual noise in time domain may be expressed as

$$\begin{aligned} e(n) &= s(n) * \left(\overline{d}(n) + y(n)\right) = s(n) * \overline{e}(n) \\ &= s(n) * \left[\sum_{i=0}^{N-1} h(i)x(n-i+q) + y(n)\right] \\ &= s(n) * \left[\sum_{i=0}^{q-1} h(i)x(n-i+q) + \sum_{i=q}^{N-1} h(i)x(n-i+q) + \sum_{p=1}^{P} y_p(n)\right] \end{aligned} \tag{6}$$

where the time-domain representation of $\overline{D}(z)$ is $\overline{d}(n)$. Analysis of (6) depends on the value of q, and the results include the following three cases.

Case I. If the secondary path estimates $\overline{S}(z)$ have minimum phase, that is $q = 0$, then (6) is reduced to

$$e(n) = s(n) * \left[\sum_{i=0}^{q-1} h(i)x(n-i+q) + \sum_{p=1}^{P} y_p(n)\right] \tag{7}$$

It has been seen that the first term of (7) is a linear combination of present and past samples of the reference input. Therefore, minimizing of residual error power $E\left\{|e(n)|^2\right\}$ is similar to linear system identification for linear primary path. On other situation, where the primary path which exhibits nonlinear behavior $q = 0$ indicates the secondary path estimate with minimum phase, then the second term of (7) is zero. It shows that minimization of residual error power $E\left\{|e(n)|^2\right\}$ is identification of nonlinear system. Thus, Volterra filter with DVSS-LMS can improve the control performance.

Case II. If the transfer function of secondary path has in non-minimum phase, the second term of (7) yields a prediction term. However, if the input signal is a linear noisy process, then the higher-order Volterra filter would not offer any advantage over others.

Case III. If $q > 0$, which indicates the secondary path with non-minimum-phase and the signal $x(n-i+q)$ of (7) is a nonlinear process, the conventional LMS algorithm is neither suitable for nonlinear input nor it is efficient predictor. On the other hand, if the input signal is a nonlinear noise process, the proposed DVSS-LMS Volterra controller offers better performance irrespective of whether the reference noise is predictable or unpredictable.

4 Conclusion

The proposed paper has a study of second- and third-order Volterra filtered LMS algorithm-based nonlinear active noise controller structure. The proposed method examines nonlinear ANC systems for a situation, where a nonlinear reference noise processes to be mitigated with a minimum and non-minimum-phase secondary path. The LMS algorithm is derived using various variable step-size LMS basic structure to perform faster and better than already existed algorithms for nonlinear ANC. Its performance in terms of normalized mean square error (NMSE) has been compared with the already existing XLMS algorithm. Thus, the Volterra filter with LMS-based nonlinear ANC is better option for active noise cancelation than other algorithms, both in terms of steady-state mean square error and computational complexity.

References

1. V.J. Mathews, Adaptive polynomial filters. IEEE Signal Process. Mag. **8**(3), 10–26 (1991)
2. S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1986)
3. S.C. Douglas, Fast implementation of the Filtered-X LMS and LMS algorithms for multichannel Active Noise Controls. IEEE Trans. Speech Audio Process. **7**, 454–465 (1999)
4. A. Carini, G.L. Sicuranza, Filtered-x affine projection algorithms for Active Noise Controls using Volterra filters. *EURASIP* J. Appl. Signal Process, 1841–1848, (2004 Dec)
5. A. Rai, A.K. Kohli, Adaptive polynomial filtering using generalized variable step-size least mean pth power (LMP) algorithm. Circuit Syst. Signal Process. (2014 June)
6. G.L. Sicuraranza, A.C. Carini, Filtered-X affine projection algorithm for multichannel Active Noise Controls using second-order Volterra filters. IEEE Signal Process. Lett. **11**(11), 853–857 (2004)
7. L. Tan, J. Jing, Adaptive Volterra filters for active control of nonlinear noise processes. IEEE Trans. Signal Process. **49**(8), 1667–1676 (2001)
8. P. Strauch, B. Mulgrew, Active control of non linear noise processes in a linear duct. IEEE Trans. Signal Process. **46**(9), 2404–2412 (1998)

Part V
Fluid Dynamics

Estimation of Dispersion in Mildly Curved Channel Flow with Absorbing Boundaries Using Multiquadric Radial Basis Function (MQRBF) Method

Bhanumati Panda and Sushil Kumar

Abstract Numerical dispersion is estimated using multiquadric radial basis function (MQRBF) method in an incompressible viscous flow in a mildly curved channel with constant curvature where the solute undergoes irreversible heterogeneous reactions with the absorbing boundaries. The MQRBF scheme is used to solve the convection–diffusion equation to study the dispersion phenomena in a mildly curved channel with constant curvature. The influence of the combined effect of the Peclet number, absorption parameter and curvature ratios on longitudinal dispersion is discussed. It is also observed that the concentration distribution is more skewed with increasing curvature. The shape parameter is tuned with increase of the curvature, Peclet number and absorption parameter, and MQRBF scheme performs well for low Peclet numbers.

Keywords MQRBF · Peclet number · Absorbing boundaries · Shape parameter Dispersion

1 Introduction

The first fundamental study of dispersion was initiated by Taylor [1] who shows that the transport of a solute is due to the interaction between advection and molecular diffusion in a viscous liquid flowing through a straight tube. He also obtained the effective diffusivity of dispersion coefficient, i.e. $D_{\text{eff}} = \frac{a^2U^2}{48D}$, where D represents the molecular diffusion coefficient, U is the average (mean) velocity, and

B. Panda (✉)
Department of Applied Science & Humanities, I.T.S Engineering College, Greater Noida 201308, UP, India
e-mail: jhununeel@gmail.com

S. Kumar
Department of Applied Mathematics, School of Vocational Studies and Applied Sciences, Gautam Buddha University, Greater Noida 201310, UP, India
e-mail: sushil12@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_22

a is the radius of the tube. Aris [2] has presented the method of moment analysis for the solute in a streamwise direction and studied that the rate of increase of the variance of the solute distribution is proportional to the sum of molecular diffusion coefficient D and Taylor's dispersion coefficient D_{eff}. He also obtained the approximate solution $D_{\text{eff}} = D + \frac{a^2U^2}{48D}$. Barton [3] resolved some technical difficulties of Aris method of moments and found the solutions for second and third moments of the solute dispersion. Many researchers have extensively studied the dispersion phenomena [4–11].

Using Taylor's dispersion phenomena, Gupta and Gupta [12] studied the dispersion of contaminant in a heterogeneous reaction in a channel flow for a very large time. [10, 13–15] has studied the boundary reaction in a shear flow in channel or tube. Kumar and Jayaraman [16], using method of moments, showed the effect of curvature of dispersion of solute in a curved channel.

The main objective of this paper is to estimate the numerical dispersion of mildly curved channel with absorbing boundaries using MQRBF method, where the effect of curvature is considered only through velocity profile. The scheme used to study the numerical estimation of dispersion is multiquadric radial basis function (MQRBF) method originally proposed by Ronald Hardy in [17] which is a mesh-free method. Franke [18] compared various approximation methods and concluded that the Hardy's multiquadric method is best. Kansa [19] used the method to solve the differential equation. Sarra [20] solved the advection–diffusion reaction equation using local MQRBF method for complexly shaped domain. Panda et al. [21, 24] used the multiquadric and inverse MQRBF method to analyse the numerical dispersion of convection–diffusion equation. The mathematical formulation of the governing equation along with the velocity profile is discussed in Sect. 2. The numerical procedure and the transformation used to convert the infinite physical domain to finite computational domain are discussed in Sect. 3. Methodology used for the present study is given in Sect. 4. Sections 5 and 6 discuss the result and discussion for the absorbing boundaries followed by salient conclusion.

2 Mathematical Formulation

Steady plane laminar flow of a viscous incompressible fluid with slow variation is considered in curved channel with constant width '2a' with the velocity $u(\eta)$, and the detailed discussion is given in Goldstein [22], Nunge et al. [23] and Kumar and Jayaraman [16]. The non-dimensional form of the velocity profile is given as:

$$\frac{u(\eta)}{U_{\text{A}}} = 2(1-k)\frac{\left\{\left[\frac{(k+1)^2}{4k}\ln\left(\frac{k+1}{k-1}\right)\right]\left[\frac{k\eta+1}{1-k}-\frac{1-k}{1+k\eta}\right]-\left(\frac{1+k\eta}{1-k}\right)\ln\left(\frac{1+k\eta}{1-k}\right)\right\}}{\left\{1-\frac{(k+1)^2(1-k)^2}{4k^2}\left[\ln\left(\frac{1+k}{1-k}\right)\right]^2\right\}} \quad (1)$$

where $U_A = \frac{-1}{2k^2}\frac{\partial p}{\partial s}\left\{-\frac{(1-k^2)^2}{8k^2}\left[\ln\left(\frac{1+k}{1-k}\right)\right]^2 + \frac{1}{2}\right\}$ represents the non-dimensional average (mean) axial velocity in a curved channel with curvature k and $\frac{\partial p}{\partial s}$ represents the pressure gradient in axial direction.

The non-dimensional form of convection–diffusion equation for solute concentration $C(s, \eta, t)$ is given as follows:

$$\frac{\partial C}{\partial t} + Pe\left(\frac{u}{(1+k\eta)}\frac{\partial C}{\partial s}\right) = \frac{1}{(1+k\eta)}\left(\frac{\partial}{\partial s}\left(\frac{1}{1+k\eta}\frac{\partial C}{\partial s}\right) + \frac{\partial}{\partial \eta}\left((1+k\eta)\frac{\partial C}{\partial \eta}\right)\right) \tag{2}$$

where Pe represents the Peclet number. For mildly curved channel, the steady-state concentration equation can be written as:

$$Pe\left(u\frac{\partial C}{\partial s}\right) = \frac{\partial^2 C}{\partial s^2} + \frac{\partial^2 C}{\partial \eta^2} \tag{3}$$

with the following initial and boundary conditions:

$$C(0, \eta) = f(\eta) \tag{4}$$

$$\frac{\partial C}{\partial \eta} \pm \beta C = 0, \quad \text{at } \eta = \pm 1 \tag{5}$$

$$C(\pm\infty, \eta) = 0 \tag{6}$$

where β is the absorption parameter.

3 Transformation

The transformation [10] is used to convert from infinite domain to finite computational domain which is given as:

$$s = \frac{1}{2a}\log\frac{1+\phi}{1-\phi}, \quad \eta = \Omega, \ -1 < \phi < 1, \ 0 \le \Omega \le 1 \tag{7}$$

where a is the stretching factor.

The transformed concentration Eq. (3) can be written as

$$Pe(u\frac{\partial C}{\partial \phi})a(1-\phi^2) = \frac{\partial^2 C}{\partial \Omega^2} + a^2(1-\phi^2)^2\frac{\partial^2 C}{\partial \phi^2} - 2a^2(1-\phi^2)\phi\frac{\partial C}{\partial \phi} \tag{8}$$

$$C(0, \Omega) = g(\Omega) \tag{9}$$

$$\frac{\partial C}{\partial \Omega} \pm \beta C = 0, \quad \text{at} \quad \Omega = \pm 1 \tag{10}$$

$$C(\pm 1, \Omega) = 0 \tag{11}$$

4 Methodology

The multiquadric radial basis function (MQRBF) method is used to solve Eqs. (8–11).

Consider a general differential equation

$$Aw = f(s, \eta) \quad \text{in} \quad \Pi \tag{12}$$

$$Bw = g(s, \eta) \quad \text{in} \quad \partial \Pi \tag{13}$$

where A and B are the differential and boundary operators imposed on governing equation and boundary conditions. Here, domain Π has been discretized in N collocation points. Let $\{P_i = (s_i, \eta_i)\}_{i=1}^{N}$ be N collocation points in domain Π of which $\{(s_i, \eta_i)\}_{i=1}^{N_I}$ are interior points; $\{(s_i, \eta_i)\}_{i=N_{I+1}}^{N}$ are boundary points in domain $\partial \Pi$.

Let the concentration C can be approximated as:

$$C(s, \eta) = \sum_{j=1}^{N} w_j \psi_j(s, \eta) \tag{14}$$

where $\{w_j\}_{j=1}^{N}$ are the unknown coefficients to be determined, and $\psi_j(s_j, \eta_j)$ is a basis function which is taken here as $\psi_j = \sqrt{(s - s_j)^2 + (\eta - \eta_j)^2 + \varepsilon^2} = \sqrt{r_j^2 + \varepsilon^2}$, where ε is the shape parameter.

The detailed methodology is discussed in Panda et al. [21, 24]

5 Result and Discussions

The MQRBF is applied to estimate the longitudinal dispersion of solute in a mildly channel flow width '2a' and infinite length for different curvature ratios and absorption parameters where the effect of curvature is shown through velocity profile only. The effect of absorption parameter (β) and curvature (k) is shown for a specific Peclet number (Pe) at different cross sections of the channel. Figures 1 and 2 show the effect of β for $k = 0.5$ and $Pe = 1$. It is observed that the solute

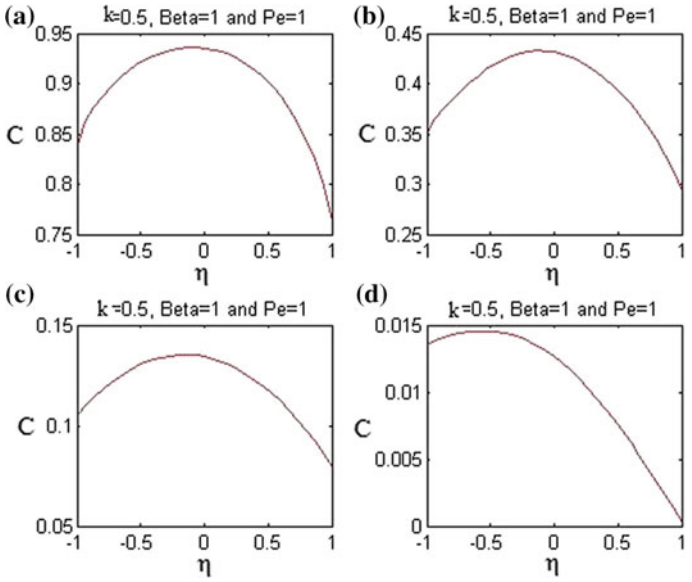

Fig. 1 Concentration C at different cross section **a** $s = 0.03$, **b** $s = 0.3$, **c** $s = 0.63$, **d** $s = 0.9$, for curvature $k = 0.5$, $\beta = 1$ and $Pe = 1$

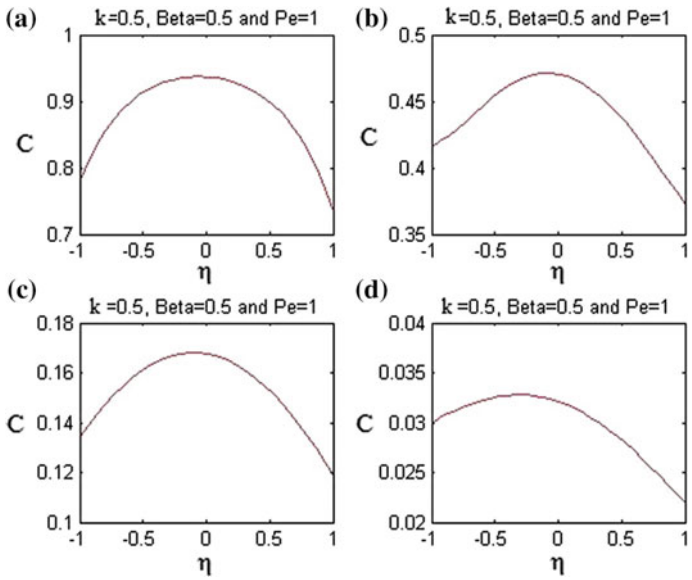

Fig. 2 Concentration C at different cross section **a** $s = 0.03$, **b** $s = 0.3$, **c** $s = 0.63$, **d** $s = 0.9$, for curvature $k = 0.5$, $\beta = 0.5$ and $Pe = 1$

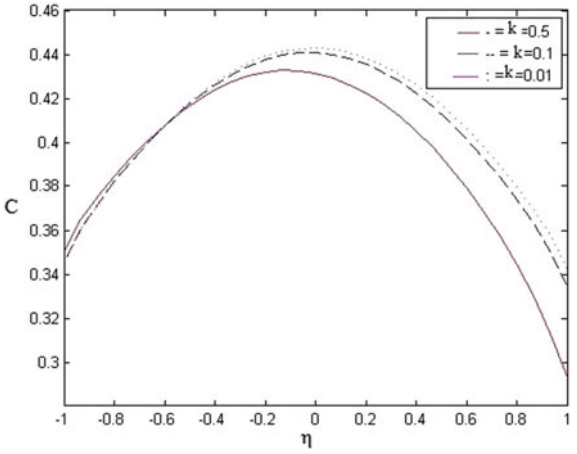

Fig. 3 Concentration C at cross section $s = 0.3$ for different curvatures

concentration decreases as β increases and also observed that solute concentration is decreased as it is away from the source. The concentration profile is more skewed as β increased and it is with conformity with Kumar and Jayaraman [16]. Figure 3 shows the concentration profile of the solute for different curvature ratio at a specific absorption parameter and Peclet number. It is observed that the skewness increases with the increase of curvature ratio. The present scheme performs well for small Peclet number. As the Peclet number increases, the methods need adaptation for desired result. Since shape parameter plays crucial role in MQRBF method, it is tuned for better accuracy.

6 Conclusions

The global MQRBF method reproduces well for small Peclet number up to a specific value for slow variation curvature. The absorption parameter and curvature both affect the solute distribution in channel flows, and the concentration is more skewed towards the inner wall. The concentration distribution is decreased with increased absorption parameter, and concentration profile is more skewed with increasing curvature.

References

1. Taylor, G., Dispersion of soluble matter in solvent flowing slowly through a tube, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society, vol. 219, pp. 186–203 (1953)

2. Aris, R., On the dispersion of a solute in a fluid flowing through a tube, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society, vol. 235, pp. 67–77 (1956)
3. N.G. Barton, On the method of moments for solute dispersion. J. Fluid Mech. **126**, 205–218 (1983)
4. M.J. Lighthill, Initial development of diffusion in poiseuille flow. IMA J. Appl. Math. **2**(1), 97–108 (1966)
5. M.E. Erdogan, P.C. Chatwin, The effects of curvature and buoyancy on the laminar dispersion of solute in a horizontal tube. J. Fluid Mech. **29**(03), 465–484 (1967)
6. W.N. Gill, R. Sankarasubhramanian, Dispersion of non-uniformly distributed time-variable continuous sources in time-dependent flow. Proc. R. Soc. Lond. A **327**, 191–208 (1972)
7. R. Smith, Longitudinal dispersion coefficients for varying channels. J. Fluid Mech. **130**, 299–314 (1983)
8. A. Mukherjee, B.S. Mazumder, Dispersion of contaminant in oscillatory flows. Acta Mech. **74**, 107–122 (1988)
9. B.S. Mazumder, R. Xia, Dispersion of pollutants in an asymmetric flow through a channel. Int. J. Eng. Sci. **32**(9), 1501–1510 (1994)
10. K.K. Mondal, B.S. Mazumder, On dispersion of settling particles from an elevated source in an open-channel flow. J. Comput. Appl. Math. **193**(1), 22–37 (2006)
11. S. Kumar, G. Jayaraman, Method of moments for laminar dispersion in an oscillatory flow through curved channels with absorbing walls. Heat Mass Transf. **44**, 1323–1336 (2008)
12. P.S. Gupta, A.S. Gupta, Effects of homogeneous and heterogeneous reactions on the dispersion of a solute in the laminar flow between two plates. Proc. R. Soc. Lond. A **330**, 59 (1972)
13. A. Purnama, Boundary retention effects upon contaminant dispersion in parallel flows. J. Fluid Mech. **195**, 393–412 (1988)
14. B.S. Mazumder, S.K. Das, Effect of boundary reaction on solute dispersion in pulsatile flow through a tube. J. Fluid Mech. **239**, 523–549 (1992)
15. G. Jayaraman, T.J. Pedley, A. Goyal, Dispersion of solute in a fluid flowing through a curved tube with absorbing walls. Q. J. Mech. Appl. Math. **51**(4), 577–598 (1998)
16. S. Kumar, G. Jayaraman, Method of moments for estimating two dimensional laminar dispersion in curved channels. Indian J. Ind. Appl. Math. **3**(1), 116–133 (2012)
17. R.L. Hardy, Multiquadric equations of topography and other irregular surfaces. J. Geophys. Res. **76**(8), 1905–1915 (1971)
18. R. Franke, Scattered data interpolation: Tests of some methods. Math. Comput. **38**(157), 181–200 (1982)
19. E.J. Kansa, Multiquadrics a scattered data approximation scheme with applications to computational fluid-dynamics i. surface approximations and partial derivative estimates. Comput. Math Appl. **19**(8), 127–145 (1990)
20. S. Sarra, A Local radial basis function method for advection-diffusion-reaction equations on complexly shaped domains. Appl. Math. Comput. **218**(19), 9853–9864 (2012)
21. B. Panda, S. Kumar, R.K. Mishra, Solving singularly perturbed problems using multiquadric/ Inverse multi-quadric radial basis function method. Indian J. Ind. Appl. Math. **7**(1), 43–57 (2016)
22. Goldstein, S., *Modern Developments in Fluid Dynamics*, vol. 1, (Dover, New-York, 1965), p. 315
23. R.J. Nunge, T.S. Lin, W.N. Gill, Laminar dispersion in curved tubes and channels. J. Fluid Mech. **51**(02), 363–383 (1972)
24. B. Panda, S. Kumar, R.K. Mishra, Numerical dispersion in small varying meandering channel using multiquadric radial basis function method. Indian J. Ind. Appl. Math. **7**(2), 262–269 (2016)

Part VI
Process Control

Model Based Predictive Control of the Four Tank System

Sankata B. Prusty, Umesh C. Pati and Kamala K. Mahapatra

Abstract This paper deals with the implementation of Model Predictive Control (MPC) in a Four Tank System (FTS). The nonlinear model of FTS has been developed from the mechanism modelling. There is coupling between the input and output and time delay in the FTS. Different control algorithms are implemented to the FTS. The objective is to track the level of liquid in tanks at the reference values. This problem is solved using different control methods such as proportional-integral-derivative (PID), MPC, and Fuzzy Modified Model Reference Adaptive Control (FMMRAC). The MPC allows closed-loop solution to the optimization problem to be obtained off-line. A general MPC control is applied to the FTS and different performance indices as well as error indices are calculated. The responses of these controllers are corroborated and the comparison of MPC with other control algorithms are presented. The MPC provides better performance than the other control algorithms. The simulation results show that good tracking performance is attained.

Keywords Proportional-integral-derivative controller · Model Predictive Control Four tank system · FMMRAC

S. B. Prusty (✉) · U. C. Pati · K. K. Mahapatra
Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela 769008, India
e-mail: sankata_prusty@yahoo.com

U. C. Pati
e-mail: ucpati@nitrkl.ac.in

K. K. Mahapatra
e-mail: kkm@nitrkl.ac.in

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_23

1 Introduction

The four tank system (FTS) is used in the laboratory to control the liquid level in tanks. There is coupling between the input and output and time delay in the FTS. Different control algorithms are implemented to the FTS. Most of the process industries have the problem of controlling the liquid level in tanks, temperature and flow. Most of the times, the liquids are used in chemical plants as well as for mixing treatment [2, 12] in the tanks. But, the liquid level must be maintained as well as the flow into tanks must be adjusted in process industries.

Several literatures are available to control the liquid level in coupled tanks. The dynamic mathematical model of the quadruple tank system has been developed by Johanson [3] for the experiment in laboratories. Both minimum and non-minimum phase systems for tank system are discussed in this work. Fuzzy logic (FL) control [1, 9, 10], neural network (NN) control [4, 14], and genetic algorithms (GA) [6] have also been employed to the tank system. The transient response of the controller are enhanced by using adaptive control methods [6]. Mohideen et al. have proposed Modified MRAC scheme for controlling the liquid level in FTS with good transient performance as well as steady-state performance. A predictive control method employed to a coupled tank system has discussed in [8]. The comparative analysis of conventional and intelligent control utilized in the process has been described by Zumberge and Passino [15].

The controller is used to adjust the variations in plant dynamics due to its capability. An adaptive method for the enhancement of the performance in case of a motor drive system has been proposed by Liu and Hsu [5]. The level of liquid in tanks can also be controlled by NN, FL, GA and their hybrid. Also, the different algorithms such as Ant Colony optimization, Particle Swarm Optimization, and Bacterial Foraging methods have been employed to control the various parameters of the process.

Predictive control methods are used to solve the control issues of the process. The Model Predictive Control (MPC) is an advanced optimization based control method for process industries since 1980 [11]. Power system balancing models are also controlled by using MPC. One of the most important advantage of MPC is that the recent timeslot are to be optimized by considering the future time slots. MPC has the capability of future event predictions and can utilize the control action accordingly. A typical application in the process industry is the tank level control system. System status and system parameters are the main causes of accuracy and precision of the liquid level system. Centralized and decoupling of quadruple tank system using MPC method has discussed by Srinivasarao and Subbaiah [13]. The generalized predictive control (GPC) which reduces the computational time has been described by Muthukumar et al. [7]. In this paper, MPC controller is implemented to control the liquid level in lower two tanks of FTS.

This paper is arranged as follows: The dynamic mathematical model of the FTS is described in Sect. 2. In Sect. 3, MPC controller algorithm is discussed. The simulation results and the descriptions about the responses presented in Sect. 4 for different controllers. Finally, the conclusion is provided in Sect. 5.

2 Dynamic Model of FTS

The schematic representation of the FTS [3] is shown in Fig. 1. The FTS has the property of multivariable interaction due to the influences of the outputs by the pumps. The valves are adjusted in such a way that the system has occurred a multivariable zero. The multivariable zero can be present at the left half plane or right half plane. Bernoulli's law and Mass balance equation are used to develop the dynamics of FTS and the system model is established according to the nonlinear mechanism. The goal of the process is to control the level of liquid in the lower two tanks at the required level. The FTS has two inputs and two outputs. Input voltages to the pump (v_1 and v_2) are considered as the input to the process and voltages from the level sensor (y_1 and y_2) are considered as the outputs. The water coming from tank 1 and tank 2 is stored in a reservoir tank present below the tanks. The liquid level in tank i is represented as h_i where $i = 1, \ldots, 4$.

The density of liquid is assumed to be constant in inlet, in outlet and in the tank to develop the mathematical model. Mass balance and Bernoulli's law are used to derive the nonlinear model of FTS are presented in Eqs. (1–4) as

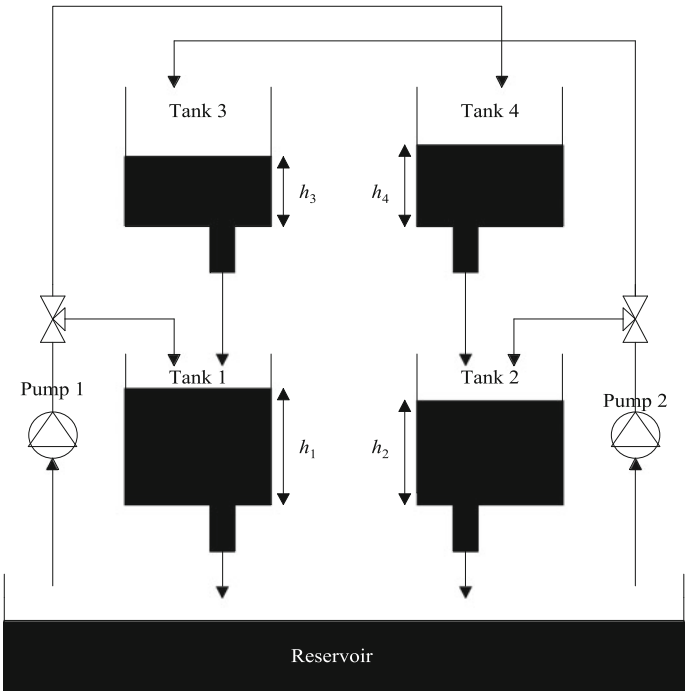

Fig. 1 Schematic representation of FTS

$$\frac{dh_1}{dt} = -\frac{a_1}{A_1}\sqrt{2gh_1} + \frac{a_3}{A_1}\sqrt{2gh_3} + \frac{\gamma_a k_1}{A_1}v_1 \tag{1}$$

$$\frac{dh_2}{dt} = -\frac{a_2}{A_2}\sqrt{2gh_2} + \frac{a_4}{A_2}\sqrt{2gh_4} + \frac{\gamma_b k_2}{A_2}v_2 \tag{2}$$

$$\frac{dh_3}{dt} = -\frac{a_3}{A_3}\sqrt{2gh_3} + \frac{(1-\gamma_b)k_2}{A_3}v_2 \tag{3}$$

$$\frac{dh_4}{dt} = -\frac{a_4}{A_4}\sqrt{2gh_4} + \frac{(1-\gamma_a)k_1}{A_4}v_1 \tag{4}$$

where

A_i Cross-sectional area of Tank *i*, $i = 1, \ldots, 4$
a_i Cross-sectional area of outlet hole
h_i Liquid level in tanks

The water is flowing into tanks through the pipe. The pump i requires a voltage v_i to run it and the flow into the tanks is $k_i v_i$. The parameters (γ_a, γ_b) are called the opening degree of the tee regulating valve. These parameters are important for finding the right half plane zero in the system. When $(\gamma_a + \gamma_b) < 1$, there is a multivariable right half plane zero. The flow of liquid into tank 1 is $\gamma_a k_1 v_1$and the corresponding flow of liquid into tank 4 is $(1 - \gamma_a)\ k_1 v_1$. Similarly, $\gamma_b k_2 v_2$ is the flow of liquid into tank 2 and $(1 - \gamma_b)\ k_2 v_2$ is the flow of liquid into tank 3. The process parameters of the FTS model are provided in Table 1 [3].

In this work, the minimum-phase characteristics of the FTS are analyzed for controlling the liquid level in tanks. The deviation variables are expressed as $H_i = h_i - \bar{h}_i$ and $u_i = v_i - \bar{v}_i$ where the steady-state values of h_i and v_i are $\bar{h}_i$ and $\bar{v}_i$. The nonlinear model of FTS are linearized as follows

$$\frac{dH}{dt} = \begin{bmatrix} -\frac{C_1}{A_1} & 0 & \frac{C_3}{A_1} & 0 \\ 0 & -\frac{C_2}{A_2} & 0 & \frac{C_4}{A_2} \\ 0 & 0 & -\frac{C_3}{A_3} & 0 \\ 0 & 0 & 0 & -\frac{C_4}{A_4} \end{bmatrix} H + \begin{bmatrix} \frac{\gamma_a k_1}{A_1} & 0 \\ 0 & \frac{\gamma_b k_2}{A_2} \\ 0 & \frac{(1-\gamma_b)k_2}{A_3} \\ \frac{(1-\gamma_a)k_1}{A_4} & 0 \end{bmatrix} u$$

Table 1 FTS model parameter values

Parameters	Value (cm^2)
a_1, a_3	0.071
a_2, a_4	0.057
A_1, A_3	28
A_2, A_4	32

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} H \tag{5}$$

where

$$C_i = a_i \sqrt{\frac{g}{2h_i}}, \quad i = 1, \ldots, 4$$

The FTS consists of four Acrylic type tanks. The liquid level in each tank is measured by the level transmitters (DPT). The FTS setup also has two control valves to adjust the flow rate of liquid into the tanks. There is a reservoir tank present below the tanks and the capacity of the reservoir tank is 75 L. To circulate the water from the reservoir tank to the tanks, two numbers of centrifugal pumps are used. The inlet flow rates into the tanks are visualized by using rotameters. The rotameters are in the range of (0–400) liters per hour. The dynamic model of FTS is required for simulation. Hence, Bernoulli's law and mass balance equation are used to derive the model of FTS are presented in Eqs. (1–4). The mathematical model is used to design the system.

3 Model Predictive Control

Various process industries use an advanced method called MPC since 1980s. MPC controllers depend on dynamic models of the process (linear empirical models obtained by using system identification methods). The basic block diagram of MPC controller is shown in Fig. 2. The MPC is required a model of the process. The future evolution of the process is found out by using the model of the process and the control signal is optimized. A model also describes the input to output behaviour of the process. There are two types of MPC calculations such as set-point calculations and control calculations which are executed at each sampling instants. MPC is the first choice on industrial applications than Internal Model control (IMC) and Smith Predictor because constrained MIMO control issues are better controlled by MPC.

A sequence of control moves are determined so that the predicted response goes to the desired reference value. In MPC controller, prediction horizon is defined as the number of predictions, P and the control horizon is defined as the number of control moves, M. A distinguishing feature of MPC is the receding horizon approach where a sequence of M control moves is determined at each sampling instant and the first input move is only executed. Then, a new sequence is determined at the next sampling instant and the first input move is only executed after the new measurements become available. This procedure is recapitulated at each sampling instant. Discrete-time models are more convenient for predictions. Hence discrete step response model is used in the MPC control calculations.

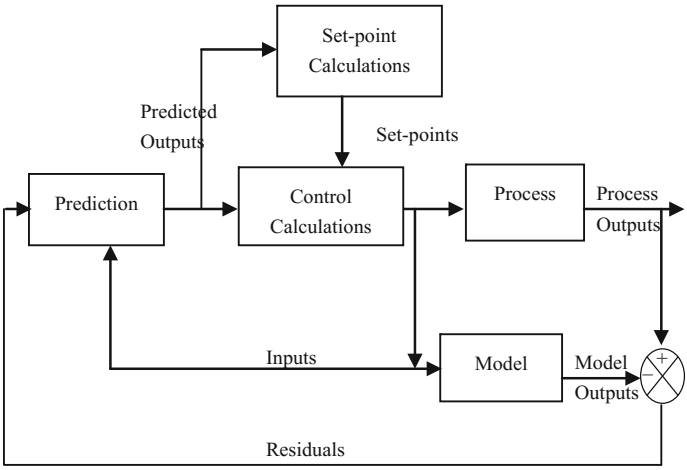

Fig. 2 Block diagram of MPC controller

3.1 Prediction for SISO Step Response Models

The one-step-ahead prediction is written as

$$\hat{y}(k+1) = \sum_{i=1}^{N-1} S_i \Delta u(k-i+1) + S_N u(k-N+1) \tag{6}$$

Similarly, the step response model of a SISO process for a *j-step-ahead prediction* is written as

$$\hat{y}(k+j) = \sum_{i=1}^{j} S_1 \Delta u(k-i+j) + \hat{y}^o(k+j) \tag{7}$$

where

$$\hat{y}^o(k+j) = \sum_{i=j+1}^{N-1} S_i \Delta u(k-i+j) + S_N u(k-N+j) \tag{8}$$

This term $\hat{y}^o(k+j)$ contains only the past control action, so, it is called as the predicted unforced response.

The MPC calculations are based on state-space models due to unified framework for both linear and nonlinear control problems. The control objective is to compute a sequence of input changes that produce the predictions as close to a reference trajectory as possible. A reference trajectory is provided to give a steady change to the desired set point.

3.2 MPC Objective Function

The control signal $\Delta \boldsymbol{U}(k)$ of rM-dimensional is determined to minimize

1. The predicted errors over the prediction horizon, P
2. The size of the control moves over the control horizon, M

An optimization objective function which is used for MPC controller is described as

$$\min_{\Delta \boldsymbol{U}(k)} \boldsymbol{J} = \hat{\boldsymbol{E}}(k+1)^{\mathrm{T}} \boldsymbol{Q} \hat{\boldsymbol{E}}(k+1) + \Delta \boldsymbol{U}(k)^{\mathrm{T}} \boldsymbol{R} \Delta \boldsymbol{U}(k) \tag{9}$$

where $\boldsymbol{Q}$ is an output weighting matrix and $\boldsymbol{R}$ is an input weighting matrix. Both $\boldsymbol{Q}$ and $\boldsymbol{R}$ are diagonal matrices. The quadratic objective function is minimized by MPC control law which can be determined mathematically as

$$\frac{\partial \boldsymbol{J}}{\partial \Delta \boldsymbol{U}(k)} = 0 \tag{10}$$

After simplification of Eqs. (9) and (10), we have

$$\Delta \boldsymbol{U}(k) = (\boldsymbol{S}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{S} + \boldsymbol{R})^{-1}\ \boldsymbol{S}^{\mathrm{T}} \boldsymbol{Q} \hat{\boldsymbol{E}}^{\boldsymbol{o}}(k+1) \tag{11}$$

The control law is then rewritten as

$$\Delta \boldsymbol{U}(k) = \boldsymbol{K}_{\boldsymbol{C}}\, \hat{\boldsymbol{E}}^{\boldsymbol{o}}(k+1) \tag{12}$$

where

$$\boldsymbol{K}_{\boldsymbol{C}} = (\boldsymbol{S}^{\mathrm{T}} \boldsymbol{Q}\, \boldsymbol{S} + \boldsymbol{R})^{-1}\ \boldsymbol{S}^{\mathrm{T}} \boldsymbol{Q} \tag{13}$$

Controller gain matrix, $\boldsymbol{K}_{\boldsymbol{C}}$ is based on the predicted error. For evaluating the value of $\boldsymbol{K}_c$ ($rM \times mP$), the dynamic matrix, $\boldsymbol{S}$ as well as the weighting matrices, $\boldsymbol{Q}$ and $\boldsymbol{R}$, are assumed to be constant. All physical systems have constraints. In MPC, the following constraints are normally defined to minimize the inequalities. Constraints in the outputs:

$$h_{\min} \le h \le h_{\max}$$

Constraints in the inputs:

$$\begin{aligned} &\Delta u_{\min} \le \Delta u \le \Delta u_{\max} \\ &u_{\min} \le u \le u_{\max} \\ &\text{where} \quad \Delta u(k) = u(k) - u(k-1) \end{aligned}$$

The MPC controller takes all these constraints into consideration when calculating the future controls.

Algorithm for MPC

1. The process output, $y(k)$ is measured and is used to estimate the disturbance.
2. The predicted unforced error, $\hat{\boldsymbol{E}}^{o}(k+1)$ is updated.

$$\hat{\boldsymbol{E}}^{o}(k+1) = \boldsymbol{Y}_{r}(k+1) - \hat{\boldsymbol{Y}}^{o}(k+1) - \phi\left[y(k) - \hat{y}(k)\right]$$

3. Solve for control moves

$$\Delta \boldsymbol{U}(k) = K_{C}\ \hat{\boldsymbol{E}}^{o}(k+1)$$

4. The first input step $\Delta \boldsymbol{U}(k)$ is only implemented.
5. Counter is updated. $k = k+1$

4 Simulation Results

Simulations are performed by taking the model derived for FTS using LabVIEW software. **T**he performance analysis is carried out for different controllers by employing a step input of amplitude 2 cm to the system. The minimum-phase characteristics of the FTS are analyzed for controlling the liquid level in tanks. The operating point values of the parameters of the FTS model are shown at Table 2. For the minimum-phase characteristics, Z_, time constants are shown in Table 3.

The physical modeling provides the transfer function matrices as

Table 2 Parameter values of the four tank system

Operating parameters	Value	Z_{-}
$\bar{h}_1, \bar{h}_2$	cm	12.4, 12.7
$\bar{h}_3, \bar{h}_4$	cm	1.8, 1.4
$\bar{v}_1, \bar{v}_2$	V	3.0, 3.0
k_1, k_2	$cm^3/V\ s$	3.33, 3.35
γ_a, γ_b		0.7, 0.6

Table 3 Time constants for minimum-phase system, Z_{-}

Time constants	Z_{-}
T_1, T_2	62.7, 90.3
T_3, T_4	23.9, 30.0

$$G_{-}(s) = \begin{bmatrix} \dfrac{2.61}{62.7s+1} & \dfrac{1.5}{(62.7s+1)(23.9s+1)} \\ \dfrac{1.41}{(90.3s+1)(30s+1)} & \dfrac{2.84}{90.3s+1} \end{bmatrix} \tag{14}$$

The response is plotted for the levels in the lower two tanks of the FTS. For the minimum-phase characteristics, the response of output level y_1 with input control signal, u_1 is shown in Fig. 3a. It comprises the responses of the set point and response of output level of tank 1 using MPC controller. The corresponding control input signal, u_1 is presented in Fig. 3b. Similarly, Fig. 3c shows the response of output level, y_2 with input control signal, u_2. Figure 3d shows it's corresponding input control signal, u_2 for the FTS. From the figure, it can be seen that the output level attains the desired reference value. The time of set point change is given at $t = 2$ s. The control signal is changed sharply at $t = 2$ s. After that the signal changes step by step and the plant output tracks the control signal accordingly.

The step response of different controllers for minimum phase characteristics is presented in Fig. 4.

The step response of output level, y_1 is shown in Fig. 4a. It comprises the responses of the set point, proportional-integral-derivative (PID) controller and MPC controller. The response of the PID controller has overshoot which is clearly shown from the figure. From the figure, it is observed that the MPC controller has no overshoot, but the PID controller has overshoot of 8.65% for the output level, y_1 and has overshoot of 15.48% for the output level, y_2.

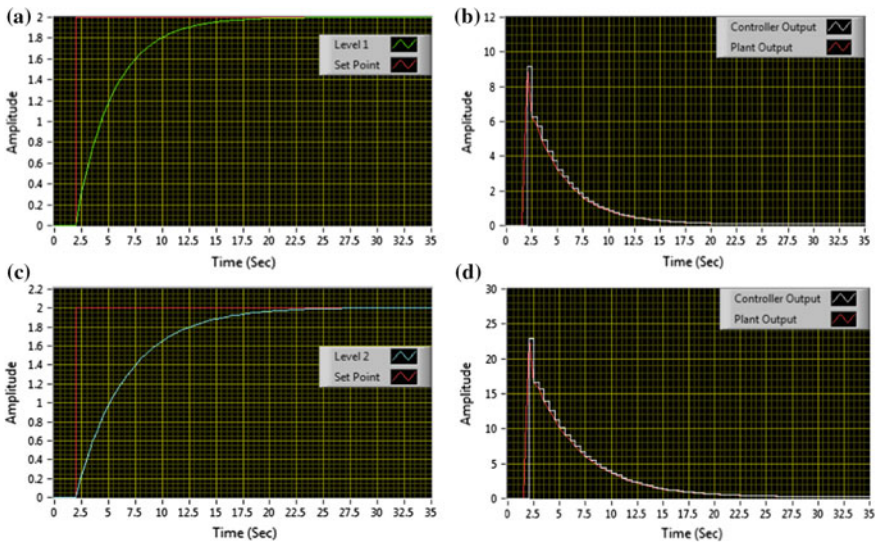

Fig. 3 Minimum-phase system: **a** response of output level y_1 with input u_1, **b** response of output level y_2 with input u_2, **c** control signal, u_1 and **d** control signal, u_2

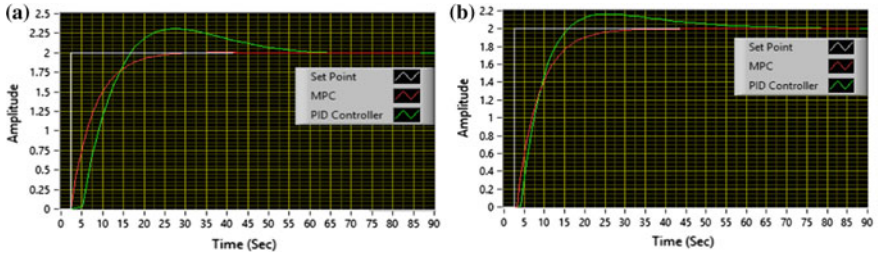

Fig. 4 Response of controllers for minimum-phase system: **a** output level, y_1, **b** output level, y_2

Table 4 Comparison of performance indices of different controllers for minimum-phase characteristics

Type		PID controller	MPC controller
Level output of tank 1	Rise time, t_r (s)	6.085	6.01
	Settling time, t_s (s)	61.09	33.07
	% Overshoot	8.65	0
	MSE	2.014	0.217
Level output of tank 2	Rise time, t_r (s)	8.81	4.04
	Settling time, t_s (s)	58.38	26.73
	% Overshoot	15.48	0
	MSE	2.32	0.28

The performance criterion for minimum-phase system characteristics are calculated for the two controllers are presented in Table 4. Four performance indices, rise time (t_r), settling time (t_s), percentage overshoot and mean square error (MSE) are compared in the table. Here, the objective is to track the output response to the desired reference model as close as possible. From the Table, it can be observed that the MPC controller has given better performance than PID controller in terms of rise time, settling time, % overshoot and MSE. For the level output of tank, the MPC controller has 45.9% less than that of PID controller in terms of settling time. Similarly, the MPC controller has 54.21% less than that of PID controller in terms of settling time for the level in tank 2. The MSE of PID controller is greater than that of the MPC controller. In case of minimum-phase characteristics, the PID controller has 89.23 and 87.83% greater MSE than that of MPC controller for output level, y_1 and output level, y_2 respectfully.

5 Conclusion

This paper describes about the control of the level of liquid in FTS using MPC controller. Mass balance and Bernoulli's law are applied for analyzing and modelling the FTS. The nonlinear model is linearized around an equilibrium point. The design procedure for MPC controller for FTS has been described. The step response of the process for minimum phase systems is studied and analyzed. Simulation results indicate that the MPC has more benefits than that of the PID controller. MPC controller exhibits stable response without overshoots for phase characteristics compared to PID controller. The MPC controller has low settling time, good robustness and also low MSE. Future work may address about the non-minimum phase characteristics of the FTS system as well as the rejection of disturbance which occurs due to parameter uncertainty.

References

1. Z. Aydogmus, Implementation of a fuzzy-based level control using SCADA. J. Expert Syst. Appl. **36**, 6593–6597 (2009). https://doi.org/10.1016/j.eswa.2008.07.055
2. B.W. Bequette, *Process Control: Modelling, Design and Simulation* (Prentice Hall, New Jersey, 2003)
3. K.H. Johansson, The quadruple-tank process: a multivariable laboratory process with an adjustable zero. IEEE Trans. on Contr. Syst. Technol. **8**, 456–465 (2000). https://doi.org/10.1109/87.845876
4. S. Kamalasadan, A.A. Ghandakly, A neural network parallel adaptive controller for dynamic system control. IEEE Trans. Instrum. Meas. **56**, 1786–1796 (2007). https://doi.org/10.1109/TIM.2007.895674
5. T.H. Liu, H.H. Hsu, Adaptive Controller design for a synchronous reluctance motor drive system with direct torque control. IET Electric Power Appl. **1**, 815–824 (2007). https://doi.org/10.1049/iet-epa:20070056
6. K.A. Mohideen et al., Real-coded Genetic Algorithm for system identification and tuning of a modified Model Reference Adaptive Controller for a hybrid tank system. J. Appl. Math. Model. **37**, 3829–3847 (2013). https://doi.org/10.1016/j.apm.2012.08.019
7. N. Muthukumar, V. Gomathi, K. Ramkumar, G. Balasubramanian, Prediction based optimal control of a quadruple tank process, in International Conference on Circuits, Power and Computing Technologies, vol. 23, pp. 685–690 (2013). doi:https://doi.org/10.1109/iccpct.2013.6529033
8. N.K. Poulsen, B. Kouvaritakis, M. Cannon, Constrained predictive control and its application to a coupled-tanks apparatus. Int. J. Control **74**, 552–564 (2001)
9. S.B. Prusty, U.C. Pati, K.K. Mahapatra, Implementation of fuzzy-PID controller to liquid level system using LabVIEW, in *Proceedings of this International Conference on Control, Instrumentation, Energy and Communication (CIEC)*, pp. 36–40 (2014). doi:https://doi.org/10.1109/ciec.2014.6959045
10. S.B. Prusty, U.C. Pati, K.K. Mahapatra, A novel fuzzy based adaptive control of the four tank system, in *Proceedings of this International Conference on Computer, Communication, Control and Information Technology (C3IT)*, pp. 1–6 (2015). doi:https://doi.org/10.1109/c3it.2015.7060161

11. D.E. Seborg, D.A. Mellichamp, T.F. Edgar, *Process Dynamics and Control* (Wiley, New York, 2010)
12. C.A. Smith, A.B. Corripio, *Principles and Practice of Automatic Process Control* (Wiley, New York, 1997)
13. P. Srinivasarao, P. Subbaiah, Centralized and decentralized of quadruple tank process. Int. J. Comput. Appl. **68**, 21–29 (2013). https://doi.org/10.5120/11656-7167
14. T. Tani, S. Murakoshi, M. Umano, Neuro-fuzzy hybrid control system of tank level in petroleum plant. IEEE Trans. Fuzzy Syst. **4**, 360–368 (1996). https://doi.org/10.1109/FUZZY.1993.327417
15. J. Zumberge, K.M. Passino, A Case study in intelligent vs. conventional control for a process control experiment. J. Control Eng. Pract. **6**, 1055–1075 (1998). https://doi.org/10.1016/S0967-0661(98)00067-7

Comparative Study of Duval Triangle with the New DGA Interpretation Scheme

Awin Gupta, Kavish Jain, Yog Raj Sood and Naveen Kumar Sharma

Abstract Fault diagnosis of power transformer for condition evaluation is being carried out by oil analysis for decades. The most efficient and easy way of condition evaluation of transformer is by evaluating the concentration of gases present in the transformer oil, done using dissolved gas analysis (DGA) over which several techniques are employed to predict the type of fault present in the transformer. Among all the techniques, Duval triangle is the most prominent. However, due to ignorance of the two combustible gases, namely ethane (C_2H_6) and hydrogen (H_2), this technique employs lower accuracy for certain faults type. The paper presents a new DGA interpretation scheme carried out through several case studies using gas ratios and their proportion limit as rules implemented in MATLAB to diagnose the fault. The gases produced during the deterioration of transformer oil and insulation paper have been examined using Duval triangle and the new DGA interpretation scheme, on the DGA samples collected from TIFAC-CORE Laboratory, NIT Hamirpur, and published papers. The result obtained from the new DGA interpretation scheme showed promising results compared to the conventional Duval triangle.

Keywords Dissolved gas analysis · Fault diagnosis · Duval triangle
New DGA interpretation scheme

1 Introduction

The power transformer is the most valuable equipment in the power generation, transmission, and distribution sector [1]. Nowadays, the demand for electricity in commercial, domestic, industrial, and agricultural sectors is increasing rapidly, the consequence of which is surge in transformer, which increases chances of fault occurrence. So for reliable operation of the transformer, it has become mandatory to

A. Gupta (✉) · K. Jain · Y. R. Sood · N. K. Sharma
Electrical Engineering Department, National Institute of Technology, Hamirpur, India
e-mail: gupta.awin2105@gmail.com

S. N. Singh et al. (eds.), *Advances in System Optimization and Control*,
Lecture Notes in Electrical Engineering 509,
https://doi.org/10.1007/978-981-13-0665-5_24

continuously monitor and diagnose the insulation condition so that it can detect the incipient fault if any [2–8]. Dissolved gas analysis (DGA) is the most prominent technique used for detecting the incipient fault occurring in the transformer.

As the temperature in the transformer reaches beyond a particular limit in the presence of oxygen (O_2) and moisture, thermal and electrical stresses are introduced degrading the physical, chemical, and electrical properties of oil. This oil degradation is the major cause of transformer failure [9, 10]. Due to these stresses, the carbon chain of oil breaks, thereby introducing hydrocarbon gases in oil as by-products, namely methane (CH_4), hydrogen(H_2), ethane(C_2H_6), ethylene(C_2H_4), acetylene(C_2H_2), and degradation of insulation paper introduces carbon monoxide (CO) and carbon dioxide (CO_2) in oil. These gases introduce further electrical and thermal faults in transformer which are categorized as per IEC 60599 and IEEE C.57.104 standards as shown in Fig. 1. The electrical faults are categorized into partial discharge (PD), discharge of low energy (D1), and discharge of high energy (D2), whereas thermal faults are categorized based on the temperature range as thermal faults of temperature $T < 300$ °C (T1), $300 < T < 700$ °C (T2), and $T > 700$ °C (T3) [11].

There are several methods for detection of incipient faults in power transformer such as ratio method, key gas method, Duval triangle, artificial intelligence-based methods. [12]. In all these methods, Duval triangle came out to be the most accurate method but the problem with it is, it also does not yield 100% accuracy.

From this point of view, a new technique is required, which can detect faults with better performance than Duval triangle. So a new DGA interpretation scheme has been employed in this paper. The new technique is based on the certain rules programmed in MATLAB which can directly provide the fault present in the transformer. This paper evaluates and compares the results obtained from Duval triangle and the new DGA interpretation scheme.

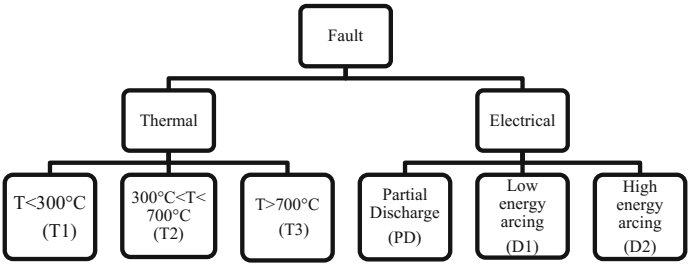

Fig. 1 Categorization of fault as per IEC 60599 and IEEE C.57.104 standards

2 Basic Knowledge

2.1 Graphical Method: Duval Triangle

The Duval triangle fault diagnostic method is a graphical approach first introduced by Michel Duval in 1974 [13]. This method uses a different set of hydrocarbon gases, placed on the three sides of an equilateral triangle representing relative proportion of gases [14]. Till date, a total of seven Duval triangles have been introduced by Duval on the basis of oil used in the equipment. The triangles 1, 2, 3 uses CH_4, C_2H_4, and C_2H_2 gases, and triangles 4 and 6 use H_2, C_2H_6, and CH_4 gases, whereas triangles 5 and 7 uses CH_4, C_2H_6, C_2H_4 gases [15]. In this paper, we are concerned with triangle 1 of mineral oil to study the type of fault present in the transformer.

The triangle is subdivided into seven zones, each zone signifying a type of fault as mentioned in Fig. 1 with one intermediate fault DT signifying mixture of electrical and thermal fault in the power transformer. The Duval triangle is shown in Fig. 2.

The relative percentage of each gas is given by

$$\%CH_4 = \frac{100(CH_4)}{CH_4 + C_2H_4 + C_2H_2} \tag{1}$$

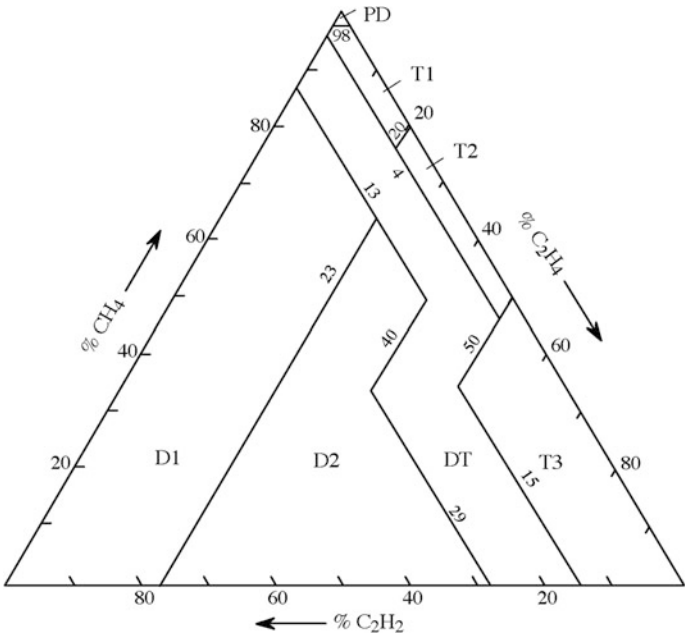

Fig. 2 Transformer fault diagnosis using Duval triangle

$$\%C_2H_4 = \frac{100(C_2H_4)}{CH_4 + C_2H_4 + C_2H_2} \tag{2}$$

$$\%C_2H_2 = \frac{100(C_2H_2)}{CH_4 + C_2H_4 + C_2H_2} \tag{3}$$

where CH_4, C_2H_4, C_2H_2 concentrations all are in ppm.

Plotting the relative percentage of each gas in the triangle provides only one point in the triangle. The location of point at a particular zone decides the type of fault present in the transformer.

2.2 New DGA Interpretation Technique

The techniques use relative proportion of four combustible gases ($\%H_2$, $\%C_2H_2$, $\%CH_4$, $\%C_2H_4$) out of seven gases present in the oil sample to detect the fault type [16]. For this, certain rules are employed derived from several case studies. Table 1 shows the rules employed in the technique, and Table 2 shows the ratios associated with the rules. The percentage relative proportion of CH_4, C_2H_2, C_2H_4 is as shown in Eqs. (1)–(3), respectively, and the percentage relative proportion of H_2 gas is given by

$$\%H_2 = \frac{100(H_2)}{H_2 + CO + CO_2 + C_2H_6} \tag{4}$$

where H_2, C_2H_6, CO, CO_2 concentrations all are in ppm.

The faults in this technique are categorized mainly into five types usually found in transformers in service as partial discharge (PD), discharge of low energy (D1), discharge of high energy (D2), thermal faults of T < 700 °C (T1/T2), and thermal faults of T > 700 °C (T3) [17].

Initially, step 1 is used to detect the PD fault type. If rule 1 is true, the fault is PD. If not, step 2 is employed in which rule 2 is checked for identifying whether fault is electrical fault (D1 or D2) or thermal fault ((T1/T2), T3). If rule 2 is true implies the fault is electrical fault type; further in step 3, rule 3 is checked. If rule 3 is true, the fault is D1 else D2. Also, if rule 2 is false, the fault lies in the thermal

Table 1 Rules employed

Fault	Rule
PD	($\%CH_4 > 96$ and $\%H_2 > 90$) or ($\%CH_4 > 96$ and $H_2 > 2500$)
D1/D2	($\%C_2H_2 > 13$ and R1 > 1.4)
D1	($\%C_2H_2 > 13$ and $\%C_2H_4 < 23.3$) or (R2 > 15)
T3	(R3 > 3.5 and R4 > 0.015) or (R5 > 0.8 and R6 > 2.6)

Table 2 Ratios involved

S. No	Code	Ratio
1	R1	(C_2H_2/C_2H_4)
2	R2	(C_2H_2/CH_4)
3	R3	(C_2H_4/C_2H_6)
4	R4	(C_2H_2/H_2)
5	R5	(C_2H_4/CH_4)
6	R6	(CH_4/H_2)

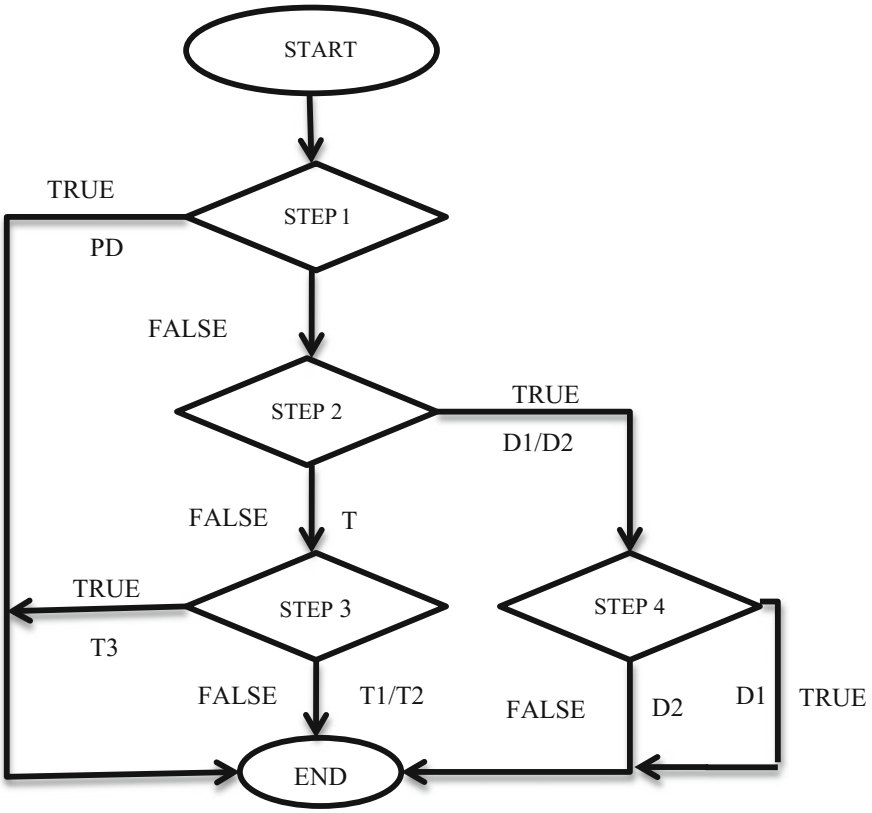

Fig. 3 Flowchart for fault diagnosis [16]

fault category; further in step 4, rule 4 is checked. If rule 4 is true, the fault is T3 else T1 or T2. Figure 3 shows the flowchart associated with the technique [16].

The fault diagnosis using the proposed technique is carried out by developing and running a MATLAB program based on the rules and the flowchart as shown in Table 1 and Fig. 3, respectively.

3 Results and Discussion

The proposed new technique is compared with the conventional Duval triangle, the most efficient technique till date. A total of 109 DGA samples have been analyzed to predict and compare the percentage accuracy of the two techniques. Out of 109 DGA samples, 29 samples and 10 samples are collected from reference [4] and reference [18], respectively. Remaining 70 samples has been collected from TIFAC-Core Laboratory, NIT Hamirpur.

The 29 DGA samples from reference [4] include several different types of oil-filled electrical equipment such as power transformers with or without on load tap changers, instrument transformer, bushing, reactor, and cable. The other 70 DGA samples are of various oil-filled working transformers currently operating in Himachal Pradesh, India.

For the new technique, a total of 7 DGA samples became a challenging task due to unavailability of one or more gases introducing (0/0) or (a/0) condition. For this, the ratio (0/0) is set to 0 and ratio (a/0) where 'a' is not zero is set to 25 in the MATLAB programming.

Table 3 shows some of the DGA samples and the fault obtained by the two techniques. Table 4 shows the number of correct diagnoses, wrong diagnosis, and the percentage accuracy of the two techniques for all the DGA samples collected.

Table 3 Some DGA samples and their fault diagnosis using (a) Duval triangle and (b) new proposed method [4, 18]

S. No	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	CO	CO_2	(a)	(b)
1	384	388	33	110	55	173	932	DT	T1/T2
2	83	52	61	2937	555	892	6081	T3	T3
3	645	86	317	110	13	74	114	D1	D1
4	12	25	21	56	98	492	3574	DT	T1/T2
5	6	23	12	11	10	451	6244	T2	T1/T2
6	41	45	147	110	11	297	1807	D2	D2
7	15	31	13	21	27	316	3487	D2	D2
8	5	23	12	32	41	248	894	DT	T3
9	1820	405	634	365	35	1010	8610	D2	D2
10	6	2990	67	26,076	29,990	6	26	T3	T3

Table 4 Diagnosis result (in percentage)

Case	No of samples	Correct diagnosis		Wrong diagnosis		% Accuracy	
		Duval triangle	New method	Duval triangle	New method	Duval triangle	New method
1	10	8	9	2	1	80	90
2	29	21	25	8	4	72.41	86.20
3	70	57	63	13	7	81.42	90
Total	109	86	97	23	12	78.89	88.99

From Table 4, it is clear that the proposed new technique has better performance than the conventional Duval triangle with 88.99% accuracy and the method can be universally accepted as a new technique for fault diagnosis.

4 Conclusion

The new proposed method uses gas ratios and proportion as their main tool to detect the type of fault in the oil-filled transformers. The method is simple and easy to use. Several steps and the rules associated with the method can easily be programmed in MATLAB for fault diagnosis. The method calculates higher accuracy compared to conventional Duval triangle which is considered as the most accurate method till date. The reason for higher accuracy is that the proposed method uses all the seven gases concentration for fault diagnosis whereas the Duval triangle uses only three gases.

References

1. N.A.B. Muhamad, N. Bashir, A.A.S. Alghamdi, A.A. Suleiman, Asset management through effective transformer diagnostics & condition monitoring, in *IEEE International Conference on Power and Energy (PECon)* (Kota Kinabalu Sabah, Malaysia, Dec 2012), pp. 212–216
2. IEEE guide for the interpretation of gases generated in oil-immersed transformers. IEEE std. C57.104-2008 (New York, USA, 2009)
3. IEC guide to the interpretation of dissolved and free gases analysis-mineral oil-impregnated electrical equipment in service. IEC60599 (Geneva, Switzerland, 1999)
4. M. Duval, A. de Pablo, Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC10 data bases. IEEE Electr. Insul. Mag. **17**(2), 31–41 (March/April 2001)
5. I.A.R. Gray, A guide to transformer oil analysis. Transformer Chem. Serv. http://www.satcs.co.za
6. G.K. Irungu, A.O. Akumu, J.L. Munda, Transformer condition assessment using dissolved gas analysis, oil testing and evidential reasoning approach, in *Proceedings of 33rd EIC 2015* (Seattle, USA, June 2015), pp. 145–149
7. Z. Mao, J. Wen, Detection of dissolved gas in oil-insulated electrical apparatus by photoacoustic spectroscopy. IEEE Electr. Insul. Mag. **31**(4), 7–14 (2015)
8. W.H. Tang, Q.H. Wu, *Condition monitoring and assessment of power transformers using computational intelligence* (Springer-Verlag limited, London, UK, 2011)
9. M. Wang et al, Review of condition assessment of power transformers in service. IEEE Electr. Insul. Mag. (2002)
10. P. Gill, Electrical *Power Equipment Maintenance and Testing*, 2nd ed. (CRC Press, 2009)
11. G.K. Irungu, A.O. Akumu, J.L. Munda, Fault diagnosticsin oil filled electrical equipment: review of duval triangle and possibility of alternatives, in *2016 Electrical Insulation Conference (EIC)* (Montréal, Qc, Canada, 19–22 June 2016), pp. 174–177
12. D.-E.A. Mansour, Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases. IEEE Trans. Dielectr. Electr. Insul. **22**(5), 2507–2512 (October 2015)

13. M. Duval, Fault gases formed in oil-filled breathing EHV power transformers—the interpretation of gas analysis data, in *IEEE PAS Conference*, Paper no C 74 476-8 (1974)
14. M. Duval, A review of faults detectable by gas-in-oil analysis in transformers. IEEE Electr. Insul. Mag. **18**(3), 8–17 (2002)
15. M. Duval, The duval triangle for load tap changers, non-mineral oils and low temperature faults in transformers. IEEE Electr. InsuI. Mag. **24**(6), 22–29 (2008)
16. X. Li, H. Wu, D. Wu, DGA interpretation scheme derived from case study. IEEE Trans. Power Delivery, 1292–1293
17. A.-X. Zhao, X.-J. Tang, Z.-H. Zhang, J.-H. Liu, The DGA interpretation method using relative content of characteristic gases and gas-ratio combinations for fault diagnosis of oil-immersed power transformers, in *Conference Proceedings of ISEIM 2014*, pp.124–127
18. A. Abu-Siada, S. Islam, A new approach to identify power transformer criticality and asset management decision based on dissolved gas-in-oil analysis. IEEE Trans. Dielectr. Electr. Insul. **19**(3), 1007–1012 (June 2012)