Lecture Notes in Electrical Engineering 656

J. Jayakumari
George K. Karagiannidis
Maode Ma
Syed Akhter Hossain *Editors*

Advances in Communication Systems and Networks

Select Proceedings of ComNet 2019

Springer

Lecture Notes in Electrical Engineering

Volume 656

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

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

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

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

China

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

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

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

USA, Canada:

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

All other Countries:

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

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

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

J. Jayakumari · George K. Karagiannidis ·
Maode Ma · Syed Akhter Hossain
Editors

Advances in Communication Systems and Networks

Select Proceedings of ComNet 2019

Springer

Editors
J. Jayakumari
Department of Electronics and Communication Engineering
Mar Baselios College of Engineering and Technology
Thiruvananthapuram, Kerala, India

Maode Ma
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore, Singapore

George K. Karagiannidis
Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki
Thessaloniki, Greece

Syed Akhter Hossain
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-15-3991-6 ISBN 978-981-15-3992-3 (eBook)
https://doi.org/10.1007/978-981-15-3992-3

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

Preface

The International Conference on Communication Systems and Networks (ComNet-2019) was organized by the Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology in association with the University of Dayton, Ohio, and with Springer as publication partner during 12th and 13th Dec 2019.

This conference aims to bring researchers, industrialists and experts from India and abroad on a single platform to interact with and discuss the latest research results, ideas, developments, and applications in all areas of advanced communication systems and networking. The topics covered include latest research in next-generation wireless technologies including 5G, new hardware platforms, recent advancements in antenna design, applications of artificial intelligence, signal processing and optimization techniques. There were three tracks—communication systems, signal processing and networking technologies.

The conference was inaugurated by **Dr. V. K. Dadhwal, Director, Indian Institute of Space Science and Technology, Valiamala.** There were invited talks by **Prof. Sharul Kamal Bin Abdul Rahim,** Faculty of Electrical Engineering, Universiti Teknologi Malaysia, **Prof. Varun Jeoti,** Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS Malaysia, **Prof. Ajit Kumar Panda,** Dean, NIST, Odisha, **Dr. T. Shanmuganantham,** Associate Professor, Department of Electronics Engineering, Pondicherry Central University and **Dr. Sujesh Sreedharan,** Engineer 'F', Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram. All the technical sessions were chaired by the experts in the field.

I hope that the proceedings will be useful for all researchers working in the relevant areas.

Dr. J. Jayakumari
Conference Chair, ComNet 2019
Professor, Department of ECE
Mar Baselios College of Engineering and Technology
Thiruvananthapuram, Kerala, India

Contents

About the Editors

Dr. J. Jayakumari is presently working as Professor, Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala. She obtained her B.E. (ECE) from M.S. University, Tirunelveli in 1994, M. Tech. (Applied Electronics and Instrumentation) in 1998 and Ph.D. in 2009 from Kerala University. Her major areas of research interests include wireless communication, coding techniques, signal and image processing. She has published more than 75 papers in international journals and conferences. She is a senior member of IEEE, Fellow of IETE, IE (I) and CET (I), Chartered Engineer of IE (I) and member of The Society of Digital Information and Wireless Communications (SDIWC). She is also an Academic Council member of the Council of Engineering and Technology (India).

Dr. George K. Karagiannidis received his university degree in 1987 and his Ph.D. degree in 1998, both in Electrical Engineering, from the University of Patras, Patras, Greece. From 2001 to 2004, he was researcher at the Institute for Space Applications and Remote Sensing, National Observatory of Athens, Greece. In June 2004, he joined the Electrical & Computer Engineering Department of Aristotle University of Thessaloniki, Greece, where he is currently Professor. His research interests include wireless communications, wireless power transfer and applications, optical wireless communications, communications for biomedical engineering, wireless security. He has published and presented more than 500 technical papers in scientific journals and international conferences. He is co-recipient of the Best Paper Award of the Wireless Communications Symposium (WCS) in IEEE International Conference on Communications (ICC'07), Glasgow, UK, June 2007, and the Best Paper Award of the IEEE VTC 2018 (Spring), Porto, June 2018. He is a Fellow IEEE, Class 2014. From January 2012 to December 2015, he was the Editor-in-Chief of IEEE Communications Letters.

Dr. Maode Ma, a Fellow of IET, received his Bachelor's degree from the Department of Computer Science and Technology in Tsinghua University in 1982, his Master's degree from the Department of Computer Science and Technology in

Tianjin University in 1991, and his Ph.D. degree in the Department of Computer Science from Hong Kong University of Science and Technology in 1999. Presently, Dr. Ma is an Associate Professor in the School of Electrical and Electronic Engineering at Nanyang Technological University in Singapore. He has extensive research interests including network security and wireless networking. He has led 25 research projects funded by government, industry, military and universities in various countries. He has been a conference chair, technical symposium chair, tutorial chair, publication chair, publicity chair and session chair for over 100 international conferences. He has been a member of the technical program committees for more than 200 international conferences. Dr. Ma has about 400 international academic publications including about 190 journal papers and 200 conference papers. He has edited 4 technical books and produced over 25 book chapters.

Dr. Syed Akhter Hossain is currently working as Professor and Head of the Department of Computer Science and Engineering at Daffodil International University, Bangladesh. Dr. Hossain obtained M.Sc. in Applied Physics and Electronics from Rajshahi University and Ph.D. in Computer Science and Engineering with outstanding distictions. He attended Erasmus Mundus post-doctoral fellowship in the area of Informatics and Industrial Engineering with University Lumiere Lyon 2 in France. He has more than 29 years of working experience in industry, education, research and training. He is actively involved in research guidance, research projects, collaborations with institutes and industries. He has more than 130 publications and his work is listed in DBLP, IEEE Explore and other research databases. He received the National ICT Award 2016 for his outstanding contribution in ICT education. He also received other international awards for teaching including other national awards for his scholastic works specially for the contribution of machine translator for Bangla Braille used by the visually impaired society. He is also member of several national and international professional bodies.

Analysis of Active Contours Without Edge-Based Segmentation Technique for Brain Tumor Classification Using SVM and KNN Classifiers

A. S. Remya Ajai and Sundararaman Gopalan

Abstract Classification of brain tumors using machine learning technology in this era is very relevant for the radiologist to confirm the analysis more accurately and quickly. The challenge lies in identifying the best suitable segmentation and classification algorithm. Active contouring segmentation without edge algorithm can be preferred due to its ability to detect shapeless tumor growth. But the perfectness of segmentation is influenced by the image enhancement techniques that we apply on raw MRI image data. In this work, we analyze different pre-processing algorithms that can be applied for image enhancement before performing the active contour without edge-based segmentation. The accuracy is compared for both linear kernel SVM and KNN classifiers. High accuracy is achieved when image sharpening or contrast stretching algorithm is used for image enhancement. We also analyzed that KNN is more suitable for brain tumor classification than linear SVM when active contouring without edge method of segmentation technique is used. MATLAB R2017b is used as the simulation tool for our analysis.

Keywords Active contouring · SVM · KNN · Gaussian filtering · Tumor

1 Introduction

Biomedical image processing and analysis using machine learning algorithms has become a vital domain in the image processing as well as machine learning research area [1]. The raw image data obtained from MRI scanners may not be directly used for analysis. It has to be enhanced, segment out the unwanted regions so that the machine learning algorithm can work only on the noise-free, relevant image segment for accurate analysis with less memory and high speed. Many machine learning

A. S. Remya Ajai (✉) · S. Gopalan
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: remya.amrita@gmail.com

S. Gopalan
e-mail: sundarg99@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_1

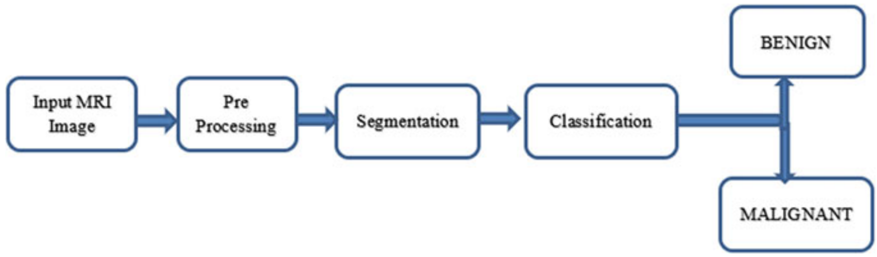


Fig. 1 Block diagram of image classifier system with machine learning approach

algorithms have formulated by the various researchers over years to identify the best method for image data processing.

In [2], an overview of the statistical feature extraction methods for classifying tumor and normal T2 MRI image is discussed. For the image segmentation, the authors in [3] found that morphological operation is faster than region growing segmentation algorithm. Anisotropic filtering is performed in [4] for breast images and observed that it preserves edges more effectively than the Gaussian filtering. Watershed-based segmentation and canny edge detection algorithms are compared in [5]. Better accuracy and precision values are obtained for watershed algorithm. Gabor filters are used for extracting features of image data in [6], and the authors found that when these features are given to SVM algorithm for classification, high accuracy of 93.54% is achieved.

Brain tumor classification using MRI images with different machine learning approaches is a subdomain in biomedical image analysis. Literatures reveal that many researches are still going on to find the best algorithm combinations for different steps in the classification of tumors as benign or malignant [7–10]. Different steps involved in brain tumor classification are given in Fig. 1.

The MRI image obtained from the scanner cannot be used directly for segmentation and analysis. The image has to be enhanced so that the tumor part can be clearly segmented out. This is the very critical step in the process. Tumor can come with vague shapes and intensities. Hence, the success of segmentation highly depends on its previous step say pre-processing or enhancement. There are several image enhancement techniques, but the selection of it for a specific application with better accuracy is very critical.

In this work, we analyze four different image pre-processing techniques, namely image sharpening, contrast stretching, Gaussian filtering and homomorphic filtering. The segmentation algorithm identified is active contouring without edges [11]. Segmentation using active contouring method and snakes are analyzed in [12, 13]. Pre-processing techniques have a significant impact on the segmentation algorithm since a perfect pre-processed image can effectively be segmented and classified. Hence, we compared KNN and linear SVM classifier algorithm which are widely used in this domain to analyze the best pre-processing technique for brain tumor classification.

2 Pre-processing Techniques

In biomedical image analysis, the raw input image cannot be used for processing directly. It needs to be enhanced before applying the segmentation so that the perfect segmentation of desired object can be obtained. There are several image pre-processing methods used in the implementation of biomedical classification algorithm using machine learning approach [14]. In this work, we analyzed the various pre-processing techniques, namely image sharpening, contrast stretching, Gaussian filtering and homomorphic filtering for brain tumor classification using T2 MRI images.

2.1 Image Sharpening

Image sharpening is an image enhanced technique in which edges and fine details in the image are concentrated [15]. By sharpening the image, the edges are well demarcated and hence help in the segmentation of specified object.

2.2 Contrast Stretching

In contrast stretching, the contrast of the image is improved by increasing the intensity ranges to span a range of desired values. It allows the full range allowed pixel values possible. Linear and nonlinear contrast stretching methods are suitable as image pre-processing steps for better segmentation [16].

2.3 Gaussian Filtering

Noise removal in an image can be achieved by a Gaussian filter. It blurs the edges and thus removes noises. This is preferred over median filtering because the basic convolution operation involved in the Gaussian filtering is faster than the sorting operation in median filter. This method is used for image recognition [17].

2.4 Homomorphic Filtering

Homomorphic filtering is an image pre-processing technique which involves a nonlinear mapping of an image to a different domain where linear filter techniques can be applied. This is followed by mapping to the original image domain [18].

3 Active Contouring Without Edge Segmentation Technique

Active contour modeling is used in image analysis for describing an image outline of a particular region within a noisy original image data. In this technique, a deformable model is tried to match an object in an image with a deformable energy minimization model. Matching is obtained by identifying the contours. It is widely used for image segmentation due to its natural ability to capture any variations in the shape of any object [13].

A simple active contour model can be defined by a set of n points V_{j}, where $j = 0, 1, 2, \ldots n - 1$. The forces in active contours contain external forces and image forces. When the model starts deforming from a closed curve, the internal and external forces are influenced to minimize the energy.

$$E = E_{\text{internal}} + E_{\text{external}} \tag{1}$$

Internal energy, E_{internal}, quantifies the smoothness of the contour. It gives control on how much the deformation needed. The parameters $\alpha(s)$ and $\beta(s)$ in Eq. (2) refer to user-defined weights.

$$E_{\text{internal}} = \frac{1}{2}\left(\alpha(s)\left\|\frac{\mathrm{d}\bar{v}}{\mathrm{d}s}(s)\right\|^2 + \beta(s)\left\|\frac{\mathrm{d}^2 v}{\mathrm{d}s^2}(s)\right\|^2\right) \tag{2}$$

The external energy (E_{external}) combines the forces due to image itself (E_{image}) and the user introduced constraint forces (E_{con}).

$$E_{\text{external}} = E_{\text{image}} + E_{\text{con}} \tag{3}$$

Constraints for external energy are mainly used for defining the snake near to the local minimum.

The target image contour is given by Eq. 4, where u_1 refers to line efficient, and u_2 refers to edge efficient.

$$E_{\text{image}} = u_1 I(x, y) + u_2 |\nabla I(x, y)|^2 + \ldots \tag{4}$$

For large values of u_1 and u_2, the contour will align to darker pixel. When their values are smaller, it will progress to bright pixels.

In the active contouring methods described above, an edge detector is used based on the image gradient. But for detecting objects with borders that cannot be clearly defined by gradients, active contour modeling may not be suitable. Since our work focuses on identifying irregular tumors from MRI images, we referred to the method of active contouring without [11] edge-based segmentation algorithm.

Similar to active contouring, deforming the contour with energy minimization is done in this method, and the convergence stops on the desired boundary but does not depend on the gradient of the image. Unlike the active contouring model, the initial boundary can be anywhere in the image.

In active contours without edge algorithm, the total force E is given by Eq. 5. $E_1(C)$ and $E_2(C)$ are the forces to shrink the contour and expand the contour, respectively.

$$E = E_1(C) + E_2(C) \tag{5}$$

$$E = \int\limits_{\text{inside}(c)} |i_0 - c_1|^2 \mathrm{d}x + \int\limits_{\text{outside}(c)} |i_0 - c_2|^2 \mathrm{d}x \tag{6}$$

In Eq. 6, c_1 represents mean of all components inside the contour c, and c_2 represents mean of all components outside the contour C. The variable i_0 represents the entire image.

The initial contour will cover the entire object and may be some background. The internal and external means are calculated and the contours continuous to shrink or expand until both means become zeros. Then, the curve fits with the object, and thus, the segmentation is performed. At this point, the energy will be minimum.

Active contour model without edges mentioned above is a specific case of minimal partition problem where best approximation of i to i_0 as a function of two mean values is done. Level set method is used to solve this minimal partition problem [11]. In level set function, the contour C has the initial value specified in Eq. 7

$$C = \{(u, v)|\varphi(u, v)\} = 0 \tag{7}$$

$$\begin{aligned} F(c_1, c_2, \varphi) = &\int\limits_{\Omega} (i_0(u, v) - c_1)^2 H(\varphi) \mathrm{d}u \mathrm{d}v \\ &+ \int\limits_{\Omega} (i_0(u, v) - c_2)^2 (1 - H(\varphi)) \mathrm{d}u \mathrm{d}v \\ &+ v \int\limits_{\Omega} |\nabla H(\varphi)| \end{aligned} \tag{8}$$

The function F shown in Eq. 8 depends on the Heaviside function $H(.)$ and the input image. For minimizing F, the derivatives are taken and equate to zeros. The variables $c1$, $c2$ and φ can be updated recursively using Eqs. 9 and 10.

$$c_1(\varphi) = \frac{\int_\Omega i_0(u, v) H(\varphi(t, u, v)) \mathrm{d}u \mathrm{d}v}{\int_\Omega H(\varphi(t, u, v)) \mathrm{d}u \mathrm{d}v} \tag{9}$$

$$c_2(\varphi) = \frac{\int_\Omega i_0(u, v)(1 - H(\varphi(t, u, v))) \mathrm{d}u \mathrm{d}v}{\int_\Omega 1 - H(\varphi(t, u, v)) \mathrm{d}u \mathrm{d}v} \tag{10}$$

$$\frac{\partial\varphi}{\partial t} = \delta\,\varphi[\breve{v}\,\mathrm{div}\left(\frac{\nabla\varphi}{|\nabla\varphi|}\right) - (i_0 - c_1)^2 - (i_0 - c_2)^2 \tag{11}$$

In the recursive function given in Eq. 11, $\delta(.)$ represents the Dirac function.

In our work, we tried this ability of active contours without edges to segment the tumor regions in an MRI data. This is quite challenging because the contours in MRI images cannot be identified properly unless the raw MRI data are perfectly pre-processed. Hence, we tried various pre-processing methods that are generally used by researchers, for segmentation and analyzed the best one which is specific to active contour without edge segmentation technique for better accuracy.

The recursive function is implemented on images taken from the OASIS dataset (URL: http://www.oasis-brains.org/), and the function is executed with various iteration levels. As shown in Fig. 2, the contour gets fitted with the border of tumor present in the selected MRI T2 image slice, when the iteration value is 250. Since minimum energy is achieved at this point, any further increase in the iteration level will not change the contour shape.

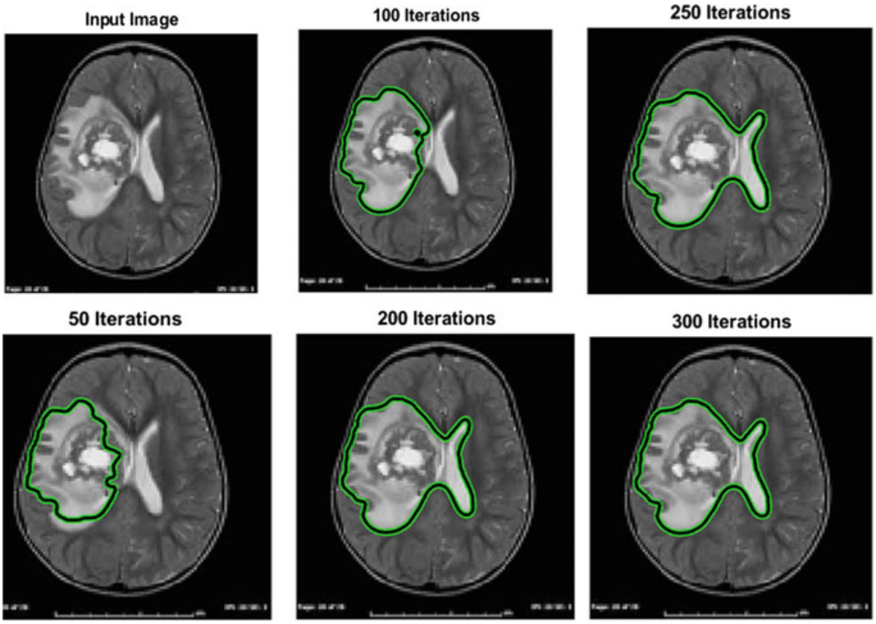


Fig. 2 Segmenting the tumor part from a pre-processed MRI image using active contouring without edge methodology for various iterations

4 Classification Algorithms

Machine learning algorithms like K-nearest neighbors (KNN) and support vector machine (SVM) are widely used for biomedical applications [19]. KNN is a simple, supervised type of algorithm which can be used for both classification and regression type of problems [20]. When KNN is used as a classifier, it identifies the objects based on their closest training data in the feature space for a set of data. For distance calculation, Euclidean algorithm can be preferred.

$$d(p,q) = d(q,p) = \sqrt{\sum_{i=1}^{n}((q_i - p_i)^2)} \tag{12}$$

SVM is also a supervised learning algorithm which is a kernel-based binary classifier. In SVM, the training data are nonlinearly projected to a higher-dimensional feature space which results in a linearly separable data. The data are classified by finding a hyperplane that separates one class of data from another class. Data points nearer to the hyperplane are called support vectors. In our training, linear kernel is used for SVM training.

5 Results and Discussions

For analyzing the best pre-processing algorithm that can be used with active contouring without edge method for segmentation, we implemented the various combinations and calculated the accuracy with SVM and KNN classifiers. The simulations are done using MATLAB R2017b. From Tables 1 and 2, it is found that the max-

Table 1 Accuracy calculated with contouring iteration 250

Pre-processing	Classification	Accuracy in percentage
Image sharpening	SVM	57.14
Contrast stretching	SVM	71.43
Gaussian filtering	SVM	71.43
Homomorphic filtering	SVM	57.14
Image sharpening	KNN	85.71
Contrast stretching	KNN	85.71
Gaussian filtering	KNN	71.4
Homomorphic filtering	KNN	71.4

Table 2 Accuracy calculated with iteration 200

Pre-processing	Classification	Accuracy in percentage
Image sharpening	SVM	42.87
Contrast stretching	SVM	57.14
Gaussian filtering	SVM	28.57
Homomorphic filtering	SVM	42.85
Image sharpening	KNN	71.4
Contrast stretching	KNN	71.4
Gaussian filtering	KNN	57.14
Homomorphic filtering	KNN	71.4

imum accuracy is obtained for KNN classifier when contrast stretching or image sharpening is used as the pre-processing technique.

The iteration level of active contouring is varied, and the accuracies are analyzed. It is interested to find from the analysis table, Table 3 that for image pre-processing using the Gaussian filtering alone, the accuracy reaches to 85.71% when the iteration level is increased to 350.

Increasing the iteration level consumes comparatively more time, and hence, the better combination of algorithms can be preferred as contrast stretching or image sharpening, active contouring segmentation with iteration count of 250 and KNN classifier.

Table 3 Accuracy calculated with iteration 350

Pre-processing	Classification	Accuracy in percentage
Image sharpening	SVM	57.14
Contrast stretching	SVM	71.4
Gaussian filtering	SVM	42.85
Homomorphic filtering	SVM	42.85
Image sharpening	KNN	85.71
Contrast stretching	KNN	85.71
Gaussian filtering	KNN	85.71
Homomorphic filtering	KNN	71.4

6 Conclusions

Various image pre-processing techniques are compared for effectively segmenting the brain tumor from MRI images using active contouring without edge method. Classification of tumors into two classes, benign and malignant, is done using linear SVM and KNN classifier algorithms. It is analyzed that KNN classifier is better than linear kernel SVM, irrespective of any of the pre-processing algorithms tried. Both image sharpening and contrast stretching proved to be the better pre-processing techniques for either KNN or SVM classifier when the segmentation technique we applied is active contouring without edge method. Based on the simulation results using MATLAB R2017b, with various pre-processing and classifier combinations, the highest accuracy is obtained as 85.71% for the KNN classifier with image sharpening or contrast stretching algorithm. For practical implementation, contrast stretching can be identified since it helps to specify the region stretch limit in the original image. Initially, the contouring algorithm code is written and executed for 250 iterations and observed the above-mentioned highest accuracy of 85.71% with KNN classifier. Then, we tried increasing and decreasing the iteration value for analysis purpose. It is observed that reducing the iteration level, say 200, decreases the accuracy. For Gaussian filtering when used with iteration of 350, the KNN accuracy is increased to 85.7%. However, the accuracy stabilizes with a particular number of iterations for other pre-processing algorithms.

References

1. Park C, Took CC, Seong J-K (2018) Machine learning in biomedical engineering. Biomed Eng Lett 8(1):1–3
2. Jhalwa N, Shah P, Sutar R (2018) A hybrid approach for MRI based statistical feature extraction to detect brain tumor index terms-brain tumor, magnetic resonance imaging (MRI), feature extraction. IOSR J VLSI Signal Process. (IOSR-JVSP) 8(2):2319–4197
3. Huma Taha H, Sufyan Ahmed S, Rasheed H (2015) Tumor detection through image processing using MRI. Int J Sci Eng Res 6(2):1692–1695
4. Gopakumar S, Sruthi K, Krishnamoorthy S (2018) Modified level-set for segmenting breast tumor from thermal images. In: 2018 3rd International conference convergence in technology I2CT 2018, pp 1–5
5. Bhagya T, Krishna A, Kanchana DS, Remya Ajai AS (2019) Analysis of image segmentation algorithms for the effective detection of leukemic cells. In: ICOEI
6. Ramanathan R, Thaneshwaran L, Viknesh V, Arunkumar T, Yuvaraj P, Soman KP (2009) A novel technique for english font recognition using support vector machines. In: ARTCom 2009—international conference on advances in recent technologies in communication and computing, vol 1, issue no. 1, pp 766–769
7. Alsadoon A, Devkota B, Elchouemi A, Prasad PWC, Singh AK (2018) Image segmentation for early stage brain tumor detection using mathematical morphological reconstruction. Procedia Comput Sci
8. Gawande SS, Mendre V (2018, January) Brain tumor diagnosis using image processing: a survey. In: RTEICT 2017—2nd ieee international conference on recent trends in electronics, information & communication technologygy, vol 2018, pp 466–470

9. Moitra D, Mandal R (2017) Review of brain tumor detection using pattern recognition techniques. Int J Comput Sci Eng 3(2):121–123
10. Jadhav PS, Bakuli M (2015, March) Brain tumor detection using MRI: a review of literature. Int J Comput Appl 4:82–87
11. Chan TF, Vese LA (2017) Active contours without edges. IEEE Trans Image Process 10(2):266–277
12. Demetri T, Kass M, Witkin A (1988) Snakes: active contour models. Int J Comput Vis 331:321–331
13. Xie X (2010) Textured image segmentation using active contours. Commun Comput Inf Sci (CCIS) 68:357–369
14. Jeyavathana RB, Balasubramanian R, Pandian AA (2016, July) A survey : analysis on preprocessing and segmentation techniques for medical images. Int J Res Sci Innov III:2321–2705
15. Archana JN, Aishwarya P (2016, August) A review on the image sharpening algorithms using unsharp masking. Int J Eng Sci Comput
16. Al-amri SS, Kalyankar NV, Khamitkar SD (2010) Linear and non-linear contrast enhancement image. Int J Comput Sci Netw Secur 10(2):139–143
17. Flusser J, Farokhi S, Höschl C, Suk T, Zitová B, Pedone M (2016) Recognition of images degraded by Gaussian blur. IEEE Trans Image Process 25(2):790–806
18. Satapathy SC, Biswal BN, Udgata SK, Mandal JK (2015) Proceedings of the 3rd international conference on frontiers of intelligent computing: theory and applications (FICTA) 2014: volume 2. Adv Intell Syst Comput 328:423–430
19. Raikwal JS (2012, July) Performance evaluation of SVM and K-nearest neighbor algorithm over medical data set. Int J Comput Appl 50(14)
20. Guo G, Wang H, Bell D, Bi Y, Greer K (2003) KNN model-based approach in classification. In: Meersman R, Tari Z, Schmidt DC (eds) On the move to meaningful internet systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture notes in computer science, vol 2888

CB-CPW Fed SRR Loaded Miniaturized U-Slot Planar Antenna for ECG Monitoring

K. Sajith, T. Shanmuganantham, and D. Sindhanaiselvi

Abstract paper describes the conductor backed coplanar waveguide (CB-CPW) fed with split ring resonator (SRR) biological planar U-slot antenna for ECG monitoring applications. The antenna is designed through bio-tissue layers like skin, fat, and muscle layers. The designed antenna covers the frequency bands of the industrial scientific medical band (ISM band) 2.36–2.48 GHz and fifth-generation medical service band (5G MSB) 3.4–3.6 GHz. The lowest value return loss is 18 dB at 2.41 GHz and 41 dB at 3.6 GHz for U-slot SRR loaded antenna. The reflection coefficient, voltage standing wave ratio; far-field plots are presented.

Keywords ECG · SRR · ISM band 5G MSB · CB-CPW · Radiation pattern

1 Introduction

The novel SRR loaded conductor backed CPW fed planar antenna technology is be of assistance to the field of healthcare diagnosis, monitor, and treatment. The SRR loaded techniques make possible the miniaturized, and low powered biological printed circuit patch for ECG sensing. The main drawback of conventional microstrip technology is higher actuation and dispersion. The novel conductor backed CPW fed enable low actuation and dispersion character. Now, the skin, fat, and muscle layer are configured under the conductor backed with SRR loaded building block as shown in Fig. 3. The conductor backed CPW structure is similar to the microstrip-fed structure,

K. Sajith (✉) · T. Shanmuganantham
Department of Electronics Engineering, Pondicherry University, Pondicherry 605014, India
e-mail: Sajithrajan999@gmail.com

T. Shanmuganantham
e-mail: shanmuga.dee@pondiuni.edu.in

D. Sindhanaiselvi
Department of Electronics and Instrumentation Engineering, Pondicherry Engineering College, Pondicherry 605014, India
e-mail: sindhanaiselvi@pec.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_2

but CB-CPW fed technique has more gain, bandwidth, and wide range of operating frequencies, and easy to fabricate in printed board circuit [1].

The ideas of anisotropic metamaterial property were first proposed by Sir Jagadesh Chandra Bose in 1868. Afterward, the mathematical explanation did by Victor Veselago in 1968. Hence the name of metamaterials is also called Veselago's material [2]. The main metamaterials structure is either circular or rectangular shaped split ring resonator. The circular-shaped split ring resonator had the circular orientation of magnetic dipole moment. The main possibility of metamaterials is a negative refractive index or negative electric and magnetic constant values [3]. The split ring resonator loaded structure enabled negative magnetic permeability but the complementary split ring resonator is shown the negative electric permittivity property [4]. The categorization of electromagnetic materials is mentioned in Fig. 1.

The medical network interfaced between the ISM band and fifth-generation medical service band shown in Fig. 2. The ISM frequency range covers from 2.36 GHz to 2.48 GHz; the fifth-generation medical service band covers from 3.3 GHz to 3.7 GHz [5]. The most case the on-body antenna configured on the skin layer of human tissue. The main function of the proposed biological printed circuit patch antenna is the renovation of the electrical nature of biological signals into the short-range electromagnetic signal. This sensed electromagnetic data interfaced between personal servers. This is short-range communication lengthened by the access points and repeater elements. The access points send data to the internetwork. This highly secure IP address medical packets receive by the medical center through wireless types of equipment like a gateway or network modems or switch [6].

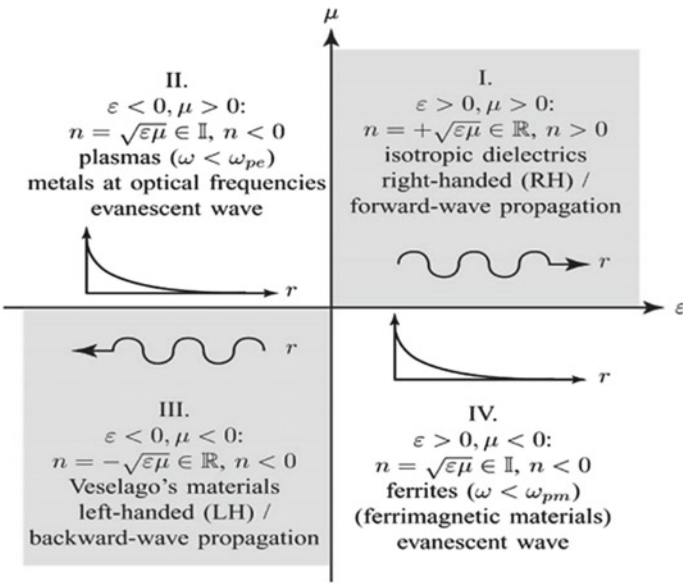


Fig. 1 The classification of metamaterial [9]

Fig. 2 Architecture of on-body communication network [5]

The paper part 1 is the preface and study of existing conventional metamaterials loaded antennas. Part 2 presents explanations of the structural sight of the proposed metamaterials and non-metamaterials loaded structure. Part 3 gives details of the simulation results of the proposed antenna parameter. The conclusion and relevance of proposed SRR loaded antenna illuminate in part 4.

2 Proposed Antenna Models

The proposed configurations consist of two prime elements. The initial parts consist of grounded coplanar waveguide fed with circular split ring resonator loaded radiating patch. After the succeeding part is to be composed of three biological surface layers are skin, fat and muscle layers, are called a medical portion of the proposed on-body device. Moreover, this design is very essential for ECG monitoring [7]. The top views of antennas are shown in Fig. 3 and bottom view of SRR loaded stricture in Fig. 5. Figure 4 mentioned on the whole sight of the anticipated antenna element.

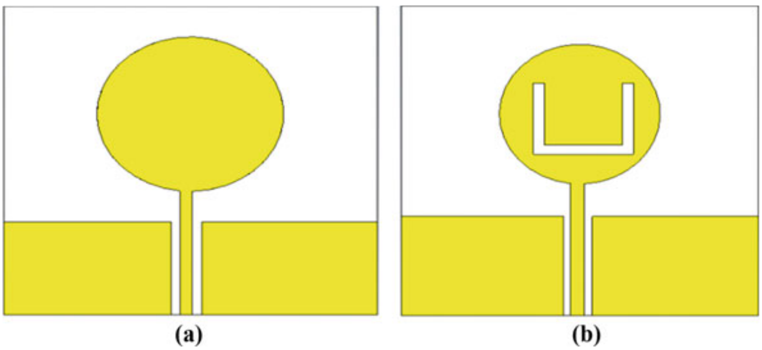


Fig. 3 Top view **a** circular shape, and **b** U-slot antenna

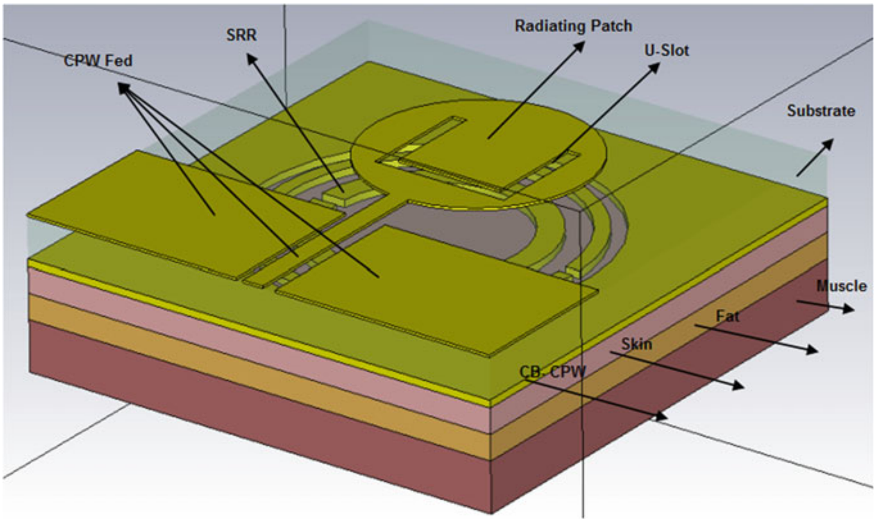


Fig. 4 The overall structure of the proposed prototype

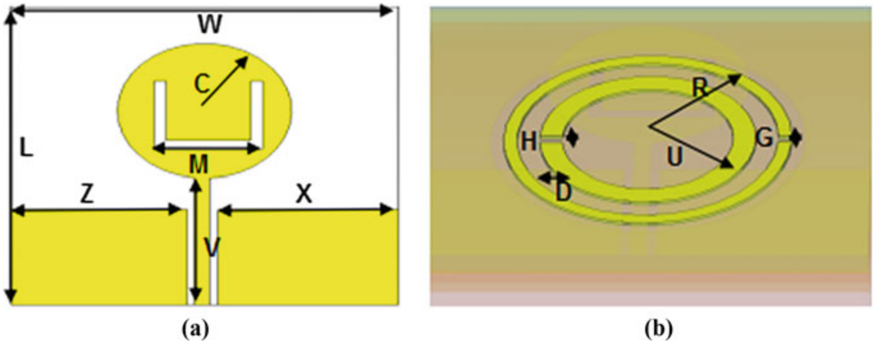


Fig. 5 **a** Dimensions of proposed antenna, **b** split ring resonator

Figures 4 and 5 declare the dimensions of proposed antennas. Table 1 points toward the dimensional parameters and Table 2 indicated the comparison with on-hand and proposed antennas.

3 Simulation Results and Discussion

The SRR loaded antenna parameter like return loss; voltage standing wave ratio, far filed gain and directivity are talk about in beneath. The other antenna parameters like efficiency, bandwidth, and beam width obtain from the above parameter. The

Table 1 Parameter and observation length

Parameter	Observation (mm)	Parameter	Observation (mm)
W	20	C	4.7
L	20	R	6
Z	9.2	U	4.8
V	8.2	D	1.2
M	3.4	H	0.5
X	9.2	G	0.5

Table 2 The compare between the propose prototype with conventional medical patch antenna

Referred papers	Size (mm^3)	Gain (dB)	BW (MHz)
[10]	1524	−16	12
[11]	1265	−25	142
[12]	790	−27	120
Circular shape $f = 2.45$ GHz	640	−10	128
Circular shape $f = 3.6$ GHz	640	−4	112
U-slot without SRR $f = 2.4$ GHz	640	−10	152
U-slot without SRR $f = 3.7$ GHz	640	−4	130
U-slot with SRR $f = 2.41$ GHz	640	−16	146
U-slot with SRR $f = 3.6$ GHz	640	−7	160

simulation result shows the return loss in terms of the reflection coefficient. Decibel scale of the reflection coefficient is negative value but return loss is always a positive value [8].

3.1 Reflection Coefficients

The proposed structure operated an industrial scientific and medical band and fifth-generation medical service band [5]. It is a dual-band antenna, and bandwidth measured from the reflection coefficient plot. The few reflection coefficient values are −28 dB at 2.45 GHz, −24 dB at 3.7 GHz for a circular patch. The −38 dB at 2.4 GHz and −24 dB at 3.68 GHz for U-slot antenna without SRR loaded technique. −18 dB at 2.41 GHz and −41 dB at 3.6 GHz for U-slot SRR loaded antennas are showing in Figs. 6, 7, and 8, respectively.

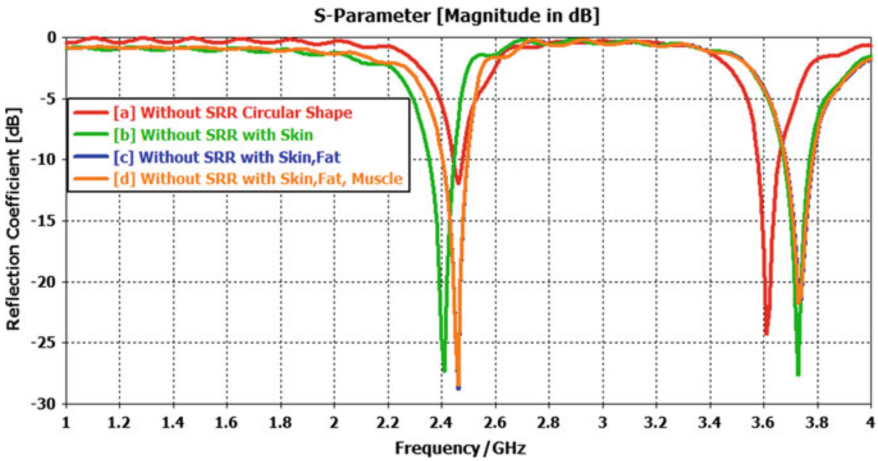


Fig. 6 Reflection coefficient plot for circular patch shape antenna

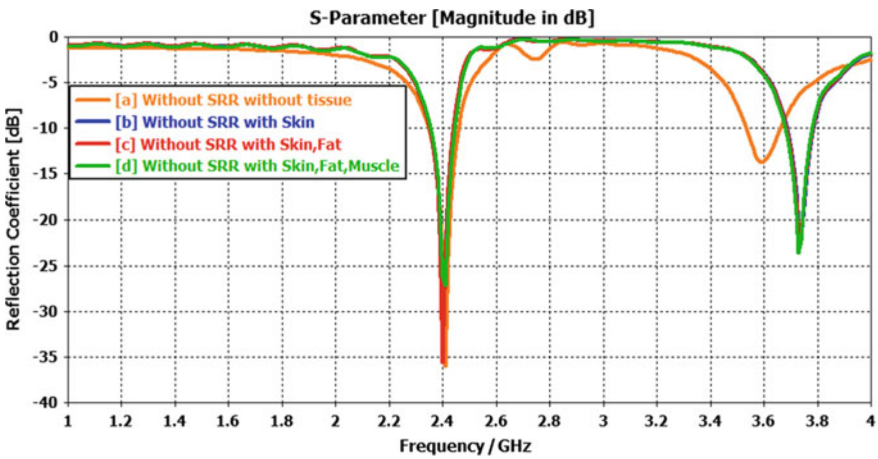


Fig. 7 Reflection coefficient plot for U-slot antenna without SRR

3.2 Voltage Standing Wave Ratio

The standing wave ratio indicated impedance matching property of the designed prototype. These values of designed antennas are 1.03 at 2.45 GHz and 1.11 at 3.7 GHz for circular shape without U-slot shown in Fig. 9. For U-slot without SRR loaded antenna 1.001 at 2.4 GHz and 1.13 at 3.7 GHz shown in Fig. 10. For U-slot SRR loaded antenna 1.3 at 2.41 GHz and 1.001 at 3.6 GHz shown in Fig. 11.

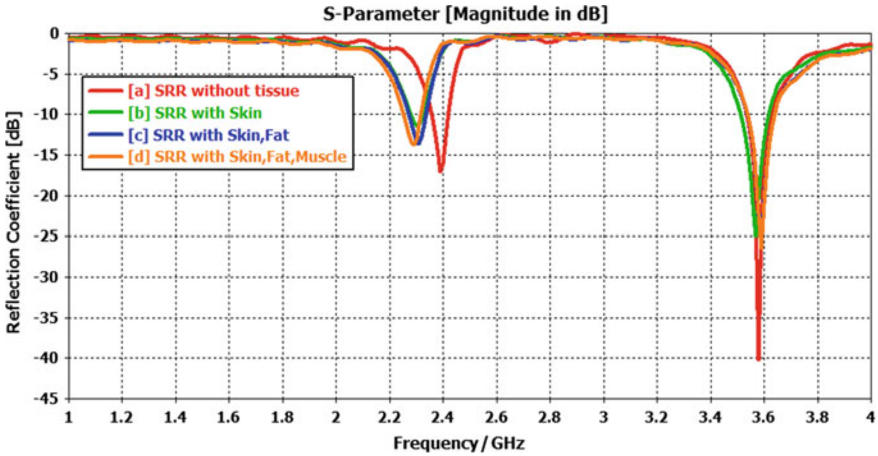


Fig. 8 Reflection coefficient plot for U-slot with SRR loaded planar antenna

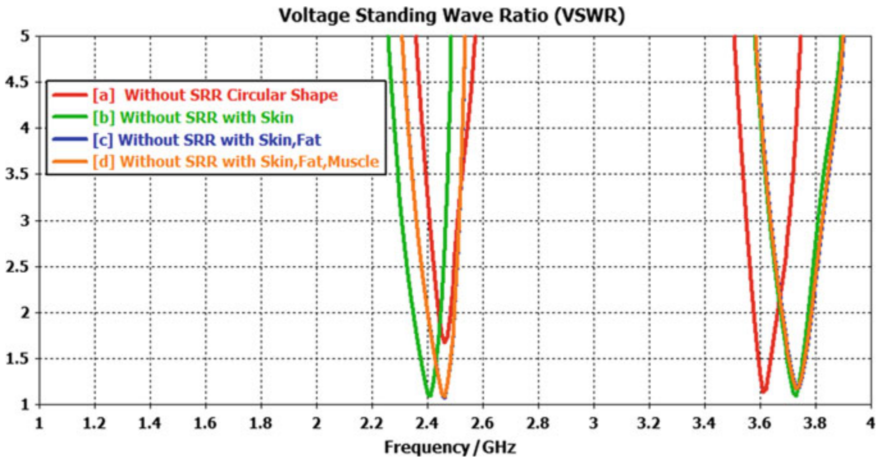


Fig. 9 VSWR plot for CB-CPW fed circular patch

3.3 Directivity and Gain

The given gain and directivity pattern are analogous to the monopole printed circuit antenna pattern, called omnidirectional radiation pattern [6]. The directivity and gain values are 2.9 dBi and −16 dB at 2.41 GHz and 4 dBi and −3 dB at 3.5 GHz are shown in Figs. 12 and 13 respectively.

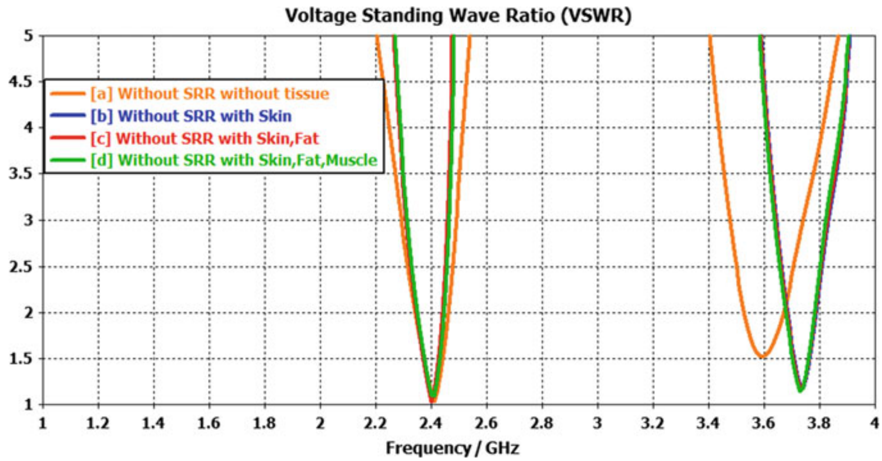


Fig. 10 VSWR plot for CB-CPW fed U-slot patch without SRR

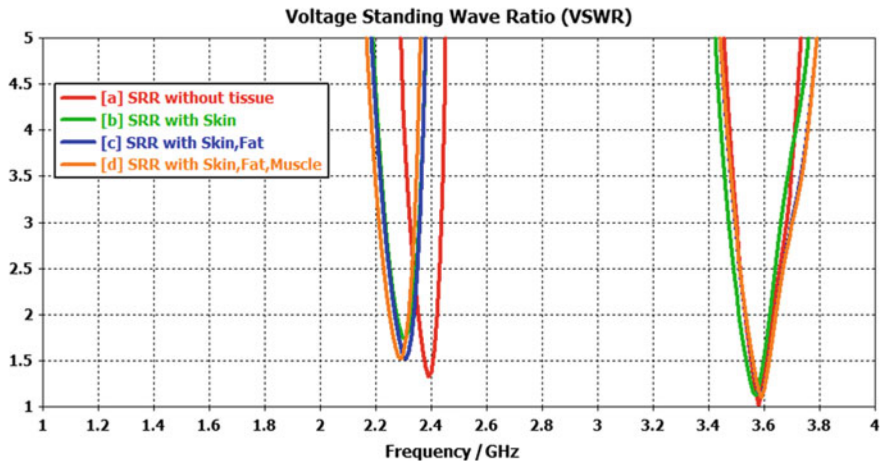
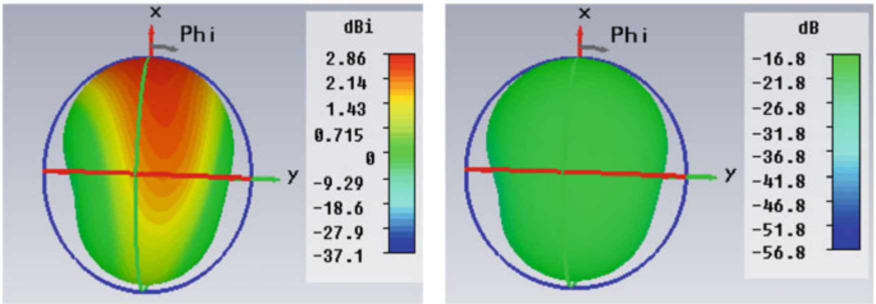


Fig. 11 VSWR plot for CB-CPW fed U-slot patch with SRR

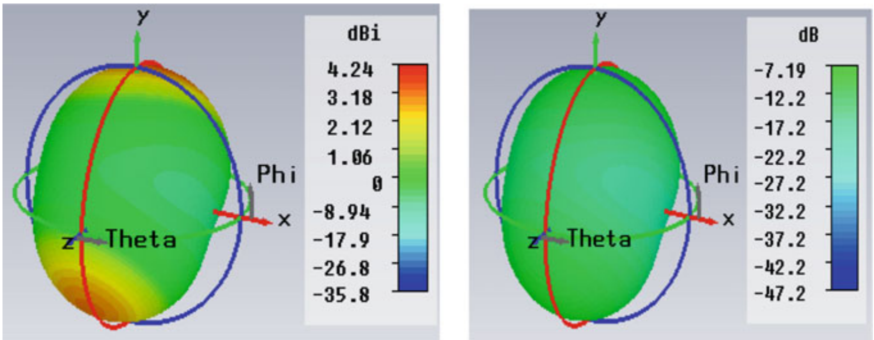
4 Conclusion

The conductor backed CPW fed SRR loaded dual-band antenna designed. The antenna first resonance in ISM band and second is fifth-generation medical service band. The total size of the proposed prototype is 640 mm^3. Furthermore, the proposed prototype model is metamaterials or SRR loaded with conductor backed CPW fed element so that it has a lower value of attenuation and dispersion concerning the conventional antennas. This prototype model configured through skin, fat, and muscle layers, for ECG monitor and other healthcare applications.



(a) (b)

Fig. 12 **a** Directivity and **b** gain plot, at 2.41 GHz U-slot patch with SRR



(a) (b)

Fig. 13 **a** Directivity and **b** gain plot, at 3.6 GHz U-slot patch with SRR

References

1. Sajith K, Shanmuganantham T (2018, April) SRR loaded CB-CPW fed antenna for wearable WBAN and interfacing applications. Int J Sci Res Sci Eng Technol (IJSRSET) 5:80–85
2. Sajith K, Gandhimohan J, Shanmuganantham T (2017, December) SRR loaded and CB-CPW fed octagonal shape antenna for wearable applications. In: IEEE international conference on circuits and systems (ICCS)
3. Sajith K, Gandhimohan J, Shanmuganantham T (2017, December) A novel SRR loaded asymmetrical CPW fed ISM band wearable antenna for health monitoring applications. In: IEEE international conference on AEMC
4. Sajith K, Gandhimohan J, Shanmuganantham T (2017, December) Design of SRR loaded octagonal slot CPW fed wearable antenna for EEG monitoring application. In: IEEE international conference on ICCS
5. Sajith K, Gandhimohan J, Shanmuganantham T (2019, May) CBCPW feed SRR loaded diamond shaped slotted radiator for on-body applications. In: IEEE international conference on IMiCPW
6. Sajith K, Gandhimohan J, Shanmuganantham T (2017, November) Design of SRR loaded asymmetrical CPW fed wearable antenna for telemetry applications. In: IEEE international

conference on antenna innovations and modern technologies for ground, aircraft and satellite application (iAIM)
7. Sajith K, Gandhimohan J, Shanmuganantham T (2017) A novel SRR loaded CPW-fed antenna for implantable ISM band applications. IEEE conference on Indian Antenna Week 2017
8. Balanis CA (2016) Antenna theory analysis and design, 4th edn. Wiley
9. Caloz C, Itoh T (2006) Electromagnetic metamaterials transmission line theory and microwave applications. Wiley
10. Fernandez S, Quvendo T (2010, April) Dual band microstrip patch antenna based on short circuital ring and spiral resonator for implantable medical device. IEEE Trans Microwave Antenna 1048–1055
11. Karacolak T, Hood AZ (2008, September) Design of dual band implantable antenna and development of skin mimicking gels for continuous glucose monitoring. IEEE Trans Microwave Theor Tech 1001–1008
12. Lee CM, Luo CH (2009, January) Band width enhancement of planar inverted F antenna for implantable biotelemetry. Microwave Opt Technol Lett 749

An Improved Resource Allocation Scheme for Heterogeneous Macro–Femto Networks

Aarathi Sankar and Jaison Jacob

Abstract The increasing demand for varied wireless applications urges the development of an efficient resource allocation mechanism that provides required rate satisfaction and interference mitigation. Femtocell deployment provides capacity and coverage improvement and service quality to indoor users. However, overlay deployment of femtocells makes resource and interference management a bit challenging. Thus, a review on the resource allocation strategies for heterogeneous macro–femto networks is conducted. A self-organized resource allocation (SO-RA) scheme was implemented to mitigate co-layer interference in the downlink transmission for a macro–femto networks. An analysis on the SO-RA scheme was conducted, considering the cross-layer interference between macrocell and femtocell. A modification to the system was proposed to mitigate the cross-layer interference by adjusting the transmit power of femto base station. The resulting modified scheme and the SO-RA scheme were compared with existing schemes like reuse-1 and simple macro network for various system parameters. The modified SO-RA scheme achieves the best performance among all the four schemes. It results in improved throughput, fairness performance, and interference mitigation compared with the SO-RA scheme.

Keywords Resource allocation · Femtocell · Macrocell · Co-layer interference · Cross-layer interference

1 Introduction

Nowadays, above 50% of voice services and 70% of information traffics originate from indoor wireless users. Despite the increasing demand in macrocell base stations, users still often suffer from quality of service degradation due to huge penetration

A. Sankar (✉) · J. Jacob
Department of Electronics and Communication, Rjagiri School of Engineering and Technology,Ernakulam, India
e-mail: aarathisankar123@gmail.com

J. Jacob
e-mail: jaison_jacob@rajagiritech.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_3

loss and indoor attenuation. Moreover, the dead zones created by macrocells due to insufficient penetration of macro signal create problems. Femto deployment is a promising methodology to solve this problem. It helps to meet high data rate requirements, provides improved capacity and coverage, and mitigates interference to a great extend [1].

Femto base stations are customer-owned, short-range, low cost and low power base stations (BSs), operating in the licensed spectrum. The femtocell deployment is a reasonable solution that satisfies rate requirement and provides interference mitigation. Femtocells on the macrocell network provide a cost-effective transmission of high-speed multimedia data and reduce the overall system costs. It benefits both users and operators. The vicinity between transmitter and receiver provides better signal quality resulting in communications with larger reliabilities and throughputs [4].

The overlay deployment of femtocells makes radio resource management (RRM) somewhat challenging because of the fact that macrocells and femtocells share the accessible radio resources. This give rise to cross-layer interference between macrocell and femtocell as well as co-layer interference between neighboring femtocells. The coordination between neighboring femtocells is fundamental to adapt an efficient resource allocation strategy. With this motive, we implement a SO-RA scheme, which helps in mitigating co-layer interference in a heterogeneous macro–femto networks. In addition to co-layer interference mitigation, the scheme also ensures fairness to users and achieves feasible trade-off between throughput and resource reuse [1].

We propose a modification to the system by analyzing the impact of cross-layer interference on the global system performance. The system shows perform badly with regard to system parameters by considering cross-layer interference. So, we adopt a method to mitigate this interference by adjusting the power of transmitting femto base station. The resulting system was compared with the implemented SO-RA scheme, macro network, and the reuse-1 scheme.

The outline of this paper is: Sect. 2 shows some of the schemes available for interference management and resource allocation. Section 3 deals with the system model of our implemented SO-RA algorithm and the modified system. Section 4 discusses the simulation results and discussions. Section 5 shows the final conclusions regarding the four systems.

2 Related Work

Several researches have been done to analyze the interference management and resource allocation in a macro–femto network. Some of these are given in [4] and [5]. The interference management in a macro–femto networks ensures two things. First, the macro users and neighboring femto clients are not affected by the interference from femto BS. Second, it achieves rate requirement satisfaction to femto users by ensuring that transmit intensity of femto base station is sufficiently high. A coordination-based approach is used, which makes decisions for the mitigation of interference and allocation of resources using the exchange of information between femto and macro base stations. Some of these coordination-based approaches for interference management are discussed below.

2.1 Resource Partitioning-Based Methods

This method is commonly known as frequency reuse-1 or essentially reuse-1. It is the simplest of all methods, where all the resources are accessible to all the femtocells, i.e., there is no confinement for their usage. It improves resource utilization efficiency, but results in an increased co-layer interference as femto density increases. The deployment is done using hybrid spectrum allocation, where inner femtocells use orthogonal allocation and outer femtocells use shared allocation.

2.2 Transmit Power Control-Based Methods

This scheme ensures constant femto BS coverage without affecting the macrocell throughput. For this, a power control mechanism is used, given in [10] and [14]. Many adaptive power control algorithms to mitigate cross-layer interference are suggested in [5] and [6].

2.3 Cognition-Based Methods

This scheme performs interference management and resource allocation cognitively. Here, the femto base stations are considered as secondary users and macro base stations as primary. The channels are determined cognitively [12]. For the management of cross-layer interference, the scheme develops autonomous algorithms. The spectrum occupied by macro users decides how efficient and beneficial is the cognitive-based approach.

We implement a SO-RA resource allocation scheme, which ensures feasible trade-off between achievable throughput and resource reuse along with the mitigation of co-layer interference to implicit levels. It also guarantees minimum fairness to the users. A modification to the system is proposed, which mitigates cross-layer interference and results in an enhanced system performance.

3 System Model

The framework comprises of seven macrocells with base station situated at the center of each macrocell. The central macrocell consists of L femtocells which are overlay deployed. The total resources of the system are divided along with frequency as well as time slots. These resources are arranged in units of physical resource blocks (PRBs) with 12 subcarriers. It is assumed that both macrocells and femtocells have N PRBs available. The SO-RA resource allocation scheme and the modification made to the system are discussed in detail below.

3.1 Self-organized Resource Allocation Scheme

The scheme mitigates co-layer interference significantly and accomplishes a general improvement in framework execution by coordination between neighboring femtocells. The algorithm involves two steps. First step is interfering neighbor set discovery and reuse-1 allocation, and second step is coordinated resource drop, which selectively drops the most interfering PRB. The second step involves two stages of a priori checks to analyze the impact of dropping PRB on the global system performance [10]. This is performed for all the femto base stations. The two steps are discussed in detail below.

3.1.1 Interfering Neighbor Set Discovery and Initial Reuse-1 Allocation

Initially, the femto users perform reference signal received power measurement. Using these measurements, femto base station computes pathloss from the neighboring femto base stations, which are then compared with specific threshold value to determine whether they may cause interference or not. Thus, each femto base station calculates a set of neighboring femto base stations that cause significant interference to its femto users. After this, resource allocation is done using reuse-1 scheme [12]. It is preferred because using this scheme signaling overhead reduces to significant amounts.

3.1.2 Coordinated Resource Drop

Here, each femto base station selectively drops the most interfering PRB or the PRB with minimum SINR without affecting the system throughput and overall performance.

Each femto base station calculates the SINR experienced on all PRBs and selects the one on which it encounters least SINR. The SINR for femto user l on PRB n is given by,

$$\mathrm{SINR}_{\mathrm{l}}^{\mathrm{n}} = \frac{\mathrm{P}_{\mathrm{Ftx}}^{\mathrm{l,n}}.\mathrm{PL}_{\mathrm{l,l}}}{\mathrm{I}_{\mathrm{femto}} + \mathrm{I}_{\mathrm{macro}} + \mathrm{N}_{\mathrm{o}}} \tag{1}$$

where $P_{\mathrm{Ftx}}^{l,n}$ denotes transmit power of femto base station l on PRB n, $PL_{l,l}$ is the path loss between femto user l and macro base station l, I_{femto} is the interference from femto base station, I_{macro} is the interference from macro base station, and N_o is the additive white Gaussian noise.

Almost certainly, least SINR is because of high interference on that PRB. Dropping that PRB brings about minimum rate loss, resulting in a decreased co-layer interference. To ensure that the choice of dropping PRB does not adversely influence the overall system performance, femto base station performs two stages of a priori checks before dropping that PRB.

Level-I a priori check: It involves analyzing the performance of serving femtocell upon dropping the PRB. It starts with the calculation of throughput on PRB for the serving femto base station. For femto base station l, the throughput on PRB n is,

$$R_l^n = B\,.\log_2(1 + \mathrm{SINR}_l^n) \tag{2}$$

where B is the bandwidth of PRB.

The total throughput of the base station l is given by,

$$R_l = \sum_{n=1}^{N} B\,.\log_2(1 + \mathrm{SINR}_l^n)\,.\,x_{n,l} \tag{3}$$

The new throughput value is calculated by taking the difference of Eqs. 2 and 3.

$$R_{l,\mathrm{new}} = R_l - R_l^n \tag{4}$$

The new throughput value is compared with the specified threshold, which will be already set as some 7 Mbps. It is done to ensure that the new throughput, after dropping the PRB is within the prescribed limits. It will give us the decision for dropping PRB. If it is less than the threshold value, we defers the decision of dropping that PRB and we move to the next femto cell and repeats the same procedure. Otherwise, the resulting condition favors the decision of dropping PRB. Then we move to the second level of a priori check [1].

Level-II a-priori check: It involves analyzing the performance of neighboring femtocells upon dropping the PRB. The serving femto base station asks its neighboring base stations to calculate and report their gain in throughputs after dropping the PRB. The new SINR and throughput values calculated by base station m are given by Eqs. 5 and 6.

$$\mathrm{SINR}_{m,\mathrm{new}}^{n} = \frac{P_{\mathrm{Ftx}}^{m,n}.PL_{l,l}}{I_{\mathrm{femto}} - P_{\mathrm{Ftx}}^{l,n}PL_{m,l} + I_{\mathrm{macro}} + N_o} \tag{5}$$

where $P_{\text{Ftx}}^{m,n}$ denotes transmit power of femto base station m on PRB n, $PL_{l,l}$ is the path loss between femto user l and macro base station l, I_{femto} is the interference from femto base station, $P_{\text{Ftx}}^{l,n}$ is the transmit power of femto base station l on n, $PL_{m,l}$ is the path loss between femto user m and macro base station l, I_{macro} is the interference from macro base station, and N_o is the additive white Gaussian noise.

$$R_{m,\text{new}} = \sum_{n=1}^{N} B \,.\, \log_2(1 + \text{SINR}_{m,\text{new}}^{n}) \,.\, x_{n,l} \tag{6}$$

The gain in throughput for femto BS m is given by,

$$\Delta R_m = \sum_{m} (R_{m,\text{new}} - R_m) \tag{7}$$

The serving base station calculates the total throughput gain of its neighbors, which is given by,

$$\Delta R = \sum_{m} \Delta R_m \tag{8}$$

The gain in throughput is compared with the loss in throughput because of dropping PRB. If there is an increase in throughput after dropping PRB, the femto base station makes a decision to drop that PRB, otherwise move to the next base station [1].

3.2 Modified Self-organized Resource Allocation Scheme

The SO-RA algorithm takes in to account the co-layer interference between neighboring femtocells. The whole system was constructed without considering the cross-layer interference from macrocell. So, we examine the effect of this cross-layer interference on the performance of the system. The performance was found to be poor for the analyzed system with regard to user throughput, SINR, and fairness to users because of the presence of cross-layer interference. Our aim is to reduce this interference to improve the performance. So, we adopted a method which meets this requirement by adjusting the transmit power of femto base station using the equation,

$$P_F = L_W R_f^{\alpha_f} P_M d_f^{-\alpha_f} \tag{9}$$

where P_F shows the adjusted transmit power of femto station, L_W represents the wall loss, R_f shows the distance of femto BS from the outer wall of the building, P_M is the transmit power of macro BS, d_f is the distance between macro and femto BSs, and α_f and α_f represent the path loss exponents of femto and macro BSs, respectively.

Table 1 Simulation parameters

Parameter	Value
Inter site distance	500 m
Total bandwidth	10 MHz with 50 PRBs
PRB bandwidth	180 KHz
Macro users per macrocell	10
Number of femtocell blocks	3
Active ratio ρ	0.1–1
Macro BS TX power	46 dBm
Femto BS TX power	20 dBm
Pathloss threshold	30 dB
Femto BS throughput threshold	7 Mbps
Averaging interval	1000 ms

The resulting modified system showed an improvement in system parameters and an enhanced system performance compared to that of our original SO-RA implemented system.

4 Results and Discussions

The system-level simulations were done in MATLAB. The overall process can be separated in to two main steps: femto deployment and resource allocation. Initially, a 4×4 array of femto deployed network was created over 10 km distance with one user per cell. So there was a total of 16 femtocells, each one having single user. Then, a 4×4 array of macro network was formed with one central base station and 16 users for the purpose of comparison. The resource allocation was done using reuse-1 scheme where all the resources are available to all the femtocells. Then, the SO-RA was implemented using two levels of a priori checks and the system parameters were calculated. A modification to the scheme was proposed by adjusting the power of transmitting femto BS, and the new values of system parameters were calculated.

In this paper, we compare the modified SO-RA algorithm with our implemented SO-RA algorithm, macro network, and the reuse-1 scheme. The parameters and the conditions used for simulations are given in Table 1. The simulation results of various parameters showing the comparison are given below, showing the superiority of our scheme compared with already existing other schemes.

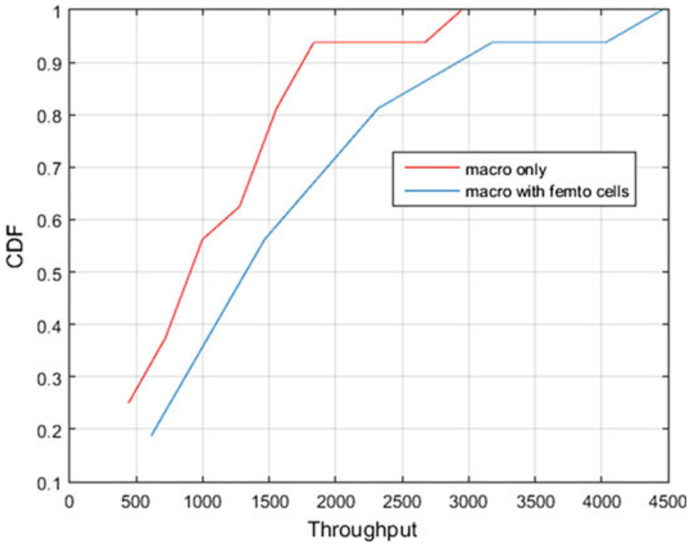


Fig. 1 Performance comparison with and without femtocell

4.1 Comparison of User Throughput

The user throughput for both macro and femto networks was calculated using Eq. 2. We review the performance by taking the CDF of throughput for both networks, and the results were plotted. The results indicate that user throughput was found to be higher with the deployment of femtocells in the network. This is due to the mitigation of interference to significant amounts by a femto network compared to macro. Consequently, SINR increases and so is the throughput. The performance examination of throughput with and without femtocells is shown in Fig. 1.

4.2 Effect of Active Ratio on Throughput

The effect of active ratio on user throughput was calculated for both femto and macro networks. It was observed that both users experience a reduction in throughput as femtocell density increases. This is because both femto and macro users experience an increase in co-layer and cross-layer interference, respectively, as number of active femto users increases. This result is shown in Fig. 2.

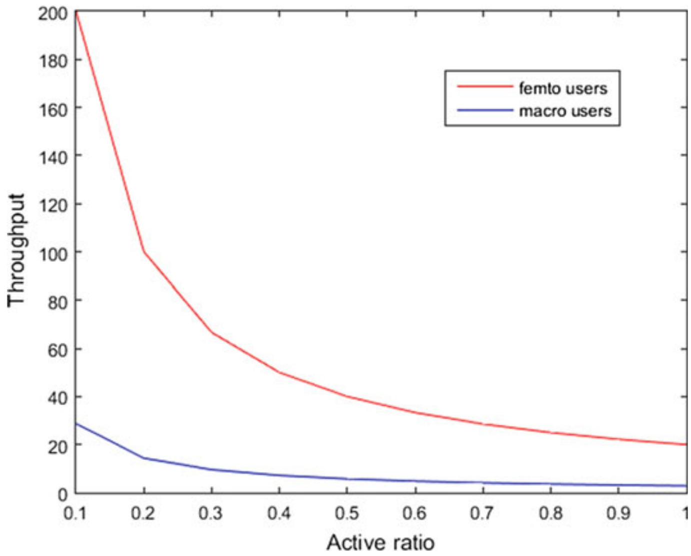


Fig. 2 Effect of active ratio on throughput

4.3 Comparison of User Throughput and SINR

The SO-RA algorithm was implemented using two levels of a priori checks, and its impact on the overall system performance was evaluated. An analysis on the system performance with respect to SINR and user throughput was made, considering cross-layer interference from macrocell. The performance was found to be poor for the analyzed system considering cross-layer interference, compared with our original SO-RA system. This is because interference plays a significant role in SINR and throughput calculations.

So, we adopted a method to mitigate this cross-layer interference. This is done by adjusting the transmit power of femto base station using Eq. 9. The SINR and throughput values were calculated using the adjusted power of femto station. It was observed that the modified framework shows better execution regarding SINR and throughput compared with other schemes. These results are shown in Fig. 3 and Fig. 4. The increase in throughput is because the scheme ensures that net gain in throughput is always greater than loss in throughput due to dropping PRB. Also, the resource utilization in SO-RA scheme is nearly 83.7%, which is twice that of AFR scheme.

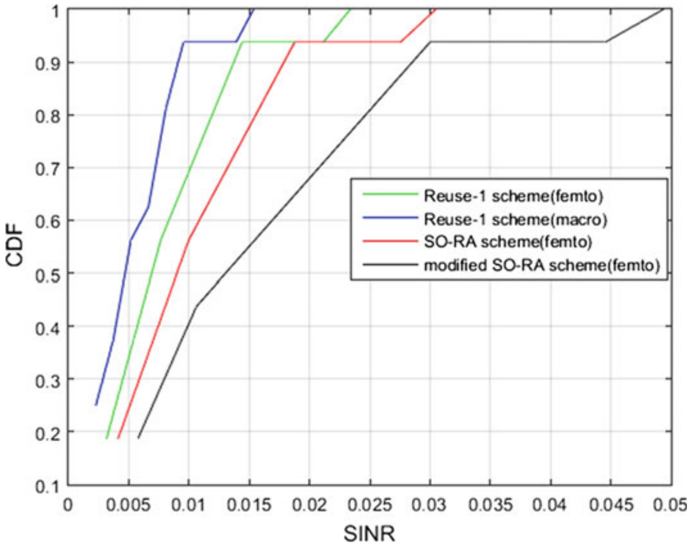


Fig. 3 SINR for femtocell users

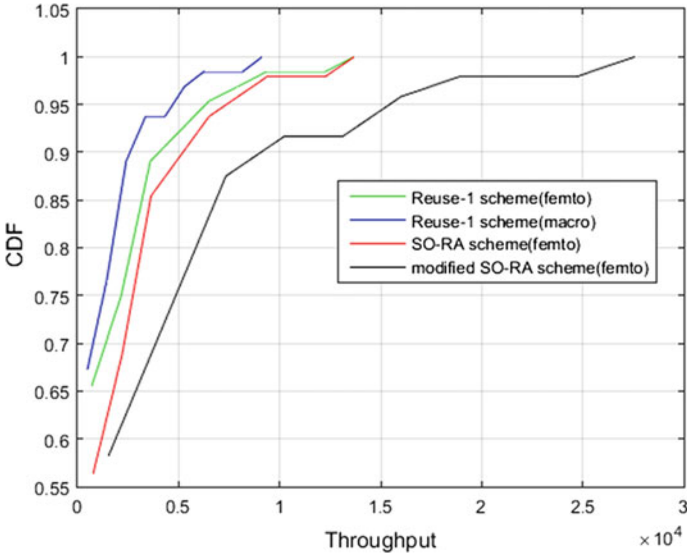


Fig. 4 Femtocell user throughput

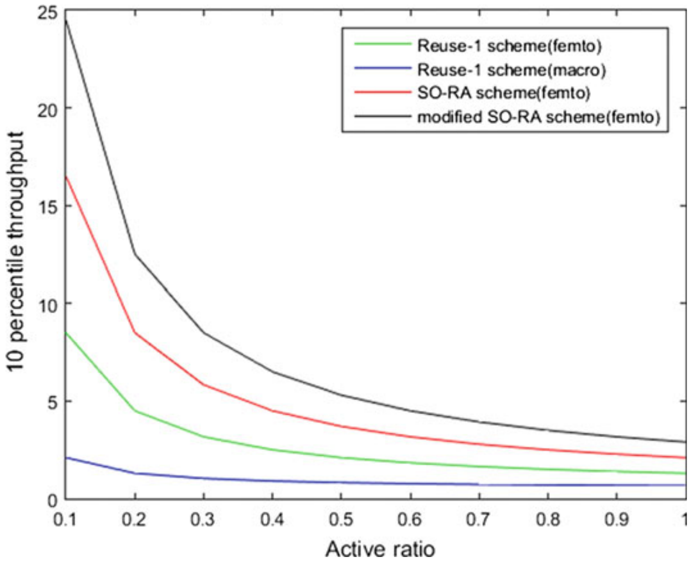


Fig. 5 10 percentile throughput of femto users

4.4 Impact of Active Ratio on 10 Percentile Throughput

Figure 5 shows the variation in 10 percentile throughput value for femto and macro users with active ratio. It was observed that there is a decrease in 10 percentile throughput value corresponding to the increase in active user density for all the four systems. However, compared with macro network and reuse-1, SO-RA scheme provides good 10 percentile throughput regardless of increased active ratio because of co-layer interference mitigation. The best 10 percentile throughput performance was observed for our modified SO-RA system, nearly 30% improvement, because of cross-layer interference mitigation.

4.5 Comparison of Fairness Performance

The fairness performance of the system was investigated using Gini fairness index. It shows the amount of inequality, and the index value lies between 0 and 1. It is computed using the equation,

$$I = \frac{1}{2L^2 R} \sum_{l=1}^{L} \sum_{m=1}^{L} |R_l - R_m| \tag{10}$$

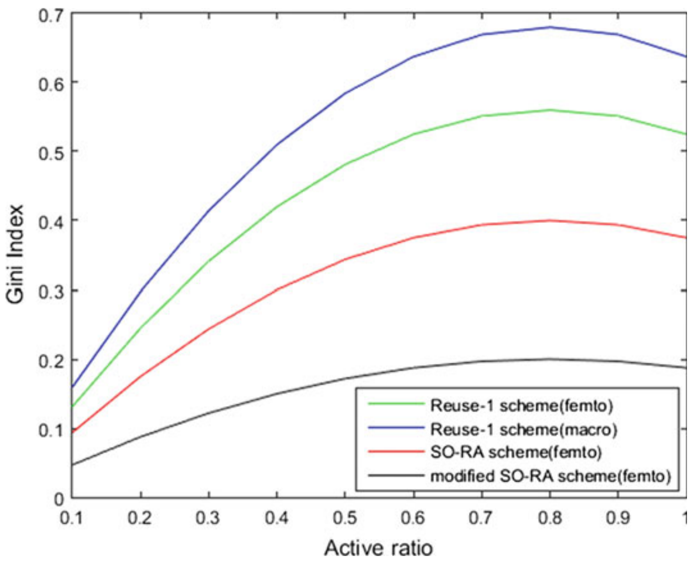


Fig. 6 Gini fairness index performance

where

$$R = \sum_{l=1}^{L} \frac{R_l}{L} \tag{11}$$

The system is perfectly reasonable if its value is 0 and unfair if its value is 1. Figure 6 shows the fairness performance plot which shows that the SO-RA scheme fulfills rate requirement and ensures fairness to femto users compared with macro and reuse-1 schemes. However, the modified SO-RA system outperforms the SO-RA scheme with regard to fairness. This is because the index value was found to be the lowest for the modified SO-RA system.

5 Conclusion

The rapid increase in wireless applications results in a huge proliferation in indoor traffic. Femtocell systems are viewed as a practical arrangement that satisfies the requests of fast voice and information traffic for indoor clients. However, an efficient resource allocation mechanism is required to meet these objectives and mitigates interference. A self-organized resource allocation scheme was implemented which decreases co-layer interference to understood levels and achieves trade-off between attainable throughput and resource utilization efficiency, compared with the competing algorithms. It also ensures fairness to all the users simultaneously by using two stages of a-priori checks and an increase in user throughput. An analysis on the system was conducted, considering the cross-layer interference between femtocell and macrocells. This system showed poor performance with respect to the system parameters. So, a modification to the system was made to mitigate interference between femtocell and macrocell by adjusting the transmit intensity of femto station. The resulting system was compared with reuse-1, macro network, and SO-RA scheme. The modified SO-RA shows better and enhanced system performance in terms of throughput, fairness to users, and interference mitigation compared with other three.

References

1. Mehta M, Rane N, Karandikar A, Imran MA, Evans BG (2014) A self-organized resource allocation scheme for heterogeneous macro-femto networks. Wiley online library , pp 330–342
2. Chandrasekhar V, Andrews JG, Gatherer A (2008) Femtocell networks : a Survey. IEEE Commun Mag 59–67
3. Perez L, Valcarce D, Roche AD, Zhang GJ (2009) OFDMA femtocells: a roadmap on interference avoidance. IEEE Commun Mag 41–48
4. Kim MS, Je HW, Tobagi F (2010) Cross-tier interference mitigation for two-tier OFDMA Femtocell networks with limited Macrocell information. In: IEEE global telecommunications conference (GLOBECOM), p 15
5. Chandrasekhar V, Andrews J (2009) Spectrum allocation in tiered cellular networks. IEEE Trans Commun 3059–3068 (2009)
6. Guvenc I, Jeong MR, Watanabe F, Inamura H (2008) A hybrid frequency assignment for femtocells and coverage area analysis for co-channel operation. IEEE Commun Lett 880–882
7. Jo HS, Mun C, Moon J, Yook JG (2010) Self-optimized coverage coordination in femtocell networks. IEEE Trans Wirel Commun 2977–2982
8. Cao G, Yang D, Ruihong A, Xuan Y, Zheng R, Zhang X (2012) An adaptive sub-band allocation scheme for dense femtocell environment. IEEE Trans Veh Technol 2194–2207
9. Cheng SM, Ao WC, Tseng FM, Chen KC (2011) Design and analysis of downlink spectrum sharing in two-tier cognitive femto networks. In: IEEE wireless communications and networking conference, pp 102–107
10. Ruihong A, Xin Z, Gen C, Ruiming Z, Lin S (2010) Interference avoidance and adaptive fraction frequency reuse in a hierarchical cell structure. IEEE Wirel Commun Netw 1–5
11. Claussen H (2007) Performance of macro- and co-channel femtocells in a hierarchical cell structure. In Personal, indoor, mobile radio communications, pp 1–5
12. Bernardo F, Agusti R, Cordero J, Crespo C (2010) Self optimization of spectrum assignment and transmission power in OFDMA femtocells. In Advanced international conference on telecommunications, pp 404–409
13. Lopez-Perez D, Ladanyi A, Juttner A, Zhang J (2009) OFDMA femtocells: a self-organizing approach for frequency assignment. In IEEE personal, indoor and mobile radio communications, pp 2202–2207
14. Liang YS, Chung WH, Ni GK, Chen IY, Zhang H, Kuo SY (2012) OFDMA femtocells: resource allocation with interference avoidance in OFDMA femtocell Networks. IEEE Trans Veh Technol 2243–2255
15. Mehta M, Aliu OG, Karandikar A, Imran MA (2011) A selforganized resource allocation using inter-cell interference coordination (icic) in relay-assisted cellular networks. ICTACT J Commun Technol 300–313

16. Zhangi H, Jiang D, Li F, Liu K, Song H (2016) Cluster-based resource allocation for spectrum-sharing femtocell networks. IEEE Access
17. Li B, Zhang X, Huang J, Ma C, Qin G (2018) A scheme of multi-domain cooperative cost resource management in outdoor + indoor Het-Net. In: IEEE international conference on communication technology
18. Bharucha Z, Saul A, Auer G, Haas H (2010) Dynamic resource partitioning for downlink femto-to-macrocell interference avoidance. EURASIP J Wirel Commun Netw 77–88

A Robust Music Composer Identification System Based on Cepstral Feature and Models

A. Revathi, D. Vishnu Vashista, Kuppa Sai Sri Teja, and R. Nagakrishnan

Abstract Music is an integral part of everyone's life. The task of recognizing a composer by listening to a musical piece is difficult for people with no knowledge of music. Even such a task is difficult for people in musical theory because every day a new music composer is born. To address this problem of classification of a musical tune for a large group of music composers, we construct an automatic system that can distinguish music composers based on their tunes composed. Classification of music based on its composer is very essential for faster retrieval of music files. In this paper, we make a large database that consists of several musical tunes belonging to several composers. We extract MFCC features and train the system using the clustering technique by developing a classification model that can accurately classify a musical tune that belongs to a music composer whose tunes are trained. The overall accuracy of the proposed system is 72.5% without discrete wavelet transform (DWT) and 77.5% with DWT.

Keywords Music information retrieval · Music composer identification · Discrete wavelet transform

1 Introduction

Modern world has seen a rise in number of musical composers, which is making the task of classifying a music composer difficult for people based on the tune heard.

A. Revathi · D. V. Vashista (✉) · K. S. S. Teja · R. Nagakrishnan
Department of ECE, School of EEE, SASTRA Deemed to be University, Thanjavur, Tamilnadu, India
e-mail: vishnuvasista99@gmail.com

A. Revathi
e-mail: revathi@ece.sastra.edu

K. S. S. Teja
e-mail: kuppasaisriteja@gmail.com

R. Nagakrishnan
e-mail: nagakrishnan@sastra.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_4

So, there is a need for a system that can classify a tune from a known list of music composers. Along with the above-mentioned task, there is also a need for faster retrieval of musical data, which is a reason for switching to this system. The system not only classifies a given tune based on the composers present in the database but also the classification requires very little test data.

The task of music information retrieval dates back to the 1960s. With the advent of computers, computer-assisted musical analysis had started to emerge. With the help of several statistical methods and information theory, the music information extraction has become easy to be processed. The primitive works were based on pitch, notes, chords, harmonics, rhythmic values and frequencies as basic information used as features for music analysis. There are other features that are playing a major role in present-day research of musical analysis like Mel-Frequency Cepstral Coefficient (MFCC), Linear Predictive Coding (LPC) and Zero-Crossing Rate (ZCR). The music composer identification system was developed by many researchers [1–8].

The feedforward neural network-based composer identification [1] has been developed. The Probabilistic Neural Networks can be used to construct the similarity matrix between the composers and analyze the Dodecaphonic Trace Vectors. the gradient and differential evaluation algorithm are used for training. Three convolutional layer-based recurrent network [2] was designed to identify the music composer. This work achieved 70% of accuracy when classifying six different composers. The automatic recognition system [3] was developed for identifying classical piano composers. The cortical algorithms can be used to reduce the large set of features and efficiently identify the composer in supervised manner. The content spectral features and hidden Markov models based [4] composer identification system were analyzed. The accuracy of the system is 5% higher than the standard spectral features-based system. The grammatical interface approach [5] can be used to identify the music composer. The language modelling technique is used [6] to capture the different style of the music composer. The five different types of classification models [7] developed for correctly classify the three different music composers. The weighted Markov chain model [8] was developed to extract the pitch frequency and inter on set intervals from the music for composer's identification.

This paper is systemized as such. In section II discussed the feature extraction techniques and K-means clustering algorithm. In section III discussed the proposed music composer identification system. Section IV presents the performance of the proposed identification system. Section V gives the conclusion of our work.

2 Materials and Methods

2.1 Database

In this work, the database is made for four music composers namely A. R. RAHMAN, ILLAYARAJA, DEVI SRI PRASAD, HANS ZIMMER. 10 sets of tunes composed by each of the above-mentioned music composers are taken. From each tune of a music composer 2500 MFCC vectors with 13 coefficients in each vector are extracted. This leads to formation of 25,000 feature vectors for each music composer which are used for training the system.

2.2 Feature Extraction

For building a system, we have large volumes of data sets which increases the computational time and required hardware for training them into the system. This problem is eliminated using a powerful technique called feature extraction. Features are like the blueprints which are smaller in size compared to that of large datasets without compromising in describing the object. In daily life, we use feature extractions for image processing, speaker identification, biomedical applications like cancer cell detection, etc. For acoustic datasets, largely used feature extraction methods are MFCC's (Mel-Frequency Cepstral Coefficients), Linear Predictive Coefficients (LPC), Linear Predictive Cepstral Coefficients (LPCC) and Relative Spectral-Perceptual Linear Predictive (RASTA-PLP). In this paper, we use MFCCs as the feature extraction technique.

2.2.1 Mel-Frequency Cepstral Coefficient (MFCC)

The first step in the music composer identification system is to extract features, i.e. identify the components of the music signal that best describe the compositions made by the music composer. This is done by extracting MFCCs vectors from the music file. The following gives an insight over-extraction of MFCCs.

Steps Involved in MFCCs Extraction
The following are the steps involved to extract MFCCs from a music file.

1. The signal is divided into number frames with less duration.
2. Periodogram estimation is done for the power spectrum of the above frames.
3. Mel filter banks are then applied to the above periodogram and energies obtained by each filter are summed.
4. The logarithmic scale is applied for the filter bank energies.
5. Further, take the DCT of the log filter bank energies.
6. Extract the first 13 DCT coefficients as a vector, this is a feature vector for a frame.

The block diagram of MFCC extraction is shown in Fig. 1. The reason for dividing the music file into a number of frames is that the signal is kept statistically stationary over the frame length. So small frame lengths of 20–45 ms are chosen for MFCCs extraction. Next, the power spectrum of each frame is calculated and the periodogram estimate gives a frequency identification in the frame which is similar to the function

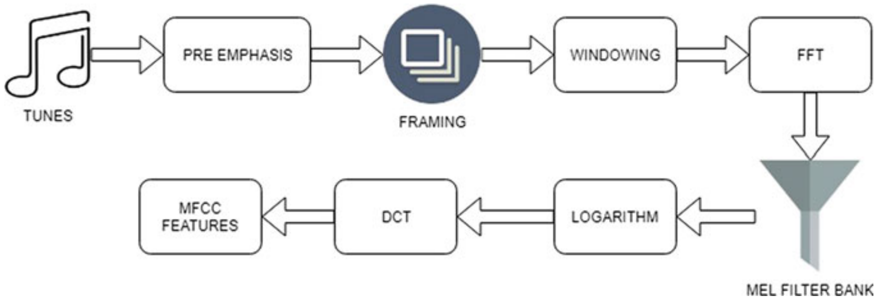


Fig. 1 MFCC feature extraction

of the cochlea (an organ in the ear) which vibrates differently based on the incoming frequencies. Mel filter bank estimates the energy present in the various frequency ranges. The first filter and gives the energy estimation near 0 Hertz. The Mel scale gives the information regarding the filter bank spacing and width later the logarithm of the obtained energies gives cepstral mean subtraction. Finally, DCT of log filter bank energies is calculated which decorrelates the energies and the first 13 coefficient s form the MFCCs vector for that frame. The first coefficient in the vector gives the frame energy.

2.3 K-Means Clustering Algorithm

K-means is a simple unsupervised learning algorithm that is used for clustering data, hence the name clustering. The algorithm follows a simple way of classifying a given data set into a certain number of clusters (assume k clusters) that are fixed a priori. The centers need to be placed in a clever manner as different placing would lead to different results. Generally, these centers are preferred to be placed far from each other. Furthermore, each point of the given data set is associated with its nearest center. After assigning each point to its nearest center the stage1 is complete. Now, K-new centroids have to be calculated as the barycenter of the clusters that are obtained from stage1. Using these new centroids, a new group of clusters has to be formed using the same data set points and the nearest centers. This leads to stage2. The above process has to be repeated for N-stages. After every iteration, we get K-new centers. But after few iterations, the centroids remain unchanged. This is the point where the algorithm has to be terminated [9]. The algorithm aims at minimizing the squared error function given by in (1)

$$J(L) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\|p_i - l_j\|\right)^2$$

$$L = \{l_1, l_2, l_3 \ldots l_c\} \tag{1}$$

p_i	values for points belonging to the data set.
l_j	centroid value.
$\|p_i - l_j\|$	Euclidean distance between p_i and l_j

Algorithmic Steps for K-means Clustering

Let $P = \{p_1, p_2, p_3 \ldots p_n\}$ be the set of data points and $L = \{l_1, l_2, l_3 \ldots l_c\}$ be the set of centers [9].

1. Select the c-cluster centers randomly.
2. Distance between each cluster center and data points are calculated.
3. The data points having the minimum distance from a particular cluster center among other distances are assigned to that cluster center.
4. New cluster centers are calculated as in (2):

$$v_i = \left(\frac{1}{c_i}\right) \sum_{j=1}^{c_i} p_i \tag{2}$$

   where 'c_i' represents the number of data points in the ith cluster.
5. Again, calculate the distance of each data point with newly obtained cluster center.
6. Iterate from step3 till the stage where centers remain unchanged.

3 Proposed Music Composer Identification System

The proposed multi-composer identification system is shown in Fig. 2.

3.1 Pre-processing Technique

The raw data set as such is not advisable to be used as a training dataset for the system. A good pre-processing technique leads to better results. The work done uses the single level discrete wavelet transform as a pre-processing technique.

3.1.1 Single Level Discrete Wavelet Decomposition

Daubechies Wavelet transform (db2) is a technique used to decompose a signal into a linear combination of scaling function and wavelet function. This reduces the noise components present in the tunes leading to an increase in the quality of tunes. The Fig. 3 shows the wavelet decomposition.

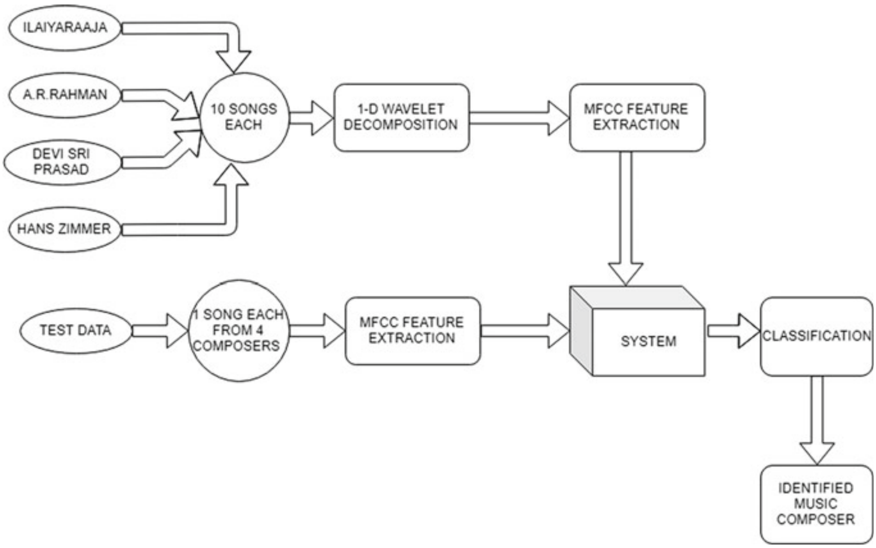


Fig. 2 Proposed music composer identification system

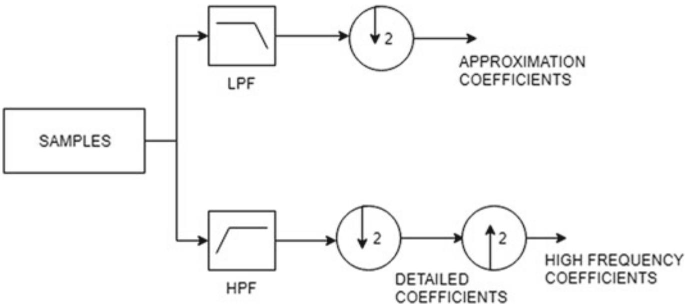


Fig. 3 1-D wavelet decomposition

The music are converted into the frames. Each frame having 882 samples with frame length of 20 ms and sampling rate of 44,100 samples/second that are passed through the low pass and high pass filters and then it is downsampled by 2. The coefficients that are produced as outputs from low pass filters are called approximation coefficients. The 441 approximation coefficients are produced from the low pass filters after downsampling. The coefficients that are produced as outputs from high pass filters are called detail coefficients. The 441 detail coefficients are produced from the high pass filters after downsampling.

Table 1 Confusion matrix—without DWT

Composer's	Rahman	DSP	Ilayaraja	Hanszimmer	RA (%)	FRR (%)
Rahman	44	4	2	0	88	12
DSP	1	42	1	6	84	16
Ilayaraja	2	13	35	0	70	30
Hans zimmer	0	19	7	24	48	52
FAR	6.38%	46.5%	22.22%	20%		

RA Recognition Accuracy, *FRR* False Rejection Ratio, *FAR* False Acceptance Ratio

3.2 Working of Proposed System

- MFCC features are extracted from the training dataset.
- Extracted features are formed into 256 clusters using K-means algorithm.
- The obtained MFCC vectors from test data are divided into 50 sub-datasets each containing 400 vectors.
- For each sub dataset mentioned above the mean distance between clusters and vectors is obtained for each music composer.
- The respective music composer who gets the least mean distance is the identified composer for that sub data set.
- The above procedure is applied for 50 data sets.
- The count at the end of 50 subtests for each composer is obtained.
- The respective composer whose count is maximum is the real composer of the given test data.

4 Results and Discussion

4.1 Performance of the System Without DWT

The system modelling has been done with four music composers in the trained database and the test results have been analyzed without DWT as a pre-processing technique and the results are shown in Table 1 and performance of the system is shown in Fig. 4.

4.2 Performance of the System with DWT

The system modelling has been done with four music composers in the trained database and the test results have been analyzed with DWT as a pre-processing

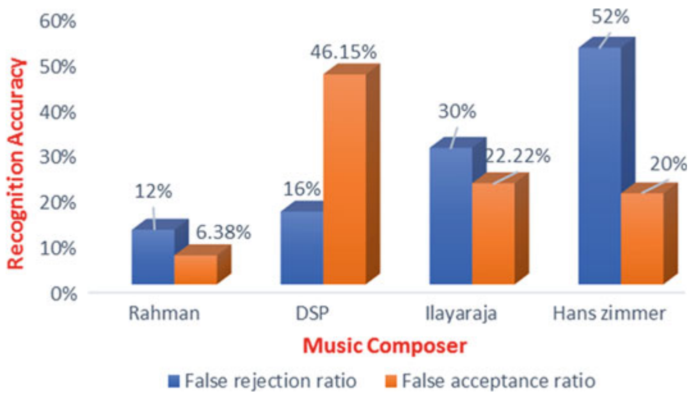


Fig. 4 Performance of the system without DWT

technique and the results are shown in Table 2 and performance of the system is shown in Fig. 5.

Table 2 Confusion matrix—with DWT

Composer's	Rahman	DSP	Ilayaraja	Hanszimmer	RA (%)	FRR (%)
Rahman	44	3	2	1	88	12
DSP	0	50	0	0	100	0
Ilayaraja	10	7	27	6	54	46
Hans zimmer	0	16	0	34	68	32
FAR	18.52%	34.21%	6.896%	17.07%		

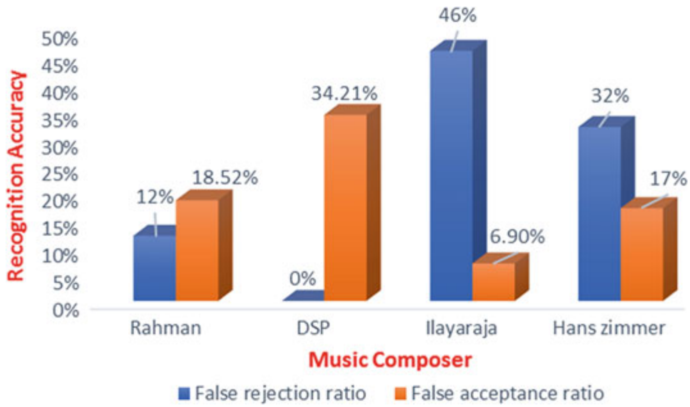


Fig. 5 Performance of the system with DWT

Table 3 Correlation analysis

Composers	Correlation coefficient between composers (without DWT)	Correlation coefficient between composers (with DWT)
Illayaraja—Rahman	0.4578	0.6795
Illayaraja—DSP	0.1056	0.1558
Illayaraja—Hans zimmer	0.4852	0.5782

4.3 Analysis

The obtained accuracy of overall system without DWT = 72.5%. The obtained accuracy of overall system with DWT = 77.5%. There were better improvements in other parameters like false rejection ratio, false acceptance ratio after usage of wavelet decomposition as a pre-processing technique DWT.

4.4 Correlation Analysis

The correlation coefficients are used to analyze the strength of the relationship between two variables. Correlation between two random variables rf_{xy} is the covariance of the two variables normalized by the variance of each variable. The correlation coefficient can be mathematically expressed as in (3)

$$rf_{xy} = \frac{\mathrm{cov}(x, y)}{\sqrt{\mathrm{var}(x)}\sqrt{\mathrm{var}(y)}} \tag{3}$$

where $\mathrm{cov}(x, y)$ is the covariance between music composer 'x'and another music composer 'y'. $\mathrm{var}(x)$, $\mathrm{var}(y)$ are the variance of the two composers x and y. The correlation coefficient between Illayaraja and other composers is shown in Table 3.

The correlation coefficient value between Illayaraja and other composers is increased, when transformed into the wavelet domain. Hence, the accuracy in the identification of the Illayaraja's composition is reduced after transforming into wavelet domain.

5 Conclusion

This work helps in the classification of music composers based on the tunes. This makes an integral part of song classification applications which provide us different archives based on music composers. This project lays a foundation on a basic understanding of extracting the features from the music files and creating a system by using classification algorithm where the extracted features are trained. Along with the feature extraction the project also gives an overview of improving the system accuracy and performance by using a pre-processing technique DWT of the music files. Having a large data set of every music composer and efficient pre-processing techniques prove to be powerful tools for music composer classification. This work finds applications in music information retrieval systems where a small duration of music is enough to get the classification of the music composer.

References

1. Kaliakatsos-Papakostas MA, Epitropakis MG, Vrahatis MN (2010) Musical composer identification through probabilistic and feedforward neural networks. In: Di Chio C et al (eds) Applications of evolutionary computation. Evo Applications 2010. Lecture notes in computer science, vol 6025. Springer, Berlin, Heidelberg
2. Micchi G (2018) A neural network for composer classification. In: International society for music information retrieval conference (ISMIR 2018), Paris
3. Hajj N, Filo M, Awad M (2018) Automated composer recognition for multi-voice piano compositions using rhythmic features, *n*-grams and modified cortical algorithms. Complex Intell Syst 4(1):55–652018
4. Bartle A (2012) Composer identification of digital audio modeling content specific features through markov models
5. Geertzen J, van Zaanen M (2008) Composer classification using grammatical inference. In: Proceedings of the MML 2008 international workshop on machine learning and music held in conjunction with ICML/COLT/UAI, pp 17–18
6. Hontanilla M, Pérez-Sancho C, Inesta JM (2013) Modeling musical style with language models for composer recognition. Pattern recognition and image analysis. Springer, Berlin, pp 740–748
7. Herremans D, Martens D, Sörensen K (2016) Composer classification models for music-theory building. In: Meredith D (ed) Computational Music analysis. Springer, Cham
8. Hu Z, Fu K, Zhang C (2013) Audio classical composer identification by deep neural network
9. Mathew AR, Anto PB, Thara NK (2017) Brain tumor segmentation and classification using DWT, Gabour wavelet and GLCM. In: 2017 international conference on intelligent computing, instrumentation and control technologies (ICICICT), Kannur, pp 1744–1750

Spectral Efficiency-Energy Efficiency Tradeoff Analysis for a Carrier Aggregated 5G NR Based System

Syama Sasikumar and J. Jayakumari

Abstract For 5G and beyond 5G systems, many technologies such as Carrier Aggregation (CA) have been suggested to improve the Spectral Efficiency (SE). Along with SE, Energy Efficiency (EE) is another key parameter to be considered while designing such systems. Judiciously using available power will not only help in reducing operational costs but also make the system more environment friendly by reducing greenhouse gas emissions. From an operator's point of view, providing better data rates to users as well as reducing operational costs is important. This paper focuses on the tradeoff between SE and EE for a system operating in the FR1 frequency band (5G) and also using CA. Detailed analysis on the effect of various parameters such as 5G NR numerology, user velocity (Doppler shift), static and signal processing power consumption of base station and power amplifier efficiency at the base station in the tradeoff between SE and EE is performed. From the simulation results, it can be observed that while varying Doppler shift has very little impact, other parameters such as numerology, static and signal processing power consumption and power amplifier efficiency greatly affect the tradeoff between SE and EE. Tradeoff improves for lower subcarrier spacing, lower static and signals processing power consumption and higher power amplifier efficiency. For simulations, a TDL-A channel model is used. A base station power consumption model which takes into account both static and signal processing power consumption and power amplifier efficiency are considered in order to provide a more practical analysis of EE.

Keywords Spectral efficiency · Energy efficiency · Carrier aggregation · 5G · Tradeoff

S. Sasikumar (✉) · J. Jayakumari
Department of ECE, Mar Baselios College of Engineering and Technology, APJ Abdul Kalam Technological University, Trivandrum, Kerala, India
e-mail: syamanair02@gmail.com

J. Jayakumari
e-mail: jayakumari.j@mbcet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_5

1 Introduction

The huge increase in user density (monthly data traffic worldwide is expected to be 50 petabytes by 2021 for smart phones alone) and demand for higher data rates has propelled the fast evolution of Information and Communication Technologies (ICT). The official launch of 5G technology is expected to happen in 2020. Few of the many drivers for 5G are mmWave technology, spectrum sharing, CA, use of unlicensed spectrum, massive Multiple Input Multiple Output (MIMO) and Cloud Radio Access Network (C-RAN) [1]. CA technology, a part of 3GPP release 10, is already implemented in Long Term Evolution (LTE) networks in a limited scale. CA increases the bandwidth provided to a single user by combining more than one component carrier (CC) with lower bandwidth, thereby increasing data rates. An overview of CA for LTE-advanced networks is presented in [2]. Various aspects such as design in both physical and higher layers, deployment, power control, etc. are discussed in detail. Several research works have focused on various aspects of CA. Power allocation for a CA-based Cognitive Radio Network (CRN) is investigated in [3]. Selection of sub-channels for a CA system is performed in [4]. Performance of user equipment (UE) s across different layers in a CA heterogeneous network (HetNet) is studied in [5].

SE and its variants such as Area Spectral Efficiency (ASE), Cell Spectral Efficiency (CSE) and Peak Spectral Efficiency (PSE) have long been an important performance metric for many systems. Recently, EE has also emerged as another crucial parameter for analyzing the performance of a system. Since it is equally important to satisfy users by providing good data rates, as well as reduce the system's operational expenses and contain its impact to the environment, tradeoff between SE and EE should be studied and efforts should be made for improvement. Many works in literature have addressed the problem of SE-EE tradeoff. Joint optimization of EE and ASE from the base station's viewpoint for an ultra-dense network is performed in [6]. Firefly algorithm is employed for optimization. Maximization of EE and SE in a HetNet which shares radio resources is performed in [7]. Multi-Objective optimization for considering SE-EE tradeoff for Device-to-Device (D2D) communications is performed in [8]. Optimization is done using ε-constraint method with robustness. Zhang et ai. [9] deals with power allocation for EE-SE tradeoff in both low and high vehicular traffic density situations. A relatively new parameter called Economic Efficiency (ECE) is used for evaluating the performance of the optimization technique. EE-SE tradeoff in D2D networks within a single cell with uplink channel sharing is addressed in [10]. Tsilimantos et al. [11] investigates EE-SE tradeoff in orthogonal cellular networks. A theoretical framework is proposed for optimum allocation of the resources, bandwidth and power.

In this work, we perform an analysis on the effect on important parameters such as 5G NR numerology, user velocity, static and signal processing power consumptions and power amplifier efficiency at the base station on the SE-EE tradeoff in 5G FR1 frequency band based system employing CA. Such insight into the variation of the tradeoff will contribute to devising techniques for improving the relationship between both parameters.

The remaining content of the paper is organized as follows. Section 2 describes the system and channel models, Sect. 3 discuss the parameters used for the analysis. Section 4 presents the simulation results and the inferences drawn from them. Finally, conclusions and possibilities for future work are mentioned in Sect. 5.

2 System and Channel Models

This section describes the system and channel models chosen for performing the analysis.

2.1 System Model

Figure 1 depicts the system model used for the work. A next-generation node B (gNB), i.e. the base station, capable of CA sends three CCs each with a different bandwidth to user equipment (UE). Therefore, the UE enjoys the benefits of the combined bandwidth of all the three CCs.

Bandwidth of each CC is given by the relation [12],

$$B_i = n\_\mathrm{RB}_i * 12 * \mathrm{SCS}_i \tag{1}$$

where, i is the CC number, n_RB_i and SCS_i are the number of downlink resource blocks and subcarrier spacing, respectively, for the ith CC.

The final aggregated bandwidth given to the UE is [12],

$$\mathrm{BW\_CA} = \mathrm{FE\_high} - \mathrm{FE\_low} \tag{2}$$

where FE_high and FE_low are the higher and lower edge frequencies the expressions for which can be found in [12]. Signal-to-noise ratio (SNR) for the kth subcarrier of ith CC can be expressed as,

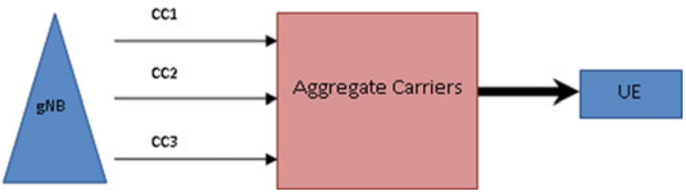


Fig. 1 System model

$$\gamma_k = \frac{h_k P_D(k,i)}{N0_i B_k} \tag{3}$$

h_k, $P_D(k, i)$, $N0_i$ and B_k are the subcarrier channel gain, dynamic power allocated to kth subcarrier, noise power spectral density and subcarrier bandwidth of the ith CC.

2.2 Channel Model

A tapped delay line (TDL) channel model specified in [13] is used. Out of the many variants of the TDL model, TDL-A model with 23 taps is selected for performing the analysis. It is a simple model which can be used for non-line of sight (NLOS) evaluations. The path gains of the channel for CC-1 is shown in Fig. 2.

3 Parameters for Analysis

This section describes in detail the various parameters used for analysis in this work.

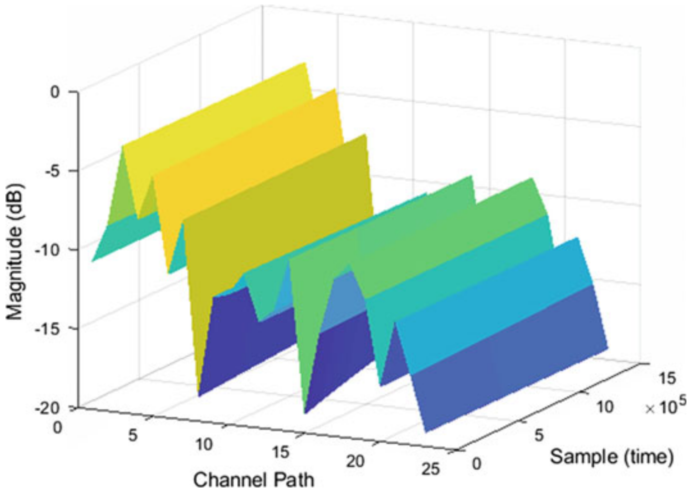


Fig. 2 Pathgains of TDL-A channel for CC-1

Table 1 5G NR numerology

μ	SCS
0	15
1	30
2	60
3	120
4	240

3.1 5G NR Numerology

Unlike 4G, where only 15 kHz subcarrier spacing and therefore a constant 180 kHz bandwidth was allowable for a resource block, 5G NR supports a flexible design with various Orthogonal Frequency Division Multiplexing (OFDM) numerologies as depicted in Table 1 [14]. Different subcarrier spacing may cater to different scenarios. For example, a lower subcarrier spacing and thereby a higher OFDM symbol duration may be better when frequency selective fading is experienced while a higher subcarrier spacing and lower symbol duration may be better in conditions where fast fading is experienced.

SCS is the subcarrier spacing and it is related to μ according to the following relation.

$$\text{SCS} = 2^{\mu} * 15(\text{kHz}) \tag{4}$$

3.2 User Velocity

In the case of mobile equipment users, mobility is an important factor for consideration. Pedestrian users may have very low velocity while vehicular users may be comparatively fast-moving. Mobility of the user is crucial since it affects an important parameter called the Doppler frequency shift. The maximum Doppler frequency shift is given by

$$f_{D\text{max}} = \frac{v * c}{f_c} \tag{5}$$

where v, c and f_c are the user velocity, velocity of light and carrier frequency respectively.

3.3 Power Amplifier Efficiency, Static and Signal Processing Power Consumptions at the Base Station

The dynamic power allocated to the transmission of the component carrier is not the only parameter that should be considered while evaluating EE. The efficiency of the power amplifier at the base station determines its power consumption. Signal processing, DC to DC conversion, baseband unit operations etc. consume power in addition to transmission power. Also, even when there is no signal transmission, the base station might consume some power for purposes such as cooling. A simple model of base station power consumption for the considered system [9] is given below.

$$P_{\text{tot}} = \varepsilon \sum_{i=1}^{n\_\text{CC}} \sum_{k=1}^{n\_\text{SC}_i} P_D(k, i) + P_c \tag{6a}$$

$$P_c = P_{\text{sp}} + P_{\text{static}} \tag{6b}$$

where, $1/\varepsilon$ is the power amplifier drain efficiency, P_{sp} and P_{static} are the signal processing and static power consumptions at the base station.

3.4 Spectral Efficiency and Energy Efficiency

SE, a measure of data rate, is calculated as follows.

$$\text{SE} = \frac{C_{\text{tot}}}{\text{BW\_CA}} (\text{bits/s/Hz}) \tag{7}$$

Total system capacity,

$$C_{\text{tot}} = \sum_{i=1}^{n\_\text{CC}} \sum_{k=1}^{n\_\text{SC}_i} C(k, i) \tag{8}$$

where n_CC is the number of component carriers and n_SC_i is the number of subcarriers of the ith component carrier given by,

$$n\_\text{SC}_i = n\_\text{RB}_i * 12 \tag{9}$$

since each resource block of 5G consists of 12 subcarriers.

Subchannel capacity of kth subcarrier of ith CC,

$$C(k, i) = B_k \log_2(1 + \gamma_k) \tag{10}$$

EE, a measure of efficient utilization of the resource power, can be calculated as,

$$EE = \frac{SE}{P_{tot}} \tag{11}$$

4 Simulation Results and Inferences

This section presents the results of simulations and analyses them to draw inferences. Table 2 presents the simulation parameters. Since the number of slots and symbols per subframe vary with numerology, they are not mentioned in the table. Figure 3 shows the aggregated signal with the combined bandwidth of all CCs.

Table 2 Simulation parameters

Number of CCs	3
Number of downlink resource blocks	160, 140, 120
Modulation	Windowed orthogonal frequency division multiplexing (W-OFDM)
N0	−90 dBm for all CCs
Center frequency of first CC	450 MHz (FR1 frequency band of 5G) [13]

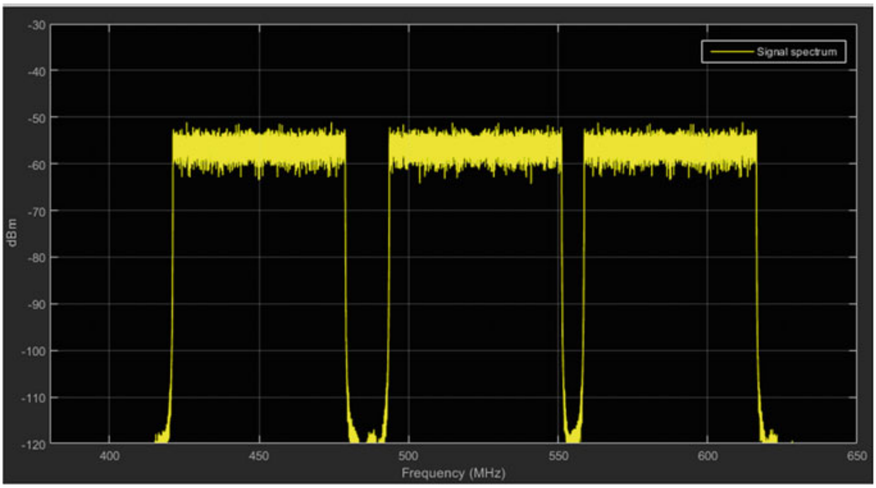


Fig. 3 Aggregated signal

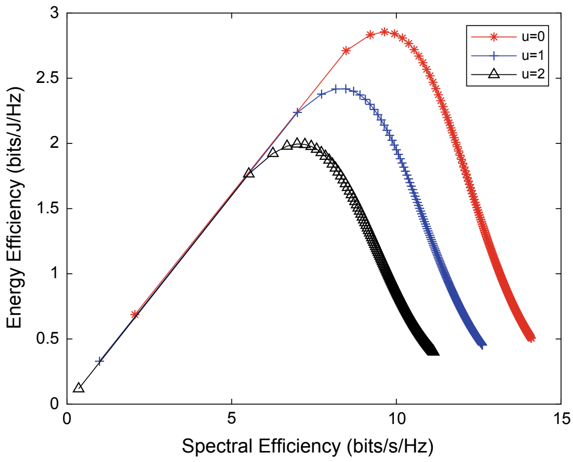


Fig. 4 SE-EE tradeoff with varying numerology

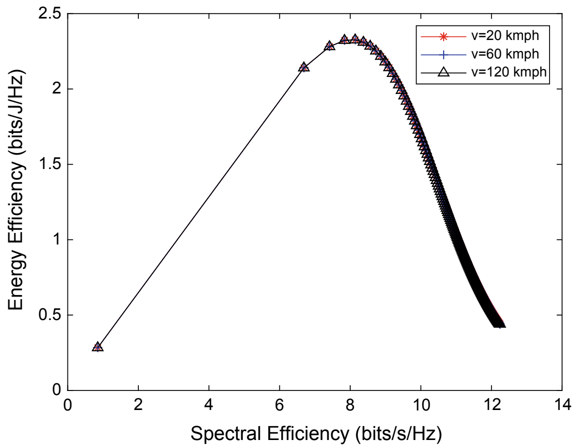
4.1 *Varying Numerology*

Figure 4 shows the effect of varying numerology on SE-EE tradeoff. Tradeoff is plotted for varying values of the factor μ. It can be observed that lower subcarrier spacing improves the tradeoff. EE improves almost by 20% with single point decrease in μ for the same SE. This behaviour may be attributed to the dependency of the effect of noise on the bandwidth. Increasing μ increases the bandwidth as can be observed from (1) and (4). Increase in bandwidth increases the effect of noise and thus negatively impacts the tradeoff.

4.2 *Varying User Velocity*

Figure 5 shows the effect of varying user velocity on SE-EE tradeoff. Simulations are performed for a low-velocity user (20 km/h—slow-moving vehicle), medium velocity user (60 km/h—fast-moving vehicle) and a high-velocity user (120 km/h—high-speed vehicle). It can be observed that though Doppler shift caused by moving users is undesirable, it does not seem to impact the tradeoff. The reason behind such behaviour may be explained as follows. The coherence time for the same users is approximately 31 ms, 10 ms, 5 ms, respectively, for the parameters mentioned in Table 2. OFDM symbol duration chosen for the system is 33.33 μs which is well within the coherence time for all three types of users and hence the increasing speed has little impact on EE-SE tradeoff for the chosen carrier frequency.

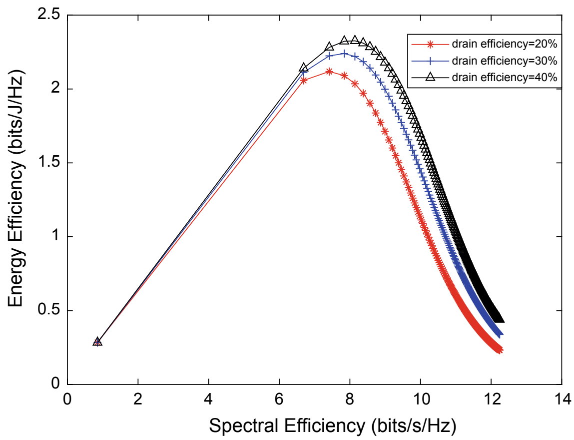
Fig. 5 SE-EE tradeoff with varying user speed



4.3 Varying Power Amplifier Efficiency

Figure 6 shows the effect of varying power amplifier efficiency on EE-SE tradeoff. Power amplifier efficiency is a measure of DC power converted to RF power. It can be observed that higher power amplifier efficiency improves tradeoff (almost 12% improvement in EE for the same SE for 10% increase in power amplifier efficiency). This happens since higher power amplifier efficiency would mean less power consumption by the power amplifier and hence higher EE.

Fig. 6 SE-EE tradeoff with varying power amplifier efficiency



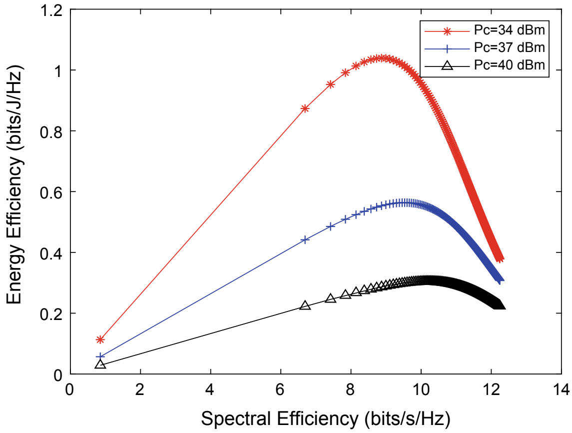


Fig. 7 SE-EE tradeoff with varying static and signal processing power consumption

4.4 Varying Static and Signal Processing Power Consumptions at the Base Station

Figure 7 shows the effect of varying static and signal processing power consumption on EE-SE tradeoff. It can be observed that with 3 dBm reduction in Pc, more than 100% improvement is obtained in EE for the same SE and the tradeoff improves considerably.

5 Conclusions and Future Work

Analysis on the effect of parameters 5G NR numerology, user velocity, static and signal processing power consumption of the base station and power amplifier efficiency at the base station on EE-SE tradeoff for a CA system working in 5G NR frequency band was conducted in this work. It can be concluded that while user velocity has little impact on the tradeoff, other parameters affect it significantly. Decreasing the numerology factor μ by 1 increases EE by almost 20% for the same SE. Increasing the power amplifier efficiency improves EE by 12% for the same SE and decreasing the static and signal processing power consumption of base station even by 3 dBm improves the EE by more than 100% for the same SE. In future, the results of this analysis can be used in designing methods to improve EE-SE tradeoff for 5G and beyond 5G systems using CA. Also, the study can be extended to HetNets and multi-user networks.

Acknowledgements This work is funded by Centre for Engineering Research and Development (CERD), A P J Abdul Kalam Technological University, Kerala, India.

References

1. Morgado A, Huq KM, Mumtaz S, Rodriguez J (2018) A survey of 5G technologies: regulatory, standardization and industrial perspectives. Digital Commun Netw 4(2):87–97. https://doi.org/10.1016/j.dcan.2017.09.010
2. Shen Z, Papasakellariou A, Montojo J, Gerstenberger D, Xu F (2012) Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications. IEEE Commun Mag 50(2):122–130. https://doi.org/10.1109/mcom.2012.6146491%5b3]. Lee JS (1981) Refined filtering of image noise using local statistics. Comput Graphic Image Process 15:380–389
3. Diamantoulakis PD, Pappi KN, Muhaidat S, Karagiannidis GK, Khattab T (2017) Carrier aggregation for cooperative cognitive radio networks. IEEE Trans Veh Technol 66(7):5904–5918. https://doi.org/10.1109/tvt.2016.2635112
4. Tsinos CG, Foukalas F, Khattab T, Lai L (2018) On channel selection for carrier aggregation systems. IEEE Trans Commun 66(2):808–818. https://doi.org/10.1109/tcomm.2017.2757478
5. Alorainy A, Hossain MJ (2019) Cross-layer performance analysis of downlink multi-flow carrier aggregation in heterogeneous networks. IEEE Access 7:23303–23318. https://doi.org/10.1109/access.2019.2899114
6. Luo Y, Shi Z, Bu F, Xiong J (2019) Joint optimization of area spectral efficiency and energy efficiency for two-tier heterogeneous ultra-dense networks. IEEE Access 7:12073–12086. https://doi.org/10.1109/access.2019.2891551
7. Aydin O, Jorswieck EA, Aziz D, Zappone A (2017) Energy-spectral efficiency tradeoffs in 5G multi-operator networks with heterogeneous constraints. IEEE Trans Wirel Commun 16(9):5869–5881. https://doi.org/10.1109/twc.2017.2716948
8. Hao Y, Ni Q, Li H, Hou S (2018) Robust multi-objective optimization for EE-SE tradeoff in D2D communications underlaying heterogeneous networks. IEEE Trans Commun 66(10):4936–4949. https://doi.org/10.1109/tcomm.2018.2834920
9. Zhang R, Li Y, Wang C, Ruan Y, Fu Y, Zhang H (2018) Energy-spectral efficiency trade-off in underlaying mobile D2D communications: an economic efficiency perspective. IEEE Trans Wirel Commun 17(7):4288–4301
10. Bhardwaj A, Agnihotri S (2018) Energy- and spectral-efficiency trade-off for D2D-multicasts in underlay cellular networks. IEEE Wireless Commun Lett 7(4):546–549. https://doi.org/10.1109/lwc.2018.2794353
11. Tsilimantos D, Gorce J, Jaffres-Runser K, Poor HV (2016) Spectral and energy efficiency trade-offs in cellular networks. IEEE Trans Wirel Commun 15(1):54–66. https://doi.org/10.1109/twc.2015.2466541
12. GPP TS 38.104(2019) V 15.5.0. Available at https://www.etsi.org/docdeliver/etsi_ts/138100_138199/138104/15.03.00_60/ts_138104v150300p.docx
13. ETSI TR 138901(2018) V14.3.0. Available at https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/14.03.00_60/tr_138901v140300p.pdf
14. ETSI TS 138211(2018) V15.2.0. Available at https://www.etsi.org/deliver/etsi_ts/138200_138299/138211/15.02.00_60/ts_138211v150200p.pdf

An Image Restoration Method for Outdoor and Its Application to Under Water Using Improved Transmission Map and Airlight Estimation

D. Eesha and Siddappaji

Abstract Dehazing is an important image restoration technique to remove the presence of Haze from a hazy image. Recent dehazing algorithm is not sufficient to remove Haze from the given outdoor or underwater hazy images. Therefore, an efficient dehazing algorithm is needed for the removal of Haze. Initially, multiple image dehazing methods are used to remove Haze and these dehazing methods have many drawbacks such as, multiple image methods cannot be applied to dynamic scenes and cannot provide results instantly. In order to overcome drawbacks of multiple image dehazing methods, Single image dehazing methods are introduced which are based on some important observations or priors. One such single image dehazing technique is dark channel prior. The thickness of Haze and airlight is estimated using dark channel prior. Guided Filter technique is used to refine the transmission map. But the estimated Haze thickness is inaccurate because of the usage of minimum operator in dark channel prior method. To improve the estimation of Haze thickness, the edge collapse based repair is used after dark channel prior and guided filter technique. This paper presents the time-efficient dehazing of outdoor images with patch size of 25 $\times$ 25 and airlight of 3% and this principle is applied to remove Haze in underwater images. The experimental result shows a better result for both outdoor and underwater images.

Keywords Haze · de-Haze · Outdoor and underwater · Haze thickness · Airlight

1 Introduction

The atmospheric particle interacts with light and causes scattering and absorption of light, which results in poor visibility. Haze is a weather condition caused by a particle in the atmosphere called Aerosol. The sources for Haze include volcanic

D. Eesha (✉) · Siddappaji
Department of ECE, BMS College of Engineering, Bengaluru, India
e-mail: eesha.ece@bmsce.ac.in

Siddappaji
e-mail: siddu.ece@bmsce.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_6

ashes, foliage exudation, and combustion products in outdoor and sand, minerals, and plankton in underwater. Many automatic systems like surveillance, intelligent vehicles, etc., always assume that the input images are of good quality and have clear visibility. If the quality of the images is poor because of Haze, then these applications will not work efficiently. Therefore, removal of Haze is very important.

In order to remove Haze in outdoor and underwater, many dehazing methods use dark channel prior. However, the usage of two minimum operators reduces the performance of this algorithm. To improve the estimation of Haze thickness, the edge collapse based repair algorithm is used after dark channel prior and guided filter. In this paper, this concept is implemented with a patch size of 25×25 to improve the quality of image in outdoor conditions and is applied to remove Haze from underwater images.

2 Related Work

In order to perform dehazing, initially, multiple image dehazing methods [1, 2] were developed which require information from more than one image to perform dehazing.

These multiple image methods are not time-efficient and they cannot be applied to dynamic scenes. So, single image dehazing methods are developed which are successful as they are based on stronger assumptions, observations or priors.

A single image dehazing algorithm is proposed [3] based on the observation that the contrast of Haze-free image must be greater than the Haze image. Based on this principle, the authors have removed the Haze by maximizing local contrast of the input hazy image. This principle is visually compelling but may not be physically valid. Similarly, in paper [4], another important single image dehazing method is proposed based on the observation that in an outdoor Haze-free image at least one color channel will have zero or low intensity in the non-sky regions. This observation is called as ‘dark channel prior’. This dehazing method uses a dark channel prior to estimating the transmission map and to refines the estimated transmission map using soft matting in order to eliminate halo-artefacts. This method also estimates airlight and finally all estimated results are used to remove the Haze.

The usage of soft matting makes the dehazing algorithm time inefficient. In order to reduce the execution time, the authors have used a guided filter [5] to perform the refining of the estimated transmission map. As a result, the execution time is reduced and quality of the image is increased. Even though halo-artefacts are removed, the Haze thickness in the transmission map may be underestimated because of the usage of two minimum operators while finding a dark channel.

To overcome this problem, the authors in paper [6] have proposed an edge collapse-based repair technique, which can properly estimate the Haze thickness.

The dark channel prior to condition with smaller changes is applied to underwater condition because of its good performance in outdoor. Some of the important underwater dark channel prior based dehazing algorithm was presented in [7–11]. These underwater dehazing algorithms change the dark channel prior according to underwater condition and use refining methods like soft matting or guided filter to remove halo-artefacts. However, these methods fail in proper estimation of the transmission map because of the usage of two minimum operators while calculating the dark channel. Therefore, edge collapse based repair as suggested in paper [6] for outdoor conditions. Here, this technique is applied to remove Haze in underwater conditions.

3 Dehazing Algorithm

In order to perform the dehazing of an image, a dark channel prior based dehazing method is applied for outdoor and underwater images. The dehazing method estimates the Haze thickness by using dark channel prior (DCP) for outdoor hazy images and Underwater Dark Channel Prior (UDCP) for underwater hazy images. The estimated Haze thickness is refined using a guided filter and repaired using edge collapsed repair. The reason for refining the transmission map is to improve Haze thickness measurement and to remove halo-artefacts, which is present in the image. Here, dark channel is used for calculating atmospheric light and finally Haze is removed. The Fig. 1 shows the flow diagram for dehazing in outdoor and underwater images.

The following steps are to be carried out to remove Haze in outdoor and underwater conditions.

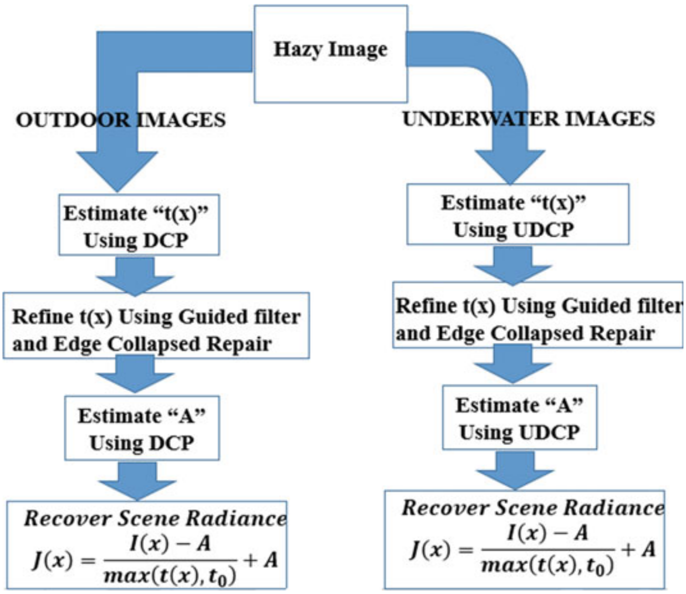


Fig. 1 Methodology for dehazing both outdoor and underwater images

3.1 Providing Input to the Dehazing System

An outdoor or underwater hazy image is given as input to the dehazing algorithm.

3.2 Obtain Dark Channel

The dehazing algorithm taken the hazy image as the input and calculates its dark channel using below equation.

For an outdoor image (J) the dark channel is given by,

$$J^{\text{dark}}(x) = \min_{y \in \Omega(\mathrm{x})}\left(\min_{\mathrm{c} \in \{\mathrm{r,g,b}\}} J^{c}(y)\right) \tag{1}$$

Here, minimum operator min c∈{r, g, b} is applied on every pixel of the input image and min $y \in \Omega(x)$ is one more minimum operator applied on a patch.

For an underwater image (J), the Underwater DCP is given by

$$J^{\text{dark}}(x) = \min_{y \in \Omega(\mathrm{x})}\left(\min_{\mathrm{c} \in \{\mathrm{g,b}\}} J^{c}(y)\right) \tag{2}$$

Here, minimum operator min c∈{g, b} is applied on green and blue channel of the input image and min $y \in \Omega(x)$ is a minimum operator applied on a patch.

In this work, the patch size is 25 × 25 considered. An experimental survey is performed to determine the effective patch size to restore the image. The results obtained for the patch size of 25 × 25 is better than other patch sizes. Earlier methods [4, 5, 6] had considered 15 × 15 as patch size.

To understand the calculation of dark channel, consider an example where J is a hazy image and J^{c} is color channel (Fig. 2) and $\Omega(x)$ is a local patch cantered at x.

From dark channel prior, we know that for an image which is Haze-free, the dark channel is completely dark. Similarly, for an image that is affected by Haze, the dark channel is not dark because of added atmospheric light.

3.3 Transmission Map Estimation

The following steps are followed to estimate the transmission map. Here the derivation is done by assuming atmospheric light is known prior.

Consider Haze Image Equation

In image processing, the formation of Haze image is described by Koschmieder's model [12] as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{3}$$

a. Haze free image

b. For each pixel, calculate min minimum of (r, g, b) values

c. Minimum filter is applied on(4b)

Fig. 2 Understanding calculation of dark channel

The above equation is normalized for each color channel independently.

$$\frac{I^c(x)}{A^c} = t(x)\frac{J^c(y)}{A^c} + 1 - t(x). \tag{4}$$

Here, transmission t(x) in a local patch is assumed as constant.

Calculate the dark channel on both side of normalized Haze image equation
For outdoor hazy image:

$$\min_{y \in \Omega(x)}\left(\min_{c \in \{r,g,b\}} \frac{I^c(y)}{A^c}\right) = \tilde{t}(x) \min_{y \in \Omega(x)}\left(\min_{c \in \{r,g,b\}} \frac{J^c(y)}{A^c}\right) + 1 - \tilde{t}(x) \tag{5}$$

For underwater hazy image:

$$\min_{y \in \Omega(x)}\left(\min_{c \in \{r,g,b\}} \frac{I^c(y)}{A^c}\right) = \tilde{t}(x) \min_{y \in \Omega(x)}\left(\min_{c \in \{g,b\}} \frac{J^c(y)}{A^c}\right) + 1 - \tilde{t}(x) \tag{6}$$

Here $\tilde{t}(x)$ is a constant. So take this separately. As J indicates a Haze-free image, its dark channel is equal to zero.

Re-arrange the equation
After making dark channel of J as zero and re-arranging the equation, the final transmission map is obtained. The final equation for transmission map is given by,

For outdoor hazy image:

$$\tilde{t}(x) = 1 - w \min_{y \in \Omega(x)} \left(\min_{c \in \{r,g,b\}} \frac{I^c(y)}{A^c} \right) \tag{7}$$

For underwater hazy image:

$$\tilde{t}(x) = 1 - w \min_{y \in \Omega(x)} \left(\min_{c \in \{g,b\}} \frac{I^c(y)}{A^c} \right) \tag{8}$$

The complete removal of Haze from input image makes the image appear unnatural. So, keep a small amount of Haze. So a parameter 'w' is introduced in the equation. The parameter 'w' is application dependent. Here the value of parameter 'w' is taken as 0.95.

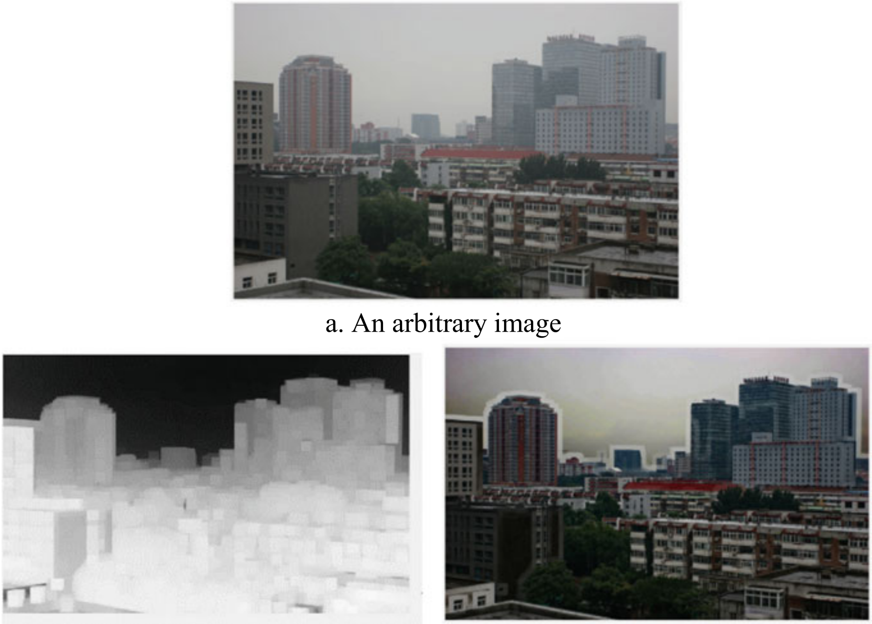
3.4 Atmospheric Light Estimation

From dark channel of input hazy image, atmospheric light is estimated. The earlier dehazing methods have considered top 0.1% brightest pixels from dark channel. In the present method top 3% brightest pixels from dark channel is considered. An experimental survey is performed to determine the effect of atmospheric light on recovery of the image. The performance of the system is better for 3% airlight compared to 0.1% airlight.

3.5 Image Visibility Recovery Without Refining Transmission Map

Once the transmission map and atmospheric light for outdoor or underwater hazy image is obtained, the Haze can be removed by using the equation given below,

$$J(x) = \frac{I(x) - A}{\max\left(t_{r(x,y)}, t_0\right)} + A \tag{9}$$



a. An arbitrary image

b. Transmission map of fig.3a

c. Recovery using transmission map

Fig. 3 Recovery of an image using transmission map which is not refined

Here, the transmission $t(x)$ becomes zero if the distance of the scene is very large. Therefore, parameter t_0 introduced to provide lower bound for the transmission and its typical value is 0.1. A is used to improve the brightness of dehazed image.

If the above equation is used to perform the recovery of the image (Fig. 3a) where $t(x)$ (Fig. 3b) is the transmission map obtained in the previous step, then the result obtained (Fig. 3c) is not good because halo-artefacts are not removed and recovery of the image is incomplete. So refine the transmission map.

3.6 Refinement Using Guided Filter

Guided filter [13] is used to remove halo-artefacts in the hazy image. To remove halo-artefacts, soft matting [14] can also be used. However, it is not time-efficient. So guided image filtering method is used.

$$t_{r(x,y)} = t\max\left(\frac{t(x, y)}{t\max}\right)^{\gamma} \tag{10}$$

3.7 Edge Collapse Based Repair

The combination of dark channel prior and guided filter will estimate and refine the transmission map (indicates Haze thickness) respectively. However, the estimated Haze thickness is inaccurate because of the usage of two minimum operators while calculating the dark channel. Therefore, edge collapse based repair as suggested in paper [6] for outdoor conditions. Here, this technique is applied to remove Haze in underwater conditions. According to [6], it is observed that edge information provided by a hazy image is lesser than the Haze-free image. This method uses this edge attenuation as an advantage. We know that the edge information present in an image is given by the entropy of gradient magnitude. So this entropy can be considered as clue for thickness of the Haze present in the hazy image. Using this clue, the Haze thickness can be adjusted. So the output of the guided filter ($t(x, y)$)is repaired using [6]. According to [6] the repaired transmission map is given by,

$$t_{r(x,y)} = t_{\max}\left(\frac{t(x, y)}{t_{\max}}\right)^{\gamma}, \tag{11}$$

'γ' represents a self-adjusting parameter that changes dynamically with different Haze images (see below table and respective hazy images) and it can be used for repairing the transmission map efficiently (Table 1).

Table 1 Value of gamma for different scenario

Scenario	Building	Temple	Trees
Gamma	1.0376	1.1641	1.0120

Building

Temple

Trees

a. An arbitrary image

b. Tramsmission map of fig.6a c. Recovery using transmission map

Fig. 4 Recovery of an image using refined transmission map

3.8 Image Visibility Recovery with Refined Transmission Map

After refining the transmission map using guided filter and repaired using edge collapse based repair, if the recovery of the image is done, the halo-artefacts are removed and recovery of the image is complete which is shown in Fig. 4.

4 Results and Discussion

4.1 Comparison of Dehazing Algorithms

Outdoor Conditions

This section presents the comparison of improved dehazing algorithms with the existing algorithms such as Zhang et al.'s method [5], and Chen et al.'s method [6].

In order to evaluate the results, the following criteria are used. Here, different Statistical matrices like Peak Signal to Noise (PSNR) Ratio, Mean Squared Error (MSE), etc. are used to evaluate a dehazing algorithm. Refer [15] for more information about result evaluation methods.

The below table and Fig. 5 shows the comparison of different dehazing algorithm (Table 2).

The dehazing method from Zhang et al. [5] uses dark channel prior to and guided filter with patch size of 15 × 15 and airlight of top 0.1%. Whereas the dehazing method from Chen et al. [6] uses dark channel prior, guided filter and edge collapse repair with patch size of 15 × 15 and airlight of top 0.1%.It can be observed that Zhang et al. [5] method is subset of edge collapse based method [6], where the method [6] uses [5] and edge collapse repair. From the above table, it can be seen that the value of result evaluation methods produced by method [6] with no edge collapse repair (same as Zhang et al. [5]) is lesser. Therefore, to improve Chen et al. [6] method patch size of 25 × 25 and airlight of top 3% is considered. So, the improved method uses dark channel prior, guided filter and edge collapse repair with patch size of 25 × 25 and airlight of top 3%.

Underwater Conditions

The improved dehazing method is compared with other dehazing methods like Nascimento et al.'s method [9] (UDCP) and Chen et al.'s method [6]. The below table and Fig. 6 shows the comparison of different dehazing algorithm (Table 3).

The dehazing method from Nascimento et al. [5] uses underwater dark channel prior and guided filter with patch size of 15 × 15 and airlight of top 0.1%. Whereas the dehazing method from Chen et al. [6] is applied to underwater conditions which

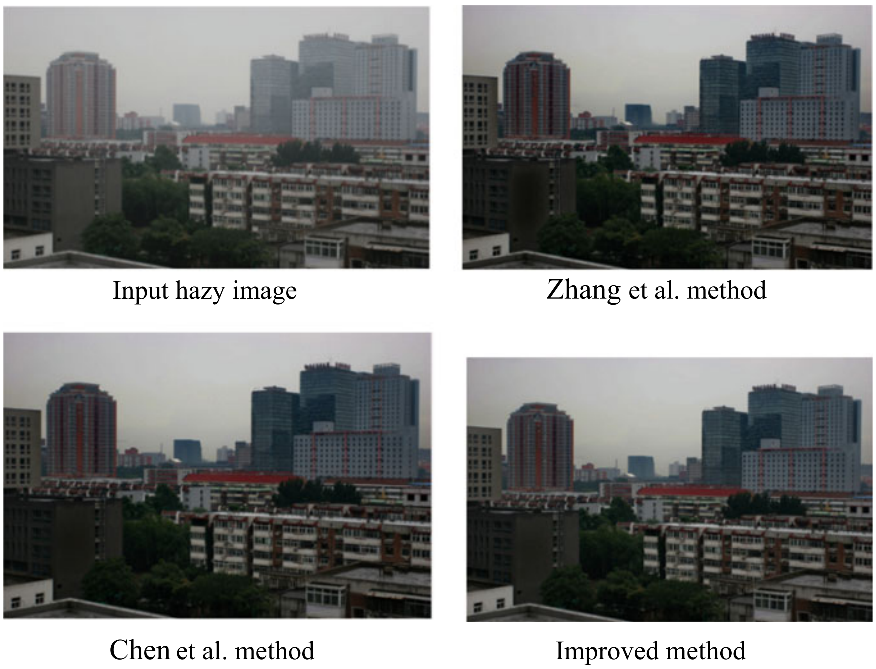


Fig. 5 Comparison of outdoor dehazing algorithms

Table 2 Comparison of different outdoor dehazing algorithm

Result evaluation method	Zhang et al. [5]	Chen et al. [6]	Improved method
PSNR	18.1939	17.7919	18.7899
MSE	985.5865	1081	859.1857
NCC	0.8289	0.8218	0.8409
AD	25.5898	26.7313	23.5871
SC	1.3836	1.3986	1.3540
MD	90	95	86
NAE	0.2351	0.2456	0.2167

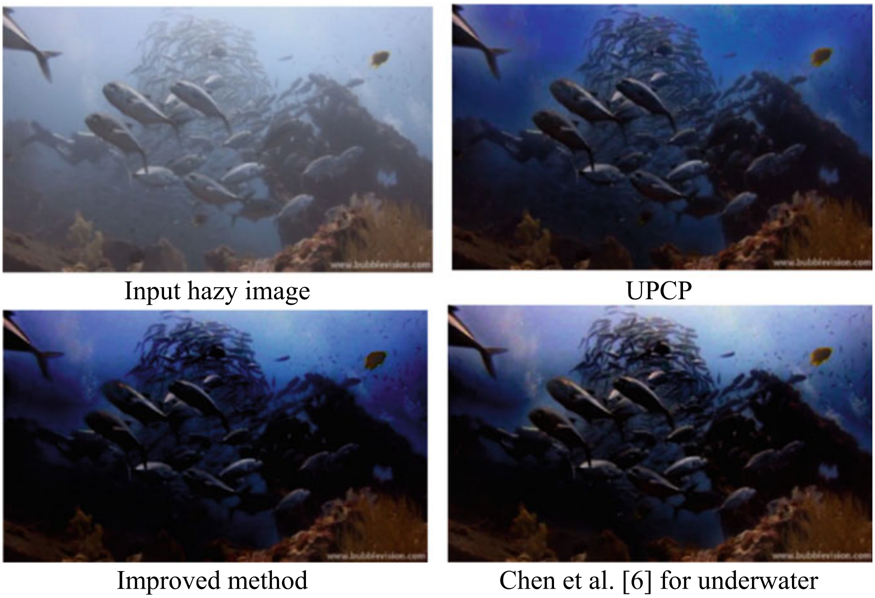


Fig. 6 Comparison of Underwater dehazing algorithms

Table 3 Comparison of different underwater dehazing algorithm

Result evaluation method	UDCP	Chen et al. [6] applied to underwater	Improved method
PSNR	9.5734	8.7061	10.1531
MSE	7173	8759	6277
NCC	0.3708	0.3403	0.4884
AD	79.9435	88.5606	73.1448
SC	6.0466	5.0531	2.8767
MD	151	157	147
NAE	0.6585	0.7294	0.6033

Table 4 Processing speed for different sized images

Image size (pixels)	194 × 259	400 × 600	850 × 540	768 × 1024
Time consumption (in seconds)	2.851893	8.081258	12.761	23.962164

uses dark channel prior, guided filter, and edge collapse repair with patch size of 15 × 15 and airlight of top 0.1%. It can be observed that Nascimento et al. [5] method is subset of edge collapse based method [6] applied to underwater, where the method [6] uses method [5] and edge collapse repair.

From the above table, it can be seen that the value of result evaluation methods produced by method [6] with no edge collapse repair (same as Nascimento et al. [5]) is lesser. Therefore, to improve Chen et al. [6] method patch size of 25 × 25 and airlight of top 3% is considered. So, the improved method uses underwater dark channel prior, guided filter, and edge collapse repair with patch size of 25 × 25 and airlight of top 3%.

Inference from the above two comparison
The above two comparisons prove that the improved method produce higher quality images compared to other two dehazing algorithms. So the improved method is an efficient dehazing algorithm.

4.2 Processing Speed

The below table indicates the processing speed for the improved dehazing method for different sized images. From the table, it can be concluded that the algorithm is time-efficient (Table 4).

5 Conclusion

In this work, a new dehazing approach to restore quality of image for outdoor and underwater image is presented. The dark channel prior is applied for estimation of Haze thickness with a patch size of 25 × 25 and uses guided filter to refine Haze thickness. Then edge collapse repair is used for accurate estimation of Haze thickness. The principle considers top 3% of brightest pixels for airlight estimation. This principle is applied to underwater images to restore the quality of the image. The simulation results show that the present approach is efficient compared to previous dehazing algorithms.

References

1. Chenault DB, Pezzaniti JL (200) Polarization imaging through scattering media. In: Proceedings of SPIE, vol 4133
2. Narasimhan SG, Nayar SK (2003, June) Contrast restoration of weather degraded images. In: IEEE transactions on pattern analysis and machine intelligence, vol 25, issue no 6
3. Tan R (2008, June) Visibility in bad weather from a single image. In: Proceedings of IEEE conference computer vision and pattern recognition
4. He K, Sun J, Tang X (2011, December) Single image haze removal using dark channel prior. IEEE Trans Pattern Anal Mach Intell 33(12):2341–2353
5. Zhang Q, Li X (2015) Fast image dehazing using guided filter. In: Proceedings of ICCT 2015
6. Chen B-H, Huang S-C (2016, September) Edge collapse-based dehazing algorithm for visibility restoration in real scenes. IEEE J Disp Technol 12(9)
7. Chao L, Wang M (2010) Removal of water scattering. In: 2010 2nd international conference on computer engineering and technology, vol 2, pp 35–39
8. Carlevaris-Bianco N, Mohan A, Eustice R (2010) Initial results in underwater single image dehazing. In: Oceans 2010 Mts/IEEE Seattle. IEEE, pp 1–8
9. Drews P Jr, do Nascimento E, Moraes F, Botelho S, Campos M (2013, December) Transmission estimation in underwater single images. In: 2013 IEEE international conference on computer vision workshops, pp 825–830
10. Cheng CY, Sung CC, Chang HH (2015, October) Underwater image restoration by red-dark channel prior and point spread function deconvolution. In: 2015 IEEE international conference on signal and image processing applications (ICSIPA), pp 110–115
11. Łuczynski T, Birk A (2017) Underwater image haze removal with an underwater-ready dark channel prior. 978-0-692-94690-9, MTS
12. Koschmieder H (1924) Theorie der Horizontalen Sichtweite. Beitr Phys Freien Atmosphare 12:171–181
13. He K, Sun J, Tang X (2013, June) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 35(6):1397–1409
14. Levin A, Lischinski D, Weiss Y (2008, February) A closed-form solution to natural image matting. IEEE Trans Pattern Anal Mach Intell 30(2)
15. Memon F, Unar MA, Memon S (2015, October) Image quality assessment for performance evaluation of focus measure operators. Mehran Univ Res J Eng Technol 34(4). ISSN 0254-7821

A Novel Approach to Ultrasound Image Thresholding Using Phase Gradients

Revathy Sivanandan and J. Jayakumari

Abstract Apart from being simple and cost-effective compared to other imaging techniques, ultrasound (US) imaging provides a non-invasive, ionization-free approach for diagnosis of tumours. However, US images are inherent with blurring and granular speckle noise which poses as a problem for effective diagnosis. The images also contain intensity inhomogeneities which make detection of tumours (similar to surrounding tissues) difficult. The commonly used segmentation approach is thresholding with Otsu thresholding being mostly used. However, due to intensity of inhomogeneities, normal non-tumour regions are also categorized as tumour regions in standard Otsu method. In this paper, we adapt a new method for thresholding of ultrasound images using phase gradients. Phase gradients have been used in microscopy techniques to enhance the visibility of low contrast structures within the cell body. As a pre-processing step, the image is enhanced using neutrosophic image enhancement and despeckled using shearlets. Neutrosophic approach is used considering the fuzzy nature of ultrasound images. The enhanced despeckled image is thresholded using an adaptive Otsu thresholding where the constraints for thresholding is derived also from texture, gradient, phase and phase gradient apart from mean and between class variance. The performance was evaluated using metrics like SSIM, MSE, SNR and accuracy revealing that the proposed method shows better results than the conventional Otsu thresholding. The proposed method was tested on US images with different echogenecities (hypoechoic, hyperechoic, anechoic, isoechoic) and showed promising results in identifying the tumour regions.

Keywords Ultrasound · Thresholding · Neutrosophic image · Shearlet · Gabor filter · Phase gradient

R. Sivanandan (✉) · J. Jayakumari
Department of ECE, Mar Baselios College of Engineering and Technology, APJ Abdul Kalam Technological University, Trivandrum, Kerala, India
e-mail: revathysivan.official@gmail.com

J. Jayakumari
e-mail: jayakumari.j@mbcet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_7

1 Introduction

US imaging is emerging as the primary diagnostic tool for detection of tumours. Its features include non-invasiveness, pain-free and less expensive than tomography or resonance imaging, ionisation-free, ease of applicability, etc. However, the images contain multiplicative noise component called speckle (resulting from backscattered echoes) along with Gaussian noise and blurring from the underlying US system which affects the resolution and contrast quality [1]. To model the US imaging system, a 2D convolution model employing tissue reflectivity (TR) image and system point spread function (PSF) is commonly used [2]. For an US imaging system, let u be the original TR-US image, H be the blurring/ convolution operator resulting from the PSF, η_1 be an additive Gaussian noise, η_2 be an multiplicative noise and f be the obtained degraded/ noisy image which satisfies the formulation:

$$f = H \otimes u.\eta_2 + \eta_1 \tag{1}$$

Since H is introduced by the US system, most literature does not consider it into effect and the effect of additive noise is less than that of the multiplicative one, the model thus reduces to:

$$f = u.\eta_2 \tag{2}$$

In short, u is affected by the blurring operator H and is then contaminated by multiplicative specular noise η_2. The objective of US restoration is to estimate u from f. Generally, there are two types of approaches for despeckling of US images. The first type concentrates on filters that perform on US images directly like Lee filter [3], Kuan filter [4], SRAD filter [5], wavelet filter [6], etc. The second type works in the homomorphic cepstral domain by converting the speckle to an additive noise through logarithmic transformation and further filtering in the log domain [7]. The final restored image is obtained back by exponential transformation. By applying log-transformation on Eq. 2,

$$\log(f) = \log(u) + \log(\eta_2) \geq g = v + n \tag{3}$$

where $g = \log(f)$, $v = \log(u)$ and $n = \log(\eta_2)$. The probability density function of n for a shape parameter α is given by Rayleigh distribution as:

$$\mathrm{pdf}(n) = \frac{n}{\alpha^2}\exp\left(\frac{-n^2}{2\alpha^2}\right) \tag{4}$$

It is necessary to understand what the obtained images represent and to clearly delineate the required region of interest. In US for tumour detection, the underlying tumours show various echogenecities [8], viz. isoechoic (tumour intensities similar

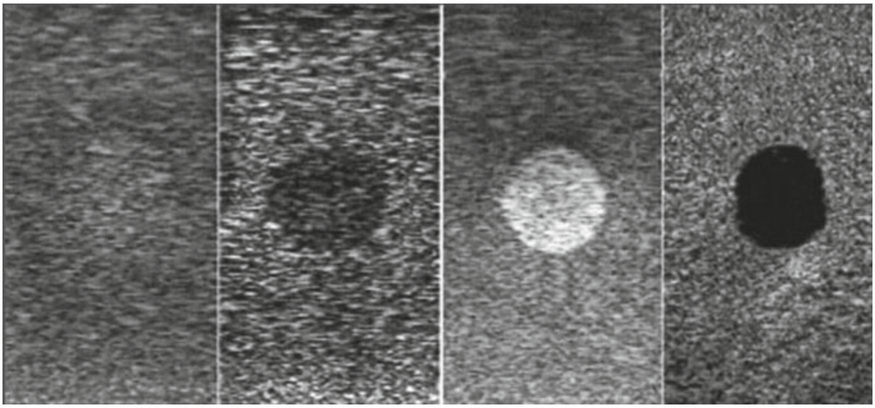


Fig. 1 Demonstration of various US echogenecities corrupted with specular noise (from left to right, isoechoic, hypoechoic, hyperechoic and anechoic)

to surrounding tissue intensities), hypoechoic (tumour intensities different than surrounding tissue intensities), hyperechoic (tumour showing bright white intensities) and anechoic (tumour showing black intensities) as shown in Fig. 1 [9]. Among these, isoechoic and slightly hypoechoic tumours are difficult to be detected and require adept image processing.

Methods dealing in grouping image pixels into labelled regions are broadly called segmentation, and among those, thresholding is labelling into two groups (not exclusively) where one represents the objects of interest and the other the non-objects or background. There is no method as such which deals with thresholding of US images with echogenicity variations. Methods of thresholding can be categorized on the basis of histograms, clusters, entropy, attributes, objects, etc. Since tumour regions exhibit higher entropies or randomness, most literature focuses on methods employing entropy.

Kapur et al. [10] uses entropy-based algorithms which conclude that the threshold value is obtained when the total of background and foreground object entropies reaches a maximum value. Abutaleb et al. [11] extended the work using 2D entropies calculated from 2D histograms, while He et al. [12] used 2D histograms and multiresolution analysis. Ordinary thresholding performs poorly for noisy and inhomogenous data for which fuzzy set theory can be used to handle image vagueness. Prasad et al. [13] used fuzzy C-partitions based on entropic criteria to obtain optimum thresholding while Lopes et al. [14] employed fuzzy techniques for thresholding without any entropy function so as to reduce time complexity. Li [15] proposed a thresholding function based on artificial bee colony algorithm which finds the threshold based on the maximum entropy. Most of these methods fail to give satisfactory results in the presence of noise.

The widely used Otsu thresholding [16] is an automatic optimal global thresholding used for its effectiveness, simplicity and low computational complexity and is based on features obtained from intensity histograms. But while separating the foreground object of interest, most of the unwanted background is still retained in the results. To refine this, Liu et al. [17] adapted a two-step modified Otsu-based thresholding to constrain the obtained range of thresholds so as to extract the object. And to further refine the results, morphological operations (dilation using a disc structuring element of radius 9 and erosion using a disc structuring element of radius 3) were performed on the results of the previous step. It can be seen that a fixed morphological operation after Otsu thresholding cannot be applied to arbitrary US images where the tumour can be of any size or echogenecity.

In the present work, we present a simple yet effective procedure to threshold US images using constraints derived using mean, gradient, texture and phase gradients on the image. It is shown that the proposed method works even in US images with cases of posterior acoustic shadowing (PAS). We do not suggest that the proposed method is superior to other thresholding methods, just that in US images with varying echogenecities, the proposed method will yield less unwanted regions (foreground objects) as compared with entropy thresholding or Otsu thresholding. We also make a contribution regarding the pre-processing of US images using neutrosophic approach which will be explained later.

We briefly summarize the remaining contents: In Sect. 2, we present the pre-processing steps adopted in this work and the proposed thresholding method. In Sect. 3, the results and the performance are discussed. In Sect. 4, we make the concluding remarks and scope for future research.

2 Background and Methodology

This section highlights and describes the proposed scheme of the method which includes US image pre-processing (neutrosophic enhancement and shearlet despeckling) and US thresholding based on phase and phase gradient, and the included subsections describes the different steps of the procedure.

2.1 Pre-processing: Neutrosophic Enhancement

Neutrosophy, a branch of philosophy introduced by Florentin Smarandache in 1980, studies neutralities and indeterminacies and defines as "a concept "*A*" in relation to its opposite, "*Anti-A*" and that which is not A, "*Non-A*", and that which is neither "*A*" nor "*Anti-A*", denoted by "*Neut-A*"". Transformation of an image into neutrosophic domain is to separate it into three layers or three images, namely foreground objects, boundary objects and background [18]. Due to the fuzziness or randomness exhibited by US images, neutrosophic-based processing can yield significant results.

Accordingly, "a neutrosophic set "*A*" in X (a space of points or objects with an element denote as x) given as $A = \{(x, T(x), I(x), F(x))|x \in X\}$ is characterized by a truth membership function T, an indeterminacy membership function I and a ¼

F and $T(x)$, $I(x)$, $F(x)$ vary in $(-0, 1+)$. A pixel in the neutrosophic domain can be represented as $p(T, I, F)$, meaning the pixel is $t\%$ true in the bright pixel/foreground set, $i\%$ indeterminate/edge and $f\%$ false/background, where t varies in T, i varies in I and f varies in F, respectively". A Cartesian image coordinate pixel $g(i, j)$ maps to neutrosophic domain as $g(i, j) = \{T(i, j), I(i, j), F(i, j)\}$ where

$$T(i, j) = \frac{\bar{g}(i, j) - \bar{g}_{\min}}{\bar{g}_{\max} - \bar{g}_{\min}} \tag{5a}$$

$$I(i, j) = \frac{\delta(i, j) - \delta_{\min}}{\delta_{\max} - \delta_{\min}} \tag{5b}$$

$$F(i, j) = 1 - T(i, j) \tag{5c}$$

and

$$\bar{g}(i, j) = \frac{1}{w * w} \sum_{m=i-w/2}^{i+w/2} \sum_{n=j-w/2}^{j+w/2} g(m, n) \tag{6a}$$

$$\bar{g}_{\max} = \max(\bar{g}(i, j) \tag{6b}$$

$$\bar{g}_{\min} = \min(\bar{g}(i, j) \tag{6c}$$

$$\delta(i, j) = abs(g(i, j) - \bar{g}(i, j)) \tag{7a}$$

$$\delta_{\max} = \max(\delta(i, j)) \tag{7b}$$

$$\delta_{\min} = \min(\delta(i, j)) \tag{7c}$$

Hence, the image becomes a 3D matrix in the neutrosophic domain, and the individual components for an US tumour image are given in Fig. 2. To summarize, the conversion of an image to neutrosophic image is as:

1. Read the image.

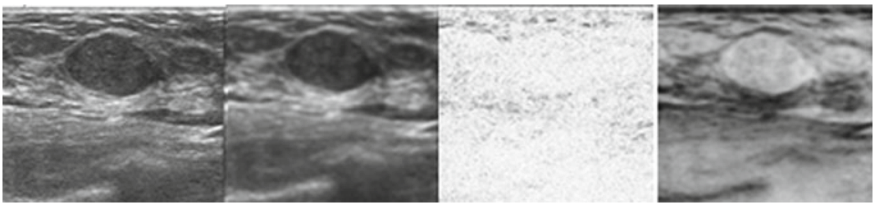


Fig. 2 (From left to right) original US image, truth (T) image, indeterminate (I) image and false (F) image; I image shows the speckles present in the US image

2. For each pixel, compute its local mean intensity and its maximum and minimum values (Eqs. 6a–6c).
3. Compute absolute divergence between each pixel intensity and its local mean intensity and its maximum and minimum values (Eqs. 7a–7c)
4. Construct the truth matrix T, indeterminate matrix I and the falseness matrix F (Eqs. 5a–5c).

After the image has been converted to the neutrosophic domain, enhancement is performed on each matrix individually [19]. The enhancement method is twofold, i.e., on one hand, the method tries to denoise the image, and on the other, the image contrast is improved. To do so, firstly a power law transformation (gamma correction) is performed on the truth matrix. Secondly, a Gaussian filter is used on the indeterminate matrix to remove any additive noise and enhance the edges. Thirdly, the falseness matrix is subjected to a logarithmic transformation to enhance its corresponding details.

$$\bar{T}(i, j) = cT^{\gamma}(i, j) \tag{8a}$$

$$\bar{I}(i, j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(I(i, j) - \mu)^2}{2\sigma^2}\right) \tag{8b}$$

$$\bar{F}(i, j) = c \ln(1 + F(i, j)) \tag{8c}$$

The neutrosophic image entropy is defined as summation of entropies of the three sets. It is used to quantify the degree of indeterminacy in images. Strength of the correlation between truth and false sets with indeterminate set are influenced by the distribution of the pixels and the entropy of the indeterminate set. The enhancement is iterated till the entropy associated with indeterminate set is less than a predefined threshold. Finally, the enhanced truth image is converted back to the Cartesian image. A pixel in Cartesian domain is given as:

$$f(i, j) = f_{\min} + (f_{\max} - f_{\min})\bar{T}(i, j) \tag{9}$$

where $f_{\min}$ and $f_{\max}$ are the minimum and maximum values in the set $\bar{T}$.

The specular noise associated with US imagery is multiplicative in nature and follows Rayleigh or Gamma distribution [20]. Hence, we use a Rayleigh model on the indeterminate set (Eq. 10) for noise removal rather than Gaussian filter, and by implementing so, we have obtained improved signal-to-noise ratios.

$$\bar{I}(i, j) = \frac{I(i, j)}{\sigma^2} \exp\left(\frac{-I(i, j)^2}{2\sigma^2}\right) \tag{10}$$

The iteration threshold is chosen as a small value <1. Smaller the value, better the enhancement but at increased computation time. So as a compromise, we choose the iteration threshold that optimizes the time, but this may result in some speckle not being eliminated. Therefore, the transformed enhanced image is further despeckled using shearlets to remove any more existing noise which may degrade our thresholding efficiency.

2.2 Pre-processing: Shearlet Despeckling

Shearlets can effectively capture the geometry of images and are more efficient for representing images with edges. They are a multiscale framework that allows encoding anisotropic features than compared with wavelets and preserves significant detail information in despeckled outputs [21, 22]. In US imagery, detail preservation is necessary and shearlet basis functions are supported in the frequency domain, and thus, finer details like texture information can be retained after restoration.

The discrete shearlet system [23] is defined as:

$$\mathrm{SH}(\psi) = \left\{\psi_{j,k,m} = |\det A|^{j/2}\psi(S^k A^j - m) : j, k \in Z, m \in Z^2\right\} \tag{11}$$

where $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$ is the anisotropy matrix for multiscale partitions/resolution change, $S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ is the shear matrix for directional analysis/orientation change and *a, s, t* are the scale, orientation and location variables, respectively. The discrete shearlet transform can be denoted as:

$$SH_\psi f(j, k, m) = \left\langle f, \psi_{j,k,m} \right\rangle \tag{12}$$

Shearlet transform implies filtering the signal in a pseudo polar grid and further filtering by 1D band-pass filter banks without further sampling operations. Discrete Shearlet transform [23] is a combination of Laplacian pyramid (LP) and directional filter and non-subsampled shearlet transform (NSST) [24]. Denoising applications uses non subsampled LP (NSLP) [24] rather than LP which can improve its effectiveness. NSLP analysis [25, 26] is done iteratively as:

$$\mathrm{NSLP}_{j+1} f = \left(\mathrm{Ah}_j^1 \prod_{k=1}^{j-1} \mathrm{Ah}_k^0\right) f \tag{13}$$

where f is the image, $\mathrm{NSLP_{j+1}}$ is detail coefficients at scale $j + 1$, and $Ah_k^0$and Ah_j^1 are low and high pass filters at scale j and k.

NSST algorithm for a resolution scale j and number of directions D_j can be summarized as:

1. Decompose f into a low pass image f_a^j and a high pass image f_d^{j} using NSLP

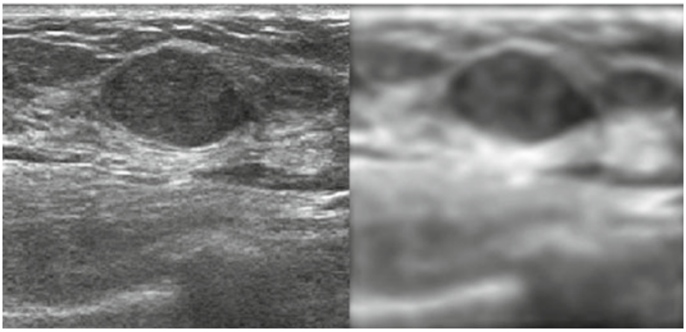


Fig. 3 (From left to right) original US image and the enhanced and shearlet denoised US image

2. $\widehat{f}_d^{\,j}$ is computed in pseudo polar grid to get Pf_d^j
3. Band-pass filtering is applied to Pf_d^j to obtain $\left\{\widehat{f}_{d,k}^{\,j}\right\}_{k=1}^{D_j}$
4. Inverse FFT is applied in pseudo polar grid to obtain NSST coefficients $\left\{\widehat{f}_{d,k}^{\,j}\right\}_{k=1}^{D_j}$

For our work, we have considered the log transform method (Eq. 2) for multiplicative noise (speckle) removal similar to [27]. Obtain an estimate of f from the noisy image g using soft thresholding on NSST coefficients of g. The threshold levels are given by $\tau_{j,k} = c_j \sigma_{\gamma_{j,k}}$ where the noise standard deviation at scale j and shear directional band k is given as $\sigma_{\gamma_{j,k}}$ and scaling parameter is given as c_j. Five levels of the NSLP decomposition is applied, viz. eight 32×32 shear filters is used for the first two coarser scales and sixteen 16×16 shear filters is used for the third and fourth finer scales and so on. The denoised image is converted back using exponential transformation. A simulation result of such a denoised US image is given in Fig. 3. The denoised enhanced image is then passed through Gabor filters for extraction of textural features which is used as one of the constraining parameters of our thresholding algorithm.

2.3 Gabor Filter for Texture Extraction

Gabor filters are extensively used in areas of texture-based segmentation, feature extraction and classification [28, 29]. A Gabor function can be defined as a sinusoidal modulated Gaussian with a spread of σ_x and σ_y in the x and y directions and a modulating frequency of u_0. Its impulse response is given as:

$$h(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right\} \cos(2\pi u_0 x) \tag{14}$$

Equation 14 shows an orientation of 0° w.r.t x-axis. Arbitrary rotations of the filter can be obtained by rotating the function. The frequency u_0 and the rotation angle θ define the filter centre location. By tuning the two values to different centre locations, multiple filters can be created that span the entire domain.

Gabor wavelets can be utilized to extract texture features where a mother wavelet (Eq. 15) can be used to generate a set of wavelets and the whole image is given as input to the wavelet set [30, 31] which is characterized by the choice of proper set of frequencies and orientations that covers the domain and captures texture information as much as possible.

$$\phi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y}\left\{\frac{-1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j U_h x\right\} \tag{15}$$

$$h(x, y) = \left(\frac{U_h}{U_l}\right)^{\frac{-s}{s-1}} \phi(X, Y) \tag{16}$$

$$X = \left(\frac{U_h}{U_l}\right)^{\frac{-s}{s-1}} \left[(x - x_0)\cos\left(\frac{d\pi}{N_d} + (y - y_0)\sin\left(\frac{d\pi}{N_d}\right)\right)\right. \tag{17a}$$

$$Y = \left(\frac{U_h}{U_l}\right)^{\frac{-s}{s-1}} \left[-(x - x_0)\sin\left(\frac{d\pi}{N_d} + (y - y_0)\cos\left(\frac{d\pi}{N_d}\right)\right)\right. \tag{17b}$$

Here, $s = 1,2, \ldots, N_s$ is scaling parameter, and $d = 1,2, \ldots, N_d$ is direction parameter of wavelets; (x_0, y_0) is filter centre in spatial domain, and U_h and U_l are maximum and minimum values of centre frequencies. In our work, we have chosen four scales each with eight different orientations to capture as much echogenic texture from the denoised images as possible.

2.4 Thresholding

A summary of the widely used optimal global thresholding, Otsu method [16] is explained. For an image of L grey levels, n_i denotes number of pixels with grey value i and total number of pixels $N = \sum_{i=0}^{L-1} n_i$. The probability density distribution of i is given as $p_i = \frac{n_i}{N}$ and $0 \leq p_i \leq 1$. The image pixels can be divided into F (foreground) and B (background) classes by specifying a threshold τ. Then, F specifies the pixels within $[0, 1, \cdots, \tau]$, and B represents the pixels $[\tau + 1, \tau + 2, \ldots, L - 1]$. The mean grey level, the between class variance and the optimal threshold are given as:

$$\mu_T = \sum_{i=0}^{L-1} i p_i \tag{18}$$

$$\sigma^2 = \varpi_0(\mu_0 - \mu_T)^2 + \varpi_1(\mu_1 - \mu_T)^2 \tag{19}$$

$$\tau_{\text{opt}} = \arg\max \sigma^2 \tag{20}$$

Since US images have poor resolution and contrast, Otsu thresholding will not yield satisfactory results for tumour region thresholding, despite of contrast adjustment operations. Experimentally, Otsu method could not distinguish between tumour region and certain surrounding tissues of similar echogenecity/ intensities.

For effective thresholding, we adopted multiple thresholds each based on Otsu method and applied this method on images derived as a result of texture, phase, gradient and phase gradient operations on the enhanced despeckled image. Frequency domain operations are offering an alternative approach in image enhancement, denoising, etc., where the image is represented in terms of its magnitude and phase spectra. In such cases, the phase of an image contains important information regarding the edges and avoiding the phase during FFT reconstruction which will result in significant information loss. Although not essentially the same, the concept of phase gradients is used in (Hoffman nodulation contrast) microscopy to enhance or highlight transparent/ translucent details that are embedded within the cell [32]. We have extended this terminology to detect similar tumour echogenecities to that of surrounding tissues and thus refine the delineation of our region of interest. The steps can be briefed as:

1. Compute the simple Otsu global threshold, say *thresh1*.
2. Compute the global threshold based on its Gabor texture image, say *thresh2*.
3. Transform the image to its Fourier domain and then reconstruct the image using phase information alone. Compute the global threshold of this reconstructed phase image, say *thresh3*. Also, compute the threshold of the gradient of phase reconstructed image (phase gradient), say *thresh4*.
4. Label the image as foreground pixels only when the image pixels are less than the minimum of *thresh1*, *thresh2*, *thresh3* and *thresh4*. Rest of the pixels is labelled as background.

2.5 Proposed Work

The proposed work aims at improving the thresholding of US images with tumours irrespective of their echogenecities and is compared with the standard Otsu optimal global thresholding. The work was applied on 40 B-mode US images (18 hypoechoic, 8 isoechoic, 8 anechoic and 6 hyperechoic tumours each) from www.ultrasoundcases.info. The proposed thresholding algorithm was implemented by using MATLAB (R2017b) on a system having clock speed of 2 GHz and with 4 GB memory. Our proposed workflow can be briefed as:

1. Pre-processing
   Input-Noisy US image with tumour

- Do neutrosophic processing and enhancement on the noisy image to obtain a denoised image focusing on region of interest (tumour/foreground)
- Do further denoising using NSST

Output-Denoised US image

2. Extraction of relevant features from texture, gradient and phase
Input-Denoised US image

- Perform Gabor filtering at different scales and orientations to capture maximum information regarding tumour texture
- Compute the gradient of the image
- Transform the image to its Fourier domain and then reconstruct the image using phase information alone. Also, calculate the gradient of this image also.

3. Thresholding

- Compute the individual grey level thresholds of the feature images obtained in the previous step.
- Also, compute the global Otsu threshold.
- Obtain a suitable common threshold; say the minimum of the obtained thresholds.
- Threshold the denoised image using the above constraint.

Output-Thresholded binary image

3 Experimental Results and Evaluation

The results show significant improvement in thresholding using the proposed methodology. Using MATLAB, Otsu method and proposed method were applied on US images of different echogenecities, and the simulation results are shown in Figs. 4,

Fig. 4 (From left to right) original hyperechoic tumour US image, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

5, 6, 7 and 8. The improvement can be understood visually itself which is further substantiated by the quantitative measures shown in Figs. 9, 10 and 11. In most cases, for simple Otsu thresholding, the tumour region was connected to the neighbouring tissue regions, thus reducing the thresholding efficiency.

To compute the differences between the optimal Otsu thresholding and our proposed thresholding, we have evaluated few quantitative measures, namely

- SSIM: structural similarity index, a quality assessment index, is calculated between simple Otsu threshold and ground truth and proposed method and ground truth as defined in [33]. SSIM is comparably higher for the proposed method.

Fig. 5 (From left to right) original anechoic tumour US image, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

Fig. 6 (From left to right) original hypoechoic tumour US image, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

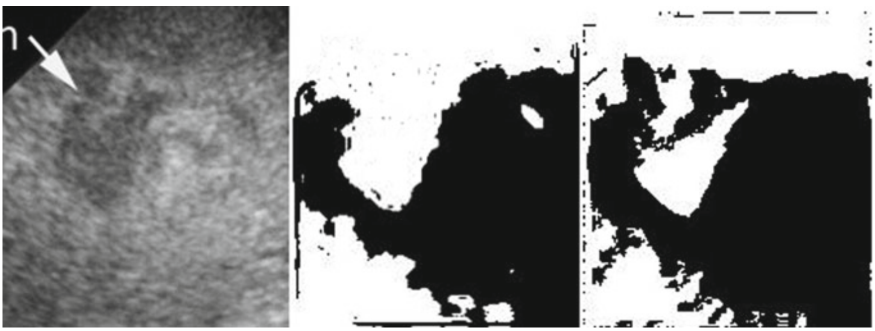


Fig. 7 (From left to right) original slightly hypoechoic/marginally isoechoic tumour US image, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

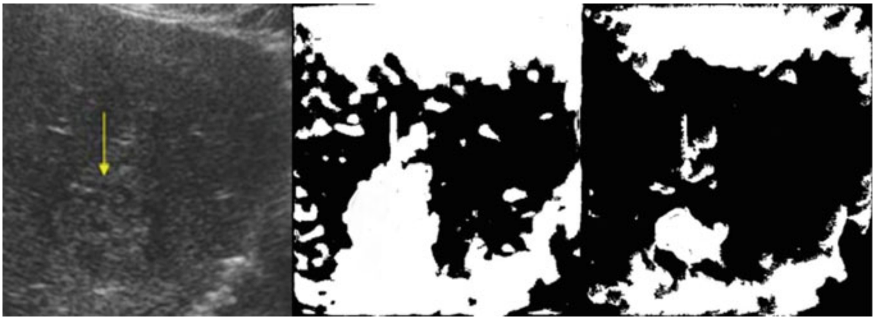


Fig. 8 (From left to right) original isoechoic tumour US image, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

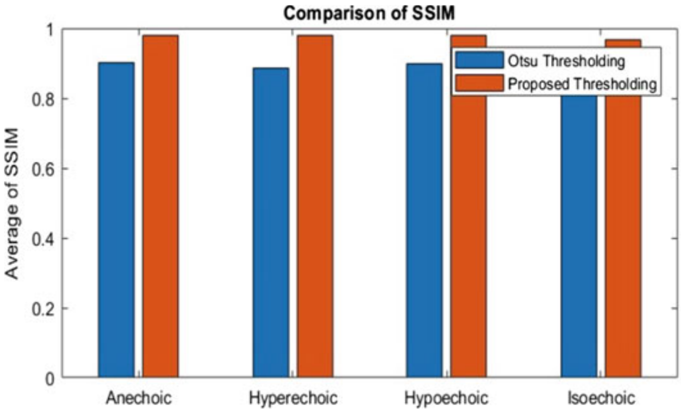


Fig. 9 Comparison of SSIM between simple Otsu threshold and proposed method

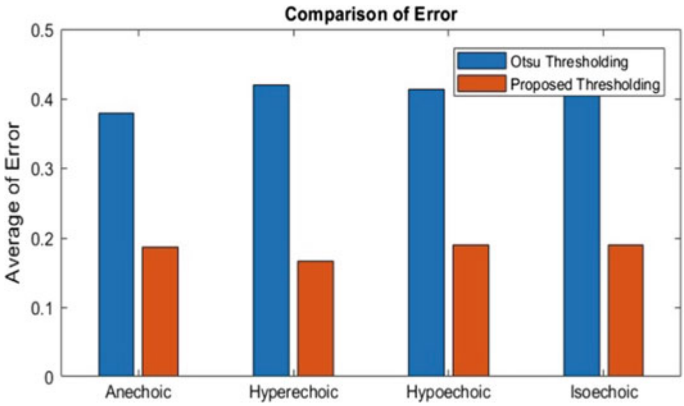


Fig. 10 Comparison of MSE between simple Otsu threshold and proposed method

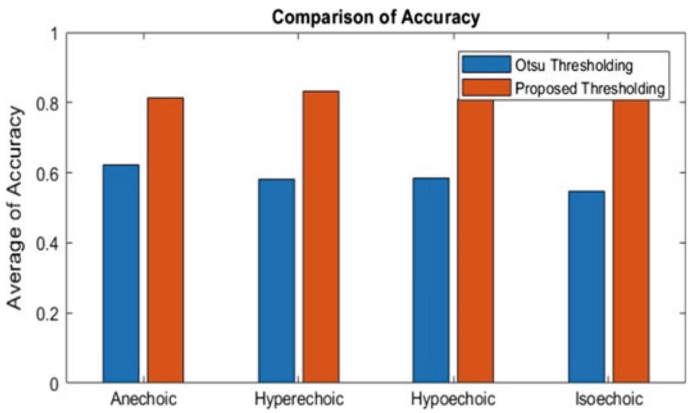


Fig. 11 Comparison of accuracy between simple Otsu threshold and proposed method

- MSE: mean square error between each of the two thresholding methods compared with the ground truth yielded lower error for the proposed method.
- Accuracy of each thresholding method compared with the ground truth. Proposed method yields better accuracy.

Numerical evaluations based on above measures is presented for Otsu method and proposed method. A visual evaluation of proposed method with other existing methods for a tumour with complicated boundaries is shown in Fig. 12. The proposed method yields better thresholding than the compared methods (mean thresholding, maximum entropy thresholding, Otsu thresholding).

PAS is an US image artefact observed for highly malignant tumours, where the tumour tissues produce a certain shadow effect on the neighbouring regions. Existing literature seldom takes these PAS cases into effect, and existing thresholding methods incorrectly label the shadow region as the tumour too. In our method, we are able to refine this problem since the shadows differ in texture, phase and gradients even though they may have the same intensities as the neighbouring tumour and is evident from a simulation result shown in Fig. 13.

The neutrosophic enhancement explained in [18] models the noise as additive Gaussian alone. We have extended the modelling of the indeterminate image to Rayleigh model as well to handle the effects of speckle noise. Significant improvement in SNR is obtained in doing this and is shown in Fig. 14.

4 Conclusions

An adaptive Otsu thresholding method using texture, phase and phase gradients along with between class variance for US image thresholding is presented. The obtained US image is neutrosophically enhanced and denoised using shearlets during

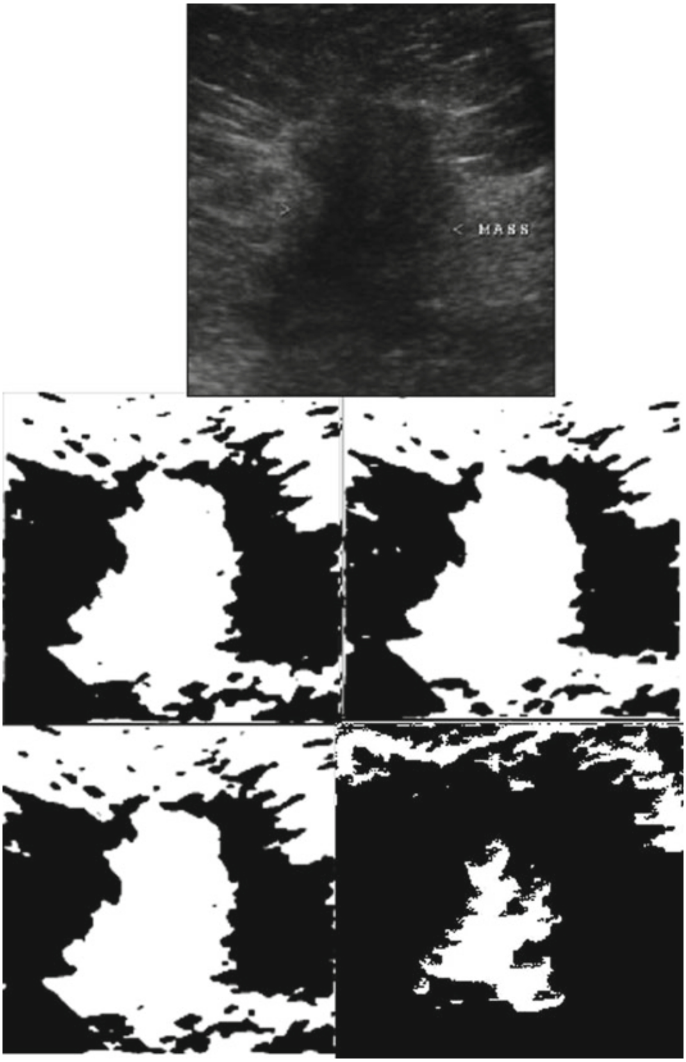


Fig. 12 Comparison of different thresholding methods, (from top, clockwise) original US image with complicated boundaries and shadows, maximum entropy thresholding, proposed method, Otsu thresholding and mean thresholding

pre-processing. Significant Gabor features are obtained from the denoised image. Otsu-based thresholding is done where the thresholding constraints are derived from variances, texture, phase and phase gradients. The results highlight improvement over present optimal global thresholding methods. The proposed thresholding method can yield better mask initialization outputs for segmentation purposes. The features include less morphological operations to refine the thresholded output for mask

Fig. 13 (From left to right) US image showing PAS, thresholded image using Otsu thresholding, thresholded image using proposed thresholding

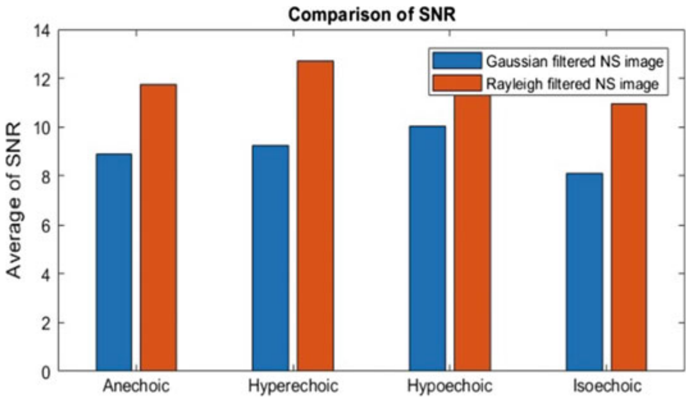


Fig. 14 Comparison of SNR between neutrosophic enhancement using Gaussian and Rayleigh modelling of noise

creation and work even in cases of US images with PAS and iso/ slightly hypo echoic tumours. The proposed method guarantees to help in diagnostic procedures where the radiologist can decide the region of interest with minimum processing. The future work aims at using this proposed algorithm for initial mask creation which can be used for segmentation of tumours using deformable models.

References

1. Godman JW (1976) Some fundamental properties of speckle. J Opt Soc Am 66(11):1145–1149
2. Ma J, Dimet Le (2009) Deblurring from highly incomplete measurements for remote sensing. IEEE Trans Geosci Remote Sens 47(3):792–802
3. Lee JS (1981) Refined filtering of image noise using local statistics. Comput Graphic and Image Processing 15:380–389
4. Kuan DT, Sawchuk AA, Strand TC, Chavel P (1987) Adaptive restoration of images with speckle. IEEE Trans Acoust Speech Signal Process 35(3):373–383
5. Yu Y, Acton ST (2002) Speckle reducing anisotropic diffusion. IEEE Trans Image Process 11(11)

6. Ruikar SD, Doye DD (2011) Wavelet based image denoising technique. Int J Adv Comput Sci Appl 2(3):49–53
7. Zong X, Laine AF, Geiser EA (1998) Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. IEEE Trans Med Imaging 17:532–540
8. Baris I, Boezaart AP (2010) Ultrasound: basic understanding and learning the language. Int J Shoulder Surg 4(3):55–62
9. Manoj KK, Wing HK (2019) 43—Ultrasound-guided regional anesthesia. a practice of anesthesia for infants and children, 6th edn, pp 988–1022
10. Kapur JN, Sahoo PK, Wong AKC (1985) A new method for gray-level picture thresholding using the entropy of the histogram. Comput Vis Graph Image Proc 29(3):273–285
11. Abutaleb AS (2009) Automatic thresholding of gray-level pictures using two-dimensional entropy. Computer Vision, Graphics, and Image Processing
12. He C, Xing J, Li J, Yang Q, Wang R (2015) A new wavelet threshold determination method considering interscale correlation in signal denoising. Math Prob Eng 15
13. Prasad MS, Divakar T, Rao BS, Raju N (2012) Unsupervised image thresholding using fuzzy measures. Int J Comput Appl 27(2):32–41
14. Lopes NV (2010) Automatic Histogram Threshold using Fuzzy Measures. IEEE Transactions on Image Processing, vol 19, no 1
15. Li L, Sun L, Guo J, Han C, Zhou J, Li S (2017) A quick artificial bee colony algorithm for image thresholding. Information, MDP
16. Otsu N (1978) A threshold selection method from gray-level histogram. IEEE Trans Syst 8:62–66
17. Liu L, Li K, Qin W, Wen T, Li L, Wu J, Gu J (2018) Automated breast tumor detection and segmentation with a novel computational framework of whole ultrasound images. Med Biol Eng Compu 56(2):183–199
18. Salama AA, Smarandache F (2014) Introduction to image processing via neutrosophic techniques. Neutrosophic Sets Syst 5
19. Salama AA, Smarandache F (2018) Neutrosophic approach to grayscale images domain. Neutrosophic Sets Syst 21
20. Aysal TC, Barner KE (2007) Rayleigh-maximum-likelihood filtering for speckle reduction of ultrasound images. IEEE Trans Med Imaging 26(5):712–727
21. Kutyniok G, Labate D (2009) Resolution of the wavefront set using continuous shearlets. Trans Am Math Soc 361(5):2719–2754
22. Easley GR, Labate D, Colonna F (2009) Shearlet based total variation for denoising. IEEE Trans Image Process 18(2):260–268
23. Dahlke S, Steidl G, Teschke G (2010) The continuous shearlet transform in arbitrary space dimensions. J Fourier Anal Appl 16:340–364
24. Easley G, Labate D, Lim WQ (2008) Sparse directional image representations using the discrete shearlet transform. Appl Comp Harmon Anal 25:25–46
25. da Cunha AL, Zhou J, Do MN (2006) The nonsubsampled contourlet transform: theory design and applications. IEEE Trans Image Process 15(10):3089–3101
26. Hou B, Zhang X, Bu X, Feng H (2012) SAR image despeckling based on nonsubsampled shearlet transform. IEEE J Sel Top Appl Earth Observations Remote Sens 5(3)
27. Abazari R, Lakestani M (2019) Non-subsampled shearlet transform and log-transform methods for despeckling of medical ultrasound images. Informatica 30(1):1–19
28. Porat M, Zeevi YY (1998) Localized texture processing in vision: analysis and synthesis in the gaborian space. IEEE Trans Biomed Eng 2:115–129
29. Haralick RM (1979) Statistical and structural approaches to texture. Proc IEEE 67(5):786–804
30. Weldon TP, Higgins WE, Dunn DF (1996) Efficient Gabor filter design for texture segmentation. Pattern Recognit 29(12):2005–2015
31. Mirzapour F, Ghassemian H (2003) Texture feature extraction in satellite images using Gabor wavelets. MVIP

32. Sarafis V (2013) phase imaging in plant cells and tissues. Biomed Opt Phase Microsc Nanoscopy
33. Wang Z, Simoncelli EP, Bovik AC (2004) Multiscale structural similarity for image quality assessment. Conf Sig Syst Comput 2(2):1398–1402

Investigation of Techniques to Recognize Optimal Power Structuring of Vedic Multiplier

P. Anitha and P. Ramanathan

Abstract Low power and high speed digital systems are essential for enhancing battery life of portable devices such as smartphones and digital computers. The integral part of any arithmetic and logic unit is adder. When compared to addition, subtraction and multiplication require more hardware resources and processing time. Low power consumption, delay and process variation parameters need to be taken care while designing the integrated circuit. In our proposed work, improved version of Vedic multiplier is designed and implemented by using CSA based on NEDFF. The proposed design offers low power dissipation and high speed. The power and delay results of existing and proposed multipliers are taken by using micro wind tool with technology of 90 nm. The experimental results signify that proposed Vedic multiplier using a CSA based on NEDFF provides 50% improvement in performance.

Keywords Carry select adder (CSA) · Partial product reduction (PPR) · Vedic mathematics (VM) · Negative edge triggered D flip-flop (NEDFF)

P. Anitha (✉)
Department of ECE, Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India
e-mail: anithangptech@gmail.com

P. Ramanathan
Department of ECE, Madanapalle Institute of Technology and Science, Madanapalle, Andra Pradesh, India
e-mail: pramanathan2509@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_8

1 Introduction

Digital multiplication is the most extensively used operation in signal processing. The demand for higher throughput, increased performance, reduced area and power dissipation are foremost requirement for many applications. Vedic Mathematics—“An ancient Indian System of calculation” rediscovered from Vedas between 1911 and 1918 by Sri Bharati Krishna Tirthaji Maharaj. Urdhwa and Triyakbhyam is one of the multiplication algorithms among sixteen sutra. Vedas embodies the principles of algebra for empirical accuracy. Vedic mathematics becomes very useful for engineering applications such as complex computations and signal processing by ensuring speed and accuracy. Multiplier architecture is sub-divided into three stages. In the first stage, partial products are generated by an AND array. The process of generation of partial product starts from LSB to MSB. The second stage namely partial product compression has critical path delay that decides the overall speed of the multiplier. The complexity of an array multiplier is O(n) whereas that of a tree multiplier is O(log2n), where “n” is number of bits in the operand. Tree algorithm provides high speed multiplication, but wiring overhead is a major problem. CSA reduces considerable amount of power dissipation and wiring overhead. This work is focused on the optimization of partial product addition (PPA) using negative edge triggered D flip-flop based on carry select adder.

2 Literature Review

In the literature, various types of multipliers are reported to obtain low power consumption and performance enhancement. G.R. Gokhale and P.D. Bahisonade have proposed a high-speed Vedic multiplier using binary to excess one converter-based CSA [1]. This multiplier performs better when compared to booth multiplier. Lokesh M. Bhalmeet al proposed 4×4 Vedic multiplier with ripple carry adder having less delay than 4×4 Vedic multiplier realized using 4-bit traditional booths multiplier [2]. CLAs are the fastest adders, but it increases structural complexity. Damarla Paradhasaradhi et al. proposed area efficient square root CSA by sharing the common Boolean logic (CBL) term, and the duplicated adder cells in the conventional CSA are removed [3]. Kumari DP et al. have proposed Switching Transistor based D Flip Flop (STDFF) but, this STDFF uses dual clock pulse generator [4]. Anitha Ponnusamy et al. have proposed partitioned data multiplier using NSDFF, which increase computational complexity [5]. Richa Chauhan et al. have proposed Vedic multiplier using CSA with common Boolean logic which offers high computation speed [6]. Ganesh Kumar et al. proposed modified carry save adder to improve power and delay constraints [7]. High-speed adder offers high computation speed with increased structural complexity. When compared to existing methodologies

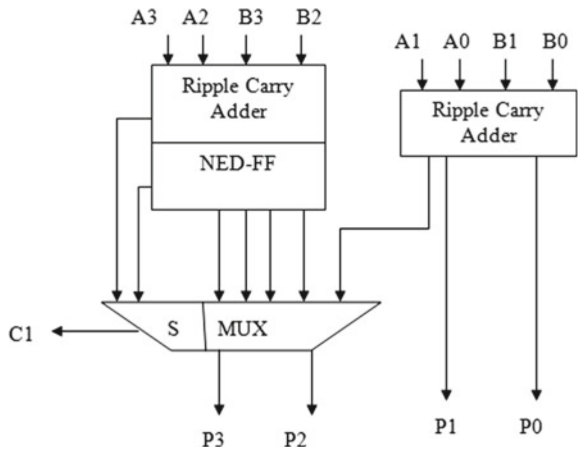


Fig. 1 Block diagram NEDFF-based CSA

such as BEC-based CSA and regular DFF-based CSA, NEDFF has very less structural complexity [8]. The total number of transistors required to construct NEDFF is 8T.

3 Carry Select Adder

Figure 1 shows logic diagram of 4-bit CSA using NEDFF. The CSA using NEDFF offers minimal critical path delay and reduced total power consumption. This CSA performs addition whenever the clock signal of NEDFF goes low. Until that, the output will get simultaneously affected by their applied input.

4 Proposed Vedic Multiplier

The proposed 4 × 4 Vedic multiplier uses "UrdhvaTiryakbhyam Sutra" (algorithm). It is a unique approach. This multiplication algorithm is used to perform multiplication in crosswise and vertically (criss-cross multiplication addition) to improve the computation speed by generating and adding partial products simultaneously. In the existing multiplier, generated partial products are compressed by 4-bit ripple carry adder. The carry propagation delay is the major disadvantage of ripple carry adder. The proposed 4 × 4 multiplier comprises four 2 × 2 multipliers, and three 4-bit RCA is replaced by three 4-bit CSA based on NEDFF. Implementation of the proposed 4 × 4 Vedic multiplier using NEDFF-based CSA is shown in Fig. 4. The computation speed of this algorithm is very high for all types of numerics.

In first stage, the generated partial products of 4 bit input sequence is divided into four 2 bit sequences such as a[1:0] & b[1:0], a[1:0] & b[3:2], a[3:2] & b[1:0] and a[3:2] & b[3:2], respectively. The product is calculated in parallel and crosswise methodology since it requires same quantity of time to calculate the sum of the multiplier. In this multiplier, partial product compression operation in the column is performed by using CSA-NEDFF.

4.1 UrdhvaTiryakbhyam Algorithm

Let us assume A0 = 4, A1 = 3, A2 = 2, A3 = 1. Similarly, B0 = 4, B1 = 3, B2 = 2, B3 = 1. Firstly, least significant bit (LSB) of multiplicant is multipied with least significant bit (LSB) of mulitiplier. Secondly, vertically and crosswise multiplication have been carried out in the following steps 2, 3, 4, 5, 6 and 7, respectively. At same time, generated partial products are added, and then, the carry values will be forwarded to next higher stages (Fig. 2).

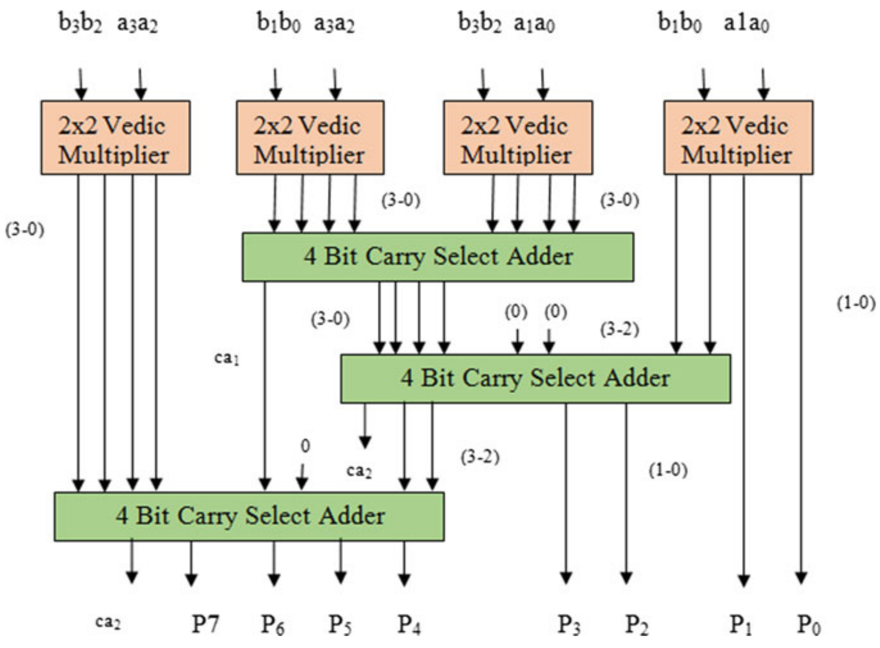


Fig. 2 Block diagram of proposed 4 × 4 Vedic multiplier

Step 1: (A0 * B0) = 4 * 4 = 16
previous carry = 0
Therefore 16 + 0 = 16
S_0 = 6 ; C_0 = 1

Step 2: (A0 * B1) + (A1 * B0) = 12 + 12 = 24
Prevoius carry = 1
Therefore 24 + 1 = 25
S_1 = 5 ; C_1 = 2

Step 3: (A0 * B2) + (B1 * A1) + (A2 * B0) = 8 + 9 + 8 = 25
Previous carry = 2
Therefore 25 + 2 = 27
S_2 = 7 ; C_2 = 2

Step 4: (A3 * B0) + (A2 * B1) + (A1 * B2) + (A0 * B3) = 4 + 6 + 6 + 4 = 20
Previous carry = 2
Therefore 20 + 2 = 22
S_3 = 2 ; C_3 = 2

Step 5: (A3 * B1) + (A2 * B2) + (A1 * B2) = 3 + 4 + 3 = 10
Previous carry = 2
Therefore 10 + 2 = 12
S_4 = 2 ; C_4 = 1

Step 6: (A2 * B3) + (A3 * B2) = 2 + 2 = 4
Previous carry = 1
Therefore 4 + 1 = 5
S_5 = 5 ; C_5 = 0

Step 7: (A3 * B3) = 1 * 1 = 1
previous carry = 0
Therefore 1 + 0 = 1
S_6 = 1 ; C_6 = 0

Final Product is $C_6S_6S_5S_4S_3S_2S_1S_0 = 1522756$.

Line diagram of UrdhvaTiryakbhyam Sutra, generated partial product and reduction of PPA using CSA is shown in Figs. 3, 4 and 1, respectively. The area and overall total power consumption of proposed Vedic multiplier is minimized when compared with existing Vedic multiplier. The length and width of transistors of pull up transistor are increased for power trading. The proposed Vedic multiplier is designed using micro wind tool, and simulation of the multiplier is carried out in 90 nm technology. Implementation of proposed Vedic multiplier using negative triggerd D flip flop is shown in Fig. 6.

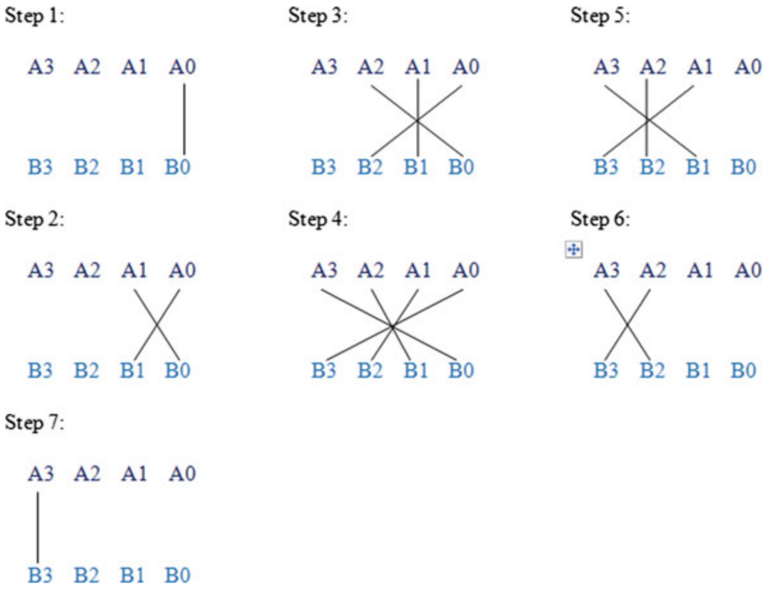
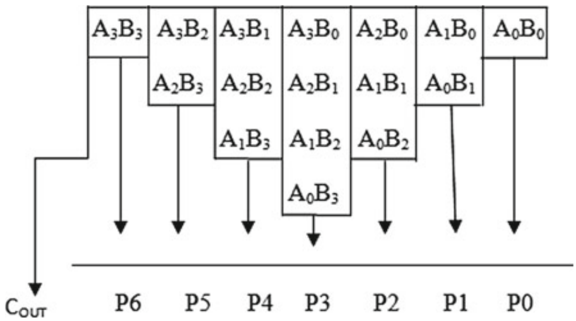


Fig. 3 Line diagram of 4 × 4 Vedic multiplier using UrdhvaTiryakbhyam Sutra

Fig. 4 Generated partial products



5 Results and Discussion

The parametric results of CSA using RCA, carry skip and carry select adder are listed in Table 1. It infers 43.76 and 46.75% of improvement of power delay product (PDP) in NEDFF-based carry select adder when compared to carry skip adder and ripple carry adder-based Vedic multipliers. It is observed that the number of transistors required to implement NEDFF-based CSA is least when compared to other multiplier structures. A comparison graph of existing and proposed multipliers are shown in Fig. 5. The details of transistor utilization in each multiplier are reported in Table 2 Comparison of previous literature results are reported in Table 3.

Table 1 Power analysis result of various multipliers

Types of multiplier design	Parametric analysis		
	Total power consumption (mW)	Delay (ns)	PDP (Ws1)
4 × 4 ripple carry adder-based VM	0.325	3.7	1.202
4 × 4 carry skip adder-based VM	0.214	4.6	0.984
Proposed 4 × 4 Vedic multiplier	0.229	2.3	0.526
Proposed 8 × 8 Vedic multiplier	0.458	4.9	2.242

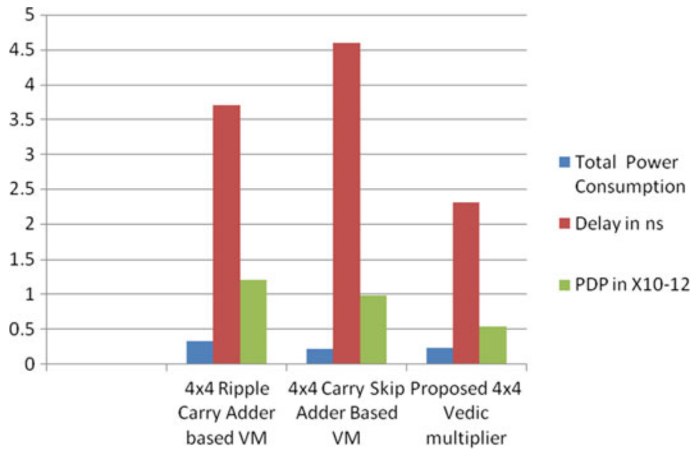


Fig. 5 Comparison graph between proposed and existing multipliers

Table 2 Transistors utilization details

Type of multiplier	Total number of transistors used
4 × 4 ripple carry adder-based VM	240
4 × 4 carry skip adder-based VM	216
Proposed 4 × 4 Vedic multiplier	240
Proposed 8 × 8 Vedic multiplier	232 18 FA + 7 D-FF + 8 Mux (1 FA = 8T, 1 D-ff = 8T, 1 Mux = 4T)

6 Conclusion

In this paper, we presented an optimal power structuring of Vedic multiplier using CSA-NEDFF. The aim is to provide suitable and appropriate multiplier with very low power dissipation for battery life portable devices. The results reveal that the proposed Vedic multiplier using NEDFF-based consumes very less power. This multiplier

Table 3 Comparison of results with previous literatures

Simulation tool	Author	Size of the multiplier	Power (mW)	Delay (ns)
Xilinx 14.1i	Richa Chauhan, M. Zahid Alam	16 × 16	–	30.89
Xilinx 14.1i	G.R. Gokhale, M.S.P.D. Bahirgone	8 × 8	–	45.67
Synopsis	G. Ganesh Kumar, Subhenu K Sahoo	8 × 8	0.358	1.04

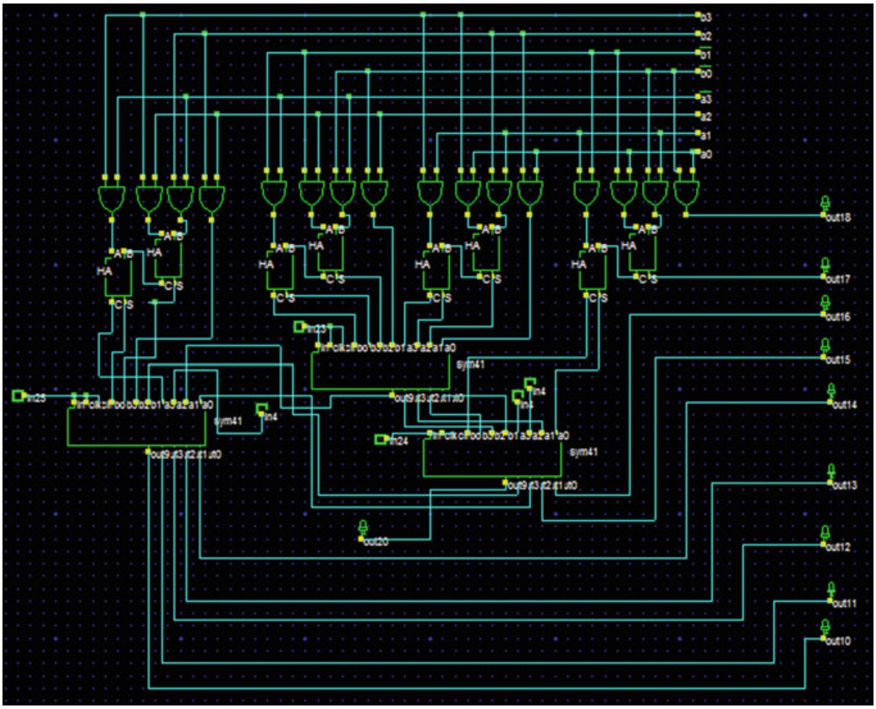


Fig. 6 Implementation of 4 × 4 Vedic multiplier using negative edge triggered D flip-flop

provides better trade-off between power and silicon area. This multiplier eradicates worst case time delay.

Acknowledgements I would like to submit my sincere thanks to the management and our beloved Principal Dr. J. Janet, Sri Krishna College of Engineering and Technology, Coimbatore, for providing necessary support and encouragement.

References

1. Gokhale, GR, Bahisonde PD (2015) Design of vedic-multiplier using area-efficient carry select adder. IEEE international conference on advances in computing communication & informatics, pp 576–581
2. Bhalme LM, Wakpanjar OM., Rakhunde Si (2015) High speed multiplier using vedic mathematics technique. Int J Adv Res Comput Sci Softw Eng 5(2):114–118
3. Paradhasaradhi D, Anusudha K (2013) An area efficient enhanced SQRT carry select adder. J Eng Res Appl 3(6):876–880
4. Kumari DP, Rao RSP, Bhaskar BV (2012) A future technology for enhanced operation in flip-flop oriented circuits. Int J Eng Res Appl 2(4):2177–2180
5. Anitha P, Ramanathan P (2014) Area efficient high speed adder using negative edge triggered D flip flop. Appl Mech Mater 573(187):187–193. https://doi.org/10.4028/www.scientific.net/AMM.573.187
6. Chauhan R, Alam MZ (2016) Design of 16-bit vedic multiplier using Kogge Stone adder for fast parallel FIR filter. Int J Adv Res Comput Sci Softw Eng 6(5):944–947
7. Kumar GG, Sahoo SK (2015) Implementation of a high speed multiplier for high-performance and low power applications. In: IEEE conference https://doi.org/10.1109/ISVDAT.2015.7208143. (978-1-4799-1743-3/15)
8. Edison AJ, Manikandababu CS (2012) An efficient CSLA architecture for VLSI hardware implementation. Int J Manag IT Eng 2(5):610–622

Power Optimization in Application-Based D2D Communication

Sereena Helen Sajeev and Rinju Mariam Rolly

Abstract The 5G networks are in their final stage and are expected to be available soon. Fifth generation has increased the energy consumption in mobile networks in order to meet the demands of users with an increase in carbon footprints. Therefore, energy efficiency in cellular networks is a big concern in order to reduce the overall environmental effects. This paper proposes a green algorithm for energy-efficient networks which encompass D2D transmission initiated between a single transmitter (PT) and many receivers (PR). Three applications are considered: conversational video, conversational voice and text. Optimal power allocation is done in order to meet the target rates of all the demanded applications. The algorithm introduces a parameter known as zone factor for determining the transmission time at each zone for the demanded applications, and this time determines the energy consumed at the user terminal and thus it affects the battery life directly. The efficiency of the green algorithm proposed is validated using MATLAB simulation.

Keywords D2D communication · Energy efficiency (EE) · Battery lifetime · Zone factor · Power optimization

1 Introduction

Digital wireless communication systems are consistently on the mission to fulfill the growing needs of humans with the evolution of 1G, 2G, 3G, 4G and now 5G. Wireless communication networks (WCN) undergo changes continuously to fulfill user demands. The fifth generation (5G) is currently under development and is expected to hit the markets around 2020. Compared to the 4G LTE technology, 5G is

S. H. Sajeev (✉) · R. M. Rolly
Department of Electronics and Communication, Rajagiri School of Engineering and Technology, Ernakulam, Kerala, India
e-mail: sereenahelensajeev@gmail.com

R. M. Rolly
e-mail: rinjum@rajagiritech.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_9

targeted to achieve high speed, low latency and low power. Fifth generation has some potential targets that include 10× peak data rate, increased network capacity, 10× energy efficiency and 10× lower latency [1, 2]. The highlights of 5G increased peak bitrate, high system spectral efficiency, mass connectivity, low battery consumption and better connectivity.

For achieving the 5G goals, network operators and carriers are planning to leverage emerging device-centric systems like device-to-device (D2D) communications, small cells and nano cells [1]. The vision of 5G WCN is an unceasing traffic demand in order to meet the speed demands of the large number of users [3], and it is accompanied with the emergence of multiple technologies in 5G network like spectrum sharing (SS) and D2D communication. D2D supports simultaneous transmissions between users, and it also improves the spectral efficiency and battery lifetime. It has lower latency, and it reduced the power consumed.

The main issue faced by smartphones is corroding of battery life due to high demand of diverse applications. Power consumption should be reduced without failing to meet the user demands. Power consumption can be reduced by transmitting at optimal power for the demanded application so that the user demands are achieved. This paper introduces a power optimization algorithm for reducing the power consumption, and thus, it helps in increasing battery life and makes the network more energy efficient.

2 Related Work

A zone-based green algorithm is proposed in [3] for increasing the battery lifetime in spectrum sharing networks using D2D communication. The proposed algorithm known as ZBGrEEn is for energy efficiency (EE) improvement and to improve battery life. ZBGrEEn considers a single primary transmitter (PT) and multiple primary receivers (PR) placed in different zones. If the PT cannot meet the target rates over the D2D links, then it transmits cooperatively through an appropriate secondary transmitter (ST) in the SS network.

In [4], it is assumed that the secondary users completely knew the information about the primary users in the network. By using dirty paper coding technique, the achieved data rate is maximum for transmitting data of primary and secondary users in [4]. By using spectrum holes in [5], primary information is forwarded with the help of secondary users. This technique not only promotes power savings but also helps in enhancing the battery lifetime of the phones [3].

In [6], detailed survey of the emerging 5G cellular technology along with its network architecture and some key emerging techniques like massive MIMO technology, device-to-device communication (D2D), spectrum sharing, etc., are presented. Also, an architecture for 5G cellular network is proposed in [6] which shows the inclusion of device-to-device communication (D2D), small cell access points, network cloud and Internet of Things (IOT) along with the 5G architecture. Various resource allocation algorithms for D2D communication are discussed in [7–13].

3 System Model

3.1 System Framework

The system framework is shown in Fig. 1. It has a single transmitter (PT) and a more than one primary receivers (PRs) P in j zones around the PT, such that, $P = \{PR_1, PR_2, \ldots, PR_v\}$ with 'z' PRs in each zones. The zones are represented as $Z = \{Z_1, Z_2, \ldots, Z_j\}$. The framework for $j = 4$ is given in Fig. 1.

The zones are divided depending upon the distance of PRs from PT. The distance considered is from 0 to 20 m. Four zones are considered Zone 1, Zone 2, Zone 3 and Zone 4 such that Zone 1 Z1 denotes 0–5 m, Zone 2 Z2 denotes 5–10 m, Zone 3 Z3 denotes 10–15 m, and Zone 4 Z4 denotes 15–20 m. By dividing the distance into zones, it helps to easily determine the required power for transmission in each zone.

3.2 System Model

OFDM is considered in which the channel is undergoing Rayleigh flat fading. Log-normal shadow fading is also considered in each zones. Let the signal transmitted be x_p from PT to $PR_{i,j}$, then the received signal is

$$y_p^i(j) = \sqrt{P_T^{t(i,j)}} h_{PT,PR_{i,j}} x_p + \eta_{PT,PR_{i,j}} \tag{1}$$

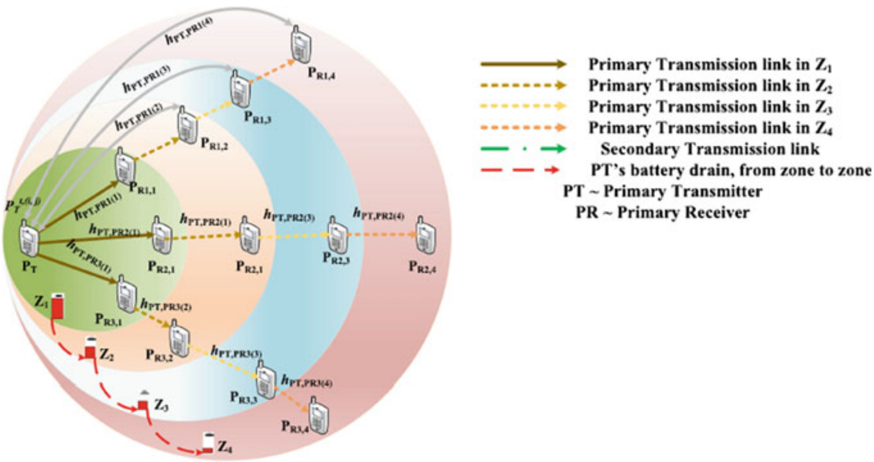


Fig. 1 Framework of zone-based D2D communication [3]

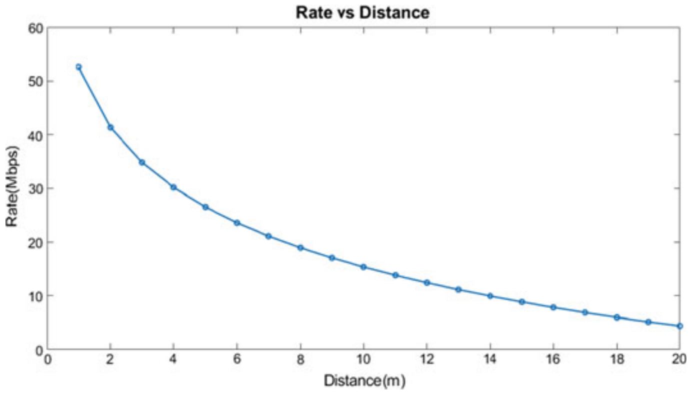


Fig. 2 Rate achievable versus distance

where $P_T^{t(i,j)}$ is the power required for transmission at PT , the channel coefficient between the PT and $\mathrm{PR}_{i,j}$ is given by $h_{\mathrm{PT},\mathrm{PR}_{i,j}}$ and $\eta_{\mathrm{PT},and\mathrm{PR}_{i,j}}$ represents the additive white Gaussian noise (AWGN) with zero mean and σ_0^2 variance.

Pathloss of D2D communication at distance d km is (from ITU Model)

$$\mathrm{PL} = 148 + 40\log_{10}d\ (\mathrm{km}) \tag{2}$$

For distance $d \leq d0$

$$\mathrm{PL} = 40\log_{10}d\ (\mathrm{km}) + 30\log_{10}f_c\ (\mathrm{Mhz}) + 49 \tag{3}$$

Channel gain at a given pathloss (PL) is

$$h_{\mathrm{PT,PR}} = 10^{-\frac{\mathrm{PL}}{10}} \tag{4}$$

The signal-to-noise ratio (SNR) at $PR_{i,j}$ is calculated as

$$\mathrm{SNR}_i^j = \frac{P_T^{t(i,j)}\left|h_{\mathrm{PT},\mathrm{PR}_{i,j}}\right|^2}{\sigma_0^2} \tag{5}$$

If the bandwidth available is W Hz, then the achievable rate can be calculated as

$$R_p^i\,(j) = W\log_2\left(1 + \mathrm{SNR}_i^j\right) \tag{6}$$

In the proposed system model, a set of applications A_p is considered such that A_1 is conversational video, A_2 is conversational audio and A_3 is Text. The target rates are $R_T^{A_1} = 13.15$ Mbps for conversational video, $R_T^{A_2} = 6$ Mbps for conversational audio and $R_T^{A_3} = 2$ Mbps for text. The applications in A_p are arranged in descending priority. The primary transmitter will serve primary receivers in the order of priority of the application demanded.

PRs in different zones are affected by pathloss, thereby impacting the QoS. As the distance between PT $\rightarrow$ PR increases, the pathloss also increases. Transmission power is dependent on PT $\rightarrow$ PR, i.e., the distance between primary transmitter and receiver. Therefore, for serving the PRs in Z_1 needs least power, followed by Z_2, Z_3 and Z_4. The relation between achievable rate and distance is shown in Fig. 2. It shows that as the distance between PT $\rightarrow$ PR increases, the achievable rate decreases.

For different values of the transmitted power $P_T^{t(i,j)}$ from 0 to 43 dBm, the corresponding data rates achieved are shown in Fig. 3. It can be seen that Z_1 can achieve high data rates at low transmission power levels than Z_2, Z_3 and Z_4, i.e., the target rates $R_T^{A_1}$, $R_T^{A_2}$ and $R_T^{A_3}$ can be achieved at smaller transmitted power levels in zones Z_1 and Z_2 than in zones Z_3 and Z_4 because of the increase in the distance between PT and PR.

Higher efficiency in usage of resources can be attained by transmitting the information at minimum or controlled power levels. This improves and extends the life of battery of the transmitting device [3].

Depending on the target rate $R_T^{A_i}$ required to meet the application demanded, power required for transmission at PT can be calculated [3]. The minimum power required for transmission $P_{\mathrm{T,opt}}^{t[i,j)}$ to meet the target rate $R_T^{A_i}$ can be obtained by using the Shannon capacity formula as

$$R_T^{A_i} = W \log_2 \left(1 + \frac{P_{T,\mathrm{opt}}^{t(i,j)} \left| h_{\mathrm{PT},\mathrm{PR}_{i,j}} \right|^2}{\sigma_0^2} \right) \tag{7}$$

The transmission power level required (optimum power) of PT can be obtained from (7) as

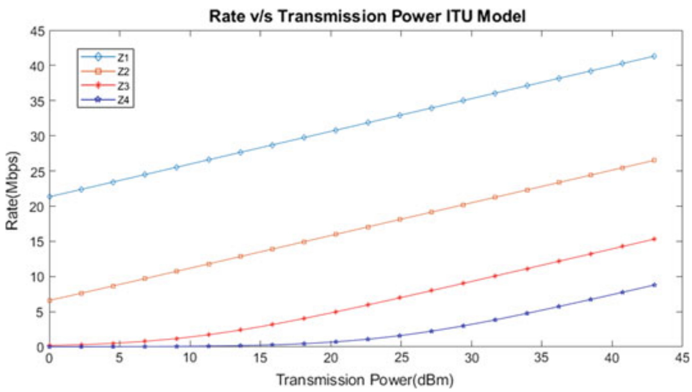


Fig. 3 Rate achievable versus transmission power

$$P_{T,\text{opt}}^{t(i,j)} = \frac{\sigma_0^2}{\left|h_{\text{PT},\text{PR}_{i,j}}\right|^2}\left(2^{\frac{R_T^{A_i}}{W}} - 1\right) \tag{8}$$

Thus, the required transmission power levels can be obtained for all the applications considered in A_p such that direct PT $\rightarrow$ PR$_{i,j}$ transmissions can be done at the optimal power $P_{T,\text{opt}}^{t[i,j)}$ for greater power savings. Power can be saved by transmitting at $P_{\text{T,opt}}^{t[i,j)}$ which is computed as

$$P_{\text{saved}(i,n)} = P_T^{t(\max)} - P_{T,\text{opt}}^{t(i,j)} \tag{9}$$

The PT will stop serving the primary receivers PRs if the transmission power level of the PR goes beyond the threshold, i.e., $P_T^{t[i,j)} < P_T^{\text{threshold}}$. This is because if it continues the transmission at very low power levels, the battery will be drained and finally stops functioning. The power optimization process is described in algorithm 1. The proposed algorithm optimizes the power levels for achieving $R_T^{A_i}$, and thus it helps in improving energy efficiency.

Algorithm 1: Optimizing Power for the system
Step 1: Input the system Parameters
Transmitter Parameter: $P_T^{t(\max)}$, $R_T^{A_i}$, A_p, W
Channel: σ_0^2
Step 2: Initialization
Number of Primary Receivers, i
Number of zones, j
Battery Capacity and Battery Energy: $I_{battery}^{capacity}$, $E_{battery}^T$
Threshold distance for pair formation, d_0
Circuit power consumption P_c
Step 3: Identifying the demanded application and power level required
/*by using loop statement*/
for l=1:j /*for zones*/
for z=1:i /*for PRs*/
Identifying the application demanded by PR from the available set of A_p
Compute $P_{T,opt}^{t(i,j)}$ to attain the desired rate, such that $P_T^{t(i,j)} = P_{T,opt}^{t(i,j)}$
if $P_T^{t(i,j)} < P_T^{t(\max)}$&& $P_{T,opt}^{t(i,j)} > P_T^{threshold}$
Direct transmission of $PT \rightarrow PR_{i,j}$can be initiated
Compute power saved $P_{saved(i)}$
else if $P_{T,opt}^{t(i,j)} < P_T^{threshold}$&& $P_{T,opt}^{t(i,j)} > P_{saved(i)}$
then Direct transmission of cannot be initiated.
End all the loops

The energy efficiency of each of the four zones with 'z' PRs and PRs demanding different application is obtained by [3] as

$$\mathrm{EE}_i^j = \frac{R_p^i(j)}{\Sigma P_T^{t(i,j)} + P_c} \tag{10}$$

The corresponding optimization problem is

$$\max_{P_T^{t(i,j)}} \mathrm{EE}_i^j \tag{11}$$

$$\sum_{i=1}^{z} P_T^{t(i,j)} < P_T^{t(\max)} \tag{11a}$$

$$P_T^{t(i,j)} > P_T^{\text{threshold}} \tag{11b}$$

$$R_p^i\,(j) \approx \max\left(R_T^{A_i}\right) \tag{11c}$$

The condition (11a) shows that the total transmission power available at the PT serving the demands of the PR in the corresponding zone should not be more than the maximum power available for transmission at the primary transmitter, and the constraint (11b) ensures that the primary transmitter's battery will not be completely drained out when serving the demand of the primary receiver. Constraint (11c) assures that the target rate will be the achievable rate of the highest application demanded. A constant circuit power consumption P_c is assumed for all the users.

Zone factor determines the time for transmission in each zone, and it can be computed as [1]

$$T_{p(i,n)}^j = \frac{\sum_{n\in A_p} p_n b_n}{R_T^{A_i}} \tag{12}$$

where p_n gives the number of packets of the files needed for the application demanded to be transmitted, and b_n denotes the number of the bits per packet. When multiple PRs demand combination of applications, then $R_T^{A_i} \rightarrow \max(R_T^{A_i})$.

The energy consumed by the primary transmitter for the transferring of files for the n applications to i PRs in zone j is [14]

$$E_{\text{consumed(i,n)}}^j = \frac{E_{\text{battery}}^T T_{i,n}^j}{3600} \tag{13}$$

Depending on the value of the energy consumed by the primary transmitter $E_{\text{consumed}(i,n)}^j$, the device will draw an equivalent amount of current from the battery $I_{\text{PT},n}^j$, given by

$$I_{\text{PT},n(j)} = \sqrt{\frac{E_{\text{consumed}(m,n)}^j I_{\text{battery}}^{\text{capacity}^2}}{E_{\text{battery}}^T}} \tag{14}$$

and it affects the battery lifetime, which is given by

$$T_b^{j,n} = \frac{I_{\text{battery}}^{\text{capacity}}}{I_{\text{PT},n}^{j}} * 0.7 \tag{15}$$

0.7 is for the effects of the external factors.

4 Results and Discussions

This section shows the simulation results obtained from the proposed power optimization scheme. The simulation is done using MATLAB R2018b. The primary transmitter is assumed to have a battery energy of 7.45 Wh. The different primary receivers demand different applications. The users in the system are undergoing Rayleigh flat fading, and channel is modeled using pathloss equations of the ITU model [15]. The simulation parameters considered are given in the Table 1. The target rate achieved is maximum for A_1 and then A_2 and A_3.

4.1 Result Analysis

The data rates that are achievable over the communication links decrease as the distance between transmitter and receiver increases. This is due to rise in pathloss. For varying value of $d_{\text{PT}\to\text{PR}_{i,j}}$ from 1 to 20 m, the data rates achieved are depicted in Fig. 4. Transmission is done at $P_T^{t(\max)}$. From the graph, it can be seen that all the

Table 1 Parameters used for simulation [7]

Parameter	Value
Carrier frequency	2 GHz
Noise power, σ_0^2	−126 dBm
Total bandwidth available, W	1.4 MHz
Maximum transmission power of PT, $P_T^{t(\max)}$	43 dBm
$P_T^{\text{threshold}}$	18 dBm
Log-normal shadow fading	4dB–8dB(in different zones)
Target rate of demanded application	A_1=13.15 Mbps A_2= 6 Mbps A_3=2 xbps
Threshold distance	20 m
Circuit power consumption	100 mW
Modulation schemes	256/128/64 QAM
File size for each application	A_1: 100 MB A_2: 50 MB A_3: 200 KB
Battery energy consumption	7.45 Wh
Battery capacity	1960 mAh

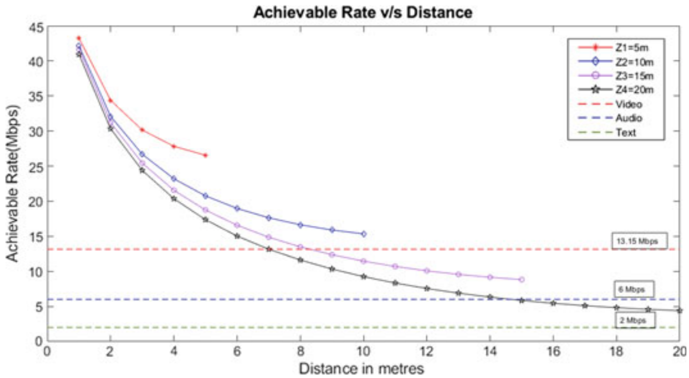


Fig. 4 Rate achieved versus distance for PT and PR in zones

four zones do not require maximum transmission power in order to achieve the target rate $R_T^{A_i}$. The maximum achievable rate in Z1 is 44 Mbps but the target rate is 13.15 Mbps only. Thus, transmission initiated at a low power level can successfully meet the target rate $R_T^{A_i}$. Transmissions done at higher power levels will consume more battery and reduce the life of battery [3]. In Z_3 and Z_4 where the distance between the transmitter and receiver is high, the target data rates that are expected to achieve are unable to meet.

4.2 Analysis of Energy Efficiency and Transmission Power

From Figs. 3 and 4, it can be seen that maximum transmission power is not required in all zones to meet the target rate $R_T^{A_i}$. Optimal power to meet the target rate can be calculated for each zone from (8). Instead of transmitting at maximum power, transmission can be done at optimal power so that power can be saved and power consumption can be reduced.

From Fig. 4, it can be seen that conversational video (13.15 Mbps) can be transmitted at much lower power levels than the maximum transmission power in Z_1 and Z_2 (up to 10m). Figure 5 shows the power saved by transmitting at optimal power level up to distance 10 m. It can be depicted from the graph that as the distance between PT and PR increases, power consumption increases and power saving reduces.

Energy efficiency can be enhanced by transmitting at optimal power levels. Figure 6 shows energy efficiency for video transmission in different zones. It can be observed from the graph that energy efficiency decreases as the distance increases because the power consumption increases with increase in distance between PT and PR. Figure 7 shows the overall energy saving of the system for video transmission up to 10m.

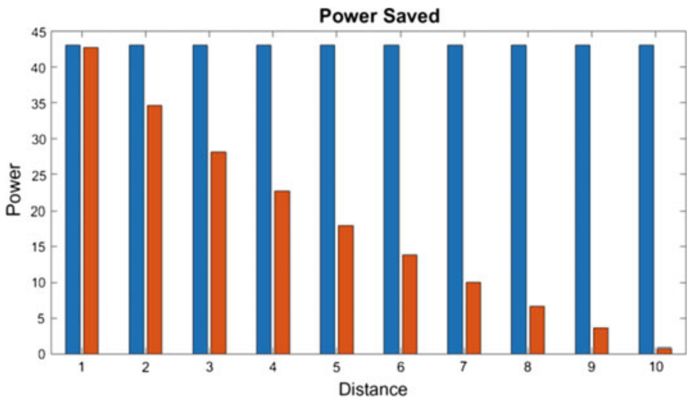


Fig. 5 Power saved

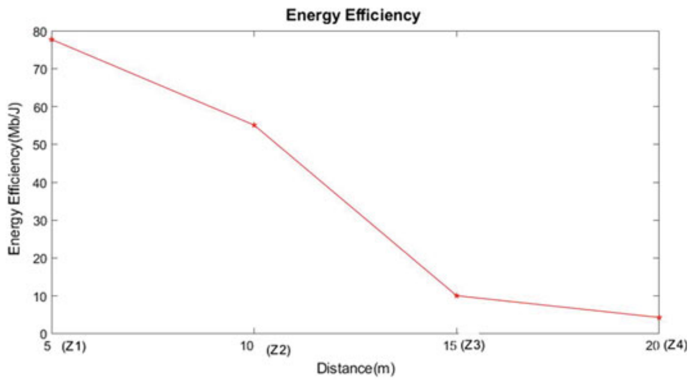


Fig. 6 Energy efficiency at each zones

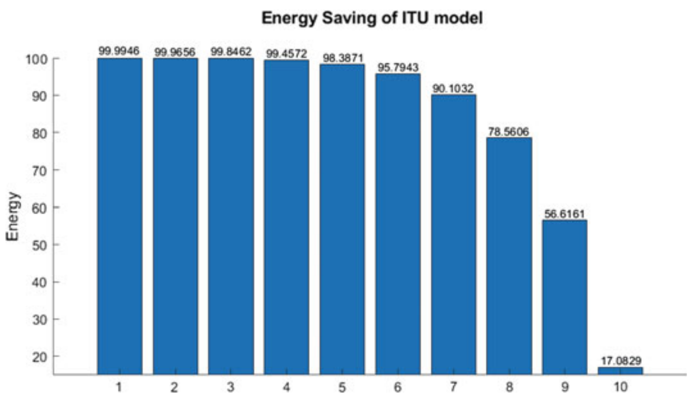


Fig. 7 Energy saving

4.3 Transmission Time and Battery Life

The transmission time needed at each zone depends on the number of data packets required to transmit the file size of the demanded application [3]. The transmission time for each zone or zone factor can be calculated from (12). It is assumed that Zone 1 and Zone 2 are demanding video, Zone 3 is demanding audio and text, and Zone 4 is demanding text. Figure 8 shows the zone factor for each zone.

Zones 1 and 2 are demanding the same application but as the distance increases, transmission time increases and hence Zone 2 has the highest zone factor. Zone 4 is demanding text only and so Zone 4 has the lowest zone factor. Zone factor plays an important role in energy consumption of each zone. Battery life of the handsets is a major concern presently because of the rising demands of the technologies. The proposed scheme augments the battery lifetime of PT.

For the battery lifetime analysis, each user in the system is assumed to have fully charged battery initially with a lifetime of 14 h. Battery lifetime of the PT is analyzed in each zone, and the results are shown in Fig. 9. It can be observed that the proposed scheme can increase the battery life up to 19.8 h, i.e., approx. 42% augmentation can be achieved.Transmission at optimal power yields greater power savings. The transmission time depends on the demanded application by the primary transmitter. Energy consumption is calculated based on the transmission time. Battery lifetime is directly affected by the energy consumption at each zone. Thus, an energy-efficient low transmission power scheme helps in prolonging battery life of the primary transmitter.

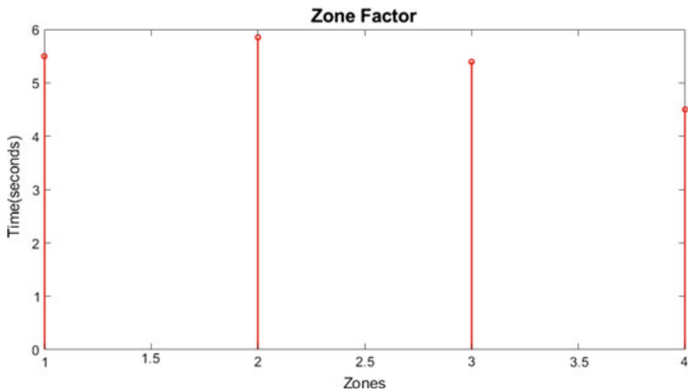


Fig. 8 Zone factor (transmission time) for different zone

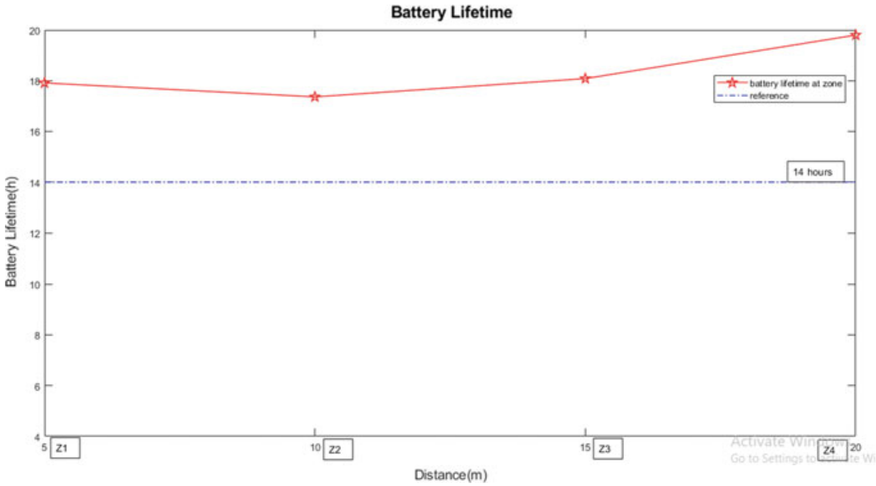


Fig. 9 Battery lifetime (in hour) for PT in zones

4.4 Comparison with Hata Model

For comparison, the pathloss has been calculated using Okumura–Hata Model as given by (16). The algorithm is then implemented using this model.

$$\begin{aligned} \mathrm{PL} = 69.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_B) \\ -C_H + [44.9 - 6.55\log_{10}(h_B)]\log_{10}(d) \end{aligned} \tag{16}$$

C_H is the antenna correction factor, h_B is the height of the antenna, and h_m is the height of mobile station.

$$C_H = 3.2\left(\{\log_{10}(11.75h_m)\right)^2 - 4.97 \tag{17}$$

Figure 10 shows the relation between achievable rate and transmission power. As the transmission power increases, achievable rate also increases. Figure 11 shows the achievable rate in different zones. All the zones do not require transmission at maximum power to meet the target rate. It can be observed that target rate can be achieved at much lower power levels than the maximum power. Power optimization is applied, and transmissions are done at optimal power for power savings. Figure 12 shows power saved for video transmission. It can be observed that as the distance increases, power saving reduces.

It can be observed that Hata model can achieve more data rate than ITU model at the given transmission power. In case of ITU model, video transmission can be done up to 10 m only but in Hata model, video transmission can be done up to 18 m. Therefore, Hata model has more power savings and energy savings compared to ITU model.

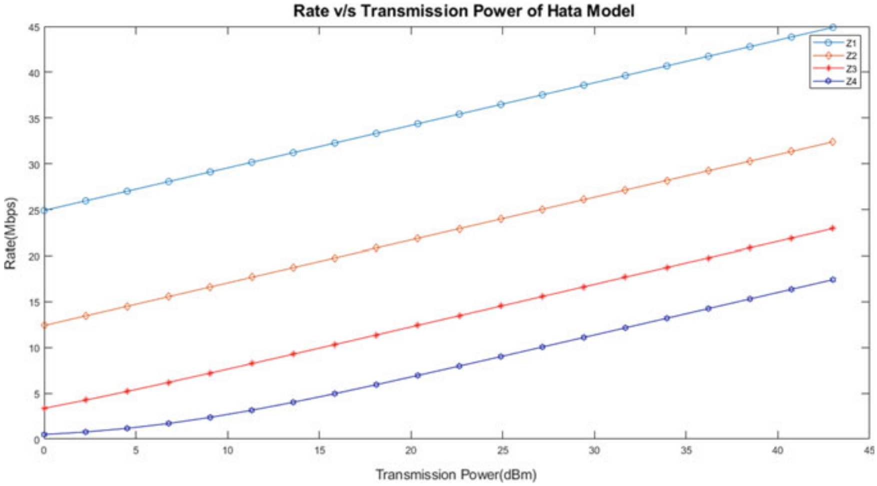


Fig. 10 Achievable rate versus transmission power

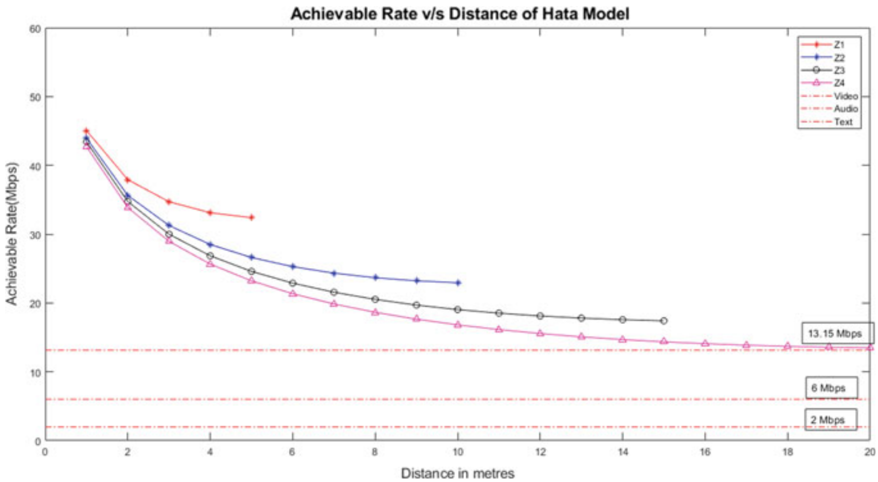


Fig. 11 Achievable rate versus distance in different zones

5 Conclusion

Systems with increased mobility of users accelerate the overall power consumption of the entire system making the system less energy efficient and prone to high draining of battery. These aspects are the major concern of the 5G WCNs. The scheme proposed here uses ITU model which involves the placement of primary receivers in a network at different distances thus by forming different zones around the primary transmitter and forms PT PR D2D links. The minimum transmission power needed

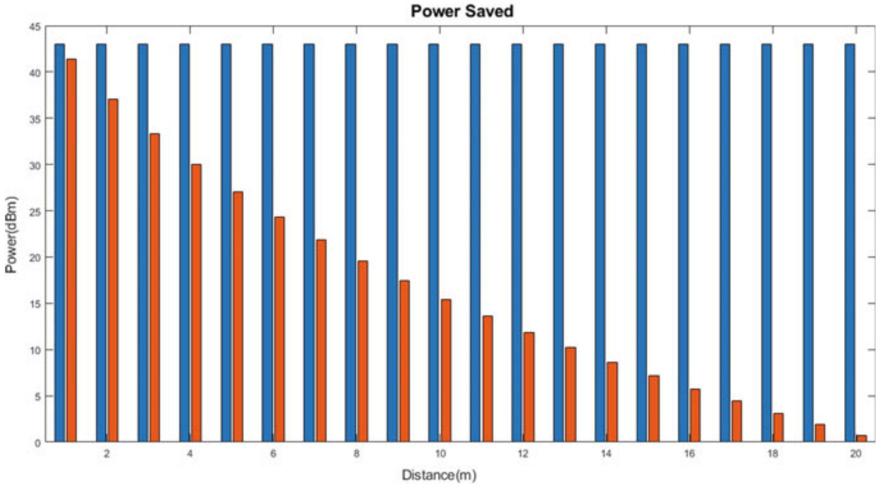


Fig. 12 Power saved

to initiate the transmission in order to meet the target rate of the demanded application is calculated for each primary receiver in different zones. Low transmission power increases the energy efficiency of the system which is clear from the results obtained. The transmission power needed to meet the target rates increases as the distance between primary transmitter and receiver increases. The achievable data rates decrease with increase in distance between primary transmitter and receiver. Power savings can be increased by transmitting at optimal power levels. As the power consumption reduces, energy efficiency increases and the system becomes more energy efficient. Energy savings can be increased by transmitting at optimal power levels. The proposed algorithm introduces zone factor which determines the transmission time for each zone, and it also determines the energy consumed at each zone for the transfer of required information. This affects the battery lifetime of the PT. Using simulations done in MATLAB, it is observed that the proposed scheme can increase the battery lifetime of primary transmitter by up to 42%. For comparison, the proposed scheme has been implemented using Okumura–Hata model. The results show that Hata model can achieve more data rate than ITU model at the given transmission power. Power savings are higher in Hata model which increases the energy savings and makes the network more energy efficient.

References

1. Shakir MZ, Ismail M, Wang X, Qaraqe KA, Serpedin E (2017) From D2D to Ds2D: prolonging the battery life of mobile devices via Ds2D communications. IEEE Wirel Commun Mag 24
2. Gandotra P, Jha RK, Jain S (2017) Green communication in next generation cellular networks: a survey. IEEE Access 5

3. Gandotra P, Jha, RK (2019) Zonal-based green algorithm for augmenting the battery life in spectrum shared networks via D2D communication. IEEE Trans Veh Technol
4. Devroye N, Mitran P, Tarokh V (2006) Achievable rates in cognitive radio channels. IEEE Trans Inf Theory 52(5):1813–1827
5. He C et al (2017) Cooperative spectrum sharing based on contract theory with optimal bandwidth and power allocation. In: IEEE wireless communication and mobile computing conference, pp 346–350
6. Gupta A, Jha RK (2015) A survey of 5G network: architecture and emerging technologies. IEEE Access 3:1206–1232
7. Gandotra P, Jha RK, Jain S (2018) Sector-based radio resource allocation (SBRRA) algorithm for better quality of service and experience in device-to-device (D2D) communication. IEEE Trans Veh Technol 67(7)
8. Andrews JG et al (2014) What will 5G be? IEEE J Sel Areas Commun 32(6):1065–1082
9. Feng D et al (2013) A survey of energy-efficient wireless communications. IEEE Commun Surv Tuts 15(1):167–178
10. Gandotra P, Jha RK (2017) Next generation cellular networks and green communication. IEEE Commun Lett 22(7)
11. Costa M (1983) Writing on dirty paper. IEEE Trans Inf Theory 29(3):439–441
12. Tawalbeh M, Eardley A (2016) Studying the energy consumption in mobile devices. Procedia Comput Sci 94:183–189
13. Su Z, Hui Y, Yang Q (2017) The next generation vehicular networks: a content-centric framework. IEEE Wirel Commun 24(1):60–66
14. Gandotra P, Jha RK, Jain S (2018) Prolonging user battery lifetime using green communication in spectrum sharing networks. IEEE Commun Lett 22(7)
15. Abrol A, Jha RK (2016) Power optimization in 5G networks: a step towards GrEEn communication. Special section on green communications and networking for 5g wireless, vol 15

View-Based and Visual-Attention-Based Background Modeling for Detecting Frequently and Infrequently Moving Objects for Video Summarization

D. Minola Davids and C. Seldev Christopher

Abstract The real-time face detection (FD) algorithm is proposed to find faces in the images as well as videos. Besides face regions, this algorithm also finds the exact localities of the face parts like lips and eyes. Initially, skin pixels are extracted centered on the rules of simple quadratic polynomial model. By introducing small modifications, this polynomial model (PM) could be applied for extracting the lips. The merits of adopting these two identical PMs are two-fold. Firstly, computation time is saved. Secondly, these extraction processes could be executed at the same time on one scan of the video or image frame. Subsequent to skin and lips, the eyes are extorted. Later, the algorithm eliminates the falsely extorted parts by validating with rules taken as of the spatial and geometrical relationships (SGR) of face parts. At last, the exact face regions are ascertained accordingly. As per the experiential outcomes, the proposed algorithm evinces preeminent task in respect of accuracy and speed for FD with huge differences in color, size, shape, expressions, and angles.

1 Introduction

Nowadays, the advancements in image processing (IP) approaches as well as the cost reduction of video/image acquisition gadgets stimulated the introduction of numerous computer vision applications. Some of such applications are vision-centered surveillances, vision-centered man to machine interfaces, and vision-centered biometrics. Face recognition is the main job that draws the interest of several researchers. Manifold works had offered some FD applications in commercial and laboratorial

D. Minola Davids (✉)
Department of Electronics and Communication Engineering, C.S.I Institute of Technology, Nagercoil, Kanyakumari, TamilNadu, India
e-mail: dminoladavids@yahoo.co.in

C. Seldev Christopher
Department of Computer Science and Engineering, St.Xavier's Catholic College of Engineering, Chunkankadai, Nagercoil, Kanyakumari, TamilNadu, India
e-mail: seldev@ieee.org

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_10

scales [1]. While designing an excellent FD system, designing the effectual algorithm is a main task to detect faces on captured videos or images. In fact, FD is as significant as recognition systems in certain applications. For instance, in a particular video transmission, the only changing foreground things in the video frames are the human faces. Consequently, avert the repetition of encoding, transmissions, and de-coding of unchanged background objects for saving the network bandwidth and reducing computations.

As FD is the first phase in the course of transmission or recognition areas, its performance will place some stern limitation on the attained system performance. FD actively involves in the segmentation of faces as of the video backgrounds. Preferably, an excellent face detector ought to precisely extort the entire faces of the images irrespective of its scales, positions, colors, poses, orientations, expressions, scales along with lighting conditions.

Nevertheless, for the existent top-notch IP technologies, it is a huge challenge. Consequently, numerous designed face detectors cope only with frontal and upright faces in well-restrained platforms [1].

Apart from accuracy, the detection speed is the vital concern. In several surveillance and video phone applications, the real-time speed is requisite which forbade numerous algorithms that accurately extort faces with maximal computation time. Certain high-speed CPUs might render an excellent hardware solution to this speed requisite; nevertheless, the elevated prices of such powerful CPUs might also cut back these systems' acceptability for the common users.

Here, a novel real-time FD is proposed, and it could precisely find the face sections and their respective parts like lips and eyes in the images. The meticulous specifications of the proposed work's capability are exhibited below.

1. Users could slant their faces right or left for around 45°:
2. Users could lower, rotate, or raise their heads only if lips and eyes are not occluded.
3. The faces sizes are restricted as of 1600 (= 40 * 40) pixels and 9216 (= 96 * 96) pixels in order to fix the resolution needs for FR engines. These values could be effortlessly adjusted if disparate resolutions are insisted.

If the natural lights consistently illuminate on the faces, the light on some regions of the face images is excluded. The fundamental conception of this work is to extort and also to confirm the preferred parts, counting skins, faces, lips, or eyes with the simplest rules. These rules could manage a huge degree of differences in faces. Owing to their (rules) ease and effectiveness, the proposed algorithm could precisely perform FD with broad variations at instantaneous speed.

The remaining sections are prearranged as: Sect. 2 proffers a short review on the associated works. Section 3 elucidates the proposed work in detail. Section 4 exemplifies the proposed algorithm's effectiveness by conducting the performance comparisons and assessments in respect of the speed and accuracy. At last, Sect. 5 briefly infers the paper.

2 Related Works

A clear-cut methodology for FD in images is the template co-relation matching as elucidated in [1–4]. A template could be modeled or sometimes learned via collecting the face patterns. At the time of matching, a template is firstly convolved to the sub-images of the inputted image to ascertain the feasible candidates in respect of previously defined distance or similarity. For managing the feasible differences of size, shape, orientation, et cetera, '2' approaches are normally employed. The primary approach was to re-size the inputted image to disparate dimensions ahead of matching. Subsequently, the 'template matching (TM)' is performed on every single re-sized image again and again. The secondary approach was to utilize manifold face templates different sizes, expressions, lightings, or orientations to match the inputted image [5, 6]. Clearly, the matching time drastically elevates centered on the dimensions and numbers of the utilized inputted images and templates.

Skin pixel extortion (SPE) was commonly done to diminish the candidate searching time. The base concept of the SPE was to develop a statistical design for the colors of skin's suchlike pixels. The blend of Gaussian distributions was a helpful probabilistic design, and it was utilized for eliminating the non-skin pixels. Other analysts suggested certain neural network (NN) methodologies for approximating complex frameworks regarding skin pixels in the images. Nevertheless, the augmented hazards of such neural designs as well raised the computational costs.

Using the recognized skin pixels, the connected skin pixels as of numerous candidate face sections were extorted. Consequently, the previously stored face templates were utilized for matching the images or sub-images in those candidate sections. Then, the undesired skin areas that have no face-like pattern were abandoned. Besides TM, the NN verifier was a variant framework for face validations. As per the experiential outcomes, this verifier normally evinces high-level variation tolerance on considering a TM. Nevertheless, the prolonged training time and the requisite of huge train-sets were the prime demerits of this verifier.

The aforesaid FD method could well be implied as image-centered (top-down) detection. A word 'top-down' indicates that the face areas are evaluated with no re-sorting to the detection of individual face parts like lips, noses, eyes, etc. Here, a variant approach, which was termed as the feature-centered (bottom-up) detection for discovering face areas, was presented. Bottom-up signifies that the exact face areas were built as of the recognized face parts. The face parts needed to be extorted before determining the correct localities of face areas. Its philosophy was that certain individual face parts like lips and eyes normally revealed visual features slightly sensitive to the aforesaid face variations. The lips and eyes are normally the darker-skin parts on a face. Such features facilitated eyes to identify the face parts easily. Therefore, the extortion of such face parts was normally steady. Additionally, as the individual parts (eyes, lips, etc.) were much smaller on considering the complete face, the processing time requisite for extortion was smaller.

3 Rule-Based Face Detection Algorithm

As per the above-discussed bottom-up detection methodology, the algorithm is planned to extort the face parts counting eyes as well as lips.

To lessen the searching regions in the images (input), the algorithm extorts the skin pixels. Nevertheless, rather than utilizing probabilistic designs, a quadratic PM is used for the given color model (CM) of skin pixels to lessen its computational time. Furthermore, this PM is as well extended to the lips extortion with a slight adjustment. Lastly, the incorrectly extorted eyes and lips are eliminated centered on the rules induced as of the common SGR amongst normal face parts. Accordingly, the final correct face sections are determined.

3.1 Rules for Skin-Color Region Extraction

The extortion of skin-color sections aims to lessen its searching time intended for potential face areas of the inputted image. For simplifying the effect of environmental light's brightness on the skin pixel sections, the proposed work espouses the chromatic color coordinate for color representations. Aimed at chromatic-CS, every pixel is signified by '2' values, like (r, g): Mathematically, the conversion process of the RGB to the chromatic-CS is stated as given (Eq. 1):

$$r = \frac{R}{R + G + B}$$

$$G = \frac{G}{R + G + B} \tag{1}$$

Here,

R, G, B—Respective intensities of the red, green, blue pixels.

As per this transformation, it is simple to perceive that the brightness variations in the considered imageries could be normalized in a chromatic-CS. Furthermore, the color representations in this 2D-chromatic-CS enable one to visualize as well as examine effortlessly while constructing the CM for the considered skin pixels. For constructing a CM, the dissemination of pixels on r-g plane is plotted in Fig. 1.

Normally, the considered skin pixels generate a small section over an r-g plane. In numerous existent methodologies, certain probabilistic designs are utilized to approximate the distribution. For instance, the Gaussian mixture is an extremely eminent model. The main downside of such probabilistic design shows the higher computation expenses in computing the given probabilities. The probabilistic computation ought to be performed for every single pixel. The computational time augments as the image's size elevates. As this algorithm aims to instantaneously recognize the faces, those computation expenses are relatively obstructive to the target. Therefore, rather than endeavoring to develop this design, a simpler approach which needs only

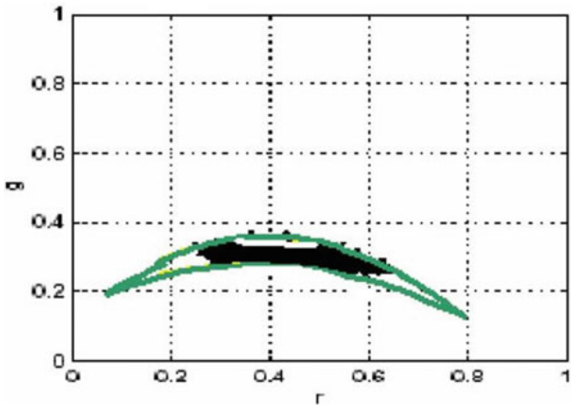


Fig. 1 Distribution of skin-color pixels on an r-g plane of chromatic-CS and the upper-lower parabolic boundaries proffered by the 2 quadratic polynomials

too low computational costs is adopted. The Soriano and Martinkauppi's approach is modified, which utilizes '2' quadratic polynomials for effectively approximating the upper-lower boundaries in this compact area that are regarded as skin locus, developed by the pixel distribution on an r-g plane. The '2' polynomials are expressed below (Eq. 2):

$$f_{\text{upper}}(r) = -1.3767r^2 + 1.0743_r + 0.1452$$

$$f_{\text{lower}}(r) = 0.776r^2 + 0.5601_r + 0.1766 \tag{2}$$

In reality, the coefficients of '2' polynomials could be evaluated utilizing least-mean squared (LMS) error minimization. Once the given boundaries of compact area are drawn, consider $(r_1; g_1)$; $(r_2; g_2)$; $y(r_\text{n}; g_{n)}$ to be 'n' sampled points on the upper boundary. Subsequently, the succeeding linear equations could be acquired as (Eq. 3):

$$\begin{aligned} a_u r_1^2 + b_u r_1 + c_u &= g_1; \\ a_u r_2^2 + b_u r_2 + c_u &= g_2; \\ a_u r_n^2 + b_u r_n + c_u &= g_n; \end{aligned} \tag{3}$$

Here,

a_u, $b_{u,}$ and c_u—'3' coefficients of quadratic PM for the upper boundary.

Rewriting the aforesaid equations in a matrix as proffered below (Eq. 4),

$$u = R^W G; \tag{4}$$

Here,

R^w—Pseudo-inverse of R and is equivalent to $(R^t R)^{-1} R^t$::

The coefficients in the quadratic polynomials for the lower boundary could well be determined in the similar way.

With the ascertained coefficients, the pixels that are in the area betwixt those '2' polynomials could be extorted using the succeeding rule by verifying their *r–g* values:

$$\text{R1} : g > f_{\text{lower}}(r) \text{ and } g < f_{\text{upper}}(r) : \quad (5)$$

In Fig. 1, it could be perceived that the above '2' inequalities proffer a crescent-like region on an r-g plane. Apart from those '2' polynomials, Soriano and Martinkauppi stated a circle on an *r-g* plane to leave out the bright-white pixels falling at the point $(r; g) = (0{:}33, 0{:}33)$ as of the region covered with those polynomials. Consequently, the exclusive circle is given as

$$\text{R2} : W = (r - 0 : 33)^2 + (g - 0 : 33)^2 \leq 0 : 0004 \quad (6)$$

There exhibits the outcome of employing rules R1 and R2, simultaneously. As per the outcomes, the '2' rules are yet imprecise to effectively filter the non-skin pixels, chiefly for blues, yellow greens, as well as oranges that drop about the left, top, and right ends of a crescent region correspondingly.

To ameliorate the outcomes, the proposed algorithm employed additional '2' simple rules like:

$$\text{R3} : R > G; G > B; \text{R4} : R - G > 45 : G > B; \quad (7)$$

Rule R3 is basically obtained as of the perception that the considered pixels are inclined to be yellow and red. Amongst the '3' channels, blue is the least intensity. With R3, the pixels in blue could be efficiently eliminated. R4 is stated to eradicate the yellow green ones.

There exhibits the enhanced outcome after bringing in R3 and also R4: However, R1 to R4 involve simpler computations. This extraction procedure is more competent than those probabilistic strategies. The final rule is summarized for extorting the considered pixels and is proffered as:

$$\begin{aligned} &S = 1 \text{ if all } R1,\ R2,\ R3,\ R4 \text{ are true} \\ &S = 0 \text{ otherwise} \end{aligned} \quad (8)$$

Here, $S = 1$ proves that this evaluated pixel is a skin pixel.

Subsequent to this extortion, gather the connected skin pixels into the compacted regions. For enhancing such compactness, the morphological dilation, as well as, erosion is executed over these extorted pixels. After that, the process ascertains the bounding boxes aimed at the joined skin pixels' parts. Those boxes turn out to be the candidate skin-color sections for additional recognition of single face parts.

3.2 Rules for Lips and Eyes Detection

Preserving every lip pixels (generally lips), it is found that the lip colors gamut as of dark red to purple beneath usual lighting conditions. As of the standpoint of a person visual insight, the lips are extremely simple to be distinguished as of the skins for any populace (i.e., races) on account of their disparate color contrasts. Centered on this, the color dissemination of the skins and lips ought to be discernible (Figs. 2 and 3).

The left figure is the inputted image, whereas the right one is the resulting image $R3$.

In fact, the experimental outcomes exhibit that the lip colors disseminate at the crescent region mainly at its lower side as stated on an r-g plane. Another quadratic polynomial discriminant function (DF) closer to the skin extortion is proffered for lip extortion. While computing the polynomial function, the succeeding '2' design's objectives are regarded for attaining maximal-speed extraction

1. Computationally, the DF is ought to be effective
2. The lip and skin pixels could be detected in '1' scan of the video or image frame (in parallel).

Centered upon those '2' objectives and the experiments concerning the dissemination of lip colors, it is found that the quadratic polynomials of a lower boundary $f_{\text{lower}}(r)$ in the Eq could be reutilized. Since the color (lips) distributes about the lower term's value in $f_{\text{lower}}(r)$, increase slowly the value of the constant to attain the upper boundary. After such evaluations, the '2' DFs for lips are defined as

$$
\begin{aligned}
l(r) &= 0:776r^2 + 0:560lr + 0:2123; \\
f_{\text{lower}}(r) &= 0:776r^2 + 0:560lr + 0:1766;
\end{aligned} \tag{9}
$$

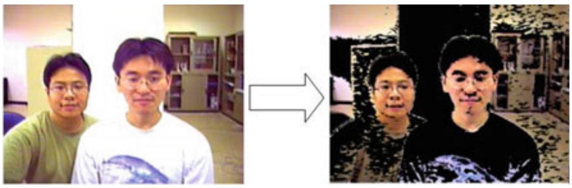


Fig. 2 Skin pixels filtering utilizing the Soriano and Martinkauppi technique

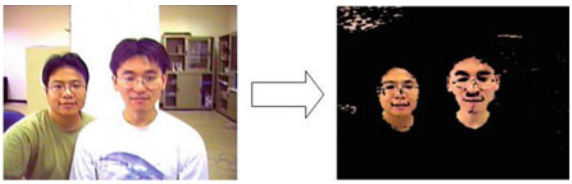


Fig. 3 Improved result of skin pixel extraction after utilizing extra rules

At the time of utilizing $l(r)$, the eyebrows, nose holes, or eyes, which are the dark pixels, are eradicated. Hence, the succeeding rule is defined for recognizing lips.

$$\begin{aligned} L &= 1 \text{ if } f_{\text{lower}}(r) \leq l(r) \text{and } R \geq 20,\, G \geq 20,\, B \geq 20 \\ &= 0 \text{ otherwise} \end{aligned} \tag{10}$$

$L = 1$ signifies the evaluated pixel is a lip's pixel. The cause of utilizing RGB in place of chromatic-CS to leave out the above-said dark parts is that the chromatic-one is inappropriate for distinguishing the dark from bright pixels. Darker and brighter ones may encompass the similar transmuted r-g values on account of the normalization in the brightness of chromatic-CS.

Thus, computation time of the DF is less, and it is the prime advantage of utilizing suchlike mathematical methods in skin and lips detections. Additionally, both are detected on '1' scan of the video or image. Consequently, this design highly followed the aforesaid objectives. Subsequent to the detection of lips, the proposed algorithm labels the connected parts for grouping the connected lip pixels into candidate lip sections.

This proposed algorithm does not handle the faces tilted right or left for >45°: Consequently, the lips are always presumed to be underneath the eyes position in a face. Utilizing this presumption, numerous infeasible combinations for eyes and lips would be eliminated in later SGR verifications of face parts. The algorithm eradicates the incorrectly extorted parts in the component-centered verifications process (CVP).

Concerning the extortion of eye pixels (simply eyes), it is perceived that the eye parts are the darkest one in the face below usual lighting condition. Nevertheless, it does not entail that eye pixels are in black at all times. On the off chance that a dark eye pixel encompasses RGB values of (1; 1; 1): then chromatic-CS of $r = g = 0{:}333$ is obtained. Nevertheless, for any bright-white pixel whose RGB values are (v; v; v) for a higher value of v, the transmuted r-g vector is $r = g = 0{:}333$. Consequently, the polynomial DF in chromatic-CS is inappropriate for extorting eyes. Hence, the algorithm takes another way to extort the darker parts. The eyes are extorted via a threshold operation (=20) on the histogram-equalized gray-scale image transmuted as of the original image color. Furthermore, the bounding boxes of skins are utilized to dispose of the incorrectly extorted dark parts. This technique is very effectual. Albeit certain dark components are incorrectly recognized, the succeeding CVP will eradicate them.

3.3 Rules for Component Verifications and Face Region Determination

In the extracted feasible lips and eyes, certain false candidates may exist. For removing these incorrectly extorted parts, the proposed algorithm executes CVP centered on the rules taken as of the common SGR amongst face parts. The proposed algorithm utilizes the succeeding rules,

1. Consider $p_{eR}(x_{eR}; y_{eR})$ as the center of right eye and $p_{eL}(x_{eL}; y_{eL})$ as the center of left one. The angle betwixt the horizontal line and $p_{eL}p_{eR}$ (line) is ought to be within [45; 45]. It is utilized to eradicate the faces that are tilted right or left for >45°
2. Without losing generality, presume $x_{eL} \leq x_{eR}$ Consider the center of the lip as be the center of the lip component and $p_{mp}(x_{mp,}\ y_{mp})$ the point of y_mÞ and the point of projection on the $p_m(x_m;\ y_m)$ If m_e is the slope of $p_{eL}p_{eR}$, then the succeeding spatial rules amongst eyes and lips should be satisfied:

The first rule indicates that the projection of the lip center should fall betwixt the eyes on condition that the face's tilt angle is not high ($|m_e| < 0{:}2$): The second and the third rules eliminate the cases in which the lips' position is inconsistent with the head's tilt direction. The projected points could not be in the left region of the left eye, that is, $x_{mp} \times x_{eL}$: If the head tilted left ($0{:}2 \leq m_e \leq 1$) and also vice versa.

3. Consider L(p,q) as the length of pq (line segment) and $p_{ec}(x_{ec};\ y_{ec})$ as the mid-point of the '2' eyes.
4. Consider $W(C_i)$ as the width of bounding box of part 'i'. Here, the succeeding geometries rules on the face parts should be satisfied.

$$\frac{1}{1:3} \leq \frac{W\left(C_{\text{Left - eye}}\right)}{W\left(C_{\text{Right - eye}}\right)} \leq 1:3;$$
$$0:6 \leq \frac{W\left(C_{\text{Left - eye}}\right)}{W(C_{\text{Lin}})} \leq 1:1 \text{ and } 0:6 \leq \frac{W\left(C_{\text{Right - eye}}\right)}{W(C_{\text{Lin}})} \leq 1:1;$$
$$0:25 \leq \frac{W\left(C_{\text{Left - eye}}\right)}{L(p_{eL}p_{eR})} \leq 0:4 \text{ and } 0:25 \leq \frac{W\left(C_{\text{Right - eye}}\right)}{L(p_{eL}p_{eR})} \leq 0:4; \tag{11}$$

The aforementioned rules help to eradicate the noise. However, the eyebrows may found connected from the eyes in the extorted parts. For those scenarios, the algorithm concerns the connected eyebrow and eye as a noise. These actions slightly affect the precision of the eyes' positions but not affect the accuracy of the extorted eyes. Therefore, the faces' detection accurately is not notably affected.

Grounded upon the verification of the said '4' SGR rules for lips and eyes, the candidate face triangles on the skins could be extorted to process further. In the proposed algorithm, the detailed differences of lips and eyes could be attained, could highly lessen the total of extorted triangles. Additionally, the supposition that the lips are just below the eyes in one's face is also obliging in eliminating the infeasible combinations.

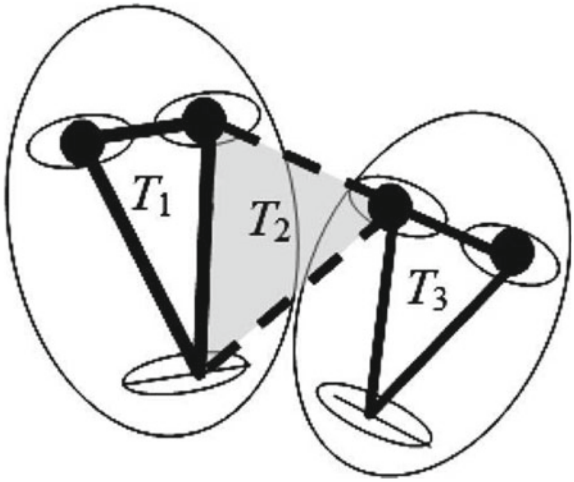


Fig. 4 Case of confusing triangles in face component verifications

3.4 The Arbitration of Confusing Eye–Lip Triangles

All noise parts could be eliminated after CVP. Nevertheless, there might be certain confusing combinations that are requisite to be managed further. Those confusing combinations aroused as of the triangles which comprise the eyes of one face and the lips of another. As in Fig. 4, there are '3' feasible triangles like T_1; T_2; T_3 for face parts, wherein T_2 implies a wrong one. For addressing this uncertainty, a strategy is proposed to eliminate the confusing triangles. Initially, termed skin color ratio (which is SCR) is proffered for the settlement in the bewildering triangles. Now, the SCR is evaluated as

$$\mathrm{SCR}\left(\Delta_{p1p2p3}\right) = \frac{N_{\mathrm{skin}\,j\,\Delta p1p2p3}}{A\left(\Delta_{p1p2p3}\right)}; \tag{12}$$

Here, A($\Delta p_1 p_2 p_{3)}$ indicates the area of the triangle named $\Delta p_1 p_2 p_3$ as well as $N_{\mathrm{skinj}\Delta p1p2p3}$ signifies the total skin-color pixels within the $\Delta p_1 p_2 p_3$. If SCR is too low, then $\Delta p_1 p_2 p_3$ is likely to be a triangle that contains 3 facial components from scattered or distinct skin-color areas. For suchlike triangles, SCR is very probable to be surplus since the high rate of non-skin pixels normally comes as of the non-skin background which segregates in the existent faces (as in T2 of Fig. 4). Therefore, SCR stands as a preeminent sign for the settlement of bewildering triangles.

To gauge $N_{\mathrm{skinj}\Delta p1p2p3}$, it requires specific computations. Luckily, the commenced cost is not too high since the total bewildering triangles are normally too less subsequent to the CVP. For hoarding computation time, a fast technique was utilized to gauge $N_{\mathrm{skinj}\Delta p1p2p3}$. This technique scrutinizes every detected skin pixel to perceive if it is inside the $\Delta \mathrm{p}_1 \mathrm{p}_2 \mathrm{p}_3$: If so, then $N_{\mathrm{skinj}\Delta p1p2p3}$ is augmented by '1'. Like this, the

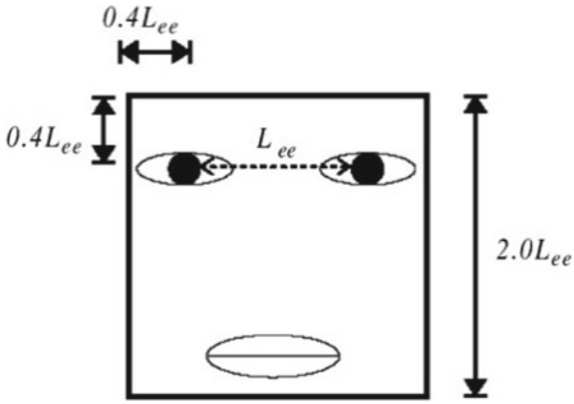


Fig. 5 Determination of the face region

time-consuming algorithms can well be averted to trace the interior pixels of every bewildering triangle.

Here, triangle T2 is a poor confusing triangle.

Only the simple triangle interior checks are to be done over every confusing triangle for all skin pixels. As per the simple geometry, provided the triangle vertices like $p_1(x_1; y_1)$; $p_2(x_2; y_2)$; and $p_3(x_3; y_3)$, the rule for the triangle internal check of any $p(x; y)$ point is

$$\begin{aligned} C_1 &: x(y_2 - y_1) + y(x_1 - x_2) + (x_2 y_1 - x_1 y_2) > 0; \\ C_2 &: x(y_3 - y_2) + y(x_2 - x_3) + (x_3 y_2 - x_2 y_3) > 0; \\ C_3 &: x(y_1 - y_3) + y(x_3 - x_1) + (x_1 y_3 - x_3 y_1) > 0; \end{aligned} \tag{13}$$

If (all C_1; $C_{2,}$ and C_3 are true) or (all C_1; $C_{2,}$ and C_3 are false),

Then, $p(x; y)$ is inside the triangle $\Delta p_1 p_2 p_3$.

Subsequent to the triangular extortion of face parts, the ultimate face areas could be effortlessly determined centered on the triangles. Figure 5 evinces the bounding box dimension for the extorted face area. In Figure-5, L_{ee} indicates the distance betwixt the '2' eyes. L_{me} signifies the distance from the center point of a line linking the '2' eyes to that of the extorted lips. The concerned bounding box is rotated in respect of the line's (linking 2 eyes) tilt angle.

4 Performance Evaluation

At this moment, the proposed algorithm is executed on a computer with 128 M RAM and PentiumIII 800 CPU for evaluating the performance. This system has '2' operation modes like (i) online and (ii) offline modes. The online operation mode is intended to recognize real-time faces on video frames taken as of a PC camera. But, the offline mode is modeled for detecting faces in still images. For assessing the proposed algorithm's speed and accuracy, a test set containing 1000 imageries is prepared. Amongst them, 815 imageries having [320 × 240] dimension of pixels are attained by the ORITE VQ-681 USB PC camera, while the other 185 are gathered as of WWW. Aimed at the '815' images attained, they comprise merely Chinese. For evaluating the robustness of the extraction of skin pixel, 185 face images for diverse races are gathered as of WWW. These '2' sorts of test imageries comprise single- and multi-face images with different orientations, sizes, expressions, and tilt angles.

Occlusion of face parts: The SGR amongst face parts is grounded upon the CVP. If any lip or eye parts are missing on account of body occlusion, wearing, head rotation, et cetera, then the system normally fails in the CVP.

Terrible illumination condition: Consider the light in the normal state, i.e., the light is not a color-biased one, and it evenly illuminates on the faces. If the presumptions are breached, then it directs to failure in the extortion of lips and skins. Specifically, the incorrect extortion of lips always brings erroneous detection outcomes.

Deprived imaging quality: Certain images gathered as of WWW have poor or no quality on account of the compression. Sometimes, face features as well become unclear owing to the compression.

5 Concluding Remarks

As per the experiential outcomes, the proposed algorithm evinces high-level performances in accuracy as well as speed. Generally, the applications and uses are in well-restrained circumstances with system utilization and also in environmental control. The proposed system could be again enhanced in accuracy as well as speed by further refinement and simplification of the system model. Still, there is a limitation in utilizing the proposed algorithm. For which, the light condition has to be normal. Contrarily, the FD could be done using more enhanced component-centered detection as well as verification process for the imperfect face parts. As a future enhancement, this algorithm can be extended for video summarization process and tracking for surveillance.

References

1. Pentland A, Moghaddam B, Stamer T, Oliyide O, Turk M (1994) View-based and modular eigenspaces for face recognition. In: IEEE conference on computer vision and pattern recognition, Seattle, pp 84–91
2. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Machine Intell 19(7):711–20
3. Er MJ, Wu SQ, Lu JW, Toh HL (2002) Face recognition with radial basis function (rbf) neural networks. IEEE Trans Neural Netw 13(3):697–710
4. Kim HC, Kim D, Bang SY (2002) Face recognition using the mixture-of-eigenfaces method. Pattern Recogn Lett 23(13):1549–1558

5. Gao YS, Leung MKH (2002) Face recognition using line edge map. IEEE Trans Pattern Anal Mach Intell 24(6):764–779
6. Rong D, Su GD, Lin XG (2002) Face recognition algorithm using local and global information. Electron Lett 38(8):363

Analysis of Segmentation Algorithms for Detection of Anomalies in MR Brain Images

A. Lenin Fred, S. N. Kumar, Parasuraman Padmanabhan, Balazs Gulyas, and H. Ajay Kumar

Abstract The role of segmentation is vital in image processing for the extraction of region of interest. In the perspective of medical images, the region of interest corresponds to anatomical organs or anomalies such as tumor, cyst. This work analyses various algorithms for the analysis of MR brain images. The clustering-based segmentation technique was found to be efficient and for the validation of results, performance metrics like Jaccard Index, rand index, false positive, and false negative are used. The segmentation algorithms are tested on real-time MR brain and brain web database images. The algorithms are developed in Matlab 2010a and the goal of this research work is to guide the researchers for choosing appropriate algorithm for the analysis of MR brain images.

Keywords Clustering · Region of interest · Thresholding · Region growing

A. Lenin Fred · H. A. Kumar
Mar Ephraem College of Engineering and Technology, Elavuvilai, India
e-mail: leninfred.a@gmail.com

H. A. Kumar
e-mail: ajayhakkumar@gmail.com

S. N. Kumar (✉)
Amal Jyothi College of Engineering, Kanjirappally, Kerala, India
e-mail: appu123kumar@gmail.com

P. Padmanabhan · B. Gulyas
Cognitive Neuroimaging Centre, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
e-mail: ppadmanabhan1@gmail.com

B. Gulyas
e-mail: balazs.gulyas@ntu.edu.sg

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_11

1 Introduction

Image segmentation is defined as the technique of delineation of the desired Region of Interest (ROI) and in the scenario of the medical field, ROI corresponds to an anatomical organ or anomalies like tumor or cyst. In general segmentation algorithms are categorized into contextual and non-contextual techniques. In the contextual technique, the spatial relationship between the pixels is not considered and ROI extraction relies on a global attribute such as gray value. Examples of the contextual technique are classical thresholding, adaptive thresholding, color thresholding, etc. The non-contextual technique relies on the spatial relationship between the pixel values. Examples of non-contextual techniques are region growing, clustering, etc. The segmentation algorithms are validated by suitable metrics based on the availability of gold standard image used for the efficiency assessment.

In the medical field, the rapid development of the non-invasive imaging technology led to the new dimension in the analysis and examining the anatomy and functions of the human body. MRI, EEG, PET, and FMRI are the different modalities of images for the visualization of the different organs and tissues inside the human body [1]. The analysis of the brain imaging data by manual process is complex and time-dependent and a difference in opinion with respect to each examining observer [2], hence there is a large scope in the quantitative analysis of MR brain imaging-related studies. The computer algorithms help the physicians in the better representation of anatomical structures and the desired region of interest [3]. These technologies have a significant role in the diagnosis, treatment planning, localization of ROI, volume quantification and computer-assisted surgery [4].

In the MR imaging, the identification of the brain structures is essential in neuro studies for the mapping of the functions, study anatomy and brain development and for the analysis of neurodegenerative brain disorders. Brain image segmentation has an important role in medical imaging for detecting, diagnosing, treatment planning, and so on [5]. There are numerous variants of automatic and semi-automatic segmentation algorithms for medical image segmentation [6]. Because of the noise incurred during acquisition, weak boundaries, image inhomogeneity, poor contrast the segmentation algorithms fail to generate better results. For better clinical diagnosis, accurate segmentation is necessary [7, 8]. Many image segmentation algorithms are used for the segmentation of the brain structures such as thresholding, region growing, watershed, Atlas-based, contour models, Markov random field model, and clustering [9, 10]. Section 2 highlights the segmentation related works in brain anomalies detection. Section 3 describes the choice of segmentation algorithm for MR brain images and finally conclusion is drawn in Sect. 4.

2 Segmentation Algorithms for Brain Anomalies Detection

A wide number of segmentation algorithms are there for medical images. In [11, 12], a detailed analysis has been performed on segmentation algorithms for CT, MR, and mammogram images. Some of the widely used algorithms for the detection of anomalies in MR brain images are as follows.

2.1 Thresholding

Otsu suggested a nonparametric and unsupervised threshold selection from zeroth and first-order moments for the segmentation of gray-level images [13]. Kalavathi suggested a method for the optimal selection of the threshold value from the Otsu multiple thresholding techniques for the brain tissue segmentation and obtained better results in terms of non-overlapping measures [14]. Olivo proposed a histogram characterization technique with the wavelet transform to obtain zero crossing and local extrema which helps in coarse to fine variation analysis. The variation in histogram was used for the automatic selection of threshold [15]. Harris et al. suggest a segmentation technique for MR brain images based on supervised thresholding of gray values by estimating the brightness of the image and segments the GM and WM of the brain [16]. Sandhya et al. introduced a novel technique for the extraction of WM, GM, and CSF by multilevel thresholding integrated with electromagnetism optimization algorithm. This method gives better results when compared with the classical clustering techniques coupled with optimization algorithms [17].

2.2 Region Growing

Sandor proposed region growing algorithm for the segmentation of CSF and brain tissues of CT brain images [18]. Kayle et al. developed a semi-automatic technique for the segmentation of 3D MR brain images; the seed point is initially selected for segmentation which achieves better results when compared with the manual segmentation of 3D brain image [19]. Pohle developed a fully automatic segmentation algorithm based on adaptive region growing that characterizes the region to be segmented by the homogeneity criterion and the shape properties of the region [20]. Xuan integrated region growing algorithm with edge detection for the ROI extraction in MR brain images which eliminates the problem of over-segmentation [21]. Tang et al. proposed a $T2$ weighted MR brain image segmentation with multiresolution edge detection based on structural connectivity and intensity inhomogeneity, and intensity threshold is selected automatically for the segmentation of brain tissues [22].

2.3 K-*means Clustering*

Sharma et al. discussed various types of unsupervised segmentation based on the clustering [23]. Abras et al. proposed a modified *K*-means algorithm for the automatic segmentation of CSF, WM, GM, and background [24]. Madhukumar et al. made a qualitative comparison of FCM and *K*-means algorithm with histogram guided initialization of the center pixel for the brain segmentation and the results reveal that *K*-means performs better than FCM algorithm [25]. Yan et al. proposed an adaptive *K*-means algorithm for the extraction of GM, WM, and CSF of 3D MR brain images. This method adapts the local intensity variation of the ROI and it is insensitive to shading effect [26]. Vemuri et al. proposed adaptive *K*-means segmentation algorithm with multiresolution wavelet transformation for fast convergence by the Gibbsian prior and β computation. This algorithm segments bone, WM, GM, CSF, and background with assigned gray values [27]. Ng et al. proposed a hybrid algorithm with *K*-means and watershed transformation. The *K*-means algorithm initially clusters the image regions and edges were filtered by Sobel operator and segmentation is done by watershed transformation [28].

2.4 *Fuzzy C-Means Clustering*

Yoon et al. proposed FCM algorithm for the extraction of GM and WM of MR brain images, which reduces the computation time and the algorithm was evaluated with symmetry measures [29]. Nevin A. Mohamed et al. introduced an improved FCM algorithm with the parameters of Markov Random Field (MRF) which is less sensitive to noise that eliminates the problem of over-segmentation in classical FCM algorithm [30]. Balafar suggested a novel multiscale medical image segmentation algorithm based on FCM that minimizes the problem of noise and intensity inhomogeneity in the MR images [31]. Chuang et al. proposed a novel FCM algorithm by the inclusion of the spatial information in the membership function for clustering which includes neighborhood of each pixel. This method is homogeneous and less sensitive to noise [32]. Szilagyi et al. introduced a γ factor in the FCM which enhances the speed for computation and applied it to the bias-corrected FCM that results in good quality segmentation of brain [33]. Pham modified FCM with the inclusion of objective function employed in Adaptive Fuzzy *C*-means algorithm (AFCM) and was less sensitive to noise. The efficient results are produced, when compared with Fuzzy and Noise Tolerant Adaptive Segmentation Method (FANTASM) and classical FCM [34]. Taherdangkoo et al. improved the segmentation performance of the FCM algorithm by the ABC algorithm with the changes in the variables λ and ζ that improves the performance of membership function [35].

2.5 Hidden Markov Random Field (HMRF)

Held et al. developed an unsupervised algorithm for the 3D segmentation of brain with Markov random field algorithm. Simulated annealing and iterated conditional modes are also implemented which classifies efficiently GM, WM CSF, scalp-bone, and background [36]. Bricq et al. developed a Hidden Markov Chain (HMC) that considers neighborhood information for the ROI extraction in multimodal brain MR images [37]. Ibrahim et al. introduced Hidden Markov Models (HMMs) for the 3D MRI segmentation that handles complexity in segmentation and artifacts, [38]. Chen et al. developed region-based hidden Markov random field model (RBHMRF) for the segmentation of the MR brain images. The results reveal that RBHMRF performs better when compared with HMRF and FGM algorithms [39]. Zhang et al. proposed HMRF model along with EM algorithm which segments the brain tissues accurately based on spatial neighborhood pixels. The HMRF-EM is compared with FCM-EM model which is slow in computation by 10% [40]. Guerrout et al. developed a segmentation technique for MR images using HMRF model with MAP criterion [41]. Ruan et al. proposed an unsupervised method for the segmentation of MR brain images based on stochastic gradient algorithm and fuzzy MRF model [42].

2.6 Watershed Segmentation

Parvati et al. coupled Marker-Controlled Watershed algorithm and region growing algorithm for grayscale images [43]. Grau et al. suggested a modified version of the watershed algorithm to overcome the disadvantages of over-segmentation; noise sensitivity and thin edges. The prior probability estimation incorporated in the algorithm performs better for knee cartilage and WM/GM segmentation in MR images [44]. Haris et al. developed a hybrid method for the segmentation of multidimensional MR image with edge-preserving partition by watershed transform and region adjacent graph for the contour extraction. This method performs even better with low SNR medical images [45]. The watershed and thresholding algorithms are coupled for the delineation of tumor in MR brain images [46]. The watershed with texture-based region growing algorithm yields efficient results for facial MR images [47]. The watershed segmentation was used for the delineation of breast tumors in ultrasound images and for classification; self-organizing map (SOM) is used [48].

3 Choice of Segmentation Algorithm for MR Brain Images

The clustering algorithm was found to be efficient in the analysis of MR images of brain. Also, the clustering technique was found to be better when compared with other classical segmentation algorithms and are depicted in results and discussion.

(i) Fuzzy C-means Clustering

The FCM is a soft clustering approach, while K-means clustering is a hard clustering approach. The objective function of the FCM algorithm is defined as follows and is represented in Eq. below.

$$J = \sum_{x=1}^{P} \sum_{y=1}^{C} O_{x,y}^{f} d\left(u_x, v_y\right) = \sum_{x=1}^{P} \sum_{y=1}^{C} O_{x,y}^{f} \left\| u_x, v_y^2 \right\|$$

where C is the number of clusters and in the context of image processing, it represents the number of segmented regions; P is the number pixels; f is the fuzzifier intensity greater than 1; O_{xy} is the membership function of the attribute u_x to cluster j.

The O_{xy} satisfy the following conditions represented below

$$O_{xy} \in [0, 1]; \sum_{y=1}^{C} O_{xy} = 1$$

O_{xy} is defined as the fuzzy membership of the pixel u_x in the yth cluster; v_y is the yth cluster center.

The membership function and cluster center is updated and is expressed as follows:

$$O_{xy} = \frac{1}{\sum_{c=1}^{C} \left(\frac{D_{xy}}{D_{xc}} \right)^{\frac{2}{f-1}}} = \frac{1}{\sum_{c=1}^{C} \left(\frac{\left\| u_x, v_y \right\|^2}{\left\| u_x, v_c \right\|^2} \right)^{\frac{1}{f-1}}}$$

and

$$V_x = \frac{\sum_{x=1}^{N} O_{xy}^{f} u_x}{\sum_{x=1}^{N} O_{xy}^{f}}$$

where D_{xy} and D_{xc} are the distance measures.

(ii) K-Means Clustering

K-means clustering is an unsupervised self-learning algorithm. K-means clustering consists of K number of classes which is user-defined. The pixels are grouped into any one of the classes based on similarity and it is termed as hard clustering algorithm. The distance between cluster centroid of each class and each pixel is calculated by Euclidean distance metrics. The cluster centroid is initialized randomly and is iteratively updated. The preprocessing is required for classical K-means segmentation algorithm, since it is sensitive to noise.

The steps in classical K-means algorithm as follows:

1. The number of clusters and centroids are initialized randomly.

2. Estimate the distance between the cluster centroids and each pixel $I(x, y)$

$$D = \|I(x, y) - C_k\|^2$$

3. Each centroid defines a cluster, group pixels with minimum Euclidean distance to the cluster center.
4. Cluster centers are updated

$$C_k = \frac{\sum_{x \in k} \sum_{y \in k} I(x, y)}{n \in k}$$

5. Repeat steps 3 and 4, until convergence criteria are attained; difference in Euclidean distance is less than the threshold value.

(iii) Gaussian Mixture Model

The Gaussian Mixture Model algorithm is also based on soft clustering approach. It is a generative model in which, each cluster generates mean and variance. The pixel intensity values are represented in terms of probability density function represented as Gaussian mixture models.
The steps for Gaussian Mixture Model algorithm is summarized below

1. Random Initialization of the probability distribution values in all clusters

$$\mu_n = \frac{n+\max(i)+1}{j+1}$$
$$\sigma^2 = \max(i) + 1$$
$$p(C_n) = \frac{1}{j}$$

2. The probability distribution function is represented as follows, where the cluster C_n is represented by Gaussian distribution $N(\mu_n, \sigma_n)$ and of y_i represents the probability belonging to any class C_n

$$p(C_n|y_i) = \frac{p(y_i|C_n) * p(C_n)}{p(y_i)}$$
$$p(y_i|C_n) = \frac{1}{\sqrt{2\pi\sigma}} * \exp\left(\frac{-(y_i - \mu_n)^2}{2\sigma_n^2}\right)$$
$$p(y_i) \sum_p p(y_i|C_n) * p(C_n)$$

where $p(y_i|C_n)$ is likelihood, $p(C_n)$ is prior knowledge and $p(y_i)$ is evidence.

3. Update the mean, standard deviation, and probability values.

$$\mu_n = \frac{\sum_i p(C_n|y_i) * y_i}{\sum_i p(C_n|y_i)}$$
$$\sigma_n = \frac{\sum_i p(C_n|y_i) * (y_i - \mu_n)^2}{\sum_i p(C_n|y_i)}$$
$$p(C_n) = \frac{\sum_i p(C_n|y_i)}{n}$$

4. Repeat step 2 and 3 until the convergence criteria is achieved. The convergence criteria is defined by threshold value for the minimal change in the values of mean and standard deviation for the iteration number specified.

4 Results and Discussion

The widely used segmentation algorithms in medical imaging like thresholding, region growing, edge detectors, clustering, and watershed algorithms are validated on real-time and database MR images. The algorithms are developed in Matlab 2010a and for performance validation of clustering algorithms, brain web database images are used, since gold standard images are there in the database. The input MR images are depicted in Fig. 1.

The thresholding is a simple segmentation algorithm for the extraction of region of interest and it is applicable for the images with simple objects. Adaptive thresholding is a widely used technique in many applications. In this work, for thresholding, two techniques are employed; Classical thresholding and Wellner's adaptive thresholding. The parameters of the adaptive thresholding are local window size, threshold value, and filtering. The input medical images acquired are always subjected to noise, in general, the CT images are degraded by Gaussian noise, MR images are degraded by rician noise and ultrasound images are degraded by speckle noise. For filtering mean and median filtering options are incorporated in the thresholding algorithms. The threshold value changes based on the input image and universal threshold value cannot be chosen. The thresholding results are depicted in Fig. 2.

The Wellner's adaptive thresholding is another thresholding technique employed for the medical images. The threshold value is changed in accordance with the local mean or median values in the image.

The parameters of the Wellner's adaptive thresholding are filter size, filter type and a constant value that relies on the mode which is determined as percentage or fixed amount of local average or median gray values. The constant value is positive for the extraction of dark objects and is negative for the extraction of white objects. The constant value is set within the range −20 to +20. The thresholding algorithm is

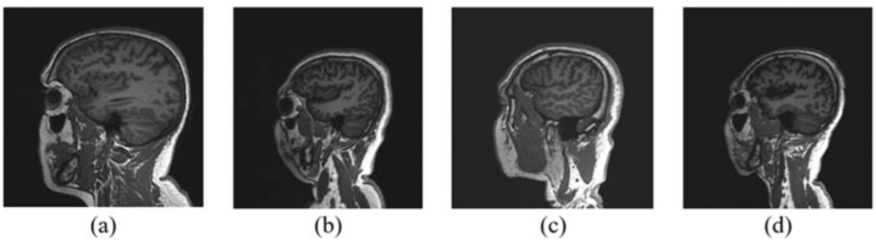


Fig. 1 Input MR brain images (ID1-ID4)

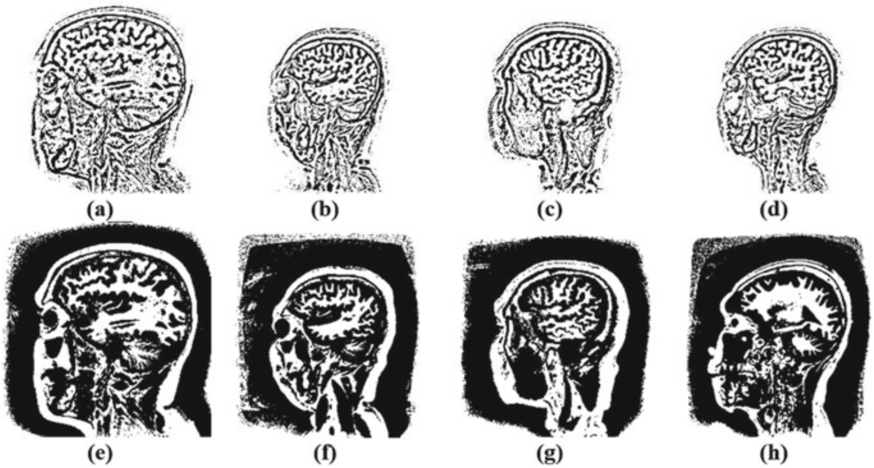


Fig. 2 Classical thresholding in row 1 and Wellner's adaptive thresholding in row two segmentation results corresponding to dataset (ID1-ID4)

sensitive to noise and hence an appropriate preprocessing technique is vital prior to segmentation. In the cases of images with complex objects, multilevel thresholding is employed to yield efficient results. The optimization algorithms are also coupled with the thresholding technique to yield robust results.

The region growing explores the neighborhood connectivity of pixels and adds the pixels to the selected seed point based on gray-level similarity. The seed points have to be specified and it varies for each input image. The manual selection of seed point is crucial in region growing algorithms and automated region growing algorithms based on optimization techniques are there for seed point selection. Figure 3 depicts the region growing segmentation results.

Edge detection is also a simple and classical segmentation algorithm that traces the boundary of objects in the image. In this work, the following edge detectors; Sobel, Prewitt, Roberts, Log, and Canny are tested on medical images. The canny edge detector produces superior results when compared with the other classical techniques

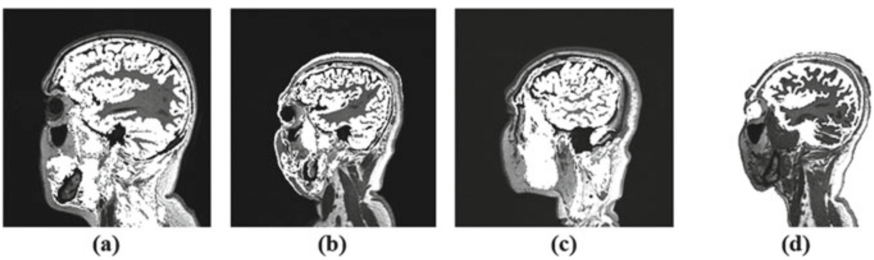


Fig. 3 Region growing segmentation results corresponding to dataset (ID1-ID4)

in terms of the computation time and the quality of edges detected. The canny edge detector is efficient for noisy images too, the parameters of canny edge detector are standard deviation of Gaussian filter and threshold values ($T1$ and $T2$). Higher value of σ implies the localization of larger edges. The classical edge detectors are based on the gradient operator and are an initial step in object detection and recognition. Figures 4, 5 and 6 depict the edge detector algorithms output.

The clustering is a widely used data mining algorithm and its role is prominent in medical image processing. Here, the classical FCM, K-means, and GMM clustering are employed. The FCM segmentation results are depicted in Fig. 7. The K-means segmentation results are depicted in Fig. 8 and Gaussian Mixture Model results are depicted in Fig. 9.

The HMRF model and EM algorithm framework were also applied for MR brain images. In [49], HMR-EM segmentation model was initially proposed for the MR brain images. The HMRF-EM algorithm is termed as an edge preservation model that generates a prior segmentation result by K-means clustering algorithm. The prior segmentation output supplies the parameters for the MAP algorithm and EM algorithm. The parameters $X(0)$ for the MAP algorithm and $\theta\hat{}[(0)]$ for the EM

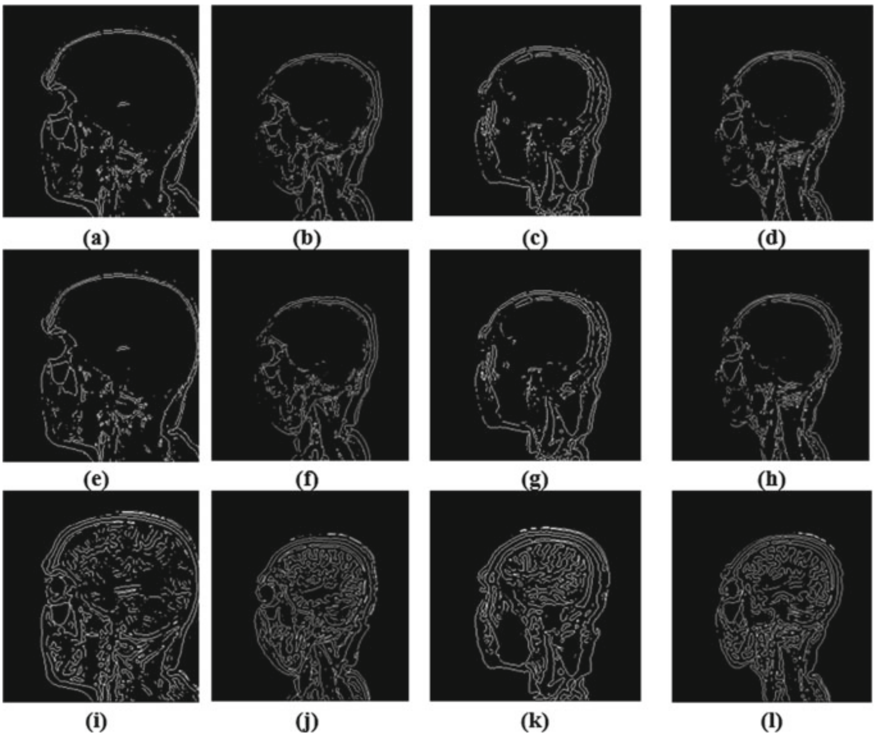


Fig. 4 Sobel edge detector (row 1), Perwitt edge detector (row 2), log edge detector (row 3) segmentation results corresponding to dataset (ID1-ID4)

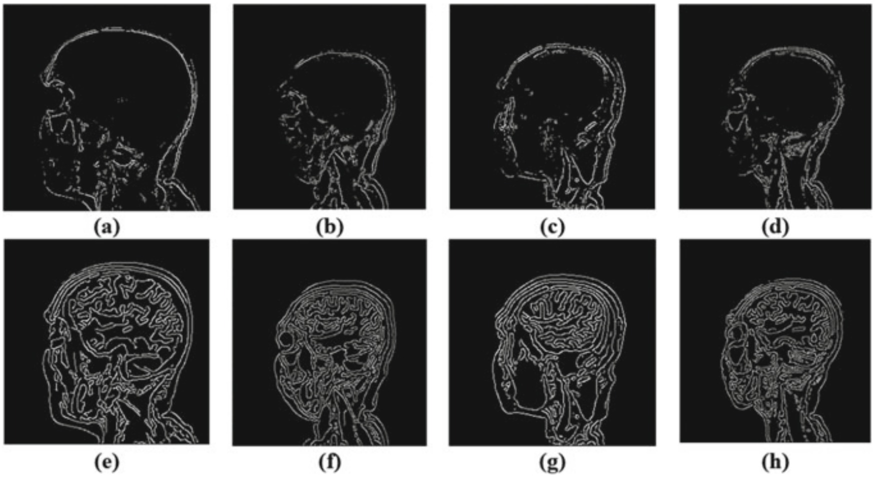


Fig. 5 Robert edge detector (row 1) and Canny edge detector (row 2) segmentation results corresponding to dataset (ID1-ID4)

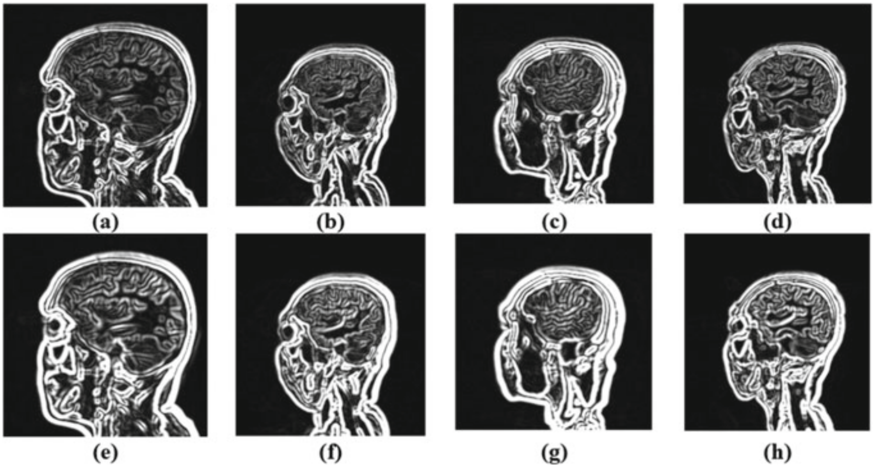


Fig. 6 Gauss gradient edge detector segmentation results corresponding to dataset (ID1-ID4)

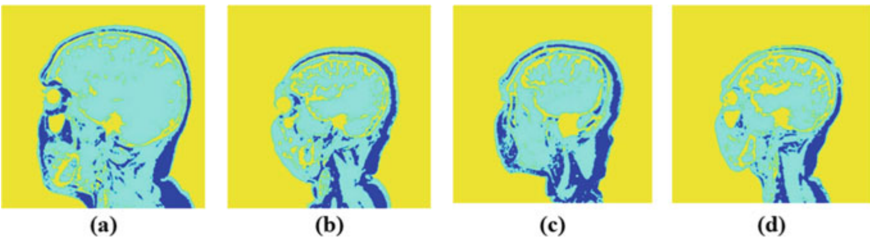


Fig. 7 Fuzzy *C*-means segmentation results corresponding to dataset (ID1-ID4)

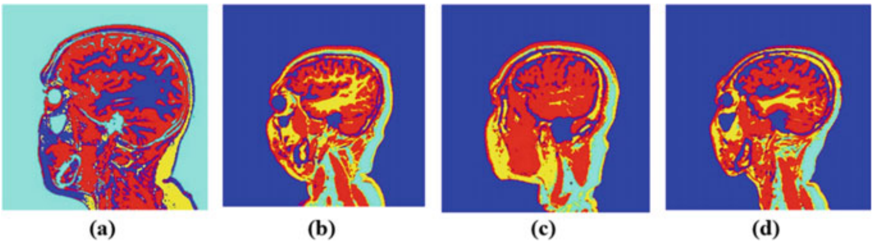


Fig. 8 *K*-means segmentation results corresponding to dataset (ID1-ID4)

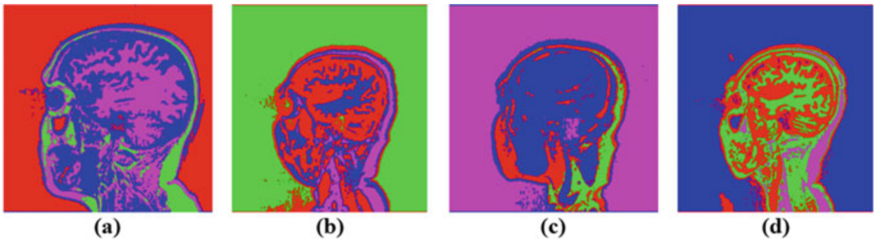


Fig. 9 GMM segmentation results corresponding to dataset (ID1-ID4)

algorithm are determined from the *K*-means clustering algorithm. In the proposed work, for the *K*-means clustering algorithm, number of clusters = 4 is chosen and the number of iterations for EM and MAP algorithm is set to 10. The hidden Markov random field segmentation results are depicted in Fig. 10.

The following algorithms are considered for qualitative analysis; Thresholding-I, Region-Based-II, Edge-based-III, Clustering-V, Watershed-VI, Markov Random Fields-X. The characteristics of segmentation algorithms are evaluated based on the parameters characteristics [12] and numerical values are assigned for each characteristic. The comparative analysis of segmentation algorithms are depicted in Table 1.

The performance analysis reveals that clustering technique is having a high score, when compared with other algorithms. Hence the clustering technique is found to be

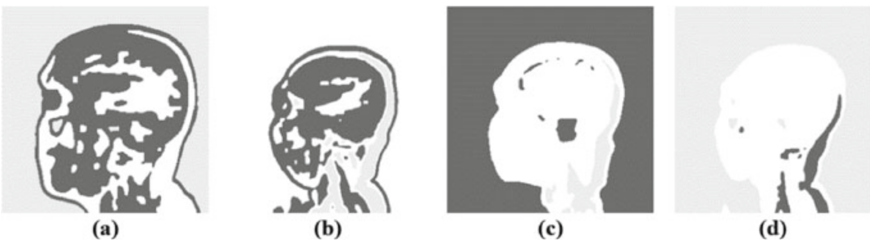


Fig. 10 HMRF segmentation results corresponding to dataset (ID1-ID4)

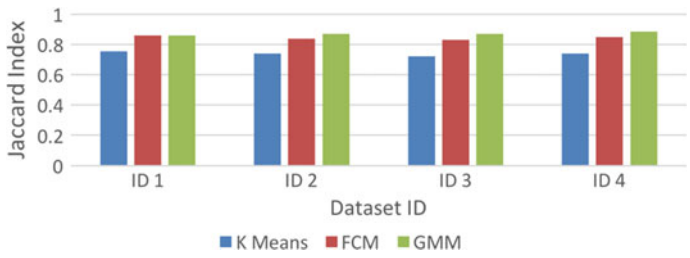


Fig. 11 Performance plot for Jaccard Index

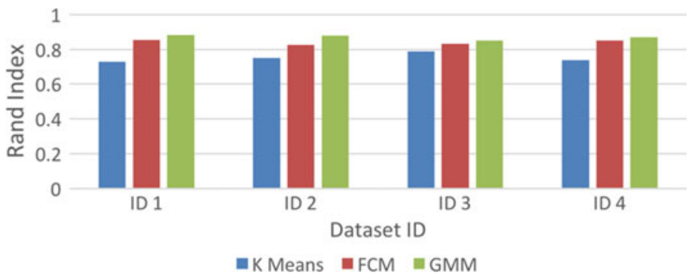


Fig. 12 Performance plot for rand index

Table 1 Comparative analysis of segmentation algorithms

Characteristics of algorithms	Segmentation algorithms					
	I	II	III	V	VI	X
Spatial information	1	2	1	2	2	3
Region continuity	2	2	1	3	2	3
Noise immunity	1	2	2	2	1	2
Parameters selection	2	1	3	2	2	1
Complexity	3	2	3	2	2	1
Computation time	3	2	3	2	2	2
Accuracy	1	2	1	3	2	3
Overall score	13	13	14	16	13	15

efficient for the delineation of anatomical organs and anomalies in MR brain images. The three clustering algorithms *K*-means, FCM, and GMM results are depicted here for real-time medical images. For the validation of segmentation algorithms, ground truth images are needed, hence the clustering techniques are also tested on brain web data base images (ID1, ID2, ID3, and ID4) and the results are depicted in Fig. 13.

The Jaccard Index is proportional to the amount of spatial overlap between groundtruth image and the segmented image. The rand index estimates the similarity of pixels in the segmented image and groundtruth image. The value of Jaccard

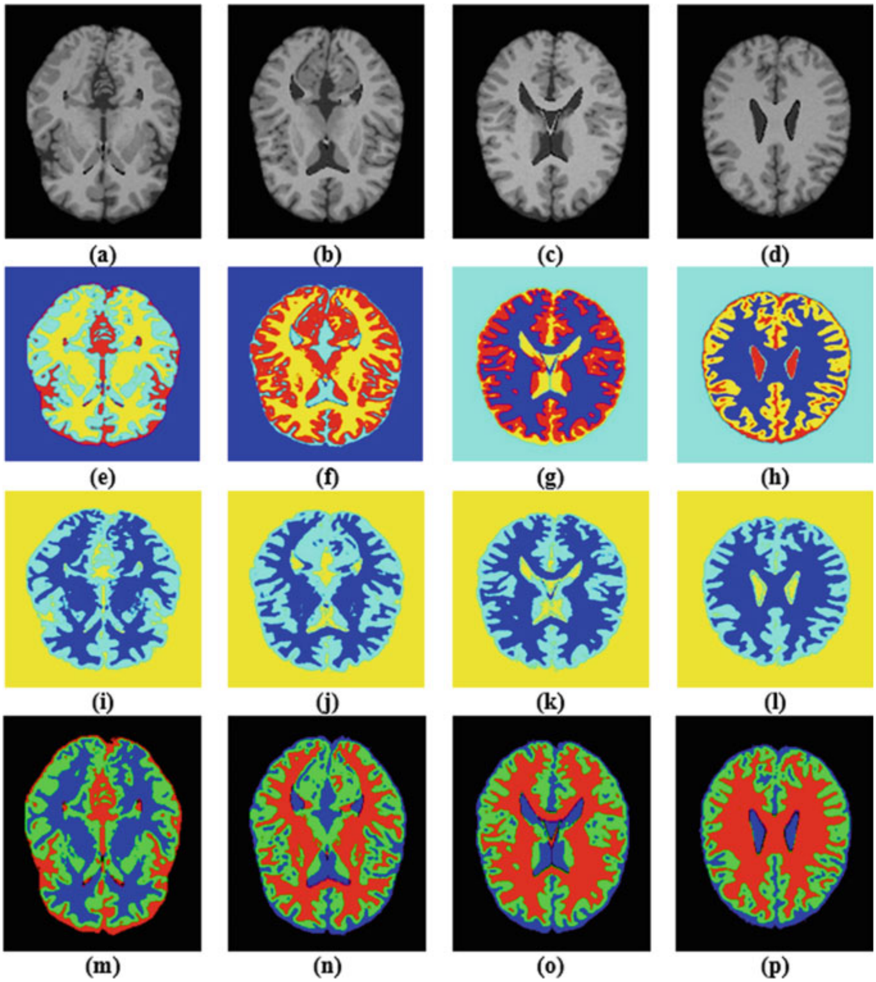


Fig. 13 Clustering algorithms results on brain web data base images; **a**–**d**: Input MR brain images (ID1-ID4), **e**–**h**: *K*-means clustering, **i**–**l**: FCM, **m**–**p**: GMM

Index and Rand Index range for 0 (Poor matching) to 1 (Perfect matching). The performance metrics plots of segmentation algorithms are depicted in Figs. 11, 12, 14 and 15 (Fig. 13).

The false positive is a measure of improper segmentation as ROI when it is not the ROI. The false negative is a measure of improper segmentation not as ROI when it is really ROI.

$$\mathrm{JI} = \frac{|\mathrm{TP}|}{|\mathrm{TP}| + |\mathrm{FN}| + |\mathrm{FP}|}$$

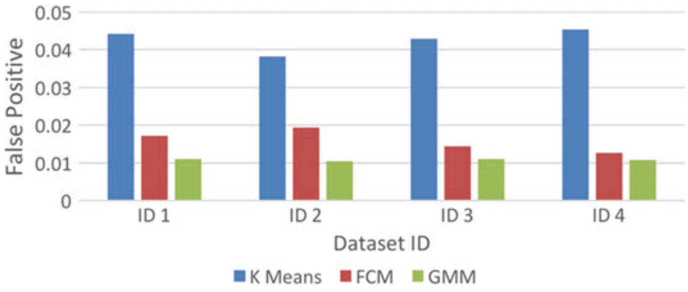


Fig. 14 Performance plot for false positive

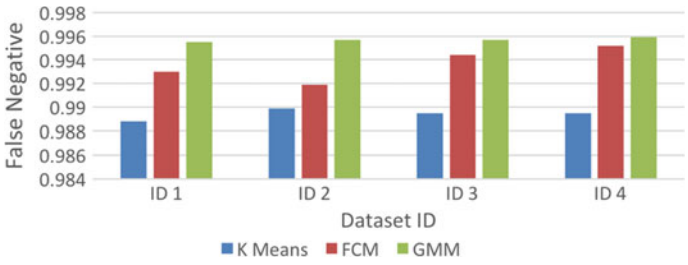


Fig. 15 Performance plot for false negative

$$\mathrm{RI} = \frac{|\mathrm{TP}| + |\mathrm{TN}|}{|\mathrm{TP}| + |\mathrm{FN}| + |\mathrm{TN}| + |\mathrm{FP}|}$$

5 Conclusion

This research work analyses various segmentation algorithms on medical images. The qualitative and quantitative analysis has been performed on various algorithms and tested on real-time and benchmark images. For the validation of segmentation algorithms, performance metrics are used. The quantitative analysis reveals that clustering is an appropriate algorithm for the MR brain images. The three variants of clustering algorithms are considered in this research work; *K*-means, FCM, and GMM algorithms. The GMM and FCM clustering algorithm generate promising results. This research work thus analyses various segmentation algorithms and its characteristics thus pave a way for the selection of an appropriate algorithm.

Acknowledgements The authors would like to acknowledge the support provided by Nanyang Technological University under NTU Ref: RCA-17/334 for providing the medical images and supporting us in the preparation of the manuscript.

References

1. Yasmin M, Mohsin S, Sharif M, Raza M, Masood S (2012) Brain image analysis: a survey. World Appl Sci Journal 19(10):1484–1494
2. Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophys J 76(2):691–708
3. Ritter F, Boskamp T, Homeyer A, Laue H, Schwier M, Link F, Peitgen HO (2011) Medical image analysis. IEEE Pulse 2(6):60–70
4. Williamson SJ, Lü ZL, Karron D, Kaufman L (1991) Advantages and limitations of magnetic source imaging. Brain Topogr 4(2):169–180
5. Bandyopadhyay SK (2011) A survey on brain image segmentation methods. J Glob Res Comput Sci 2(2)
6. Senthilkumaran N, Rajesh R (2011) Brain image segmentation. Int J Wisdom Based Comput 1(3):14–18
7. Pham DL, Xu C, Prince JL (2000) Current methods in medical image segmentation. Annu Rev Biomed Eng 2(1):315–337
8. Khan W (2013) Image segmentation techniques: A survey. J Image Graphics 1(4):166–170
9. Balafar MA, Ramli AR, Saripan MI, Mashohor S (2010) Review of brain MRI image segmentation methods. Artif Intell Rev 33(3):261–274
10. Norouzi A, Rahim MS, Altameem A, Saba T, Rad AE, Rehman A, Uddin M (2014) Medical image segmentation methods, algorithms, and applications. IETE Tech Review 31(3):199–213
11. Kumar SN, Muthukumar S, Kumar H, Varghese P (2018) A voyage on medical image segmentation algorithms. Biomed Res (0970–938X)
12. Kumar SN, Fred AL, Varghese PS An overview of segmentation algorithms for the analysis of anomalies on medical images. J Intell Syst
13. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66
14. Somasundaram K, Kalavathi P (2013) Brain tissue segmentation in MR brain images using otsu's multiple thresholding technique. In: Proceeding of international conference on computer science and education, Colombo, pp 639–642
15. Olivo JC (1994) Automatic threshold selection using the wavelet transform. CVGIP: Graph Model Image Process. 56(3):205–218
16. Harris GJ, Barta PE, Peng LW, Lee S, Brettschneider PD, Shah A, Henderer JD, Schlaepfer TE, Pearlson GD (1994) MR volume segmentation of gray matter and white matter using manual thresholding: dependence on image brightness. Am J Neuroradiol 15(2):225–230
17. Sandhya G, Babu Kande G, Savithri TS (2017) Multilevel thresholding method based on electromagnetism for accurate brain MRI segmentation to detect white matter, gray matter, and CSF. BioMed Res Int 2017
18. Sandor T, Metcalf D, Kim YJ (1991) Segmentation of brain CT images using the concept of region growing. Int J Biomed Comput 29(2):133–147
19. Justice RK, Stokely EM (1996) 3D segmentation of MR brain images using seeded region growing. In: Proceedings of 18th annual international conference of the IEEE engineering in medicine and biology society, vol 3. IEEE, pp 1083–1084
20. Pohle R, Toennies KD (2001) Segmentation of medical images using adaptive region growing. In Medical imaging 2001: image processing, vol 4322. International Society for Optics and Photonics, pp 1337–1347
21. Xuan J, Adali T, Wang Y (1995) Segmentation of magnetic resonance brain image: integrating region growing and edge detection. In Proceedings., international conference on image processing, vol 3. IEEE, Washington, DC, pp 544–547
22. Tang H, Wu EX, Ma QY, Gallagher D, Perera GM, Zhuang T (2000) MRI brain image segmentation by multi-resolution edge detection and region selection. Comput Med Imaging Graph 24(6):349–357
23. Sharma N, Aggarwal LM (2010) Automated medical image segmentation techniques. J Med Phys/Assoc Med Phys India 35(1):3

24. Abras GN, Ballarín VL (2005) A weighted k-means algorithm applied to brain tissue classification. J Comput Sci Technol 5
25. Madhukumar S, Santhiyakumari N (2015) Evaluation of k-means and fuzzy C-means segmentation on MR images of brain. Egypt J Radiol Nucl Medicine 46(2):475–479
26. Yan MX, Karp JS (1994) Segmentation of 3D brain MR using an adaptive K-means clustering algorithm. In: Proceedings of 1994 IEEE nuclear science symposium-NSS'94, vol 4. IEEE, Norfolk, VA,, pp 1529–1533
27. Vemuri BC, Rahman S, Li J (1995) Multiresolution adaptive K-means algorithm for segmentation of brain MRI. In: International computer science conference. Springer, Berlin, pp 347–354
28. Ng HP, Ong SH, Foong KW, Goh PS, Nowinski WL (2006) Medical image segmentation using k-means clustering and improved watershed algorithm. In: 2006 IEEE southwest symposium on image analysis and interpretation. IEEE, Denver, pp 61–65
29. Yoon OK, Kwak DM, Kim DW, Park KH (1999) MR brain image segmentation using fuzzy clustering. In: FUZZ-IEEE'99. 1999 IEEE international fuzzy systems. Conference proceedings (Cat. No. 99CH36315), vol 2. IEEE, Seoul, pp 853–857
30. Mohamed NA, Ahmed MN, Farag A (1998) Modified fuzzy c-mean in medical image segmentation. In: Proceedings of the 20th annual international conference of the IEEE engineering in medicine and biology society. Vol 20. Biomedical engineering towards the year 2000 and beyond (Cat. No. 98CH36286), vol 3. IEEE, Hong Kong, pp 1377–1380
31. Balafar MA, Ramli AR, Saripan MI, Mahmud R, Mashohor S, Balafar M (2008) New multi-scale medical image segmentation based on fuzzy c-mean (FCM). In: 2008 IEEE conference on innovative technologies in intelligent systems and industrial applications. IEEE, Cyberjaya, pp 66–70
32. Chuang KS, Tzeng HL, Chen S, Wu J, Chen TJ (2006) Fuzzy c-means clustering with spatial information for image segmentation. Comput Med Imaging Graph 30(1):9–15
33. Szilagyi L, Benyo Z, Szilágyi SM, Adam HS (2003) MR brain image segmentation using an enhanced fuzzy c-means algorithm. In: Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society (IEEE Cat. No. 03CH37439), vol 1. IEEE, Cancun, pp 724–726
34. Pham DL (2001) Robust fuzzy segmentation of magnetic resonance images. In: Proceedings 14th IEEE symposium on computer-based medical systems. CBMS 2001. IEEE, Bethesda, pp 127–131
35. Taherdangkoo M, Yazdi M, Rezvani MH (2010) Segmentation of MR brain images using FCM improved by artificial bee colony (ABC) algorithm. In: Proceedings of the 10th IEEE international conference on information technology and applications in biomedicine. IEEE, Corfu, pp 1–5
36. Held K, Kops ER, Krause BJ, Wells WM, Kikinis R, Muller-Gartner HW (1997) Markov random field segmentation of brain MR images. IEEE Trans Med Imaging 16(6):878–886
37. Bricq S, Collet C, Armspach JP (2008) Unifying framework for multimodal brain MRI segmentation based on hidden markov chains. Med Image Anal 12(6):639–652
38. Ibrahim M, John N, Kabuka M, Younis A (2006) Hidden markov models-based 3D MRI brain segmentation. Image Vis Comput 24(10):1065–1079
39. Chen T, Huang TS (2005) Region based hidden markov random field model for brain MR image segmentation. In: WEC (2):233–236
40. Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 20(1):45–57
41. Guerrout EH, Mahiou R, Ait-Aoudia S (2013) Medical image segmentation on a cluster of PCs using Markov random fields. Int J New Comput Arch Appl (IJNCAA) 3(1):35–44
42. Ruan S, Bloyet D, Revenu M, Dou W, Liao Q (2002) Cerebral magnetic resonance image segmentation using fuzzy Markov random fields. In: Proceedings IEEE international symposium on biomedical imaging. IEEE, pp 237–240

43. Parvati K, Rao P, Mariya Das M (2008) Image segmentation using gray-scale morphology and marker-controlled watershed transformation. Discret Dyn Nat Soc 2008
44. Grau V, Mewes AU, Alcaniz M, Kikinis R, Warfield SK (2004) Improved watershed transform for medical image segmentation using prior information. IEEE Trans Med Imaging 23(4):447–458
45. Haris K, Efstratiadis SN, Maglaveras N, Katsaggelos AK (1998) Hybrid image segmentation using watersheds and fast region merging. IEEE Trans Image Process 7(12):1684–1699
46. Ng HP, Huang S, Ong SH, Foong KW, Goh PS, Nowinski WL (2008) Medical image segmentation using watershed segmentation with texture-based region merging. In: 2008 30th annual international conference of the IEEE engineering in Medicine and Biology Society. IEEE, pp 4039–4042
47. Mustaqeem A, Javed A, Fatima T (2012) An efficient brain tumor detection algorithm using watershed & thresholding based segmentation. Int J Image Graph Signal Processing 4(10):34
48. Huang YL, Chen DR (2004) Watershed segmentation for breast tumor in 2D sonography. Ultrasound Med Biol 30(5):625–632
49. Alecci M, Zhang Y, Brady JM, Jezzard P, Smith S (2000) Image-based evaluation of ap-ion'B1 field correction and its effect on MRI tissue segmentation. Proc Intl Sot Mag Reson Med 8

HEP-2 Specimen Image Segmentation and Classification Using GLCM and DCT Based Feature Extraction with CNN Classifier

C. C. Manju and M. Victor Jose

Abstract Effective Human Epithelial-2 (HEp-2) cell image classification can encourage the determination of numerous autoimmune system diseases Different computerized image processing methods can be successfully connected to perform HEp-2 segmentation and classification process. This work, a programmed framework for HEp-2 segmentation and classification utilizing image processing ideas is utilized. HEP-2 cell image classification (CIC) utilizing convolutional neural network (CNN) with Gray-Level Co-Occurrence Matrix (GLCM) and Discrete Cosine Transform (DCT) feature extraction (HEP-CIC-CNN-DCT-GLCM) is proposed in this work to improve the presentation of classification precision. Adaptive Gamma Correction (ADC) technique is used as preprocessing technique to improve the differentiation of the cell image for the segmentation process. Median Filtering technique method is utilized to expel noise from the image if any kind of distortion or cracks occurred while acquisition or transmission. Data augmentation is utilized in the training stage to improve the viability of training process. Different arrangement of data is created by pivoting pictures in various points to get more sample images to perform great training. Proposed technique can classify six classes such as Homogeneous, Speckled, Nucleolar, Centromere, Nuclear Membrane, and Golgi. Mean Class Accuracy (MCA) of about 96.56%, which is each a lot higher contrasted with past work related to accessible ICPR 2014 dataset.

Keywords Human Epithelial-2 (Hep-2) cell · Convolution neural networks · Gray-level co-occurrence matrix · Discrete cosine transform · Mean class accuracy · Adaptive gamma correction · Median filter

C. C. Manju (✉) · M. Victor Jose
Noorul Islam Centre for Higher Education, Thuckalay, India
e-mail: manjucc@gmail.com

M. Victor Jose
e-mail: mvictorjose@yahoo.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_12

1 Introduction

Pattern recognition methodologies are comprehensively employed in the world of drug for the headway of Artificial Intelligence systems. Immune system illnesses are established to be connected with the event of autoantibodies in patient bodily fluid [1]. The anti-nuclear anti-bodies (ANA) test is routinely used to break down connective tissue infections (CTD, for instance, crucial lupus erythematosus (SLE), and Sjogren's Syndrome. The most elevated quality level for playing out this test is the deviant immunofluorescence (IIF) convention using human epithelial sort 2 (HEp-2) cells on account of the declaration of a wide extent of antigens on HEp-2 cells. In any case, the show is time and work genuine. Likewise, there is a high intra-and between research center variety of the test [2].

Indirect immunofluorescence (IIF) on human epithelial type 2 (HEp-2) is that the brilliant normal strategy for ANA check. The pathologist watches pictures, containing refined cells of the HEp-2 cell line, increased under a fluorescence magnifying instrument and makes an end subject to two things, (1) the intensity of the fluorescence (negative, moderate, or positive) (2) sort of the fluorescence recoloring structure. As such, ANA test is enthusiastic as it depends upon the ability of the pathologist. ANA testing is best performed using the IIF technique with HEp-2 cells as the substrate. Regardless, it is enthusiastic and requires computerization. ANA testing yields staining patterns of HEp-2 cells, these recoloring examples are connected with invulnerable framework infections. Thus, the HEp-2 cell staining pattern classification calculation is the center of the ANA test computerization structure [3].

Starting late there has been creating energy for showing automated pattern order frameworks for microscopy pictures. The results from these casing works may offer an inexorably target classification which would improve result consistency and resolve any irregularities in the emotional assessments. Distinctive HEp-2 cell arrangement strategies have been represented and most rely upon covering windows with comparative sizes, which are called patch-based picture expert processing. The strategy first concentrates diverse visual features, for instance, local binary pattern (LBP) and scale-invariant feature transform (SIFT) from each picture patch, and a short time later addresses them through each picture patch, and afterward speaks to them through various changes, for example, privately obliged direct coding extraordinary pyramid coordinating, autonomous segment investigation, non-parametric Bayesian model, and so forth. The performance of the classification can be improved by using support vector machine or similar kind of classifiers [4]. Because of superior, CNN is generally a utilized system for classification problems.

In this work, a highly accurate HEP-2 cell image classification utilizing DCT and GLCM feature extraction with convolutional neural network is proposed to improve the classification accuracy. Six different types of classes can be classified by the proposed method highly accurately. HEP-2 cell images are collected from standard database and used with data augmentation technique is to generate more tranining data. Data augmentation techniques are employed to generate different data on different angles. The major contribution of this work is as follows:

(1) *Image acquisition and augmentation*: HEP-2 cell images are collected from standard database and used with data augmentation technique is to generate various training data to perform best training.
(2) *Preprocessing and Feature Extraction*: Adaptive gamma correction and median filtering techniques are used to increase the pixel quality in the spatial domain.
(3) *Algorithm Validation*: Proposed method is validated with various performance metric, for example, exactness and confusion metrics parameter to prove the efficiency.

The arrangement of this proposed work is as follows. Section 2 summarizes the previous works related to this paper. Section 3 gives detailed information about the proposed techniques which include the block diagram explanation and mathematical model, etc. In Sect. 4, result and discussion for proposed work with respective to the conventional classification techniques which is proposed previously. The final section concludes the paper with a detailed summary of HEP-CIC-CNN-DCT-GLCM.

2 Related Work

Deep Convolutional Neural Network [5]: presented the use of CNN with used 22 layered CNN with four pooling layers to separate more high-level features for cell image characterization. Contrasted with the current CNN-based HEp-2 order techniques, a system with more profound architecture is used to perform classification.

Binary Tree algorithm [6]: proposed a pattern classification framework for confused HEp-2 cell arrangement was de-marked and tried on ICPR 2016 HEp-2 dataset. From binary picture I B (s, t) territory of pixel, the normal gray-level, and fractal measurement for the limits were calculated. These gray levels were isolated into eight proportional intervals t, by then lower l and upper u breaking point was set to make the paired pictures. Binary Tree was utilized as a classification conspire.

Random Forest [7]: exhibited a hybridization approach for a class of FS methodology, viz. the channel-based FS procedures. The proposed strategy considers simply filter systems for the hybridization, as these methodologies are possibly Adhoc, i.e., they give feature situating, anyway the assurance of a perfect course of action feature is chosen by an experimental limit on highlight positioning (score) positioning (score). The paper displayed a limit of Random Forest (RF) and Random Uniform Forest (RUF) for deciding for HEp-2 cell picture arrangement.

Linear SVM [8]: paper present a local structure really as the differential vector of the encompassing pixels within one called as smaller scale Texton and model it as a parametric likelihood process with GMM. In addition, the proposed strategy removes the Fisher vector as demonstrated by the model parameters, which would be more discriminant for image portrayal and can be united with a linear classifier, (for example, a linear SVM).

Nearest Convex Hull Classifier [9]: introduced a cell classification framework involved a dual-region codebook-based descriptor joined with the Nearest Convex

Hull Classifier. The structure parts a cell picture into little patches, which are then accumulated into sets speaking to the inward and edge regions of the cell. The examination evaluated different variations of the descriptor on two straightforwardly available datasets: ICPR HEp-2 cell classification challenge dataset and the new SNPHEp-2 dataset.

Previously a number of methodologies proposed to perform HEP-2 cell image segmentation and classification. Most of the methods are not flexible with databases. When different images from database are used, the classification accuracy effected. CCN based method which is proposed in this work is stable when compared to the conventional works due to the deep learning capability. Detailed information about the proposed work is as follows.

3 HEP-CIC-CNN-DCT-GLCM: HEP-2 Cell Image Classification (CIC) Using Convolutional Neural Network (CNN) with Gray-Level Co-occurrence Matrix (GLCM) and Discrete Cosine Transform (DCT)

In this section, detailed information about proposed work is explained. Proposed deep CNN-based HEp-2 cell image classification system consists of two main sections such as CNN training and testing. Training and testing sections have the following image processing blocks (1) preprocessing, (2) data augmentation, (3) feature extraction, (4) CNN training, and (5) classification or CNN testing modules. Figure 1 shows the block diagram for the proposed method with various components and blocks, which can perform singular activity required for image segmentation and classification. For training stage, physically segmented images are preprocessed and corresponding labels are created and feed to the DCNN module. To improve the cell texture, GLCM is utilized in the gray-level pixels. Further, the pictures are trimmed and rotated to create different samples from single image to produce additionally training data. To extract feature GLCM and DCT methods are utilized in the individual cell pattern. The extracted features are from GLCM and DCT cascaded and fed to DCNN. Trained DCNN modules used for testing process with segmentation module. The watershed segmentation technique is used for the segmentation. Individual modules are feed to the trained DCNN module to predict the classes.

3.1 Image Preprocessing

A reasonable image preprocessing steps are performed to improve the pixel quality before main processing. That takes the normal for image into idea is crucial for deep CNNs to get extraordinary inside component depiction and classification. The brightness, noise level, and contrast of the HEp-2 cell image given by the ICPR

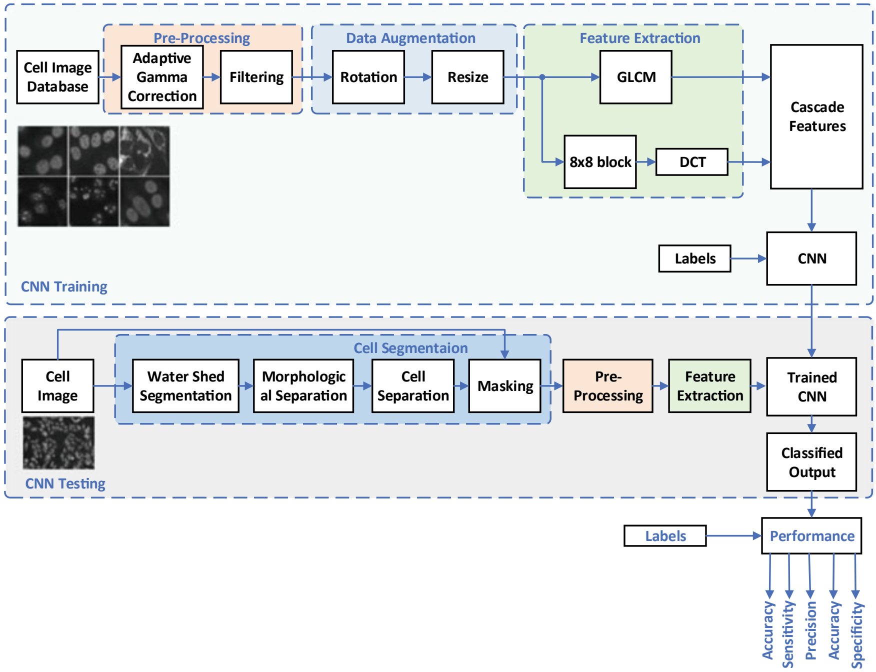


Fig. 1 Block diagram for the proposed method

2014 challenge (ICPR 2014 dataset) differ enormously. The preprocessing consists of Adaptive Gamma Correction, Filtering techniques which are further explained below sections.

Adaptive Gamma Correction To improve the contrast level of the picture, Adaptive Gamma Correction (AGC) Method is connected as preprocessing procedure. Adaptive gamma correction (AGC) strategy can progressively decide an intensity transformation function as indicated by the qualities of the input image. To begin with, pixel intensities must be scaled from the range (0, 255)–(0, 1.0). The mathematical model for adaptive gamma correction is given in (1).

$$O = I^{(1/G)} \tag{1}$$

where I is the input pixel matrix and G is gamma esteem. The yield picture O is then downsized to the range (0, 255). Gamma esteems < 1 will move the picture towards the darker finish of the range while gamma esteems > 1 will cause the picture to seem lighter [10]. A gamma estimation of $G = 1$ will have no effect on the information picture: Fig. 2 shows the original and gamma correction ($G = 1.5$) applied image.

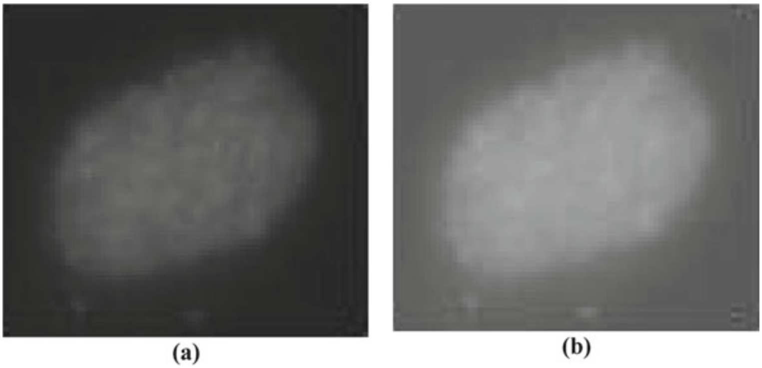


Fig. 2 **a** Original image, **b** gamma-corrected image

Filtering For exact classification, the picture ought to have ideal complexity and less noise. Noise is a subjective assortment of picture intensity and evident as a noteworthy part of grains in the image. It may convey at the period of getting or picture transmission. Noise suggests, the pixels in the image show particular power regards instead of real pixel regards that are gotten from picture. Noise departure figuring is the path toward removing or on the other hand lessening the noise from the image. The noise ejection figuring decline or empty the detectable quality of noise by smoothing the entire picture leaving regions close separation limits. In any case, these procedures can obscure fine, low separation nuances. Filters are utilized to expel noise from the picture. In our proposed work middle separating technique is utilized to evacuate undesirable dot noise. Median Filter is basic and all the more dominant nonlinear filter. It is utilized for diminishing the measure of intensity between one pixel from another pixel. The inside pixel estimation of the picture is supplanted by middle estimation of block.

3.2 Data Augmentation

HEp-2 cell picture given by the ICPR 2014 is used for our work. Two data augmentation plans were utilized to expand the size of training dataset and balance the picture volumes of various cell designs. In this dataset different sized images are available in various contrasts. The picture is trimmed for the six-classification procedure. Rotation of pictures is done at different angles such as 0°, 45°, 90°, 180°, and resized each picture to 28×28 to get consistency for the CNN preparing and diminish computational expense.

3.3 Feature Extraction

For feature extraction technique, Gray-Level Co-Occurrence Matrix (GLCM) and Discrete Cosine Transform (DCT) methodologies are used. GLCM method is used for texture feature extraction of the image. For frequency extraction from image, DCT technique is applied. The detailed explanation of feature extraction is given below:

Discrete Cosine Transform (DCT) DCT is an orthogonal transformation that is in all respects generally utilized in picture compression and is broadly acknowledged in the multimedia standards. DCT has a place with a group of 16 trigonometric transformations. The type-2 DCT changes a block of image of size $N \times N$ having pixel powers $s(n1, n2)$ into a changing array of coefficients $S(k1, k2)$, portrayed by the accompanying condition.

The changed array $S(k1, k2)$ got through (2) is additionally of the size $N \times N$, same as that of the original image block. It ought to be noted here that the change area indices $k1$ and $k2$ demonstrate the spatial frequencies in the ways of $n1$ and $n2$ individually. $k1 = k2 = 0$ relates to the normal or the DC component and all the staying ones are the AC segments which compare to higher spatial frequencies as $k1$ and $k2$ increment. From computational contemplations, it might be noticed that immediate use of the above condition to process the changed exhibit requires $O(N^4)$ calculations [11].

$$S(k1, k2) = \sqrt{\frac{4}{N^2}} C(k1)C(k2) \sum_{n1=0}^{N-1} \sum_{n2=0}^{N-1} s(n1, n2) * \cos\left[\frac{\pi(2n_1 + 1)k1}{2N}\right] \cos\left[\frac{\pi(2n_2 + 1)k2}{2N}\right] \quad (2)$$

where $k_{1,}k_2, n_1, n_2 = 0, 1, \ldots, N - 1$ and

$$C(k) = \begin{cases} 1/\sqrt{2} \text{ for } & k = 0 \\ 1 = \text{otherwise} \end{cases} \quad (3)$$

Gray-Level Co-occurrence Matrix (GLCM) Picture made out of a few pixels and every pixel having their very own intensity level. Gray Level Co-occurrence matrix [12] is a technique for organizing the pixels with different intensity levels. It is the oldest style second-request statistical methodology for texture examination. GLCM improves the accuracy level by picking feasible quantitative level for early analysis. In the underlying advance, the chief request statistical textural examination feature information of the image was isolated and frequencies of the gray level at arbitrary picture positions were assessed without considering neighbor pixels. In the subsequent development, the second-request textural examination feature were removed by considering neighbor pixels the factual highlights were isolated using

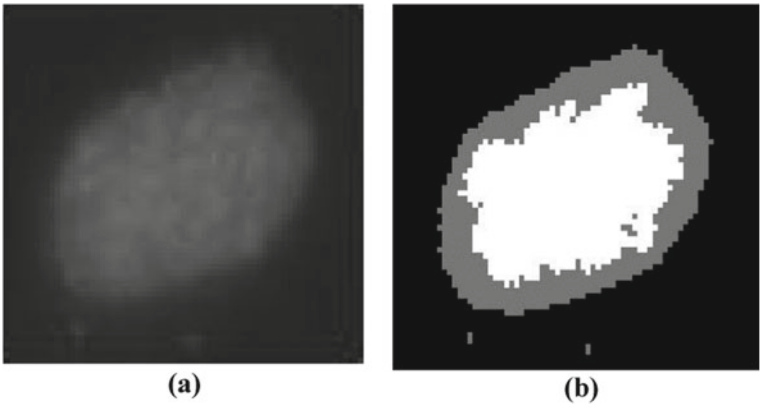


Fig. 3 GLCM input and output **a** input HEP-2 Cell, **b** GLCM of input image

GLCM, similar to shrewd known gray-level spatial dependence matrix (GLSDM). In GLCM methodology, gray-level co-occurrence matrix has the textural highlights, for example, differentiate, relationship, vitality, homogeneity, entropy was gotten from LL and HL sub gatherings of beginning four degrees of wavelet disintegration. The textural highlights extricated are recorded underneath. Figure 3 demonstrates the handled GLCM cell picture.

3.4 Network Architecture

Convolutional Neural Network (CNN, or ConvNet) is an exceptional kind of multi-layer neural framework, planned to see visual examples really from pixel pictures with irrelevant preprocessing. A genuine decision of system design is significant to CNNs. Our profound CNN shares the fundamental engineering of the old-style LeNet-5. The proposed technique utilizes five convolution layer previous layers will rehash 32 times and last will 64 times, 128, 256, 512. Convolution layers are within the structure square of a Con-volitional Network that does a large portion of the computational really troublesome work. The CONV layer's parameters include a lot of learnable channels. Each filter is little spatially (along width and stature), anyway connects through the full significance of the information volume. To speed up the training process and decrease the amount of memory used, we utilized pooling layer between layers. In our method, max pooling is used. In max pooling, a window passes over an image according to a set stride (how many units to move on each pass. The last layer is the classification layer. Classification layer consists of one or two fully connected layers. The yield layer of a CNN is responsible for delivering the likelihood of each class (every digit) given the input image.

3.5 CNN Training

Due to the non-calculated property of the cost furthest reaches of CNNs. sensible setting arrangement of training parameter is fundamental for the smooth join to the classification. CNN is parameterized by the loads and inclinations of different convolution layers and completely related layers. In our Deep CNN, for feature extraction, GLCM and (Discrete Cosine Transform) DCT techniques are used. GLCM is a texture element extraction method. The DCT resembles the discrete Fourier change: it changes a sign or picture from the spatial area to the recurrence space. The element extricated through GLCM and DCT is fell to prepare CNN. Ordinary of this the two yields are urged to prepare CNN.

3.6 CNN Testing

To classify test picture above preprocessing strategy is associated. The picture is then forward-impelled through the system, and the likelihood of this cell for each class is obtained. To further improve the power of course of action, we select four down to practically indistinguishable CNNs after the planning strategy ends up stable and use them in general for characterization portrayal by following. The foreseen class is the one having the most extraordinary yield probability found the center estimation of over the four probabilities.

Image segmentation is the division of a picture into regions or classifications, which contrast with different objects or parts of object. Every pixel in an image is assigned to one of a portion of these classes. An exceptionally straightforward procedure of segmenting cells from the info image is the thresholding system. Watershed Approach is a proficient strategy for cell segmentation from the entire picture, for the most part, the foundation of cell is dark in shading and dependent on the limit area it very well may be effectively isolated. Beforehand more specialists are utilized watershed way to deal with fragment the cell from input images. A blend of watershed and morphological activities-based strategy is connected to portion the cell from the picture. The watershed algorithm, in its standard use, is fully automatic. For morphological task, a blend of enlargement and disintegrations are due to channel the little spots and dots from the picture.

4 Result and Discussion

In this section detailed explanation of outcome of the work is described. Phase one describes the details of data set used. Second phase describes the Data Augmentation and third explains the result analysis. The analysis is done in MATLAB R2018a in I5 system with 6 GB RAM.

4.1 Dataset

This dataset contains 13,596 training cell pictures, and the test set is spared by the test facilitators and not released now. The cell images are expelled from 83 model pictures gotten by monochrome high special range cooled micros-duplicate camera fitted on an amplifying focal point with a plane-Apochromat 20×/0:8 objective point of convergence and a LED edification source. Each image has a spot with one of the six staining patterns: Homogeneous, Speckled, Nucleolar, Centromere, Nuclear Membrane, and Golgi. Figure 4 shows the six classes of cell types.

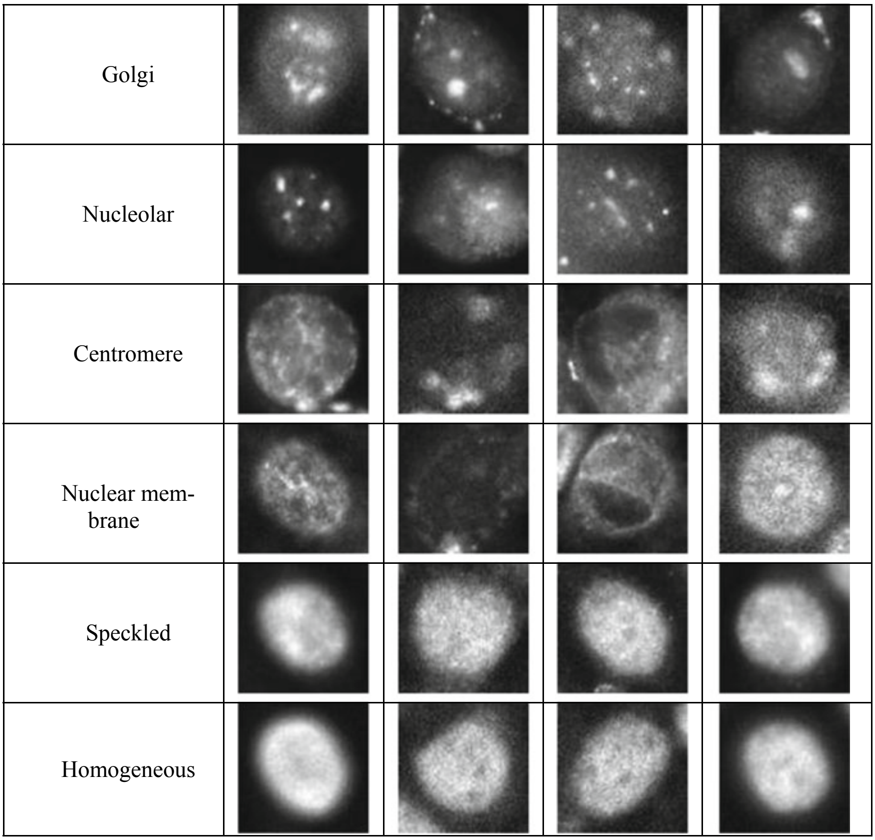


Fig. 4 Various type of cell

4.2 Adequacy of Data Augmentation

To get to more cell picture with accessible tests and reduce potential over-fitting of the prepared CNNs, we develop the preparation set by rotating each cell picture for 360°, with the movement of 0°, 45° and 90°,180°, independently. Thusly, the preparation set is stretched out by 1525 and on numerous occasions, and they are used to prepare the CNNs, independently. To improve the quality of our structure, we select four CNNs identifying with the 75th, 85th, 95th, and 100th ages after the framework adapting progress toward getting to be stable. A test image will encounter a comparative turn process and be commonly orchestrated by the four CNNs.

4.3 Result Analysis

The Deep CNN is prepared with ICPR 2014 dataset obtained with various research facility settings. CNN is prepared at various augmented cell (cell turned at 0°, 45°, 90°, 180°). All the trainable system parameters are invigorated for the adjusting process. To indicate straightforwardness and productivity change is practiced for 10 epochs which take on a very basic level less time than the 150 age spent in training CNN-Standard.

Mean Class Accuracy (MCA) for 20 epochs for CNN finetuning appears on Fig. 5 for different revolution point of the cell picture. As showed up by the line of "No revolution", CNN finetuning does not work admirably at the beginning. By the by, it gets up to speed rapidly in a few ages and achieves a fantastic exhibition in 10 epochs Fig. 6 shows the confusion matrix for the proposed method with various classes. From figure plainly the proposed framework is giving better execution as far as exactness. Every segment determines the exhibition of exactness for each

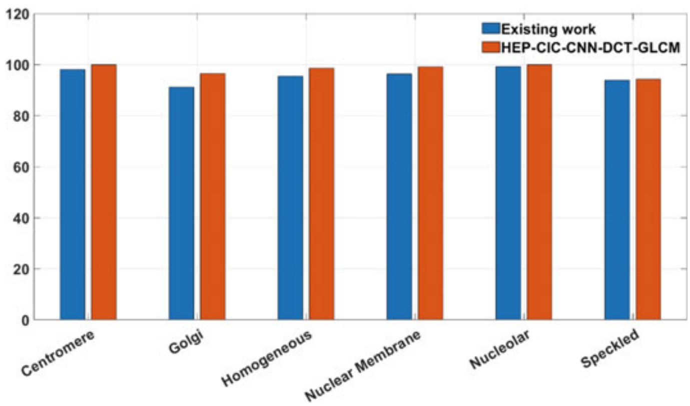


Fig. 5 MCA for 20 epochs after CNN finetuning at various angle steps

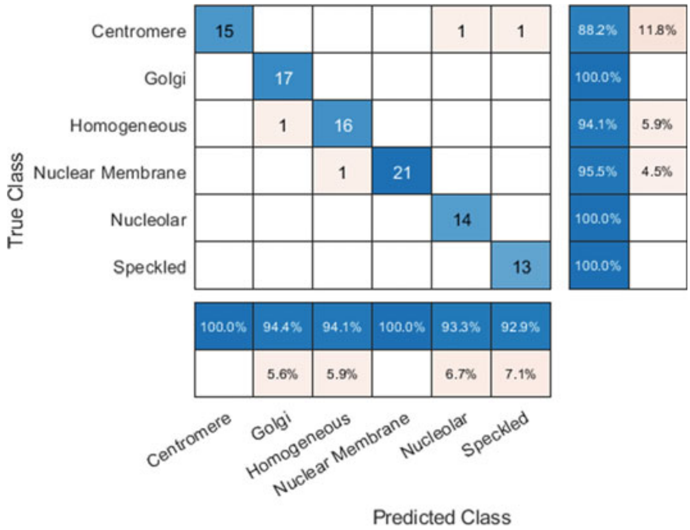


Fig. 6 Confusion matrix for proposed method with various class

class. From figure obviously every individual class demonstrating great execution. These outcomes exhibit the high effectiveness of the versatility of our CNN-based framework, particularly thinking about that there are unique classes of pattern over these datasets.

5 Conclusion

This paper represents an automatic HEp-2 cell staining pattern arrangement framework with deep convolutional neural frameworks. We give a point by point depiction on parts of this structure, carefully talk about different key issues that could influence its classification execution, and report a couple of captivating disclosures got from our investigation. In this work, a new architecture based on CNN with GLCM and DCT based feature extraction technique is suggested to classify the cell pattern. To generate efficient training data, rotation, and scaling with different angle is used. Different execution measures incorporate disarray framework and exactness is determined to approve the presentation. Precision results it is evident that the suggested framework is giving great execution as far as exactness and computational when contrasted with different works. The proposed work is prepared and tried in ICPR 2014 dataset and got MCA of 96.56%, which is the best exactness contrasted with past works.

References

1. Theodorakopoulos I (2014) Hep-2 cells classification via sparse representation of textural features fused into dissimilarity space. Pattern Recognit 47(7):2367–2378
2. Hobson P (2015) Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset. Artif Intell Med 65(3):239–250
3. Divya BS (2018) Human epithelial type-2 cell image classification using an artificial neural network with hybrid descriptors. IETE J Res 1–12
4. Ensafi S (2016) Accurate HEp-2 cell classification based on sparse coding of super pixels. Pattern Recognit Lett 82:64–71
5. Jia X (2016) Deep convolutional neural network based HEp-2 cell classification. In: 2016 23rd international conference on pattern recognition (ICPR), IEEE
6. Divya BS (2018) Human epithelial type-2 cell categorization using hybrid descriptor with binary tree. J Ambient Intell HumanizedComput 1–8
7. Gupta V (2018) Feature importance for human epithelial (HEp-2) cell image classification. J Imaging 4(3):46
8. Han XH (2014) Fisher vector of micro-texton for HEp-2 staining pattern classification. IFAC Proc Vol 47(3):3575–3580
9. Wiliem A (2013) Classification of human epithelial type 2 cell indirect immunofluorescence images via codebook-based descriptors. In: 2013 IEEE workshop on applications of computer vision (WACV), IEEE
10. Rahman S (2016) An adaptive gamma correction for image enhancement. EURASIP J Image Video Process 2016(1):35
11. Sorwar G (2001) Texture classification based on DCT and soft computing. In: 10th IEEE international conference on fuzzy systems, Cat. No. 01CH37297, vol 2. IEEE
12. Malegori C (2016) GLCM, an image analysis technique for early detection of biofilm. J Food Eng 185:48–55

Real-Time Traffic Signal Management System for Emergency Vehicles Using Embedded Systems

Cyriac Jose and K. S. Vijula Grace

Abstract The main objective of real-time traffic signal management system (RTSMS) is to save the life of people using cost-effective system. The proposed system is a straightforward and effective operable framework for the real-time embedded emergency vehicle (EV) management applications. In everyday life, there is consistent issue to a crisis vehicle, for example, rescue vehicle, ambulance and fireman to go through traffic light since it was red and it's exasperating the drive. This circumstance is frequently happening since resistance from a regular citizen, some of the time drivers don't have an encounter with this circumstance so they will hang tight for the traffic light to turn green. The RTSMS coordinates with the driver and assistants him in a cleared a path by keeping up a vital separation from a deferment in the surge hour gridlock. In this work, a new architecture for the emergency vehicle management system using an embedded system is proposed. This system is used to communicate between EV and traffic signal control unit. The processing time between the signal turn and decision making can be minimized by using the proposed hardware. The proposed work is simulated and tested with real-time hardware with different operating conditions.

Keywords Real-time traffic signal management system · Emergency vehicle · Microcontroller · Embedded system · On-board unit · Road side unit

1 Introduction

The fast advancement of metropolitan town areas into smart town areas and day by day expanded in number of vehicles utilizing the constrained street systems substructure causes an increment in rush hour gridlock support difficulties, for example, traffic clog, mishaps, and transmission EV [1]. For a smart city, a clever, secure and

C. Jose (✉) · K. S. Vijula Grace
Noorul Islam Centre for Higher Education, Thuckalay, India
e-mail: kcyriacjose@gmail.com

K. S. Vijula Grace
e-mail: vijulasundar@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_13

coordinated traffic control framework is required to guarantee smooth activity of everyday traffic just as oversee emergency vehicle development with the negligible postponement. Among accessible traffic control frameworks, none can right now consequently clear traffic clog if there should arise an occurrence of an emergency, permitting vehicles like ambulances, fire motors, and police vehicles to pass [2]. With the exponential development in a number of vehicles out and about, there is a solid requirement for Intelligent Traffic and Transport Systems for medical clinics, fire stations and police headquarters.

Survivability in risky conditions is directly associated with the reaction time of crisis organizations. In the specific occasion of heart failure, traffic deferrals of as small as three minutes can part the chances of a patient's survival capacity [3]. Notwithstanding authorization empowering crisis vehicles to progress through red traffic signals, security methods, holding up lines of vehicles, and pinnacle hour stop up moderate the development of these vehicles through the unions as they should ensure cross-traffic has ended to empower them to safely proceed. Intersection red traffic flag on occupied convergences while reacting to an emergency is the most perilous situation emergency administrations faculty faces every day.

Preceding achieving the crossing point, the emergency group may have just managed the flighty responses of drivers who are lined at the red traffic signal and have endeavoured to move out of their way, or the emergency team may have even needed to go on an inappropriate side of the street to pass traffic lined once more from the convergence. When they achieve the red sign and endeavour to cross the convergence, they may need to explore past people on foot crossing the street on a green walk signal, drivers who are going through a green sign and uninformed an emergency administrations vehicle is attempting to cross the convergence and the erratic conduct of drivers attempting to stay away from the emergency administrations vehicle. These situations place emergency administration faculty and the network at an expanded danger of coincidental damage or passing [4].

Our framework makes some assumptions also consider some situations it assumes that all the traffic signal lights are smart that is they are equipped with wireless communication systems and the controlling units also it neglects the transmission delay in the communication systems. It considers three criteria of emergency situations they are high, medium and low. When the high emergency situation arises, the system should accept the request quickly as soon as possible. For the high-level emergency situation, the traffic controller allows rerouting the traffic change the speed limit, change the behaviour of the traffic lights and also allows to modify the traffic policies. In a medium level situation, the traffic controller allows to reroute the traffic and also switch the traffic signals. In a low emergency situation, the system is allowed to change the traffic lights only.

2 Literature Review

Zheng and Ling [5] have presented emergency transportation arranging in a fiasco help store network the executives. The paper proposes a coordinated model of emergency transportation arranging, and build up an agreeable improvement strategy for productively taking care of the issue. The arrangement strategy utilizes three applicable positioning criteria for fluffy variable assessment, separates the incorporated issue into a lot of sub-parts, utilizes MOTS to enhance sub-answers for transportation task designation and asset distribution, utilizes MOGA to simultaneously improve populaces of sub-solutions for delivery scheduling and vehicle steering, and unites the outcomes to manufacture total arrangements of the coordinated issue. Test results demonstrate that the proposed strategy displays critical execution advantage in taking care of complex emergency transportation planning issues.

Li et al. [6] have proposed a progressive structure for intelligent traffic on the board in savvy urban areas. In this framework builds up a hierarchical multi-agent system (MAS) based operating system that reacts adequately to rising worries with traffic inefficiencies in keen urban communities. The proposed system strikes a harmony between nearby decisions and worldwide intelligence prerequisites to control traffic signal settings that can decrease potential traffic congestion and accelerate traffic streams. The usage of the proposed MAS improves the city's ability to deal with the traffic stream autonomously and respond proactively when unanticipated changes happen in rush hour gridlock streams. This secluded and summed up structure is adaptable to the organization of common traffic control components in which individual MAS operators are refreshed or supplanted as activity necessities differ, which guarantees the adaptability and extensibility of the proposed system.

Qi et al. [7] have presented a deterministic and stochastic Petri net (PNs) to plan an emergency traffic light control framework for crossing points furnishing emergency reaction to manage mishaps. As indicated by blocked intersection segments, as portrayed by unique PN models, the comparing emergency traffic light methodologies are intended to guarantee the security of a crossing point. The participation among traffic lights/offices at those influenced convergences and streets is represented. For the upstream neighbouring crossing points, a traffic-signal-based emergency control approach is intended to help counteract mishap prompted huge scale clog. Gridlock recuperation, live lock counteractive action, and compromise procedures are created. This framework utilizes PNs to demonstrate and plan an ongoing traffic emergency framework for convergences confronting mishaps. It very well may be utilized to improve the best in class progressively auto collision the executives and traffic security at crossing points.

Li et al. [8] propose a lot of calculations to configuration sign planning plans by means of profound support learning. The inside idea of this approach is to set up a deep neural network (DNN) to get acquainted with the Q-limit of fortress picking up from the tried traffic state/control inputs and looking at traffic system execution yield. In light of the got DNN, locate the proper sign planning arrangements by certainly displaying the control activities and the difference in framework states.

3 Proposed Method

This framework encourages the framework to coordinate with the driver and aides him in a cleared way staying away from deferral in the rush hour gridlock. Through GPS the car influxes and barricades are perceived by the framework and another safe course will be explored to the driver for achieving the goal within less time. A remote alarm will be given to the traffic superintendents if there should arise an occurrence of any trouble for the E-Vehicles to make room. The working model for the proposed method is shown in Fig. 1. The basic block diagram for the proposed method is shown in Fig. 2. In this method, we have used LPC1768 microcontroller which is from ARM CORTEX M3 family with 32 KB of flash memory, which acts as the brain for the system. Here the LCD display is used to know the current location and also used to know the messages from the control station. Initially, the microcontroller sends a message to the control station by using the GSM modem to clear the road. The minimum set of data contains information about the incident, including time, precise location, the direction the vehicle was travelling, and vehicle identification.

The traffic controller unit is responsible for clearing the road to make the way for the emergency vehicle. The traffic controller has the communication unit, data analyzing unit and controlling unit. The communication unit is responsible for receiving and sending data to the vehicle and the traffic signal lights. The data analyzing unit is a typical microcontroller that analyzes the received data and takes the necessary

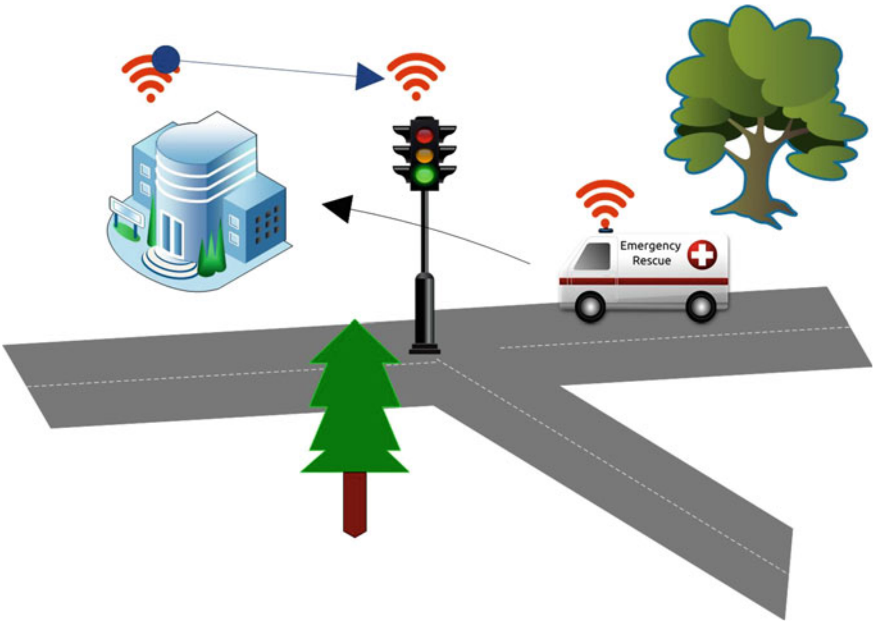


Fig. 1 Working model of the proposed system

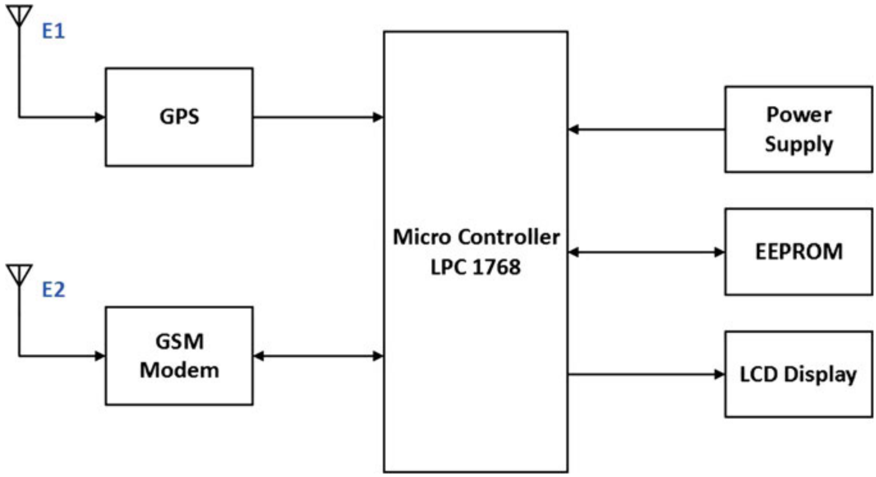


Fig. 2 Block diagram

actions. When the communication unit receives a request from an emergency vehicle it immediately transfer the data into the data analyzing unit extract the time, direction, incident and vehicle identification from the request and validate the request if the received data are valid then it accepts the request and send acknowledge to the vehicle then it sends the control commands to the traffic signal lights. When receiving a command from the traffic controller the traffic signal light switches on the green light till the emergency vehicle crosses the point. After that, the traffic controller switches into the normal working mode.

3.1 System Hardware

Cortex-M3 Figure 3 shows the circuit diagram used for the signal transmission ans reception in proposed system. The Cortex-M3 processor, in light of the ARMv7-M engineering, has a hierarchical structure. It coordinates the focal processor center, called the CM3Core, with cutting edge framework peripherals to empower incorporated capacities like intruding on control, memory insurance and framework debug and follow. These peripherals are exceptionally configurable to permit the Cortex-M3 processor to address a wide scope of utilizations and be all the more firmly lined up with the framework prerequisites. The Cortex-M3 center and the incorporated segments have been explicitly intended to meet the necessities of insignificant memory execution, decreased pin count and low power utilization.

GSM Module SIM 800 GSM modem is used to send and receive messages by using the wireless network. SIM 800 GSM module supports Quad-band. It is very compatible, flexible and plug and play type. It uses RS232 serial communication to

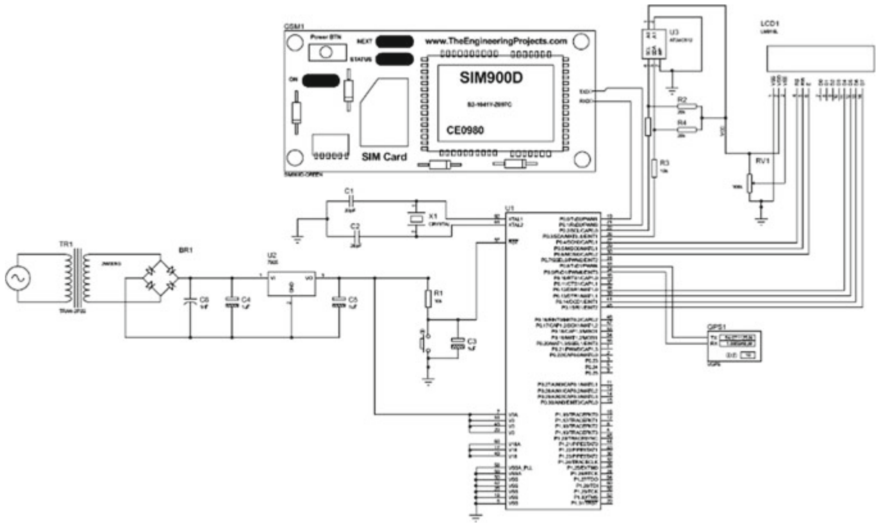


Fig. 3 Overall circuit diagram

interface with the microcontroller. It also supports voice services, data, SMS, GPRS and also TCP/IP stack (Fig. 4).

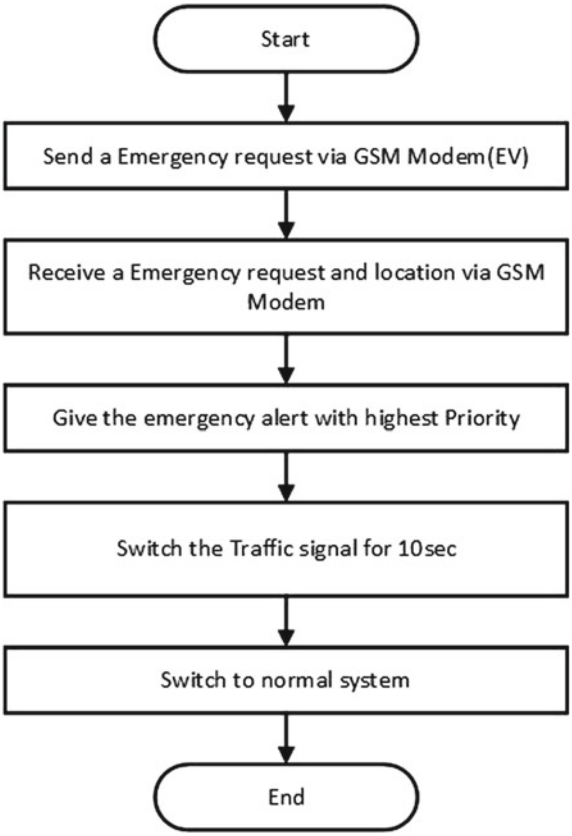
GPS Module (L80) GPS receiver (L80) is a device that is capable of calculating the exact geographical position of the device by processing the information from the GPS satellites. L80 GPS module is very compatible and it has the patch antenna and also it is power efficient. The main advantage of L80 is it can capable of fixing the position very quickly, it supports antenna switching function also it has always located technology.

AT24C512 is a non-volatile memory used to store the phone numbers that are used by the microcontroller. As the LPC1768 has small amount of memory it is very difficult to store the phone numbers in the memory of the microcontroller to overcome that problem AT24C512 is used. AT24C512 provides 524,288 bits of electrically erasable memory.

4 Results

The entire system is designed and simulated using Proteus software the software for the microcontroller is coded by using embedded c and that is built by using the Keilμ V5. Then it is dumped into the LPC1768 by using the SPI protocol through flash magic. For the controlling purpose we used HPLUS EX development board from CoiNel Technology Solutions, it is optimized to save development time and has a

Fig. 4 Flowchart for the overall Process



few important peripheral interfaces assembled for evaluation and testing. PCB for the system is designed by using Ki CAD V5.1.2. The manual exertion with respect to the traffic police officer is saved by using automatic applications. Human mediation is limited in light of the fact that the whole framework is computerized. Right, when a crisis vehicle approaches this system, it is viably distinguished by the structure as an emergency vehicle and traffic light exchanging is activated. This model acquaints a path with an arrangement with the answer to execute the idea of green waves in urban territories. Figure 5 shows the hardware model of the proposed method. The general structure is practical and has focal points over the customary progressions.

Figure 6 shows the logging information from the on-board unit that is the request from the unit to clear the traffic. This message includes the direction of traffic to be cleared (3rd line) the next one is vehicle identification number (line 5 and 6), the next information is the exact location of the vehicle which is given by the GPS module (18th line) after sending these information's the system will update its position for every 30 s also it will update the distance from the traffic signal lights.

Figure 7 shows the data receives from the GPS module. GPS data is displayed in different message formats over a serial interface. There are standard and non-standard

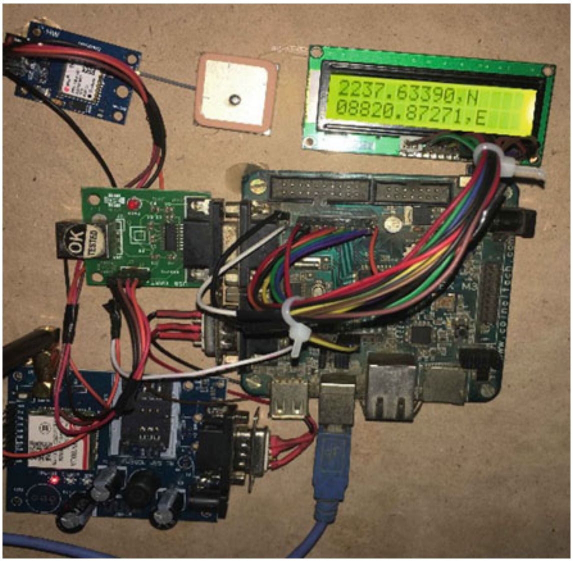


Fig. 5 Hardware model

formats. Nearly all GPS receivers' output is NMEA data. The NMEA standard is formatted in lines of data called sentences. Each sentence contains various bits of data organized in comma-delimited format. For example

$$\$GPGGA,010, 614.400, 4315.7323, N, 07953.7377,$$
$$W, 5, 3.65, 152.9, M.-34.7, M,, *61$$

The above GPGGA sentence (from Fig. 7 1st line) contains the following:

Time: 010614.400 is 01:06 and 14.400 s in Greenwich meantime
Longitude: 4315.7323, N is latitude in degrees. Decimal minutes, north
Latitude: 07953.7377, W is longitude in degrees. Decimal minutes, west
Number of satellites seen: 05
Altitude: 152.9 m.

5 Conclusion

In this work, a real-time traffic signal management system is implemented and tested in real-time environment using various hardware components. For instance, the sensors that are fixed over the emergency vehicle, fire motors and on squad cars empower the traffic signals to atomize their entry. By using the proposed system, emergency vehicle management can be achieved highly efficient way. Due to the availability of on-board units in all EV, the practical implementation of this system is also very

```
Compare direction=1
Exec time for check direction=862
PREEMPTING THE TRAFFIC ----W
PREEMPTING VEH----
1234
Inside store2_prem_vehi_info:search no-0
W
Course=14.64Executtion time preme inside fun=1055
Executtion time preempt=1100
Executtiontime=1461
5
5
1302
30000
Switch2-preempt
inside recv_msg
length-34
1234,12.917750,77.681600,5,185.235
First item in string: 1234
Second item in string: 12.917750
Third item in string: 77.681600
Fourth item in string: 5
Fifth item in string: 185.23
12.917751
77.681594
5.000000
185.229995
1234
Distance Towards traffic light-0.009815
Stage c- Preempt--Premption Vehicle=YES
Inside Check veh cross
```

Fig. 6 Logging information from on-board unit

easy when compared to the other complex computational units. Different advantages of RTLMSEV can lead to using in the real-time environment to manage all kind of EV's. The proposed system can be extended and used in autonomous vehicle.

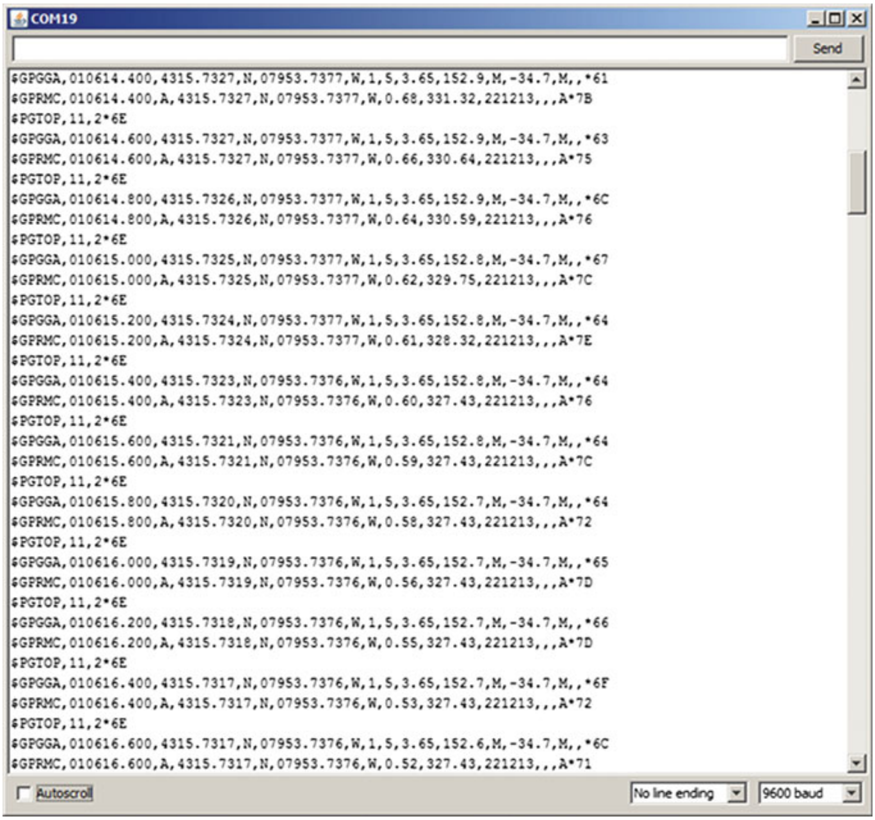


Fig. 7 GPS coordinates from GPS module

References

1. Chakraborty PS, Tiwari A, Sinha PR (2015) Adaptive and optimized emergency vehicle dispatching algorithm for intelligent traffic management system. Procedia Comput Sci 57:1384–1393
2. Chowdhury A (2016) Priority-based and secured traffic management system for emergency vehicle using IoT. In: 2016 international conference on engineering & MIS (ICEMIS). IEEE
3. Djahel S, Salehie M, Tal I, Jamshidi P (2013) Adaptive traffic management for secure and efficient emergency services in smart cities. In: 2013 IEEE international conference on pervasive computing and communications workshops (PERCOM workshops). IEEE
4. Sumi L, Ranga V (2018) Intelligent traffic management system for prioritizing emergency vehicles in a smart city. Int J Eng (IJE) IJE TRANSACTIONS B: Appl 31(2):278–283
5. Zheng YJ, Ling HF (2013) Emergency transportation planning in disaster relief supply chain management: a cooperative fuzzy optimization approach. Soft Comput 17(7):1301–1314
6. Zhiyi L et al (2017) A hierarchical framework for intelligent traffic management in smart cities. IEEE Trans Smart Grid 10(1):691–701

7. Liang Q et al (2015) Emergency traffic-light control system design for intersections subject to accidents. IEEE Trans Intell Transp Syst 17(1):170–183
8. Li L et al (2016) Traffic signal timing via deep reinforcement learning. IEEE/CAA J Autom Sin 3(3):247–254

Reduction of PAPR in Optical OFDM Signals Using PTS Schemes Without Side Information

A. Binita and P. P. Hema

Abstract Visible light communication provides bandwidth efficiency, secured communication and brightness along with data transmission simultaneously. Optical orthogonal frequency-division multiplexing (optical OFDM) in VLC systems can accomplish high information rate while transmission guarantees high dependability over the multipath fading condition; consequently, it has been embraced as a standard strategy in different communication systems that function as wireless. For decreasing peak-to-average power ratio of optical orthogonal frequency-division multiplexing signals, two partial transmit sequence (PTS) schemes without side information (SI) are proposed. Since recognizable phase alteration is connected to the components of every rotating vector, without transmitting SI, distinguish a rotating vector. The maximum likelihood (ML) detector is utilized to partition SI from the signal that is being obtained at the receiver and recuperate the information data stream. The Euclidean separation between the input information signal constellation and the signal constellation that is being rotated using the phase offsets is being abused by this ML detector. It is researched how to pick good phase offsets for implanting SI, by doing pairwise error probability (PEP) examination.

Keywords Optical orthogonal frequency-division multiplexing (Optical OFDM) · Peak-to-average power ratio (PAPR) · Partial transmit sequence (PTS) · Subblock partitioning · Side information (SI) · Pairwise error probability (PEP)

A. Binita (✉) · P. P. Hema
Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India
e-mail: binitaaaron@gmail.com

P. P. Hema
e-mail: pp.hema@mbcet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_14

1 Introduction

This visible light communication (VLC) that paved way to optical wireless communication has sprouted out as a promising technique to radio frequency because of its distinguishing advantages like unpredictable bandwidth, reduced power consumption, less interference in the electromagnetic spectrum and security in communication [1]. For high data rate excellency and maximum reliability in multipath fading, we use orthogonal frequency-division multiplexing (OFDM) in optical wireless communication (OWC) which enables transmission of data in Gigabit range [2].

Orthogonal frequency-division multiplexing (OFDM) has various independent orthogonal adjacent subcarriers that are being spaced at a very small distance to each other. When the subcarriers are added up in time domain on the performance of IFFT, high PAPR occurs says [3]. High PAPR is the result of phase differentiated subcarriers. At each point of instance, the value of PAPR varies with respect to each subcarrier at various phase values. When the value of PAPR at every point achieves a peak threshold at the same time, the output signal envelope rises causing a ‘peak’ in the signal output. Since OFDM signal has numerous subcarriers which are being modulated independently, the peak value of considered OFDM signals will be much greater than its average value [1]. For the OFDM signals considered in LTE systems, the PAPR is approximately 12 dB. The high PAPR causes the degradation in the performance of power amplifiers at the transmitter side, since operation at such large back off must be done as in [3].

High peak amplitudes with an immense dynamic range are the main disadvantages of optical OFDM signals; thus, it requires radio frequency power amplifiers with large peak-to-average power ratio than single-carrier systems due to leakage of the discrete Fourier transform as quoted in [4].

PAPR gives the relationship of the maximum power of a sample signal/symbol in each optical OFDM transmit symbol to the average power of that optical OFDM signal/symbol as described in [3]. In simple terms, PAPR is the ratio of peak power to the average power of a signal. The SI unit of PAPR is decibels (dB).

In a multicarrier system, if different subcarriers present in the system are out of phase with each other, the peak-to-average power ratio occurs. At different phase values, they are different with respect to each other, at each instant. The output envelope will suddenly shoot up which causes a ‘peak’ in the output envelope, when all the points achieve the maximum value simultaneously.

Since many subcarrier components are added via an inverse fast Fourier transformation (IFFT) operation, the transmit signals in an optical orthogonal frequency-division multiplexing (optical OFDM) system can have high peak values in the time domain. As a result, when compared to single-carrier systems, OFDM systems are known to have a high peak-to-average power ratio (PAPR) as seen in [5, 6]. In fact, the high PAPR decreases the signal-to-quantization noise ratio (SQNR) of the analog-to-digital converter (ADC) and digital-to-analog converter (DAC) and is one of the most detrimental aspects in an optical OFDM system. The efficiency of the power amplifier in the transmitter is also being degraded. Since the efficiency of the power amplifier is critical due to the limited battery power in a mobile terminal, the PAPR problem is more of a concern in the uplink. Optical OFDM signal PAPR is approx. 12 dB in LTE system. The PAPR of an optical OFDM signal of peak amplitude ‘s’ is given by,

$$\mathrm{PAPR}(s) = \max\left(|s|^2\right)/E\left(|s^2|\right) \tag{1}$$

where s is the peak amplitude of the optical OFDM signal, $\max(|s^2|)$ is the peak power and $E(|s^2|)$ is the average power.

As a solution to this high PAPR of OFDM signals, various schemes have been suggested like joint channel estimation [7], pilot symbol pattern [8], multi-points square mapping [5], cyclic shifted sequences [6], clipping [9], selective mapping (SLM), tone injection (TI), partial transmit sequence (PTS) schemes, etc. Among all these methods, PTS scheme is well suitable because we could reduce the PAPR of OFDM signals with least distortion of the output signals.

This paper encapsulates two schemes using PTS without transmission of side information (SI). These PTS schemes provide a very less complex detection of rotating vector that is being chosen in the receiver side. The PTS schemes endorse the selected signal index which is referred to as the side information (SI) by choosing rotating vectors into sequences of the signal through noticeable phase offsets given to elements of each rotating vector. Phase offset selection and their number (i.e., the number of required phase offsets) is done by taking into account the subsequences and their number (i.e., the total number of subsequences) of PTS scheme; the factors with which the phase is being rotated are being shifted using the offsets.

The extraction of side information from the signal that is obtained at receiver and the recovery of the data sequence are done using a maximum likelihood detector. The ML detector utilizes the Euclidian distance that is present in between the constellation of the input signal and the constellation of the signal that is being rotated using the phase offsets. Pairwise error probability (PEP) analysis helps us to choose suitable phase offsets for encapsulating the side information.

The PTS scheme will not alter the pilot tones; thus, the efficiency in the spectrum usage and the accuracy in channel estimation are achieved at the same time. This PTS scheme does not increase the number of constellation points, but it just rotates the original signal constellation using specific number of phase offsets, and hence, it proves to be computationally efficient.

The remaining framework of this work is as follows: in Sect. 2, the conventional PTS scheme is being depicted. In Sect. 3, the transmitter of an ACO-OFDM is being depicted. In Sect. 4, the transmitter section of HACO-OFDM is being discussed. In Sect. 5, the novel PTS scheme in HACO-OFDM signal and the ML detector for the recovery of the original signal from the received signal is being discussed. Section 6 is the performance comparison using the simulation results of the ACO-OFDM, PAM-DMT and the HACO-OFDM signals using PTS schemes. Section 7 encapsulates the inferences.

2 PTS Schemes

PTS scheme is a relevant PAPR reduction technique used in the case of HACO-orthogonal frequency-division multiplexing (HACO-OFDM). But, in the case of recovery of data, the receivers should possess side information, e.g., the phase offset values obtained from transmitters. Thus, we get large computational inefficiency which is said in [10].

The basic theme behind PTS is that the blocks of data are being divided into orthogonal/overlap-free subdivided blocks of data which do not depend on the rotating factor. The rotation element gives the data in the time domain with comparatively low amplitude. The basic theme of this technique is the division of the actual/initial optical OFDM data symbol into divided sub-data which can be transmitted via the subdivided blocks of data. These are then being multiplied by weighing factors which differ in the phase rotation factor unless we choose the optimum value with least PAPR as said in [10]. The sequence of data '*X*' in the frequency domain is being divided into '*v*' small sequences. These are being transmitted as overlap-free equal-size subblocks. The size of each subblock of data maybe '*N*', with *N*/*V* nonzero elements in each subdivided blocks of data. Hence, the PAPR is being decreased in the HACO-OFDM signal on using the PTS scheme. The major disadvantage of this technique is that the search complexity is being elevated exponentially as the number of subblocks increases. This conventional scheme is depicted in Fig. 1.

The PTS schemes discussed below provides efficiency in computations regarding the searching of sequences with minimum PAPR, spectrum usage; also, accuracy in channel estimation is achieved at the same time (i.e., using the below discussed PTS Schemes we could find the sub-sequence with minimum PAPR, also the PTS Schemes helps in efficient spectrum usage and accuracy in channel estimation too). The compared PTS schemes are computationally efficient also.

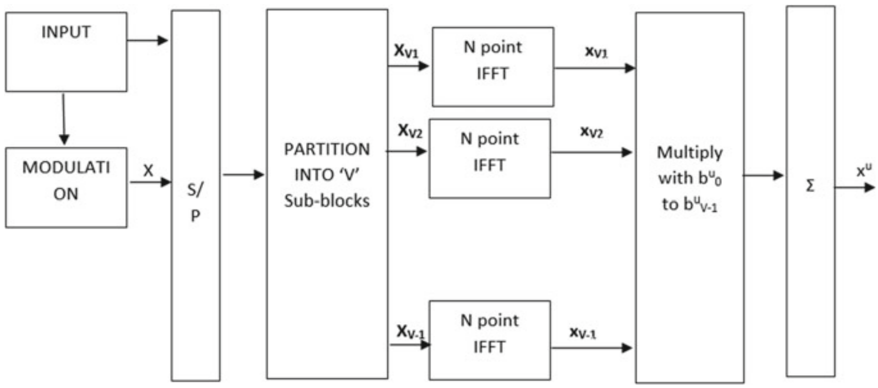


Fig. 1 Block diagram of conventional PTS scheme

3 ACO-OFDM Sequence

The input data stream which is in the time domain is being subjected to N-QAM modulation, and the resultant frequency domain data stream is being pinned on to the odd indices of the IFFT block within the permissible frequency band to get the frequency domain output values. Consider real-valued outputs of IFFT to enhance the efficiency in computations. Instead of taking the real outputs, we can take the Hermitian symmetry using positive and negative pairs of frequency. This method creates a no power output at imaginary values says [6].

The parallel stream of outputs from IFFT is being converted to serial data stream using a parallel-to-serial converter. The serial data bits can now be converted into analog signal using a digital-to-analog converter. Clipping of negative signal component can be done either before or after the DAC.

When negative samples are set to zero, the even subcarriers reduce intermodulation distortion. A DC component is generated as a result of clipping, and this DC component is being considered as the mean optical power or non-aided self-biasing since it is at an optimum level even without a control unit as in [6].

Figure 2 depicts the ACO-OFDM generation, and Fig. 3 is the diagrammatic representation of the clipping process that accompanies. Similar is the generation of PAM-DMT signal; the only difference is using PAM modulation for mapping the input bits.

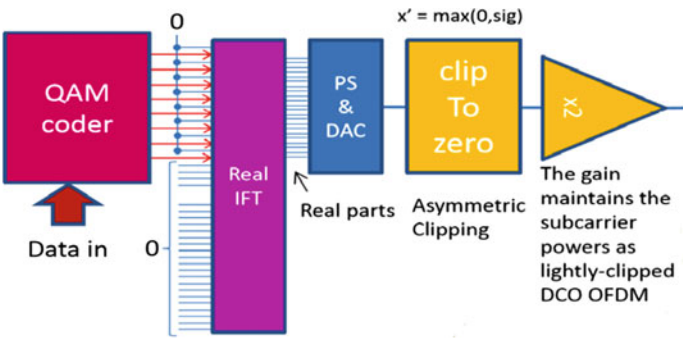


Fig. 2 Block diagram of odd-subcarrier ACO-OFDM system

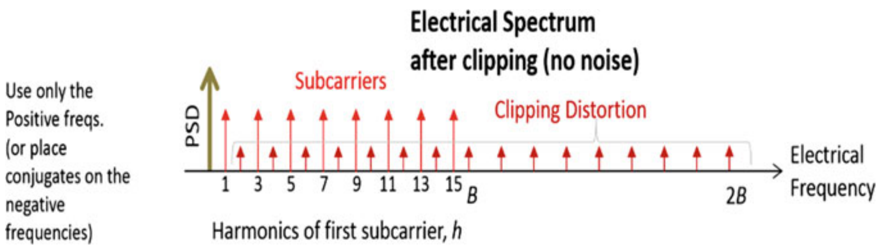


Fig. 3 Illustrative spectrum after the clipping process

4 HACO-OFDM

In optical wireless communication (OWC) systems, the transmitted signals will be positive due to the peculiar operating nature of the optical modulator, that is, the optical modulator gives only digital signals as output because there are only two operational levels basically a ‘zero’ or a ‘one’ which can be visibly indicated using the ON or OFF state of an LED. For ensuring the positivity of signals, unipolar modulation schemes are preferred. Conventional scheme for such modulation is direct current offset optical (DCO) OFDM, but it gives away the output of the bipolar signal as the positive part plus the DC components (the DC component is the rudiment of the negative part). This additional DC component adds to the efficiency degradation in power since the DC component renders no information as depicted in [4].

OFDM channel when clipped asymmetrically reduces the needed optical power in the case of specific data rates. The reason behind this is that, during DC biasing a bipolar OFDM, no power is being discarded. Thus, power efficient schemes such as asymmetrically clipped optical (ACO) OFDM and pulse-amplitude-modulated discrete multitone (PAM-DMT) that make use of the advantage of Fourier transform and clipping away of the negative signal component to obtain positive or unipolar time domain signals with minimum wastage of information have emerged. ACO-OFDM as well as the PAM-DMT when used alone causes spectral inefficiency because they utilize just half of the total number of subcarriers. For the improvement of frequency spectrum of operation (i.e., to get bandwidth efficiency), we can concatenate the ACO-OFDM signal subcarriers and the PAM-DMT signal subcarriers. This concatenation results in the generation of the hybrid asymmetrically clipped optical (HACO) OFDM signal as in [11] (Fig. 4).

Odd indices of the HACO-OFDM are occupied by the subcarriers of ACO-OFDM which are in a unipolar manner (i.e., positive signals are only being transmitted by cutting the even subcarriers at zero level) at the cost of spectrum usage, and the even indices of the HACO-OFDM are occupied by the subcarriers of PAM-DMT.

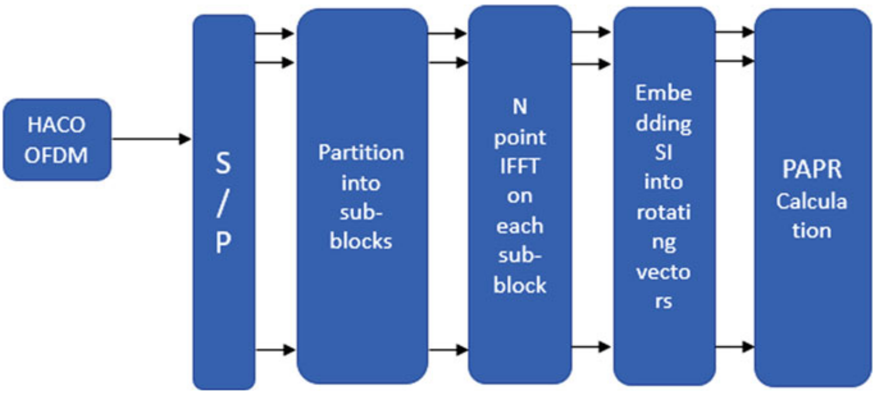


Fig. 4 PTS schemes in HACO-OFDM

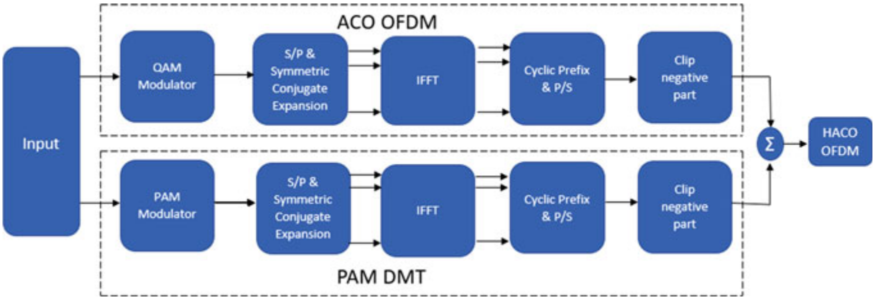


Fig. 5 Block diagram of a HACO-OFDM transmitter

ACO-OFDM has lowest cost at spectrum efficiency that is less than 2 bits/Hz, while PAM-DMT has the highest cost, and the cost increases with the spectrum usage. Figure 5 shows the block diagram of a HACO-OFDM transmitter. The upper half depicts the ACO-OFDM generator and the lower portion depicts the generation of a PAM-DMT signal according to [11].

The generated HACO-OFDM signal is converted into parallel data blocks. The parallel data stream is being partitioned into subblocks of data. N point IFFT operation is being done on the partitioned subblocks. The side information is embedded onto rotating vectors chosen according to the phase offset vectors; this is depicted in Fig. 4 and is discussed in detail in [11].

5 PAPR Reduction Without Side Information Using PTS Scheme

The use of low-multifaceted nature of the chosen rotating vector at the recipient is performed utilizing the PTS scheme without SI. The SI is being inserted into various identifiable rotating vectors into the elective signal constellations. For giving noticeable phase offsets onto the components of each rotating vector, we utilize the PTS scheme without side information. In particular, by taking into account the quantity of sub-streams of the input sequence of PTS scheme, the number of phase offsets must be legitimately picked. The values that are being used for the rotation of phase are moved using the above-mentioned phase offset values. An ML detector is thus being utilized to remove SI present in the receiver output signal, and the sequence of information is being recovered. The Euclidean separation between the constellation of signal that is given as input and the constellation of signal adjusted or tuned using phase factors is being used by the ML detector. It is investigated how to pick rotating vectors for embedding SI, by doing pairwise error probability (PEP) examination. Likewise, in light of PEP, the degradation in performance is brought about by malice in SI detection as in [12].

The pilot tones are not changed using the PTS schemes without side information. That is, by the PTS scheme without side information, both efficiency in the utilization of the bandwidth and exact identification of the channel parameters are accomplished. Likewise, the extension of signal constellation is not finished by the proposed schemes. Instead, on choosing a modest amount of phase factors, the depicted schemes rotate the signal constellation. In this way, to examine a smaller search space, the depicted schemes show low efficiency in computations. The phase offsets and the subblock phase offset vectors are portrayed which are being utilized so that the SI can be encapsulated into rotating vectors. The degradation in performance brought about by failure in the detection of SI is investigated, and an ML identifier for PTS scheme without side information is additionally being clarified.

In an optical OFDM system, where N is the number of subcarriers from a signal constellation Q of size q, a symbol subsequence that is being fed as the input $X = [X_0\ X_1 \cdots X_{N-1}]$ is comprised of N complex symbols. By applying reverse Fourier transform (IFFT) to the input X as $x = \text{IFFT}(X)$, a discrete OFDM signal sequence $x = [x_0\ x_1 \cdots x_{N-1}]$ is created, that is,

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{j2\pi \frac{nk}{N}} \tag{2}$$

An information symbol sequence arrangement X is being subdivided as V disjoint symbol subsequences $X_v = [X_{v,0}\ X_{v,1} \cdots X_{v,N-1}]$, $0 \le v \le V - 1$ in the referenced PTS scheme. All symbols are zero with the exception of N/V information symbols, for every symbol subsequence X_v. By imposing IFFT to every symbol subsequence X_v, the signal subsequence $x_v = [x_{v,0}\ x_{v,1} \cdots x_{v,N-1}]$, called a subblock, is being produced. By a rotating constant b_v^u with size one, chosen from a variable B, each subblock x_v is duplicated which is essentially $\{\pm 1\}$ or $\{\pm 1, \pm j\}$. Additionally, the size or number of elements of B is signified as $|B| = W$ which is called its cardinality. The uth alternate signal sequence $x^u = [x_0^u x_1^u \cdots x_{N-1}^u]$ is acquired on summing these V phase rotated subblocks as,

$$X^u = \sum_{v=0}^{V-1} b_v^u X_v \tag{3}$$

Let the uth rotating vector be $b^u = \left[b_0^u b_1^u \cdots b_{V-1}^u\right]$, where $b^u \in B^V$ and $b_0^u = 1$. The optimum rotating vector used for transmission is $b^{\tilde{u}} = \left[b_0^{\tilde{u}} b_1^{\tilde{u}} \cdots b_{V-1}^{\tilde{u}}\right]$. This is being done in order to choose the minimum PAPR signal sequence.

5.1 Encapsulation of Side Information

To encapsulate the SI into the rotating factors, and to dispense with the transmission of SI, shift their phase with proper offset values. For this reason, initially, phase offset which is a V-tuple vector is being characterized as, $S^u = \left[S_0^u S_1^u \cdots S_{V-1}^u\right]$ where $S_v^u \in \{0, 1, \ldots, Z\}$, $0 \leq v \leq V - 1$ and $0 \leq u \leq U - 1$ where, Z is the identifiable nonzero phase offset values and $S_v^u = 0$ depicts a zero-phase offset as depicted in [12]. The product of each rotating factor b_v^u, $0 \leq v \leq V - 1$, by $e^{j\theta_{S_v^u}}$, is to encapsulate the SI into the uth rotating vector b^u using u which is a variable, where the phase offset value is $\theta_{S_v^u}$ and is depicted by S_v^u as $0 \leq \theta_{S_v^u} < 2\pi$. Then, rotating vector which is being uth modified is depicted as b^u,

$$b_u = \left[b_0^u b_1^u \cdots b_{V-1}^u\right] \tag{4}$$

$$\text{i.e.,} \quad b^u = [b_0^u e^{j\theta_{S_0^u}} b_1^u e^{j\theta_{S_1^u}} \cdots b_{V-1}^u e^{j\theta_{S_{V-1}^u}}]$$

and the signal sequence which is uth modified is x^u is being depicted in the PTS scheme without side information, and it is expressed as,

$$X^u = \sum_{v=0}^{V-1} \bar{b}_v^u X_v \tag{5}$$

To the obtained optical OFDM signal, FFT is being applied and the phase modified symbol for kth input symbol X_k of vth symbol sequence X_v, $R_k = b_u X_k$. On modifying the signal constellation Q by $\theta_{\bar{S}_v^u}$, where $0 \leq \theta_{\bar{S}_v^u} < 2\pi$ and $\bar{S}_v^u$ is the vth component of uth phase offset vector picked for sending, we get $b_u X_k$ which is a symbol in the signal constellation Q. The distinguished symbol is b_v^u. R_k does not return on the signal constellation Q when $S_v^u = \bar{S}_v^u$, where b_v^u* is the complex conjugate of b_v^u; this is criteria for the choice of the nonzero phase offsets. The receiver could identify the correct SI utilizing this method seen in [12]. Specifically, the receiver could identify the SI by watching the degenerate Euclidean separation of identified symbol which is a part of the constellation signal denoted by Q. U the number of phase modifying vectors, which is being depicted by using U phase tuning vectors S^u. Clearly, Z must be chosen with the end goal that it should appropriately adjust U tuning vectors on utilizing phase tuning vectors which are V-tuple,

$$(Z + 1)^V \geq U = W^{V-1} \tag{6}$$

At this point, by utilizing a detector which is called the 'maximal likelihood detector', that makes use of the Euclidean separation in between first and signal constellations that are being rotated, the receiver can analyze the SI and recuperate the input information symbol sequence X.

5.2 PTS Scheme II

To minimize complexity in detection of the signal in the receiver side, there should be a reduction in the number of nonzero phase shifts. Using $Z = 1$ for $B = \{\pm 1\}$, minimization in the number of phase shifts for PTS scheme I occurs which is being inferred as $Z = 1$, and this improves detection of signals at the reconstruction part. But, for $B = \{\pm 1, \pm j\}$, the least nonzero phase shifts for the PTS scheme I must stick onto a value larger than one. For low detection complexity, PTS scheme II is introduced, that uses $Z = 1$ for $B = \{\pm 1, \pm j\}$. Here, use the binary phase shifts $\theta_0 = 0$ and $\theta_1 = \pi/4$. If the value of $Z = 1$, the Euclidean separation of the constellations of signals, namely Q and $Q\pi/4$, is being enhanced for the QAM modulations basically 4-QAM, 16-QAM and 64-QAM. PTS scheme II also divides the subblocks to assign various phase shifts staking into account the linearity of the output of the FFT block. The vth subblock x_v is partitioned into an even sequence $x_v^e = \left[x_{v,0}^e x_{v,1}^e \cdots x_{v,N-1}^e\right]$ and an odd subblock $x_v^0 = \left[x_{v,0}^0 x_{v,1}^0 \cdots x_{v,N-1}^0\right]$ by employing IFFT to every even and odd symbols in the frequency domain equipped with N/2 zeros in an interleaved manner as in [12].

5.3 Maximal Likelihood Detector in the PTS Scheme Without Side Information

The receiver discovers the index $\bar{u}$ irrespective of utilizing SI and recuperates the symbol sequence X which is being fed as the input, from the received signal created by the adjusted rotating vector $\bar{b}^u$. The received symbol R_k which is a part of the subcarrier k, which has a place with the vth symbol subsequence X_v, in frequency domain is depicted as, $R_k = H_k\ R_k = H_k \bar{b}_v^u X_k + N_k$ where H_k is the output response in frequency and N_k an AWGN sample at the subcarrier k with the variations per measurement $N0/2$. An assumption in the immaculate channel state information (CSI) is being done, that is, quasi-static Rayleigh fading channel is being chosen, and the receiver is made known of all the H_k's statistically autonomous and impeccably. Hence, the ML finder for the PTS scheme without side information, utilizes the Euclidean partition present in between rotated signal constellations and the initial signal constellation. Let the vth received symbol subsequence $R_v = [R_{v,0}\ R_{v,1} \cdots R_{v,N-1}]$, $0 \le v \le V-1$. The maximal likelihood detector of the P-PTS scheme I works in two stages. To begin with, the fractional metric for each received symbol subsequence R_v comparing to X_v is resolved as,

$$D_{v,S_{uv}} = \sum_{k \in \mathbb{I}_v} \min_{X\prime k \in Q} \left| R_{v,k} e^{-j\theta_{S_v^u}} - \hat{H}_k X_k' \right|^2 \tag{7}$$

where $|\cdot|$ means the extent of a complex number or its magnitude, H_k is the evaluated channel response and $\mathbb{I}_v = \left\{\mathbb{I}_{v,0}, \mathbb{I}_{v,1}, \ldots, \mathbb{I}_{v,N-1}\right\} \subset \{0, 1, \ldots, N-1\}$ is defined as the collection of N/V information images encapsulated in the subsequence of the vth symbol, X_v. $\mathbb{I}_v$ is resolved utilizing the subblock partitioning scheme. We get $D_{v,S^u_{v,0}}, D_{v,S^u_{v,1}}, \ldots, D_{v,S^u_{v,N-1}}$, for every ith symbol subsequence. Additionally, a vector $\bar{X}_{S^u_{vv}} = \left[\bar{X}_{S^u_{vv,0}} \bar{X}_{S^u_{vv,1}} \ldots \bar{X}_{S^u_{vv,N-1}}\right]$, $0 \le u \le U-1$, which minimizes the Euclidean separation between $R_{v,k}e^{-j\theta_{S^u_v}}$ and $\hat{H}_k X'_k$ denotes the corresponding X_k for every S^u_v, where $\bar{X}_{S^u_{vv}}$ is the point in Q which is the constellation for X_k. Clearly, $\bar{X}_{S^u_{vv}}$ consists of $N(V-1)/V$ zeros for the expelling index positions and N/V recognized information symbols with indices in $\mathbb{I}_v$. This process is rehashed for all v, $0 \le v \le V-1$. Second, calculate $D_u = \sum_{v=0}^{V-1} D_{v,S^u_v}$, $0 \le u \le U-1$ by utilizing the partial matrices D_{v,S^u_v}'s, and the index $\bar{u}$ is analyzed by seeking D^u with the minimum value among U metric values D_u as,

$$u = \arg \min_{0 \le u \le U-1} D_u \tag{8}$$

Therefore, the obtained sequence of the symbol that is being inputted at the receiver is finally depicted as,

$$\hat{X}^{\hat{u}} = \sum_{v=0}^{V-1} b^{\hat{u}}_v * \hat{X}^{S^u_v} \tag{9}$$

which is the sum of V detected subsequences of the symbols multiplied $b^{\hat{u}}_v *$, $0 \le v \le V-1$ to de-rotate it. Figure 6 shows the schematic blueprint of the utilized detector which maybe maximal likelihood for the PTS scheme that embeds the side information. Z must be less than U, if the minimum Z is chosen. Therefore, the general detection sophistication of the PTS I to discover $\hat{u}$ and the recognized symbol sequence $\hat{X}_u$ is $(Z+1)qN$ $|\bullet|^2$ tasks if the real additions are disregarded in light of the fact that the complexity of $|\bullet|^2$ activity is a lot bigger than that of real addition as in [12].

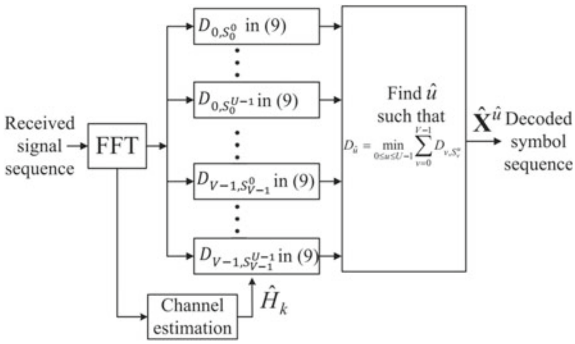


Fig. 6 Block diagram of an ML estimator

6 Simulation Results

Simulation outcomes are used to compare the functioning of both PTS schemes in ACO-OFDM signals, PAM-DMT signals and HACO-OFDM signals. The simulation results depicted using the compared methods show that the reduction of PAPR using the PTS scheme II without side information is more efficient in all the compared signals. The theoretical PAPR of HACO-OFDM signals in an LTE system is approximately 12 dB. The reduction of PAPR using the PTS technique 1 is 10 dB; when the PTS technique 2 is being used, the PAPR is further being reduced to just 5.5 dB as depicted in Fig. 12. Even though the PAPR reduction of PAM-DMT signals using PTS scheme II is less compared to HACO-OFDM signals, HACO-OFDM signals are preferred in wireless communication because of its bandwidth efficiency. The graphs are plotted between the PAPR values at each original subsequence obtained numerically and the probability of PAPR values of each modified subsequence with respect to the threshold value.

Figure 7 depicts the generated ACO-OFDM signal sequence, Fig. 8 represents the generated PAM-DMT signal sequence and Fig. 9 depicts the generated HACO-OFDM signal sequence.

Based on the simulation processes that are being carried out in Fig. 10, the results depicted in Fig. 10 show that the PAPR of the generated ACO-OFDM signal is being reduced using the new PTS scheme I, and it is reduced further using PTS scheme II.

Based on the simulation processes that are being carried out, the results depicted in Fig. 11 show that the PAPR of the generated PAM-DMT signal is being reduced using the new PTS scheme I, and it is reduced further using PTS scheme II.

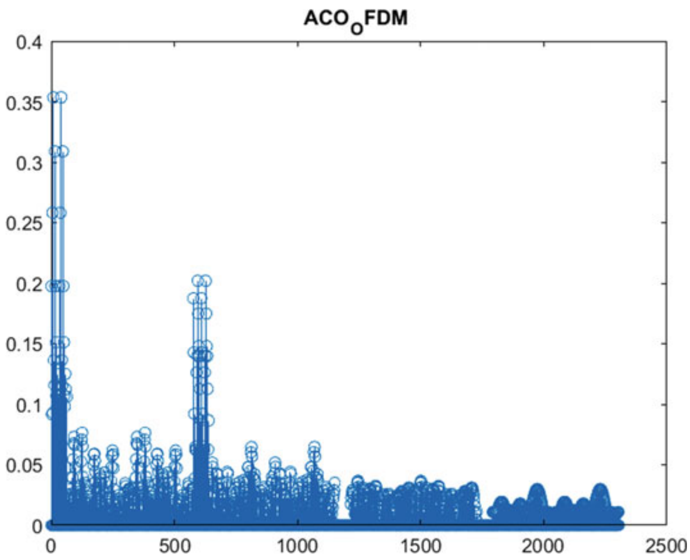


Fig. 7 Generated ACO-OFDM sequence

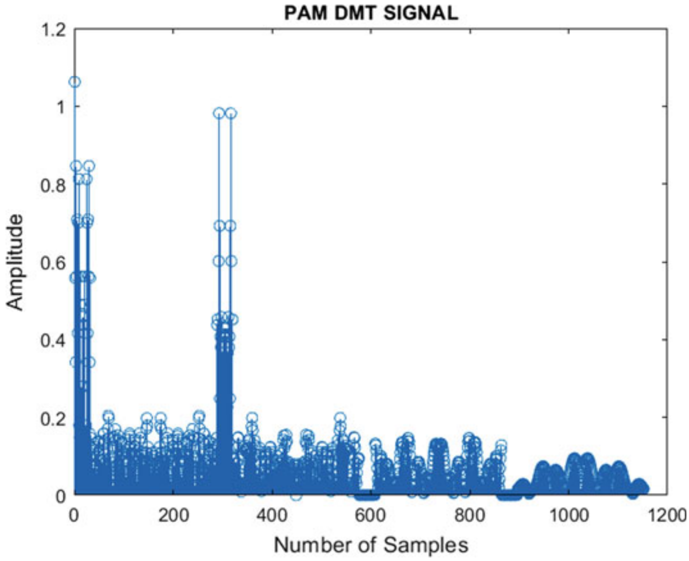


Fig. 8 Generated PAM-DMT sequence

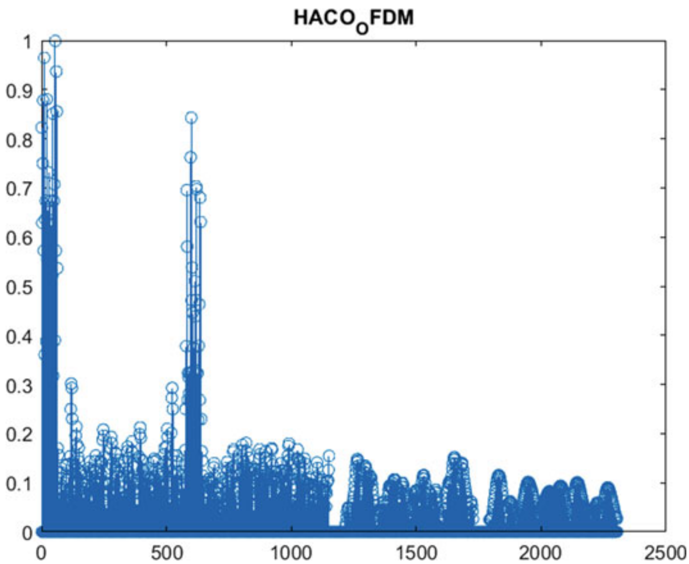


Fig. 9 Generated HACO-OFDM sequence

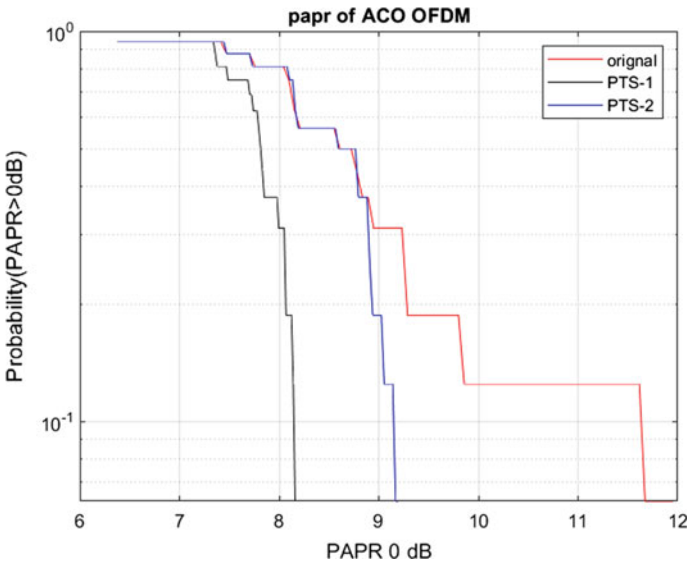


Fig. 10 Comparison of PAPRs of ACO-OFDM sequence using PTS schemes I and II

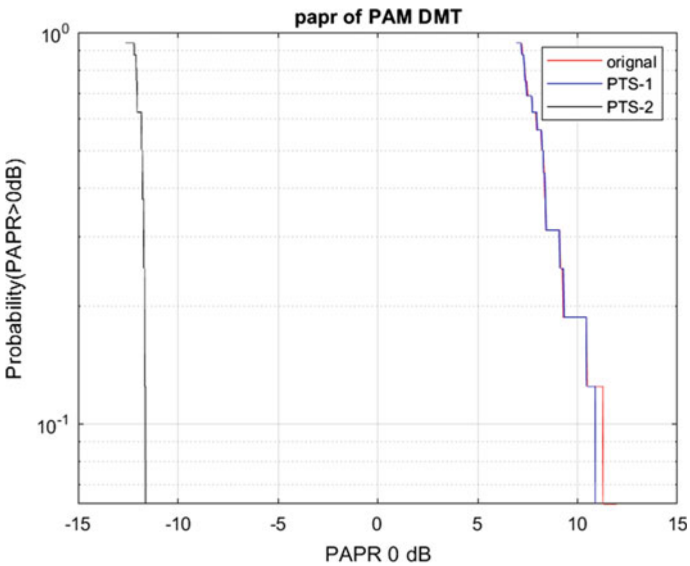


Fig. 11 Comparison of PAPRs of PAM-DMT sequence using PTS schemes I and II

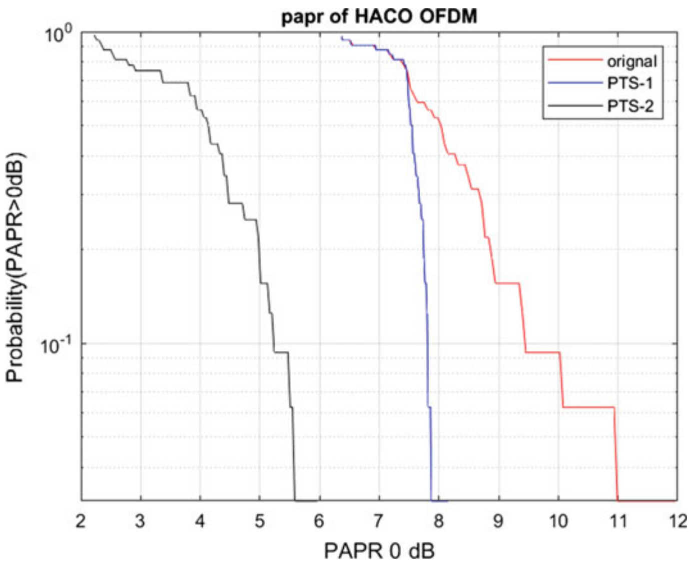


Fig. 12 Comparison of PAPRs of HACO-OFDM sequence using PTS schemes I and II

Based on the simulation processes that are being carried out, the results depicted in Fig. 12 show that the PAPR of the generated HACO-OFDM signal is being reduced using the new PTS scheme I, and it is reduced further using PTS scheme II.

7 Conclusion

In this paper, two PTS plans without SI (schemes which do not transmit the side information), distinguishing a rotating vector on the grounds that the phase offset is identifiable and can be imposed onto the components of each rotating vector are used for lessening the PAPR of ACO-OFDM signals, PAM-DMT signals and HACO-OFDM signals and enhancing the throughput. The maximal likelihood detection for the depicted PTS schemes is obtained, to obtain side information of the rotating factor and recoup the information sequence at the output side. The PAPR execution of two depicted PTS schemes is irrelevantly disintegrated, and the traditional PTS with impeccable SI is unmistakably being portrayed utilizing the simulation results.

References

1. Jovicic A, Li J, Richardson T (2013) Visible light communication: opportunities, challenges and the path to market. IEEE Commun Mag 51(12):26–32
2. Khalid AM, Cossu G, Corsini R, Choudhury P, Ciaramella E (2012) 1-Gb/s transmission over a phosphorescent white LED by using rate adaptive discrete multitone modulation. IEEE Photon J 4(5):1465–1473
3. Nassim Aarab M, Otman C (2020) Chapter 19: MIMO-OFDM for wireless systems: an overview. Springer, Berlin
4. Li B, Xu W, Zhang H, Zhao C, Hanzo L (2017) PAPR reduction for hybrid ACO-OFDM aided IM/DD optical wireless vehicular communications. IEEE Trans Veh Technol
5. Zhou Y, Jiang T (2009) A novel multi-points square mapping combined with PTS to reduce PAPR of OFDM signals without side information. IEEE Trans Broadcast 55(4):831–835
6. Kim KH (2016) On the shift value set of cyclic shifted sequences for PAPR reduction in OFDM systems. IEEE Trans Broadcast 62(2):496–500
7. Guan L, Jiang T, Qu D, Zhou Y (2010) Joint channel estimation and PTS to reduce peak-to-average-power radio in OFDM systems without side information. IEEE Signal Process Lett 17(10):883–886
8. Kim H, Hong E, Ahn C, Har D (2011) A pilot symbol pattern enabling data recovery without side information in PTS-based OFDM systems. IEEE Trans Broadcast 57(2):307–312
9. Ochiai H, Imai H (2000) Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited OFDM systems. IEEE J Sel Areas Commun 18(11):2270–2277
10. Goel A, Gupta P, Agrawal M (2013) Generalized M-2M mapping scheme for SLM and PTS based OFDM systems without side information. Wirel Pers Commun
11. Ranjha B, Kavehrad M (2014) Hybrid asymmetrically clipped OFDM based IM/DD optical wireless system. J Opt Commun Netw 6(4):387–396
12. Joo H-S, Kim K-H, No J-S, Shin D-J (2017) New PTS schemes for PAPR reduction of OFDM signals without side information. IEEE Trans Broadcast

Effectiveness of Wilson Amplitude for the Detection of Murmur from the PCG Records

P. Careena, M. Mary Synthuja Jain Preetha, and P. Arun

Abstract Methods employed for the automatic detection of valve ailments from the heart records is a skilled task in cardiology. However, automated approaches, proposed for envisaging the cardiovascular diseases greatly depend on the features mined from the heart sound. The analysis of Phonocardiogram (PCG) signals, offers adequate information about the functioning of the heart. Feature extraction techniques in the time domain have analytical simplicity and less computational complexity. In this paper, the effectiveness of the feature namely Wilson amplitude for the detection of murmur from heart signal is investigated and is tested with different threshold (5–50 mV) levels. It is found that the Wilson amplitude at 20 mV threshold is capable to detect the murmur from PCG signal with 88.33% accuracy, 76.67% sensitivity and 100% specificity than other thresholds.

Keywords Heart abnormality · Murmur · PCG signal · Statistical significance · Wilson amplitude · Time domain feature

1 Introduction

The feature extraction is the primary step involved in any artificial intelligence system. In the system meant for automated analysis, the signal processing techniques engaged for extracting the features shows the main role. This mainly involves

P. Careena (✉)
Department of Electronics and Communication Engineering, Amal Jyothi College of Engineering, Kanjirapally 686518, India
e-mail: careenaarun@gmail.com

M. Mary Synthuja Jain Preetha
Department of Electronics and Communication Engineering, Noorul Islam University, Nagercoil 629180, India

P. Arun
Department of Electronics and Communication Engineering, St. Joseph's College of Engineering and Technology, Palai 686579, India

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_15

fault diagnosis of mechanical/electrical systems using their vibration data, detection of diseases by analyzing various biological signals, etc. The problems can be detected, accurately traced and even their type can be recognized by using the features, extracted via suitable signal processing procedures. Moreover, the features selected via the appropriate feature extraction technique should be statistically significant and computationally efficient. The features extracted will be of temporal, spectral-domain or spectrotemporal domain. Out of these, temporal features are simpler and analytically viable than features in other domains because they do not involve complex domain transformation.

Coronary artery disease (CAD) is one of the reasons for mortality and morbidity worldwide. Based on the report of the World Health Organization (WHO), in 2016 almost 17.9 million humans expired due to CVDs, representing 31% of all global deaths. Out of this, 85% are due to heart attack and stroke [1]. The biological records like Electrocardiogram (ECG), Electroatriogram (EAG), Electroventriculogram (EVG), intracardiac electrogram (EGM), Electromyogram (EMG), echocardiogram, photoplethysmogram (PPG) and Phonocardiogram (PCG) can be used for analyzing the heart functionality. Of this, the methodologies based on the features extracted from PCG records may be used to detect murmur from heart sound [2].

A few methods that incorporate Wilson amplitude to identify problems of numerous sectors are presented in the literature. Naji et al. [3] proposed a method to identify best feature spaces that separate among postural tasks involving different trunk flexion by analyzing the preprocessed EMG signal. Nayana and Geethanjali [4] estimated the Wilson amplitude of the bearing vibration as one of the features for roller element bearing fault diagnosis. Prior to this computation, the vibration data were preprocessed. The feature was given into the SVM classifier. For hand gesture recognition, Zhang et al. [5] collected sEMG (surface electromyography) signals from six forearm hand muscles during grasping or pinch tasks. These signals were segmented via empirical mode decomposition. The muscle activity was feature extracted using zero crossings and Wilson amplitude of the first four resulting intrinsic mode functions. Then a feature vector was framed by using these features and is applied to a classifier combined with support vector machine and genetic algorithm. Xi et al. [6] proposed a method for activity monitoring and fall detection using sEMG. The feature extraction using the above feature has been done after preprocessing. In the analysis of surface electromyography (sEMG) signals, Angkoon et al. [7] selected Wilson amplitude at a threshold of 5 mV as the best feature that tolerates with white Gaussian noise. Waris and Kamavuako [8] estimated the Wilson amplitude of the EMG signal to observe the outcome of threshold selection on the classification for prosthetic use in medical applications. Samuel et al. [9] put forth a system for the arm movement classification using pattern recognition. EMG signals were recorded from the residual arms of eight amputees while performing different upper limb movements. To rate the efficiency, the features like the absolute value of the summation of square root, mean value of the square root and the absolute value of the summation of the preprocessed EMG signal have been compared with that of the Wilson amplitude.

Bi et al. [10] measured Wilson amplitude of the preprocessed EMG signal for motor intention prediction of continuous human upper limb motion for the human–robot collaboration.

In this paper, the effectiveness of employing the time domain feature known as Wilson amplitude for the detection of murmur from heart sound by examining their statistical significance and the separability offered among them is investigated on the PCG records collected from the Pascal heart sound challenge database [11]. The highlights of this study are (i) The statistical significance and the separability of Wilson amplitude is tested and the separability given by this to discriminate normal/murmur is evaluated; (ii) It is comparatively simple as the proposed work uses only a fundamental temporal feature.

The investigation, mathematical design and the particulars of the dataset tested are given in Sect. 2. In Sect. 3, the statistical significance and separability given by the features to differentiate normal and the murmur is investigated.

2 Methodology

The schematic of the procedures contains in the calculation of Wilson amplitude is shown in Fig. 1.

The PCG records are preprocessed before the computation of Wilson amplitude. In the preprocessing, to attain proper sampling rate, the signal is upsampled by a factor 2, the amplitude normalization is done between −1 and +1 and to eliminate the frequency components below 10 Hz by incorporating a high pass filter. As stated earlier, the Wilson amplitude of the preprocessed PCG signal is measured to examine its ability to identify the presence of murmur in the PCG signal. The feature evaluation is done after extracting this feature.

The normalized upsampled PCG signal is given as

$$X_n(t) = \frac{X_i(t)}{\max|X_i(t)|} \tag{1}$$

where '$X_i(t)$' is the PCG signal (sampling rate '1/fs' and 'N' samples) processed between $1 \leq n \leq N$. The PCG records are normalized (−1 and +1) to regulate the amplitude of the signal as in (1).

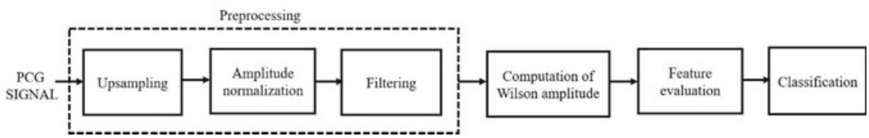


Fig. 1 Schematic of the experimental design

The feature Wilson amplitude is the number of times that the difference among two successive amplitudes go above a certain threshold [6]. It can be formulated as,

$$\text{Wilson amplitude} = \sum_{t=1}^{N-1} f\left(\left|x_{n(t-1)} - x_{n(t)}\right|\right) \tag{2}$$

$$f(x) = \begin{cases} 1, & \text{if}\, x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

In this work, the effectiveness of Wilson amplitude at threshold of 50, 40, 30, 20, 10 and 5 mV is examined on the preprocessed PCG signal. The method is tested on a total of 60 records (30 normal and 60 murmurs) collected from the Pascal heart sounds challenge database. Each selected PCG records have a duration of more than 8 s and are upsampled by 2 i.e.; the sampling frequency 'fs' of the dataset is 4 kHz. A high pass filter with cut-off frequency 10 Hz is utilized to eliminate the frequency components less than 10 Hz.

The wave pattern of PCG records analogous to normal and murmur data set are shown in Fig. 2a–d. The wave pattern of both types of signal is completely different

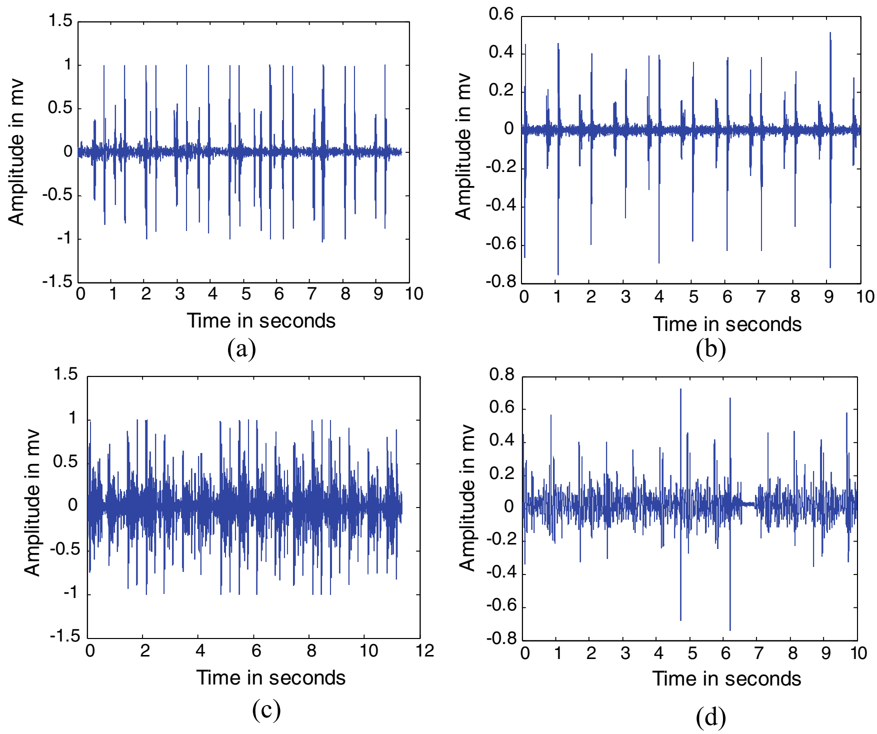


Fig. 2 Wave pattern of two classes of PCG records **a** and **b** normal heart sound, **c** and **d** murmur

from each other. They have different amplitudes and randomness parameters.

The Kolmogorov–Smirnov test is used as the statistical evaluation test to analyze the efficacy of Wilson amplitude to discriminate normal and murmur. The separability offered by the feature is examined by the histogram. Matlab® is the software employed for feature extraction and statistical evaluation.

3 Results

In this paper, the effectiveness of employing Wilson amplitude for the detection of murmur from heart sound is investigated on the Phonocardiogram signal availed from the Pascal heart sound challenge database. The wave shape of the preprocessed signal is given in Fig. 3a–d. The wave pattern of both classes of heart signal is entirely different especially in terms of their amplitude and randomness.

Table 1 shows the range and numerical values of waveform length compliant to normal and murmur.

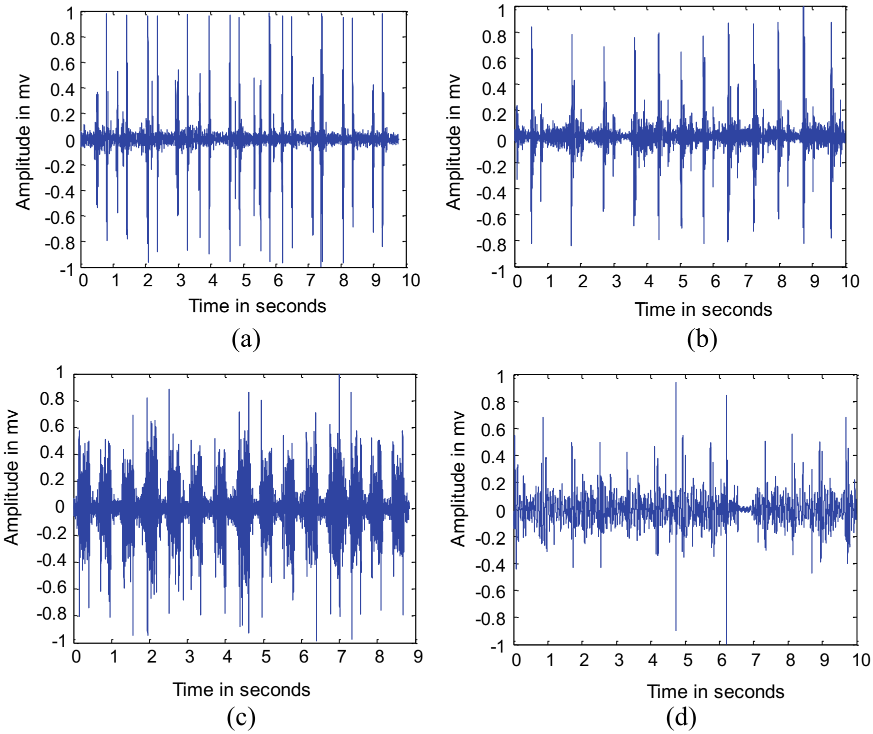


Fig. 3 Wave shape of preprocessed PCG records **a** and **b** normal heart sound, **c** and **d** murmur

Table 1 Range and numerical values of the feature conforming to normal and murmur

		Threshold					
		5 mV	10 mV	20 mV	30 mV	40 mV	50 mV
Normal	Range	5347–24,663	824–11,643	92–4725	9–2622	0–1618	0–989
	Mean ± SD	14,754.1 ± 5335.88	5492.83 ± 2806.01	2030.33 ± 1358.53	986.43 ± 728.18	552.2 ± 448.05	336.13 ± 302.67
Murmur	Range	7136–93,934	3768–75,881	291–61,360	80–52,160	30–44,799	8–38,375
	Mean ± SD	38,780.5 ± 20,448.74	24,149.73 ± 16,850.01	13,454.13 ± 13,491.86	8688.07 ± 11,292.39	6050.83 ± 9705.83	4437.2 ± 8387.24

Table 2 Kolmogorov–Smirnov Test—'H' and 'P' values of Wilson amplitude of PCG records analogous to normal and murmur

H and P values	Threshold					
	5 mV	10 mV	20 mV	30 mV	40 mV	50 mV
Chi square value (H)	1	1	1	1	1	1
Probability value (P)	2.62×10^{-7}	5.59×10^{-8}	6.11×10^{-8}	1.14×10^{-6}	4.64×10^{-6}	1.76×10^{-5}

It is observed (from Table 1) that, the range of normal signal is 5347–24,663, 824–11,643, 92–4725, 9–2622, 0–1618 and 0–989 for a threshold of 5 mV, 10 mV, 20 mV, 30 mV, 40 mV, and 50 mV, respectively. The range of murmur is 7136–93,934, 3768–75,881, 291–61,360, 80–52,160, 30–44,799 and 8–38,375, for a threshold of 5 mV, 10 mV, 20 mV, 30 mV, 40 mV, and 50 mV, respectively. For the same threshold levels, the numerical values of normal are 14,754.1 ± 5335.88, 5492.83 ± 2806.01, 2030.33 ± 1358.53, 986.43 ± 728.18, 552.2 ± 448.05 and 336.13 ± 302.67, respectively. The numerical values of murmur are 38,780.5 ± 20,448.74, 24,149.73 ± 16,850.01, 13,454.13 ± 13,491.86, 8688.07 ± 11,292.39, 6050.83 ± 9705.83 and 4437.2 ± 8387.24 for thresholds of 5 mV, 10 mV, 20 mV, 30 mV, 40 mV and 50 mV, respectively.

By examining the range of murmur it is clear that they are derestricted to a wide range related to that of normal heart sound for all the said threshold levels. That means, the numerical value of the feature Wilson amplitude matches to normal PCG records limited to a narrow range than that of murmur. These outward differences among the magnitude and the range of feature extracted from the PCG records give the prospect of Wilson amplitude to discriminate normal and as murmur.

The Kolmogorov–Smirnov test is used as the statistical evaluation test to measure the separability of Wilson amplitude to discriminate normal and murmur. The 'H' and 'P' values of these features for different thresholds are provided in Table 2.

The H values obtained from the K-S Test is '1' for all thresholds. The 'H' values are 2.62×10^{-7}, 5.59×10^{-8}, 6.11×10^{-8}, 1.14×10^{-6}, 4.64×10^{-6} and 1.76×10^{-5} for thresholds of 5 mV, 10 mV, 20 mV, 30 mV, 40 mV and 50 mV, respectively.

As already mentioned, the separability offered by the feature for the qualitative assessment is done via histogram. The histogram of Wilson amplitude conforming to normal heart sound and murmur for thresholds 5 mV, 10 mV, 20 mV, 30 mV, 40 mV and 50 mV are shown in Fig. 4a–f, respectively.

In the histogram shown in Fig. 4, the histograms of all the thresholds exhibits overlay among that of both classes of heart signal in specific areas. The histogram analogous to normal heart signal is lying adequately apart spatially from that of murmur with less separability. It has also found that the overlap of the two classes of heart sounds is minor at the thresholds 20–30 mV than all other threshold levels.

The performance parameters like accuracy, sensitivity and specificity of Wilson amplitude to classify murmur/normal are also calculated for various threshold and are given in Table 3.

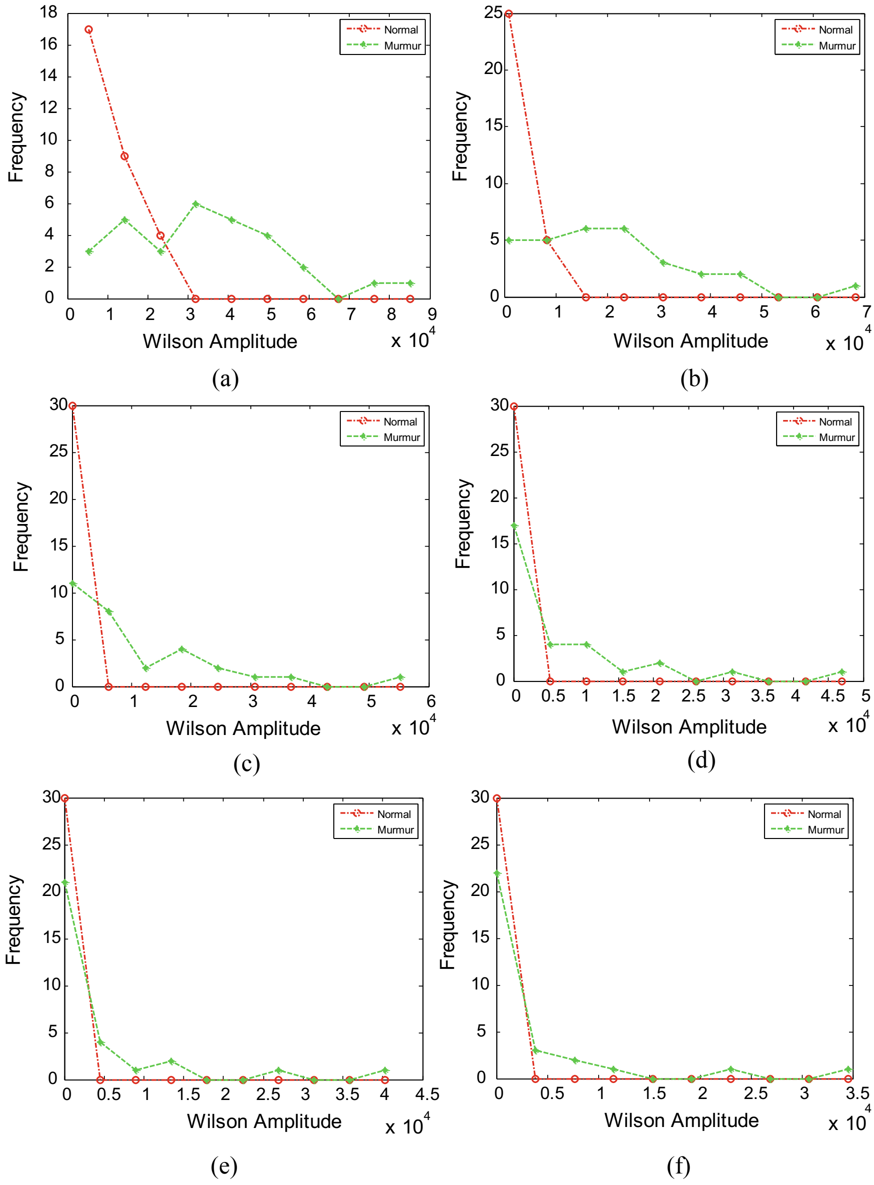


Fig. 4 The histogram of Wilson amplitude of normal and murmur

Table 3 Performance parameters of Wilson amplitude of heart signal equivalent to normal and murmur at different thresholds

Parameter (in %)	Threshold					
	5 mV	10 mV	20 mV	30 mV	40 mV	50 mV
Accuracy	81.67	86.67	**88.33**	80	78.33	78.33
Sensitivity	83.33	80	**76.67**	80	70	76.67
Specificity	80	93.33	**100**	80	86.66	80

It has observed from Table 3 that, to distinguish murmur and normal from the PCG records the Wilson amplitude offers better performance at a threshold of 20 mV. The same is represented bold italics in the fourth column of the above table. The Wilson amplitude at 20 mV threshold is given accuracy of 88.33%, sensitivity of 76.67% and a specificity of 100%.

4 Conclusions

In this paper, the effectiveness of Wilson amplitude at different thresholds like 5, 10, 20, 30, 40 and 50 mV for the detection of murmur from heart signal was investigated on the preprocessed PCG records acquired from Pascal heart sound database. The feature was statistically assessed via Kolmogorov–Smirnov test and the separability among the feature to distinguish murmur and normal was examined by means of histogram. It has found that Wilson amplitude at 20 mV threshold relates to normal and murmur exhibited a '*P*' value of 6.11×10^{-8}. The same has offer better accuracy (88.33%), sensitivity (76.67%) and specificity (100%) than all other thresholds. The technique comprising the Wilson amplitude as a feature may resolve the complications related with the physical auscultation. The possibility of this feature to classify the murmur phenotypes can also be considered as a future extension.

References

1. WH Organization (2017) Cardiovascular diseases. www.who.int/mediacentre/factsheets/fs317/en/. Last accessed 13 Aug 2019
2. Samanta P, Pathak A, Mandana K, Saha G (2019) Classification of coronary artery diseased and normal subjects using multi-channel phonocardiogram signal. Biocybern Biomed Eng 36:426–443
3. Naji M, Firoozabadi M, Kahrizi S (2012) Evaluation of EMG features of trunk muscles during flexed postures. In: 2012 19th Iranian conference of biomedical engineering (ICBME). Tehran, pp 71–74
4. Nayana BR, Geethanjali P (2017) Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal. IEEE Sens J 17:5618–5625

5. Zhang Y, Liao Y, Wu X, Chen L, Xiong Q, Gao Z, Zheng X, Li G, Hou W (2018) Non-uniform sample assignment in training set improving recognition of hand gestures dominated with similar muscle activities. Front Neurorobotics 12:1–12
6. Xi X, Tang M, Miran SM, Luo Z (2017) Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors. Sensors 17:1–20
7. Angkoon P, Chusak L, Phukpattaranont P (2008) EMG feature extraction for tolerance of white gaussian noise. In: International workshop and symposium on science and technology, pp 178–183
8. Waris A, Kamavuako EN (2018) Effect of threshold values on the combination of EMG time domain features: Surface versus intramuscular EMG. Biomed Signal Process Control 45:263–273
9. Samuel OW et al (2018) Pattern recognition of electromyography signals based on novel time domain features for amputees' limb motion classification. Comput Electr Eng 67:646–655
10. Bi L, Genetu Feleke A, Guan C (2019) A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration. Biomed Signal Process Control 51:113–127
11. Bentley P, Nordehn G, Coimbra M, Mannor S, Rita G (2011) The pascal classifying heart sounds challenge, sponsored by PASCAL

Analysis of Electromagnetic Field Variance in Random PCB Model Using 2D Stochastic FDTD

Jinu Joseph and R. Kiran

Abstract This paper describes the electromagnetic field variance estimation in a random printed circuit board (PCB) model due to the unreliability in permittivity and conductivity of the substrate material using a stochastic finite difference time domain (S-FDTD) method. Traditional FDTD method is updated for mean and variance estimation of the electromagnetic field through the random medium by Taylor series approximation and delta method. The correlation coefficient between the constitutive properties is assumed as to bound the variance field. The 2D S-FDTD method is used for simulating variance of fields in PCB model, and the results are validated using Monte Carlo results.

Keywords Finite difference time domain (FDTD) · Stochastic FDTD (S-FDTD) · Delta method · Variance · Monte Carlo simulation

1 Introduction

The numerical analysis of complex structures is estimated using the finite difference time domain (FDTD) method and is a common technique for modelling computational electrodynamics. As with this electromagnetic simulation, the field components are solved at desired instants of space and time. For random media, permittivity and conductivity vary statistically due to uncertainty in material composition, variation among individuals for biological tissues, temperature and frequency dependence, etc. Electromagnetic field in a model depends upon the electrical properties of the medium under consideration. Traditional FDTD uses mean values of electrical properties and simulates average field values. It finds numerous applications including bioelectromagnetics, atmospheric studies, circuit boards and waveguides.

J. Joseph (✉) · R. Kiran
College of Engineering Trivandrum, Thiruvananthapuram, Kerala, India
e-mail: jinujoseph113@gmail.com

R. Kiran
e-mail: kiranr@cet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_16

If the medium is random, the variation in permittivity and conductivity leads to variation in electromagnetic fields. To overcome this issue, multiple simulation method is considered with random inputs. The Monte Carlo method [1] is used as the golden standard to compute the parameters in statistics. The variation of the electromagnetic field is achieved by combining the Monte Carlo method with the traditional FDTD method, Monte Carlo FDTD (M-FDTD) method. Monte Carlo method is a golden standard in the field of statistical analysis, and it uses thousands of simulation by choosing values of electrical properties randomly in each iteration along with FDTD and field values of individual simulations are collected. The recorded data is post processed to obtain the statistical values such as mean and variance of electromagnetic fields. The method is applicable in calculating the specific absorption rate (SAR) due to cell phone/human body interactions [2]. Even if the method gives an accurate result, it demands extensive computational requirements. As an alternative, the stochastic FDTD (S-FDTD) method can be used for electromagnetic field analysis within a fraction of time compared to the M-FDTD method.

S-FDTD method is very efficient for random medium analysis, which is derived from the traditional FDTD method. The function with random variables is expanded using Taylor's series, and the mean field values are derived by providing suitable approximations to the function. Variance of fields is derived using delta method, which is an extended form of Taylor's series approximation in terms of variance. The derived equations for the electric and magnetic field can be used to directly estimate mean and variation of fields at every point in space and time [3].

In this paper, we proposes a 2D PCB model with random substrate material, and it contains conductive traces. The interaction of plane wave with this structure is analysed in FDTD domain. The known mean and standard deviation of electrical properties of substrate medium are used for estimating the variance of electromagnetic field variance inside the model. The S-FDTD method is used for directly estimating the variance of field components. The results are validated using Monte Carlo FDTD results.

This paper is structured as follows: Sect. 1 is the introduction. Section 2 described the methodology used, Sect. 3 shows the simulation requirements, and Sect. 4 includes results obtained and its discussion. Section 5 deals with the conclusion.

2 Methodology Description

The finite difference time domain method starts with Maxwell's equations. The electric and magnetic vector fields for a given instant of space and time are derived using the set of equations (1).

$$\nabla \times E = -\partial B/\partial t, \nabla \times H = \partial D/\partial t + \sigma E \tag{1}$$

Equation (1) is processed in such a way that the E and H fields are chosen to develop the grid system in FDTD method. The 2D grid system using stochastic finite difference time domain method is used for field variance simulation in the proposed random PCB model.

2.1 2D Space Modelling

S-FDTD is derived from Yee's FDTD method [4]. In this method, the time-dependent Maxwell's equations are discretized using central difference approximations to the space and time partial derivatives. The resulting finite difference equations are then solved, and the electric field vector components in a volume of space are solved at a given instant in time; then the magnetic field vector components in the same spatial volume are solved at the next instant in time; the process is repeated until the desired electromagnetic field characteristics are obtained.

The 2D problem space gets discretized both in time and space, and electric and magnetic fields are get updated at $\Delta t/2$ and Δt time intervals, respectively. It can be achieved by discretization of Maxwell's differential equation by using central difference approximation. The updated equations for TM mode propagation are given as follows:

$$\begin{aligned} H_x^{n+\frac{1}{2}}\left[i, j+\frac{1}{2}\right] = &\frac{1-\frac{\sigma\Delta_t}{2\mu}}{1+\frac{\sigma\Delta_t}{2\mu}} H_x^{n-\frac{1}{2}}\left[i, j+\frac{1}{2}\right] \\ &-\frac{1}{1+\frac{\sigma\Delta_t}{2\mu}}\frac{\Delta_t}{\mu\Delta_y}\left(E_z^n[i, j+1]-E_z^n[i, j]\right) \end{aligned} \tag{2}$$

$$\begin{aligned} H_y^{n+\frac{1}{2}}\left[i, j+\frac{1}{2}\right] = &\frac{1-\frac{\sigma\Delta_t}{2\mu}}{1+\frac{\sigma\Delta_t}{2\mu}} H_y^{n-\frac{1}{2}}\left[i, j+\frac{1}{2}\right] \\ &-\frac{1}{1+\frac{\sigma\Delta_t}{2\mu}}\frac{\Delta_t}{\mu\Delta_x}\left(E_z^n[i, j+1]-E_z^n[i, j]\right) \end{aligned} \tag{3}$$

$$\begin{aligned} E_z^{n+1}[i, j] = \frac{1-\frac{\sigma\Delta_t}{2\epsilon}}{1+\frac{\sigma\Delta_t}{2\epsilon}} E_z^n[i, j] + \frac{1}{1+\frac{\sigma\Delta_t}{2\epsilon}}\Bigg(\frac{\Delta_t}{\epsilon\Delta_x}\left\{H_y^{n+\frac{1}{2}}\left[i+\tfrac{1}{2}, j\right]-H_y^{n+\frac{1}{2}}\left[i-\tfrac{1}{2}, j\right]\right\} \\ -\frac{\Delta_t}{\epsilon\Delta_y}\left\{H_x^{n+\frac{1}{2}}\left[i, j+\tfrac{1}{2}\right]-H_x^{n+\frac{1}{2}}\left[i, j-\tfrac{1}{2}\right]\right\}\Bigg) \end{aligned} \tag{4}$$

where the E_z field component at the particular instant of space and time is estimated using the associated E_z, H_x and H_y components. The updation of fields for a 2D plane with boundary regions is represented in Fig. 1. Along x and y directions, space steps are taken as Δx and Δy. The direction of field components updation is also illustrated in the figure. The variance of field components in S-FDTD is updated as same as the updation of mean field components in FDTD.

2.2 Mean and Variance Approximation

The statistical parameters such as average and variance values of field components using S-FDTD are derived from traditional FDTD equations by Smith and Furse [5].

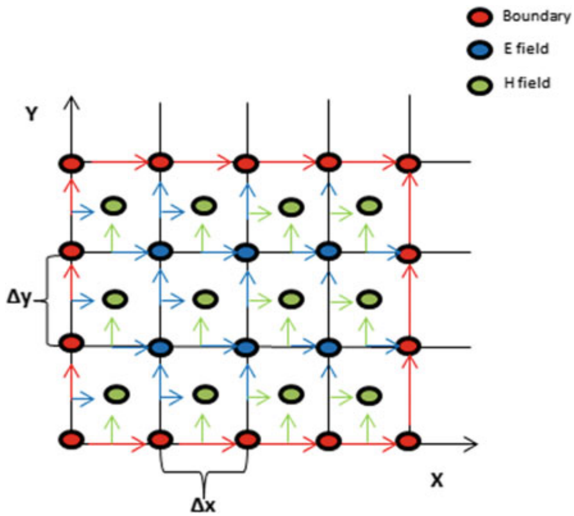


Fig. 1 2D field updation

The mean field values are same as that of FDTD equations as (2)–(4), derived by using Taylor's series approximation of random variables. The variance of field is derived by using delta method and random variable identities. The variance of electric field is derived as in Eq. (5). Where $\hat{\sigma}$ and ε are the conductivity and permittivity of the medium, which are random variables. σ represents the variance function. The correlation terms are taken as one for providing an upper bound to variance estimation and taken as 0.5 for providing a lower bound to variance estimation.

$$
\begin{aligned}
&\sigma\left\{E_z^{n+1/2}(i-1/2, j+1/2)\right\} = \frac{2\varepsilon - \Delta t\hat{\sigma}}{2\varepsilon + \Delta t\hat{\sigma}} \times \sigma\left\{E_z^{n-1/2}(i-1/2, j+1/2)\right\} \\
&+ 4\Delta t\left(\hat{\sigma}\rho_{\varepsilon,\mathrm{E}}\sigma\{\varepsilon\} - \varepsilon\rho_{\hat{\sigma},E}\sigma\{\hat{\sigma}\}\right)\frac{1}{(2\varepsilon + \Delta t\hat{\sigma})^2} \times E_z^{n-1/2}(i-1/2, j+1/2) \\
&+ 2\Delta t\frac{1}{(2\varepsilon + \Delta t\hat{\sigma})} \times E_z^{n-1/2}(i-1/2, j+1/2) \\
&+ 2\Delta t\frac{1}{(2\varepsilon + \Delta t\hat{\sigma})} \times \left(\left(\frac{\sigma\left\{H_y^n(i, j+1/2)\right\} - \sigma\left\{H_y^n(i-1, j+1/2)\right\}}{\Delta x}\right)\right. \\
&- \left(\frac{\sigma\left\{H_x^n(i-1/2, j+1)\right\} - \sigma\left\{H_x^n(i-1/2, j)\right\}}{\Delta y}\right) \\
&- \left(2\sigma\{\varepsilon\}\rho_{\varepsilon,\mathrm{H}} + \Delta t\sigma\{\hat{\sigma}\}\rho_{\hat{\sigma},\mathrm{H}}\right)\frac{1}{(2\varepsilon + \Delta t\hat{\sigma})} \\
&\left.\times\left(\frac{H_y^n(i, j+1/2) - H_y^n(i-1, j+1/2)}{\Delta x} - \frac{H_x^n(i-1/2, j+1) - H_x^n(i-1/2, j)}{\Delta y}\right)\right)
\end{aligned}
\tag{5}
$$

3 Simulation Requirements

3.1 Numerical Model

The spatial steps along x and y direction are taken as 1 mm × 1 mm. The proposed random PCB structure composed of FR-4 substrate material and three conductive traces is shown in Fig. 3. The electrical parameters such as permittivity and conductivity of the PCB substrate are random in nature, and Gaussian probability distribution function is assumed to represent these variations. According to this, the parameters of the substrate material at 2 GHz is taken as follows: permittivity(mean) = 4.4, permittivity(standard deviation) = 0.8, conductivity(mean) = 2.2, and conductivity(standard deviation) = 0.2. A 2D slice of PCB layer structure is used to simulate the field variance. The probability distribution function for dielectric parameters used in each iteration is shown in Fig. 2, which is the permittivity distribution used in the model.

Similar to permittivity, random values of conductivity are also generated with a specific mean and standard deviation. For a defined set of iterations, these random values are picked simultaneously for statistical parameters estimation using the FDTD equations, and the mean and standard deviation values of substrate parameters are directly used in S-FDTD equations.

3.2 Simulation Setup

The 2D S-FDTD method is used here to estimate the variance of electric field in a 2D slice of PCB structure. The proposed random model is simulated with sine plane wave source. The computational domain is made of 40 × 40 cells. The spatial steps

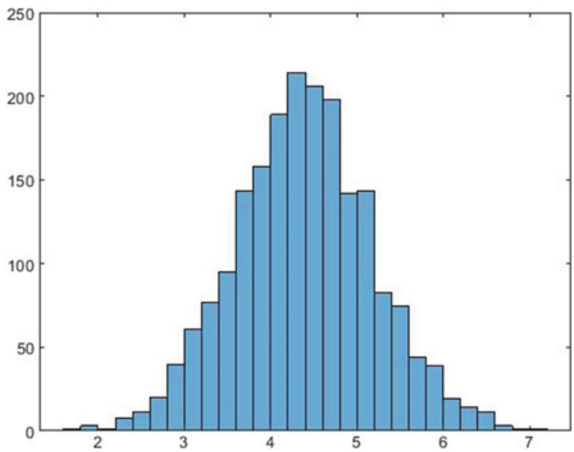


Fig. 2 Random permittivity distribution with mean = 4.4 and standard deviation = 0.8

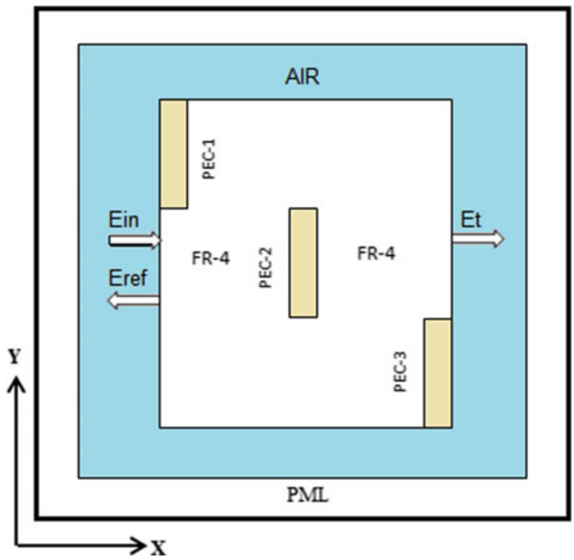


Fig. 3 Two-dimensional slice of PCB model

are equal to 1 mm for both x and y directions. The time step is set as $\Delta x/2c_0$, where c_0 is the speed of light in free space. This meets the Courant stability criteria used in FDTD method [6]. The dielectric dimension is taken as 10×10 cells initially, and 20×20 dimension of dielectric space is then considered for comparing the simulation performance. The PEC regions of the PCB model have a dimension of 2×4 cells.

The random PCB model is met with an incident sine plane wave with amplitude of 1V/m at 2 GHz from left side in TMz mode. The model is surrounded by air. The PML thickness is taken as three cells for both mean and variance to prevent reflections and limit the waveforms within the medium.

3.3 PML for Variance in S-FDTD

The statistical parameters such as mean and variance of electromagnetic fields are basically appeared as waves in S-FDTD method [7]. In order to avoid the reflections into problem space, absorbing boundary conditions should be applied for mean field equation and individual PML for variance equation [2].

4 Simulation Results and Discussion

The proposed model is simulated using MATLAB. The random PCB model shown in Fig. 3 with a sine plane wave impinging from left at a frequency of 2 GHz is considered. The electric field variance at each space instants is observed and plotted.

The S-FDTD simulation is implemented using two approximations for correlation coefficients in order to bound the variance of electric field values. Initially, the variance estimation is done by assuming correlation coefficient values as one, and the resultant waveform is plotted as shown in Fig. 4. By setting the value of correlation coefficients equal to one results in the variance peaks to an upper boundary of the domain space.

Then, second analysis was simulated by putting the correlation coefficient values as 0.5, where the PCB model is exposed to a sine plane wave at 2 GHz and the electric field variance is plotted. By setting the value of correlation coefficients equal to 0.5 outcomes the variance field values of the model with a lower boundary.

The multiple simulations were conducted using FDTD for 2000 iterations, which were considered as the standard for all the simulations. Then, the variance obtained using S-FDTD with the approximations of the correlation coefficient as 1 and 0.5 is compared with the standard results.

In the simulated result, plane wave incident from left at a point in the space and the dielectric is placed at the points between 10 and 30 in X steps.The resultant graph shows near zero field variance at PEC positions 10, 19, 28 with 2 cell width in x plane.

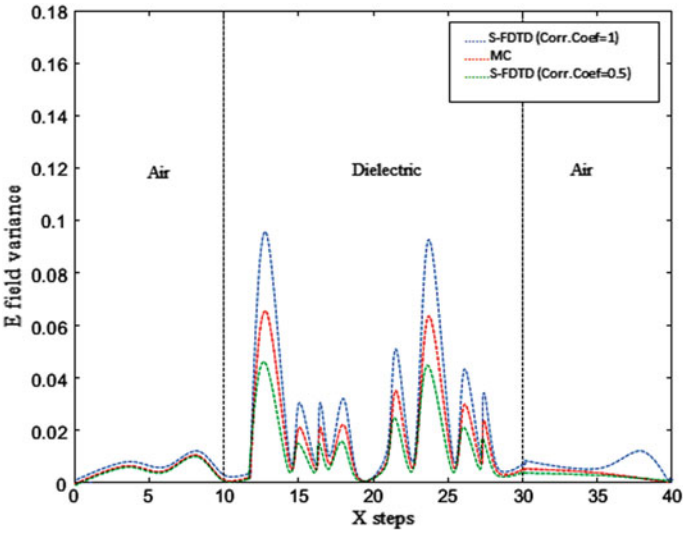


Fig. 4 E field variance in random PCB model with 20×20 dielectric

Table 1 Performance comparison

Dimension	M-FDTD	S-FDTD
40×40 cells with 10×10 dielectric	3 h 34 min	2 min 10 s
40×40 cells with 20×20 dielectric	6 h 43 min	3 min 44 s

The FDTD iterations used in the Monte Carlo method are time-consuming while S-FDTD method requires a fraction of time compared to Monte Carlo. To complete the 2000 runs in Monte Carlo analysis, it takes more than 3.5 h for a dielectric structure of 10×10 cells with Intel(R) Core(TM)i5- 2410M CPU, 2.30 GHz, 4.00 GB of RAM computer. For a 20×20 dimension dielectric structure, it takes near double time. The performance comparison of two dielectric dimensions is summarised in Table 1. Using S-FDTD instead of Monte Carlo, the simulation takes few minutes only. The field variances simulated using Monte Carlo method are taken as the benchmark. The S-FDTD variance estimation using the approximation of correlation coefficient as 1 and 0.5 binds the simulated results with the standard result having corresponding upper and lower values of variance wave.

The Stochastic FDTD method performs single simulation using the updated variance equation. Compared to multisimulation method which uses thousands of simulation, the developed method is very efficient, and it is applicable in the simulation of complex structures. The random three-dimensional structures and multidimensional structures can be analysed more easily using S-FDTD simulation method.

5 Conclusion

A 2D S-FDTD method for the variance field calculation of a PCB plane model with random substrate has been developed.The electric field variance is analysed for a sine plane wave excitation. Due to the variation in the dielectric properties of the substrate medium, there is a deviation in field values, and it appears as a variance wave. By approximating the field components and constitutive parameters that are highly correlated, the upper bound of variance is calculated. Another approximation is taken as field components are less correlated, a lower bound of variance is calculated, and the results are validated using Monte Carlo results. The S-FDTD method offers huge savings in computational time and memory with considerable accuracy. In the future, the method can be extended to analyse electromagnetic field variance in random 3D structures and multilayer random PCB model.

Acknowledgements The authors of this paper are grateful to the staff and students of College of Engineering Trivandrum for their guidance and support.

References

1. Sadiku MNO (2001) Numerical techniques in electromagnetics, 2nd edn. CRC Press, Boca Raton, FL
2. Masumnia-Bisheh K, Ghaffari-Miab M, Zakeri B (2017) Evaluation of different approximations for correlation coefficients in stochastic FDTD to estimate SAR variance in a human head model. IEEE Trans Electromag Compat 59(2):509–517
3. Smith SM, Furse C (2012) Stochastic FDTD for analysis of statistical variation in electromagnetic fields. IEEE Trans Antennas Propag 60:3343–3350
4. Kane Y (1966) Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. IEEE Trans Antennas Propag 14:302–307
5. Smith SM, Furse C (2011) Stochastic FDTD. Ph.D. dissertation, University of Utah, Salt Lake City, UT
6. Sullivan DM. Electromagnetic simulation using the FDTD method. IEEE press series on RF and Microwave Technology
7. Smith SM, Furse C (2011) "A stochastic FDTD method for statistically varying biological tissues" presented at the 2 IEEE AP-S International Symposium on Antennas and Propagation and USNC/CNC/URSI, Spokane, WA

An FDTD Method for the Transient Terminal Response of a PCB Trace Illuminated by an Electromagnetic Wave

M. S. Saheena and R. Kiran

Abstract This paper develops a novel generalized model for the terminal response of a printed circuit board trace, illuminated by an electromagnetic field using a finite difference time domain (FDTD) method. One-dimensional transmission line model has been widely used in time domain simulation, but this model cannot simulate accurately the electromagnetic effect on such structures. Therefore, FDTD is used to study the accurate effect. This desertion develops the transient analysis of PCB by deriving the coupling of the incident field with the geometrical structure using total-field/scattered-field FDTD formulation and then obtained the recursive relation of the electric field and magnetic field for the visualization of the propagation over the problem space. The geometrical structure under study is considered within a perfectly matched layer absorbing boundary grid.

Keywords Finite difference time domain (FDTD) · Total-field/scattered-field (TF/SF) · Transmission line model (TLM) · Perfectly matched layer (PML)

1 Introduction

Modeling of electromagnetic coupling plays an important role in the design and performance of the PCB trace driven by an external electromagnetic field. There are many numerical methods are available for such modeling like method of moments (MoM), finite element method (FEM) in frequency domain and transmission line model (TLM), finite difference time domain (FDTD) in time domain. The FDTD is a full wave transient method, which is found to be simple in terms of concept as well as implementation. While considering the electromagnetic compatibility (EMC) in electronic devices, the coupling effect of the propagating plane wave to the printed

M. S. Saheena (✉) · R. Kiran
College of Engineering Trivandrum, Thiruvananthapuram, Kerala, India
e-mail: saheena1994@gmail.com

R. Kiran
e-mail: kiranr@cet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_17

circuit board (PCB) becomes an essential area of investigation. So FDTD is adopted as the method to study the EM wave interaction as well as the transient analysis of such complicated structures.

There are many practical cases where electromagnetic interference occurred, which will cause to induce terminal voltage at the trace ends where sensitive lumped components may be connected. MarcoLeone et al. put forward an analytical solution for the terminal response of the PCB trace exposed to the external field, which is completely based on TLM model [1]. According to them, voltage and current at the trace end were calculated without considering the secondary reflection from the conducting parts. But nowadays due to the advancement in memory capability, FDTD becomes the advanced tool for such analysis, which gives the direct time domain solution without any frequency transformation. The propagation of electromagnetic wave through air and material media is studied by Sullivan [2] using FDTD.

Kung and Chuah designed and developed a simulating software based on 3D FDTD to study the EM field interaction with the PCB assembly [3]. Gvozdic et al. describe an explicit algorithm based on FDTD, which shows the EM wave propagation in air media with conductive structure (PEC) [4]. Where TF/SF FDTD method is adopted and also considered the secondary reflection from the PEC. Mattey et al. examined the pulse propagation over the ground plane of PCB and made the complex visualization of electric field [5]. Jiang et al. made the simulation of radiation from the PCB board using finite difference method thus mainly focused on the radiation intensity and its distributions [6].

In this paper, we propose a novel method for the interaction of a plane wave with a PCB trace, which is loaded with sensitive lumped components within the FDTD domain bounded by PML. The plane wave excitation for the PCB trace is formulated by TF/SF to avoid the difficulties of using either hard source or the initial condition approaches used for scattering problems. This method employs the volumetric sampling of unknown electric and magnetic field vectors within the structure of interest, and also this paper put forward the 3D visualization of the electric field over the entire domain of observation.

This paper is organized as follows. Section 1 is the introduction. Section 2 described the methodology used, Sect. 3 shows the simulation setup, and Sect. 4 includes results obtained and its discussion. Section 5 deals with the conclusion.

2 Methodology Description

2.1 Building a 3D Model

In this section, we propose a classical formulation method based on Yee's algorithm [2], for the construction of a PCB trace. PCB assembly usually contains two conductive patches of PEC and some lumped components. Three-dimensional model can be considered as a combination of cubes having similar size. In FDTD method, such

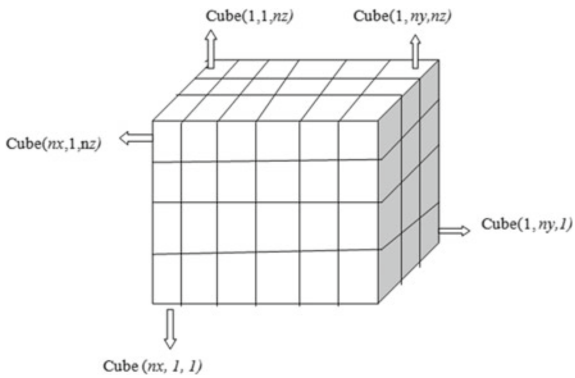


Fig. 1 Discretization of 3D space

cubes are referred as Yee's cell. Where nx, ny and nz denote the number of yee cell in x, y and z direction, respectively. Combination of them together makes the realistic approximation of the PCB assembly.

The 3D problem space gets discretized both in time and space; electric and magnetic fields get updated at $\Delta t/2$ and Δt time intervals, respectively. It can be achieved by the discretization of Maxwell's differential equation by using central difference approximation [7]. The updated equations are given as follows (Fig. 1).

$$H^{n+\frac{1}{2}}_{r(i,j,k)} = H^{n-\frac{1}{2}}_{r(i,j,k)} - \frac{\Delta t}{\mu} \nabla \times E^{n}_{r(i,j,k)} \tag{1}$$

$$E^{n+1}_{r(i,j,k)} = E^{n}_{r(i,j,k)} + \frac{\Delta t}{\varepsilon_{r(i,j,k)}} \left[\nabla \times H^{n+\frac{1}{2}}_{r(i,j,k)} - J^{n+\frac{1}{2}}_{r(i,j,k)} \right] \tag{2}$$

where r can be x, y, z

$$\nabla \times E^{n}_{x(i,j,k)} = \frac{E^{n}_{z(i,j+1,k)} - E^{n}_{z(i,j,k)}}{\Delta y} - \frac{E^{n}_{y(i,j,k+1)} - E^{n}_{y(i,j,k)}}{\Delta z}$$
$$\nabla \times H^{n+\frac{1}{2}}_{x(i,j,k)} = \frac{H^{n+\frac{1}{2}}_{z(i,j,k)} - H^{n+\frac{1}{2}}_{z(i,j-1,k)}}{\Delta y} - \frac{H^{n+\frac{1}{2}}_{y(i,j,k)} - H^{n+\frac{1}{2}}_{y(i,j,k-1)}}{\Delta z}$$

2.2 Incorporation of Lumped Element

The electric and magnetic fields over Yee grid are shown in Fig. 2. While considering the PEC patches on the PCB electric field over the trace edge permanently set to zero. During the introduction of lumped components to the trace ends, the electric field gets modified to some extent as follows [3].

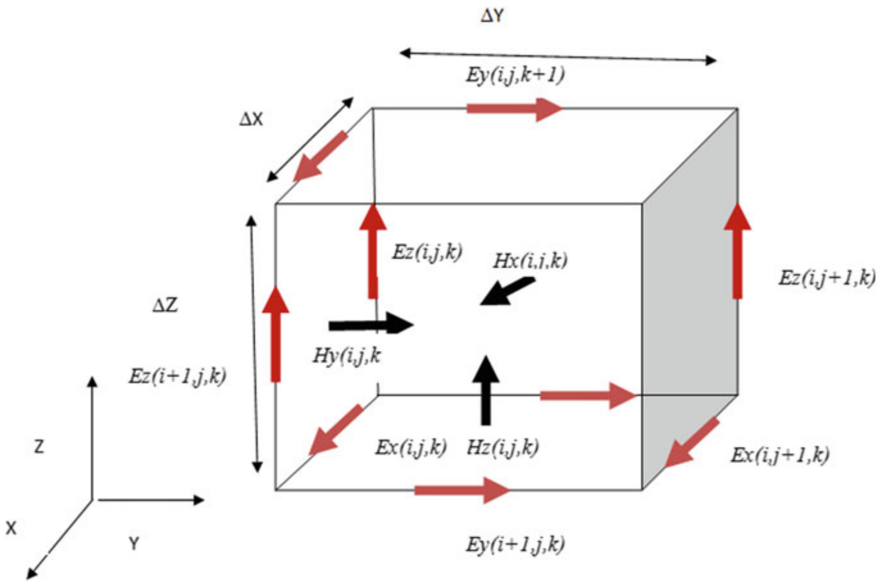


Fig. 2 Single Yee cell with electric and magnetic field distribution

1. Resistor R incorporated along z direction

$$E^{n+1}_{z(i,j,k)} = \left(\frac{1 - \frac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}}{1 + \frac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}} \right) E^{n}_{Z(i,j,k)} + \left(\frac{\frac{\Delta t}{\varepsilon}}{1 + \frac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}} \right) \nabla \times H^{n+\frac{1}{2}}_{z(i,j,k)} \tag{3}$$

where

$$\nabla \times H^{n+\frac{1}{2}}_{z(i,j,k)} = \frac{H^{n+\frac{1}{2}}_{y(i,j,k)} - H^{n-\frac{1}{2}}_{y(i-1,j,k)}}{\Delta x} - \frac{H^{n+\frac{1}{2}}_{x(i,j,k)} - H^{n-\frac{1}{2}}_{x(i,j-1,k)}}{\Delta y}$$

2.3 *TF/SF FDTD Formulation*

TF/SF formulation provides an infinite plane wave source by introducing an incident field at TF/SF boundary, and in this case wave is generated at the XZ plane as shown in Fig. 3. Where *ia*, *ja*, *ka* arr TF/SF generating boundary and *ib*, *jb*, *kb* are the terminating boundary. The structure for the analysis is placed inside the total-field as shown.

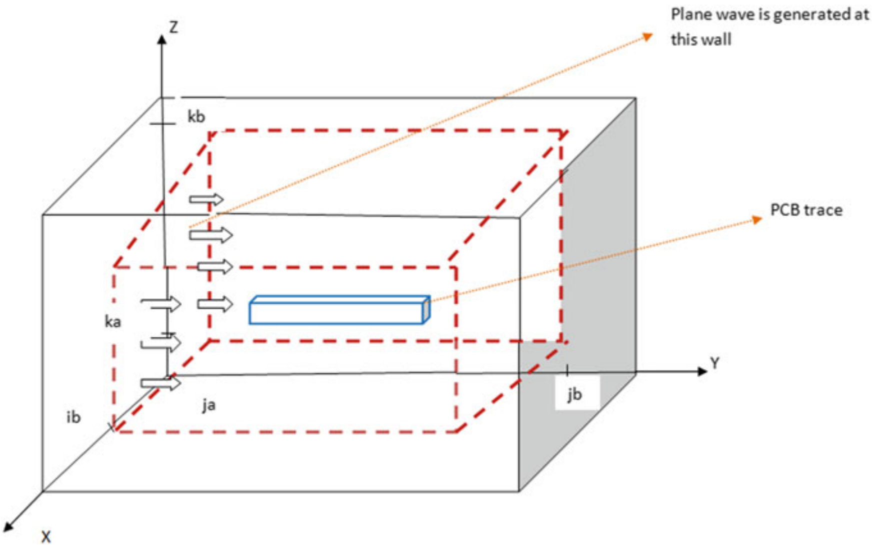


Fig. 3 PCB inside TF/SF FDTD grid

Due to the limitation of memory requirement, the calculation region should be truncated by some boundary. But this intentionally inserted boundary sometimes causes the reflection problems, so that we use the PML boundary condition [2].

The wave generated at the XZ plane of the boundary will be added to the flux density(D) or to H field and subtracted on the other side by subtraction from D or H field [2]. During the updating of field values of a point inside, the total-field may use field values outside total-field and also a point lying outside may require the use of point inside. So there are four places where some modifications are found for updating equations as follows:

1. D_z value at $j = ja$ and $j = jb$

$$\begin{aligned} D_z\,(i, j_a, k) &= D_z\,(i, j_a, k) + 0.5 \cdot H_{x\_\mathrm{inc}}\,(j_a - 1/2) \\ D_z\,(i, j_b, k) &= D_z\,(i, j_b, k) - 0.5 \cdot H_{x\_\mathrm{inc}}\,(j_b + 1/2) \end{aligned} \tag{4}$$

2. H_x value just outside $j = ja$ and $j = jb$

$$\begin{aligned} H_x\,(i, j_a - 1/2, k) &= H_x\,(i, j_a - 1/2, k) + 0.5 \cdot E_{z\_\mathrm{inc}}\,(j_a) \\ H_x(i, jb + 1/2, k) &= H_x(i, jb + 1/2, k) - 0.5 \cdot E_{z\_\mathrm{inc}}(jb) \end{aligned} \tag{5}$$

3. H_y just outside $i = ia$ and $i = ib$

$$\begin{aligned} H_y\,(i_a - 1/2, j, k) &= H_x\,(i_a - 1/2, j, k) - 0.5 \cdot E_{z\_\text{inc}}(j) \\ H_y\,(i_b + 1/2, j, k) &= H_x\,(i_b + 1/2, j, k) + 0.5 \cdot E_{z\_\text{inc}}(j) \end{aligned} \tag{6}$$

4. D_y field at $k = ka$ and $k = kb$

$$\begin{aligned} D_y(i, j + 1/2, ka) &= D_y(i, j + 1/2, ka) - 0.5 \cdot H_{x-\text{inc}}(j) \\ D_v(i, j + 1/2, kb + 1) &= D_y(i, j + 1/2, kb + 1) + 0.5 \cdot H_{x-\text{inc}}(j) \end{aligned} \tag{7}$$

3 Simulation Setup

In order to calibrate the 3D FDTD analysis for the terminal response of PCB trace with lumped loads and for the field visualization, we set up a simple model as shown in Fig. 4. The PCB trace having conductive patches and resistive components placed over the total-field is shown in Fig. 5. TF/SF problem space is bounded by the PML. The 3D FDTD grid has cube cell of side $\Delta x = \Delta y = \Delta z = 0.01$ mm. The TF/SF boundary is set along XZ plane at $ia = ja = ka = 7$. The number of time step is set as 1000, and discretized interval is chosen according to Courant condition [2]. The total grid size is assumed to be $60\Delta x * 60\Delta Y * 60\Delta Z$. This setup is boundered over a PML region.

The incident field is generated from TF/SF boundary propagating along the Y direction with HX and EZ components. The wave frequency is taken as 400 MHz.

Incident pulse = $\sin(2 * pi * f0 * t)$

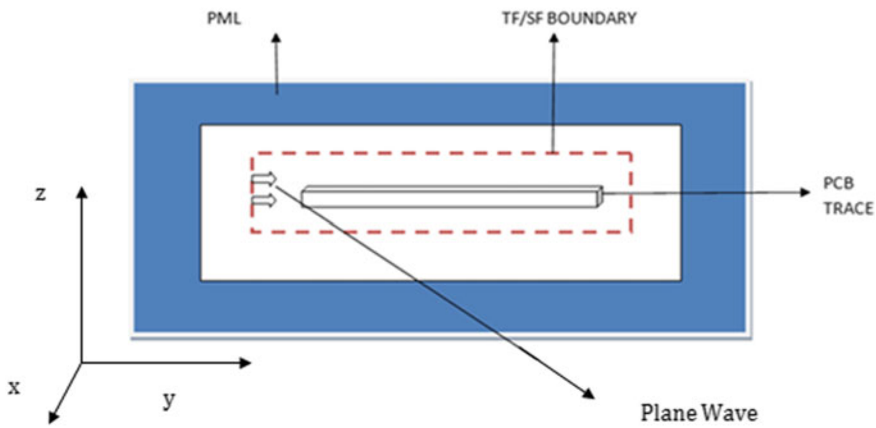


Fig. 4 TF/SF formulation in 3D FDTD

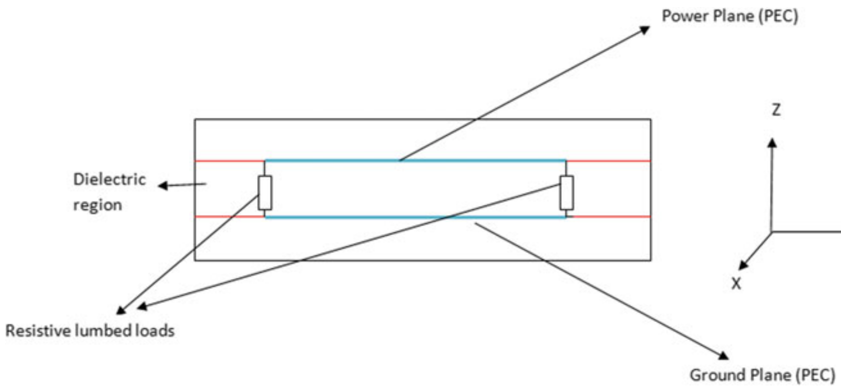


Fig. 5 Longitudinal cross section of a PCB trace

Within the 3D space PCB trace arranged with a dielectric FR4, the grid position of the dielectric along x axis is (13, 20), y axis along (10, 30) and z axis between (10, 15). The power plane having grid position for x axis is between (13,10), y axis along (12–29), z axis over position 16. The ground plane is placed parallel to power plane at the bottom side so that the z axis position becomes 10. The traces are terminated with the resistance of values 50 Ω at locations ([18, 15, 14]) and ([18, 28, 14]).

4 Simulation Results and Discussion

The terminal responses on the trace loads are simulated using MATLAB. The results are illustrated in the following figures for the resistance of 10 Ω. Figures 6 and 7 show the voltage induced on the resistance value of 10 Ω on the trace end at the location ([18, 15, 14]) and ([18, 28, 14]). There is delay and reduction in magnitude for the induced voltage on the second terminal compared with the first due to the propagation and secondary reflection effect of the PEC. After a particular number of transient cycle, the voltage gets in to a steady state.

The FDTD algorithm is validated by the OpenEMS platform [7]. Comparison of terminal responses for both load is given in Figs. 6 and 7.

5 Conclusion

A 3D FDTD method for the transient analysis of a PCB trace illuminated by an external field over a TF/SF domain has been developed. The transient response is studied for a sine plane wave excitation. From the simulation, it can be concluded

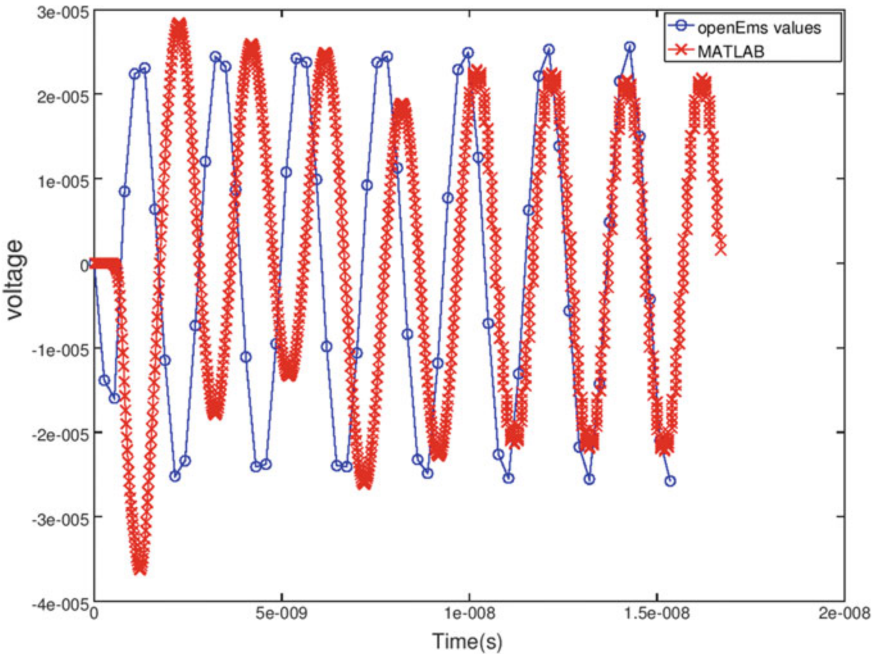


Fig. 6 Comparison of terminal response of PCB trace at the source end for 10 Ω resistor with OpenEms

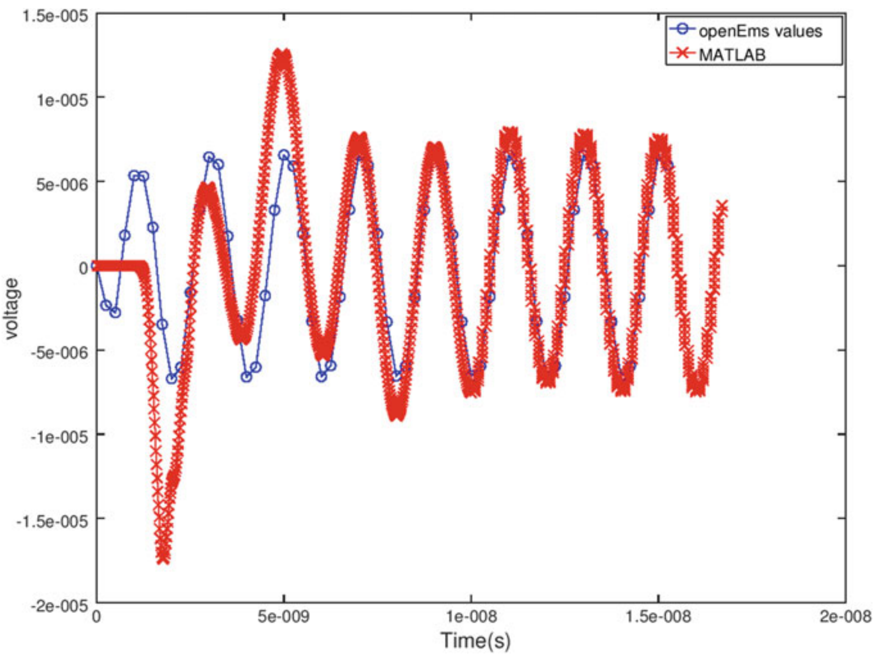


Fig. 7 Comparison of terminal response of PCB trace at the load end for 10 Ω resistor with OpenEms

that the terminal voltage induced over the resistive components at the two ends contains two stages: transient and steady-state stage. There is a delay and magnitude reduction occurred between the two end responses due to the reflection effect due to the conductive patches of PEC.

Acknowledgements The authors of this paper are grateful to the staff and students of College of Engineering Trivandrum for their guidance and support.

References

1. Leone M, Singer HL (1999) On the coupling of an external electromagnetic field to a printed circuit board trace. IEEE Trans Electromag Compat 41(4):418–424
2. Sullivan DM, Electromagnetic simulations using FDTD, 3rd edn., pp 90–100
3. Kung F, Chuah HT (2005) A finite difference time domain (FDTD) software for simulation of printed circuit board (PCB) assembly. In: PIER 50, pp 299–335
4. Gvozdic BD, Durdevic DŽ (2014) FDTD simulations of TF/SF plane waves in the presence of PEC scatterers, pp 163–166
5. Mattey NL, Edwards G, Hood RJ. A time domain approach to power integrity for printed circuit boards. Electrical & Optical Systems Research, pp 1–11
6. Jiang Y, Li L-W, Li E-P (2002) Design and analysis of printed circuit boards using FDTD method for the 20-H Rule. In: MIT alliance annual symposium, Singapore
7. Liebig T (2019) openEMS—an introduction and overview. In: Fosdem 19

Synthesis of Pseudorandom Number Generator by Combining Mentor Graphics HDL Designer and Xilinx Vivado FPGA Flow

Geethu Remadevi Somanathan, Ramesh Bhakthavathchalu, and M. Krishnakumar

Abstract Pseudorandom number generators are used in cryptographic as well as VLSI testing applications. Linear Feedback Shift Registers (LFSR) are circuits that can be used as pseudorandom number generators. This paper proposes a modified reseeding method for LFSR and also presents a design flow for the implementation. The work is done by combining Mentor Graphics HDL Designer FPGA flow and Xilinx Vivado. Different bit lengths of LFSR are generated using Mentor Graphics HDL Designer and synthesized using both Mentor Graphics Precision RTL synthesizer and Xilinx Vivado. The implementation targets Virtex-7 FPGA.

Keywords Pseudorandom number generator · Linear feedback shift register · Reseeding LFSR

1 Introduction

A pseudorandom number generator (PRNG) generates a sequence of numbers, which has the properties of random numbers. Sequences generated by PRNG is not truly random, instead, is determined by a starting state. This starting state is known as initial seed value, which may or may not be random in nature. PRNGs are vital for cryptography, gaming, and VLSI testing applications.

Based on the seed value, PRNGs generate a sequence of numbers approximating the properties of random numbers. Many numbers are generated in a short time and can also be replicated later if the seed used in the sequence is known. Linear

G. R. Somanathan (✉) · R. Bhakthavathchalu
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: geethurs@am.amrita.edu

R. Bhakthavathchalu
e-mail: rameshamrita@gmail.com

M. Krishnakumar
KSEB Ltd, Thiruvananthapuram, India
e-mail: krishnanow@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_18

Feedback Shift Registers (LFSR) are circuits that can be used as pseudorandom number generators since the area overhead is small and due to the quality of numbers generated [1]. Basic working and classifications of LFSR are not mentioned in this paper and can be found in [2]. The paper also discusses the two main classifications, internal feedback, and external feedback LFSR. Depending on the output sequences generated by an LFSR they are also classified as maximum-length or complete LFSR [2–5]. Generation of maximum-length sequences by utilizing the feedback connections is proposed in paper [6], where a method for calculations is also provided. A comparative study of the different circuits developed from Linear Feedback Shift Register (LFSR) which are used for generating the deterministic and pseudorandom numbers is also mentioned in [6]. Papers [7–9] present the utilization of LFSR as a pseudorandom number generator. Various PRNGs used for security purpose to generate SSL connections, encryption keys are compared in [7]. Paper [10] proposes a method for building a transformation matrix for Cyclic Redundancy Check (CRC) using LFSR. Fundamental property required in cryptography is randomness, the precision of which determines the power of the encryption algorithm [11]. The difference between the two basic types true random number generator (TRNG) and pseudorandom number generator (PRNG) and the mathematical modeling can be found in [7, 12]. Papers [7, 13] also discusses the different types of PRNGs and its application in a different spectrum-like internet security, data security, key generation, and secured connections. Recently, the research trend inclines towards chaos-based PRNG [14], where the output sequences show better randomness results than conventional PRNGs. This paper presents a design flow method and analysis of pseudorandom number generator based on LFSR. The proposed design flow takes less time in designing and verifying an LFSR. The method can be used in the design flow and RTL synthesis of Verilog/VHDL implementations.

2 Proposed Implementation Flow

The proposed design flow is implemented in Fig. 1, which shows the proposed architecture for pseudorandom number generation. This design considers a modified reseeding mechanism implemented in [15]. The design in [15] consists of a normal up counter and LFSR, where the ‘0000’ state of up counter is used in basic design but not used as a seed value since it creates a lock state for the LFSR. LFSR used in the paper [15] generates maximum-length sequences. This paper uses a counter, which omits the zero state of the counter in the up sequence. Output of the counter is used as seed values for the LFSR. Seed values will be supplied to the LFSR depending on the value in enable signal. Depending on the seed value given to LFSR, a set of pseudorandom numbers are generated. In this architecture, enable value ‘0’ changes the counter output states and thus, the seed values given to the LFSR. Enable value ‘1’ generates the pseudorandom numbers depending on the seed value at that time. In this work, the LFSRs considered are of maximum-length and so, the number of pseudorandom numbers generated for each seed will be 2^{n-1} [4, 16].

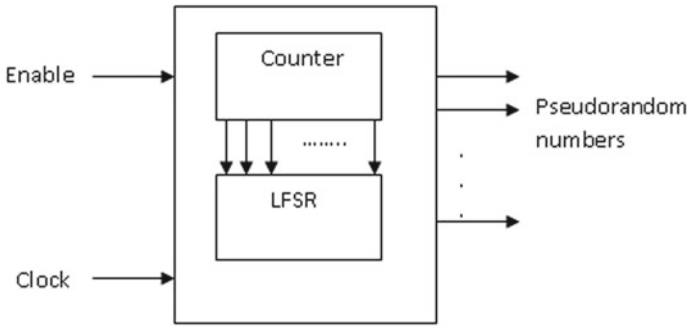


Fig. 1 Proposed architecture for pseudorandom number generation

3 Implementation Flow

2-bit to 4-bit length of the proposed architecture is implemented using Mentor Graphics HDL Designer [17]. Characteristic polynomials for the LFSRs used in the design are given from Eqs. (1)–(3).

$$x^2 + x + 1 \tag{1}$$

$$x^3 + x^2 + 1 \tag{2}$$

$$x^4 + x^3 + 1 \tag{3}$$

The pseudorandom numbers according to the polynomials given in Eqs. (1)–(3) are calculated. Mentor Graphics HDL Designer is used to design the proposed architecture using these pseudorandom numbers and generate the implementation. Counter output which forms the seed value and the corresponding pseudorandom numbers generated are given from Tables 1, 2 and 3. It can be verified from these tables that the length of the sequences are of maximum-length, 2^{n-1} [15]. Table 1 lists the output of counter and resultant pseudorandom numbers generated by 2-bit LFSR. Table 2 shows the output pseudorandom sequences for the 3-bit LFSR given by Eq. (2). Similarly, Table 3 shows output pseudorandom sequences for 4-bit LFSR given by Eq. (3).

Table 1 2-bit pseudorandom pattern sequences

Counter output	Pseudorandom numbers generated	No. of output sequences
01	01, 10, 11, …	3
10	10, 11, 01, …	3
11	11, 01, 10, …	3

Table 2 3-bit pseudorandom pattern sequences

Counter output	Pseudorandom numbers generated	No. of output sequences
001	001, 100, 010, 101, 110, 111, 011, …	7
010	010, 101, 110, 111, 011, 001, 100, …	7
011	011, 001, 100, 010,101, 110, 111, …	7
100	100, 010,101, 110, 111, 011, 001, …	7
101	101, 110, 111, 011, 001, 100, 010, …	7
110	110, 111, 011, 001, 100, 010, 101, …	7
111	111, 011, 001, 100, 010, 101, 110, …	7

State diagram for the above-mentioned bit lengths is used as input for the HDL designer. State diagram for the 3-bit version of the proposed system obtained from HDL designer is given in Fig. 2. Similar implementations using state diagrams are used for 2-bit and 4-bit length of the proposed system.

The generated Verilog files of the proposed systems are synthesized using Mentor Graphics Precision RTL synthesizer as well as Xilinx Vivado.

The generated Verilog files of the proposed systems are synthesized using Mentor Graphics Precision RTL synthesizer as well as Xilinx Vivado.

4 Experimental Results

The generated Verilog design file from HDL Designer is used in Xilinx Vivado to generate RTL. RTL schematic diagram of the proposed system in 4-bit is given in Fig. 3.

Figures 4 and 5 show the RTL schematic of proposed system for 3-bit and 2-bit, respectively. The device targeted is Virtex-7 and is implemented for a frequency of 100 MHz; with a delay for 5 ns. The post-implementation utilization percentiles for all the 3-bit lengths are compared. It can be seen from Fig. 6 that the utilization percentile is the same for all the three. This value will change when the number of states or number of bits increases.

Proposed architectures are simulated for a frequency of 100 MHz with a fall delay of 5 ns. Table 4 shows the total on-chip power for different bit lengths. Figure 7 shows its variation and this clearly shows an increase in the power consumed when the number of bits is increasing.

Table 3 4-bit pseudorandom pattern sequences

Counter output	Pseudorandom numbers generated	No. of output sequences
0001	0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, …	15
0010	0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, …	15
0011	0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, …	15
0100	0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, …	15
0101	0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, …	15
0110	0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, …	15
0111	0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, …	15
1000	1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, …	15
1001	1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, …	15
1010	1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, …	15
1011	1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, …	15
1100	1100, 0110, 1011, 0101, 1010, 1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, …	15
1101	1101, 1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, …	15
1110	1110, 1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, …	15
1111	1111, 0111, 0011, 0001, 1000, 0100, 0010, 1001, 1100, 0110, 1011, 0101, 1010, 1101, 1110, …	15

Worst Negative Slack (WNS) corresponds to the worst slack of all the timing paths for max delay analysis. This can be positive or negative. A negative slack indicates a problem in which the path violates a required setup time. Worst Hold Slack (WHS) corresponds to the worst slack of all the timing paths for min delay analysis. Like WNS this value can be positive or negative. A negative slack indicates a problem in which the path violates a required hold time [18, 19]. WNS and WHS are given in Tables 5 and 6. Both WNS and WHS are positive in different cases. Figures 8 and 9 show the decrease in slack when bit length is increasing.

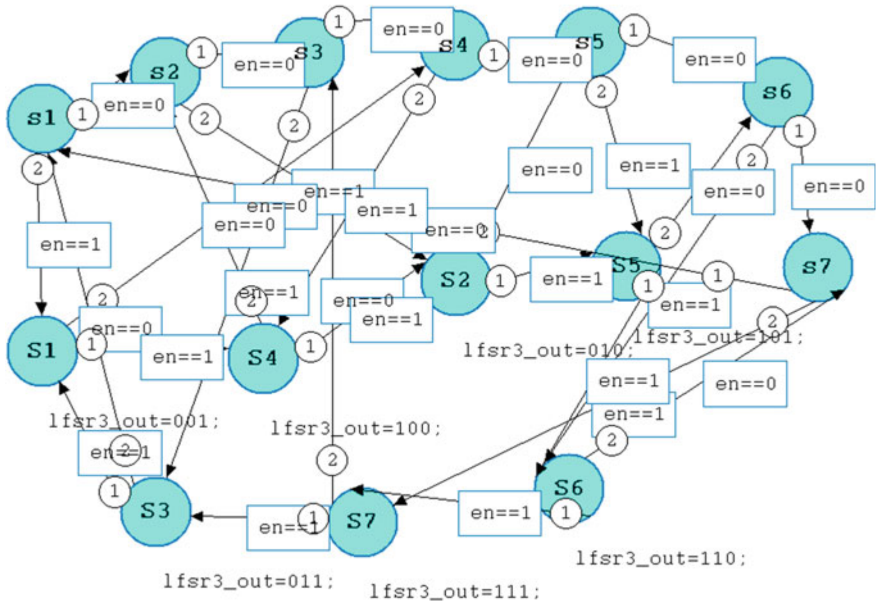


Fig. 2 3-bit state diagram for proposed method

The slice logic for the proposed logic is also compared. Figure 10 shows the variation in slice logic for the different bit lengths and Table 7 shows the corresponding values.

5 Conclusion and Future Scope

LFSR is used as an efficient pseudorandom pattern generator and several modified versions are implemented. This paper presents a study and analysis of the design and implementation of pseudorandom number generators by combining Mentor Graphics HDL Designer FPGA flow and Xilinx Vivado. Different bit lengths of LFSR are generated using Mentor Graphics HDL Designer and synthesized using both Mentor Graphics Precision RTL synthesizer and Xilinx Vivado. The implementation targets Virtex-7 FPGA. The work is carried forward by increasing bit lengths for improved verification. After validation, the possibilities of reseeding can be incorporated in the proposed flow.

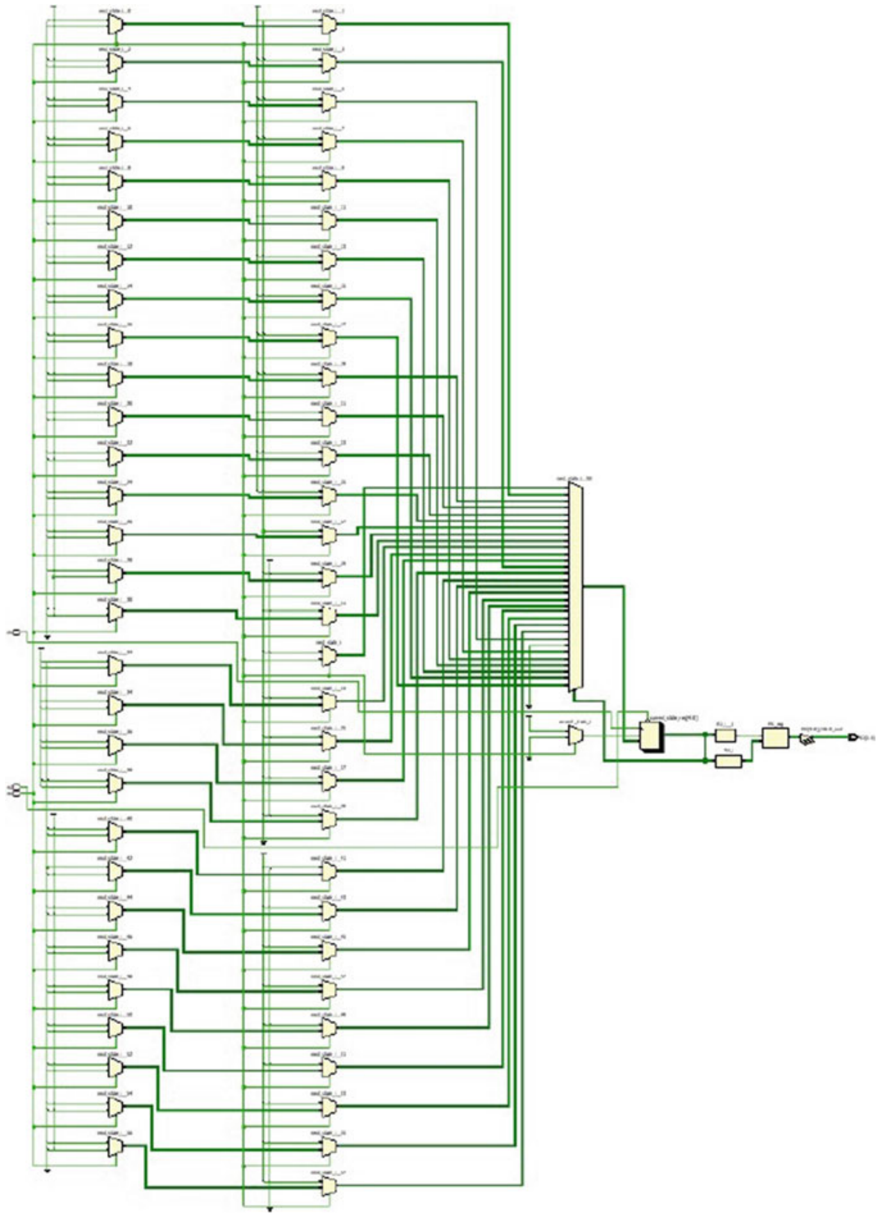


Fig. 3 Schematic of proposed system for 4-bit

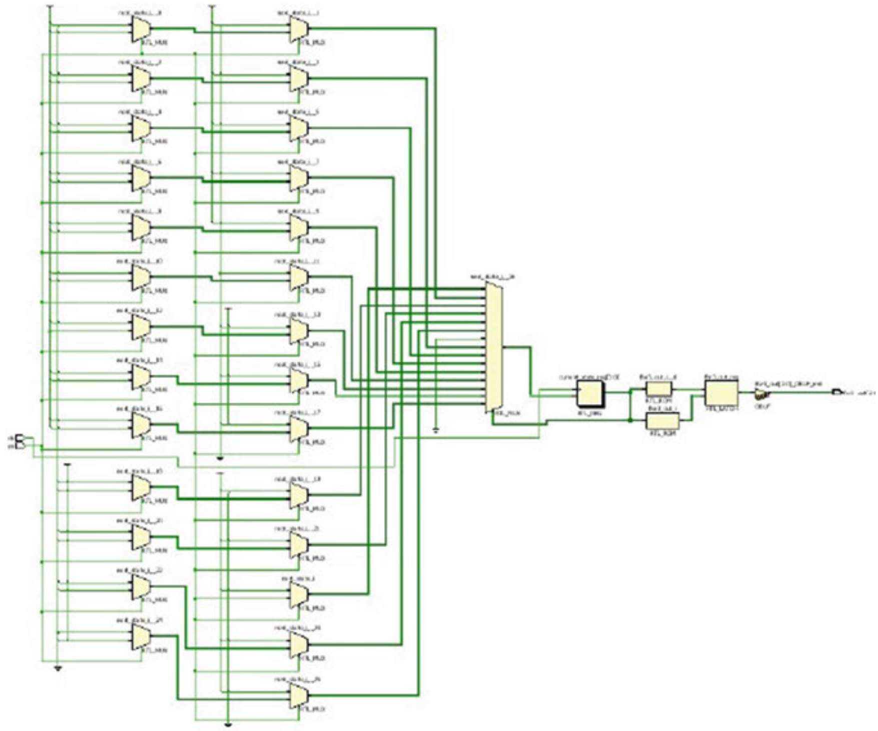


Fig. 4 Schematic of proposed system for 3-bit

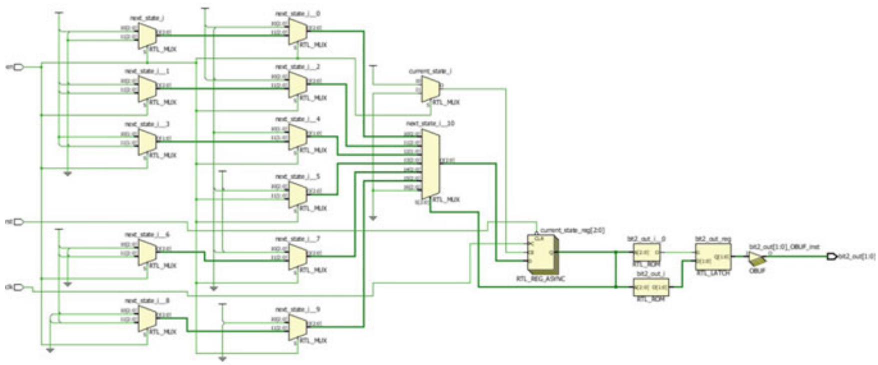
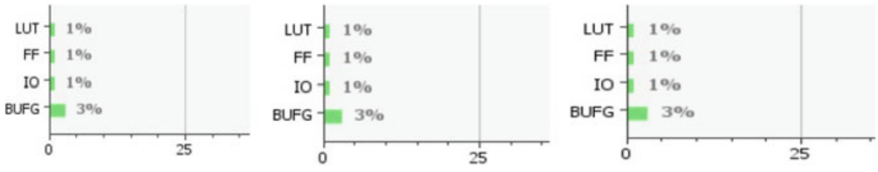


Fig. 5 Schematic of proposed system for 2-bit



(a) 2-bit (b) 3-bit (c) 4-bit

Fig. 6 Post-implementation percentile

Table 4 Total on-chip power

Proposed architecture (bits)	Total on-chip power (W)
2	0.182
3	0.182
4	0.184

Fig. 7 Bit length versus total on-chip power

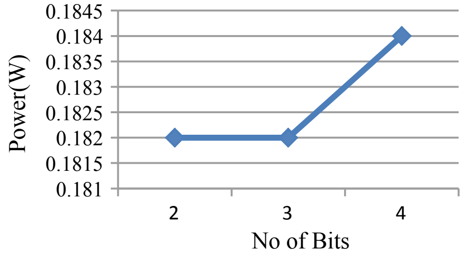


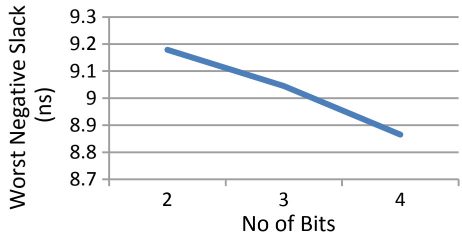
Table 5 Worst negative slack

Proposed architecture (bits)	WNS (ns)
2	9.179
3	9.045
4	8.865

Table 6 Worst Hold Slack

Proposed architecture (bits)	WHS (ns)
2	0.282
3	0.258
4	0.175

Fig. 8 Bit length versus WNS



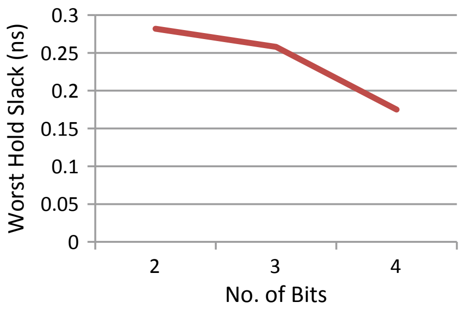


Fig. 9 Bit length versus WHS

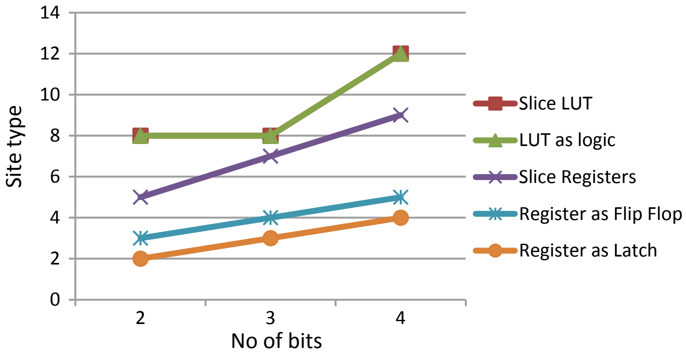


Fig. 10 Site type versus bit length

Table 7 Slice logic of 2, 3, 4-bits

Site type	2-bit	3-bit	4-bit
Slice LUTs	8	8	12
LUT as logic	8	8	12
Slice registers	5	7	9
Register as flip flop	3	4	5
Register as latch	2	3	4

References

1. Youhua Shi ZZ (2003) Multiple test set generation method for LFSR-based BIST. In: Proceedings of the ASP-DAC Asia and South Pacific design automation conference, pp 863–868. https://doi.org/10.1109/ASPDAC.2003.1195138
2. Dhingra S Comparison of LFSR and CA for BIST
3. Devika KN, Bhakthavatchalu R (2017) Design of reconfigurable LFSR for VLSI IC testing in ASIC and FPGA. In: International conference on communication and signal processing, 6–8 Apr 2017. India Design, pp. 928–932

4. Bushnell M, Agrawal V (2000) Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits
5. Wang L-T, Wu C-W, Wen X VLSI test principles and architectures
6. Mukherjee N, Rajski J, Mrugalski G, Pogiel A, Tyszer J (2011) Ring generator : an ultimate linear feedback shift register. Computer (Long Beach Calif) 64–71. https://doi.org/10.1109/MC.2010.334
7. Divyanjali TA (2015) A dissection of pseudorandom number generators. In: 2nd international conference on signal processing and integrated networks, SPIN 2015, pp 318–323. https://doi.org/10.1109/SPIN.2015.7095369
8. Stpie R, Walczak J (2013) Comparative analysis of pseudo random signals of the LFSR and DLFSR generators
9. Datta D, Datta B, Dutta HS (2017) Design and implementation of multibit LFSR on FPGA to generate pseudorandom sequence number. In: Proceedings of 2nd international conference on 2017 devices for integrated circuit, DevIC 2017, pp 346–349. https://doi.org/10.1109/DEVIC.2017.8073966
10. Prabha LS, Geethu RS (2018) Architecture of parallel CRC encoder using state space transformations. In: 2018 9th international conference on computing, communication and networking technologies, ICCCNT 2018. IEEE, pp 1–7. https://doi.org/10.1109/ICCCNT.2018.8493862
11. Tehranipoor M (2012) Introduction to hardware security and trust. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-8080-9
12. Gergely AM, Crainicu B (2017) A succinct survey on (Pseudo)-random number generators from a cryptographic perspective. In: 2017 5th international symposium on digital forensic and security, ISDFS 2017, pp 1–6. https://doi.org/10.1109/ISDFS.2017.7916504
13. Gayoso CA, González C, Arnone L, Rabini M, Moreira JC (2013) Pseudorandom number generator based on the residue number system and its FPGA implementation. In: Proceedings of the argentine school micro-nanoelectronics, technology and applications, EAMTA 2013
14. Garcia-Bosque M, Perez-Resa A, Sanchez-Azqueta C, Aldea C, Celma S (2019) Chaos-based bitwise dynamical pseudorandom number generator on FPGA. IEEE Trans Instrum Meas 68:291–293. https://doi.org/10.1109/TIM.2018.2877859
15. Dilip PS, Somanathan GR, Bhakthavatchalu R (2019) Reseeding LFSR for test pattern generation. In: Proceedings of the 2019 IEEE international conference on communication and signal processing, ICCSP 2019, pp 921–925. https://doi.org/10.1109/ICCSP.2019.8698025
16. Wang W-L, Lee K-J (2002) An efficient deterministic test pattern generator for Scan-Based BIST Environment
17. Hoare RR, Tung SC (2003) Combining mentor graphics' HDL designer FPGA flow with a reconfigurable system on a programmable chip, educational opportunity or insanity? In Proceedings of the 2003 IEEE international conference on microelectronic systems education, tomorrow's microsystems designers, MSE 2003, pp 128–130. https://doi.org/10.1109/MSE.2003.1205286
18. Xilinx Inc (2015) Vivado design suite user guide-design analysis and closure techniques
19. Xilinx Inc (2015) Vivado design suite user guide-using the Vivado IDE

Analysis on Extraction of Common Foreground Object by Co-Segmentation

K. S. Dhanya, N. Naveen, and Jacob Jaison

Abstract Co-segmentation is a subclass of segmentation used to segment out the common objects by assigning multiple labels to segment the targeted common things. Methods existing so far do not exhibit competitive results for any specific task. The key issue of the proposal selection-based co-segmentation problems lies in mining consistent information shared by the common targets. Due to uncertainty of the shared features, it usually requires manually selecting features or feature learning performed beforehand. The main goal is the comparison of an existing and proposed co-segmentation method. Proposed method (second method) focuses on reduction of running time. For each image, a multi-search strategy extracts target individually. Experiments are orchestrated on public dataset MSRC and iCoseg. Performance evaluation of both the methods is compared.

Keywords Co-segmentation · Object recognition · Graph cuts · Saliency maps

1 Introduction

Co-segmentation kingpin on common object in multiple images, while simple segmentation kingpin on segregating an image into regions by allocating labels. In generic class segmentation, extra training datasets should be needed for investigation of a model or abruptly finding the parameters, co-segmentation can regain self-executing object segmentation with priors from the images to be segmented. Co-segmentation can be represented as a modification process. With less manual workload, co-segmentation can accurately segment common objects from multiple

K. S. Dhanya (✉) · N. Naveen · J. Jaison
Rajagiri School of Engineering and Technology, Kochi 682039, India
e-mail: dhanyaks16@gmail.com

N. Naveen
e-mail: naveenn@rajagiritech.edu.in

J. Jaison
e-mail: jaison_jacob@rajagiritech.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_19

related images compared with normal segmentation methods. In an unsupervised manner, co-segmentation can be accomplished by appending foreground similarity constraint and grouping the data capability and can utilize both the inurn and intra-image priors and flourish to acceptable for so many multimedia and computer vision solicitations unlike the segmentation [1, 2].

The main goal is the comparison of two co-segmentation methods. Experiments are orchestrated on public dataset MSRC and iCoseg. In first method, the idea of [3] is used. The key point of the indefinite number of common targets involved co-segmentation problem lies in adaptive determining the number of targets, which requires fully extracting the potential targets and then mining the consistent relationships shared by the common targets. Targets are generated using object proposal generation method [4]. The weight coefficient corresponds to the proposal similarities, and a completely connected graph is constructed using weight selecting algorithm. Multiple common targets of each image are individually extracted using an adaptive label threshold T. When the multi-search process terminates, the final segmentation is done by collecting all the selected proposals. Second method, graph cut methods state image segmentation as a label assignment problem. So, this method segregates the image domain by assigning each pixel as a label (graph cuts), then finds saliency map.

2 Related Work

In the sector of image co-segmentation and numerous algorithms submitted by number of researchers, they still have crucial restrictions. Majority of the existing algorithm does not generally address the co-labelling issue of multiple images because it is very complicated and time-consuming process for the colossal dataset. In colossal image datasets, image co-segmentation is rigorous, owing to an additional energy term inter-image consistency is executed, which is not absolute for simple segmentation. When no genuine prior information is given, complication arises from the ambivalence connecting the foreground object as well as the background. Probability of losing the similarity information among images in non-identical subsets is there if clustering methods are used for co-segmentation. From the large clumps of images, clustering is not a satisfactory procedure, because it circumvents to straightly co-segment whole image set [1].

Co-segmentation method [5] where first they segment the original images into a different number of logical regions, attained by combining superpixel-based segmentation, object detection and saliency-based segmentation methods. Next, they outline a digraph the local region resemblance by feature distance and saliency map. They use co-saliency strategy and obtain more faultless saliency map. Matching fabricated digraph, co-segmentation can be attained by choosing set of nodes with maximal sum of weights. By this way, they develop the co-segmentation problem as a shortest path problem in the endmost step, and they employ powerful programming method to decode this condition [5].

Despite having many algorithms and approaches, enormous need of an optimal solution should be there. The main goal is the comparison of two co-segmentation methods. Experiments are orchestrated on public dataset MSRC and iCoseg. In first method, the idea of [3] is used. Second method, graph cut method [6] and saliency map [7] from the output of graph cuts are employed as a co-segmentation scheme.

3 Co-segmentation Method 1

The block diagram of the first co-segmentation scheme is manifested in Fig. 1. The key point of the indefinite number of common targets involved co-segmentation problem lies in adaptively determining the number of targets, which requires fully extracting the potential targets and then mining the consistent relationships allocated by the prevalent targets. Extracting the potential targets means discovering all the implied objects, and mining the consistent relationships is required to determine the common interesting target that reappears in every image. Targets are generated using object proposal generation method. The weight coefficient corresponds to the proposal similarities, and a unreservedly connected graph is constructed using weight selecting algorithm. Multiple prevalent targets of each image are individually extracted employing a adaptive label threshold T. When the multi-search action breaks off, the final segmentation is done by collecting all the selected proposals.

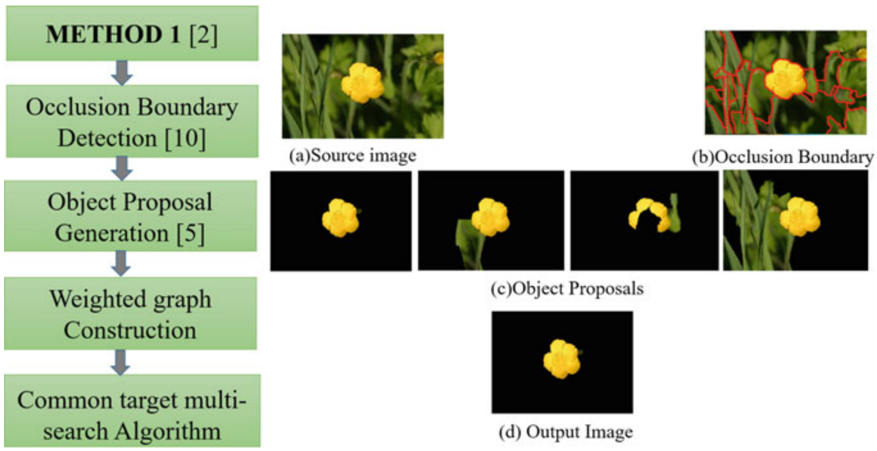


Fig. 1 Block diagram of co-segmentation method 1

3.1 Occlusion Boundary Detection

First train classifiers which foresee occlusion and figure/ground labels for hypothesized boundary fragments. A variety of cues represents statistics over the boundary. In spite of the fact that each classifier employs the features, like the similarity of region colours, which is used for occlusion versus non-occlusion detection, the other one is the signed difference of surface confidences, which is more helpful for figure/ground labelling. If features of the regions that are separated by the boundary is considered, it will be helpful. Adjacent regions having homogeneous colours or well-aligned are more likely to correlate with the same object. For figure/ground labelling, an important cue is the position, where the lower region is more likely to be foreground. Region cues are found out after representing the image in L*a*b* space. Long, smooth boundaries with strong colour or texture gradients are considered in boundary cues. Occlusion boundaries and figure/ground labels are highly predictive of the surface estimates. To recover surface information, average confidence for each region, the event of T-junctions are used as cues. Then, using CRF model and segmentation from boundary likelihoods, occlusion boundaries are detected [8].

3.2 Object Proposal Generation

The quality of object proposal pool directly impacts the performance of proposal selection-based co-segmentation method. Moreover, the proposed multi-search strategy even puts forward higher requirements on it. The measurement of the standard of proposal pool usually contains two aspects, the diversity and the representativeness. The diversity means the proposal pool should cover as many objects as possible, and the representativeness means the proposal pool should contain as few candidates as possible for each object. Encouraging the diversity of proposal pool is aimed at avoiding the omission of the interesting targets. And the representativeness is more required in multi-search strategy, whose line up is to avoid repeatedly extracting the same target as well as to reduce the storage and computation burden. To obtain the object proposal pool for each image, [4] is adopted. In this category-independent object proposal generation method, after a large number of proposals are achieved, a scoring mechanism that combines appearance features and overlap penalty is raised for proposal ranking. Finally, the proposals corresponding to different potential objects with accurate object segmentations will be top-ranked, which highly aligns with the above-mentioned requirements for object proposal pool.

3.3 Weighted Graph Construction

The similarity between two proposals is defined as,

$$S(p_i^u, p_j^v) = \alpha . S(h_{i,c}^u, h_{j,c}^v) + (1 - \alpha) S(h_{i,s}^u, h_{j,s}^v) \tag{1}$$

where $h^u_{i,c}$ and $h^u_{i,s}$ are the normalized colour and bag of histogram (BoF) [9] of proposal p^u_i, respectively.The histogram similarity is defined as follows,

$$S(h_p, h_q) = \sum_{b=1}^{\text{length}(h)} \min(h_p(b), h_q(b)) \tag{2}$$

$\alpha(0 \le \alpha \le 1)$ is the weight coefficient for the two histogram similarity terms. For different image groups. Therefore, the expected feature weights for these two image groups are greatly different. Taking the idea from [10] and design an iterative weights setting mechanism for the features α is firstly initialized as 0.5 to give the two features equal importance. In order to add adaptivity and flexibility, the idea from [10] is taken and design an iterative weights setting mechanism for the features in Eq. (2). The iterative algorithm is summarized in Algorithm 1. α is firstly initialized as 0.5 to give the two features equal importance. And then, according to Eq. (2), can construct a completely connected graph by linking every two proposals with their edge being weighted by their similarity. By solving the sub-problem from [3] on a loopy belief propagation, achieve the initial label assignments x. We then update α by maximizing the following objective function with the initial proposal labels x,

$$\alpha' = \arg \max_{0 \le \alpha \le 1} \sum_{x^u_i = x^v_j = 1} W(x^u_i, x^v_j, \alpha) - \lambda.\sigma^2(x, \alpha) \tag{3}$$

where λ is the weight and is empirically set to 100 in experiments, $\sigma^2(x, \alpha)$ is the variance of the similarities between the currently selected proposals, which can be calculated as

$$\sigma^2(x, \alpha) = \sum_{x^u_i = x^v_j = 1} (W(x^u_i, x^v_j, \alpha) - \bar{W})^2 \tag{4}$$

$\bar{W}$ is the mean value of all the similarities between those selected proposals. The intuitive intention of designing such an object function is to encourage the selected common targets to be globally consistent while keeping a low variance to make the similarity metric more reasonable and representative. After achieving the new feature coefficient α, reconstruct the weighted graph and retrieve new label assignment. Repeat such iterative loop until α does not change any more. Such an adaptive feature weight selecting method greatly extends the application range of co-segmentation approach and makes users get rid of the tedious manual feature selecting procedure [3].

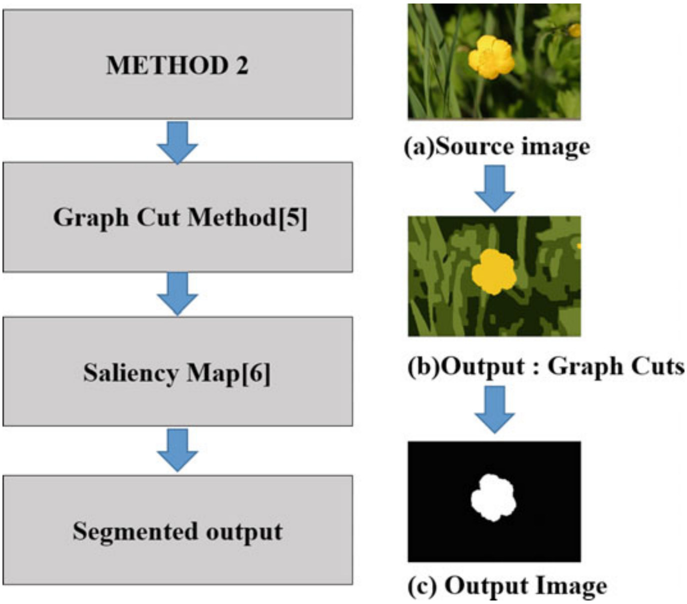


Fig. 2 Block diagram of co-segmentation method 2

3.4 Common Target Multi-search Algorithm

Initialize, label assignment x, feature weight coefficient α from weighted graph construction algorithm.Before every search, remove the discovered proposals in previous loops from the target candidate set P_i. Multiple common targets of each image are individually extracted using a adaptive label threshold T which is the weight of edge of proposal pair.

4 Co-segmentation Method 2

4.1 Graph Cut Method

Let $G = \langle N, E \rangle$ be a weighted graph, where N the set of vertices (nodes) and E the set of edges. N contains a node for each pixel in the image, and two additional nodes are named as terminals. One is the source, and next one is named as sink. There is an edge $e_{u,v}$ among any two well-defined vertices u and v. Mainly, there are two categories of edges. The first of these kind of edges is called n-links which connect the neighbouring pixels within the image, and the next kind of edge is the t-links which is used to connect the terminal nodes with the neighbourhood nodes (Fig. 2).

In this kind of graph, every one edge is allocated with a non-negative weight, named as cost A cut $X \subset E$ is a set of edges verifying:

- terminals are separated in the graph $G(X) = \langle N, E/X \rangle$
- no subset of X separates terminals in $G(X)$

The removal of cut separates the terminals into two induced subgraphs, and this means set of edges covers a cut. Furthermore, this cut is minimal in the sense that none of its subsets distinguish the terminals into the same two subgraphs. The minimum cut hurdle incorporates the discovery of a cut X, in a given graph, with the lowest cost. The cost of a cut, denoted $|X|$, amounts the sum of its edge weights. By fixing the weights of graph G suitably, one can make use of swap moves from combinatorial optimization to evaluate efficiently minimum cost cuts parallel to a local minimum of functional. For each couple of labels, swap moves note the minimum cut in the subgraph [6].

4.2 Saliency Maps

An image that conveys each pixel's unique quality is a saliency map. Goal is to make a different delineation of an image into relatively more easier to scrutinize. After performing Graph Cut , nodes are assumed as superpixels. Nodes are assumed as superpixels. Each node is represented as superpixel by SLIC algorithm [11]. The image becomes a close-loop graph, where superpixels as nodes. Based on affinity matrices, similarity to background and foreground queries are ranked from these nodes. Affinity matrix naturally captures spatial relationship information.

$$f^* = (D - \alpha K)^{-1} y \tag{5}$$

Let $y_u(0) = 1$ if u is a query point, else $y_u(0) = 0$:

- Connect the two nearest points iteratively until a connected graph $G(N, E)$ is obtained where E the set of edges.
- Form the affinity matrix K defined by $k_{uv} = e^{\|x_u - x_v\|^2/2\sigma^2}$ if an edge determined $e(u, v) \in E$ or else $k_{uv} = 0$
- Compute the matrix D, the diagonal matrix with $D_{uu} = \sum_{v=1}^{m} k_{uv}$. As an important cue for saliency detection, inverse matrix $f^* = (D - \alpha K)^{-1} y$ can be regarded as a complete affinity matrix.

The complication with graph labelling is the results with sensitivity of favoured seeds obtained. All the background and foreground seeds should be effortlessly generated via background priors and grading background queries (or seeds). By local clumping cues extricated from the entire image, the proposed algorithm [7] gives rise to clearly well-defined boundaries of salient objects and uniformly apotheosis the entire salient regions.

5 Experiments and Results

Experiments are done on MSRC and iCoseg dataset. The dataset iCoseg instigated in [12] accommodates assorted cluster of images. In [12], the authors handed down the iCoseg dataset in an synergistic co-segmentation substructure, and to the optimum of consciousness, this dataset was never worn in a extensively automatic environs. Each clump contains images of indistinguishable objects from the same grade or same object exemplar. iCoseg is a formidable dataset because the objects are distorted, there are some differences in terms of frame of reference and illumination, and in some exemplar, only visibility is lump of the object. To test co-segmentation systems, this contradicts significantly with the images used. Some images are excluded images in some of the groups to make it feasible. The results of different dataset of co-segmentation method 1 with weight coefficient 0.5 as constant and the α of the second scheme as 0.99 are shown in Figs. 3, 4, 5, 6 and 7.

The results are manifested in Tables 1 and 2 of both the methods. Table 1 consists of the comparison of time taken by two methods for different datasets and threshold of method 1 in common target multi-search algorithm, and Table 2 shows the comparison of two schemes with their precision and recall. Feature extraction is more in method 1. Proposed method (method 2) outstands performance in running time. The processor used is Intel(R) Core (TM) i5-8250UCPU, 1.60 GHZ, 1.80 GHZ, 8 GB RAM. Bar graph of comparison between running time of method 1 and 2 is shown in Fig. 8. With the variation of α (weight coefficient, 0.5), the threshold value of common target multi-search algorithm changes in method 1. With the variation of α

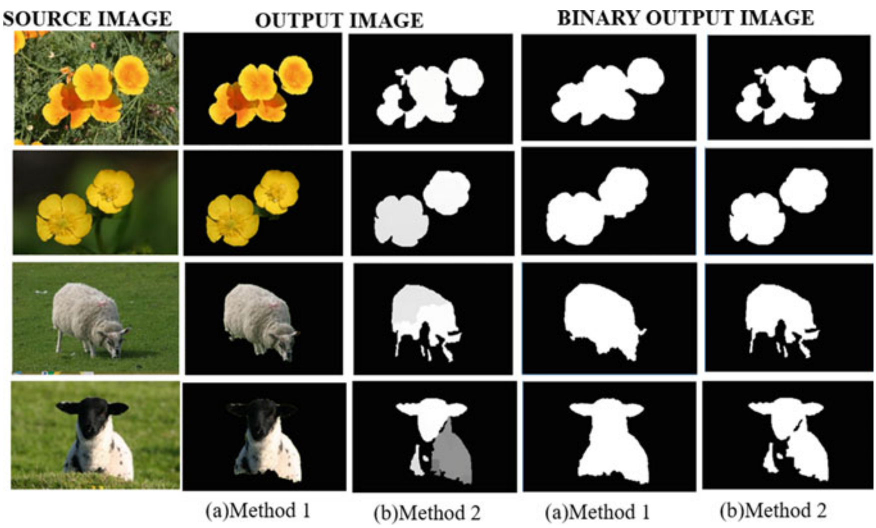


Fig. 3 Dataset: sheep (MSRC), flower (MSRC), method 1: [$\alpha = 0.5$, Threshold = 1.2, 1.18], method 2: [$k = 4$, $\alpha = 0.99$]

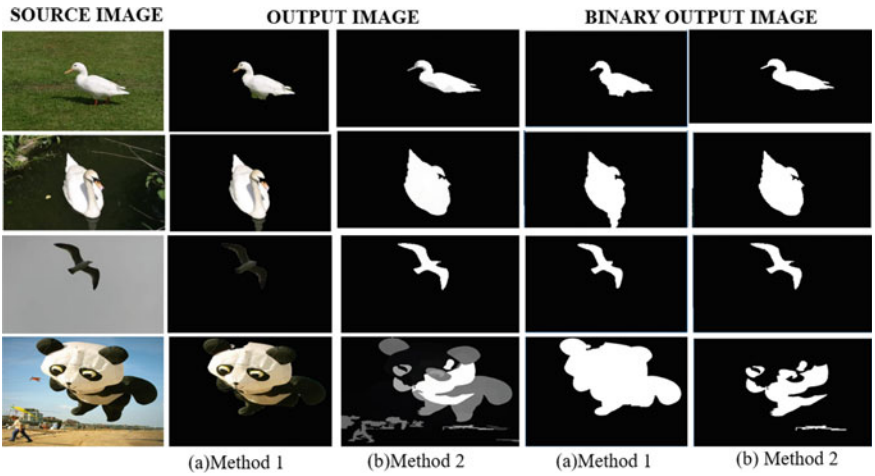


Fig. 4 Dataset: bird (MSRC), panda (iCoseg), method 1: [$\alpha = 0.5$, Threshold = 1.29, 1.6], method 2: [$k = 4$, $\alpha = 0.99$]

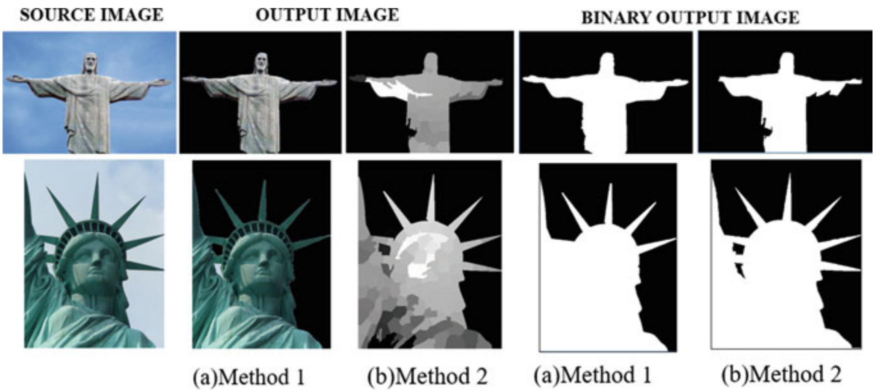


Fig. 5 Dataset: christ (iCoseg), statue (iCoseg), method 1: [$\alpha = 0.5$, Threshold = 1.9, 1.7], method 2: [$k = 4$, $\alpha = 0.99$]

(control the balance of two items in manifold ranking cost function), the precision and recall change of clumps of images from the datasets are plotted in Figs. 9 and 10. Out of ten different clumps of images, the precision and recall come best when α value is 0.99.

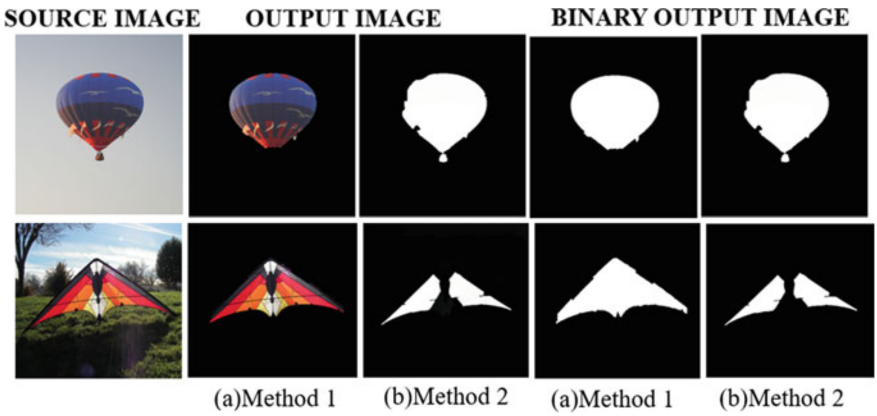


Fig. 6 Dataset: balloon (iCoseg), kite (iCoseg), method 1: [$\alpha = 0.5$, Threshold = 0.92, 1.54], method 2: [$k = 4$, $\alpha = 0.99$]

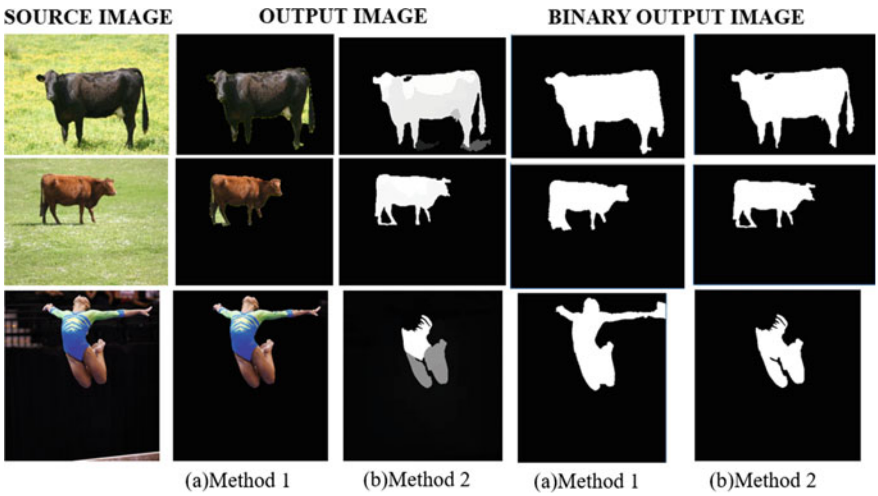


Fig. 7 Dataset: cow (MSRC), gymnastics (iCoseg), method 1: [$\alpha = 0.5$, Threshold = 1.16, 0.42], method 2: [$k = 4$, $\alpha = 0.99$]

6 Conclusion

This is a comparison work between an existing and proposed methods for co-segmentation of images. Here, results of formidable dataset are shown. The dataset used has the objects present in large deviations of viewpoint, scale and illumination. Method 1 is a co-segmentation approach for extracting common targets by proposal selection-based methods. Method 2, which is the proposed one, is a co-segmentation

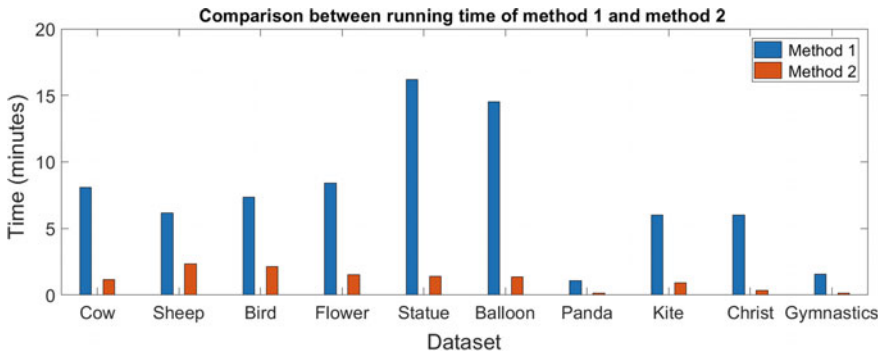


Fig. 8 Bar graph of running time between method 1 and method 2 of different datasets

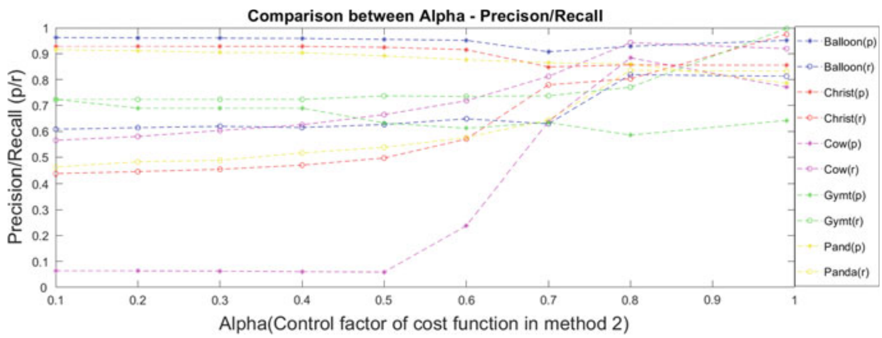


Fig. 9 Plot 1 between α and precision/recall of method 2

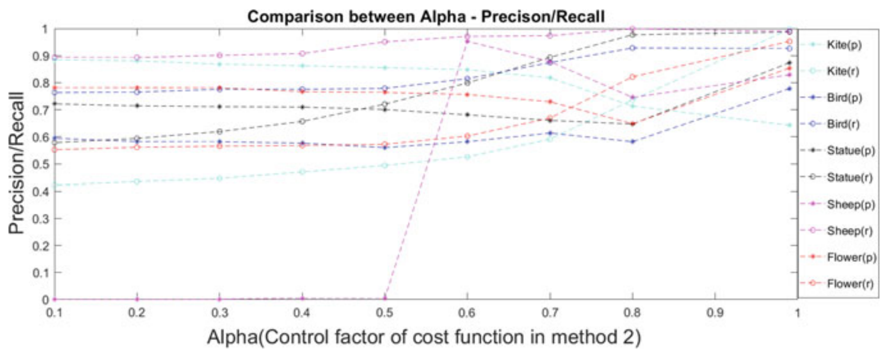


Fig. 10 Plot 2 between α and precision/recall of method 2

approach which is a combination of graph cut and saliency maps. Proposed method outstands performance in running time.

As future work,will try to apply this system to applications such as object localization for assistive robotics, image data mining.

Table 1 Threshold value of method 1 and time taken by co-segmentation method 1 and method 2

S. No.	Dataset [no: of images]	Time (min)		Threshold of method 1
		Method 1	Method 2	
MSRC				
1	Cow [16]	8.08	1.18	1.1615
2	Sheep [11]	6.18	2.36	1.2240
3	Bird [17]	7.35	2.14	1.2941
4	Flower [11]	8.41	1.51	1.1808
iCoseg				
5	Statue [16]	16.19	1.41	1.7164
6	Balloon [16]	14.54	1.35	.9190
7	Panda [3]	1.08	0.16	1.6602
8	Kite [3]	6.02	0.9	1.5400
9	Christ [7]	6.01	.33	1.9027
10	Gymnastics [3]	1.57	0.14	0.4285

Table 2 Comparison of precision and recall of co-segmentation method 1 and method 2

S. No.	Dataset [no: of images]	Precision		Recall	
		Method 1	Method 2	Method 1	Method 2
MSRC					
1	Cow [16]	0.8919	0.7710	0.9575	0.9195
2	Sheep [11]	0.9223	0.8290	0.9863	0.9900
3	Bird [17]	0.8482	0.7780	0.9095	0.9265
4	Flower [11]	0.8604	0.8542	0.9663	0.9529
iCoseg					
5	Statue [16]	0.8885	0.8737	0.9215	0.9878
6	Balloon [16]	0.9676	0.9514	0.9771	0.8127
7	Panda [3]	0.8287	0.7861	0.9751	0.8325
8	Kite [3]	0.7503	0.5205	0.9959	0.9961
9	Christ [7]	0.9854	0.8550	0.9378	0.9750
10	Gymnastics [3]	0.8333	0.6424	0.9684	0.9973

References

1. Daryanto T, Arif S, Yang S (2017) Survey: recent trends and techniques in image co-segmentation challenges, issues and its applications. IJCSSE 6(5):99
2. Vicente S, Rother C, Kolmogorov V (2011) Object cosegmentation. In Proceedings of IEEE conference on computer vision and pattern recognition, Providence, RI, USA, pp 2217–2224
3. Li K, Zhang J, Tao W (2016) Unsupervised co-segmentation for indefinite number of common foreground objects. IEEE Trans Image Process 25(4):1898–1909
4. Endres I, Hoiem D (2014) Category-independent object proposals with diverse ranking. IEEE Trans Pattern Anal Mach Intell 36(2):222–234
5. Meng F, Li H, Liu G, Ngan KN (2012) Object co-segmentation based on shortest path algorithm and saliency model. IEEE Trans Multimedia 14(5):1429–1441
6. Ayed IB (2011) Multiregion image segmentation by parametric kernel graph cuts. IEEE Trans Image Process 20(2):545–557
7. Yang C, Zhang L, Lu H, Ruan X, Yang M-H (2013) Saliency detection via graph-based manifold ranking. In: IEEE conference on computer vision and pattern recognition
8. Hoiem D, Efros AA, Hebert M (2011) Recovering occlusion boundaries from an image. Int J Comput Vis 91(3):328–346
9. Fei-Fei L, Perona P (2005) A bayesian hierarchical model for learning natural scene categories. In: Proceedings of IEEE conference on computer vision and pattern recognition, vol 2, pp 524–531
10. Meng F, Li H, Ngan KN, Zeng L, Wu Q (2013) Feature adaptive co-segmentation by complexity awareness. IEEE Trans Image Process 22(12):4809–4824
11. Achanta R, Smith K, Lucchi A, Fua P, Susstrunk S (2010) Slic superpixels, Technical report, EPFL. Tech. Rep. 149300
12. Batra D, Kowdle A, Parikh D, Luo J, Chen J (2010) iCoseg: interactive co-segmentation with intelligent scribble guidance. In: CVPR
13. Ding J-J, Lin C-J, Lu I-F, Cheng Y-H (201) Real-time interactive image segmentation using improved superpixels. IEEE, pp 740–744
14. Joulin A, Bach F, Ponce J (2012) Multi-class cosegmentation. In Proceedings of ieee conference on computer vision and pattern recognition, Providence, RI, USA, pp 542–549
15. Zhang Y, Li X, Gao X, Zhang C (2016) A simple algorithm of superpixel segmentation with boundary constraint. IEEE Trans Circ Syst Video Technol 27:1502–1514

A Novel Strategy for Data Transmission in Aerospace Vehicle Using Space-Based TTC Network—Telemetry via INSAT

Sherly Joy, K. S. Smitha, Mini Sreekumar, D. Sheba Elizabeth, S. Sanoj Kumar Roy, and K. K. Mukundan

Abstract In space missions, geographically separated TTC stations support data reception between launch vehicle and earth in real time. GSO-based INSAT segment can act as telemetry data relay platform between the host vehicle and the ground station. It has better coverage for supporting TTC requirement. Telemetry via INSAT (TIST) is used to overcome the visibility constraints of conventional TTC ground stations, ship-borne terminals and air-borne terminals, especially in the descent phase. TIST covers events in such scenarios without any hindrance by transmitting data to GSO satellite and receiving it back at ISRO earth stations. It employs digital modulation with channel coding for establishing the link from the launch vehicle to satellite. The satellite transponder upconverts the signal and transmits the information in real time to ISRO ground station. The system was tested in the RLV-TD mission of ISRO and transmitted the critical vehicle parameters in real time in order to evaluate the performance of the vehicle. This scheme is the first of its kind in ISRO launch vehicle, where communication is established between a moving launch vehicle and a single satellite in GSO orbit.

Keywords GSO · TDV · Telemetry · TIST · DRT

1 Introduction

Telemetry data transmission in launch vehicles is an automated communication process where data related to vehicle health or scientific data is collected and transmitted to the receiving station for analysis. Hitherto, the Indian Space Research Organization (ISRO) employs ground-based systems for providing telemetry support for launch vehicle programme. For missions like Human in Space Programme (HSP), more than forty ground stations are required across the globe for 100% coverage. This

S. Joy (✉) · K. S. Smitha · M. Sreekumar · D. Sheba Elizabeth · S. Sanoj Kumar Roy · K. K. Mukundan
Vikram Sarabhai Space Centre, ISRO, Thiruvananthapuram, India
e-mail: sherly_joy@vssc.gov.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_20

approach has its own limitations in terms of line of sight constraints, shorter visibility period, station handing over protocol complexity and more number of stations with the associated installation and maintenance cost.

Space-based systems are used as relay platform for communicating between the launch vehicle and ground station [1]. This is done by transmitting data to satellite in geostationary orbit (GSO) and this in turn is retransmitted to ground stations which are in line of sight with GSO. Three geostationary satellites can provide coverage over the entire earth [1]. A GSO-based system is composed of three segments: Technology Demonstrator Vehicle (TDV), space and ground. A relevant example in this regard is NASA's Tracking and Data Relay Satellite System (TDRSS) [2, 3] which provides TTC support not only to LEO and MEO satellites but also to expendable launch vehicles. Presently, India is not having a dedicated network like TDRSS. However, with the available transponder in GSO (INSAT) satellite, communication is realized to cater to the non-visibility constraints in the ground-based system.

The conventional telemetry system in launch vehicle employs analog (phase) modulation as it offers carrier acquisition even at lower CNR. However, in reentry missions like RLV, in near touchdown regimes, non-visibility exists, especially when the elevation angle of receiving ground station antenna is less than 2°. To overcome this situation, a communication link via a GSO is considered. Digital modulation is employed as it offers better immunity to noise, less occupied bandwidth and security. Hence, digital modulation overcomes the bandwidth limitation in the satellite transponder. As the data rate planned is in the order of kbps, BPSK modulation is chosen rather than higher modulation schemes like QPSK and OQPSK which are usually employed at higher data rates in Mbps, to inherit the advantage of bandwidth efficiency. Also BPSK has 3 dB CNR advantage compared to QPSK/OQPSK.

This paper describes the RF communication system—Telemetry via INSAT (TIST), which provides communication link between TDV and ground station using the GSO-based platform. TDV is India's maiden reusable launch vehicle (RLV), a winged body configuration which has been developed to prove some of the critical technologies of RLV concept. TIST utilizes the existing data relay transponder in INSAT-3A and INSAT-3D for transmitting the critical telemetry data to ground station, especially in the splashdown phase where visibility constraints from ground and ship-borne stations prevail. The first section describes frequency selection criteria, link design, link visibility study and end-to-end system details. The second section deals with the system hardware design of TDV. The paper concludes with a detailed discussion of flight performance and future scope.

2 Frequency Selection Criteria

Frequency selection is one of the major criteria in finalizing the communication link. It depends on the availability of transponder, bandwidth allotment and the spacecraft antenna coverage. The transponders in INSAT operate in UHF, S-band, C-band and Ku-band. Receiver footprints of spacecraft antenna are based on their service requirements. Transponder in UHF band has global coverage whereas S-band covers India and surrounding ocean regions. C-band transponders have different coverage for different missions and Ku-band has limited Indian coverage. For transmitting data from the vehicle via INSAT, it is necessary to consider the following aspects.

- Transmitter frequency should result in minimum power from the TDV per bit of data.
- TDV antenna with hemispheric radiation pattern or with steerable pattern which is in LOS with GSO DRT antenna provides maximum coverage.
- The frequency used should be compatible with ITU allocations and should not cause any interference to the spacecraft in geostationary orbit.

The data relay link performance is mainly limited by the uplink C/N_o which is dependent on TDV transmit power and frequency, the gain of the TDV and INSAT antennas. At higher frequencies, for a given antenna aperture size, onboard the spacecraft, the increased path loss is compensated by the increased antenna gain. But for a lower frequency, for the same aperture, higher beam width is available, which can cover most of the events of the TDV. This can be represented in the form of Eq. (1), assuming all other terms remain the same.

$$C/N_o = \text{EIRP} - L_p + G/T \tag{1}$$

where

C/N_o Carrier-to-noise density ratio
EIRP Effective Isotropic Radiated Power
L_p Path loss
G/T Receiver gain-to-noise temperature ratio of the spacecraft receiver

For a given C/N_o requirement, the minimum power from the transmitter is obtained where the sum of path loss and G/T is minimum. Path loss is considered for 43,000 km which is for zero degree elevation in GSO orbit. For various frequency bands in INSAT system, the different parameters are listed in Table 1.

For lower data rate, UHF provides optimum performance and the existing DRT is capable of supporting this requirement. The DRT system in the INSAT can accept signal in the frequency range of 402.65–402.85 MHz. Total bandwidth available in this transponder is 200, 10 kHz per carrier. The signal received by the DRT from the TDV is upconverted and is transmitted back to earth in C-band. C-band ground station receives the signal, demodulates the data and distributes to the destination by

Table 1 Parameters for INSAT system

Frequency (MHz)	Path loss (dB)	G/T (dB/K)	Total loss (dB)
UHF (402.75)	−177.0	−19.0	−196.0
S-band (2680)	−192.0	−7.5	−197.5
C-band (6400)	−198.6	−5.0	−204.0

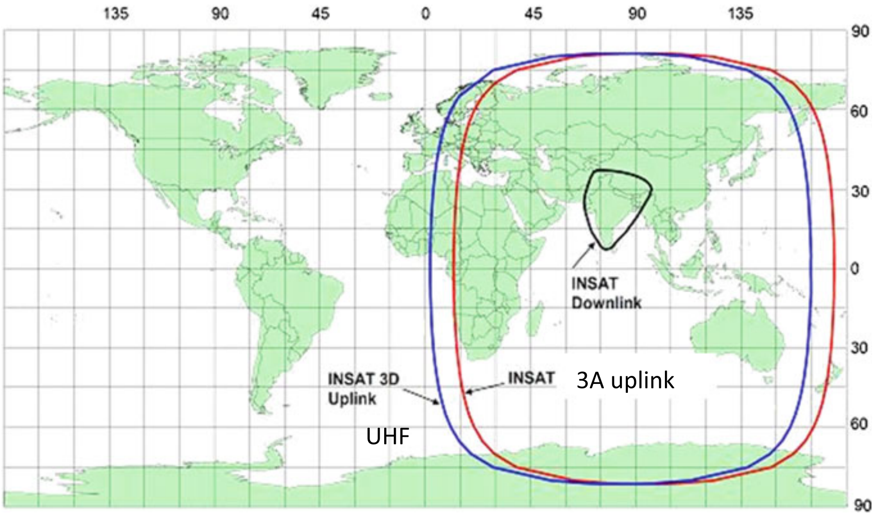


Fig. 1 Coverage of INSAT DRT transponder

conventional data links. Figure 1 shows the coverage area for the UHF uplink and C-band downlink carriers.

3 Link Budget and Visibility Studies

Link calculation is carried out based on the parameters of the TDV, INSAT transponder and ground segment. With 10 kHz bandwidth per slot in the satellite payload, only low bit rate of data transmission is possible. TDV's transmitter output is selected based on the frequency of operation, satellite *G*/*T*, DC power requirement, thermal management and package footprint availability. To get maximum time visibility to INSAT during the launch, hemispherical pattern is desired for the TDV antenna which makes a compromise between the gain and beam width. Turbo coding at the baseband level is selected to minimize the EIRP requirement and BPSK modulation at RF level to accommodate maximum possible data rate in the allotted bandwidth. Digital modulation technique of BPSK has advantage in C/N_0 and E_b/N_0 requirement for low bit rates. The important parameters of the communication system are summarized in Table 2.

For launch vehicle telemetry application, bit error rate (BER) allowed is 1×10^{-6} for which E_b/N_0 requirement is 10.6 dB. The link budget is shown in Table 3.

TIST antenna system consisting of four radiating elements (microstrip patch) is mounted on the leeward side of the vehicle such that it is facing to the INSAT for the major duration of the flight, especially in the landing phase. Visibility of TIST antenna from INSAT 3A/3D is computed using Systems Tool Kit (STK) software

Table 2 System parameters

Parameter	Value
INSAT to ground station	
Tx frequency	4505 MHz
Satellite EIRP	+39 dBm
Rx system G/T	27 dB/K
TDV to INSAT	
Tx frequency	402.73 MHz
Power output	10 W
Tx antenna gain	+4 dBi, 30° Beam width
Rx system G/T	−19 dB/K

Table 3 Typical link budget

Overall C/N_0	45 dB Hz	Coding gain	7 dB
Data rate	1 kbps	E_b/N_0	19 dB
Channel symbol rate	2 kbps, 1/2 rate coding	Link margin	8.4 dB

for the full trajectory of RLV-TD considering its attitude. The visibility is computed for the half cone angle which is the angle between the vector pointing upward in the antenna direction and the vector from RLV-TD to INSAT satellite. It is noticed that throughout the ascent and descent phases, the TIST antenna direction changes due to vehicle pitching. The time period for which half cone angle less than the 3 dB beam width (<30°) is considered as the visibility period. Figures 2 and 3 show the geometry of half cone angle between TIST antenna and INSAT and its value in RLV-TD mission, respectively. The difference in visibility is due to the location of the

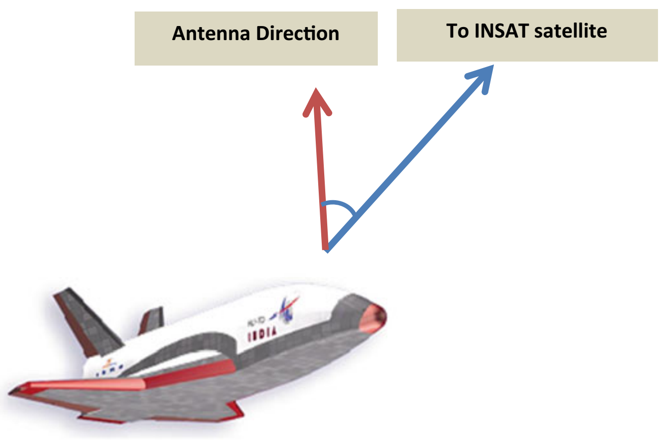


Fig. 2 Geometry of half cone angle between TIST antenna and INSAT

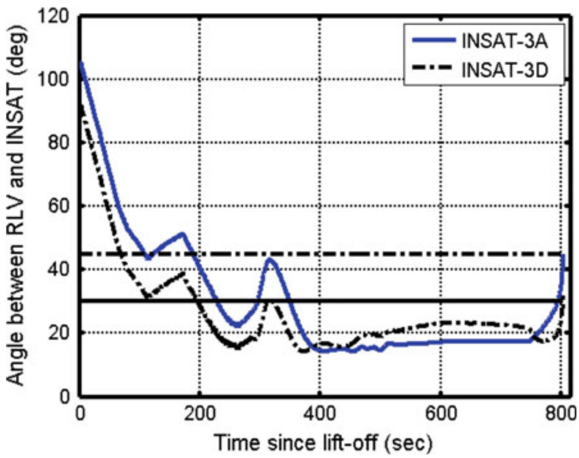


Fig. 3 Angle between TIST antenna and INSAT for RLV mission

satellites. In link calculation, slant range of 43,000 km is considered (for 36,000 km orbit) in order to take care the additional path loss. The limiting factor is the onboard and satellite antennas look angle.

4 Total System Description and Design

The total system is divided into three segments: INSAT, ground and TDV segment. The simplified block diagram is shown in Fig. 4.

4.1 INSAT Segment

Presently, data relay transponder (DRT) in the INSAT operating in UHF band is incorporated for real-time hydro-meteorological data collection from unattended platforms located at land and river basins. The data is then relayed in extended C-band to ground station. INSAT-3A and INSAT-3D satellites located at 93.50° E and 82.00° E, respectively, are having DRT. Data collection system (DCS) and satellite-aided search and rescue (SAS and R) system are two different services combined in one payload. Frequency range allotted for DRT system is 402.65–402.85 MHz, whereas downlink frequency is in extended C-band (4505 MHz). Subsystems like receiver, transmitter antennas, upconverter, local oscillator and SSPA are shared by both systems. The short backfire antenna receives both signals which are oppositely polarized, i.e. DRT receiving antenna with left-hand circular polarization and SAS and R with right-hand circular polarization.

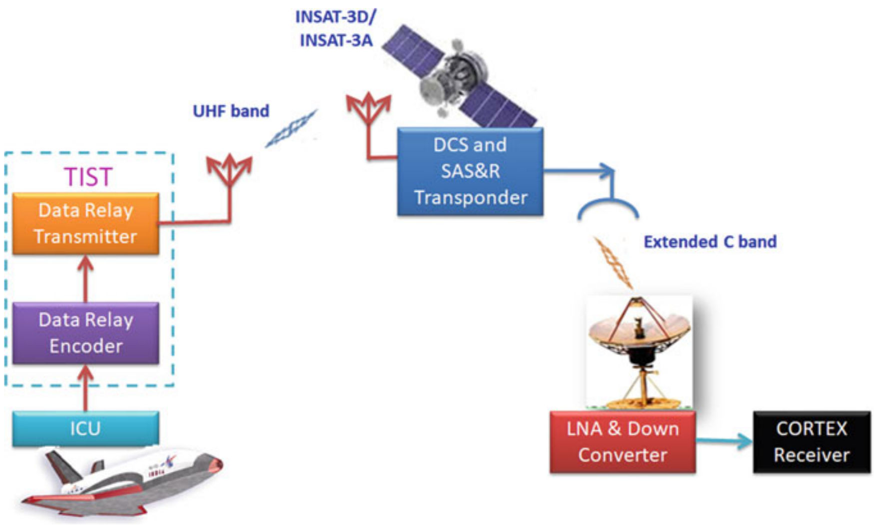


Fig. 4 Simplified block diagram

4.2 Ground Segment

The C-band ground stations at MCF, Hassan/ ISTRAC TTC stations can receive the downlink signal from INSAT. The signal received in extended C-band is downconverted to 70 MHz and is given to the BPSK receiver. This is an integrated command, ranging and telemetry system for satellite ground stations. It is capable to demodulate wide data BPSK/QPSK signal and has inbuilt CCSDS Turbo, Viterbi and RS decoders. This system also provides soft decision decoding.

4.3 TDV Segment

The TDV segment consists of the UHF transmitter modulated with turbo-coded data from the baseband system. The output is transmitted to INSAT using suitable antenna with necessary beam width and gain whose parameters are listed in Table 2. The antenna is properly located to give line of sight visibility to INSAT satellite during the required period of mission.

Design: TIST electronic system is assembled in a single box (four decks) consisting of low power transmitter (LPT) module, high power amplifier (HPA) module, baseband module and power supply (PS) module. Figure 5 shows the block diagram of the total system.

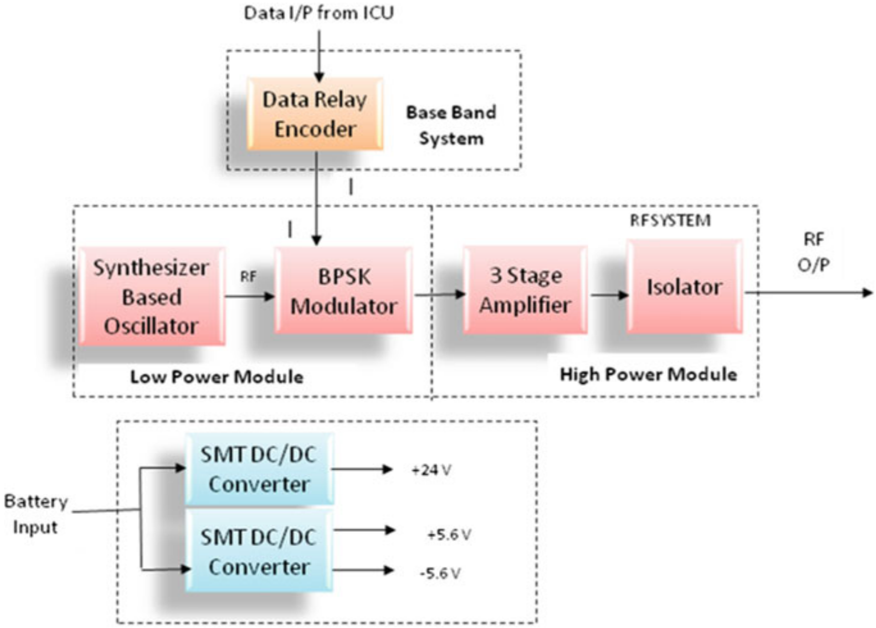


Fig. 5 Total system block diagram

4.4 Baseband System

In TIST, the essential data like TDV latitude, longitude, altitude, angle of attack and bank angle from the main telemetry unit is selected. Data relay encoder (DRE) receives and decodes PCM Bi-phase data from the main telemetry unit and selects specified data for coding. The selected data is turbo coded as per CCSDS recommendation and given to the BPSK transmitter. Coding gain improves the link margin.

4.5 BPSK Transmitter

Bandwidth available in the INSAT transponder is only 10 kHz, and to utilize the maximum bandwidth for data, the transmitter carrier is realized with temperature compensated crystal oscillator as basic reference oscillator. The improved performance requirements such as frequency stability over temperature and ageing are also considered while selecting the oscillator to make the translation errors minimum. The carrier is realized using fractional synthesizer technique with Hz range resolution so that if any interference is observed for the frequency of operation, it can be easily switched to another frequency. Adjacent channel interference in DRT

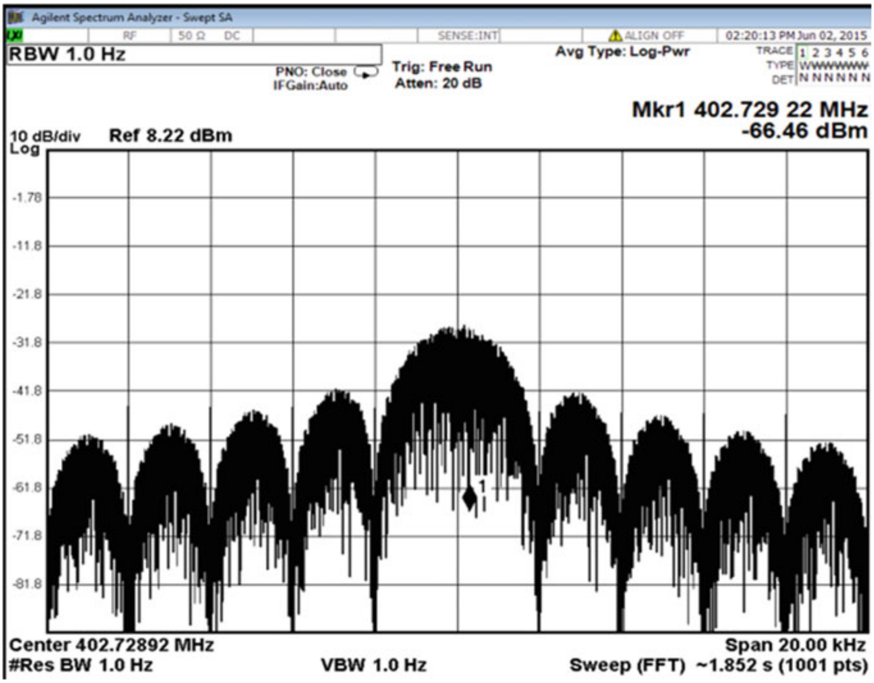


Fig. 6 BPSK spectrum of TIST

spectrum is a problem as it is crowded. Filtering is impractical at the transmitter frequency because the quality factor of the filter must be very high. To limit the transmitting spectrum, prior to modulation, the bit sequence is passed through a linear phase four-pole filter with cutoff frequency at the symbol rate [4].

Digital modulation techniques increase the hardware complexity but use the limited frequency spectrum more efficiently by allowing more users in the given bandwidth. The BPSK modulation is carried out using double-balanced mixer [4]. The output is filtered and linearly amplified using MMIC amplifiers in the low power stages and enhancement mode MOSFETs in the final stages. Frequency spectrum of the BPSK modulated carrier for 1 kHz NRZ data with ½ rate turbo coding is shown in Fig. 6.

4.6 Power Supply

Total system is powered using two in-house developed SMT DC/DC converters, which takes care of power supply variation from 26 to 36 V DC with inbuilt EMI filter complying electromagnetic interference as per MIL-STD-461C. Phase noise of the basic carrier is very important in the lower bit rate condition (1 kHz). The carrier generator circuits are powered using low dropout regulators (LDOs). In order to get better efficiency in the power amplifier stage, the devices are operated with a 24 V high efficiency (82%) converter [5]. Higher efficiency is achieved by turning on and off the gate driver MOSFET circuit much faster to reduce the switching loss.

4.7 Antenna System

The major challenge in the design is achieving medium gain (+4 dBi) with wider beam width. Also the antenna should not be a protruding type as per vehicle configuration and needs to be encapsulated with flexible external insulation (FEI) for thermal protection. The antenna is configured with left circular polarization (LCP). Four elements of microstrip patch array (2 × 2 configurations) are fed in equal amplitude and equal phase using a UHF 4-way power divider. Corner-chopped square microstrip patches realized on 60 mil thick, 2.94 dielectric constant RT Duroid 6002 substrate with 200 °C temperature capability is used. Tuning shift (3.4 MHz up) is absorbed in the design by considering the dielectric constant of FEI. Figure 7 shows the typical radiation pattern.

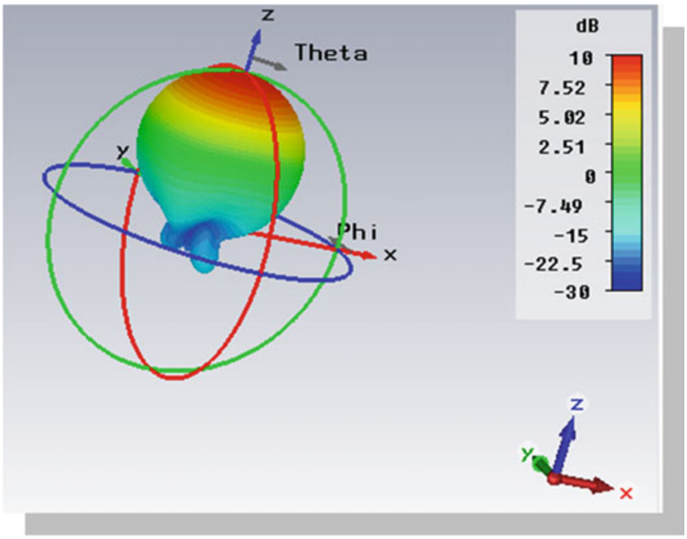


Fig. 7 3D radiation pattern of antenna

5 Performance in Flight

DRT transponders in two satellites: INSAT-3D and INSAT-3A are employed for data transmission from TDV to ground stations. Master Control Facility, Hassan (MCF)-1, MCF-2 and ISTRAC, Bangalore stations, received the TIST data. The communication system maintained carrier lock and provided telemetry data with E_b/N_0 greater than 11 dB for the intended duration. The power supply monitoring outputs, which were monitored at ground in real time through conventional telemetry link, show that the package was live till the end of mission. TIST has performed as expected, and all the assigned parameters have been received at ground stations.

5.1 Future Scope

The use of UHF can cover more events of the mission with minimum transmitter power but limits the data rate due to allocation of lower bandwidth for such services. For higher data rates (Mbps), it is preferred to go for higher frequency range like Ka-band, but the onboard antenna on INSAT as well as on the TDV are to be made to track each other. This requires the use of electronically steerable phased array antennas for TDV and spacecraft.

6 Conclusion

TIST was tested in RLV-TD mission in addition to the conventional telemetry system. The test results were compared with the existing system and found to be normal. TIST offered communication via satellite and provided data till touchdown. The advantages offered by a GSO satellite in providing the TTC support to launch vehicles in terms of wider visibility at lesser cost and complexity, compared to conventional multiple ground stations, have been successfully experimented in RLV-TD. For lower data rate transmission, UHF frequency provides optimum performance. If complete vehicle data is required to be transmitted from TDV in flight, it is necessary to go to higher frequency band with steerable phased arrays both in TDV and spacecraft.

Acknowledgements The authors thank Dr. K. Sivan, Chairman, ISRO, for his special interest in the development of TIST. They express gratitude to Shri. S. Somanath, Director, VSSC, for his relentless support in all activities. Also they thank team MCF and team ISTRAC for their constant support.

References

1. Maral G, Bousquet M Satellite communications systems—systems, techniques and technology, 5th edn. Wiley, Hoboken
2. Baohua K, Ligang F, Li M, Yuheng L, Xiaoshen X, Liyu C (2015) TDRSS resource scheduling modes and service modeling based on mission planning. In: 2015 IEEE advanced information technology, electronic and automation control conference (IAEAC), Chongqing
3. Dandan F, Benjin L, Gang FL, Shenghua G, Xiaohan J (2015) Design of global emergency communication system based on geosynchronous TDRSS. In: 2015 international conference on wireless communications & signal processing (WCSP), Nanjing
4. Haykin S (1988) Digital communications. Wiley, Hoboken
5. Erickson RW, Maksimovic D (2001) Fundamentals of power electronics, 2nd edn

Sherly Joy joined VSSC/ISRO in 1982. She obtained B.Tech in Electronics and Communication from Kerala University. Presently she is the Division Head of RFSD. She is involved in the design and development of RF systems.

K. S. Smitha joined VSSC/ISRO after obtaining M.Tech from Anna University in 2004. She is engaged in the design and development of RF systems for ISRO space programmes.

Mini Sreekumar joined ISRO in 1988. She is involved in the design and development of RF Communication Systems related to TTC for ISRO space programmes.

D. Sheba Elizabeth joined ISRO with M.E. degree from Anna University. She is involved in the design & development of Baseband Telemetry Systems for ISRO launch vehicle programmes.

S. Sanoj Kumar Roy joined ISRO in 2005. He is involved in the design and development of DC. DC Converter for ISRO launch vehicle programmes.

K. K. Mukundan obtained B.Tech in Electronics Engineering from REC, Calicut. He has expertise in antenna systems for radars and launch vehicles. Presently he is the Group Director of RF Systems Group.

An Improved Image Inpainting Technique Using Fuzzy Hard C-Means Algorithm

R. F. Liji, P. Sreejaya, M. Sasikumar, and Kim J. Seelan

Abstract A novel algorithm which can repair video sequence deprived of any artefacts that are present in numerous such prevailing technique is proposed in this paper. Specifically, a video inpainting is proposed that detects moving object in a background where water is falling from a fountain. The input video is converted to frames. Using the Canny edge detection, the edges of the objects in the frames are found out. These edges are classified into different clusters using fuzzy hard C-means algorithm so that inpainting can be done effectively. The classified edges are patch-matched with the most similar pixels. The resultant inpainted video when viewed gives a very good result with minimum time frame.

Keywords Canny edge · Clustering · Fuzzy hard C-means clustering · Image inpainting · Patch match

R. F. Liji (✉)
Department of Electronics and Communication Engineering, John Cox Memorial CSI Institute of Technology, Thiruvananthapuram, Kerala, India
e-mail: lijirf1216@gmail.com

P. Sreejaya
Department of Electrical and Electronics Engineering, College of Engineering, Thiruvananthapuram, Kerala, India
e-mail: sreejaya@ieee.org

M. Sasikumar
Department of Electronics and Communication Engineering, Marian Engineering College, Thiruvananthapuram, Kerala, India
e-mail: drmsasikumar@yahoo.com

K. J. Seelan
Department of Mechanical Engineering, John Cox Memorial CSI Institute of Technology, Thiruvananthapuram, Kerala, India
e-mail: kimseelan@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_21

1 Introduction

Video inpainting assists to fill in the lost portions or to substitute the undesired portions of the video sequence. The method has countless applications like surveillance of the surrounding, processing the video and editing of video in post-production in the movie industry, etc. One of the most important tasks in video inpainting is to mending a moving object under a still background or removing an object from a moving background. Digital image inpainting is done to remove or restore damaged images, basically. Recently, inpainting is done in a vast area of image processing like wireless image and video transmission, recovery of edges, segmentation zooming, special effects in video, visual psychology, etc. As an advancement of these, inpainting is applied in video compression at the encoder and decompression at the decoder end.

A good inpainting should minimize the loss of information in the original image [1]. Information and complexity are closely interrelated. In order to make things easier for a complex system, degree of uncertainty is used to explain the complex work. The statements obtained after this abridged arrangement are not as much precise or certain, but then again their significance to the unique system is completely preserved. Thus, the loss of information that arises during reduction of complexity of the system to a manageable level is expressed using uncertainty principle. Both complexity and information are connected with the uncertainty principle. In order to simplify real lifetime problems, various types of uncertainty principles are used. Fuzzy sets and fuzzy measures are two such uncertainty principles which can be used for characterization of mathematical formulation and investigations available.

In Sect. 2, we discuss the existing methods, followed by the proposed method in Sect. 3 and sample experimental results are discussed in Sect. 4. Section 5 concludes the paper with insight to future work.

2 Existing Methods

The term 'image inpainting' was coined by [2], and the author introduced simultaneous fill-in of numerous regions using lines that connect points of equal brightness called isophote in the selected regions. The inpainting algorithm was initially used to inpaint black-and-white images and a lot of algorithms were implemented in it. Then, inpainting technique was used in colour still images. Video images with constant background and non-moving objects in it were selected for inpainting in the next set of development of image inpainting. [3], discusses differnt existing inpainting techniques available now. [4], explains how inpainting can be effectively utilized for compression of images and video, which is an emerging technique. [5], introduced fuzzy C-means clustering for inpainting in color still images. The results from the paper [5] show great improvement when compared to basic models like CDD-based [6], TV-based [7], diffusion-based [8], MAP-based [9], exemplar-based [10, 11], etc. In [12], the author has introduced video inpainting. The paper uses motion state estimation in order to select a moving human and then remove the human from the frames of the selected video and has tried to inpaint the missing area with the background. But the output obtained after inpainting is not a clear one. Moreover, the video selected does not contain complex motion of human. The feature points were not selected automatically. The inpainted regions are not visually plausible in [12]. In [13] a self-adaptive inpainting method was proposed. The inpainting was done according to the different structural information of the object a corresponding, similar image blocks are selected to build self-adaptive groups. From this group, pixels are selected for inpainting. All the existing methods selected videos, which were less complicated with either still background or still object and moving background. The proposed method inpaints video with a moving object in front of moving background. For selecting the edges of the objects accurately, we have used the Canny edge detection along with fuzzy algorithm.

3 Proposed Method

The proposed video inpainting algorithm uses the Canny edge detection and fuzzy hard C-means (FHC-m) clustering. Algorithm for the proposed method is as given below:

1. Read the video file to be inpainted.
2. Covert to frames.
3. Detect the edges of object in each frame using the Canny edge detection.
4. Form the border of the detected edge.
5. Draw an occlusion around an object.
6. Calculate the image pyramid of an (colour) image volume and its occlusion volume.
7. Calculate the linear index/indices of the neighbourhood of a pixel *p*.
8. Retrieve a spatio-temporal pyramid of features for the purpose of video inpainting.
9. Use fuzzy hard C-means algorithm for clustering the detected object.
10. Remove the detected object.
11. Patch match in the inpainted region.
12. Interpolate the displacement field.

The above algorithm can be solved using following steps:

Step 1 Video input is read.
Step 2 The input video is converted to frames.
Step 3 Canny edge detection is done.
Step 4 Fuzzy hard C-means clustering is done.
Step 5 Patch match is done in the fuzzified output for smoothening

The Canny edge detector practices a many-stage procedure to discover an extensive variety of edges and boundaries in images or frames. The algorithm [1] can be used to get valuable structural data from diverse visual items, and it affectedly decreases the amount of data to be administered. In order to satisfy the criteria for edge detection, the algorithm has to identify as many edges shown in the image as possible. The optimum function in the Canny edge detector is defined by the totality of four exponential terms and estimated by the initial derivative of a Gaussian value. The process of the Canny edge detection algorithm is as follows:

1. Apply adaptive filtering in order to remove the noise and to smooth the image [14]. Using the gradient values $U_x(x, y)$ and $U_y(x, y)$, calculate

$$D(x, y) = \sqrt{U_x^2(x, y) + U_y^2(x, y)} \text{ and weight; } w(x, y) = \exp\left(-\frac{\sqrt{D(x, y)}}{2a^2}\right),$$

   where a is the amplitude of the edge.
   Then adaptive filtering gives

$$p(x, y) = \frac{1}{N} \sum_{i=-1}^{1} \cdot \sum_{j=-1}^{1} g(x + i, y + j) w(x + i, y + j);$$

   where $N = \sum_{i=-1}^{1} \cdot \sum_{j=-1}^{1} w(x + i, y + j)$.
2. As the second step we have to discover the intensity gradients of the frames [15]. In order to detect horizontal edges, vertical edges and diagonal edges four filters are used in this algorithm. The first derivative in the horizontal and vertical directions is U_X and U_Y, respectively.
3. Put on non-max suppression to develop a frame which is clear of spurious reactions of edge detection. It is an edge thinning technique.
4. A double threshold is applied to determine possible edges.
5. As a final step, only strong edges connected are selected.

After detection of the edges, FHC-m algorithm [16] is used to classify and cluster the detected edges of various objects in the video. The result of inpainting is, therefore, better than all the existing inpainting techniques. Experiments conducted show that relationships form recognizable or classifiable structure. By detecting the edges, the data are classified conferring to the similarity of patterns and characteristics. This is known as clustering. The iteration for clustering continues till a group of data that seems plausible in a structural and physical perspective.

There are hard (or crisp) and soft (or fuzzy) clustering [17]. The data used in the experiment is unlabelled data, so absolute measure of clustering validity does not exists.

The objects in the video frame are classified using FHC-m algorithm. Each data in the frames of video is mapped to a single data cluster only.

Table 1 A comparison of proposed method with existing methods

Parameters	Proposed method	Existing methods	
		[12]	[18]
Image type used for inpainting	Video with moving object in moving background	Video image with moving object in still background	Still image
Edge detection used	Canny edge	Motion state estimation	Exemplar-based
Clustering	Fuzzy hard c-means	Motion state classification	Nil
Restoration	Patch match	Motion state prediction	Fuzzy
Visual effect	High quality	Good	Better

Patch match is used to fill the border areas smoothly and effectively. The pixels from the neighbouring area are selected for the best patch. And thus, the steps for inpainting is completed.

Table 1 gives a comparison of the existing techniques with the proposed method, and here, we can see that the proposed method is giving a better output.

4 Experimental Results

Experiment is conducted using MATLAB 2016a version on a laptop with CPU specification Intel Core i5 7th Gen, 2.5 GHz and 4 GB RAM.

Four frames of the input original video are shown in Fig. 1. Figure 1a and d is the first and last frames of the input video, respectively. Figure 1b, c is the in between frames of the input original video where we can see the position of the human has changed. Figure 2 shows four different frames corresponding to the borderline detection using FHC-m algorithm. Figure 3 shows the different frames after inpainting is done. It took 33 s to remove the object and to inpaint the 2.06 MB video. The video selected for inpainting is human walking in front of a water fountain. The object removed is the human, in our experiment. The video selected is a complicated one. As water is flowing from the fountain, the edges of flowing water, the edges of the fountain and the edges of moving human are identified. A borderline of the moving human is identified. The detected edges are grouped into different clusters according to their similarity features. Then, according to the best match, the pixels are inpainted.

Another video example taken has three jet ski bikes with people and one water surfing person. Here, using Fuzzy Hard C- Means, two jet ski bikes are identified and they are removed. The portions left out after removing are filled with neighbouring pixels, which is water using the FHCM inpainting technique. It is a great task to remove and inpaint the area with pixels of water, but with fuzzy hard c-means algorithm it is done. Figures 4, 5 and 6 show the input video frames, the occlusion

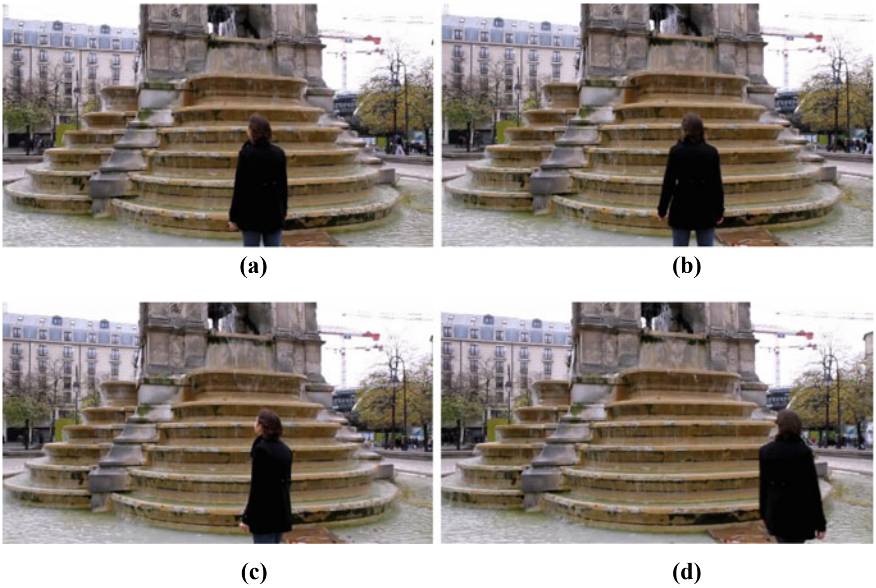


Fig. 1 **a** The first frame, **b** and **c** show the inbetween frames of video while the human is moving and **d** last frames of the input video selected for inpainting

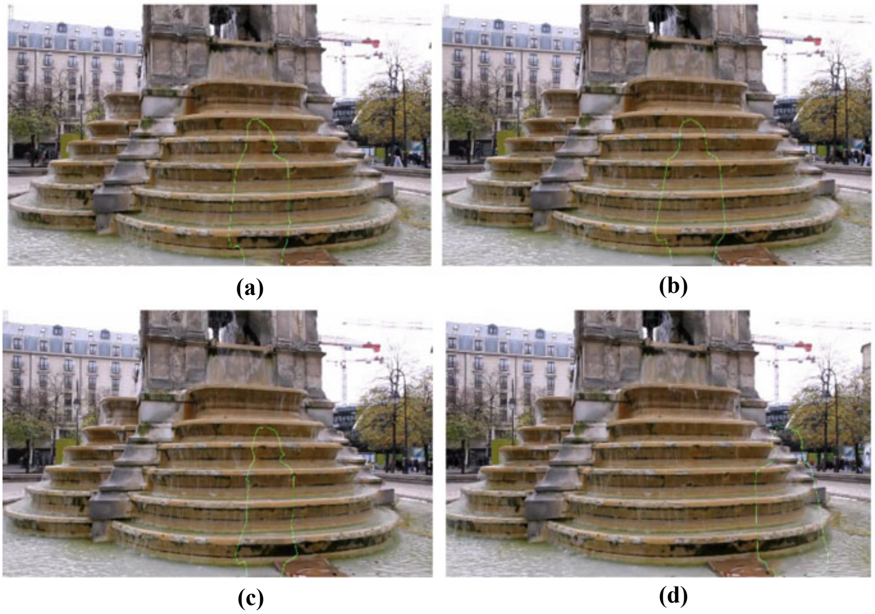


Fig. 2 **a** and **d** The first and last frame of the moving human's border detected and finalized using FHC-m algorithm. **b** and **c** to the borderlines of human being with respect to Fig. 2b and c

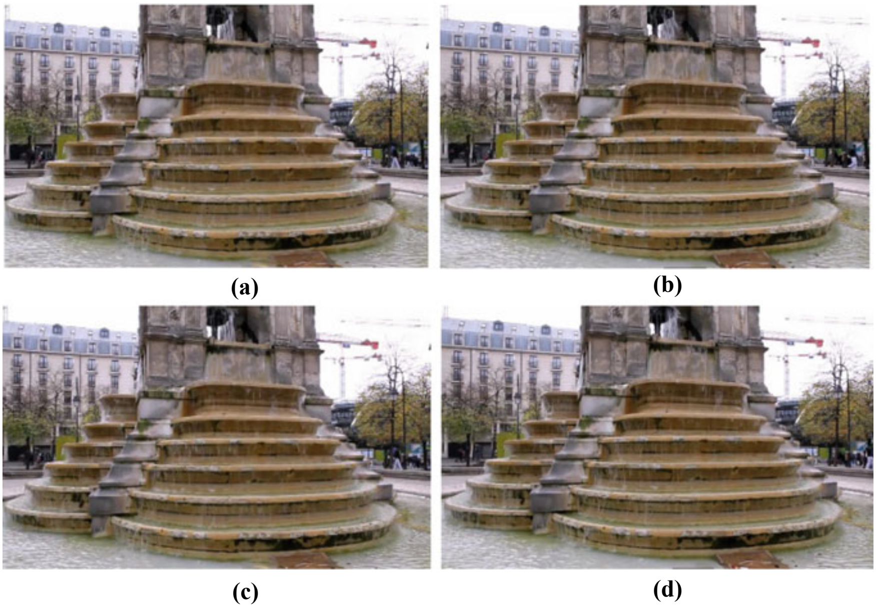


Fig. 3 **a**–**d** The frames after inpainting is applied to the video in order to remove the moving human

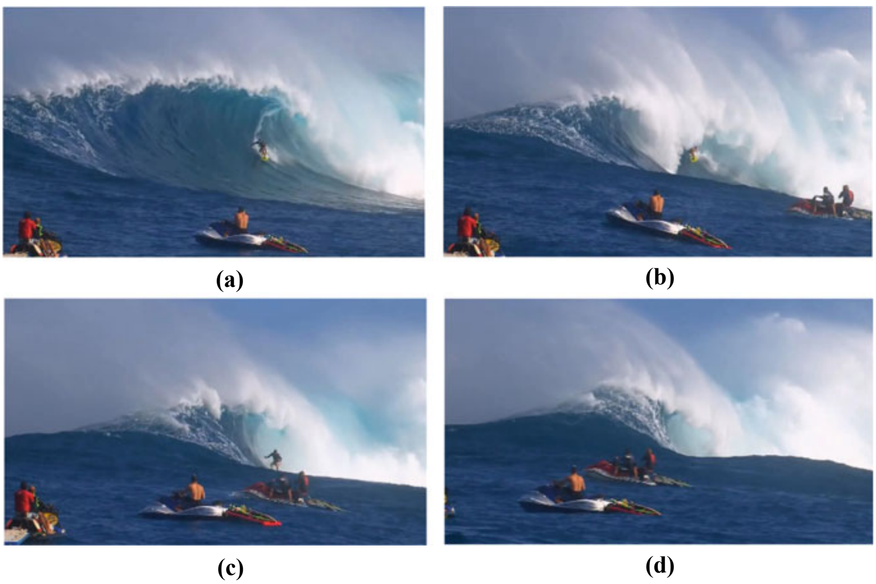
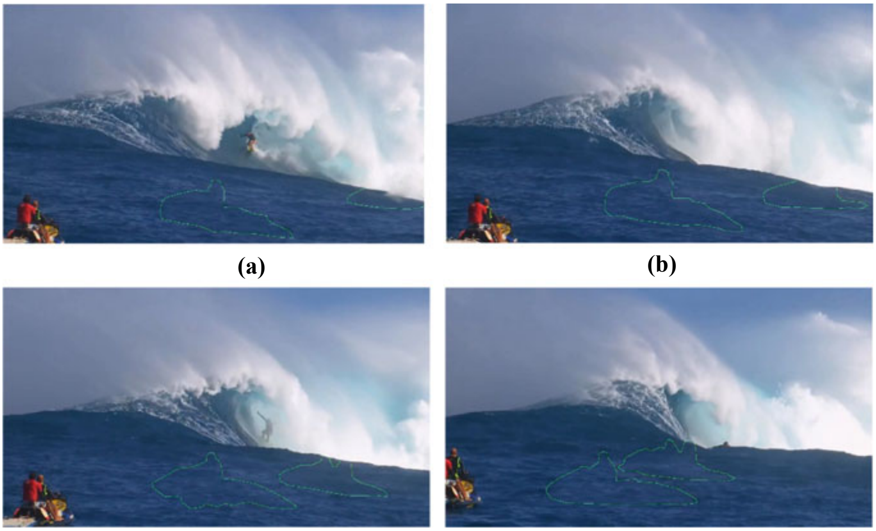
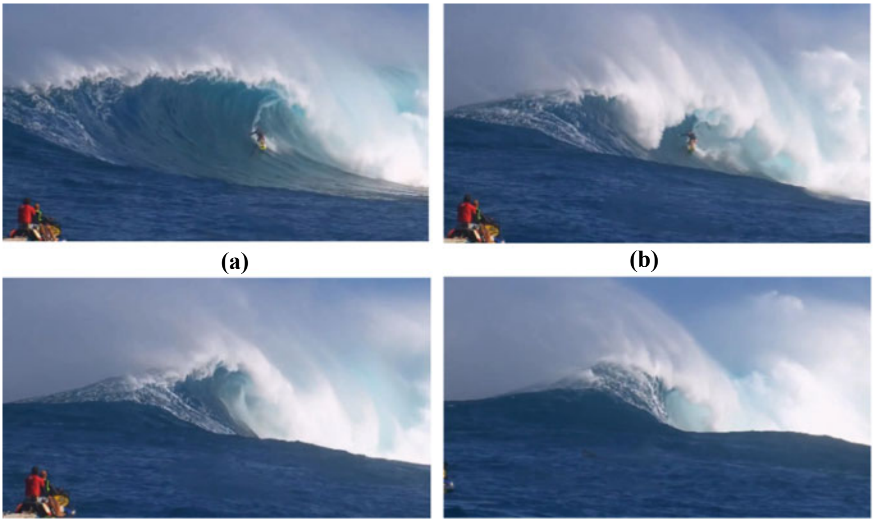


Fig. 4 **a**–**d** The frames from the input video, where at first only two jet ski bikes are seen along with the water surfer. And in the last frame shown has two jet ski bikes



(a) (b) (c) (d)

Fig. 5 **a**–**d** The frames from the borderline video, where at first only two jet ski bikes are seen along with the water surfer. And in the last frame shown has two jet ski bikes



(a) (b) (c) (d)

Fig. 6 **a**–**d** The frames from the ouput video, where the two jet ski bikes are removed from the video and inpainting is done with the help of neighbouring pixel with fuzzy hard c-means algorithm

Table 2 Frames and their properties

Figure	# of frame	Dimension	Bit depth	Size (kB)
Figure 1a	1	960 × 544	24	88.7
Figure 1b	22	960 × 544	24	87.3
Figure 1c	62	960 × 544	24	88.0
Figure 1d	104	960 × 544	24	86.1
Figure 2a	1	960 × 544	24	92.1
Figure 2b	35	960 × 544	24	92.9
Figure 2c	111	960 × 544	24	92.8
Figure 2d	207	960 × 544	24	91.0
Figure 3a	1	960 × 544	24	90.4
Figure 3b	22	960 × 544	24	91.0
Figure 3c	110	960 × 544	24	91.1
Figure 3d	207	960 × 544	24	89.7

selection frames and the inpainted frames taken from the videos.

Table 2 summarizes the details of the some of the frames obtained from our experiment. From the table, we can understand that there is no compromise in the size and dimension of the frame even after the inpainting is done. We can see that the size has improved after inpainting is done on the applied video frame. Therefore, the clarity of the video has improved after inpainting. The bit depth also has not changed.

5 Conclusion and Future Work

This paper proposes a novel fuzzy hard C-means algorithm-based inpainting for object removal, missing block completion in digital video with moving background and moving object. This is inspired from the recent progress of the research in the fields of image inpainting and fuzzy logic.

Digital video inpainting implemented using our algorithm gives a high-quality video output after inpainting. The video selected for inpainting is with lot of information to segregate. Information from the video are segregated using FHC-m clustering and better patch match is done. When the inpainted video of our work is compared to the already existing works [5, 12], it shows a superior quality of inpainted output and it gives a high visually plausible result. It is also seen that the inpainted video size has increased showing that there is an increase in pixel size and resolution and thereby improvement in the image quality.

In future, we will further investigate the application of deep learning in inpainting of digital video with moving background and multiple moving objects.

References

1. Klir GJ, Foger TA Fuzzy sets, uncertainity, and information. PHI Learning Private Limited, New Delhi
2. Bertalmio M, Sapiro G, Caselles V, Ballester C (2000) Image inpainting. In: Proceedings of SIGGRAPH, pp 417–424
3. Liji RF, Sasikumar M (2018) An exploration of image inpainting techniques. Int J Eng Adv Technol 8(2):135–138
4. Liji RF, Sasikumar M, Sreejaya P, Seelan KJ (2019) Image inpainting in the field of image compression: an investigation. In: 1st international conference on recent scientific research in engineering and technology, pp 155–159
5. Zhang L, Kang B, Li X, Zhang D (2015) Initialization of image inpainting area using fuzzy c-means clustering. J Inf Comput Sci 3129–3135
6. Xu Z, Lian X, Feng L (2008) Image inpainting algorithm based on partial differential equation. In: ISECS international colloquium on computing, communication, control and management, pp 120–124
7. Zhao Z, Ye X (2010) A new image multi-level inpainting method. In: IEEE 3rd international conference on advanced computer theory and engineering, pp V3-384–V3-388
8. Lin C-S, Leou J-J (2011) Image inpainting using multiscale salient structure propagation. In: International conference on multimedia and signal processing, pp 201–204
9. Biradar RL, Kohir VV (2013) A novel image inpainting technique based on median diffusion. Sadhana 38(4):621–644
10. Qin Y, Wang F (2010) A curvature constraint exemplar-based image inpainting. IEEE
11. Liji RF, Sasikumar M, Sreejaya P, Seelan KJ (2019) A comparative study a comparative study and analysis of Lattice Boltzmann method and exemplar method for still color image inpainting technique. In: ICICICT
12. Wang H, Li H, Li B (2007) Video inpainting for largely occluded moving human. In: ICME, pp 1719–1722
13. Chen L, Wu J (2018) Image inpainting algorithm based on self-adaptive structural group sparse representation. In: 13th IEEE conference on industrial electronics and applications
14. Deriche R (1987) Using Canny's criteria to derive a recursively implemented optimal edge detector. Int J Computer Vision 1:167–187
15. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8(6):679–698
16. Zhou P, Ye W, Wang Q (2011) An improved Canny algorithm for edge detection. J Comput Inf Syst 7(5):1516–1523
17. Ross TJ Fuzzy logic with engineering applications, 2nd edn. Wiley, New York, pp 362–391
18. Cislariu M, Gordan M, Vlaicu A (2011) A Fuzzy set generalization of the exemplar-based image inpainting. ACTA Tehnica NAPOCENNIS 52(2):54–59

An Improved EEG Acquisition Protocol Facilitates Localized Neural Activation

Jerrin Thomas Panachakel, Nandagopal Netrakanti Vinayak, Maanvi Nunna, Angarai Ganesan Ramakrishnan, and Kanishka Sharma

Abstract This work proposes improvements in the electroencephalogram (EEG) recording protocols for motor imagery through the introduction of actual motor movement and/or somatosensory cues. The results obtained demonstrate the advantage of requiring the subjects to perform motor actions following the trials of imagery. By introducing motor actions in the protocol, the subjects are able to perform actual motor planning, rather than just visualizing the motor movement, thus greatly improving the ease with which the motor movements can be imagined. This study also probes the added advantage of administering somatosensory cues in the subject, as opposed to the conventional auditory/visual cues. These changes in the protocol show promise in terms of the aptness of the spatial filters obtained by the data, on the application of the well-known common spatial pattern (CSP) algorithms. The regions highlighted by the spatial filters are more localized and consistent across the subjects when the protocol is augmented with somatosensory stimuli. Hence, we suggest that this may prove to be a better EEG acquisition protocol for detecting brain activation in response to intended motor commands in (clinically) paralyzed/locked-in patients.

Keywords EEG · Motor imagery · CSP · Protocol · Somatosensory cues · Motor planning

J. T. Panachakel (✉) · A. G. Ramakrishnan · K. Sharma
Department of Electrical Engineering, Indian Institute of Science, Bangalore, India
e-mail: jerrinp@iisc.ac.in

A. G. Ramakrishnan
e-mail: agr@iisc.ac.in

K. Sharma
e-mail: kanishka@iisc.ac.in

N. N. Vinayak · M. Nunna
Department of Computer Science and Engineering, PES University, Bangalore, India
e-mail: nandagopalnv@pesu.pes.edu

M. Nunna
e-mail: maanvinunna@pesu.pes.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_22

1 Introduction

Electroencephalogram (EEG) remains vastly popular for recording the activity of the brain due to its relatively low cost, high time resolution and easy availability in hospitals and research institutions. Recently, researchers have used EEG as a tool for understanding the neural response to motor commands in clinically unresponsive patients [7], giving impetus to the research on protocols under which EEG data is acquired. The problem we address in this work is the design of an effective protocol that improves the quality of the signal recorded from the subjects. This work proposes two approaches—use of somatosensory cues, and the enforcement of a protocol in which motor imagery of a specific limb is followed by the actual movement of the limb, in contrast to just imagining the movement, which has been used conventionally.

2 Methods

2.1 EEG Recording Setup

The EEG data was sampled at 1000 Hz from an ANT Neuro Eego™ Mylab amplifier using the EEGCA64-500 montage and the 10/10 electrode placement system with "CPz" electrode as reference electrode. A 64-channel cap with an electrooculogram (EOG) channel was used for acquisition. EOG channel was not used for independent component analysis (ICA) artifact removal, since all the participants were instructed to keep their eyes closed throughout the experiment, except when they were under rest. The subjects were seated on a wooden chair, in a well-ventilated room. The subjects were also instructed to rest their feet on wooden support to ensure that there was no electrical contact with the floor.

2.2 Subjects for the Study

A total of seven healthy subjects participated in the experiments (five males and two females, all right handed, mean age of 22.5 years, $\sigma = 3.4$). Some of the participants took part in more than one protocol. The protocol was approved by the Institute Human Ethics Committee of the Indian Institute of Science, Bangalore. The participants signed an informed consent form before taking part in the experiments. Subjects were labeled 1 through 7, and protocols were labeled A through E. Thus, a label "B4", for example, refers to the subject 4 under protocol B.

Fig. 1 Screenshot of the web app used for timing the auditory cues. The web app was used to improve the consistency of trial durations

2.3 Web Application for Controlled Timing

For the purpose of maintaining consistent intervals of time for each of the trials of the experiment and to administer cues with as less human error as possible, a web application was developed. The application was programmed to display event labels—“LEFT”, “RIGHT”, “REST” and “START”. To prevent the subjects from predicting the sequence of events, the presenting of one of the labels “LEFT” and “RIGHT” was randomized, with “REST”/“START” necessarily following the event label. The duration for each event was also specifically coded-in, depending on the protocol in effect, to ensure that the trials were consistent in duration throughout the experiment. The web app was used by one of the experiment coordinators in-charge of delivering cues (auditory/somatosensory) as per the protocol in effect, as detailed in Sect. 3. A pause button was also made available to facilitate the subjects to choose the number of successive trials before taking rest (Fig. 1).

2.4 Preprocessing of the EEG Data

The acquired EEG was preprocessed using EEGLAB [8]. The 50 Hz line noise was removed using notch filter applied from 49 to 51 Hz and then, the data was bandpass filtered from 8 to 30 Hz. This was performed, since the changes in EEG signals due to motor imagery are more visible in mu and beta bands [14, 16]. From the literature, it is known that motor imagery activation in mu and beta bands is visible in the central (Rolandic) region of the brain, primarily consisting of “C3”,“Cz” and “C4” electrodes. When the arm is moved (or imagined to have moved), mu rhythm changes contralateral to that arm movement (imagination) [12]. The frequency band of Rolandic beta rhythms varies from subject to subject, while falling in the broad range of 14–30 Hz [11].

2.5 Algorithm to Identify the Primary Region of Activation

It is known from the existing literature that for motor imagery, a minimum of eight channels are needed to obtain optimal performance [18]. However, increasing the number of channels introduces the curse of dimensionality. Thus, we use a spatial filtering algorithm that combines several channels into one, using weighted linear combinations, through which features are extracted.

The common spatial pattern (CSP) algorithm linearly transforms the multi-channel EEG signal into a low-dimensional subspace such that the variance of the EEG signal from one class (a group of channels) is maximized while that from the other class (the group of remaining channels) is minimized [10]. Mathematically, the CSP algorithm extremizes the following objective function:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T X_1 X_1^T \mathbf{w}}{\mathbf{w}^T X_2 X_2^T \mathbf{w}} = \frac{\mathbf{w}^T C_1 \mathbf{w}}{\mathbf{w}^T C_2 \mathbf{w}} \tag{1}$$

where T denotes the matrix transpose, X_i is the matrix containing the EEG signals of class i, with data samples as columns and channels as rows, $\mathbf{w}$ is the spatial filter, and C_i is the spatial covariance matrix of class i.

CSP finds spatial filters such that the variance of the filtered data is maximal for one class while simultaneously minimal for the other class. This effectively highlights those regions of the brain that contribute most to a particular action (imagery/actual movement) by comparing it to the activation as obtained during the other action. CSP is especially effective in brain–computer interface (BCI) applications that use oscillatory data like EEG, since their most important features are band power features. The CSP algorithm is computationally efficient and easy to implement. However, CSP does have some limitations. It is not robust to noise or non-stationarities and may not generalize well to new data, especially when there is very less data available. Since we are mainly interested in the spatial regions of the brain that are associated with the imagery in consideration, we decided to use the filters produced by the Lagrangian CSP algorithm and implemented it using Fabien Lotte's RCSP-Toolbox for MATLAB [13].

3 Experimental Protocols Explored

The scope of this work was to design a protocol that would return consistent results in terms of the area of activation as expected based on previous work carried out on similar motor imagery tasks. In the literature, all the reported works use either audio or visual cues to the subjects, following which they were expected to imagine the motor action (motor imagery) for the cue given [4, 5, 15]. The most popular datasets in use in this domain are the BCI competition datasets [2]. In the BCI competition III dataset, the subjects had to imagine either closing the right arm into a fist or wiggling the toes on the right foot. To validate our work, we tested our processing pipeline on the standard datasets available, and the obtained results were comparable to those obtained by Lotte and Guan [13].

The audio and somatosensory cues as detailed in Table 1 were delivered by one of the experiment conductors seated directly in front of the subject, using the web app for timing. After a set of 25–30 trials, the web app was paused in order to let the subject take rest for approximately 2–3 min. Depending on their fatigue level, the subject had a choice to request for less number of trials to be conducted for the next batch of trials.

Table 1 List of cues and the corresponding actions

Type	Cue	Task	Protocol(s)
Auditory	ARM	Imagine closing your right arm into a fist	A, B
Auditory	FOOT	Imagine wiggling the toes on your right foot	A, B
Auditory	REST	Take rest	A, B
Auditory	LEFT	Imagine closing your left arm into a fist	C
Auditory	RIGHT	Imagine closing your right arm into a fist	C
Auditory	START	Perform actual action of previously imagined motor action	B, C
Somatosensory	Tap on the LEFT inner wrist	Imagine closing your left arm into a fist	D
Somatosensory	Tap on the RIGHT inner wrist	Imagine closing your right arm into a fist	D
Somatosensory	Tap on the LEFT/RIGHT knee	Take rest	D
Somatosensory	Poke LEFT inner wrist with the blunt side of a pen	Imagine closing your left arm into a fist	E
Somatosensory	Poke RIGHT inner wrist with the blunt side of a pen	Imagine closing your right arm into a fist	E
Somatosensory	Poke LEFT palm with the blunt side of a pen	Close your left arm into a fist (actual action)	E
Somatosensory	Poke RIGHT palm with the blunt side of a pen	Close your right arm into a fist (actual action)	E

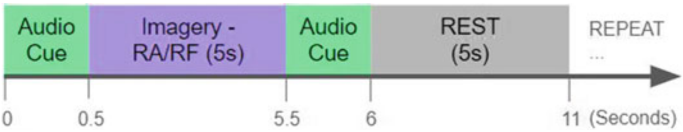


Fig. 2 Graphical representation of the timings of protocol A. An auditory cue lasting less than 0.5 s was given to ask the subject to start imagining the motor movement. After a period of 4 s, another auditory cue was given asking the subject to stop the imagination

3.1 Protocol A: RA-RF with Auditory Cues

To begin with, the first three subjects (A1, A2, A3) were made to follow the protocol as described in BCI competition III dataset IVa's description as closely as possible. This was to serve as a baseline for the work covered in this paper. The protocol included three types of events—imagery task of right arm (RA) or right foot (RF), and a rest event. Subjects were given rest for 5 s after every imagery task, which itself lasted for 5 s. The order of the imagery task was randomized to ensure that the subjects were unable to predict the sequence of cues. The graphical representation of the protocol timings is shown in Fig. 2. The feedback obtained from the subjects on this protocol was that it was often difficult to imagine the motor action without actually executing it and even harder to focus on the imagery for a duration as long as 5 s. The subjects described their imagination of the action merely as a visual image of the action, which does not strictly correspond to the motor imagery/planning that we were looking for. Taking this feedback into account, changes were made to the protocol, giving rise to newer protocols explained below.

3.2 Protocol B: RA-RF with Actual Motion and Auditory Cues

EEG data of two subjects, B4 and B5, was recorded under this protocol. This protocol has four types of events—two imagery tasks of right arm (RA) and right foot (RF), two actual motion tasks of right arm and right foot. Subjects were given auditory cues "arm" and "foot" to perform the corresponding imagery tasks repeatedly for 4 s. Followed by each imagery task, an actual action of the previously imagined task was asked to be performed before 3 s. The auditory cue for the actual action was "start". The graphical representation of the protocol timings is shown in Fig. 3.

The changes introduced in this protocol were in consideration of the feedback obtained from subjects who participated in protocol A. As seen in Sect. 4.1.2, the introduction of actual movement after the imagery task does produce better results.

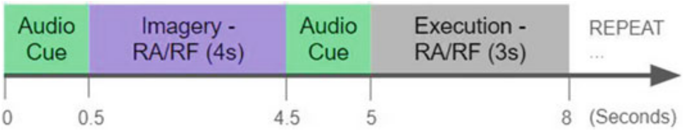


Fig. 3 Graphical representation of the timings of protocol B. An auditory cue lasting less than 0.5 s was given to ask the subject to start imagining the motor movement. After a period of 4 s, another auditory cue was given asking the subject to actually perform the action once

3.3 Protocol C: LA-RA with Actual Motion and Auditory Cues

Using this protocol, five subjects' (C1, C3, C4, C5, C6) EEG data was recorded. This protocol has four types of events—two imagery tasks of left arm (LA) and right arm (RA), two actual motion tasks of left arm and right arm. Subjects were given auditory cues "left" and "right" to perform the corresponding imagery tasks which lasted for 4 s. Followed by each imagery task, an actual action of the previously imagined task was asked to be performed once within 3 s. The auditory cue for the actual action was "start". Since we know that the sensory cortex coincides with the motor cortex in the central region of the brain, we thought introducing localized somatosensory cues will further improve our results and realized the next protocol. The graphical representation of the protocol timings is shown in Fig. 4.

3.4 Protocol D: LA-RA with Somatosensory Cues

One subject (D2) was recorded using this protocol. This protocol has three types of events—two imagery tasks of left arm (LA) and right arm (RA) and one rest event. Somatosensory cues were given on the subject's outer wrist in the form of a gentle tap to let the subject know that he/she should perform the imagery task of the corresponding hand. The imagination was asked to be carried out for 4 s, and a tap on the knee of the same side of the body as the previous cue was given to let the participant stop imagination and take rest. The graphical representation of

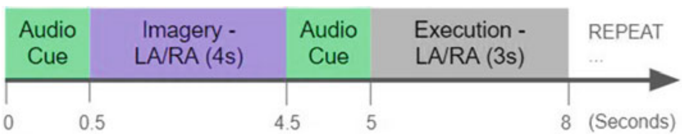


Fig. 4 Graphical representation of the timings of protocol C. An auditory cue lasting less than 0.5 s was given asking the subject to start imagining the motor movement. After a period of 5 s, another auditory cue was given asking the subject to actually perform the action once

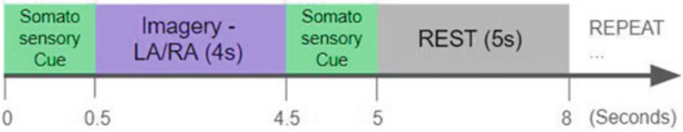


Fig. 5 Graphical representation of the timings of protocol D. Tap on the wrist was the somatosensory cue for starting the imagination, and tap on the knee was the cue for stopping the imagination and taking rest

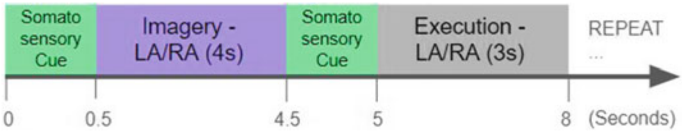


Fig. 6 Graphical representation of the timings of protocol E. Pressing the inner wrist by the blunt side of a pen was the somatosensory cue for starting the imagination and that of the palm was the cue for performing the actual action

the protocol timings is shown in Fig. 5. This protocol was to test if the introduction of somatosensory cues would affect the quality of results obtained. As observed in Sect. 4.2.1, there is a benefit to using somatosensory cues.

3.5 *Protocol E: LA-RA with Actual Motion and Somatosensory Cues*

EEG data of two subjects (E1, E6) was recorded using this protocol. This protocol has four event types—two imagery tasks of left arm (LA) and right arm (RA), two actual motion tasks of left arm and right arm. The cues that were administered to indicate the onset of imagination or task of actual motion were somatosensory in nature and given using the blunt side of a pen. Upon inducing somatosensory stimulus on the subject's inner wrist, the subject was asked to perform the corresponding imagery task, which lasted for 4 s. The somatosensory cue given on the palm told the subject to perform the actual action. The graphical representation of the protocol timings is shown in Fig. 6.

4 Results

We compare the performance of the tested protocols by visually inspecting the filters produced by the CSP algorithm. The criteria of evaluation are the correctness of regions highlighted in the filters, with reference to what we know about the expected activation of the brain for the motor imagery in consideration. It may be noted that the sign of the elements in the spatial filer vectors is not significant and only the absolute value matters. Also, all the topoplots were obtained using the EEG data corresponding to motor imagery alone, although some protocols had actual movements included in them. There are broadly two types of imagery in the protocols tested—arm and foot.

4.1 RA-RF Based Protocols

According to literature [13], the CSP filter plots for clenching of right arm into a fist when compared to wiggling of right toes on the foot (RA-RF protocol) indicate activation on the left hemisphere of the brain along electrodes "C3", "C5" and "CP3". The CSP filter plots for wiggling of right toes on the foot when compared to clenching of right arm are along the central electrode "Cz" (Fig. 7).

4.1.1 Protocol A: RA-RF with Auditory Cues

As shown in Fig. 8, the topoplots obtained highlight the regions as expected for the imagery pair RA-RF. However, the regions are not defined very distinctly. This was an observation common to all subjects who participated in this protocol. This could be because the subjects found it difficult to imagine motor action without actually performing the action. This is also likely affected by the fact that the subjects were unable to stay focused on the imagery for extended periods of time. Taking these into account, changes were made to the protocol.

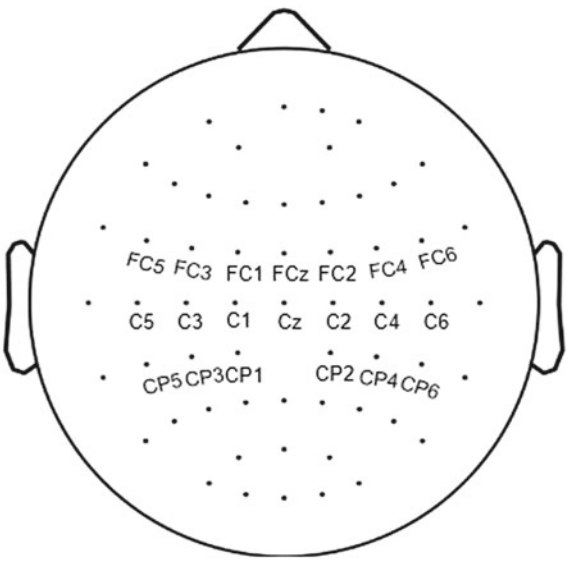


Fig. 7 10/10 EEG electrode placement system. The dots correspond to the positions of the electrodes on the scalp. The names of the relevant electrodes are also given

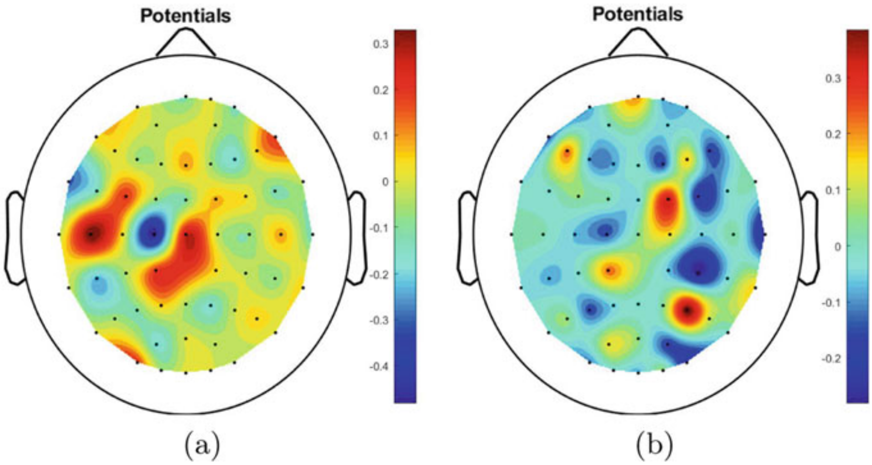


Fig. 8 CSP filter obtained for subject A3 for **a** right arm and **b** right foot motor imageries under protocol A. Under this protocol, three events were there, namely imagery of right arm movement, imagery of right foot movement and rest state

4.1.2 Protocol B: RA-RF with Auditory Cues, Followed by Actual Motion

As seen in Fig. 9, the introduction of actual motion into the protocol following every single trial of imagery did result in better results. The topoplots exhibit more definitive regions, and the subjects also reported that it was easier to stay focused through the trials, since they were instructed to perform the actual movement.

4.2 LA-RA-Based Protocols

For the clenching of a hand into a fist, we expect activation predominantly centered around the central line electrodes ("C5","C3", "C1", "Cz", "C2", "C4" and "C6"). This activation is also known to show contralateralization with respect to the left or right arm in imagery. In other words, for imagery related to the left arm, the right side of the brain is activated more and vice versa for the right hand. This property of lateral symmetry of activity for imagery of the left or right arm helps in evaluating visually the performance of the protocols in question.

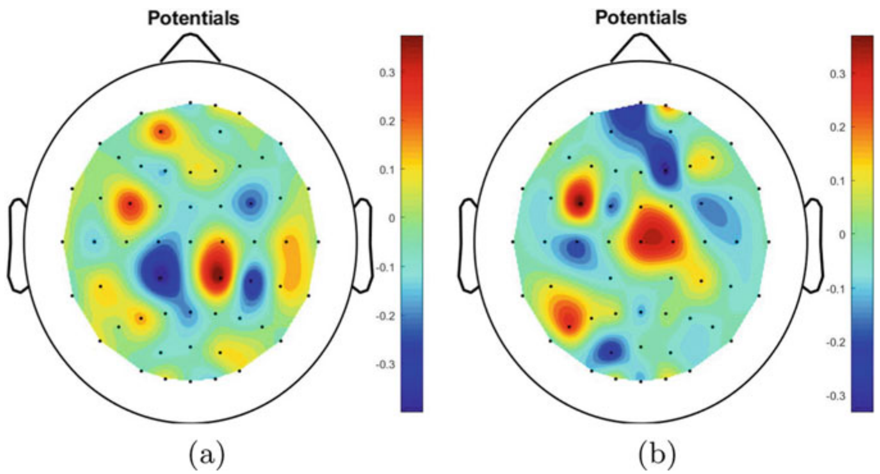


Fig. 9 CSP filter obtained for subject B4 for **a** right arm and **b** right foot motor imageries, under protocol B. Under this protocol, four events were there, namely imagery of right arm and right foot movement, and actual movement of right arm and right foot

4.2.1 Protocol C: LA-RA with Auditory Cues, Followed by Actual Motion

For this protocol and those to follow, a reduced electrode set was considered to reduce dimensions. This was decided based on our understanding of the regions within which the two motor imagery (left arm and right arm) are expected to lie. The topoplots of the filters (Fig. 10), obtained with this protocol, are very close to the ones reported in the literature.

4.2.2 Protocol D: LA-RA with Somatosensory Cues

As seen in the topoplots in Fig. 11, the introduction of localized somatosensory cues as detailed earlier helps the subject to stay focused during the imagery and leads to better results.

4.2.3 Protocol E: LA-RA with Somatosensory Cues, Followed by Actual Motion

Considering the benefits of somatosensory cues from protocol D, protocol E was brought about to put protocols C and D together. This was tested on two of the five subjects, who participated in protocol C. The improvements do seem to stack up as seen in the topoplots for this protocol (Fig. 12). The activation is exactly where they are expected, and the plots themselves have sharp and defined regions, which are considered favorable traits.

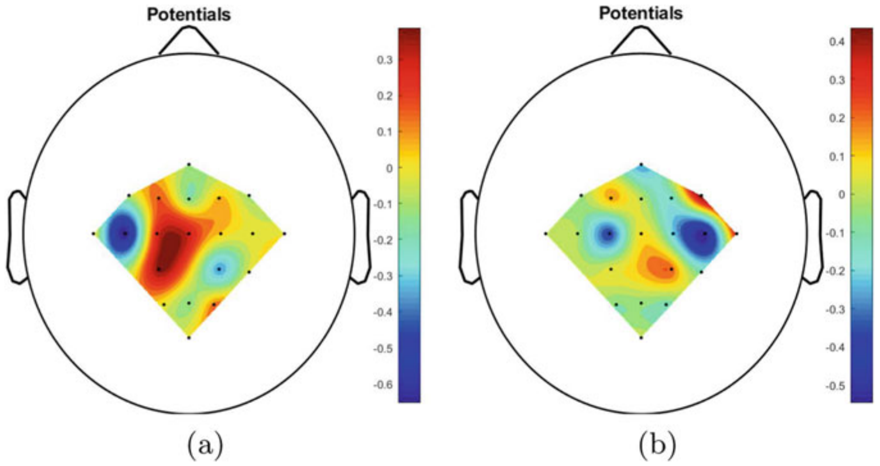


Fig. 10 CSP filter obtained for subject C6 for **a** right arm and **b** left arm motor imageries, under protocol C. Under this protocol, four events were there, namely imagery of right arm and left arm movement, and actual movement of right arm and left arm. Also, only a reduced set of electrodes were considered for this protocol

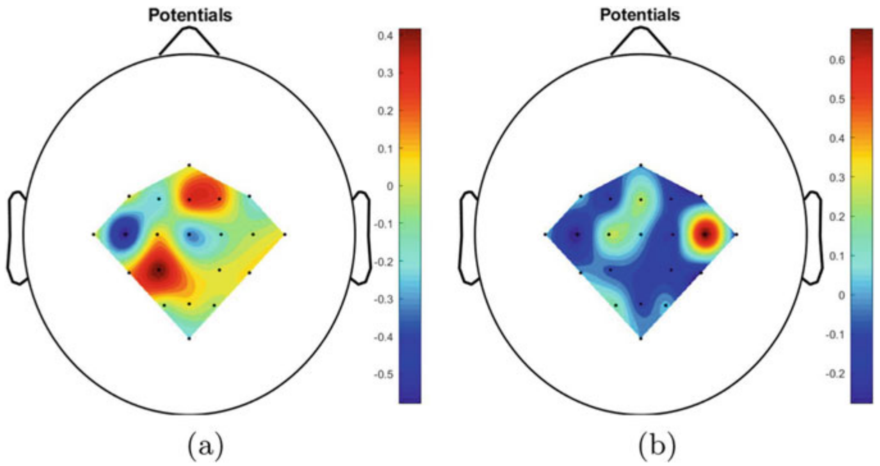


Fig. 11 CSP filter obtained for subject D2 for **a** right arm and **b** left arm motor imageries, under protocol D. Under this protocol, three events were there, namely imagery of right arm and left arm movement, and rest state. Instead of auditory cues, localized somatosensory cues were used. Also, only a reduced set of electrodes were considered for this protocol

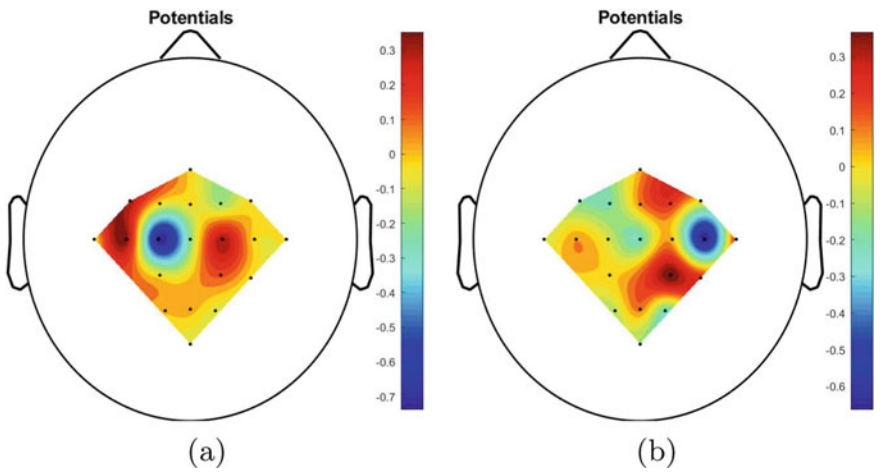


Fig. 12 CSP filter obtained for subject E6 for **a** right arm and **b** left arm motor imageries, under protocol E. Under this protocol, four events were there, namely imagery of right arm and left arm movement, and actual movement of right arm and left arm. Instead of auditory cues, localized somatosensory cues were used in this protocol. Also, only a reduced set of electrodes were considered for this protocol

5 Discussion

With four different segments of protocols, we have found CSPs for motor imagery with auditory cues, with auditory cues followed by actual movement, somatosensory cues and, somatosensory cues followed by actual movement. It is interesting to note that with somatosensory cues, the regions getting activated for motor imagery are more focused as is evident from the CSP topoplots. The regions activated are in sync with the previous experiments based on EEG [14, 15] and fMRI [9, 19] to localize motor imagery activity.

In our experiment, motor imagery with auditory cues for right foot was concentrated more sclosely in the center in topoplot (near “Cz”). This result is in sync with functional near infrared spectroscopy (fNIR) study [1] where bilateral activations were visible for limb movement imagery. Saimpont et al. have reported that auditory cues are helpful for greater visualization of motor imagery [17]. The regions activated in our study are intraparietal sulcus (IPS), supramarginal gyrus and precentral sulcus which were also activated during motor imagery in fMRI-based experiment [19]. The increased activity of the medial prefrontal cortex (mPFC) (Sensorimotor area) has been reported for motor imagery. Significant ipsilateral activations were seen in brain region represented by “P4”/“CP4” channels in patients who recovered from motor neuron disease [6]. Somatosensory cues were shown to improve motor imagery performance measured with motor evoked potentials [3].

Anterior and posterior Brodmann area (BA) 4 are involved during motor imagery as reported earlier in fMRI study on healthy subjects [19]. However, with 64 channel EEG acquisition systems, it is hard to report focused cluster-based activation in BA 4 due to the low spatial resolution.

6 Conclusions

The results obtained certainly show promise. Changing the protocol to include actual movement following the imagery helps improve the results. This arises from the fact that healthy subjects are unlikely to have been in a situation where their brain exercised motor planning (i.e., attempted to move a certain limb) but was unable, or did not want to move the limb. Thus, healthy subjects find it difficult to imagine only the planning, without actually performing the associated movement later. On the other hand, allowing the subjects to move after the trial ensures that the trial captures the activity of the brain during planning.

The results also seem to indicate that somatosensory cues stimulate the subjects better, which results in more focused motor imagery. With an evolutionary perspective, somatosensory and motor areas are in close association, which leads to diffused activation in pre and postcentral gyrus areas. This helps not only in inducing activity in the same part of the brain, but also makes it easier for the subject to localize. A subject, who is touched on his hand, would be able to easily resolve where he was touched and will be able to quickly imagine motor actions on the same hand. This resolution based on localization, we speculate, results in a better quality of motor imagery.

7 Limitations

Due to approximation of locations in EEG cap, signals from “C3” and “C4” reflect the activity from both the motor and the somatosensory cortex. Making a distinction between motor and somatosensory activity was not feasible with EEG.

References

1. Batula AM, Mark JA, Kim YE, Ayaz H (2017) Comparison of brain activation during motor imagery and motor movement using fNIRS. Comput Intell Neurosci 2017
2. Blankertz B, Müller KR, Curio G, Vaughan TM, Schalk G, Wolpaw JR, Schlögl A, Neuper C, Pfurtscheller G, Hinterberger T et al (2004) The BCI competition 2003. IEEE Trans Biomed Eng 51(6):1044–51

3. Bonassi G, Biggio M, Bisio A, Ruggeri P, Bove M, Avanzino L (2017) Provision of somatosensory inputs during motor imagery enhances learning-induced plasticity in human motor cortex. Sci Rep 7(1):9300
4. Brandl S, Höhne J, Müller KR, Samek W (2015) Bringing BCI into everyday life: motor imagery in a pseudo realistic environment. In: 2015 7th international IEEE/EMBS conference on neural engineering (NER). IEEE, pp 224–227
5. Brandl S, Müller KR, Samek W (2016) Alternative CSP approaches for multimodal distributed BCI data. In: 2016 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, pp 003742–003747
6. Bundy DT, Souders L, Baranyai K, Leonard L, Schalk G, Coker R, Moran DW, Huskey T, Leuthardt EC (2017) Contralesional brain-computer interface control of a powered exoskeleton for motor recovery in chronic stroke survivors. Stroke 48(7):1908–1915
7. Claassen J, Doyle K, Matory A, Couch C, Burger KM, Velazquez A, Okonkwo JU, King JR, Park S, Agarwal S et al (2019) Detection of brain activation in unresponsive patients with acute brain injury. N Engl J Med 380(26):2497–2505
8. Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134(1):9–21
9. Hanakawa T, Immisch I, Toma K, Dimyan MA, Van Gelderen P, Hallett M (2003) Functional properties of brain areas associated with motor execution and imagery. J Neurophysiol 89(2):989–1002
10. Jerrin TP, Ramakrishnan A, Ananthapadmanabha T (2019) A novel deep learning architecture for decoding imagined speech from EEG. In: IEEE Austria international biomedical engineering conference (AIBEC 2019). IEEE
11. Kropotov JD (2010) Quantitative EEG, event-related potentials and neurotherapy. Academic, San Diego
12. Kropotov JD (2016) Functional neuromarkers for psychiatry: applications for diagnosis and treatment. Academic, San Diego
13. Lotte F, Guan C (2010) Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. IEEE Trans Biomed Eng 58(2):355–362
14. McFarland DJ, Miner LA, Vaughan TM, Wolpaw JR (2000) Mu and beta rhythm topographies during motor imagery and actual movements. Brain Topogr 12(3):177–186
15. Pfurtscheller G, Neuper C (2001) Motor imagery and direct brain-computer communication. Proc IEEE 89(7):1123–1134
16. Pfurtscheller G, Brunner C, Schlögl A, Da Silva FL (2006) Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage 31(1):153–159
17. Saimpont A, Malouin F, Tousignant B, Jackson PL (2013) Motor imagery and aging. J Mot Behav 45(1):21–28
18. Sannelli C, Dickhaus T, Halder S, Hammer EM, Müller KR, Blankertz B (2010) On optimal channel configurations for SMR-based brain-computer interfaces. Brain Topogr 23(2):186–193
19. Sharma N, Jones PS, Carpenter T, Baron JC (2008) Mapping the involvement of BA 4a and 4p during motor imagery. Neuroimage 41(1):92–99

A Centre of Gravity-Based Preprocessing Approach for Feature Selection Using Artificial Bee Colony Algorithm on High-Dimensional Datasets

M. G. Bindu and M. K. Sabu

Abstract The process of feature selection has a high impact on data mining tasks such as classification and clustering. Removing irrelevant, noisy and redundant data not only increases the quality of the task but also reduces the computational complexity and execution time. Nature-inspired algorithms have tackled the problem of feature selection efficiently. But when applying on a high-dimensional dataset, the metaheuristic algorithms have difficulty to converge. In this paper, an existing artificial bee colony algorithm for feature selection is modified by incorporating a data preprocessing step to reduce the size of the input dataset. The preprocessing step computes the centre of gravity vectors corresponding to the original dataset to form a smaller dataset. The artificial bee colony algorithm works on this smaller dataset for feature selection. The proposed method generates better results with less time and complexity when compared to the existing algorithms.

Keywords Swarm intelligent algorithms · Artificial bee colony algorithm · Feature selection · High-dimensional dataset · Centre of gravity of a dataset

1 Introduction

Data mining involves finding out useful patterns hidden in a large collection of data. Since this is an era of data explosion, there will be a huge amount of data available for mining. The data will have irrelevant, redundant features too. Irrelevant data may mislead the mining task, while redundant data will create inconsistency and memory wastage. So it is important to select and use only the relevant data from what is available and avoid the redundant and incomplete one. This is achieved by selecting a relevant subset of features from the original data known as feature selection. It

M. G. Bindu (✉) · M. K. Sabu
Department of Computer Applications, Cochin University of Science and Technology, Kochi, Kerala, India
e-mail: bindumg@cusat.ac.in

M. K. Sabu
e-mail: sabumk@cusat.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_23

will save a large amount of time and space. Feature selection will also improve the quality of data mining tasks.

But, finding a minimal subset of relevant features out of a dataset is categorized as an NP-hard problem. Nowadays, it is handled much effectively using a category of artificial intelligent algorithms known as swarm intelligence techniques [1]. Swarm intelligent (SI) algorithms are a collection of bio-inspired algorithms. In nature, there are several groups of social insects, which collectively work to achieve a single goal like foraging for food, prey evading, etc. Individuals belonging to such a group are called agents. A self-organized population of homogeneous agents interacting with each other and with their environment is called swarm. Natural swarm systems exhibit high efficiency in solving complex problems. Swarm intelligence is an emergent collective intelligence of a group of agents. The computational modelling of SI techniques are already successful at producing fast, robust and low-cost solutions for problems in different domains like image and data analysis, optimizations, machine learning, etc. SI techniques gained popularity since they are nature-inspired, scalable, self-organized, robust and stable. Moreover, these algorithms show the highest degrees of adaptability. The best-known SI algorithms are ant colony algorithm, artificial bee colony algorithm, bat algorithm, firefly algorithm, etc.

The artificial bee colony (ABC) algorithm was proposed by Karaboga and Akay [2]. It is a population-based stochastic algorithm, which has outperformed several existing evolutionary algorithms [3–6]. ABC algorithm was initially proposed to solve numerical optimization problems. Many improved versions of the original ABC algorithm had come up and used for feature selection, clustering, vehicle routing, image processing, etc. [7]. Some of them for feature selection are briefly discussed here. A binary version of the ABC algorithm was proposed by Mauricio Schiezaro and Pedrini [8], for feature selection. It outperformed several traditional feature selection algorithms and generated better results. Uzer et al. [9] proposed a feature selection algorithm based on ABC to diagnose liver diseases. It has proved to select features faster than the traditional method. Yavuz and Aydin [10] proposed an angle-modulated ABC algorithm for feature selection. It proved quite effective for high-dimensional datasets. Suguna and Thanushkodi [11] have used a rough set-based approach to find the reduct of a medical dataset, and then on these reducts, ABC algorithm is applied to find the minimal subsets of features in medical datasets. The results were outperforming. Shanthi and Bhaskaran [12] suggested a modified ABC, which improved the exploitation ability of the ABC algorithm by considering the global best solution in the search equation. They got the highest accuracy in the for mammogram image classification.

Though feature selection using swarm intelligent algorithms is efficient, they take much time to converge for high-dimensional datasets. ABC algorithm is nothing different. It got always criticized for its low convergence. Several studies have been done to improve the convergence of the ABC algorithm. Worasucheep [13] used a hybrid ABC algorithm incorporating the concept of differential evolution to improve its convergence. The paper suggests taking benefit of differential evolution mutation strategies combined with ABC. The results showed that the convergence speed of the ABC algorithm has improved considerably. Alshamlan [14] used a co-ABC method to find the highly correlated subsets of genes from microarray subsets. They used a preprocessing step to filter out noisy features and applied ABC algorithm for feature selection. Results showed that the co-ABC algorithm provided the best accuracy results with fewer informative genes. It is suggested to be great if we reduce the size of the input dataset before applying the ABC algorithm for feature selection [14].

This paper proposes a preprocessing step to reduce the size of a high-dimensional input dataset by generating a representative reduced dataset. The concept of computing centre of gravity (CoG) vectors of a dataset [15] is used here to reduce the size of a high-dimensional dataset. The reduced dataset can be used in place of the original high-dimensional dataset as input to the ABC algorithm for feature selection. In this paper, we demonstrate how the proposed size reduction method helps the ABC algorithm to perform better on a high-dimensional dataset.

Remaining part of the paper is organized as follows. Part 2 discusses the related concepts used in this study. It discusses the original ABC algorithm and its binary version for feature selection in detail. The section also gives a detailed description of the concept of centre of gravity and how the centre of gravity vectors is calculated for a dataset. Part 3 describes the proposed work in detail. Part 4 gives the experimental set-up. Part 5 reports the results of the proposed method and discusses them in detail. Part 6 concludes the paper.

2 Related Concepts

2.1 Centre of Gravity of a Dataset

In physics, CoG is an imaginary point where the total weight of the body is concentrated so that if supported at this point, the body would balance perfectly. This idea is used to find the CoG line of a set of points on a plane [15]. On a plane, the CoG line of a set of points lies in such a way that the sum of perpendicular distance from all the points to this imaginary line is zero. Thus, CoG line lies in the midway from these points.

This concept is adapted to reduce the number of records in a multiclass high-dimensional dataset. Initially, the dataset is divided into different partitions based on their target classes. Each partition is further divided into smaller slices of fixed size. A CoG vector is calculated corresponding to each slice. Finally, all the CoG vectors are grouped to form a reduced dataset. The number of records in the new dataset is far less than that of original dataset since one slice of the original dataset is replaced with a single vector. This eases further computations on that dataset, and computation time is also reduced.

Suppose there are p attributes and m classes in a multiclass high-dimensional input dataset. Divide the dataset class-wise to make m different partitions. Let the number of records in each partition be n. Each partition is then divided to form k small slices of size $\lfloor n/k \rfloor$ or less. To calculate the CoG vector corresponding to one data slice,

consider the data slice as a matrix with $\lfloor n/k \rfloor$ number of rows and p number of columns as given by (1). Let the value $\lfloor n/k \rfloor$ be r.
i.e.

$$D_{r\times p} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1p} \\ d_{21} & d_{22} & \cdots & d_{2p} \\ & & & \vdots \\ d_{r1} & d_{r2} & \cdots & d_{rp} \end{bmatrix} \tag{1}$$

Form a new matrix $D'_{r\times p}$ by applying (2) on each element of $D_{r\times p}$ for x ranging from 1 to r.

$$d'_i(x) = \left| \sum_{j=1}^{n/k} \bigl(d_j(i) - d_x(i)\bigr) \right| \tag{2}$$

where $i = 1, 2, 3, 4, \ldots, p$.

$D'_{r\times p}$ will be the matrix given by (3)

$$D'_{r\times p} = \begin{bmatrix} d'_{11} & d'_{12} & \cdots & d'_{1p} \\ d'_{21} & d'_{22} & \cdots & d'_{2p} \\ & & & \vdots \\ d'_{r1} & d'_{r2} & \cdots & d'_{rp} \end{bmatrix} \tag{3}$$

After forming the new matrix D', find the indices of minimum values along each column in D'. Form a row vector which contains the row number of minimum value along each column of the matrix D' using (4).

$$\text{index} = \operatorname{argmin}\bigl(d'_c\bigr) \tag{4}$$

where $c = 1, 2, 3, 4, \ldots, p$.

Then using these indices, take out values corresponding to these indices from original dataset slice using (5)

$$\text{CoG}[i] = D[\text{index}[i]][i] \quad \text{for } i = 1, 2, 3, 4, \ldots, p \tag{5}$$

Each column of the original data slice is then represented by these corresponding elements from CoG vector. After calculating the CoG vector for a slice, the slice is replaced with the CoG vector. This is repeated for all slices in the dataset. For each class in a dataset, 'k' number of CoG vectors are computed, i.e. if there were 'n' objects for one class in the original dataset, they will get replaced by 'k' CoG vectors, where $k \ll n$. Each class in the dataset is thus represented by lesser number of objects. All CoG vectors are grouped together to form a reduced dataset. The dimension of dataset is thus reduced from $m \times n$ to $m \times k$.

2.2 Artificial Bee Colony Algorithm

Among SI algorithms, ABC is an effective metaheuristic algorithm. It simulates the social living of bees and their honey-collecting behaviour. Here, the artificial bees are the agents searching for good-quality food sources. The nectar amount at the source indicates the quality of the food source. The process begins when the bees start their search for food sources. They return to the beehive after collecting information about food sources and perform a waggle dance to share the information with other bees. More bees will be allocated to explore the best-quality food sources among them. They will find out the best-quality food source.

ABC algorithm works with four components. They are the food sources, employed bees, onlooker bees and scout bees. Each food source represents a possible solution to the problem. Initially, employed bees find out the food sources and share information about the quality of food sources with other bees in the hive. The number of employed bees will be equal to the number of food sources. Onlooker bees select good-quality food sources from the available ones and explore their neighbours. When one food source is explored for a certain number of times and could not find a better quality neighbourhood from it, then it is abandoned. Scout bees will find new food sources to replace the abandoned ones. The pseudo-code for ABC algorithm used in optimization problems is given below.

Pseudo-code for ABC optimization is as follows:

1. Initialization of food sources
2. Repeat
3. Employed bee phase
4. Onlooker bee phase
5. Scout bee phase
6. Memorize the best food source
7. Until maximum number of cycles reached.

Initialization of food sources The ABC algorithm puts forward a random initialization of food sources given by (6)

$$x_{ij} = x_j^{\min} + r\ \text{and}(0, 1)(x_j^{\max} - x_j^{\min}) \tag{6}$$

where $i = 1, 2, \ldots, N$, N is the number of food sources, $j = 1, 2, \ldots, M$, M is the number of optimization parameters.

Employed Bee Phase The employed bees are made to work on the food source created by (6). Employed bees explore the neighbourhood of food sources given by (7)

$$h_{ij} = x_{ij} + \emptyset_{ij}\left(x_{ij} - x_{kj}\right) \tag{7}$$

For each food source x_i, a neighbour h_i is determined. Ø is a real number between 1 and −1. Indices j and k are random variables. When the neighbours are explored, each food source is evaluated for its fitness as given by (8)

$$\text{fitness}_i = \begin{cases} \frac{1}{1+f_i} & \text{if} f_i \geq 0 \\ 1 + \text{abs}(f_i) & \text{if} f_i < 0 \end{cases} \tag{8}$$

Here, f_i is the cost function.

Onlooker Bee Phase From the fitness information of food sources provided by the employed bees, onlooker bees calculate the probability of a food source being selected by them using (9)

$$\text{Prob.}_i = \frac{\text{fitness}_i}{\sum_{i=1}^{n} \text{fitness}_i} \tag{9}$$

The onlooker selects a food source with higher probability value for further exploration. Once they select a food source for exploration, they work by (7). Each food source will be explored for a predefined number of times. Even after that if the bee could not find a neighbour of better quality, the food source is abandoned.

Scout bee phase Once a food source is abandoned, a scout bee will be assigned to create a new food source randomly.

2.3 ABC Algorithm for Feature Selection

ABC algorithm has been used in many optimization problems efficiently. It was Schiezaro and Pedrini [8] who suggested a feature selection method based on the ABC algorithm. They suggested a binary version of the original ABC algorithm. The algorithm performs well on different UCI datasets. It reduced the number of features used for classification considerably without compromising on the accuracy. When compared to the popular feature selection techniques, the feature selection using the ABC algorithm showed better results.

In the ABC algorithm for feature selection, food sources are represented in the form of bit vectors instead of real numbers. Bit vector will be having a size equal to the number of features in the dataset. Each bit vector will be representing a feature subset for classification. A '1' value in the bit vector denotes that its corresponding feature is to be considered for classification, while features corresponding to zeros in the bit vector are not used during the classification. Classification accuracy obtained using the feature subset represented by a food source is considered as the fitness of that food source. Aim of the algorithm is to find the minimal subset of features that yields maximum quality.

Initially, bit vectors are created equal to the number of food sources. Each bit vector is initialized with an only one-bit value set to '1' so that it represents a single feature. All food sources are submitted to the classifier, and accuracy is saved as food quality. Employed bees are used to explore the neighbourhood of initial food sources. Neighbour food sources are generated by randomly updating the bit values to 1. Neighbours having higher quality are added as new food sources. Food sources are abandoned when they are exhausted with neighbourhood exploration. From the fitness information collected from employed bees, onlooker bees select food sources with better quality for exploration. Scout bees generate new food sources to replace each abandoned food source by randomly initializing bit vectors. The algorithm stops when a food source of desired quality is found or termination criteria are met. Since feature selection aims at improving the classifier performance, accuracy is taken as the fitness function to measure the quality of a food source.

3 Proposed Method

The proposed method works in two phases. The first phase is for data preprocessing, and the second phase is for feature selection. The first phase uses the concept of calculating the centre of gravity vectors to reduce the size of a high-dimensional input dataset. After the execution of the first phase, a representative reduced dataset created and the original dataset is replaced by the reduced set. This reduced dataset is given as the input for the second phase. ABC performs feature selection using the reduced dataset and generates results. The proposed method works as given by Algorithm 1.

Algorithm 1 A proposed preprocessing step for feature selection using the ABC algorithm.

Input: High-dimensional input dataset

Output: Minimal subset of features with the corresponding accuracy.

1. Load the high-dimensional input dataset.
2. Divide the dataset into partitions class-wise.
3. For each partition in the dataset,
    a. Divide each partition into several slices of equal size.
    b. For each data slice in partition.
        i. Calculate the CoG vector corresponding to the slice of data.
        ii. Replace the slice with corresponding CoG vector.
4. Group CoG vectors together to form a reduced dataset.
5. The reduced dataset is given as input into the ABC algorithm for feature selection.

6. Initialize food sources by creating bit vectors.
7. Train the classifier using the initial bit vectors, and store their accuracy as fitness.
8. Explore the neighbourhood of initial food sources using employed bees.
9. Calculate the fitness of neighbourhood food sources using the classifier.
10. Compare the neighbour's accuracy with the accuracy of the original food source.
11. If the neighbour has higher fitness than the original food source, then it is considered as a new food source.
12. If the maximum number of explorations on a food source has reached and no better neighbouring food source has been found, then abandon that food source.
13. Among the available food sources, onlooker bees find the best-quality food sources and start exploring their neighbourhood.
14. For each abandoned food source, a scout bee creates a new food source randomly.
15. Once all onlookers are all distributed, the global best food source is found.
16. When the desired accuracy and a predefined number of iterations are reached, stop the process.
17. Return the accuracy with the selected features and execution time.

4 Experimental Framework

4.1 Datasets

The performance of the proposed method is evaluated using five numerical datasets. They are available publicly in the UCI data repository [16]. UCI provides datasets from various knowledge domains with various numbers of attributes and different dimensions. We have selected datasets with a larger number of records as we need to evaluate the effect of size reduction on classification performances. Detailed descriptions of the datasets are given in Table 1.

Table 1 Dataset description

Dataset		No. of instances	No. of attributes	No. of classes
D1	Avila	20,867	10	12
D2	Letter recognition	20,000	17	26
D3	Pen-based recognition of handwritten digits	10,992	16	10
D4	Electrical grid stability	10,000	14	2
D5	Anuran Calls	7195	22	60

4.2 Classification Set-up

Performance of the proposed method is evaluated based on the classification accuracy, execution time needed and the number of features used for classification. Random forest classifier is used for classification. The proposed algorithm is implemented in Python.

5 Results and Discussion

Table 2 shows the classification performances given by different datasets on applying existing ABC algorithm for feature selection. The datasets are given as inputs to the ABC algorithm for feature selection without any size reduction. The classification accuracy, execution time and the number of features selected are observed. The same datasets are preprocessed with the proposed size reduction technique to generate reduced datasets. Table 3 compares the sizes of the original dataset with the datasets obtained after size reduction. Then, the reduced datasets are given as inputs into the ABC algorithm. Table 4 gives the classification performances of reduced datasets for the proposed method.

Figure 1 is used to demonstrate the effectiveness of the proposed method. Classification accuracies obtained for the same datasets for two approaches are plotted. Figure 1 compares the accuracy obtained for the two approaches.

Table 2 Results with existing method

Dataset	Accuracy	Execution time (s)	No. of features selected
D1	53.66	117.02	8
D2	72.3	375.05	11
D3	88.55	240.38	12
D4	90.9	111.23	9
D5	59.38	370.66	16

Table 3 Size comparison of original and reduced dataset

Dataset	Size of original dataset	Size of reduced dataset	Percentage of size reduction	Time taken for preprocessing (s)
D1	20,867	4176	79.99	13.00
D2	20,000	2015	89.93	18.06
D3	10,992	3668	66.63	12.78
D4	10,000	1977	80.23	7.99
D5	7195	1440	79.98	9.86

Table 4 Results with proposed method

Dataset	Accuracy	Execution time (s)	No. of features selected
D1	75.72	20.20	8
D2	88.43	48.24	10
D3	97.00	40.48	12
D4	95	28.83	8
D5	64.46	100.64	15

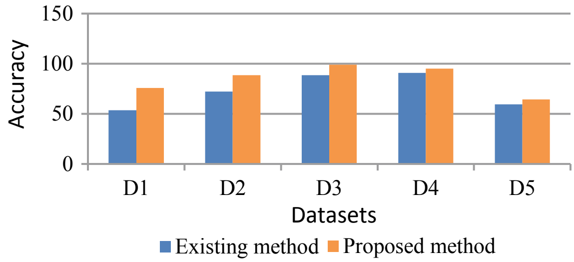


Fig. 1 Graph comparing accuracies

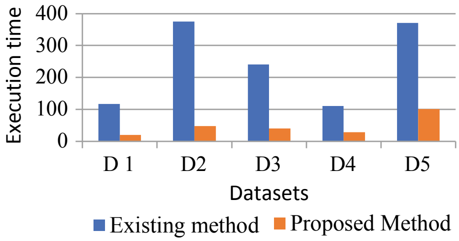


Fig. 2 Graph comparing execution time

For a better understanding of the improvement in execution time, a graph is plotted with the execution times from two approaches which form Fig. 2.

Though ABC algorithm performs well for feature selection, it is slow to converge on high-dimensional datasets. The results with the existing method on high-dimensional datasets prove the same. The proposed method, by generating a reduced dataset, solves this problem. It reduces number of records in a high-dimensional multiclass dataset without losing much information content. The reduced dataset represents original one perfectly with lesser number of records. From the results, it is observed that execution time for ABC algorithm has reduced considerably after applying the proposed preprocessing step for existing ABC algorithm. By avoiding over-fitting, classification accuracy has also increased. Thus, overall performance of ABC algorithm can be improved with the proposed method on high-dimensional datasets. Results show that the proposed method has little effect on deciding the number of selected features for classification.

6 Conclusion

In this paper, we tried to interpret the effectiveness of dimensionality reduction of high-dimensional datasets on the performance of the ABC algorithm for feature selection. The results obtained from executing the proposed method shows that the size reduction helps ABC algorithm to converge faster and thus execution time for the ABC algorithm is reduced. Though the size of a high-dimensional dataset is reduced, the proposed method preserves the information content of the original dataset by generating perfect representative vectors. Classification accuracy is also increased after applying the proposed method. Thus, the proposed method can be strongly suggested as an efficient technique to accelerate the convergence of ABC algorithm on a high-dimensional dataset. It does not improve the number of features selected for classification.

References

1. Brezočnik L, Fister I Jr, Podgorelec V (2018) Swarm intelligence algorithms for feature selection: a review. Appl Sci 8(9):1521
2. Karaboga D (2005) An idea based on honey bee swarm for numerical optimization. Erciyes University, Kayseri, Turkey, Technical Report-TR06
3. Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Global Optim 39(3):459–471
4. Karaboga D, Basturk B (2008) On the performance of artificial bee colony (ABC) algorithm. Appl Soft Comput 8(1):687–697
5. Karaboga D, Basturk B (2009) Acomparative study of artificial bee colony algorithm. Appl Math Comput 214(1):108–132
6. Karaboga D (2010) Artificial bee colony algorithm. Scholarpedia 5(3):6915
7. Xiang WL, Meng XL, Li YZ, He RC, An MQ (2018) An improved artificial bee colony algorithm based on the gravity model. Inf Sci 429:49–71. https://doi.org/10.1016/j.ins.2017.11.007
8. Schiezaro M, Pedrini H (2013) Data feature selection based on artificial bee colony algorithm. EURASIP J Image Video Process 1:47. https://doi.org/10.1186/1687-5281-2013-47
9. Uzer MS, Yilmaz N, Inan O (2013) Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification. Sci World J 2013(419187):10
10. Yavuz G., Aydin D (2016) Angle modulated artificial bee colony algorithms for feature selection. Appl Comput Intell Soft Comput. https://doi.org/10.1155/2016/9569161
11. Suguna N, Thanushkodi KG (2011) An independent rough set approach hybrid with artificial bee colony algorithm for dimensionality reduction. Am J Appl Sci 8(3):261–266. https://doi.org/10.3844/ajassp.2011.261.266
12. Shanthi S, Bhaskaran VM (2014) Modified artificial bee colony based feature selection: a new method in the application of mammogram image classification 3(6):1664–1667
13. Worasucheep C (2015) A hybrid artificial bee colony with differential evolution. Int J Mach Learn Comput 5(3):179–186
14. Alshamlan HM (2018) Co-ABC: correlation artificial bee colony algorithm for biomarker gene discovery using gene expression profile. Saudi J Biol Sci. https://doi.org/10.1016/j.sjbs.2017.12.012

15. Charlotte B, Kambhampati C (2008) Autonomous clustering using rough set theory. Int J Autom Comput 5. https://doi.org/10.1007/s11633-008-0090-3
16. https://archive.ics.uci.edu/ml/datasets.php

Design and Analysis of Metamaterial Loaded Microstrip Slotted Patch Antenna for Wireless Applications

K. Sajith, T. Shanmuganantham, and D. Sindhahaiselvi

Abstract The development in the presentation of novel microstrip slotted patch antenna geometry by the use of a circular SRR at the rear side of the ground is proposed. This improved performance has been achieved in the terms of significant addition of bands, better reflection coefficient and enhanced bandwidth and gain at the different resonant frequencies in the final stage design. The prototype has been built through the iterations of a novel geometry, introduction of slots in the second stage and finally the insertion of a split-ring resonator of circular shape at the rear side of the dielectric substrate in the third stage. Coaxial feeding technique has been employed to excite the structure at its fundamental frequency of operation. The slots embedded into the patch structure and the modified ground geometry not only scale the number of resonant frequencies, the corresponding gain and bandwidth but also provide for impedance matching at the different sideband frequencies of resonance. The anticipated prototype possibly will be used for radiolocation purpose at the design frequency of 2.73 GHz and also at the sideband frequencies of 9.11 and 9.32 GHz, fixed satellite at 4.47 and 8.81 GHz, mobile applications except aeronautical mobile at 5.67 and 7.81 GHz and for fixed mobile applications at 7.1 GHz. The final two design stages have been fabricated, and an untried verification has naked their performance to be in close approximation to the simulated ones.

Keywords Antenna parameter · SRR · HFSS-15 · Coaxial feed · Slotted patch

---

K. Sajith (✉) · T. Shanmuganantham
Department of Electronics Engineering, Pondicherry University, Pondicherry 605014, India
e-mail: Sajithrajan999@gmail.com

T. Shanmuganantham
e-mail: shanmuga.dee@pondiuni.edu.in

D. Sindhahaiselvi
Department of EIE, Pondicherry Engineering College, Pondicherry 605014, India
e-mail: sindhanaiselvi@pec.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_24

1 Introduction

The domain of antenna engineering has been completely revolutionized with the outbreak of microstrip patch antennas which continue to widen their range of applications. The three-layer microstrip patch antenna structures with a dielectric substrate sandwiched between the metallic plates of ground and patch on either side, the advantage of smaller size, low profile, higher gain, wider bandwidth. A survey of the available literatures indicates an extensive use of rectangular and circularly shaped patch antennas at an initial stage for various applications owing to their simple design and easy computation. Shanmuganantham et al. in [1] designed a microstrip rectangular patch antenna that resonated at 6.29 GHz providing a gain of just −1.95 dBi (below 0) and a bandwidth of 110 MHz. Roy et al. in [2] developed a reference rectangular patch antenna that resonated at 5.7 GHz with a gain of 4.7 dBi and bandwidth of 182 MHz. The slotting of the patch geometry has been argued to result into multiband performance. It was found that when the rectangular patch antenna in [1] was embedded with slots, the structure resonated at 5.68 GHz with a bandwidth of 200 MHz and a gain of 2.47 dBi and at 9.42 GHz with a bandwidth of 250 MHz and a gain of 5.58 dBi, respectively. Embedding slots in the patch structure have been observed to result into multiband performances with improved gain [3, 4]. Shanmuganantham et al. in [5] designed a novel shaped patch structure for fixed satellite applications. This structure displayed dual-band resonance (2.7 and 4.57 GHz) and provided respective gains of 1.6 and 4.9 dBi. Other multiband techniques use multiple elements, fractals and split-ring resonators. Raghavan et al. [6] proposed a compact size multiband antenna for wireless applications. The patch element was loaded with circular split-ring resonator, while multiple rings of rectangular split-ring resonator were embedded on the rear side of the ground plane. The structure resonated at 3.1 GHz, 4.4 GHz and 4.8 GHz with respective return loss of −11.96 dB, −12.46 dB and −16.15 dB, respectively. Thus, the proposed structure exhibited triple-band resonance with acceptable values of corresponding return loss. Shanmuganantham et al. [7] presented a SRR loaded antenna that capitulated with a −10 dB at 1.88 GHz, 6.54, 7.88, 12.2 and 15.08 GHz providing good S11 for the functional frequency bands. Shanmuganantham et al. [8] presented a compact CPW feed leaf-shaped Koch fractal antenna loaded with circular split-ring resonator that resonated at multiple frequencies including 1.977, 3.19, 6.32, 8.814, 11.87, 15.69 and 18.3 GHz with a −10 dB return loss. A total of five frequency bands with significantly improved parameters of mismatching ratio, gain and bandwidth at the corresponding frequencies were recorded in the final stage design.

The design proposed in this article is developed using a novel shaped patch (combination of four identical circular patches) at the initial stage. The novelty of the patch structure lies in the assumption of a geometrical shape other than the conventionally used rectangular and circular patches. This stage design resonates at four different frequencies (3.1, 4.43, 7.11 and 7.91 GHz) with corresponding reflection coefficient values of −37.11 dB, −13.12 dB, −17.89 dB and −15.42 dB, respectively, gain values of 3.1 dBi, 1.2 dBi, 0.54 dBi and 0.99 dBi, respectively and attained bandwidth of 90 MHz, 88 MHz, 110 MHz and 62 MHz, respectively. Thus, the novelty of the patch may also be attributed to the yielded results (multiband operation with significant reflection coefficient, gain and bandwidth) despite the fact that it has not been modified in any sense. Modification of the patch in the second stage by introduction of circular slots not only shoots the number of resonant bands to seven but also improves the aforementioned performance indicators. The antenna now has resonating frequencies of 2.7 GHz, 2.9 GHz, 5.56 GHz, 6.9 GHz, 7.65 GHz, 8.62 GHz and 9.26 GHz with the respective reflection coefficients of −28.16 dB, −15.26 dB, −19.66 dB, −16.28 dB, −12.23 dB, −14.49 dB and −21.05 dB, gain values of 4.4 dBi, 2.8 dBi, 5.26 dBi, 6.66 dBi, 9.1 dBi, 8.85 dBi and 7 dBi, respectively, and the bandwidth values of 95 MHz, 96 MHz, 115 MHz, 108 MHz, 85 MHz, 70 MHz and 66 MHz, respectively. Use of ground slots has often been considered apt for broadbanding as well as for multiband resonance with improved reflection coefficient and gain. The ground has thus been slotted using CSRR that get stamped over the rear part of the dielectric substrate. The fixing of circular split-ring resonator provides a further increment in the number of resonating bands by two. The circular SRR loaded patch geometry is recorded to resonate at 2.73 GHz, 4.47 GHz. 5.6 GHz, 7.1 GHz, 7.81 GHz, 8.1 GHz, 8.78 GHz. 9.11 GHz and 9.32 GHz, respectively, with the respective reflection coefficient values of −28.01 dB, −17.55 dB, −20.88 dB, −18.2 dB, −13.4 dB, −16.2 dB, −23 dB, −12.99 dB and −15.68 dB, respectively, respective gain values of 7 dBi, 4 dBi, 7.79 dBi, 7.54 dBi, 11.01 dBi, 7 dBi, 8.2 dBi, 9 dBi and 11.1 dBi and the corresponding bandwidth values of 98 MHz, 102 MHz, 118 MHz, 111 MHz, 88 MHz, 79 MHz, 88 MHz, 79 MHz and 75 MHz, respectively. The structure is thus far superior to those in the reviewed literatures in terms of number of resonant frequencies, the corresponding reflection coefficients, bandwidth and gain. Also, the novelty of structure lies in its unique geometry that yields far better results than its conventional counterparts. The Federal Communications Commission has allocated different applications to different functional frequency bands, while the bands extending between 2.7–2.9 GHz, 9–9.2 GHz and 9.3–9.5 GHz have been allocated for radiolocation, those between 4.4 and 4.5 GHz for fixed satellite applications and the ones between 5.65–5.725 GHz and 7.75–7.9 GHz for mobile applications except aeronautical mobile and the one between 7.075 and 7.145 GHz for fixed satellite applications. The proposed prototype may be used for radiolocation purpose at the design frequency of 2.73 GHz and also at the sideband frequencies of 8.81, 9.11 and 9.32 GHz, fixed satellite at 4.47 GHz, mobile applications except aeronautical mobile at 5.67–7.81 GHz and for permanent mobile applications at 7.1 GHz.

A brief overview of the microstrip patch antennas, the conventionally used patch shapes, their limitations, a gradual shift to the novel shaped patch structures and techniques used for enhancement of gain and bandwidth together with multiband performance has been discussed in Sect. 1. Section 2 describes the proposed antenna configuration that has been built through different iterations. The results of the different stages of the design process have been discussed in the final section.

2 Overview of Antenna Geometry

The designed configuration evolved through the iterations of a novel patch design followed by introduction of slots and the subsequent insertion of SRR at the rear side of the substrate. This structure has been designed and simulated over high frequency simulation software 15. The substrate used in the three design stages is 60 mm × 60 mm economical FR4 epoxy that is 1.6 mm thick and the electrical permittivity of 4.4. The coaxial feed that employs outer and inner conductors is composed of a perfect electric conducting material. The two conducting layers are separated by a dielectric Teflon layer (relative permittivity = 2.76) to provide insulation between them. The dimension of the conductor layers is so adjusted to give characteristic impedance for the coaxial feed to be 50 Ω using the following equation:

$$Z_o = \left(130/\sqrt{\varepsilon_r}\right) * \ln\left(\frac{b}{a}\right) \tag{1}$$

To match the load (antenna) impedance to this characteristic impedance of feed which would serve as the input impedance, a quarter-wave transformer is considered. Thus, impedance matching is achieved at the fundamental design frequency that now has characteristic impedance of 50 Ω. The modified patch and the impress of circular split-ring resonator on the back side of the dielectric substrate result in impedance matching at the other sideband frequencies.

2.1 Concept of Patch Design

The illustrate Fig. 1 indicated designed structure that has obtained from the union of four identical circular elements. The dimensions of the structure have been indicated below. The radius of the single circular element was fixed using the following equation.

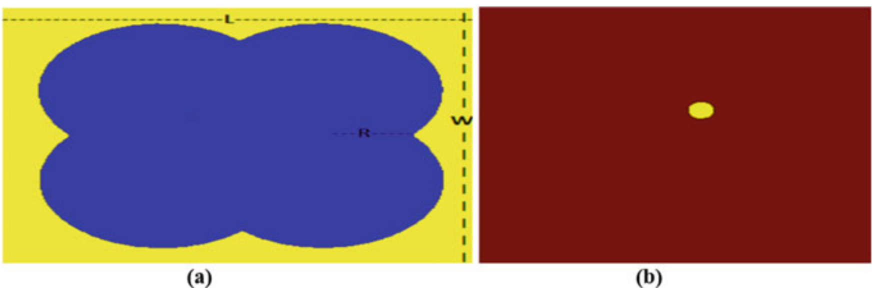


Fig. 1 Stage 1 prototype. **a** Front view, **b** back view

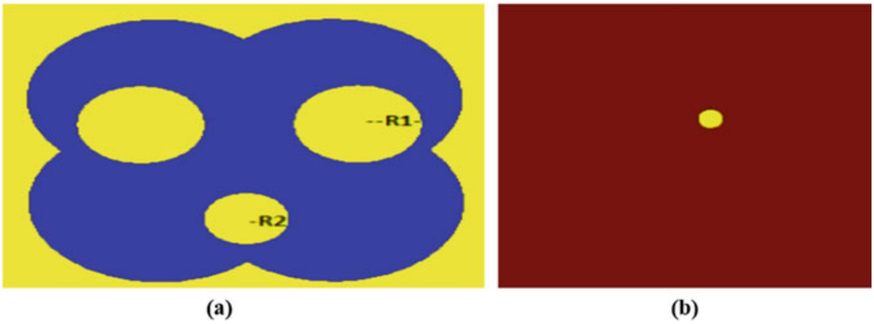


Fig. 2 Stage 2 prototypes. **a** Top view, **b** back view

$$a = \frac{F}{\sqrt{\left\{1 + \frac{2h}{\pi\,\varepsilon_r\,F}\left[\ln\left(\frac{\pi F}{2h}\right) + 1.7726\right]\right\}}} \tag{2}$$

Then

$$F = \frac{8.791 \times 10^9}{f_r\,\sqrt{\varepsilon_r}} \tag{3}$$

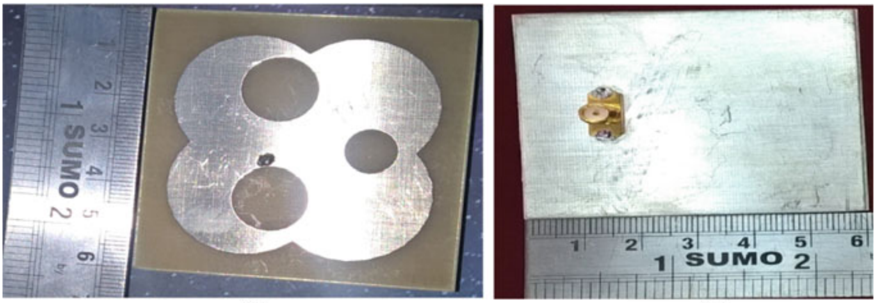
A combination of four such circular elements of equal radius was experimented to view the change in the results.

2.2 Design of Slots on Patch

The initially designed structure in Fig. 1 was slotted at appropriate locations shown in Fig. 2 for increasing the number of resonant bands added with improved bandwidth characteristics and almost matched impedance at the sideband frequencies. The fabricated prototype is shown in Fig. 3.

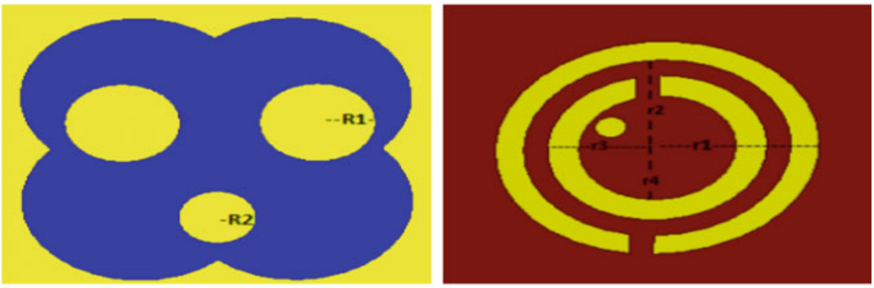
2.3 SRR Design

Dual metallic rings constantly embracing gaps on contradictory sides as in Fig. 4 are impressed on the back face of the dielectric substrate. These split-ring resonators provide for further increment in the number of resonant frequencies, improved radiation characteristics of gain and almost matched impedance at the sideband frequencies of resonance. The fabricated structure is shown in Fig. 5.



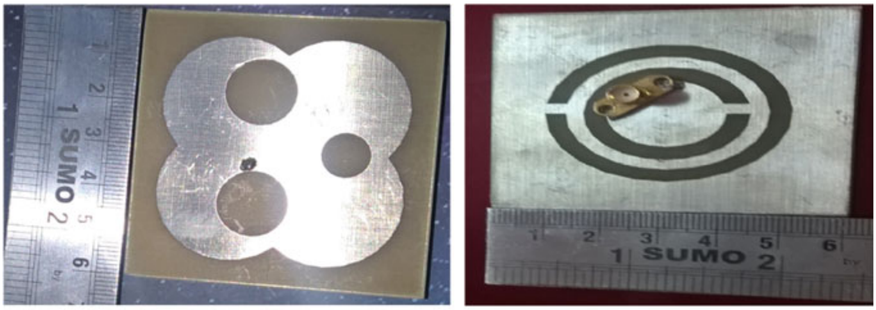
(a) (b)

Fig. 3 Stage 2 fabricated prototypes. **a** Top view, **b** back view



(a) (b)

Fig. 4 Stage 3 prototypes. **a** Top view, **b** bottom view



(a) (b)

Fig. 5 Stage 3 fabricated prototypes. **a** Top view, **b** bottom view

3 Results and Discussion

Various design stages were analyzed over high frequency simulation software version-15 for the different performance indicators including different antenna

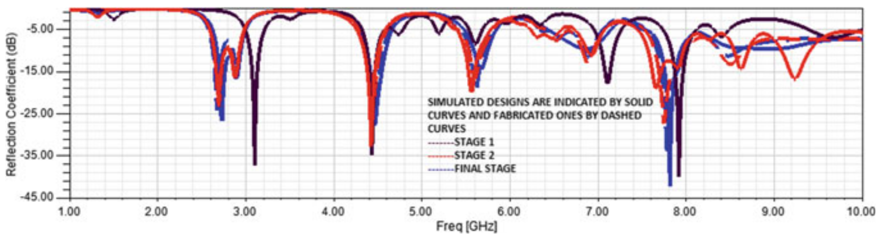


Fig. 6 Return losses verse frequency plot for various design steps 1

parameters. The designs have also been tested in anechoic chamber for the radiation pattern plots. With the received set of values, the calculated data has been plotted aligned with the simulated data on the same graph shown in Fig. 6. The design thus yields parametric excellent results compared to SRR loaded structures reviewed in the literatures. The number of bands, attained return loss is far behind the proposed structure shown in Fig. 7. The polar electric field and magnetic field plot for the different design stages and their fabricated structure are shown in Fig. 8. The results for the different design stages are also given in Table 1.

A sole circular patch is known to resonate at only one frequency. The primarily designed microstrip patches antenna geometry produced by the alliance of four indistinguishable circular patch elements is experimental to resonate at four different frequencies. This design has resonant frequencies of 3.1 GHz, 4.43 GHz, 7.11 GHz and 7.91 GHz, respectively, with the respective reflection coefficients of −37.11 dB, −13.12 dB, −17.89 dB and −15.42 dB, gain values of 3.1 dBi, 1.2 dBi, 0.54 dBi and 0.99 dBi, respectively, and bandwidth of 90 MHz, 88 MHz, 110 MHz and 62 MHz, respectively.

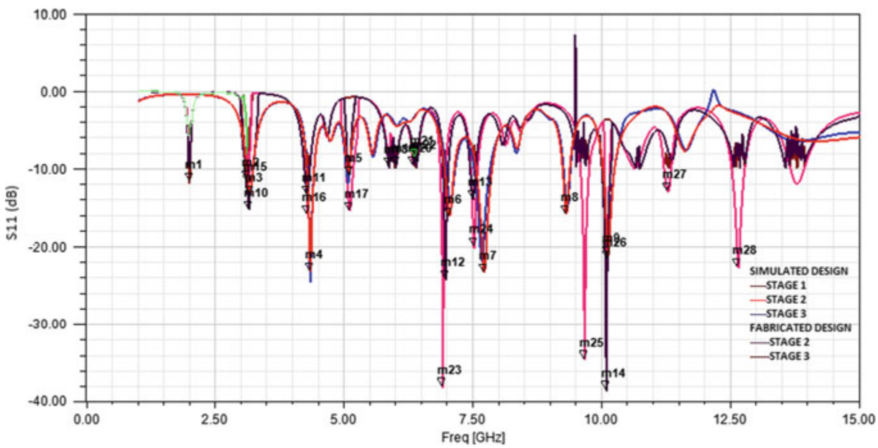


Fig. 7 Return losses verse frequency plot for various design steps 2

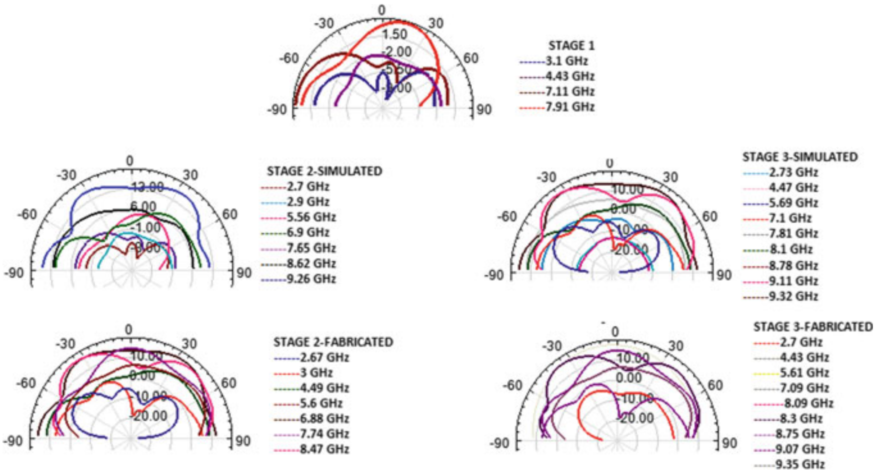


Fig. 8 Polar E&H plot of various design steps

The slot stage 1 patch geometry at appropriate locations results into a multiband performance in the stage two designs together with an improvement the performance indicators of enhanced bandwidth. The reflected in the fabricated counterpart of the stage two design which resonates at 2.67 GHz, 3 GHz, 4.49 GHz, 5.6 GHz 6.88 GHz, 7.74 GHz and 8.47 GHz, respectively, with the respective return loss of −26.08 dB, −15.22 dB, −19.98 dB, −16.11 dB, −12.20 dB, −14.8 dB and −21 dB, respective gain values of 4 dBi, 2.5 dBi, 5.22 dBi, 6.1 dBi, 9 dBi, 7.07 dBi and 6.61 dBi and bandwidth values of 92 MHz, 96 MHz, 113 MHz, 105 MHz, 86 MHz, 74 MHz and 63 MHz, respectively.

Use of ground slots has often been considered apt for broadbanding as well as for multiband resonance with improved reflection coefficient and gain. The ground has thus been slotted using CSRR that get stamped over the rear part of the dielectric substrate. A further improvement in gain and bandwidth is observed upon the insertion of SRR at the bottom part of the dielectric substrate. The circular SRR loaded patch geometry is recorded to resonate at 2.73 GHz, 4.47 GHz. 5.6 GHz, 7.1 GHz, 7.81 GHz, 8.1 GHz, 8.78 GHz. 9.11 GHz and 9.32 GHz, respectively, with the respective reflection coefficient values of −28.01 dB, −17.55 dB, −20.88 dB, −18.2 dB, −13.4 dB, −16.2 dB, −23 dB, −12.99 dB and −15.68 dB, respectively, respective gain values of 7 dBi, 4 dBi, 7.79 dBi, 7.54 dBi, 11.01 dBi, 7 dBi, 8.2 dBi, 9 dBi and 11.1 dBi and the corresponding bandwidth values of 98 MHz, 102 MHz, 118 MHz, 111 MHz, 88 MHz, 79 MHz, 88 MHz, 79 MHz and 75 MHz, respectively. The results so obtained are closely approximated by those obtained for the fabricated counterpart upon experimental verification. The fabricated structure resonates at 2.7 GHz, 4.43 GHz, 5.61 GHz, 7.09 GHz, 8.09 GHz, 8.3 GHz, 8.75 GHz, 9.07 GHz and 9.35 GHz with reflection coefficient values of −27.78 dB, −17.50 dB, −20 dB, −18.19 dB, −13.23 dB, −16.11 dB, −23.2 dB, −13.5 dB and −15.16 dB, gain values of 6.94 dBi, 3.2 dBi, 7.5 dBi, 7.21 dBi, 10.99 dBi, 6.21 dBi, 8.19 dBi,

Table 1 Parametric results of different design stages 1

Design stage	No. of bands	Resonant frequency (GHz) f1/f2/f3…	Reflection coefficient (dB) S11 at f1/S11 at f2/…
1	4	3.1/4.43/7.11/7.91	−37.11/−13.12/−17.89/−15.42
2. Simulated	7	2.7/2.9/5.56/6.9/7.65/8.62/9.26	−28.16/−15.26/−19.66/−16.28/−12.23/−14.49/−21.05
2. Fabricated	7	2.67/3/4.49/5.6/6.88/7.74/8.47	−26.08/−15.22/−19.98/−16.11/−12.20/−14.8/−21
3. Simulated	9	2.73/4.47/5.67/7.1/7.81/8.1/8.78/9.11/9.32	−28.01/−17.55/−20.88/−18.2/−13.4/−16.2/−23/−12.99/−15.68
3. Fabricated	9	2.7/4.43/5.61/7.09/8.09/8.3/8.75/9.07/9.35	−27.78/−17.50/−20/−18.19/−13.23/−16.11/−23.2/−13.5/−15.16

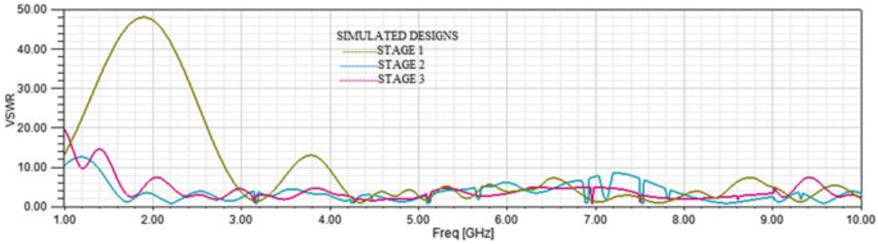


Fig. 9 SWR plot for various design steps

7.55 dBi and 9.88 dBi, respectively, and the bandwidth values of 98 MHz, 99 MHz, 112 MHz, 115 MHz, 87 MHz, 72 MHz, 83 MHz, 78 MHz and 75 MHz, respectively. Thus, all the objectives of multiband performance, reduced return loss, improved directive gain and resonance width, proper impedance matching are well served at all the design stages. The SWR (in terms of voltage) plot of the three simulated design steps is shown in Fig. 9.

The stage 1 design has a VSWR value of 1.27 at 3.1 GHz, 1.15 at 4.43 GHz, 1.49 at 7.11 GHz and 1.99 at 7.91 GHz. Thus, the condition that VSWR value should lie between 1 and 2 is satisfied at all the resonant frequencies. The second stage design has a VSWR value of 1.53 at 2.7 GHz, 1.9 at 2.9 GHz, 1.92 at 5.56 GHz, 1.99 at 6.9 GHz, 1.82 at 7.65 GHz, 1.8 at 8.62 GHz and 2 at 9.26 GHz. The final stage design is seen to resonate with VSWR values of 1.99 at 2.73 GHz, 1.76 at 4.47 GHz, 1.91 at 5.67 GHz, 1.99 at 7.1 GHz, 2 at 7.81 GHz, 2 at 8.1, 1.56 at 8.78 GHz, 1.75 at 9.11 GHz and 2 at 9.32 GHz. Thus, the VSWR condition is satisfied in every design stage.

The current distributions over the patch for the different resonant frequencies in the final stage design are shown in Fig. 10. These are in accordance with the cavity model of microstrip patch antenna that requires minimum current to flow over the surface of the patch.

The numerical values of performance indicators of different design stages and their fabricated counterpart are shown in Tables 1 and 2. The recorded set of values of different parametric indicators for the proposed prototype clearly indicate the inherent advantages of multiband performance and significant radiation characteristics of directivity and resonance width on the exploit of a circular SRR with microstrip slotted patch antenna. A closer look at the tabulated values leads to an inference that both the simulated and the fabricated results are in close agreement, and the variation in results is mainly accounted to the air gap in soldering the connectors and losses in connector cables. Table 3 draws a relationship of the parametric results of the different reference antennas that have been surveyed with those of the proposed structure.

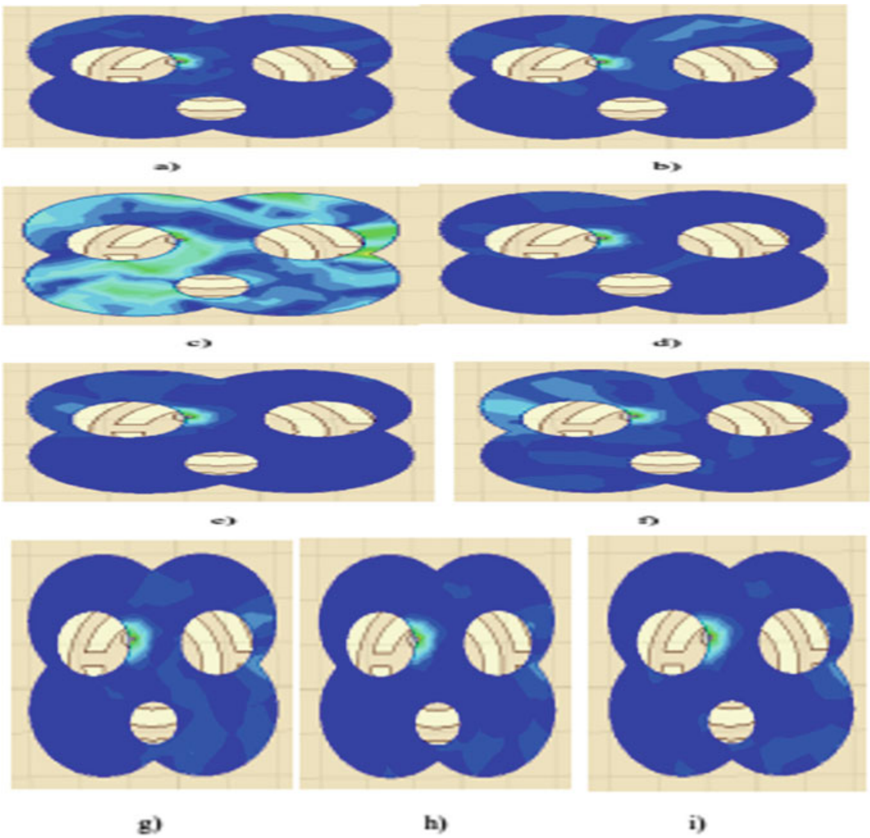


Fig. 10 Current distributions of the final stage design at **a** 2.73 GHz, **b** 4.47 GHz, **c** 5.67 GHz, **d** 7.1 GHz, **e** 7.81 GHz, **f** 8.1 GHz, **g** 8.78 GHz, **h** 9.11 GHz, **i** 9.32 GHz

Table 2 Parametric results of different design stages 2

Design stage	Bandwidth (MHz) BW1 at f1/BW2 at f2/…	Gain (dBi) G1 at f1/G2 at f2/…
1	90/88/110/62	3.1/1.2/0.54/0.99
2. Simulated	95/96/115/108/85/70/66	4.4/2.8/5.26/6.66/9.1/8.85/7
2. Fabricated	92/96/113/105/86/74/63	4/2.5/5.22/6.1/9/7.07/6.61
3. Simulated	98/102/118/111/88/79/88/79/75	7/4/7.79/7.54/11.01/7/8.2/9/11.1
3. Fabricated	98/99/112/115/87/72/83/78/75	6.94/3.2/7.5/7.21/10.99/6.21/8.19/7.55/9.88

Table 3 Comparison results with different referred papers

S. No	References	Size (mm^2)	Resonating frequency (GHz)	Nature of BW
1	[6]	16 × 19	2.6/4.8	Dual-band
2	[7]	14 × 16	1.977/3.19/6.32/8.814/11.87/15.69/18.3	Multiband
3	[8]	12 × 14	1.88/6.54/7.88/12.2/15.08	Multiband
4	Proposed	60 × 60	2.7/4.43/5.61/7.09/8.09/8.3/8.75/9.07/9.35	Multiband

4 Conclusion

The microstrip slotted patch antenna geometry by the use circular SRR at the back side of the ground is proposed. The prototype has been built through iterations of a novel geometry, introduction of slots in the second stage and finally the insertion of a circular-shaped SRR at the bottom side of the substrate in the third stage. Improved performance in terms of standard results including the multiband resonance, return loss, VSWR and gain at every subsequent stage has been recorded. The overall volume of proposed antenna is 5760 mm^3. The final two design stages have been fabricated, and an experimental verification has obtained.

References

1. Kaushal D, Shanmuganantham T, Sajith K (2017) Dual band characteristics in a microstrip rectangular patch antenna using novel slotting technique. In: IEEE international conference on intelligent computing, instrumentation and control technologies, (ICICICT), Kerala, India
2. Roy A, Choudhary PK, Anand S, Sarkar PP, Bhunia S (2014) A novel approach on miniaturization of micros trip patch antenna with loaded strips. In: International conference on electronics, communication and instrumentation (ICECI), Kolkata, India
3. Kaushal D, Shanmuganantham T Microstrip slotted caterpillar patch antenna for S, Ku and K-Band Applications. J Mater Today (in press)
4. Shanmuganantham T, Kaushal D (2017) A miniaturized rectangular slotted nameplate antenna design for satellite and radio determination applications. Electron Telecommun Res Inst J (ETRI)
5. Kaushal D, Shanmuganantham T (2016) Danger microstrip patch antenna for fixed satellite applications. In: IEEE international conference on emerging trends in technology (ICETT), Kollam, India
6. Jha N, Kumar A, Dayama V, Prudhviraj C, Pandeeswari R, Raghavan S (2016) A compact multiband microstrip patch antenna loaded with CSRR for wireless applications. In: IEEE international conference on recent trends in electronics, information and communication technology (RTEICT)
7. Elavarasi C, Shanmuganantham T (2017) Multiband SRR loaded leaf-shaped Koch fractal with a modified CPW-fed antenna. Int J Electron Lett 06(2):137–145
8. Elavarasi C, Shanmuganantham T (2018) Multiband SRR loaded koch star fractal antenna. Alex Eng J 57(03):1549–1555
9. Nayna TFA, Baki AKM, Ahmed F (2014) Comparative study of rectangular and circular microstrip patch antennas in X band. In: International conference on electrical engineering and information & communication technology (ICEEICT), Dhaka, Bangladesh

10. Kaushal D, Shanmuganantham T (2017) A vinayak slotted rectangular microstrip patch antenna design for C-band applications. Microw Opt Technol Lett 59(8):1833–1837
11. Kaushal D, Shanmuganantham T (2017) High frequency slotted robot patch antenna for multiple space applications. In: International conference on antenna test and measurement society (ATMS), Hyderabad, India
12. Kaushal D, Shanmuganantham T Design of multi utility high frequency dual band slotted android logo patch antenna using coaxial feed. In: International conference on antenna test and measurement society (ATMS), Hyderabad, India

Assistive Technology for the Blind

Aditya Nagesh, Akshara S. Vijay, Nabeel Muhammed Salim, M. C. Vaisakh, and Jisha John

Abstract Individuals who are totally visually impaired or have weakened vision have a troublesome time exploring outside the spaces that they are acquainted with. Truth be told, physical adaptation is one of the greatest difficulties for visually impaired individuals. Voyaging or simply moving down a busy road can be challenging. Due to this, numerous individuals with low vision will like to go with a companion or relative while exploring new places. Likewise, blind individuals must remember the objects in their home condition. Items like beds, tables, and seats must not be moved without notice to avoid mishaps. In the event that a visually impaired individual lives with others, every individual from the family unit must be steadily about keeping walkways clear and all things in their assigned locations. This project addresses an assistive technology, designed and developed to facilitate conversion of the real-world scenes to voice. The system contains cameras. The view at their front is captured by the camera. This image captured by the camera is given to You Only Look Once (YOLO) to identify objects and then the description is generated using long short-term memory (LSTM) and Inception V3 model. It is converted to audio. This audio is output through the speakers. The created framework is viewed as a stage forward toward the progressions in Electronic Travel Aids, and should add to the improvement of the life of people with vision misfortune.

Keywords YOLO · LSTM · InceptionV3 · Pyttsx3 · Image captioning

1 Introduction

Individuals who are visually impaired may be born with vision loss or develop a visual impairment later in life as a result of an accident or eye disease. A person who is visually impaired has a decreased ability to see, even with corrective lenses, that adversely affects his visual access or interferes with processing visual information.

A. Nagesh · A. S. Vijay · N. M. Salim · M. C. Vaisakh · J. John (✉)
Mar Baselios College of Engineering and Technology, Trivandrum, India
e-mail: jisha.json@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_25

This paper addresses an assistive technology, designed and developed to facilitate conversion of the real-world scenes to voice.

There are many related works in literature. The obstacle avoidance system using haptics and laser rangefinder also known as white cane is well known for its simplicity. But for people who have some additional disability, a wheel chair is also recommended. A virtual white cane [1] is introduced which uses a laser rangefinder to scan the environment. A haptic interface is used to present this information to the user. The virtual white cane enables the user to poke at obstacles that are meters away without actually touching the obstacles. The interface enables a simple interaction. But the limitation is in determining precise location.

In [2], a Path Force feedback belt is presented. This enables blind people to navigate on the road. There are mainly three main components of the force feedback belt design; these are the main unit with two dual video cameras, power supply which is packed into a pocket, and the belt to be worn around the user's waist. A number of cells are present in the belt to give a feedback to the user. Two video cameras are used that take video stream and a 3D model of the user's surrounding area is generated. The main environment features like borders or walls are extracted when the surrounding is tracked. On the basis of extracted features, signals are send to the force feedback belt's corresponded cells. The right path will be shown by the vibration of corresponding cells and each feature has its own vibration pattern. Range is too small and more training has to be done. Smart cane is an assistive technology for the blind which was originally proposed by the students of Central Michigan University. The development of the product was studied by Wahab et al. for object detection and producing precise instructions for navigation [3]. The developed system comprises of a fuzzy controller, ultrasonic sensors, and servo motors in order to detect objects facing the user and provide instructions accordingly via voice messages or mechanical vibrations to guide the visually impaired people. The limitations of the system include cost, inaccuracy in sensor output when it comes to water detection, and lack of power supply meter.

Silicon Eyes is an assistive navigator for the visually impaired people [4]. The system helps users to detect their current location, using technologies like Global System for Mobile Communications (GSM) and Global Positioning System (GPS), and then guides them accordingly using haptic feedback. The user is also provided with the current time, date, and also the color of the objects they are facing using voice outputs. The proposed device is integrated within a silicon glove which can be worn by the user. The limitations of the system include the lack of real-time response, low accuracy of the GPS receiver at high altitudes, inefficient haptic feedback, and limited memory. In the obstacle avoidance using auto-adaptive thresholding method [5], the system detects and calculates the distance between the obstacle and the user. It uses a notebook computer, earphone, and Microsoft Kinect camera. Kinect sensor transfers data to the notebook computer. The depth image is divided into three areas (left, right and middle) and processed to find obstacles in front. Beeps are generated in the earphone when an object is in the range of 150 cm. When the obstacle is in the range of 100 cm, a voice recommendation is given to the user suggesting him/her to take left, right or middle path. Disadvantage is that the system cannot detect and differentiate objects if the distance is 250 m.

Eye substitution is an embedded device developed for vision impaired people [6]. It is TI MSP 430G2553 microcontroller and mainly implemented using Android application. The application uses GPS, improved GSM, and General Packet Radio Services (GPRS) to get the location of the person and generate better directions. The device consists of an ultrasonic sensor that sends a sequence of pulses which reflects in presence of an object, which is captured by a receiver. The microcontroller processes and activates the motors by sending pulse width modulation. The device is light and adaptable. The usage of two sensors helps in overcoming narrow cone angle. The main drawback is that it is limited only to android.

In the remote guidance system for the blind, there are two parts—a mobile unit on the user's side and a remote operating terminal [7]. The mobile unit comprises of the camera, GPS receiver, and the headphone. The camera records the video around it and sends it to the remote operating terminal via the GSM and the Internet, along with the GPS data. The operator on the other side directs the blind person through audio signals which is send back through the same link. The user will be able to hear with the help of the headphones attached. This system always needs an operator on the other end for the user to travel. Also the distance between both should be within a particular limit. In the wearable navigation aid for blind using vibrotactile information transfer system, common mobile phone vibrators are placed in different parts of the body, with different vibration times and vibration intensities [8]. The instructions to avoid collisions were transmitted to the person via vibrotactile signals on the belt. The vibration on the left indicates the user to move to the left and signal on the right indicates the user to move to the right. The instructions where provided by a second person. As the above-mentioned work, even this system needs an aid of a second person to operate. The infrared system—Pathfinder, works when the user comes in contact with any obstacle [9]. The system emits a signal to the user which the user can consider or not. The main issue with this system is distance. Only when the user is in a particular distance from the object, will the signal be emitted, also the user's device and the object should be aligned.

In this paper, a novel assistive technology is proposed for the blind to visualize the scene in front. Section 2 describes the proposed methodology and the detailed result and analysis are provided in Sect. 3 followed by the conclusion.

2 Methodology

The methodology mentioned in the flow diagram above is just simple as it depicts. Once the blind person hears a sound around or he in necessity to know what is happening around, it is just a matter of click that can help the person to meet the goal. First, the person needs to rotate his head almost to place the line of sight toward the required scene to be detected. Next, the user has to capture the image of the scene in front. This can be done on a button click and the image will be captured. Once the image is captured, the processing begins. On passing the image through the model, description is generated and it is given as audio output to the user. It is described in the following sections.

2.1 Data Collection

Common Objects in Context Dataset (COCO) is used to train and test the model. It has 80 k train images, 40 k validation images, and 40 k test images. For each image, there are 5 descriptions. Hence, a total of 4 lakh training instances are used to train the model.

2.2 Data Cleaning

Data needs to be cleaned. The process includes removing special characters, converting all words to lower case, removing words with numbers, etc. There are around 4 lakh image captions (corpus) in the dataset. A vocabulary consisting of unique words in the corpus is created.

Large number of words occurs less than ten times. Since we are making a predictive model, we do not need to consider every word present in the vocabulary. This enables the model to turn out to be progressively make less mistakes and robust to outliers. Thus, we consider just those words which happen at least ten times in the whole corpus.

2.3 Data Preprocessing—Images

As the input to any model be given in a vector form, the input (X) to the model being images must be given in a vector form. The image is input to the You Only Look Once (YOLO) network, which detects the objects from the image. At most, eight objects are integer encoded and input to the model. Each image will have an encoded input of shape (8,) [10].

As the images need to be converted into fixed size images, before being fed into the network, we opt for transfer learning by using the Inception V3 model (convolutional neural network) created by Google Research. Inception V3 was trained on an image database organized according to the WordNet hierarchy, named ImageNet, which performs classification on over 1000 different classes of images. However, another process named automatic feature engineering is chosen, which is best suited for the model fixed-length informative vector is needed as an output and not the original output, i.e., the classes [11].

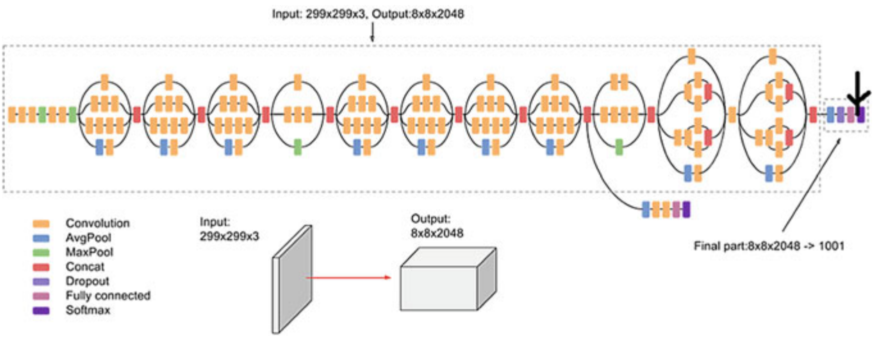


Fig. 1 Feature vector extraction (feature engineering) from Inception V3

Hence, for the convenience, the last layer called the softmax layer is removed from the model and a 2048 length vector (bottleneck features) for every image is extracted as shown in Fig. 1 [12].

Now, every input image is passed into this model for getting the corresponding 2048 length feature vector.

2.4 Data Preprocessing—Captions

Captions are being the output that the model will want to predict. So the target variables (Y) will be the captions, during the training period.

But predicting the entire caption does so not happen in a single occasion. The caption as a whole is predicted word by word. Thus, each word that is to be generated is encoded into a fixed size vector. For this purpose, two Python dictionaries namely 'wordtoix' (pronounced—word to index) and 'ixtoword' (pronounced—index to word) are created.

Every unique word in the vocabulary is represented by an integer (index). As being described above, 1652 unique words from the corpus has been generated and thus, each of these unique words will thereby be represented by an integer index between 1 and 1652.

2.5 Data Preparation Using Generator Function

This is one of the most important steps. Assume two images and their captions to train the model and the third image to test the model. Primarily, convert the two images into their corresponding feature vectors of length 2048. Let 'Image_1' and 'Image_2' be the feature vectors of the first two images, respectively.

The next step is to build a vocabulary for the first two (train) captions by adding the two tokens ‘startseq’ and ‘stopseq’ in both of them: (Assume that the basic cleaning steps have already been performed).

Each word in the vocabulary is given an index.

Now, frame it as a supervised learning problem where there is a set of data points $D = \{X_i, Y_i\}$, where X_i is the feature vector of data point ‘i’ and Y_i is the corresponding target variable.

Take the first image vector Image_1 and its corresponding caption say, ‘startseq the black cat sat on grass stopseq’. Recall that, Image vector is the input and the caption is what is needed to be predicted. The caption is predicted in the following way:

For the first time, provide the image vector and the first word as input and try to predict the second word, i.e.,

$$\text{Input} = \text{Image\_1} + \text{‘startseq’}; \text{Output} = \text{‘the’}$$

Then provide image vector and the first two words as input and try to predict the third word, i.e.,

$$\text{Input} = \text{Image\_1} + \text{‘startseq the’}; \text{Output} = \text{‘cat’}$$

This process is continued. It must be noted that, one image+caption is not a single data point but are multiple data points depending on the length of the caption.

The image is not the only input that does into the system. A partial caption is also fed to the system, which helps in the prediction of the words that follow in the sequence.

Since the sequences are processed, we employ a recurrent neural network in order to read the partial captions.

However, the actual English text of the caption is not passed, rather the sequence of indices are to be passed where each index represents a unique word. Since there is an already created index for each word, replace the words with their indices. Since batch processing is to be done, make sure that each sequence is of equal length. Hence, we need to append zero padding (0’s) at the end of each sequence.

The model uses ‘categorical cross entropy’ as the loss function and ‘adam’ as the optimizer. To update the gradients, the loss of the entire dataset is not calculated. Instead, during every iteration, the loss on a batch of data points is calculated and then updates the gradients accordingly. Since batch of data is used only at once, the entire dataset need not be stored in the memory at a time. It is necessary only to have the current batch of points in the memory. A generator function in Python is used exactly for this purpose. It is like an iterator which resumes the functionality from the point it left the last time it was called. It is trained using a fit generator with batch size as 16 and 20 epochs.

2.6 Word Embedding

The words are integer encoded. Each word in the description will be integer encoded and hence each description will have a shape of (50,) where 50 is the maximum length of description in the training set.

It uses pre-trained Global Vectors (GloVe) model to convert each word to a 200 dimension word embedding. The output of the embedding layer is passed to long short-term memory (LSTM) layer which has 256 time steps [13].

2.7 Model Architecture

The model takes three inputs.

1. The first input to the model is feature vector of the image.
   - The image is given as the input to the Inception V3 model. As the top layer of the network has been removed, the output for each image will be the feature vector in the shape (2048,).
2. The second input the model is the partial caption or partial description.
3. The third input to the model is encoded output from YOLO network (Fig. 2).

The feature vector is passed through a dense layer of 256 nodes. The output will hence be a 256 dimension.

All three 256 dimension outputs are added and passed to another dense layer of 256 nodes. The output layer is a dense layer with 6256 nodes and uses a 'softmax' activation function.

The output of each image will have a shape of (50, 6256). That is, each word is represented using an array of dimension 6256 where each value in the array ranging between 0 and 1 and the sum of all the values in the array being 1. The index of the highest probability is the integer encoding of the predicted word. The reverse mapping of the index is used to get the word.

The greedy search approach is used to predict the complete description, where the partial description is repeatedly passed to the model until 'stopseq' is predicted (Fig. 3).

2.8 Inference

Data has been prepared and model has been trained. As a final step, it is to be tested (infer) against new images, i.e., how to generate a caption for a new test image.

Input: Image vector + "startseq" (as partial caption)

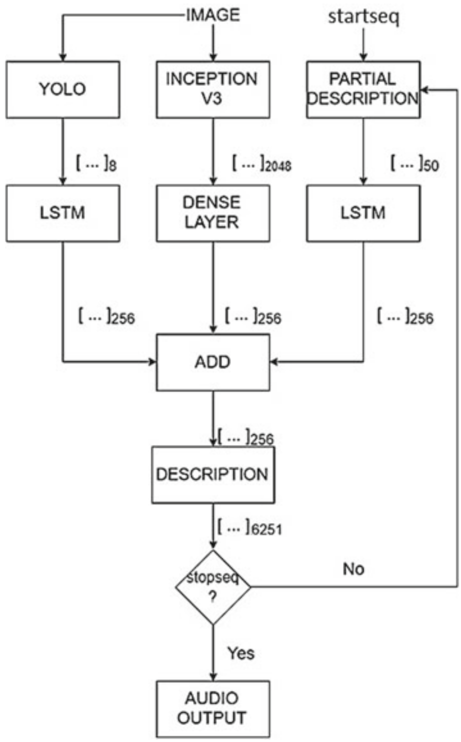


Fig. 2 Model architecture

The model avidly selects the word which has the maximum probability while given the inputs-feature vector and partial caption. This is called as maximum likelihood estimation (MLE), i.e., the word which is most suitable in describing the given input, according to the model. Sometimes this method is also known by the name Greedy Search, as the word for the maximum probability is searched greedily.

The model stops predicting when either of the two below-mentioned conditions is met:

- When encountered an 'stopseq' token meaning that the model thinks that the end of the caption is met and the caption is complete.
- When the maximum threshold of the length of the caption is met.

The model on meeting any of the above conditions, it breaks the loop and the generated caption is stated as the output.

Layer (type)	Output Shape	Param #	Connected to
input_5 (InputLayer)	(None, 50)	0	
input_6 (InputLayer)	(None, 8)	0	
input_4 (InputLayer)	(None, 2048)	0	
embedding_3 (Embedding)	(None, 50, 200)	1251200	input_5[0][0]
embedding_4 (Embedding)	(None, 8, 200)	1251200	input_6[0][0]
dropout_4 (Dropout)	(None, 2048)	0	input_4[0][0]
dropout_5 (Dropout)	(None, 50, 200)	0	embedding_3[0][0]
dropout_6 (Dropout)	(None, 8, 200)	0	embedding_4[0][0]
dense_2 (Dense)	(None, 256)	524544	dropout_4[0][0]
lstm_2 (LSTM)	(None, 256)	467968	dropout_5[0][0]
lstm_3 (LSTM)	(None, 256)	467968	dropout_6[0][0]
add_1 (Add)	(None, 256)	0	dense_2[0][0] lstm_2[0][0] lstm_3[0][0]
dense_3 (Dense)	(None, 256)	65792	add_1[0][0]
dense_4 (Dense)	(None, 6256)	1607792	dense_3[0][0]

Total params: 5,636,464
Trainable params: 3,134,064
Non-trainable params: 2,502,400

Fig. 3 Model summary

2.9 Audio Output

A text-to-speech conversion library in Python, pyttsx3 is used. This application invokes the pyttsx3.init() factory function to get a reference to a pyttsx3.Engine instance. During construction, pyttsx3.driver is initialised be the driver.

DriverProxy object is responsible for loading a speech engine driver implementation from the pyttsx3.drivers module. After construction, an application uses the engine object to register and unregister event callbacks; produce and stop speech; get and set speech engine properties; and start and stop event loops.

3 Result and Analysis

To verify the working of the above-described system, several real-time images were given as input to the model. Some of the results are given in Figs. 4, 5 and 6. Most of the results were observed to be correct. That is, most of the scenes were interpreted and correct description was generated. But some of the scenes might not be captioned exactly but the overall semantics of the scene that is the meaning of the scene will

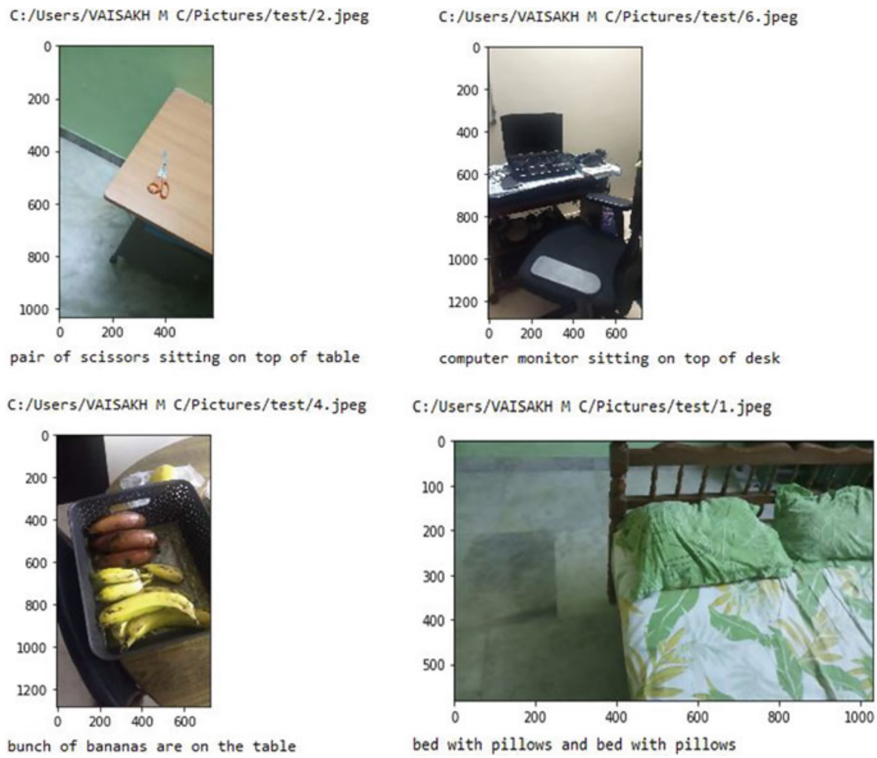


Fig. 4 Output screenshots of more accurate results

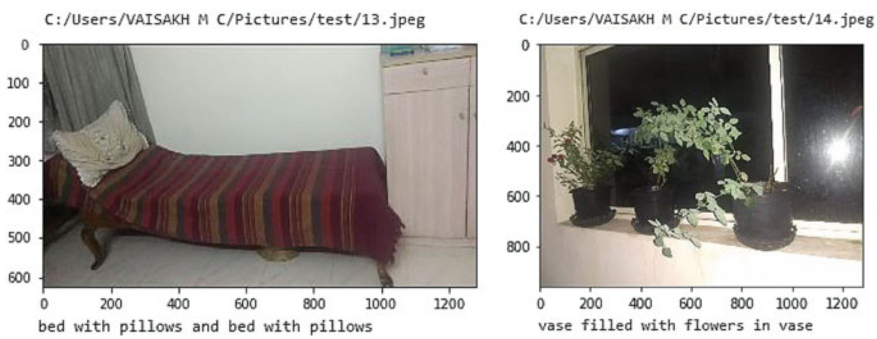


Fig. 5 Output screenshots of relatively acceptable results

be preserved. Above all, there is a good chance for erroneous description of the scene. The testing images should be meaningfully related to the training images. For instance, on the off chance that we train our model on the pictures of dogs, cats, and so on, we should not test it on pictures of waterfalls, air planes, and so forth. This

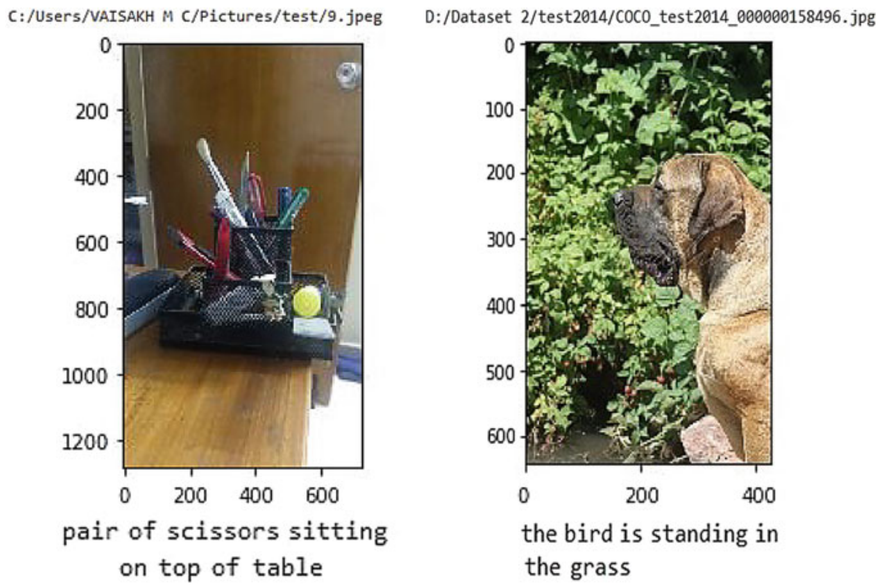


Fig. 6 Output screenshots of erroneous results

is a scenario where the training set and testing set will be entirely different. In such cases, no machine learning algorithm will give good results.

The testing images should be meaningfully related to the training images. For instance, on the off chance that we train our model on the pictures of dogs, cats, and so on, we should not test it on pictures of waterfalls, air planes, and so forth. This is a scenario where the training set and testing set will be entirely different. In such cases no machine learning algorithm will give good results.

4 Conclusion and Future Work

The world is in an amazing time in terms of finding technological solutions for people who have lost their sight. Each day seems to bring a new medical or electronic solution, making the world more accessible for people who are blind or visually impaired. A variety of tech-fueled solutions are now enabling the visually impaired to experience their world in a more customized, intuitive, and independent way, by creating solutions that are low cost and accessible to people living with some form of visual impairment. The mentioned assistive technology will be a real breakthrough in the people's life. As a future plan, we can implement this into identifying the emotions of listeners and also into measuring distance. This is beyond the current scope and can be hopefully a matter of concern in the emerging technological world.

References

1. Ahlmark DI, Fredriksson H, Hyyppa K (2013) Obstacle avoidance using haptics and a laser rangefinder. In: Proceedings of the 2013 IEEE workshop on advanced robotics and its social impacts (ARSO), Tokyo, Japan, 7–9 Nov 2013
2. Oliveira JF (2013) The path force feedback belt. In: Proceedings of the 2013 8th international conference on information technology in Asia (CITA), Kuching, Malaysia, 1–4 July 2013
3. Wahab A, Helmy M, Talib AA, Kadir HA, Johari A, Noraziah A, Sidek RM, Mutalib AA (2011) Smart cane: assistive cane for visually-impaired people. Int J Comput Sci Issues
4. Prudhvi BR, Bagani R (2013) Silicon eyes: GPS-GSM based navigation assistant for visually impaired using capacitive touch braille keypad and smart SMS facility. In: Proceedings of the 2013 world congress on computer and information technology (WCCIT), Sousse, Tunisia, 22–24 June 2013
5. Saputra MRU, Santosa PI (2014) Obstacle avoidance for visually impaired using auto-adaptive thresholding on Kinect's depth image. In: Proceedings of the IEEE 14th international conference on scalable computing and communications and its associated workshops (UTC-ATC-ScalCom), Bali, Indonesia, 9–12 Dec 2014
6. Bharambe S, Thakker R, Patil H, Bhurchandi KM (2013) Substitute eyes for blind with navigator using android. In: Proceedings of the India educators conference (TIIEC), Bangalore, India, 4–6 Apr 2013, pp 38–43
7. Baransk P, Polanczyk M, Strumillo P (2010) A remote guidance system for the blind. In: The 12th international conference on e-Health networking, applications and services, Lyon, France, 1–3 July 2010
8. Adame MR, Yu J, Möller K, Seemann E (2013) A wearable navigation aid for the blind people using a vibroctile information transfer system. In: 2013 ICME international conference on complex medical engineering, Beijing, China, 25–28 May 2013
9. Whitney GJ (1998) New technology assistive devices to assist blind and partially sighted people to travel within and use the built environment. In: 9th international conference on road transport information and control (Conf. Publ. No. 454), London, UK, 21–23 Apr 1998
10. https://towardsdatascience.com/yolo-you-only-look-once-real-time-object-detection-explained-492dc9230006
11. https://towardsdatascience.com/a-simple-guide-to-the-versions-of-the-inception-network-7fc52b863202
12. https://towardsdatascience.com/image-captioning-with-keras-teaching-computers-to-describe-pictures-c88a46a311b8
13. https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Fractal-Based Patch Antenna Design for Multi-band Applications

K Nimitha, Sanoj Viswasom, and S. Santhosh Kumar

Abstract A 1 × 3 linear Koch Snowflake antenna array is proposed for multi-band applications. The Koch array has a gain of 1.2 dB, 4.38 dB, and 5.22 dB at 5.4 GHz, 7.7 GHz, and 9.8 GHz, respectively. The number of resonant frequency of the Koch fractal increases as the number of iteration increases. The design and simulation of the Koch antenna were done in AnSoft HFSS and MATLAB. The antenna parameters, such as peak gain, bandwidth, and return loss, are analysed and presented for different iterations. The return loss obtained for all Koch antennas studied is well below −10 dB which shows good radiation characteristics. The third iteration Koch fractal antenna was fabricated using FR4 substrate and tested using Vector Network Analyser. The good agreement between the simulated and measured values validates the proposed design and satisfies the requirements for multi-band applications. Based on these results, the presented fractal-based patch antenna system can be considered as a best candidate for multi-band wireless systems.

Keywords Microstrip patch antenna · Fractal · Multi-band

1 Introduction

In today's world of wireless communication systems, demand for compact, multi-band, and wideband antennas have been increased multifold. The rapid growth from analog voice communication to the challenging 5G systems required high frequency, low cost, compact, multi-band antenna design to meet the system characteristics. With the features like low cost, small size, and availability in different shapes, Microstrip Patch Antennas (MPA) are widely used for such applications. MPA has great priority in wireless studies due to its planar configuration and easy integration with microstrip

K. Nimitha (✉) · S. Viswasom · S. Santhosh Kumar
College of Engineering Trivandrum, Trivandrum, Kerala, India
e-mail: nimithaharidas1224@gmail.com

S. Viswasom
e-mail: sanojv@yahoo.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_26

technologies. It has a dielectric substrate of $\epsilon_r < 10$ with a conductive patch on the one side and a metal ground plane on the other side. The maximum radiation is normal to the patch and the side lobes are negligible [1].

Since the current is mainly distributed along the edge of the radiating element, determining the physical dimensions of the antenna is very important. So changing the effective length of the radiator, antenna performance can be altered. Applying geometry modifications to a rectangular patch helps to improve the multi-band property [2]. In [2], Md. Imran Khan et al. proposed a multi-band patch antenna for WiMax, WiFi, WLAN, and satellite applications. This antenna structure comprises of two L shaped, three I shaped, and one F shaped patch operating in three frequency bands, namely (2.10–2.60 GHz), (3.02–5.82 GHz), and (7.00–9.50 GHz). A $19 \times 19 \times 0.787\,\text{mm}^3$ sized dual band patch antenna proposed by Yassine Jandi et al. possessed good radiation characteristics and is able to operate at 5G system bands (10.15 and 28 GHz) [3]. Therefore to achieve high gain and multiple resonant frequencies, fractal concept is introduced to the patch geometry and thereby eliminating the limitations of MPA [4].

The term fractal was introduced by Benoit Mendelbet which represents the family of complex shapes having self-similarity or self-affinity. To minimise the antenna size while keeping radiation efficiency high, fractal antennas are widely suggested [5]. Since their large electrical length is effectively packed into small areas, they are also known as structures with infinite perimeter and constant area. Due to reduction in size, area and mass, fractal antennas provide attractive solutions for single small antenna needed for different frequency ranges (multi-band operation) [6]. There are many fractal shapes: Koch curve, Sierpinski carpet, Minkowski curve, Hilbert fractal, etc. In Koch fractal, multi-band operation is achieved when coupling between sharp angles produce different current paths. In [7] Yu et al. designed a Koch Snowflake broadband antenna for wireless applications. The system combines the advantage of trapezoidal CPW-fed along with the Koch features. A Koch loop fractal antenna is experimented by R. Jothi Chitra and V. Nagarajan, using an octagonal shaped substrate over 2–30 GHz is presented in [8]. Mohsen Khalily et al. used parasitic patches stacked on the top of each inset-fed patches for Bandwidth enhancement and high gain in [9]. This mm Wave microstrip array have shown the best results for different 5G applications.

This paper presents a compact and improved Koch Snowflake antenna to support multi-band operation. The behaviour of the Koch fractal is observed on the basis of iteration order and a 1×3 linear array is designed to improve the gain and bandwidth. In the next section, design and simulation of fractal antenna using High Frequency Structure Simulator (HFSS) and MATLAB are explained. In the following sections, obtained results are compared for different iterations in terms of return loss, peak gain, number of resonant frequencies and bandwidth. Fabrication and testing results are also included. The multi-band property of Koch Snowflake antenna and bandwidth-gain improvement in Koch antenna array is concluded in the final section.

2 Proposed Work

2.1 Construction of Koch Snowflake Structure

A Koch curve can be obtained by replacing the middle third of a straight section by a bent section. In the next iteration, each edge is divided into three equal parts and middle part is substituted with a bent section. This process continues for every iteration. To generate the Koch structure, a MATLAB code is created using Iterative Function Systems (IFS). A simple Koch curve is shown in Fig. 1.

Scaling by 1/3 is represented by

$$f_1(x) = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{1}$$

The function $f_2(x)$ scales the line by 1/3, rotates 60° and translates,

$$f_2(x) = \begin{bmatrix} \frac{1}{3}\cos(\pi/3) & -\frac{1}{3}\sin(\pi/3) \\ \frac{1}{3}\sin(\pi/3) & \frac{1}{3}\cos(\pi/3) \end{bmatrix} x + \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix} \tag{2}$$

The function $f_3(x)$ scales the line by 1/3, rotates −60° and translates,

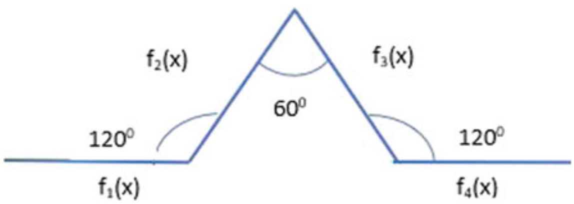
$$f_3(x) = \begin{bmatrix} \frac{1}{3}\cos(\pi/3) & \frac{1}{3}\sin(\pi/3) \\ -\frac{1}{3}\sin(\pi/3) & \frac{1}{3}\cos(\pi/3) \end{bmatrix} x + \begin{bmatrix} \cos(\pi/3) \\ \frac{1}{3}\sin(\pi/3) \end{bmatrix} \tag{3}$$

The function $f_4(x)$ scales the line by 1/3 and translates,

$$f_4(x) = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} x + \begin{bmatrix} \frac{2}{3} \\ 0 \end{bmatrix} \tag{4}$$

To generate Koch Snowflake, modifications are applied to the script of Koch curve. In Koch loop, 3 Koch curves are placed 60° to one another in a triangular shape. Only the length and iteration order was entered as inputs. The Koch Snowflake Matlab model for second iteration is shown in Fig. 2. The generated Koch loop is transferred to the HFSS environment by creating a VB Script file.

Fig. 1 Simple Koch curve



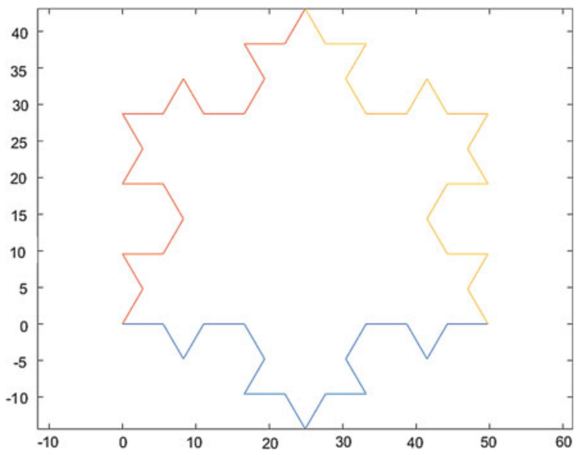


Fig. 2 MATLAB model of Koch Snowflake fractal-iteration 2

2.2 Antenna Design

The design equations used for a simple microstrip patch antenna are applied for fractal antennas also. For comparing the antenna parameters, the dimensions of patch antenna and fractal antenna should be the same. Substrate used is FR4 Epoxy of relative permittivity (ϵ_r) 4.4 and the height (h) of substrate is set to as 1.6 mm. The frequency of operation must be selected properly from the starting of design process. In this work, (f_0) is taken as 5.2 GHz. With the values of ϵ_r, h and f_0, patch antenna dimensions are calculated using the following equations.

$$L = L_{\text{eff}} - 2\Delta L \tag{5}$$

$$L_{\text{eff}} = \frac{c}{2f\sqrt{\epsilon_{\text{reff}}}} \tag{6}$$

$$\epsilon_{\text{reff}} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2}\left[\frac{1}{2f\sqrt{1 + 12\frac{h}{w}}}\right] \tag{7}$$

$$\Delta L = 0.412\frac{[\epsilon_{\text{reff}} + 0.3]\left[\frac{h}{w} + 0.264\right]}{[\epsilon_{\text{reff}} - 0.258]\left[\frac{h}{w} + 0.8\right]} \tag{8}$$

$$W = \frac{c\sqrt{2}}{2f\sqrt{\epsilon_r + 1}} \tag{9}$$

3 Simulation Results

The MPA and the proposed Koch Snowflake fractal antenna is simulated using AnSoft HFSS which works on the Finite Element Method. The antenna models drawn in HFSS and the obtained return loss and radiation pattern are presented in this section. Firstly, a rectangular microstrip patch antenna is simulated. Its results show a return loss of −38 dB and a gain of 4 dB. Patch antenna exhibits a single resonant frequency with 200 MHz bandwidth.

Figure 3a shows the HFSS structure of Koch Snowflake antenna of third iteration. The HFSS structure was obtained using a VB script file generated using MATLAB. The distribution of current will be maximum along the edges of the fractal. Since the fractal edges are less for the first and second iteration, the path for current flow is also less. So as the iteration levels increased, effective length becomes longer and number of resonant frequencies also increased. This trend can be well observed in Fig. 4. Among three different iterations, third order has better s_{11} values and resonant frequencies. The return loss values for third iteration are −10.63 dB, −13.7 dB, −22.4 dB, and −19.2 dB at resonance frequencies 4.2 GHz, 5.5 GHz, 7.8 GHz, and 9.8 GHz, respectively.

Better antenna performance can be guaranteed for a more negative return loss. A return loss value below −10 dB is sufficient for an antenna to radiate which corresponds to the Voltage Standing Wave Ratio (VSWR)value of 2. The bandwidth of an antenna is the range of frequency over which VSWR ≤ 2 or $S_{11} \leq -10$ dB. Hence, it is inferred in Fig. 4 that the reported antenna exhibits good radiation characteristics.

It is known that antenna array has got more gain and enhanced bandwidth compared to single antenna. So in order to improve the gain and bandwidth, Koch Snowflake is duplicated and a 1 × 3 linear array is obtained. The array was provided with individual ports and a 50 Ω radiation resistance. Simulation results show that the bandwidth is doubled and peak gain is also increased from 0.8 dB, 3.1 dB and 4.2 dB to 1.2 dB, 4 dB, and 5.2 dB, respectively. Figure 3b shows the simulation structure of Koch fractal array. Three-dimensional radiation pattern of single and array Koch

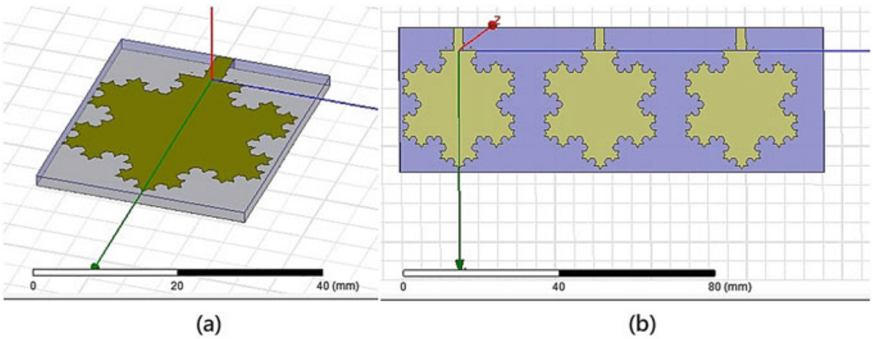


Fig. 3 Layout of third iteration Koch fractal: **a** single element **b** 1 × 3 linear array

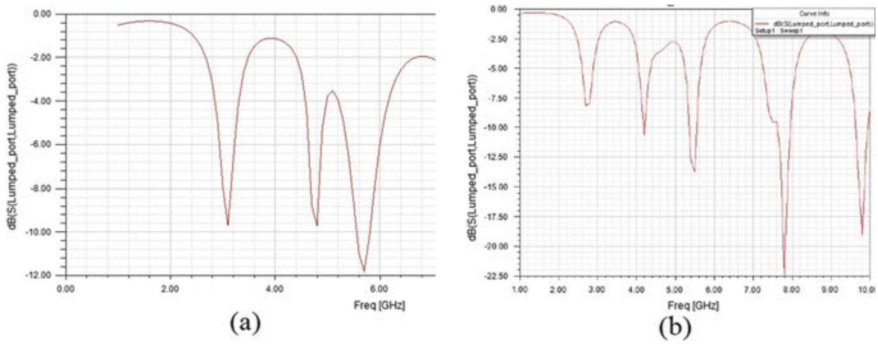


Fig. 4 Return loss of single Koch Snowflake antenna: **a** first iteration; **b** third iteration

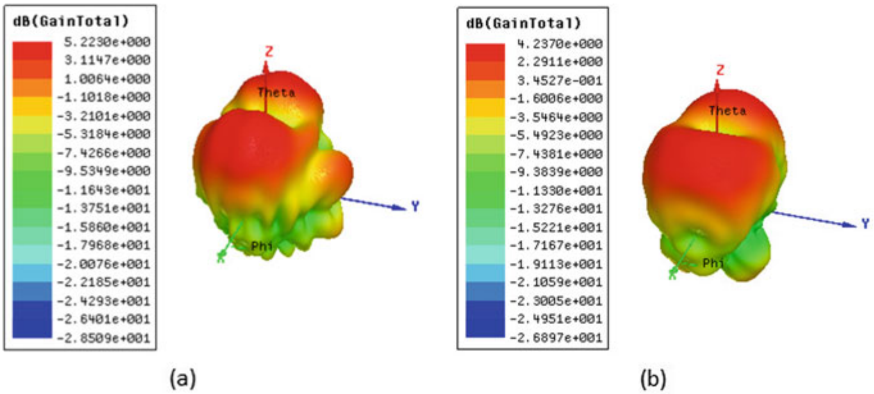


Fig. 5 3D radiation pattern at 9.8 GHz: **a** Koch fractal array; **b** single Koch fractal

fractal antenna at 9.8 GHz is shown in Fig. 5. Simulation results for Koch Snowflake fractal antenna (first, second, and third iteration) and 1 $\times$ 3 Koch Snowflake antenna array in terms of return loss, peak gain, and bandwidth are presented in Table 1.

4 Fabrication and Measured Results

A prototype of the proposed antenna was fabricated and its measured its response to verify the multi-band performance. The antenna was constructed using a 1.6 mm thick FR4 substrate having a loss tangent of 0.02 and a 30 μm thick copper layer on the two sides. The fabricated antenna and testing set up are shown in Fig. 6.

The antenna was tested using a P9373A 2 port Vector Network Analyser (VNA). SMA Female Connector (Pcb Edge Mountable) was used. The measured and simulated results are compared and observed good agreement in terms of return loss

Table 1 Antenna parameters comparison

Fractal configuration	Resonant frequency (GHz)	Bandwidth (MHz)	Return loss (dB)	Peak gain
Single—iteration 1	5.7	250.4	−11.8	1.39
Single—iteration 2	5.5	219.8	−14.7	0.65
	8.2	259.7	−18.8	4.9
	10	424.1	−22.1	3.3
Single—iteration 3	4.2	29.2	−10.63	4.237
	5.5	195.6	−13.7	0.8
	7.8	338.6	−22.4	3.87
	9.8	314.9	−19.2	4.2
Array—iteration 3	5.5	406.2	−14.9	1.2
	7.7	600	−31	4.38
	9.8	584.9	−16.6	5.22

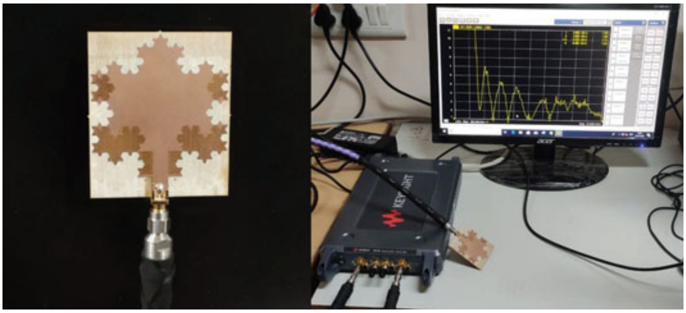


Fig. 6 Fabricated third-order Koch Snowflake antenna and measurement setup

values as shown in Fig. 7. Measured return loss values are −17 dB, −24 dB, −29.5 dB and −14.2 dB at 2.6 GHz, 3.6 GHz, 5 GHz, and 6.4 GHz respectively. Bandwidths obtained for these frequencies are 132.12 MHz, 188.54 MHz, 684.97 MHz, and 221.81 MHz, respectively.

5 Conclusion

This paper has proposed a design procedure for Koch Snowflake fractal antenna and developed a 1 × 3 linear Koch fractal antenna array of third iteration. A Koch fractal is designed in MATLAB using microstrip patch antenna design equations together

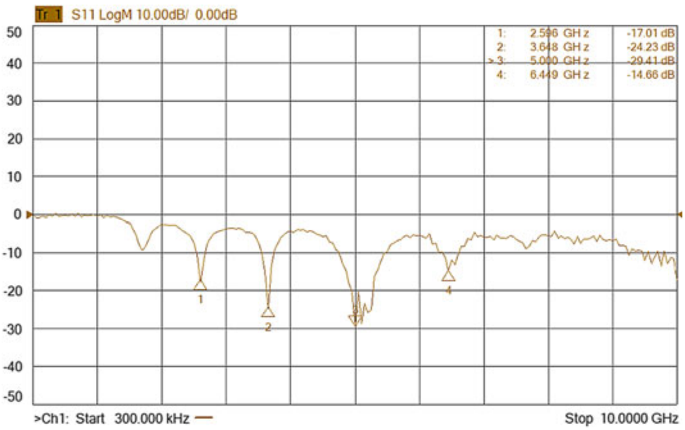


Fig. 7 Measured return loss of fabricated Koch Snowflake antenna

with IFS and is exported to HFSS. For the analysis of antenna performance, gain and return loss were plotted. It can be seen that the number of multi-bands get increased upon increasing the iteration order of Koch fractal.

The proposed antenna system has three resonant frequencies at 5.4 GHz, 7.7 GHz, and 9.8 GHz with a gain of 1.2 dB, 4.38 dB, and 5.22 dB, respectively. The return loss and gain are also improved compared to single Koch fractal antenna. Fabrication results of third iteration Koch fractal match with the simulation results. Since this fractal-based patch antenna array system ensures improved gain, multi-band and wideband features, it can be considered suitable for multi-band wireless applications.

References

1. Verma S, Mahajan L, Kumar R, Saini HS, Kumar N (2016) A small microstrip patch antenna for future 5G applications. In: 2016 5th international conference on reliability, Infocom technologies and optimization (trends and future directions) (ICRITO). IEEE, pp 460–463
2. Khan MI, Ahasan Kabir M, Hira A (2017) A simple multiband patch antenna for application in wireless communication. In: 2017 international conference on electrical, computer and communication engineering (ECCE). IEEE, pp 615–618
3. Jandi Y, Gharnati F, Oulad Said A (2017) Design of a compact dual bands patch antenna for 5G applications. In: 2017 international conference on wireless technologies, embedded and intelligent systems (WITS). IEEE, pp 1–4
4. Maharana MS, Mishra GP, Mangaraj BB (2018) Design and simulation of a Sierpinski carpet fractal antenna for 5G commercial applications. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 1718–1721
5. Mondal T, Maity S, Ghatak R, Chaudhuri SRB (2018) Compact circularly polarized wide-beamwidth fern-fractal-shaped microstrip antenna for vehicular communication. IEEE Trans Veh Technol 67(6):5126–5134
6. Hwang KC (2007) A modified Sierpinski fractal antenna for multiband application. IEEE Antennas Wirel Propag Lett 6:357–360

7. Yu Z, Yu J, Zhu C, Yang Z (2017) An improved koch snowflake fractal broadband antenna for wireless applications. In: 2017 IEEE 5th international symposium on electromagnetic compatibility (EMC-Beijing). IEEE, pp 1–5
8. Jothi Chitra R, Nagarajan V (2016) Design and development of koch fractal antenna. In: 2016 international conference on communication and signal processing (ICCSP). IEEE, pp 2294–2298
9. Khalily M, Tafazolli R, Xiao P, Kishk AA (2018) Broadband mm-wave microstrip array antenna with improved radiation characteristics for different 5G applications. IEEE Trans Antennas Propag 66(9):4641–4647

Analysis of Light Field Imaging and Segmentation on All-Focus Images

Parvathy Prathap and J. Jayakumari

Abstract Image analysis is one of the fastest-growing research areas, and within it, light field or plenoptic image processing has gained great popularity in the recent years. Plenoptic images have got the inbuilt capability of imaging an object at slightly variant viewing angles which opens up possibilities for a deeper image analysis. A light field image is basically constituted with an array of sub-aperture images that have been captured using an array of lenses. Segmenting an image into its constituent parts of interest is very crucial in an image analysis point of view. Its efficiency influences the subsequent image processing stages. Various segmentation algorithms for conventional images have been developed over the recent years. This paper analyses the field of light field imaging in terms of its acquisition and discusses the various application areas and challenges pertaining to it. It also throws light on some of the research possibilities in this field. Finally, few segmentation algorithms are analysed on how it works when applied to light field all-focus images. The segmentation parameters like jaccard, dice, sensitivity, accuracy and specificity are being analysed here. From among the available sub-aperture images, the central view all-focus image has been chosen to evaluate the segmentation results. The segmentation is being evaluated on light field saliency dataset (LFSD) which has the ground truth data of segmentation corresponding to all-focus light field images.

Keywords Light field · Plenoptic · Sub-aperture images · Segmentation · Focal stack

---

P. Prathap (✉) · J. Jayakumari
Department of ECE, Mar Baselios College of Engineering and Technology, APJ Abdul Kalam Technological University, Trivandrum, Kerala, India
e-mail: parvathy00@gmail.com

J. Jayakumari
e-mail: jayakumari.j@mbcet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_27

1 Introduction

The light rays that are present around us carry enormous information about the objects around us. Hence, proper utilization of the information present in it would open up lots of opportunities in exploring the world around us. Imaging techniques capture this information present in the light rays to a great extend but certain drawbacks of conventional imaging methods limit the amount of information extracted from these light rays. One of the major limitations present here is that the conventional techniques capture only the intensity of light rays that is being reflected back from an object while imaging it. But the fact is that, apart from the intensity information, if we are able to capture the directional information of the light rays too, then that would serve as a key to explore the enormous amount of three-dimensional information that is encapsulated in these light rays. Taking into consideration this key requirement, the light field imaging technology has now become the latest emerging trend in capturing the three-dimensional volume of information present around us. Apart from capturing the intensity information, light field imaging technology (also known by the name plenoptic imaging) captures the directional information also which helps in a better three-dimensional visualization of the objects around us. Another attraction in light field photography is the ability to change the point of focus after the image is being taken. In simple words, we can say that the act of refocusing the image regions can be done on the software side also. This opens up enormous opportunities for light field imaging in applications, wherein images of moving objects are to be captured and analysed later.

Even though valuable information gets captured in plenoptic images, the extraction of this heavy information content studded in these images is a very challenging task for the researchers to address. This extraction process is limited by various factors like the decreased spatial resolution, complications owing to the higher dimensionality of data, difficulty in proper extraction of images from the raw files. A proper analysis in this respect and the development of efficient methods to overcome these inherent limitations of plenoptic images is really necessary at this point of time. This paper basically analyses the field of plenoptic imaging in terms of its acquisition, the inherent limitations present in them, some of the possible research opportunities in this field and simulation and performance comparison of a few segmentation algorithms on some 'all in focus' light field images that are captured by light field cameras.

The remaining content of the paper is organized as follows. Section 2 describes the basic acquisition model and representation of plenoptic images; Sect. 3 discusses the major application areas of light field photography. It also discusses the possible research scope in this field. Section 4 discusses the algorithms that are considered for evaluation here. Section 5 presents the simulation results and the performance analysis of the segmentation techniques on the all-focus light field images. Finally, conclusion and possibilities for future works are mentioned in Sect. 6.

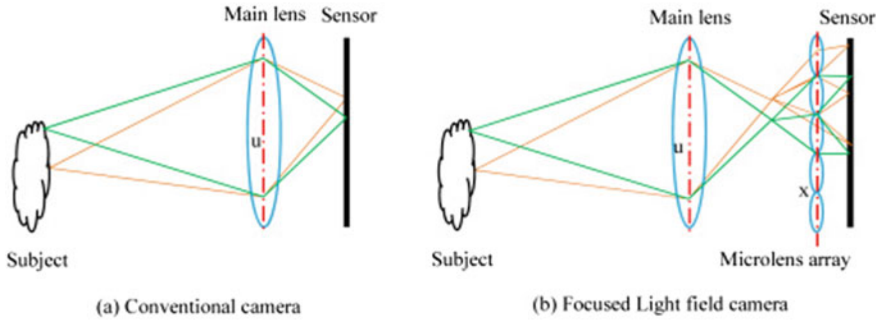


Fig. 1 Image acquisition model [1]

2 Basic Acquisition Model and Representation of Plenoptic Images

This section describes the basic acquisition model and representation of light field images.

2.1 Basic Image Acquisition Model

Figure 1 depicts how an image acquisition set-up varies in a conventional imaging method and multi-view imaging method. The basic idea of a light field image is to generate multiple views of the same scene. In order to facilitate this, a multi-lens arrangement or an array of lenses is being used to recapture the scene under different views. So, the information that is missing under one view can be made available from another view and this complementary information that is available from sub-aperture images is the greatest strength of light field imaging technology. This also helps in situations, wherein occluding structures are present. The missing details of the object of interest that is being occluded in one sub-aperture image can be recovered from another sub-aperture image.

2.2 Representation of Plenoptic Images

A plenoptic function, in its very simple form as proposed by Levoy and Hanrahan [2], can be represented as a four-dimensional function. This reduction in dimension was proposed by them considering the computational efficiency aspect and the difficulty to deal with higher dimensions. There, two planes are considered and the coordinate system is represented as, say (u, v) for the first plane and (s, t) for the second plane. So a light ray is said to intersect both the u-v plane and then the s-t plane and hence

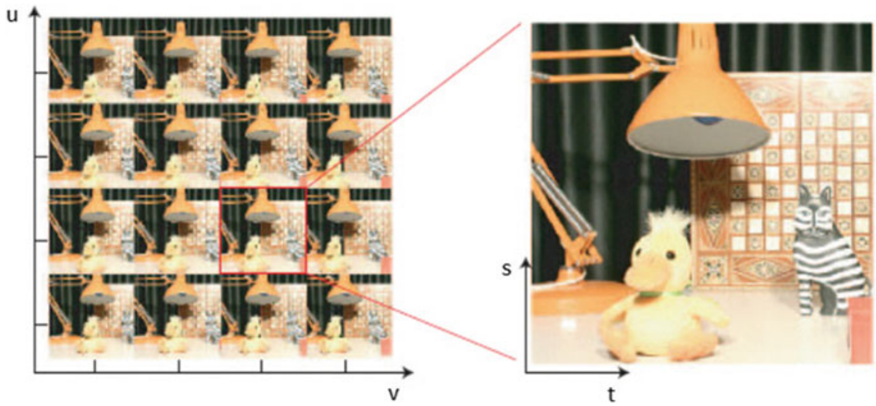


Fig. 2 Light field as an array of images [3]

can be defined using a function indicating these values, say $L(u, v, s, t)$. In normal course, the u-v plane can be considered to be the lenslet plane which indicates the position of the lens and the s-t plane gives the spatial coordinate information and both these planes together help to capture the spatial and angular information from an image. In short, we can say that the 4-D light field image can be thought of as an array of many 2-D images. The same has been shown in Fig. 2.

3 Applications and Research Scope in Plenoptic Imaging

This section discusses some of the potential application areas of light field imaging. Light field imaging can be considered as an effective technique in areas, where the concept of synthetic-aperture imaging is to be applied. This technique has got the capability to handle occluding structures with the dense information content it has. As discussed above, this imaging technique also has the capability to refocus the image after its acquisition stage and also to produce an 'all in focus' image by taking into account the complementary information present in various sub-aperture images. These capabilities make it a good choice for imaging non-stationary objects. Due to its capability to provide 3-D effects, it can be used to provide special effects to movies and photographs. It can also be used in underwater imaging, robotics, 3-D scanning, autonomous vehicle development, etc., where an accurate recording of the depth information of the scene is also desired. Light field telescopes have been developed now to enhance the capabilities of conventional telescopes. Using this technology in conjunction with conventional microscopy [4] and endoscopy methods provides newer heights of exploration in medical imaging domain too. Researchers have found its application in areas like phenotyping too. With its immense inherent capabilities of 3-D visualization, this becomes a very good choice for any object recognition problem.

The inherent capabilities of light field images come hand in hand with some inherent drawbacks too. As discussed, the light field image brings in the concept of angular resolution too. But this is being introduced at the cost of a decline in its spatial resolution. Hence, spatio-angular resolution trade-off analysis and resolution enhancement are areas, where effective methods need to be introduced [5]. The next issue in light field imaging arises because of the enormous sizes of light field image files owing to the capture of data in higher dimensions. The light field image, which is basically a collection of sub-aperture images, needs to be analysed and redundant data needs to be eliminated. For this purpose, effective methods to analyse the disparity between the various sub-aperture images need to be researched upon. The information that is redundant across multiple sub-aperture images needs to be effectively identified in this regard. Thus, light field compression becomes another research area with a great scope [6]. In addition to this, using the depth information that is inherently captured by this imaging technique, efficient depth map estimation methods can be formulated. These depth maps would be more reliable and accurate than those formed from 2-D images [7]. Occlusion handling techniques using these kinds of images would be more efficient because of the complementing information available across various multi-view images. This information helps in the proper estimation of the occluding structure and hence helps in revealing the underlying objects. Another less researched area in the field of plenoptic images is motion de-blurring in these kinds of images. Blind de-blurring technique [8] is being tried upon but proper degradation estimation and restoration are still unexplored. Post-capture refocusing, a feature that distinguishes light field imaging from conventional imaging, is another area that has scope to be worked on. The inherent refocusing software that companies like Lytro and Raytrix provide can be improvised upon to bring in more capabilities to analyse the object of interest. Segmentation of light field images is another area to be explored as conventional 2-D image segmentation methods alone may not serve the purpose of segmenting objects by bringing in a 3-D view. Other application-oriented works such as colour correction, rain removal of such image are also being researched on owing to its applications in fields like image forensics. Dansereau [9] has provided a set of tools that would be helpful in analysing the raw LFP or LFR files. It allows visualization, decoding, rectification, colour correction and basic filtering of light fields.

4 Segmentation Analyses of All-Focus Light Field Images

This section is an analysis of some of the basic image segmentation algorithms that are being used for conventional imaging.

The algorithms that are being evaluated here are the Otsu segmentation method, k-means-based segmentation and active contour-based segmentation. A brief discussion on the basic concepts of these algorithms is given below.

4.1 K-Means

This is one of the commonly used clustering methods which eventually convert the samples into different clusters based on the distance between each sample and the selected centroids [10]. 'K' can take different values depending on the number of clusters required. In this case, we have segmented the image into two categories of pixels. It is basically done as two phases. The first phase is often called by the name 'batch updates' where different iterations are done to reassign points to the cluster centroid that is nearest to it. Here, in general we can say that if the collection of various centroids in the set C is denoted by 'c_i', then the assignment of each data point 'x' to the cluster is based on the equation

$$\underset{c_i \Xi C}{\arg\min}\ \mathrm{dist}(c_i, x)^2 \tag{1}$$

where dist(.) indicates a standard distance measure. In the second phase, the data points are reassigned after the new centroid is calculated. If S_i indicates the data point assignment for each of the ith cluster centroid, then the updated centroid can be indicated as

$$c_i = \frac{1}{|S_i|} \sum_{x_i \in S_i} x_i \tag{2}$$

In the distance calculation step, various distance measures can be considered as the distance metric between two data points. They are the squared Euclidean distance, the city block distance, Hamming distance, etc. The number of iterations for this algorithm can be decided based on the requirement, the more the number of iterations, the better would be the clustering results normally. The computation time constraints also need to be considered in this regard. The process is usually stopped when the newly calculated centroid starts remaining unchanged or when the desired number of iterations has reached.

4.2 Otsu

Segmentation using Otsu's method [11] is based on the threshold that helps to reduce the intra-class variance value of the thresholded white-and-black pixels. It is being formulated as a discriminant analysis method where a particular criteria function is being chosen as a factor indicative of the statistical separation between the classes. The threshold is chosen such that the variance within a class is minimized and that between two classes are maximized. It basically works on the histogram of the image.

The weighted within-class variance basically is given as

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \tag{3}$$

Here $q_1(t)$ and $q_2(t)$ indicate the class probabilities of classes 1 and 2, respectively. The individual class variances can be indicated as

$$\sigma_1^2(t) = \sum_{i=1}^{t} [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \tag{4}$$

$$\sigma_2^2(t) = \sum_{i=t+1}^{I} [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)} \tag{5}$$

Here $\mu_1(t)$ and $\mu_2(t)$ are the class means corresponding to classes 1 and 2. Here

$$q_1(t) = \sum_{i=1}^{t} P(i) \tag{6}$$

$$q_2(t) = \sum_{i=t+1}^{I} P(i) \tag{7}$$

The total variance can be expressed as a sum of this intra-class and inter-class variances. Maximizing the interclass variance would then automatically lead to a decrease in the intra-class variance which ultimately helps to formulate an efficient threshold value.

4.3 Active Contour

Contour models generally describe the boundaries of various shapes in an image. Active contours [12], alternatively known as snakes, are a deformable model and are capable of compensating for the effects of missing boundary information. Here object boundaries are detected by using evolving curves. Given two points on a curve, the idea is to basically find a curve that connects these two points. To initiate this process, a mask needs to be defined to declare the initial state of the contour. The initial stage is to define an energy function, and this function could naturally take small values for data points located along the actual boundary. The behaviour of this algorithm tends to depend upon the choice of the energy function. Initial boundary points are got, and then the energy at boundary points is compared with the energy at points along the neighbourhood of boundaries. For each boundary point, boundary is moved on to the neighbouring point that has got the lowest energy. One iteration is said to have been completed, once the above operation is completed on all points of the boundary. This process needs to be repeated until the boundary points become static or do not tend to change. At this point, we can say that the convergence of the snake has happened and the segmentation has been completed.

5 Simulation Results and Inferences

This section presents the results of simulations and analyses them to draw inferences. Analysis was done using MATLAB R2019a version. The idea is to throw light on how some of the segmentation methodologies work on light field images. These segmentation algorithms were applied on light field images acquired from light field saliency dataset (LFSD) [8]. This data set provides focal stack images of light fields captured using the Lytro light field camera. The data set also contains the corresponding all-focus images for every scene that has been captured. The ground truth data is available for the all-focus images, and these are being used here for analysis. The analysis is done based on the parameters jaccard, dice, sensitivity, specificity and accuracy using 25 light field images from the data set. Table 1 shows the average values of the performance metrics obtained during this evaluation using various segmentation methodologies. Figure 3 shows a comparison of various performance metrics across different algorithms that are being discussed here.

Normal 2-D segmentation cannot be directly applied to the field of light field image processing, since data from more viewing angles also needs to be incorporated into the segmentation process as opposed to segmentation from conventional 2-D images. But, as a starting stage to an effective 3-D segmentation model of an image, we can

Table 1 Performance parameters

Algorithm	Jaccard	Dice	Sensitivity	Specificity	Accuracy
K-means	0.48425	0.625059	0.787398	0.72438	0.7482
Otsu	0.49469	0.63243	0.80562	0.714408	0.7439
Active contour	0.42673	0.53017	0.55442	0.83845	0.7593

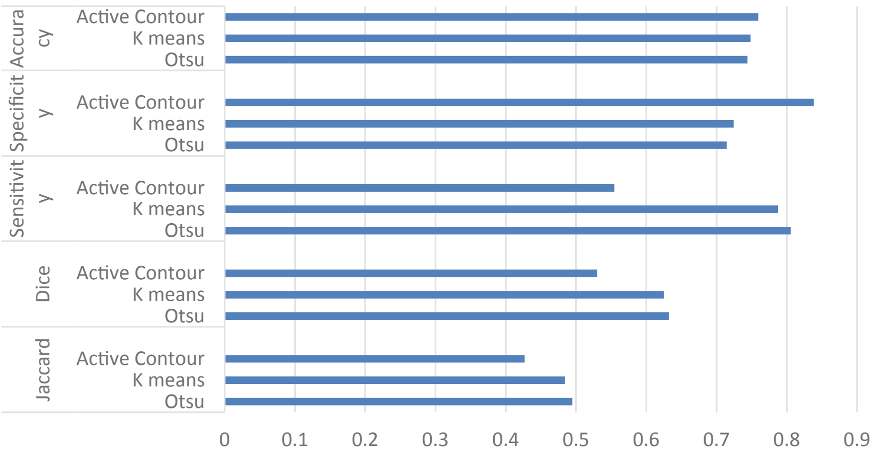


Fig. 3 Performance metric comparison

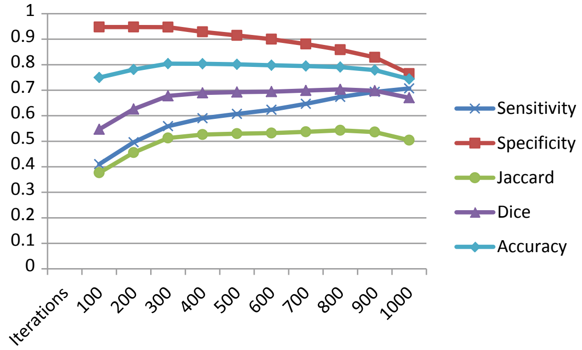


Fig. 4 Performance versus number of iterations of active contour algorithm

use an all-focus light field 2-D image and then use this in conjunction with an efficient depth estimation algorithm to extend it to a 3-D framework. It is basically with this intention that an effective segmentation method for such all-focus light field images is in demand.

From the above results, it is clear that existing algorithms do not always show an exemplary performance in terms of the metrics that are considered here for evaluation. One which is effective in detecting the true negatives (say, the active contours with pretty good specificity) are not efficient in detecting the true positives (as they have low sensitivity and Jaccard values). The active contour was initially evaluated with the default value of 100 iterations, but then the sensitivity was too low. An increase in the iterations initially had a positive effect on parameters like sensitivity, jaccard and dice. But an iteration above, say 800, did not show a significant impact on performance improvement. Hence, the algorithm was made to stop there. A plot between the number of iterations and the performance of the snake's algorithm is being shown in Fig. 4. From the graph, it can be noted that the performance improvement in the initial stages tends to cease towards the right side of the graph. Hence, further increase in the number of iterations would only increase the execution time rather than having a positive impact on performance.

Otsu's method showed above average performance in terms of the true positive rate, but it was seen that they tend to ignore many of the true-negative pixels, i.e. the background pixels were not effectively separated. The K-means algorithm (with $K = 2$) showed pretty good performance in the normal course, but in images capturing more complicated scenarios, the algorithm failed to effectively segregate object and background pixels. As the light field focal stack images tend to have lower spatial resolution compared to the conventional imaging, the existing algorithms for segmentation fail to capture certain critical information present in images depicting somewhat complicated scenarios. A similar scenario has been indicated in the following section. A light field image has been considered from the LFSD data set, and their outputs under various segmentation conditions are shown in Fig. 5. In this case, we can see that the active contour segmentation gave comparatively better results. Active contours showed better accuracy values too. But there are scenarios which

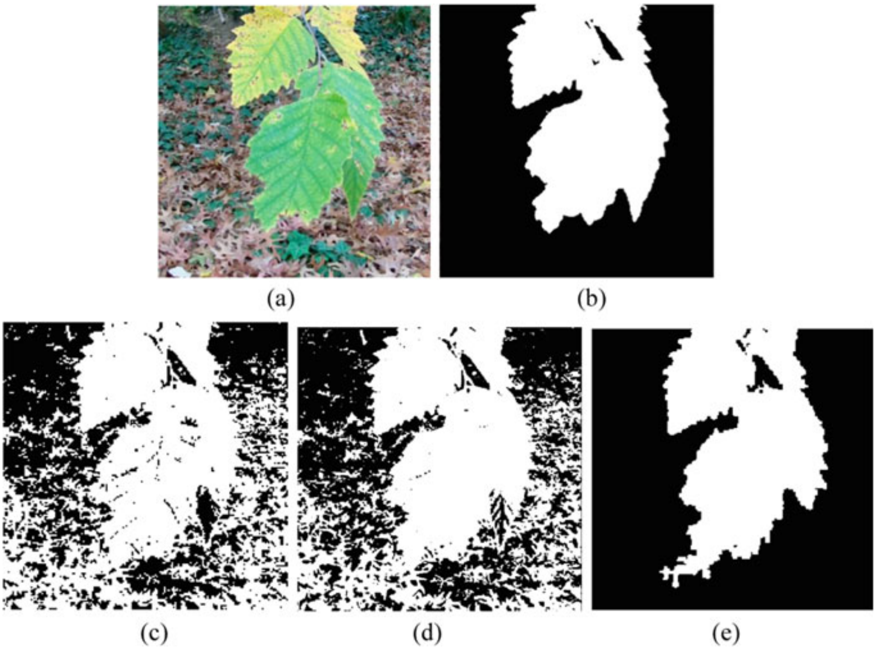


Fig. 5 **a** Input image 1, **b** ground truth, **c** Otsu segmented output, **d** K-means segmented output, **e** active contour segmented output

were not efficiently captured by any of these algorithms which further demand a more effective algorithm for such light field images. An example for this has been shown in Fig. 6.

6 Conclusion and Future Work

Light field image segmentation demands novel methods for utilizing the 3-D data available in the light field. As an initial work in this regard, it is essential to have an efficient segmentation scheme by which either the central view images or the all-focus images in the light field focal stack can be effectively segmented. An analysis in this regard shows that the performance of some of the existing methodologies degrades especially when fed with slightly challenging image scenarios. Three segmentation algorithms were evaluated here on light field all-focus images. As far as the overall performance or accuracy was concerned, the active contour algorithm showed a better performance compared to Otsu and K-means but, when it comes to segmenting complicated scenarios, a few of which that has been depicted, none of the algorithms showed a competing performance. An effective algorithm in this regard is required because this stage can be used as a foundation for a more sophisticated 3-D

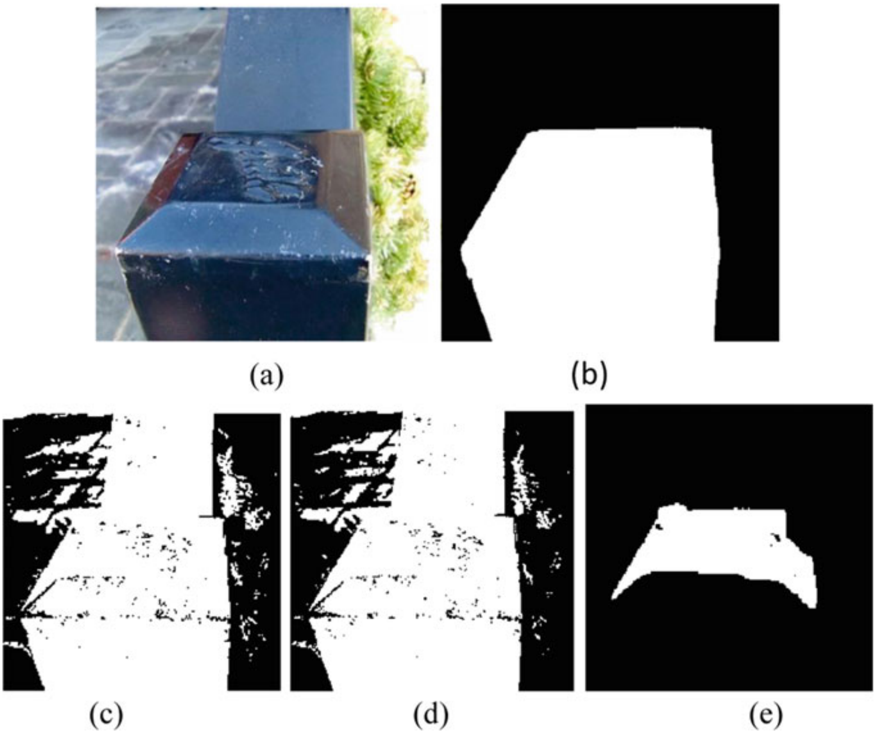


Fig. 6 **a** Input Image 2 , **b** ground truth, **c** Otsu segmented output, **d** K-means segmented output, **e** active contour segmented output

segmentation. This analysis can be used as a basis for understanding the loopholes in these algorithms, thereby throwing light on certain scenarios where it fails to achieve the desired segmented results. An improvement in this regard would be to develop an algorithm which effectively does the segmentation in areas where the existing algorithms failed. The results of this analysis would aid in designing an algorithm which achieves good values for all the performance metrics considered here. An extension in this regard would be to use this segmentation algorithm in conjunction with an efficient depth estimation framework in order to have a 3-D segmentation of the objects of interest from light field images.

Acknowledgements This work is funded by the Kerala State Council for Science, Technology and Environment (KSCSTE), Kerala, India.

References

1. Sun J et al (2016) Three-dimensional temperature field measurement of flame using a single light field camera. Opt Express 24(2):1118–1132
2. Levoy M, Hanrahan P (1996) Light field rendering. In: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp 31–42
3. Berent J, Dragotti PL (2007) Plenoptic manifolds: exploiting structure and coherence in multiview images. IEEE Sig Process Mag 24:34–44
4. Levoy M et al (2006) Light field microscopy. ACM Trans Graph (TOG) 25(3):924–934
5. Yeung HWF et al (2018) Light field spatial super-resolution using deep efficient spatial-angular separable convolution. IEEE Trans Image Process 28(5):2319–2330
6. Jia C et al (2018) Light field image compression using generative adversarial network-based view synthesis. IEEE J Emerg Sel Top Circ Syst 9(1):177–189
7. Zhu H, Wang Q, Yu J (2017) Occlusion-model guided antiocclusion depth estimation in light field. IEEE J Sel Top Sig Process 11(7):965–978
8. Chandramouli P et al (2017) Plenoptic image motion deblurring. IEEE Trans Image Process 27(4):1723–1734
9. Dansereau D (2015) Light-field toolbox for matlab (Online). http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4
10. Spath H (1985) Cluster dissection and analysis: theory, fortran programs, examples (trans: Goldschmidt J). Halsted Press, New York
11. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66
12. Lankton S (2008) Active contour segmentation. The MathWorks, Inc (online), pp 4–12

Feature Extraction Methods in Person Re-identification System: A Technical Review

C. Jayavarthini and C. Malathy

Abstract Intelligent surveillance is an emerging research area in the field of computer vision. Person re-identification is one among the tools involved in intelligent surveillance. Person re-identification is used to recognize and identify a person of interest captured by different surveillance cameras at different times and at different locations, when an input image is given. Automation of person re-identification is difficult in real time due to changes in pose, background, illumination and occlusion. Recent researchers have focused on developing discriminant, robust features, learning distance metric models or fusion of both for matching between the images of person. Our main objective is to provide the future researchers the importance of various state-of-the-art feature extraction techniques and deep learning approaches used in person re-identification, till date. Different algorithms with their strengths and accuracy percentage were summarized in a comparison table. Finally, unsolved problems in person re-id were listed that can be used as guidelines for future research.

Keywords Person re-id · Feature extraction · Discriminator · Intelligent surveillance

1 Introduction

Recently, camera-based surveillance has drawn a lot of attention from the research community. They are used to monitor human activities and track people of interest, to detect suspicious and abnormal or undesirable events in the public, to interpret the moving actions of any object or people, to provide security for high risk environments like airport, pilgrimages, shopping center's, government buildings, gallery [5], etc. Currently, surveillance is being monitored manually. Operator's job is to watch the

C. Jayavarthini (✉) · C. Malathy
SRM Institute of Science and Technology, Chennai, India
e-mail: jayavarc@srmist.edu.in

C. Malathy
e-mail: malathyc@srmist.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_28

surveillance videos to monitor and detect such abnormal activities. But human error may happen, and the required events may not be noticed by the operator. Also, it is too expensive and unreliable. But these limitations can be effectively removed with the help of automated computers. A computer vision-aided system can be used to select the part of the video that contains person of interest, person or object to be tracked. Hence, there is a need for intelligent surveillance systems to monitor humans across different cameras fixed in a public environment. Person re-identification [14] is defined as the establishment of the relationship between images of a person taken from different cameras at different times and at different locations. It can also be used to provide a connection between the disconnected tracks to enable tracking across different cameras. It can be used in different instances: (i) Given an image of criminal suspect, the intelligent surveillance system should retrieve the clips that contain image of that person, captured by the mass city surveillance videos; (ii) a kid lost its way to home, given the image of the kid, the system should retrieve the surveillance video that has the stills of the kid.

2 Existing Challenges

The automation of re-identifying a person is difficult to accomplish without human intervention due to the following reasons [5, 14]:

1. A person's appearance may vary across different cameras, at different times.
2. Extraction of robust and discriminant descriptors from unconstrained, noncooperative and uncontrolled environment.
3. The person may be occluded.
4. Body deformations will also be present.
5. The poses also may vary.
6. Variations in lighting condition also affect the system.
7. The images may be of low quality. Also the poor quality video feeds with very huge amount of unwanted information.
8. There is no knowledge of the spatiotemporal relationship between cameras.
9. Uniform clothing in educational institutions or workplaces makes re-id algorithm really challenging to identify a person.
10. Public areas are under surveillance by thousands of cameras. Technology is required to address multi-camera surveillance problem.

3 Types of Person Re-identification

3.1 Open Set

The open set re-id system can effectively recognize whether a person seen is new or already visited person [29]. The input for the person re-id system may be a probe image or a track that contains the person of interest. Usually, the probe is matched

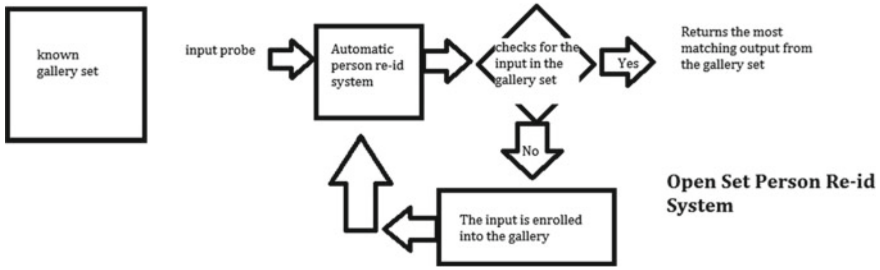


Fig. 1 Open set person re-identification

against a gallery that contains a set of images. But in the open set, the probe may not be a subset of gallery images. The gallery changes over time dynamically [19]. The probe set changes dynamically for each camera's field of view. This open set re-id is suitable for public places. It is diagrammatically explained in Fig. 1.

3.2 Closed Set

In closed set person re-identification, the gallery to be searched contains known set of images. It is diagrammatically explained in Fig. 2. There is prior knowledge about the gallery set. When the person to be queried is unavailable in the gallery, the system results in the wrong match. The closed set re-id system is unable to handle such system. But the open set re-id system will add the feature details of new person to the gallery set.

3.3 Single Shot

The input probe is the image of the person to be identified. The gallery contains different images of the same person taken in different time and location. Also, the images may be occluded. They may vary in pose and illumination.

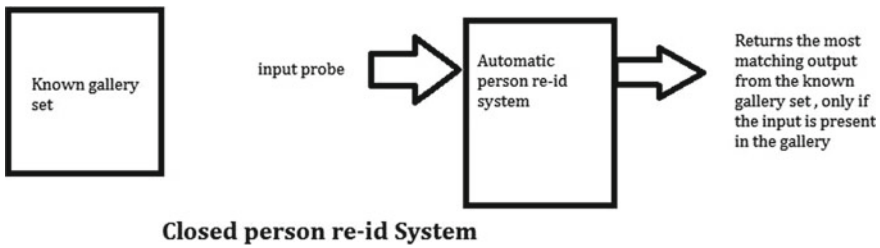


Fig. 2 Closed set person re-identification

3.4 Multi-Shot Person Re-identification

The input probe is the video track of the person to be identified. The gallery contains different video sequences of the same person taken in different time and location. They may vary in pose and illumination.

3.5 Language Based

Language can also be used for extracting effective visual cues to distinguish among different persons. Such joint image language methods can be used to increase its performance.

3.6 Long Term

It is the task of re-identifying a person from the images or videos captured with large temporal separation. It should be able to identify a person correctly even if there are changes in resolution, pose and age.

4 Steps in Person Re-identification System

The following steps are used in re-identifying a person:

1. The image/video sequence/language description of the person to be identified is given as input.
2. The robust, visual discriminant features of the person are extracted, and the descriptor id is generated.
3. The gallery is searched for the relevant input probe.
4. The contents of the gallery are ranked with a similarity score of the input with the gallery. If no similarity is found between the gallery and input probe, the probe is added to the gallery.
5. Based on the similarity score, the results are displayed to the user.

5 Related Work

Feature extraction helps in person re-identification by extracting the color, shape, texture, interest points in the human body parts and gradient features of the image regions or video track. In order to extract robust and relevant features, various methods have been proposed. The features extracted should be invariant to the view angle, poses, background clutter, occlusion, lighting conditions, etc.

5.1 Color-Based Methods

For color-based extraction, HSV or RGB histogram values were used. Sparse-based HSV color histogram matching algorithm was proposed [12] to recognize a probe person in the gallery by reducing its size. An effective dominant color descriptor (DCD) algorithm was proposed. Dominant color descriptor is based on the dominant color histogram and spatial configuration of the dominant color regions of the clothing. Probabilistic color histogram (PCH)-based visual signature was introduced [13] for re-identifying a person in the surveillance video. For each person, PCH was evaluated using fuzzy K-nearest neighbors (KNN) classifier, after applying fuzzy color quantization on selecting segments of the bounding box. The proposed descriptor can be used to discriminate and re-identify the person captured in two different cameras.

Symmetry-driven appearance-based method was presented [9] for characterizing the appearance of individuals. Overall chromatic content, recurrent local motifs, color's spatial arrangement of human body parts were used for encoding the descriptor. Random ensemble of color features (RECF) was proposed [10] for handling view point changes. The similarity function between pairs of person images was learnt using a random forest classifier. The similarity function is based on color features from RGB, CIE XYZ, normalized RGB (NRGB), YCbCr and HSV. They proved that combination of NRGB, YCbCr, HSV and CIE laboratory color spaces gave compromising results. Color patterns were learnt [2] from the samples of the image pixels that are taken with two different camera views. Similar colors were coded close to each other by learning linear transformation and dictionary encoding of pixel values together.

5.2 Texture-Based Methods

5.2.1 Covariance Descriptors

Fusion of biologically inspired features (BIF) and covariance descriptors was used for re-identifying a person [28]. Similarities of BIF with respect to neighboring scales were computed. Insists were made on the descriptors used for representing and comparing the image regions of the person [18]. BiCov needs no training images and so can be used for real-time applications. Local descriptor encoded by Fisher vector (LDFV) was introduced in [20]. The image of a person is represented as a 7D space which contains the coordinates, intensity, first-order and second-order derivatives of each pixel. Then, it is encoded in LDFV using a suitable distance metric learning approach for computing the similarity. This approach outperforms BiCov.

5.2.2 Local Binary Pattern-Based Approaches

Patch encoding local binary patterns are the fundamental properties of image texture. A method to recognize such uniform patterns was proposed by Ojala et al. [23]. Hence, LBP can be used in person re-identification. But if LBP is applied to each color channel, respectively, the relationship between the color channels was ignored. To overcome this, quaternionic LBP was proposed by Lan et al. [15, 16]. A quaternion number was used to represent each color image pixelcite. LBP features were extracted based on this representation. Then, multi-block LBP (MB-LBP) was used for person re-id. It can be computed on multiple scales in constant time using an integral image. Comparison of the sum of pixel intensities to the central rectangle determines the feature.

5.2.3 Fusion of Different Features

Color, gabor and local binary pattern were fused to form a covariance descriptor of a region in an image. Thus, it helps in discriminating the person even though there are variations in illumination, viewpoint and non-rigid body. Combination of HSV, MB-LBP and HOG features was used together for person re-id in [16]. But the interrelation between the features was not taken into account. An ensemble of localized features was proposed by Gray et al. [11]. ELF extracts the color features and texture features from the images. A feature space was designed for each image, and AdaBoost learning algorithm was used to select the feature set that uniquely distinguishes one person from another.

5.3 Human Body Part-Based Methods

5.3.1 Appearance-Based Methods

A “color position” histogram was proposed in [26, 27] for the spectral classification of appearance-based signatures. Support vector machines (SVM) were used for classifying the spectral information but the complexity increases as the training set size increases. The reduction in dimensionality can ease the learning process. An appearance model was proposed [8]. The model is based on spatial covariance regions of human body parts. A multi-component matching (MCM) framework was proposed [25]. A robust silhouette extraction algorithm was developed with a classification procedure.

5.3.2 Gradient-Based Methods

Histograms of oriented gradients (HOG) are used to automatically detect the human body parts. Discriminative human signature is generated from the correlation vectors that are connected using Riemannian manifolds [24]. A hierarchical Gaussian distribution (GOG) was proposed [21], in which both means and covariance were included. A local region in an image was modeled as a set of many Gaussian distributions.

5.3.3 Spatiotemporal-Based Methods

The open set matching problem was solved using the part-based spatiotemporal model [3]. It depends on color features that capture complementary aspects of a person's appearance and generate a compact signature of representative colors. It was computationally efficient since a signature requires low memory requirement. Another adaptive human part-based spatiotemporal model was proposed [4]. The model is based on the distribution of color features and low-level facial cues of a person of an individual across multiple frames.

5.4 Deep Learning-Based Methods

5.4.1 Convolution Neural Network

A feature extraction method based on deep convolution network was proposed by Wang et al. [30] to measure the similarities between pedestrian observations across different cameras. Further improvement can be achieved by efficient CNN architecture, combining discriminative handcrafted features with suitable metric learning method. McLaughlin et al. proposed a multitask learning method [22]. A deep convolutional network was used to characterize a person's appearance as a feature vector and labeling is done. Clothing, gender and pose of the pair of similar and dissimilar images are trained to extract the features. High accuracy can be obtained by combining identification with verification.

5.4.2 Deep Learning Approaches Using Human Body Parts

A deep learning method to associate persons captured from different cameras was proposed [31]. Decomposition of human body parts into discriminative regions, computation of similarity and aggregation of similarities were learnt through minimizing the triplet loss function. This approach does not require body part labeling information. Integrated method of local maximal occurrence (LOMO) features with multilabel multilayer perceptron (MLMLP) classifier was proposed [6]. A triplet CNN similarity learning framework was generated fusing both learned CNN features and attributes. A deep learning supervised approach based on the fusion of identity and attribute was proposed [7]. Attribute pattern learning on different body parts was done using part-specific CNN and then fused with low-level local maximal occurrence (LOMO) features. Again, triplet structure embedded CNN was merged for identification. An adaptive re-ranking (ARR) was proposed [17] to iteratively incorporate the contextual information between the deeply learnt neighboring features. Thus, non-relevant images are removed. Also, deep feature fusion (DFF) method was used to split the features into sub-features, and feature information was exchanged by a diffusion process. A deep long short-term memory approach was proposed [1] to learn full person details. Identification and ranking approach were integrated for person re-identification.

6 Comparison

See Table 1.

7 Findings

From the comparison table in Sect. 6, the following inferences were made:

1. Local discriminant fisher vector (LDFV) and spatial covariance region (SCR) have 60.04% accuracy (Rank 10) and 62.7% (Rank 10) accuracy, respectively, even with invariant illumination and occlusion. LDFV handles pose changes also.
2. With respect to changes in view point alone, ensemble of local features (ELF) achieves 70% accuracy at Rank 10, deep learning method with LOMO and MLMLP achieves 86% accuracy at Rank 5, and DL with long short-term memory (LSTM) achieves 89.4% at rank 1 itself.
3. Hierarchical Gaussian of Gaussian distribution (GOG) achieves 84.5% accuracy (Rank 10) at distinguishing between two persons with similar color and dress.
4. Histogram of gradients (HOG) achieves the highest accuracy 89% at 10^{-4} False Positive Per Window (FPPW), even with changes in background and pose.
5. Deep learning method with part-specific CNN + LOMO achieves 90.3% accuracy (Rank 5), even with changes in visual appearance, body pose and camera views.
6. Multitask learning with Siamese network architecture can be used for person re-identification with cross-dataset. It achieves 80.62% (Rank 10).
7. Integration of adaptive re-ranking and deep feature fusion considers the contextual information and hence achieves 91.5% accuracy at Rank 1 itself.

Table 1 Comparison of existing feature extraction algorithms

S. No.	Name of the technique	Year	Challenges encountered	Dataset	Rank 1	Rank 5	Rank 10
Comparison of color-based feature extraction algorithms							
1.	Probabilistic color histogram (PCH)-based visual signature	2011	Changes in pose, illumination and view point	VIPeR	20.33	25.65	30.75
2.	Random ensemble of color features (RECF)	2012	Viewpoint changes	VIPeR	16.96	30.98	42.56
3.	Dominant color descriptor (DCD) algorithm	2014	Changes in pose, viewpoint and illumination	VIPeR	30	40	50
4.	Color patterns	2016	Variant camera views and lighting conditions	VIPeR	14		45
5.	Symmetry-driven appearance-based method (SDALF)	2013	Low resolution images, occlusions, pose changes, viewpoint variations and illumination	VIPeR	19.87	38.89	49.37
6.	Sparse-based HSV color histogram matching algorithm	2017	Low image resolution and partial occlusion	VIPeR	23.19	42.03	52.17
Comparison of texture-based feature extraction algorithms							
1.	Local Binary Pattern (LBP)	2002	Variations in illumination, view angle, pose and occlusion	ETHZ, i-LDS	28.2	52.97	–
2.	Ensemble of local features (ELF)	2008	Changes in view point	Customized dataset	12		70
3.	Gabor LBP	2011	Changes in illumination, viewpoint and non-rigid body	VIPeR	13.5	–	–
4.	Local discriminant fisher vector (LDFV)	2012	Changes in pose, viewpoint, lighting and occlusions	VIPeR	22.34	30	60.04
5.	Quaternionic LBP	2014	Increasing number of persons in the test images	ETHZ, i-LDS	31.27	57.79	–
6	BiCov	2016	Variations in scale, illumination and background	VIPeR	10	22.5	33.5

(continued)

Table 1 (continued)

S. No.	Name of the technique	Year	Challenges encountered	Dataset	Rank 1	Rank 5	Rank 10
Comparison of interest regions based feature extraction algorithms							
1.	Hierarchical Gaussian of Gaussian distribution (GOG)	2010	To distinguish between two persons with similar color complexion, dress, etc.	Multiple camera tracking scenario (MCTS)	42.3	66.7	84.5
2.	Spatial covariance region descriptor (SCR)	2010	Non-overlapping multiple camera views, occlusions, large variations in both view angle and illumination	i-LDS	32.5	54.6	62.7
3.	Color position histogram + support vector machines (SVM)	2011	Identify a person in two different background	INRETS	TRR = 95%	- -	–
4.	Multi-component matching framework (MCM)	2011	Changes in illumination with low computational cost	VIPeR	20	40	56.75
5.	Spatiotemporal model based on color and facial features	2012	Multiple person re-identification with dynamically changing environment	Customized dataset	80%	–	–
6.	Adaptive Gaussians mixture model in joint spatio-colorimetric feature space	2012	Unstable lighting, distraction in motion	INRETS	TRR = 97.5	–	–
7.	Histograms of oriented gradients (HOG)	2016	Change in pose and background	MIT pedestrian database	89% at 10^{-4} FPPW	–	–
Comparison of deep learning algorithms							
1.	Deep convolution network-based feature extraction	2015	Cross camera	CUHK03 (detection)	32.92	52.98	63.94
2.	Multitask learning with Siamese network architecture	2016	Cross-dataset scenario	VIPeR	29.8	59.6	80.62
3.	Fusion of body part extraction with part representation	2017	Changes in pose and background	Market1501, VIPeR, CUHK03, CUHK01,	85.4	97.6	99.4
4.	Attribute-based human recognition using LOMO and mutilabel multilayer perceptron classifier (MLMLP)	2017	Changes in view point	VIPeR, GRID, APiS, CUHK03	55.1	86.1	93.3
5.	Adaptive re-ranking (ARR) + deep feature fusion (DFF)	2018	Considered contextual information	Market1501	91.5		
6.	Fusion of part-specific CNN and low-level local maxima occurrence (LOMO)	2018	Variations in person's look, pose and camera views	VIPer, PETA, APiS, CUHK03	65	90.3	95.1
7.	Long short-term memory (LSTM)	2018	Disjoint camera views	Market1501 CUHK03 (detection)	89.4	98.2	99.1

8 Conclusion

Person re-identification is gaining research interest nowadays. It can be used in various fields like real-time video surveillance, multimedia, forensics, suspicious activity detection, undesirable event prediction, content-based media retrieval, etc. This paper summarizes the various feature extraction methods used for person re-identification. Since feature extraction is the very first step in re-identifying a person, it has a great impact in the result. It has been found that the combination of various feature extraction methods gives compromising results. Researches are going on in the multimodal techniques to improve the person re-identification system. Fusion of pattern learning using deep learning with already existing state-of-the-art feature extraction methods was proved to improve the accuracy. Still a lot of research contributions are required in the field of identifying and tacking group of individuals, long-term person re-identification, reduction in time complexity while handling huge volume of video data.

References

1. Bai X, Yang M, Huang T, Dou Z, Yu R, Xu Y (2018) Deep-person: Learning discriminative deep features for person re-identification. arXiv:1711.10658
2. Bazzani L, Cristani M, Murino V (2013) Symmetry-driven accumulation of local features for human characterization and re-identification. Comput Vis Image Underst 117(2)
3. Bedagkar-Gala A, Shah S (2011) Multiple person re-identification using part based spatio-temporal color appearance model. In: IEEE international conference on computer vision workshops (ICCV workshops)
4. Bedagkar-Gala A, Shah S (2012) Part-based spatio-temporal model for multi-person re-identification. Pattern Recogn Lett 33:1908–1915
5. Bedagkar-Gala A, Shah SK (2014) A survey of approaches and trends in person re-identification. Image Vis Comput 32:270–286
6. Chen Y et al (2017) Triplet CNN and pedestrian attribute recognition for improved person re-identification. In: 14th IEEE international conference on advanced video and signal based surveillance (AVSS)
7. Chen Y et al (2018) Deep and low-level feature based attribute learning for person re-identification. Image Vis Comput 79:25–34
8. Corvee E, Brémond F, Thonnat M (2010) Person re-identification using spatial covariance regions of human body parts. In: 7th IEEE international conference on advanced video and signal based surveillance
9. D'Angelo A, Dugelay J (2011) People re-identification in camera networks based on probabilistic color histograms. In: Visual information processing, vol 7882
10. Forssén PE (2007) Maximally stable color regions for recognition and matching. In: IEEE conference on computer vision and pattern recognition
11. Gray D, Tao H (2008) Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: ECCV
12. Jaehoon Jung HK, Paik J (2017) Person re-identification using color name descriptor-based sparse representation. In: IEEE 7th annual computing and communication workshop and conference (CCWC)
13. Jang K, Han S, Kim I (2014) Person re-identification based on color histogram and spatial configuration of dominant color regions. arXiv:1411.3410

14. Karanam S, Wu Z, Radke RJ (2016) A systematic evaluation and benchmark for person re-identification: features, metrics, and datasets. arXiv:1605.09653
15. Lan R, Zhou Y, Tang YY, Chen CLP (2014) Person reidentification using quaternionic local binary pattern. In: IEEE international conference on multimedia and expo (ICME)
16. Li S (2011) Gabor-LBP based region covariance descriptor for person re-identification. In: Sixth international conference on image and graphics
17. Liu Y, Shang L, Song A (2018) Adaptive re-ranking of deep feature for person re-identification. arXiv:1811.08561
18. Ma B, Su Y, Jurie F (2012) BiCov: a novel image representation for person re-identification and face verification. In: British machine vision conference
19. Ma B, Su Y, Jurie F (2014) Discriminative image descriptors for person re-identification. Advances in computer vision and pattern recognition. Springer, London
20. Ma B, Su Y, Jurie F (2014) Discriminative image descriptors for person re-identification. Springer, London
21. Matsukawa T, Okabe T, Suzuki E, Sato Y (2016) Hierarchical Gaussian descriptor for person re-identification. In: The IEEE conference on computer vision and pattern recognition
22. McLaughlin N, Martinez del Rincon J, Miller PC (2016) Person reidentification using deep convnets with multitask learning. IEEE Trans Circuits Syst Video Technol 27:525–539
23. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987
24. Oncel Tuzel FP, Meer P (2008) Pedestrian detection via classification on Riemannian manifolds. IEEE Trans Pattern Anal Mach Intell
25. Satta R, Fumera G, Rolli F, Cristani M, Murino V (2011) A multiple component matching framework for person re-identification. In: International conference on image analysis and processing, ICIAP 2011
26. Truong Conga DN, Khoudoura L, Ahard C, Meurie C, Lezora O (2010) People re-identification by spectral classification of silhouettes. Signal Process 90:2362–2374
27. Truong Cong D-N, Achard C, Khoudour L (2010) People re-identification by classification of silhouettes based on sparse representation. In: 2nd international conference on image processing theory, tools and applications
28. Varior RR, Wang G, Lu J (2016) Learning invariant color features for person re-identification. IEEE Trans Image Process 25:3395–3410
29. Vidanapathirana M, Sudasingha I, Kanchana P, Vidanapathirana J, Perera I (2017) Open set person re-identification framework on closed set re-id systems. In: IEEE 2nd international conference on signal and image processing (ICSIP)
30. Wang S et al (2015) Deep features for person re-identification. In: 11th international conference on semantics, knowledge and grids (SKG). IEEE
31. Zhao L et al (2017) Deeply-learned part-aligned representations for person re-identification. In: Proceedings of the IEEE international conference on computer vision

Emotion Recognition from Speech Using Perceptual Features and Convolutional Neural Networks

A. Revathi, R. Nagakrishnan, D. Vishnu Vashista, Kuppa Sai Sri Teja, and N. Sasikaladevi

Abstract Emotional computing has played a crucial role in acting as an interface between humans and machines. Speech based emotion recognition system is difficult to be implemented because of the dataset which is containing a limited number of speeches. In this work, multi speaker independent emotion recognition encompasses the use of perceptual features with filters spaced in BARK scale and Equivalent rectangular bandwidth (ERB) and vector quantization (VQ) for classifying groups and convolutional neural network with backpropagation algorithm for emotion classification in a group. The proposed system has provided consistently better accuracy for the perceptual feature with critical band analysis done in ERB scale with overall accuracy as 86% and decision level fusion classification yielded 100% accuracy for all emotions. Speaker dependent emotion recognition system has provided 100% as accuracy for all the emotions for ERB-PLPC features and perceptual linear predictive cepstrum has given 100% as accuracy for all emotions except sad emotion.

Keywords PLPC · ERBPLPC · CNN · Emotion recognition

A. Revathi · R. Nagakrishnan · D. Vishnu Vashista · K. S. S. Teja (✉)
Department of ECE, School of EEE, SASTRA Deemed to be University,
Thanjavur, Tamil Nadu, India
e-mail: kuppasaisriteja@gmail.com

A. Revathi
e-mail: revathi@ece.sastra.edu

R. Nagakrishnan
e-mail: nagakrishnan@sastra.ac.in

D. Vishnu Vashista
e-mail: vishnuvasista99@gmail.com

N. Sasikaladevi
Department of CSE, School of Computing, SASTRA Deemed to be University,
Thanjavur, Tamil Nadu, India
e-mail: sasi@cse.sastra.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_29

1 Introduction

Modern Speech signal is considered as the acoustic signal obtained by exciting the vocal tract by quasi periodic pulses of air for voiced sounds and noise like excitation for unvoiced sounds. Speech utterances reveal the speaker's linguistic content, accent, slang and emotional state. This is cumbersome to recognize the emotions from speech with a limited set of data. Emotion recognition from speech has found applications in call centers and unmanned control of risky processes. This system would be useful for treating the mentally retarded patients and patients with depression and anxiety. Web related services, retrieval of information and synthesis of data would use this automated emotion recognition system. These systems will find a place in operating robots for the speech commands given by the emotional worker. The emotion recognition has been done using hierarchical binary classifier and acoustic and statistical features by Lee et al. [1]. Wua et al. [2] has used a novel Modulation spectral features for recognize the different emotion. Sreenivasa Rao et al. [3] developed an emotion recognition system using MFCC feature and GMM. The emotion recognition system has been developed based on new MFCC feature and neural network by Sapra et al. [4]. Koolakudi et al. [5] developed a speaker recognition system using MFCC feature and Gaussian Mixture Model in an emotional environment. Nithya Roopa et al. [6] developed an emotion recognition system based on deep learning neural networks. Huang et al. [7] has used verbal and nonverbal sounds for developing an emotion recognition system using deep learning neural networks. This work mainly encompasses the development of emotion recognition system using perceptual features and group classification by VQ based classifier and individual emotion classification in a group by convolutional neural network (CNN) classifier.

In this work, a set of features are extracted from training speech utterances and models are established as illustrative of emotions. In the test phase, classification of the group can be done based on minimum distance classifier and then classification of individual emotions is completed in a group comprising appropriate emotion models using linear binary classifier. Perceptual linear predictive cepstrum with critical band analysis done in BARK and ERB scale are used as features and they deliver complimentary proof in analyzing performance of the systems using convolutional neural networks. This work also deals with the comparative analysis between the features and investigation is done broadly to evaluate the performance of the system with utterances spoken by all the speakers taken for training and testing for speaker dependent case. Utterances for some speakers are taken for training and other speakers for testing to implement speaker independent system.

Organization of the work envisages the feature extraction technique and development of CNN models for each emotion in Sect. 2 and testing the utterance in pertinent emotion by comparing the features with CNN models in Sect. 3. Conclusion of the work is presented in Sect. 4.

2 Materials and Methods

2.1 PLPC and ERBPLPC Extraction

The extraction of perceptual linear predictive cepstrum deals with computation of spectrum using FFT technique, wrap the spectrum along the BARK and ERB frequency scales, performing convolution between the changed spectrum and spectrum of the replicated critical band masking curve, performing pre-emphasis by simulated loudness equalization, mapping between the intensity and the perceived loudness done by cube root compression, generation of linear predictive (LP) coefficients and conversion of them into LP derived cepstrum. Procedure [6] used for PLP and ERBPLP perceptual feature extraction steps are detailed below.

1. The power spectrum of the windowed speech segment is computed.
2. Make clustering to 21 critical bands in BARK scale and 35 critical bands in ERB scale for 16 kHz sampling frequency.
3. Loudness equalization and cube root compression are done to simulate the power law of hearing are performed.
4. Inverse fast Fourier transform is computed.
5. Levinson-Durbin procedure is used to perform linear predictive analysis to obtain linear predictive coefficients (LPC).
6. The LPC are converted into PLP and ERBPLP Cepstral coefficients.

Frequencies in Bark and ERB scale are related to the frequency in Hz as shown in (1) and (2)

$$f(\mathrm{Bark}) = 6 * \sinh^{-1}(f(\mathrm{Hz})/600) \tag{1}$$

$$f(\mathrm{ERB}) = 21.4 * \log 10\big(4.37\mathrm{e}^{-3} * \mathrm{f(Hz)} + 1\big) \tag{2}$$

2.2 Experimental Analysis Based on Clustering Method

The K-means clustering technique [8, 9] is used to quantize the number of training vectors into number of clusters with cluster centroid representing the group of training vectors by computing Euclidean distance. The test feature vectors are represented as v_i and stored cluster centroids are represented as v_j. The distance between v_i and v_j is computed as in (3)

$$d_k = d\big(v_i, v_j\big) = 0 \quad \text{when } v_i = v_j \tag{3}$$

The distance is minimum or zero for the test speaker and matching speaker template, while it is maximum between the test emotion and other emotion templates.

So, the minimum distance index is the identifier of the test emotion. It is represented as in (4)

$$I = \arg(\min(d_k)) \quad \text{for } 1 \leq k \leq M \tag{4}$$

2.3 Emotion Recognition Based on Convolutional Neural Network

All recognition system's performance is judged based on the database. EMO-DB Berlin database used in this work contains speech utterances spoken in seven different emotions by ten actors in the age between 21 and 35 years. Robustness of the system depends on the amount of data considered for training to create templates for emotions. This work on emotion recognition from speech contains training and testing phase. During training, the speech signal is generated by concatenating the utterances in pertinent emotion from nine speakers. Pre-emphasis is the done on the speech vector to flatten the signal in spectral domain. Frame blocking is the technique used to convert the pre-emphasized speech into frames with 16 ms duration and 8 ms frame shift. Windowing is the technique done on the frames to remove the signal discontinuities at the beginning and end of speech frames. The perceptual features with filers spaced in bark and ERB scale are extracted for each speech frame of 16 ms duration. Neural network models are created for emotions in a group. A Convolutional Neural Network is a method of deep learning, most commonly used in pattern recognition. CNN is normalized versions of multi-layer perceptron networks [11]. It can be considered as a modified version of the standard NN [12]. It consists of three layers namely, convolutional layer, pooling layer, and fully-connected layer [10]. Speech utterance is converted into feature vectors essential to be prearranged as feature maps, which is given as input to the convolutional neural network. So, it is needed to change speech features into feature maps. The so formed feature map is applied with the convolution layer and pooling layer operations, to create the activations of the neurons in respective layers, as shown in Fig. 1. The input feature maps $x_i \; i = 1, 2 \ldots, I$ are linked to feature maps $y_j \; j = 1, 2 \ldots, J$ in the convolution layer based on several weight matrix $w_{i,j}$. It can be computed as in (5)

$$y_j = \sigma\left(\sum_{i=1}^{I} x_i * w_{i,j}\right) \tag{5}$$

I, J—total number of feature maps in the input and convolutional layer.

The back-propagation algorithm can be used to update the weights in the convolution layer. There are two major different aspects in the convolution layer when compared to the standard, fully connected hidden layer. First, the local area of the input comprising of some features forms the input to the convolutional unit. Second,

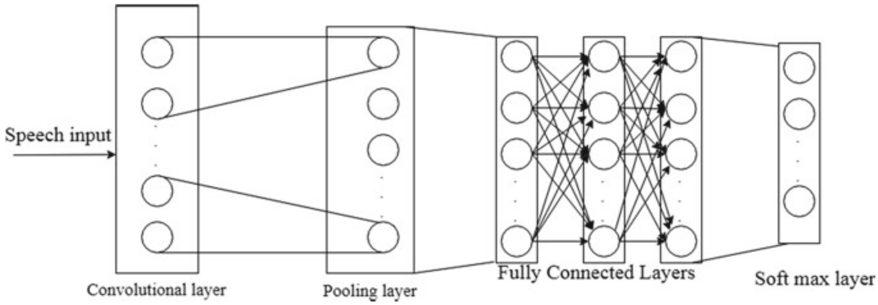


Fig. 1 Architecture of convolutional neural network

each distinct feature-map in the self-organized convolution layer receive input from different positions of the lower layer but share the respective weight of the feature map. The feature maps of the convolutional layer are fed to the pooling layer.

The pooling layer is prepared into feature maps. The quantity of feature maps is same in the convolution and pooling layer; however, every feature map is lesser. Reduction in resolution of feature maps is done by pooling function. The size of the local region is determined by pooling size. The pooling function may be averaging or maximization. The pooling layer based on maximization of pooling function is defined as in (6)

$$z_{i,m} = \max_{n=1}^{G} y_{i,(m-1)*s+n} \tag{6}$$

G—Pooling size
s—Shift size.

Pooling layer based on the average of the pooling function is defined as in (7)

$$z_{i,m} = r \sum_{n=1}^{G} y_{i,(m-1)*s+n} \tag{7}$$

r—is a scaling factor. In image recognition application, the value of G is equivalent to s. i.e., there is no overlap and spacing in between the pooling windows. In emotion recognition the value G and s adjusted independently. The generation of output is done by application of nonlinear activation function to the $z_{i,m}$.

Test speech vector is pre-emphasized by using single order high pass filter so that flattening of speech spectrum is done. After getting speech vector passing through the pre-emphasis filter stage, the vector is divided into overlapping speech segment of 16 ms with an overlap of 50%, between segments so that there is no information loss. To remove discontinuity at the start and end of each segment, multiplication of each segment with hamming window is performed. Speech segment after pre-processing would undergo FFT filtering, Loudness equalization, critical band analysis, cube root compression, Linear Prediction coefficients extraction and deriving cepstrum from the LP coefficients. These extracted features are given as input to the CNN models pertaining to the emotions and emotion is recognized correctly using minimum distance classifier. The emotion recognition accuracy is computed by counting the number of speech segments correctly classified out of entire number of segments considered for each emotion.

3 Experimental Evaluation of the Emotion Recognition System

This multi-speaker independent emotion recognition system is evaluated by using the implementation of the parallel specific emotional pattern classifier [9] and parallel group classifier to increase the system performance. Parallel emotion specific classifier and parallel group classifier is indicated in Fig. 2. The concatenated test speech vector is applied to the pre-processing techniques and the extracted feature vectors are given as input to the vector quantization (VQ) templates for identifying a group and fed to the CNN templates and based on the minimum error criterion test speech is classified as association with pertinent emotion. Group classification is done based on minimum distance classifier and individual emotion in a group is classified based on CNN based minimum mean squared error classifier. After conventional pre-processing stages, the emotional models are created from extracted perceptual features and group should be correctly classified based on minimum distance classifier. The features are applied to the emotion specific CNN models and classification of the pertinent emotion is done by calculating the number of indices correctly pointing the pertinent emotion in a group of two emotions. For testing using CNN, extracted features from test utterances in relevant emotion after group classification are applied to the CNN templates, and classification between two emotions is done by computing the indices derived as compared with target indices. Figure 3 indicates the Convolutional neural network structure for classifying emotions Anger or fear, boredom or disgust and neutral or sad.

Evaluation of the CNN based emotion recognition system with utterances uttered by speaker considered for testing being different from the utterances spoken by speakers considered for training is done for individual features and decision level fusion of features is depicted in Table 1. From Table 1, ERBPLP feature provides 0% accuracy for neutral emotion and PLP feature provides 0% accuracy for Fear and sad emotions. But decision level fusion classification among the features has provided 100% accuracy for all the emotions. Table 2 indicates the performance of the speaker dependent emotion recognition system for classifying emotions from the speech utterances of all speakers considered for training and testing.

From Table 2, in speaker dependent emotion recognition, ERBPLP provides 100% accuracy for all emotions and PLP provides 100% accuracy for all emotions excepting sad. Decision level fusion classifier has yielded 100% accuracy for all emotions

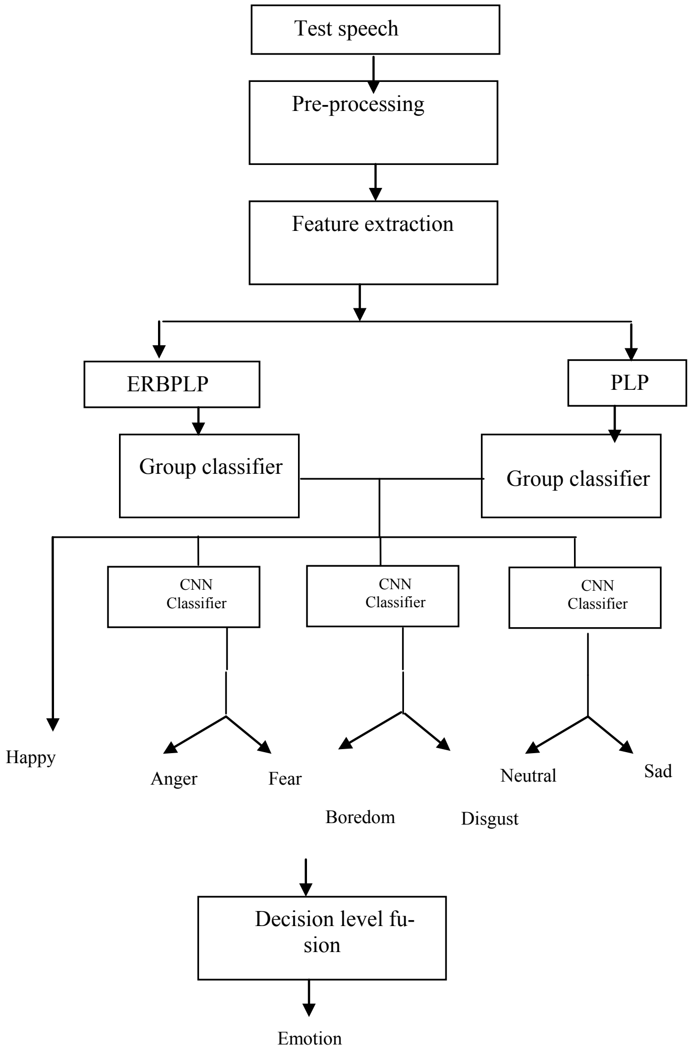


Fig. 2 Parallel and decision level fusion classifier using CNN

indicating CNN based modeling technique can be used for emotion classification using speech.

F-ratio is the parameter used to estimate the recognize system performance on the basis of feature selection. It is ratio of the variance computed between features of two speakers and variance between features of the same speaker. It is found to be monotonically increasing function of accuracy or monotonically decreasing function of misclassification of one speaker to another. F-ratio is calculated as in (8).

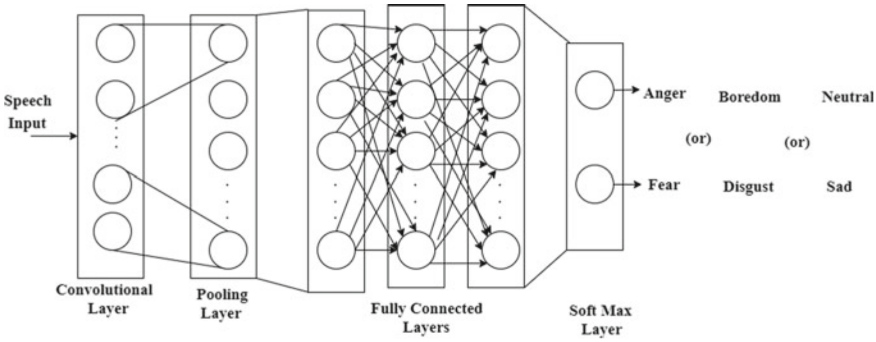


Fig. 3 CNN structure—classification of Emotions

Table 1 Performance evaluation of speaker independent system—ERBPLP and PLP, decision level fusion of features using CNN

ERBPLP			PLP		Decision level fusion (Features)	
Concerned emotion		Other emotion	Concerned emotion	Other emotion	Concerned emotion	Other emotion
Anger	166	0	166	0	166	0
Boredom	392	0	392	0	392	0
Disgust	392	0	392	0	392	0
Fear	166	0	0	166	166	0
Happy	310	0	310	0	310	0
Neutral	0	348	348	0	348	0
Sad	348	0	0	348	348	0

Table 2 Performance evaluation of speaker dependent system—ERBPLP and PLP, decision level fusion of feature using CNN

ERBPLP			PLP		Decision level fusion (features)	
Concerned emotion		Other emotion	Concerned emotion	Other emotion	Concerned emotion	Other emotion
Anger	1282	0	1282	0	1282	0
Boredom	1090	0	1090	0	1090	0
Disgust	1090	0	1090	0	1090	0
Fear	1282	0	1282	0	1282	0
Happy	1376	0	1376	0	1376	0
Neutral	1580	0	1580	0	1580	0
Sad	1580	0	0	1580	1580	0

$$F = \frac{\sigma_1^2}{\sigma_2^2} \tag{8}$$

where σ_1^2 indicates inter speaker dispersion with σ_2^2 indicating intra speaker dispersion of mean values. Figure 4 indicates the variation in F-ratio for perceptual features with critical band analysis done in BARK, and ERB frequency scales. F-Ratio and the accuracy for ERBPLP feature are relatively high as compared to PLP feature.

Figure 5 indicates the variation in accuracy with respect to the features for speaker dependent emotion recognition. It is evident that accuracy is directly proportional to F-Ratio.

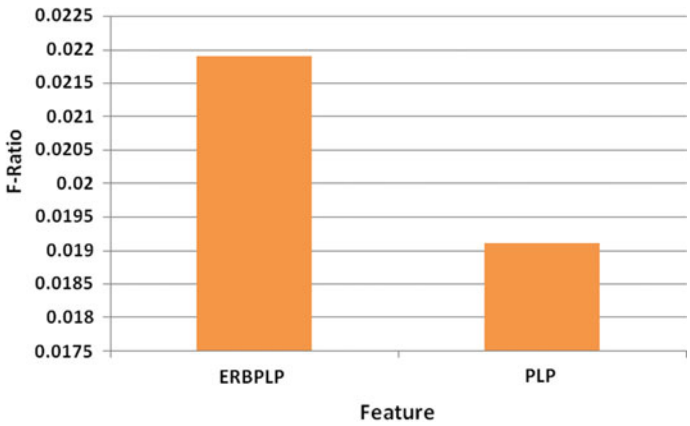


Fig. 4 F-Ratio versus features

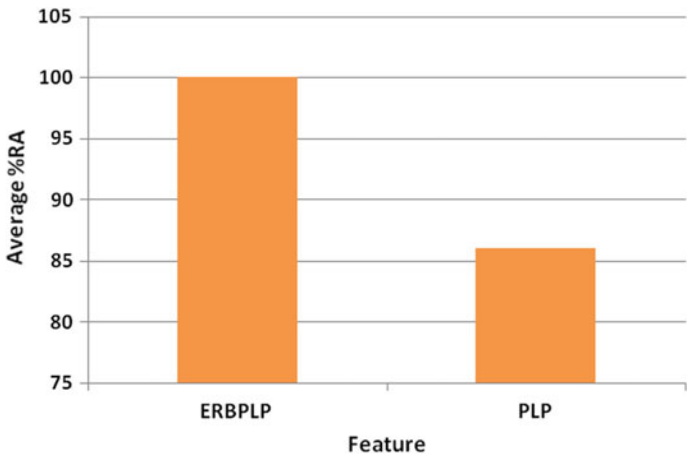


Fig. 5 Accuracy versus features

4 Conclusion

The paper provides the insight on the use of perceptual features with critical band analysis done in ERB and BARK scale and Convolutional neural network for creating templates as representative models for emotions. The extraction of features from concatenated training speeches pertaining to the emotions is performed. Such extracted features pertaining to arousal and soft emotions are used to create set of clusters representing group models. Features extracted using pertinent emotional speech utterances are applied to the CNN models with initialization of dimensions for convolution layer, pooling layer and fully connected layer and corresponding weights between layers. Iterative procedure is used to update the weights by fixing the target error and neural network models are formed for each emotion. The test speeches are concatenated during test phase, to form a test vector. This vector is fed as an input to the initial pre-processing stages and features are extracted. Features extracted from the pre-processed test speech vector are given as input to the group models for group identification based on minimum distance classifier and further classification is done in a group containing pertinent emotion models based on minimum mean squared error criterion. Perceptual features with filters spaced in ERB scale has provided the better accuracy of 86% as compared to the perceptual features with filters spaced in BARK scale for which the accuracy is found to be 72% for speaker independent emotion recognition. Decision level fusion classifier has yielded the accuracy as 100%. Speaker dependent emotion recognition system is implemented by using same set of utterances for training and testing. Perceptual features with critical band analysis done in ERB scale has provided 100% as accuracy for all emotions and perceptual features with filters spaced in BARK has given 100% as individual accuracy for all emotions except sad. Decision level fusion classifier using both the features and CNN as modeling technique has given the accuracy as 100% for all emotions. In this work, F-ratio is used to validate feature selection for assessing the emotion recognition system performance. Emotion recognition from speech would find applications in business processing centers, web analysis, medical diagnosis of patients, and automated services based on the emotional state of a speaker.

References

1. Lee C-C, Mower E, Busso C, Lee S, Narayanan S (2011) Emotion recognition using a hierarchical binary decision tree approach. Speech Commun 53:1162–1171
2. Wua S, Falk TH, Chan W-Y (2011) Automatic speech emotion recognition using modulation spectral features. Speech Commun 53:768–785
3. Sreenivasa Rao K, Kumar TP, Anusha K, Leela B, Bhavana I, Gowtham SVSK (2012) Emotion recognition from speech. Int J Comput Sci Inf Technol 3(2):3603–3607
4. Sapra A, Panwar N, Panwar S (2013) Emotion recognition from speech. Int J Emerg Technol Adv Eng 3(2). ISSN 2250–2459, ISO 9001:2008 Certified Journal

5. Koolagudi SG, Sharma K, Sreenivasa Rao K (2012) Speaker recognition in emotional environment. Commun Comput Inf Sci 305:117–124
6. Nithya Roopa S, Prabhakaran M, Betty P (2018) Speech emotion recognition using deep learning. Int J Recent Technol Eng (IJRTE) 7(4S):247–250
7. Huang K, Wu C, Hong Q, Su M, Chen Y (2019) Speech emotion recognition using deep neural network considering verbal and nonverbal speech sounds. In: ICASSP 2019–2019 IEEE international conference on acoustics, speech and signal processing (ICASSP), Brighton, United Kingdom, pp 5866–5870
8. Revathi A, Venkataramani Y (2012) Perceptual features based continuous speech recognition in additive noise environment using various modeling techniques. STM J Curr Trends Sig Process 2(3):1–15
9. Jeyalakshmi J, Revathi A, Venkataramani Y (2016) Integrated models and features based speaker independent emotion recognition. Int J Tele Med Clin Pract 1(3):271–291
10. Huang J, Li J, Gong Y (2015) An analysis of convolutional neural networks for speech recognition. In: 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), Brisbane, QLD, pp 4989–4993
11. Palaz D, Magimai Doss M, Collobert R (2015) Convolutional neural networks-based continuous speech recognition using raw speech signal. In: 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), Brisbane, QLD, pp 4295–4299
12. Abdel-Hamid O, Mohamed A, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional neural networks for speech recognition. IEEE/ACM Trans Audio Speech Lang Process 22(10):1533–1545

Implementation of Building Stability Analyzer with Earthquake Detection Using Simple MEMS Pressure Sensor

D. Sindhanaiselvi and T. Shanmuganantham

Abstract Analysis of building stability is an essential measurement process for all buildings in cities for achieving safety and serviceability. This is achieved by measuring the vibration of the building, and also, an alert is created to save the life of people who live in very tall building. It is proposed to design a stability analyzer combined with earthquake detection alarm. When earthquake is detected, the controller checks the vibration level with threshold level. If the vibration input exceeds threshold level, the controller triggers alarm and sends voice message through GSM. In the proposed paper, the MEMS pressure sensor (BMP180) is used to sense the vibration. In addition to sending data to monitoring section, it automatically opens all doors and sends data to nearby rescue station and emergency warning call to evacuate the people from high-rise building. The vibration of the building is recorded continuously for the proper maintenance and service.

Keywords Stability · Building · Pressure sensor · MEMS · Earthquake

1 Introduction

The stability of the building is monitored regularly to avoid the collapse of the building not only due to aging but also due to earthquake. This is achieved by measuring the vibration in and around the building. The most important factor of vibration is earthquake which is most disastrous natural event. Naturally, earthquake is generated due to the movement of rocks beneath the earth. This is exposed outside by vibration, shaking, and sometimes displacement of the ground which is nothing but pressure exerted.

D. Sindhanaiselvi (✉)
Department of EIE, Pondicherry Engineering College, Pondicherry 605014, India
e-mail: sindhanaiselvi@pec.edu

T. Shanmuganantham
Department of Electronics Engineering, Pondicherry University, Pondicherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_30

The disaster alert systems have been introduced in the early years to save the life and to initiate safety precautions to avoid any mishap [1–6]. In 2010, Hiroshi taji mi [7] presented a paper on theoretical analysis of earth pressure surges on basement walls on a basis of 2D-wave propagation theory. The results are given for distribution of earth pressure on wall and the coefficient of soil reaction.

In 2012, Raju et al. [8] proposed a system which focused on monitoring water level and earth vibrations via sensors with threshold and transmitted alarm notification through GSM. In 2015, Hoque et al. [9] proposed a system to save human lives and also store the data for future analysis using MEMS accelerometer which is complex in data analysis [7]. However, this is designed to detect p-wave which is difficult because of low frequency and MEMS accelerometer fails to detect such frequency. It does not use wireless transmission which makes this difficult for data transmission [10].

The concept of earthquake detection designed using MEMS accelerometer gives poor performance. It is not possible to measure the magnitude of earthquake and data logging together. However, MEMS sensors are low in cost, rugged, stable and also recommended harsh in environments such as flood level monitoring [11], indoor tracking system [12] and low pressure measurement [13–15] it is most suitable choice of sensor in the proposed design. To appreciate this, simple MEMS pressure sensors are used for sensing earthquake with an alarm facility. Also the system is approached a wireless sensor network systems using ZigBee network topology is adopted with LabVIEW as a software for monitoring and data logging purpose is used.

In the existing system [6, 12], MEMS accelerometer is used to measure the vibration of building, which sends data to monitoring section through GSM wireless technology. In the proposed project, the MEMS accelerometer is replaced by MEMS pressure sensor (BMP180), and in addition to sending data to monitoring section, it also opens all the automatic doors and sends data to nearby rescue station. The proposed system also contains wireless transmission of data through ZigBee and storage of data for future reference.

A brief overview of the disaster alert systems the conventionally used approaches and their limitations have been discussed in Sect. 1. Section 2 explains the block diagram for implementation. Section 3 describes the proposed pressure sensor implementation via I2C multiplexer using Arduino processor of hardware interfacing in detail. The different software interfacing stages with flowchart are discussed in Sect. 4, and outputs of the vibration display using LCD have been discussed in Sect. 5.

2 Overview of Block Diagram

The proposed design of implementation consists of transmitting and receiving sections in the overall block diagram as shown in Fig. 1. The Mega 2560 is a microcontroller board based on the ATmega2560. It is compatible to support the microcontroller and can be easily powered by USB or AC-to-DC adapter.

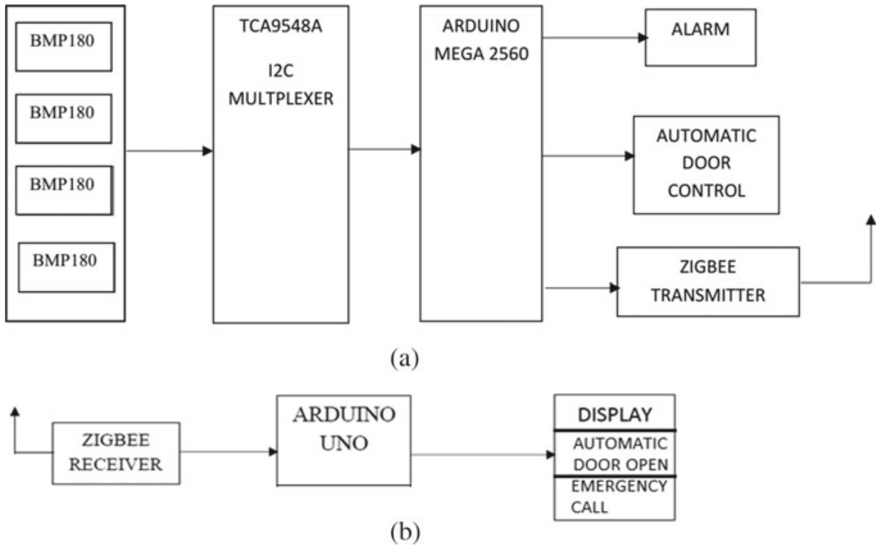


Fig. 1 Overall block diagram, **a** transmitting block, **b** receiving block

The force exerted on the building due to earthquake is related to

$$F = ma \tag{1}$$

where F is the force (N), m is the mass (Kg), and a is the acceleration due to gravity (m/sec^2). The pressure

$$P = F/A \tag{2}$$

where p is the pressure in (N/m^2), F is the force (N), and A is the area (m^2). The acceleration measurement using accelerometer is difficult and complex in analysis. Instead, a simple and equivalent method is proposed here to measure the pressure on the building using BMP180 pressure sensor. The temperature is included to improve the accuracy such that pressure is directly proportional to temperature by gas law.

3 Hardware Interfacing

3.1 BMP180 Sensor, ADC and I2C Interface with Arduino

The BMP180 is the high-precision digital pressure sensor based on piezoresistive technology. This sensor is selected since it has appreciable linearity, has good accuracy, has EMC robustness, and is stable in measurement. It measures the absolute

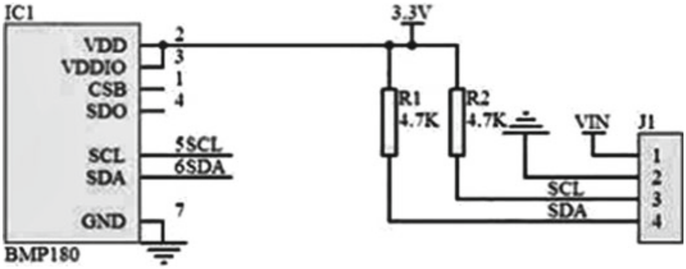


Fig. 2 BMP180 pin diagram

pressure with both altitude and weather correction and so it is highly accurate in measurement. The pin diagram of BMP180 pressure sensor is shown in Fig. 2. The BMP180 is made up of piezoresistive type, where resistance changes with respect to pressure with good gauge factor. It is internally built up with an analog-to-digital converter, control unit, and a serial I2C interface.

ZigBee chips are easily integrated with microcontrollers. The maximum distance that can be achieved by ZigBee is around 70 m. JHD12864E is a 128 × 64 graphical LCD used for display with the simple command interface using parallel interface.

In this design, four BMP180 sensors are interfaced with TCA9548A.BMP180 andI2C devices in which the two or more are inter-integrated. TCA9548A is I2C MUX device with compatibility of connecting eight devices with high switching speed. The schematics of BMP180 with TCA9548A are shown in Fig. 3.

Now, the interfacing between sensor module and Arduino Mega 2560 which is the main controller is shown in Fig. 4. It has provision for I2C communication.

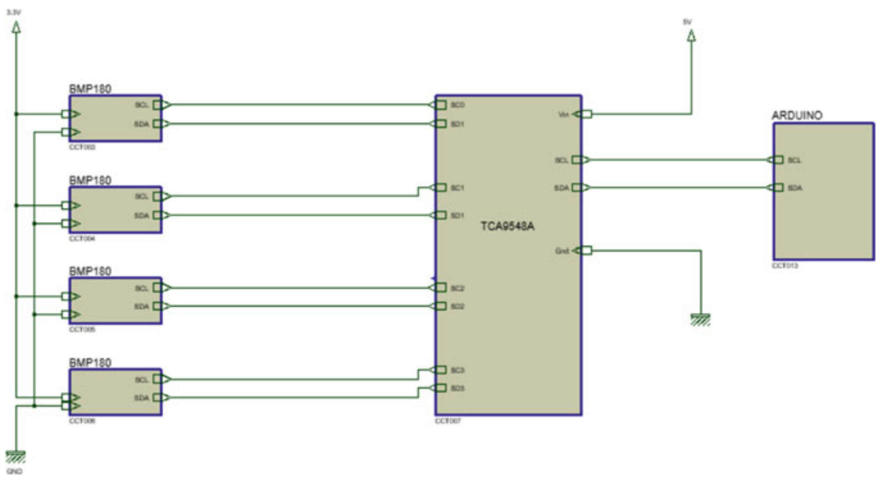


Fig. 3 Schematics of BMP180 with TCA9548A (sensor module)

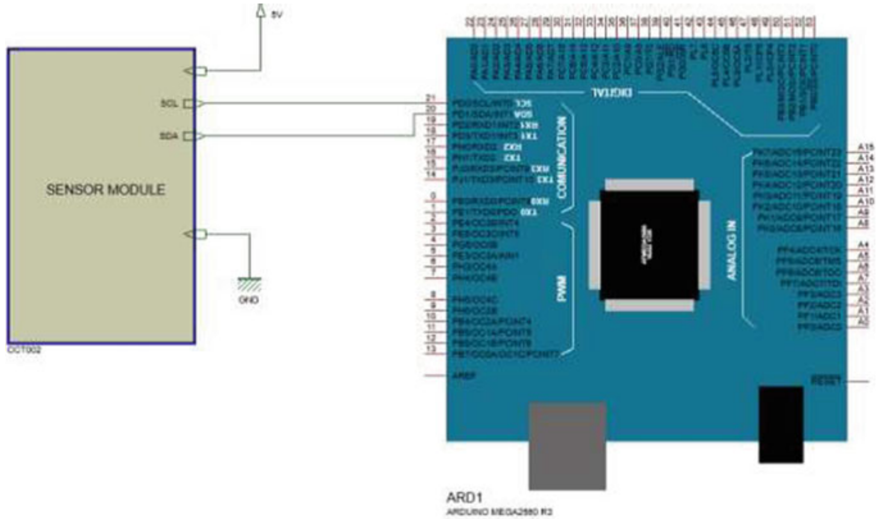


Fig. 4 Interfacing BMP180 sensor module with Arduino

3.2 *ZigBee Transmitter and GSM Interface with Arduino*

In this, the ZigBee and GSM module are interfaced with Arduino module as shown in Fig. 5. The configuration and checking are completed, and connections are made for ZigBee and GSM with Arduino; Arduino has four serial ports, and each is initialized before connecting GSM and ZigBee. In this, 14 and 15 are initialized as TX and RX. This is connected with RX and TX of GSM. Simply, RX is connected with TX of Arduino (14) and TX with Arduino RX (15).

3.3 *ZigBee Receiver and LCD Interface with Arduino*

The main blocks of the receiving section are ZigBee receiving section, Arduino, and graphical LCD. ZigBee is configured, and connections are made with Arduino. The output from the ZigBee transmitter is received through this receiver section ZigBee. The ZigBee receives the data in bytes. The data from the Arduino UNO is displayed in the graphical LCD. The ZigBee receiving pin is connected with the transmitting pin of UNO as shown in Fig. 6.

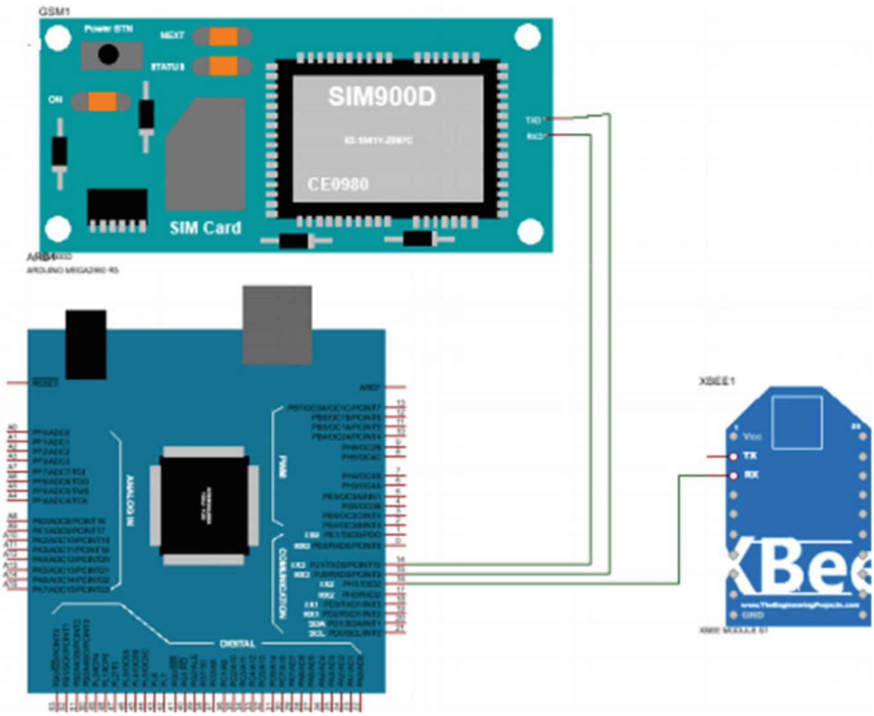


Fig. 5 Interfacing ZigBee transmitter and GSM

4 Software Interfacing

The Arduino, GSM, and the ZigBee for two-way communication are configured, and then, the data acquisition from BMP180 is done by the following flowchart as shown in Fig. 7. The algorithm starts with initializing bus scanning for an eight-bit data and sends to the input port of ADC through I2C addressing mode.

The transmitting section which includes sensing part and GSM is shown in Fig. 8. The sensor data is checked with threshold value and activated by the GSM. If the threshold is exceeded, then transmission is started for display of the data and emergency warning call or alert.

The receiving section initializes the serial port and prints the data in graphical LCD as shown in Fig. 9.

5 Results and Discussion

The building stability analyzer and earthquake detector not only make the measurement but also alert the people to evacuate from the high-rise building by automatically

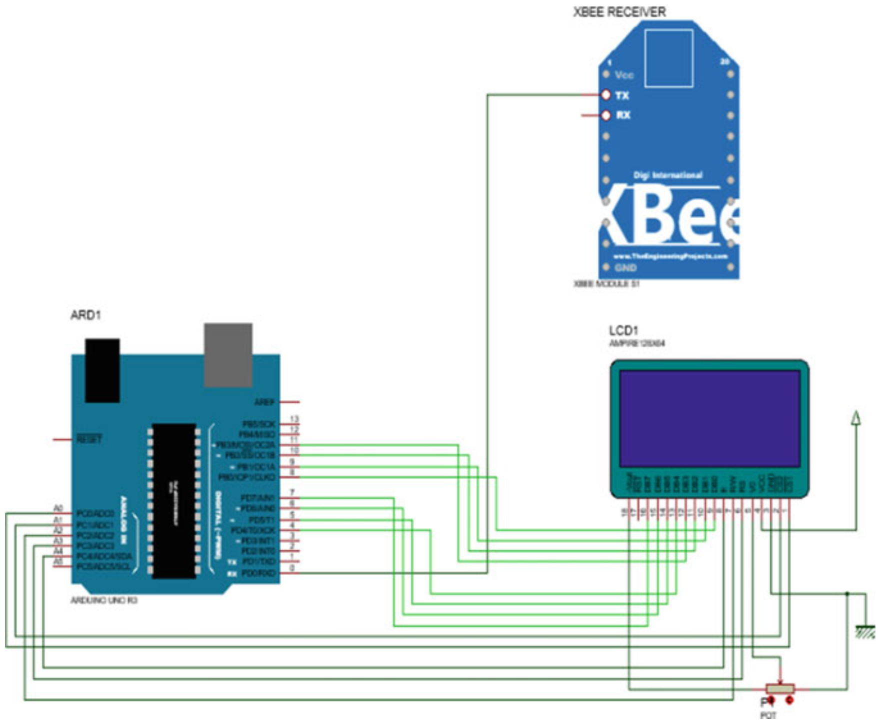


Fig. 6 Interfacing ZigBee receiver and LCD for display

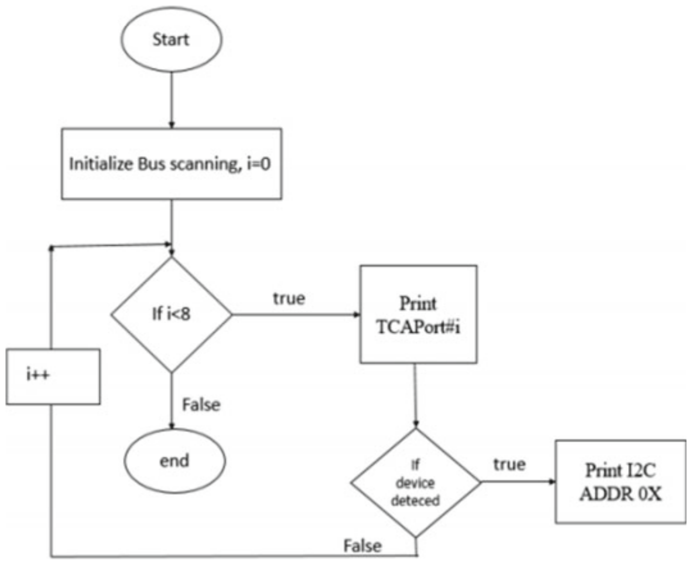


Fig. 7 Flowchart for data acquisition from sensor BMP180

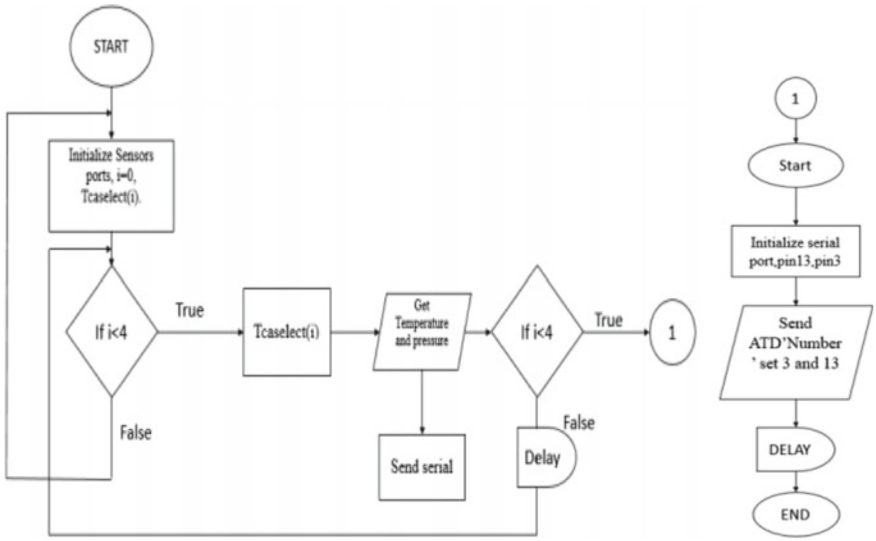


Fig. 8 Flowchart for transmitting section

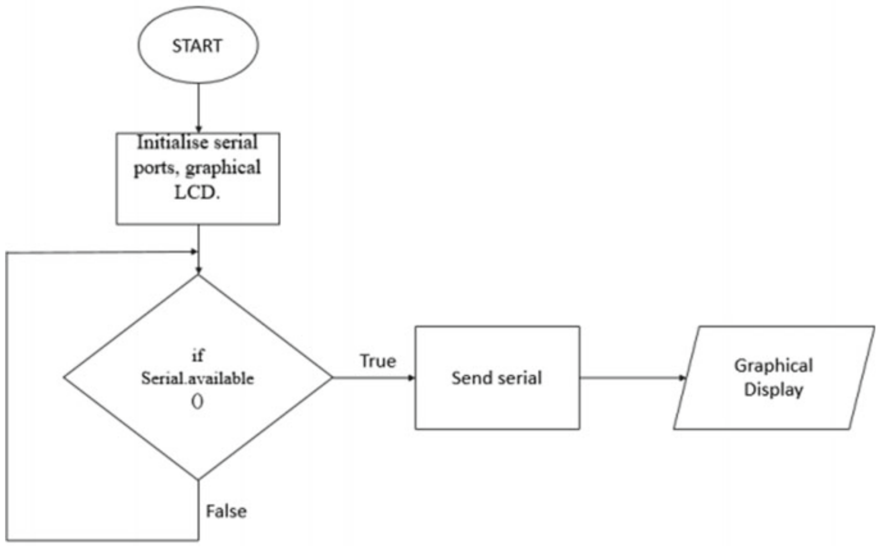


Fig. 9 Flowchart for receiving section

opening the door and sending the emergency warning call to rescue station to save the life of people. The transmitting section of the building stability analyzer is shown in Fig. 10, which consists of BMP180 sensor placed in various places of the building to improve the accuracy of measurement.

The receiving section consists of graphical LCD which continuously displays the pressure data exerted on the building on all sides of the is shown in Fig. 11.

When the pressure exceeds (the strain on the building) the threshold value, the output section sends a warning call to the rescue station and automatically opens door lock. The pressure value in the LCD display shows that if pressure exceeds 4.05 Psig, then warning call is generated which is shown in Fig. 12.

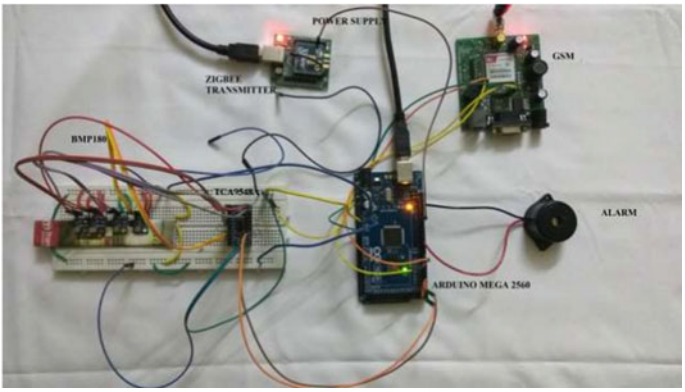


Fig. 10 Transmitting section of the building stability analyzer

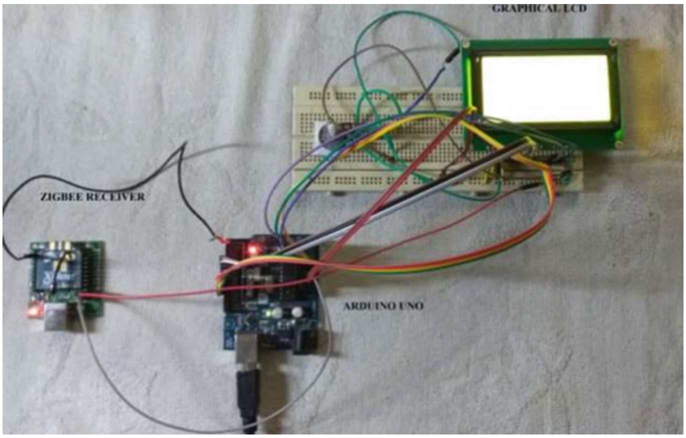


Fig. 11 Receiving section of the building stability analyzer

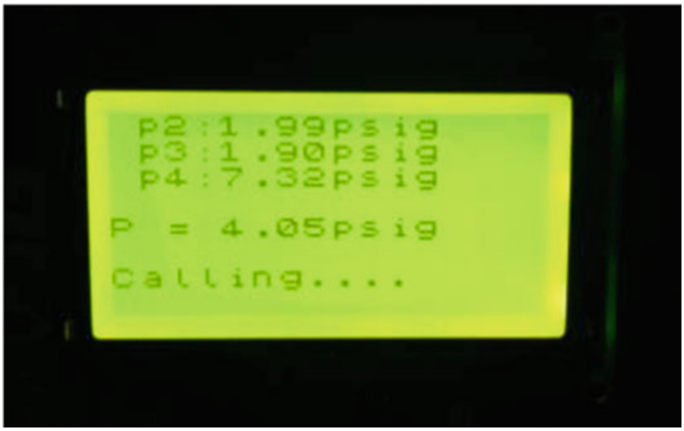


Fig. 12 Display showing the pressure exceeding the threshold and generating warning call

6 Conclusion

This paper is to appreciate the measurement of building stability along with earthquake detection by using simple pressure sensor. The accuracy of measurement is improved by placing the four BMP180 sensors in different places of the building. When the measured data exceeds the threshold value, emergency warning call is generated and an alert or alarm is generated to evacuate the people from high-rise building to avoid collapse. The automatic door opening is used to avoid the rush up of the people when earthquake strikes to avoid human delay. The continuous monitoring of stability of the building is used to have regular maintenance.

References

1. Das BM, Puri VK et al (1996) Static and dynamic active earth pressure. Geotech Geol Eng 14:353–366
2. Sirohi J, Chopra I et al (2000) Fundamental understanding of piezoelectric strain sensors. J Intell Mater Syst Struct 11(4):246–257
3. Allen RM, Kanamori H (2003) The potential for earthquake early warning in Southern California. Science 300:786–789
4. Ciang CC, Lee J et al (2008) Structural health monitoring for a wind turbine system: a review of damage detection methods. Meas Sci Technol 19(12):122001
5. Faludi R (2010) Building wireless sensor networks, 1st edn. O'Reilly, Sebastopol
6. Zheng H et al (2011) Wireless earthquake alarm design based on MEMS accelerometer. In: International conference on consumer electronics, communications and networks (CECNet)
7. Chandrakant N, Poojar S, Shenoy PD, Venogopal KR et al (2012) Middleware service oriented earthquake detection in advance using WSNs. In: Fourth international conference on advanced computing (ICoAC), pp 1–8
8. Raju B, Sagar M et al (2012) A global earthquake and flood alerting system using MEMS and GSM. Int J Sci Res (IJSR) 2132–2134

9. Hoque R, Hassan S et al (2015) Earthquake monitoring and warning system. In: International conference on advances in electrical engineering
10. Kumar A, Singh A et al (2016) A review on earth quake alarm using MEMS technology. SSRG Int J Electron Commun Eng (SSRG-IJECE) 3(8)
11. Sindhanaiselvi D, Shanmuganantham T (2018) Performance analysis of embossed diaphragm based MEMS piezo resistive pressure sensor for flood level measurement. J Mater Today 5(10):2214–7853 (Elsevier)
12. Sindhanaiselvi D, Shanmuganantham T (2018) Design and implementation of indoor tracking system using inertial sensor. J Lect Notes Electr Eng 521:465–474
13. Sindhanaiselvi D, Shanmuganantham T (2017) Double boss sculptured diaphragm employed piezoresistive MEMS pressure sensor with silicon-on-insulator (SOI). J Eng Sci Technol 12(7) (School of Engineering, Taylor's University)
14. Nallathambi A, Shanmuganantham T, Sindhanaiselvi D (2018) Design and analysis of MEMS based piezoresistive pressure sensor for sensitivity enhancement. J Mater Today 5:1897–1903 (Elsevier)
15. Sindhanaiselvi D, Ananda Natarajan R, Shanmuganantham T (2014) Performance analysis of sculptured diaphragm for low pressure MEMS sensors. Int J Appl Mech Mater 592–594:2193–2198

A Novel Technique for Low PAPR in LFDMA Systems

Lekshmi R. Nair and Sakuntala S. Pillai

Abstract Single carrier frequency division multiple access (SC-FDMA) is the standardized technology used in the Third Generation Partnership Project (3GPP) used for implementing LTE-based uplink connection. It is gaining increased popularity as the preferred technique because of its better spectral efficiency. It also has an outstanding feature of reduced PAPR because of its single carrier design. An extensive study of modern modulation techniques such as SC-FDMA has played a vital role in the design of future networks and systems. This paper analyzes the effects of selective mapping (SLM) technique on PAPR lowering of the localized frequency domain multiple access (LFDMA) network which is a kind of SC-FDMA system by means of subcarrier mapping techniques. A multistage selective mapping method is used here to bring down the PAPR of the LFDMA system. This method achieves a better reduction in PAPR than that of the traditional SLM system. The behavior of the proposed selective mapping-based LFDMA network can be obtained through finding complementary cumulative distribution function (CCDF) for PAPR characteristics. The BER feature of the presented LFDMA system is better compared to the ones reported so far.

Keywords SC-FDMA · LFDMA · IFDMA · SLM · PTS · PAPR · BER · CFO · MAI

L. R. Nair (✉)
Department of Electronics and Communication, Sree Narayana Institute of Technology, Adoor, Kerala 691554, India
e-mail: lekshmiunnikrishnan@gmail.com

S. S. Pillai
Mar Baselios College of Engineering and Technology, Trivandrum, Kerala 695015, India
e-mail: sakuntala.pillai@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_31

1 Introduction

In the recent wireless communication scenario, 5G communication system offers a wide range of mobile capabilities by providing enormous capacity, low latency, very high data rate, Internet of things facility, and many more new applications. New methods for digital broadcast are the proposed to face growing demands for increasing bandwidth requirements in wireless communications. The researchers have recommended several single-carrier transmission techniques in order to reduce PAPR in wireless communication. A new modulation technique known as single carrier frequency division multiple access which is a form of orthogonal frequency-division multiple access is found to offer a good performance in such applications [1]. It is derived from single carrier frequency domain equalization technique, and it can accommodate more number of users. The single carrier frequency domain equalization (SC-FDE) network together with concept of frequency-division multiple access (FDMA) leads to SC-FDMA network which is used in Long-Term Evolution (LTE) uplink for carrying high data rate signals.

The channel bandwidth of SC-FDMA is shared between many orthogonal subcarriers. The cyclic prefix is attached with the transmitted signal which is used as the guard time interval. The SC-FDMA link differs from the OFDMA system since that in SC-FDMA technique the signal has to pass through FFT system and then map the generated subcarrier accordingly.

SC-FDMA method has some benefits such as reduced PAPR that equips it compatible with uplink communication standard for wireless systems. SC-FDMA has a better amplitude profile, bandwidth, and speed of transmission due to its coherent single carrier property when compared with OFDMA downlink scheme [2]. Pulse shaping method using raised-cosine filter mitigates the PAPR. When underwater acoustic (UWA) channel transmission is analyzed, the results show that the average BER that can be attained using SC-FDMA is lower compared to the reported values [3].

Localized frequency domain multiple access (LFDMA) scheme and distributed frequency domain multiple access (DFDMA) scheme are two methods adopted for the subcarrier mapping in SC-FDMA systems [4]. Successive subcarriers are carrying individual consumer's signal in LFDMA system, but IFDMA system carries consumer's signal in discontinuous subcarriers. If the transmission characteristics are favorable, the LFDMA transmission can achieve frequency selective diversity as well as multiuser diversity for frequency selective fading path.

This paper presents a multistage PAPR reduction technique for LFDMA network. Here initially SLM is applied to the set of input blocks and is mapped strategically to the SC-FDMA subcarriers. As another stage, each cluster of LFDMA subcarriers will undergo SLM and generate the new set of signals for transmission. This paper analyzes LFDMA network characteristics using minimum mean square error (MMSE) equalizer after applying multistage selective mapping technique.

The organization of the paper can be presented as given: Sect. 1 presents a concept of LTE network along with its key constituents and associated with the literature survey; Sect. 2 describes the proposed methodology; and simulation outputs about BER and PAPR characteristics of LFDMA-based network are discussed in Sect. 3. Section 4 concludes the paper.

2 Methodology

2.1 Problem Statement

The OFDM system is having large PAPR due to the superimposition of all subcarrier signals. This causes the need for amplifiers with highly linear characteristics and also high synchronization accuracy for the system. The IFDMA system provides low PAPR, but the system is highly susceptible toward carrier frequency offset (CFO) similar to OFDMA system [5, 6]. Phase noise is also a responsive feature of IFDMA. Apart from this, LFDMA has advantage of huskiness toward multiple access interference (MAI) [2, 7] but it faces higher PAPR when compared to IFDMA scheme.

Partial transmit sequence (PTS) and SLM have comparable performance in terms of PAPR reduction. Though PTS has low PAPR characteristics, SLM is better than PTS when number of data vectors is taken into account. PTS shows a rise in complexity along with data vectors count. Also, the information redundancy required is higher for PTS when compared to SLM algorithm [8].

In order to overcome the main drawback of high PAPR faced by LFDMA, the proposed scheme uses a novel PAPR reduction technique for LFDMA which makes use of selective mapping technique.

2.2 Overview

The architecture of SC-FDMA system with selective mapping comprises fast Fourier transform (FFT), inverse fast Fourier transform (IFFT), subcarrier mapping, etc., in the source terminal and again discrete Fourier transform (DFT), inverse discrete Fourier transform (IDFT), subcarrier demapping, etc., in the destination terminal.

Let the data stream after modulation is given by

$$x = \left[x_0 x_1 x_2 \ldots x_{K-1}\right] \tag{1}$$

Then C copies of the data blocks are generated. Each data vector is correlated with C dissimilar sequences with length K leading to new data vectors. The phase sequence can be represented as given below.

$$B(C) = \left[b_{C,0}, b_{C,1}, \ldots, b_{C,K-1}\right] \tag{2}$$

where $C = 1, 2, \ldots$. The newly generated information block for the pth phase sequence is denoted by

$$X^{(p)} = \left[X_0 b_{p,0}, X_1 b_{p,1}, \ldots, X_{N-1} b_{p,K-1}\right] \tag{3}$$

Then these newly generated information blocks are converted to frequency domain by FFT processor which preserves the vector length same as K. The FFT processor outputs a signal that can be symbolized as follows:

$$X(u) = \sum_{u=0}^{K-1} x(u) \mathrm{e}^{-j2\pi ku/K} \tag{4}$$

In LFDMA domain, the subcarriers are mapped to V orthogonal subcarriers where $V > K$. It is then passed through V-point IFFT processor for getting time-domain information signal where $V = QK$. Q gives the extreme count for simultaneous customers which can be handled by this system. The representation of output sequence from the IFFT system is shown below:

$$x(v) = \frac{1}{V} \sum_{l=0}^{V-1} X(l) \mathrm{e}^{\mathrm{j}2\pi vl/V} \tag{5}$$

Here $\{X(l)\colon l = 0\ldots\ V - 1\}$ indicates the frequency domain signal. A cyclic prefix is appended to all the signal vectors before forwarding through the wireless medium which will act as a guard interval for the signal. Multipath transmission causes severe inter-block interference, but can be minimized by providing guard interval between signal blocks.

At destination side, after the noisy signals are received, the cyclic prefix is detached from the received signal. The signal without cyclic prefix can be written as

$$s = \sum_{a=1}^{A} H_c^a \bar{x}^a(u) + n \tag{6}$$

H_c^a is a circular matrix representing the multiple transmission path from the base station to the ath customer and n represents a $V \times 1$ vector of zero mean additive white Gaussian noise (AWGN) with variance σ^2.

Then the first stage of SLM demapping is performed on the received vector and then passed through V-point FFT to convert it to frequency base. Then demapping will be performed on the signal. The signal will undergo frequency domain equalization (FDE) process similar to the minimum mean square error (MMSE) afterward [9]. The signal coming from equalizer will be as follows:

$$x^a = wR_d \tag{7}$$

R_d is the vector corresponds to ath customer in the transmitter terminal.

Finally, K-point IDFT will change the equalized vectors into time base signal. It is then followed by the second stage of SLM demapping and the demodulation processes. The channel requirement of the network is reduced by adopting pulse shaping filters. It also minimizes the probability of deviations in output due to decoder errors [10].

2.3 LFDMA Uplink System with Selective Mapping Method

For obtaining LFDMA signals, the subcarriers in frequency base are to be mapped accordingly to obtain $\widetilde{T}_l$ can be described as follows [11]

$$\begin{aligned} \widetilde{T}_l = T_l \quad & \text{for } 0 \le l \le K-1 \\ 0 \quad & \text{for } K \le l \le V-1 \end{aligned} \tag{8}$$

T_l denotes frequency domain signals after K-point FFT. The time base signal is obtained through inverse FFT of $\widetilde{T}_l$. Let $v = W \cdot k + w$, where $0 \le w \le W - 1$ and $0 \le k \le K - 1$ [5], then

$$\widetilde{T}k = \widetilde{T}_{W.Q+w} = \frac{1}{V}\sum_{l=0}^{V-1} \widetilde{T}_l \mathrm{e}^{j2\pi lv/V} \tag{9}$$

The signals at the output of the LFDMA system will replicate the sequence applied to the system. These signals are found in the N multiple sample positions. The sequence of the input vector with separate phase factors is added together, and the value thus got will occupy the space among the sample positions. These parameters will keep the control of increase in PAPR of the LFDMA system.

2.4 Selective Mapping (SLM)

LFDMA network can utilize selective mapping (SLM) which is an easy and non-distortional method for reducing peak-to-average power ratio. The SLM method makes use of symbol scrambling. Replicas of the signals are formed with the same content, and it is multiplied using erratic phase sequences so that PAPR features are altered. The one with the lowest PAPR is found and forwarded [12]. The selected signal transmission requires a strong shield of side information for the receiver. The decoding will be easy if the receiver has got that additional information. Selective mapping technology offers reasonable for larger number of subcarriers.

The structural diagram of selective mapping methodology is given in Fig. 1. In this method, replicas of numerically independent input vectors are created at the

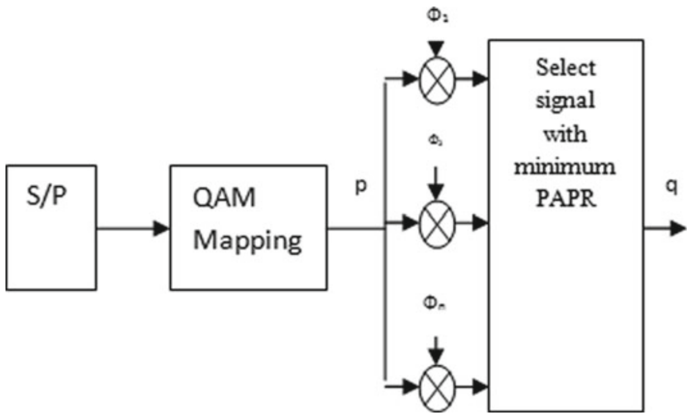


Fig. 1 Block diagram of selected mapping technique

transmitter. The signal that provides the least PAPR is chosen for forwarding to further stages.

The significance of this methodology in LFDMA is that it can handle several numbers of subcarriers which are generated due to the collaboration of FDMA and SC-FDE, and also, it does not alter the peaks present in the signal. The main challenge of this method points toward the need for additional information shield which has to be sent along with the mapped data through the channel for error-free reception of the original source data.

2.5 Peak-to-Average Power Ratio (PAPR)

PAPR is a parameter of communication network that is found from ratio of extreme value of power ($P_{\max}$) to mean value of power (P_{avg}) passing through a system. Signals with large PAPR require extremely linear power amplifiers in order to decrease intermodulation distortion [13]. Increase in PAPR will affect uplink transmission network significantly and will also lead to some power constraints. PAPR can be represented as,

$$\text{PAPR} = \frac{\text{Maximum Power}}{\text{Average Power}}$$

The proposed method is applied on LFDMA signal to balance the effects of PAPR.

2.6 Proposed Model

The proposed approach is a new way of reducing PAPR in LFDMA systems using selective mapping techniques. The multistage selective mapping model is used to symbolize and select the discrete nonlinear time-domain signal and LFDMA signal. In this design, PAPR reducing scheme is applied twice, and its BER and PAPR characteristics are observed.

In this technology, the first subsystem rotates the subcarriers once in time domain and then places them on the subcarrier class. The phase rotation sequence can be generated using complex numbers with unit magnitude. Quaternary representations such as i, $-i$, 1, -1 are commonly adopted for forming phase rotation sequence $\Phi(u)$ [13]. The increased choice of rotation can be obtained by randomizing the mapped subcarriers. After this stage, LFDMA symbol is obtained by the distribution of selected data with the lowest PAPR on subcarriers.

The second subsystem is providing a further nonlinear phase rotation for the LFDMA signal. Thus, phase rotation is applied in two stages instead of single stage. By applying the selective mapping in the second stage, this new technique can lower PAPR of transmitted time-domain sequence.

The application of the second level of selective mapping will not affect the complexity of the system [14, 15]. This method reduces the probability of occurrence of high peak signals in the time domain by a great deal. This work also recommends the optimal positions for two stages of selective mapping. Unitary, Recien, Riemann, or Hadamard phase sequence can also be used for the subsequent stages. The structural diagram for the presented LFDMA network is shown in Fig. 2. The proposed scheme is realized with the help of MATLAB 2017 software. Table 1 shows the parameters used for obtaining the resultant signal.

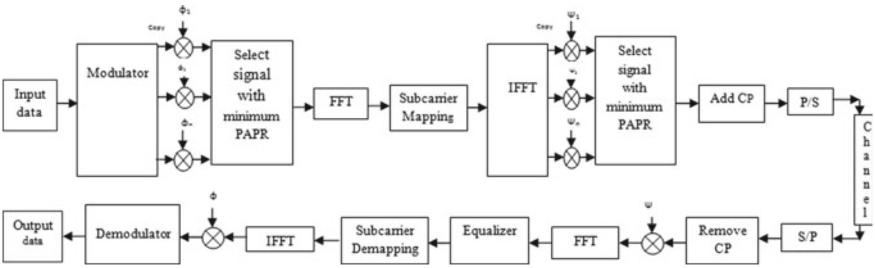


Fig. 2 Block diagram of the proposed LFDMA system

Table 1 LFDMA simulation parameters

Parameters	Values
Bandwidth	5 MHz
Frequency	15 kHz
LFDMA block size	32–128
DFT	64 bits
Subcarrier location	128–1024
Mapping	QAM/QPSK
Phase rotation	Random
Cyclic prefix	12 bits
Pulse shaping filter	RR/RC
Channel estimation	Ideal
Equalizer	MMSE

3 Results and Discussion

3.1 PAPR Analysis

The performance of the proposed LFDMA system has been analyzed by various simulations. Here the results for reducing PAPR in the proposed LFDMA system are discussed. CCDF characteristics of peak-to-average power ratio for the LFDMA systems have been analyzed for two types of pulse shaping filters. 10^5 blocks of data have been generated for calculating the CCDF of the PAPR.

Figure 3 analyzes the PAPR performance of LFDMA system with root-raised-cosine (RR) filter. It is clear that the proposed system with RR filter has the lowest PAPR that differs the LFDMA system without SLM and with RR filter by about 7 dB and SLM-based LFDMA system with RR filter by about 5.5 dB for 8-QAM.

For LFDMA system with raised-cosine (RC) filter, the PAPR is high compared to LFDMA with RR filter for both QPSK and QAM. In Fig. 4, PAPR of the proposed system with RC filter is 5.5 dB less than LFDMA system without SLM and using RC filter and 4 dB less than LFDMA system with conventional SLM using RC filter for 8 QAM.

The CCDFs of the PAPR in LFDMA network applying RR filter with different input data block sizes for 8-QAM base are shown in Fig. 5. There is a PAPR variation of approximately 0.25 dB as data block size changes.

The CCDFs obtained for PAPR in the LFDMA network using RC filter with different modulation formats are shown in Fig. 6, and the dependency of the modulation scheme can be observed from the curves. PAPR in LFDMA network for 8-QPSK and 8-QAM differs by about 0.25 dB.

Figure 7 shows PAPR characteristics of the LFDMA network with choice of phase offsets which is related to the choice of the resource unit. PAPR does not have significant change with change in phase offset values.

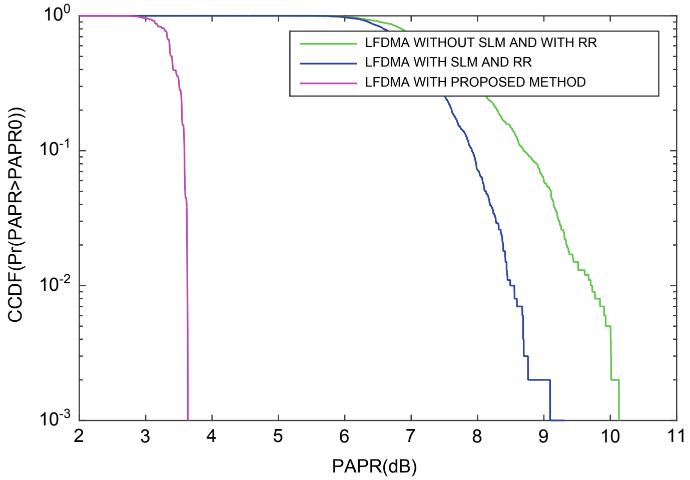


Fig. 3 CCDF characteristics of the proposed LFDMA network with RR filters

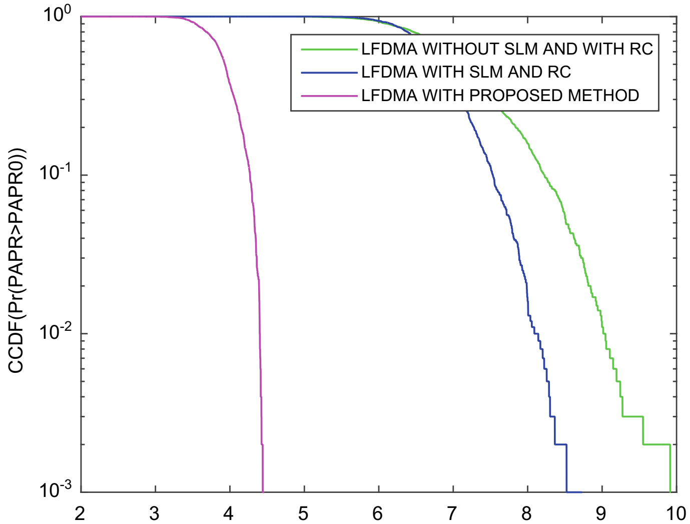


Fig. 4 CCDF performance obtained for the proposed LFDMA network with RC filters

3.2 BER Analysis

To analyze the BER preservation property of the proposed LFDMA system, BER performance of conventional LFDMA system and the proposed LFDMA system are obtained. Figure 8 shows the BER characteristics of MMSE equalizer model LFDMA

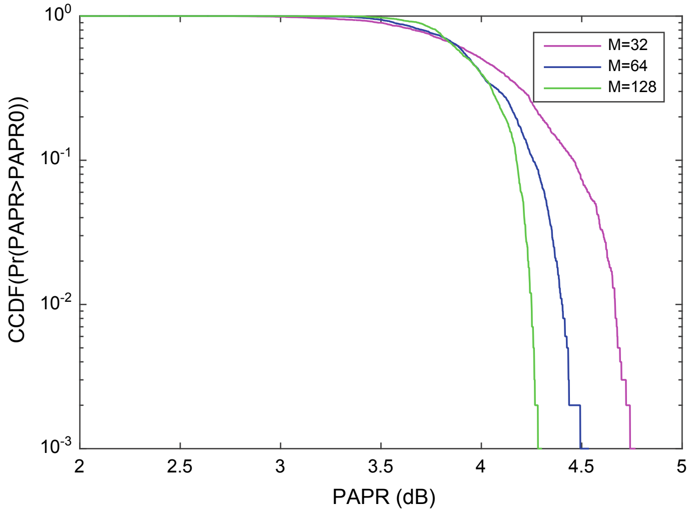


Fig. 5 CCDF performance of SLM technique for different input block size

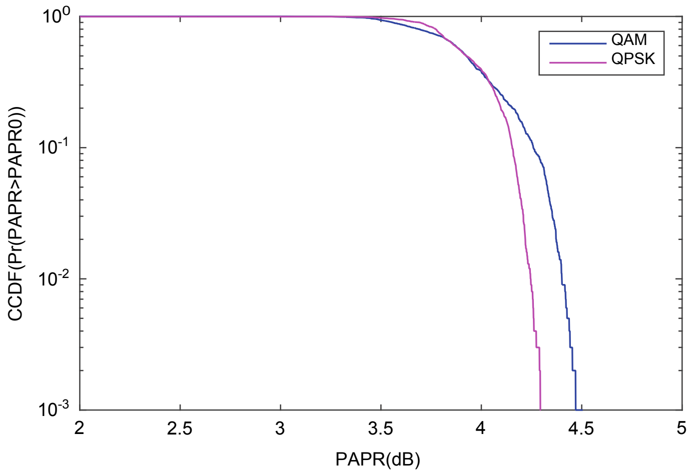


Fig. 6 Proposed LFDMA signal CCDF performance for different modulation schemes

system. It is seen that MMSE model of the proposed LFDMA system attains a slightly greater but comparable BER performance as that of general LFDMA system.

The simulations show the comparison of CCDF and BER characteristics of the proposed LFDMA network. LFDMA using the proposed SLM method is having the lowest PAPR compared to LFDMA system with conventional SLM and LFDMA system without using SLM. Complexity is slightly high for the proposed system. But when compared to PTS, as the source data vector size goes high, complexity is

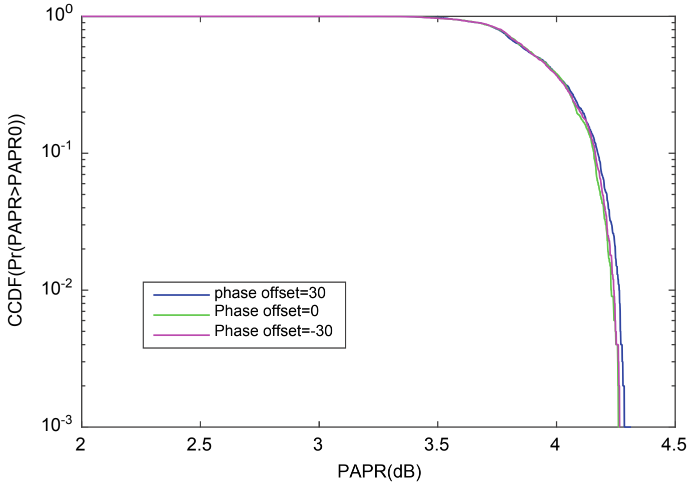


Fig. 7 CCDF performance of SLM technique for different phase offsets

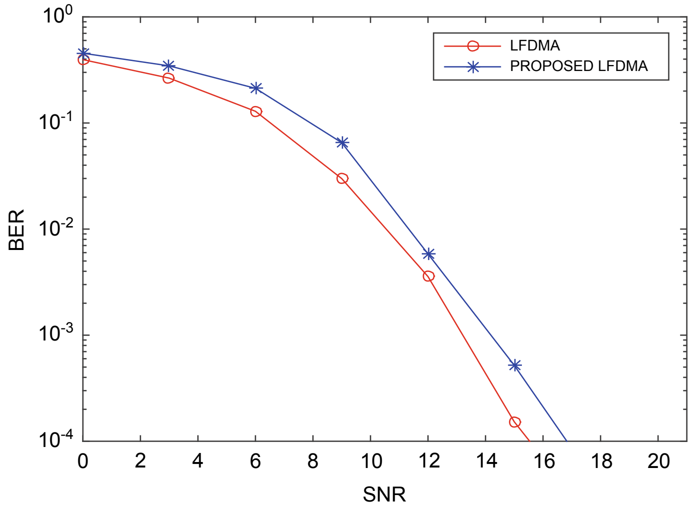


Fig. 8 Variation of BER as a function of SNR in the proposed LFDMA system

less since the phase rotation sequence trails required to obtain the lowest PAPR are reduced.

4 Conclusion

This paper presents a unique method to generate data block sets of LFDMA network by applying phase rotation to various stages of transmission. This is a simple and spectrum efficient technique based on selective mapping methodology for lowering PAPR in LFDMA network. The phase transformation is executed in two stages; it is initially employed to the modulated data blocks before FFT stage and again to the LFDMA subcarriers after IFFT stage. Here ergotic phase sequences are utilized, and the most effective LFDMA symbol with respect to PAPR value is chosen in both the stages for a drastic reduction in PAPR. The system is realized and analyzed for both RR and RC filters. Simulation results show that the proposed LFDMA system dramatically reduces PAPR of LFDMA output without causing severe BER shift when compared to traditional LFDMA with RR/RC filter. The results prove that the substantial LFDMA method explained in this paper achieves superior PAPR reduction performance compared with LFDMA-SLM while offering comparable computational complexity.

References

1. Gazda J, Drotar P, Galajda P, Kocur D (2010) Comparative evaluation of OFDMA and SC-FDMA based transmission systems. In: IEEE 8th international symposium on applied machine intelligence and informatics (SAMI), pp 177–181. IEEE, Herlany
2. Khandekar AD, Palanki R, Sutivong A (2017) Code division multiplexing in a single carrier frequency division multiple access system. U.S. patent, June
3. Dixit S, Tripathi P, Shukla M (2015) SC-FDMA-IDMA scheme for underwater acoustic communications. In: International conference on communication, control and intelligent systems (CCIS), pp 204–207. IEEE, Mathura
4. Myung HG, Lim J, Goodman DJ (2006) Peak-to-average power ratio of single carrier FDMA signals with pulse shaping. In: IEEE 17th international symposium on personal, indoor and mobile radio communications (PIMRC 06), pp 1–5.IEEE, Helsiniki
5. Myung HG, Goodman DJ (2008) Single carrier FDMA: a new air interface for long term evaluation. Wiley, Chichester, U.K.
6. Wilzeck A, Cai Q, Schiewer M, Kaiser T (2007) Effect of multiple carrier frequency offsets in MIMO SC-FDMA systems. In: International ITG/IEEE workshop on smart antennas, IEEE, Vienna, Austria
7. Estevez C, Molina J, Azurdia-Meza CA (2016) Subcarrier mapping distribution effect on single carrier FDMA transmissions. In: 8th IEEE Latin-American conference on communications (LATINCOM), pp 1–5, IEEE, Medellin, Colombia
8. Muller SH, Huber JB (1997) A comparison of peak power reduction schemes for OFDM. In: IEEE GLOBECOM'97, pp 1–5. Phoenix, AZ
9. Okuyama S, Takeda K, Adachi F (2010) MMSE frequency-domain equalization using spectrum combining for Nyquist filtered broadband single-carrier transmission. In: The 71st IEEE vehicular technology conference (VTC 2010, Spring). IEEE. Taiwan
10. Rappaport TS (2002) Wireless communications principles and practice, 2nd edn. Pearson Education, Indianapolis, IN
11. Baig I, Ayaz M, Jeoti V (2013) A SLM based localized SC-FDMA uplink system with reduced PAPR for LTE-A. J King Saud Univ Eng Sci, pp 119–123

12. Breiling M, Mller-Weinfurtner SH, Huber JB (2001) SLM peak-power reduction without explicit side information. IEEE Commun Lett 5(6):239–241
13. Harne S (2016) Zanjade: a PAPR reduction in OFDM system using phase sequence of Riemann matrix. Int Adv Res J Sci Eng Technol 3(9):56–60
14. Mohammady S, Sidek RM, Varahram P, Hamidon MN, Sulaiman N (2013) A low complexity selected mapping scheme for peak to average power ratio reduction with digital predistortion in OFDM systems. Int J Commun Syst (IJCS) 26(4):481–494
15. Baxley RJ, Zhou GT (2007) Comparing selected mapping and partial transmit sequence for PAR reduction. IEEE Trans Broadcast 53(4):797–803

Apparel CPW-Fed Antenna for Medical Anatomy Zone Web Applications

V. Sravani, S. Ashok Kumar and T. Shanmuganantham

Abstract In this paper, a coplanar waveguide (CPW)-fed antenna is presented for medical anatomy region grid. The proposed antenna is operating at 2.45 GHz ISM band frequency, and it is simulated using IE3D simulator version 15. When it works close to human tissue, large specific absorption rate will occur. Based on flexible substrate and miniaturization, SAR value is decreased up to 95%. The results of proposed antenna produce a band from 2.4 to 2.48 GHz. The simulations are fulfilled in a body for validation. The antenna proposes a fine accordance between the simulated and analysis results to examine the range of the data transmission.

Keywords CPW · ISM band · Wearable system · Body area networks

1 Introduction

In wearable systems, medical body region grid is getting essential technology telemetry, and medicinal services are the applications of MBAN [1]. This is also utilized by hospitals and outpatient surgeons. It includes monitoring hypertension and electrocardiogram. MBAN is also used for wireless communication machines, and it has low heaviness and low profile which can be used for designing periodic examination [2, 3]. For wireless communication system, antenna can be used as a key and it directs to operate in efficient manner.

V. Sravani (✉) · S. Ashok Kumar
Jyothishmathi Institute of Technological Sciences, Karimnagar, India
e-mail: sravaniveerabhathini2@gmail.com

S. Ashok Kumar
e-mail: ashokape@gmail.com

T. Shanmuganantham
Pondicherry University, Puducherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_32

A challenging task of wireless communication system is designing narrowband and wearable antenna of sustainable efficiency and presentation [4–6]. Positioning the wearable antenna near by human body causes performance reduction, since capacitive coupling between the human body and the antenna alters the operating frequencies. Transmission errors can occur because of difference in the radiation structure [7–9]. Antenna should be able to perform well close to anthropoid tissue with very light propagation towards an anatomy to observe the particular absorption rate limit. 2.4 GHz is the demand frequency of the wearable antenna which has to be operating with and without flexible materials. Coplanar Waveguide (CPW) antennas, conventional patches, slot patches and monopole antennas, subsequent versatile antenna designs with improved efficiency and performance have appeared [10, 11]. Monopoles and inverted-F fabric receiver have been well recorded, and even in small area, they may extend SAR quantity when operating close to the material attained by antenna designs. Large footprints are maybe too wide and large for MBAN applications which are the drawbacks of these designs. Semi-flexible materials are not appropriately bendable. Because of bending or poor front-to-back ratio, it may not comfortable for users.

2 Representation of Antenna Designs

Figure 1 represents antenna design over EBG-FSS which is in the form of C. This slot is fabricated to get density in size. A coplanar waveguide is attached to the below of the chip which consists of single strip of width. There are two orthogonal marks which help the receiver as floor flat surface. Fabrication complexity of the antenna is reduced by CPW feeding method because at top of the dielectric ground planes, the radiator is placed as shown in Fig. 1.

If increasing the trimming capacitor or shape of the conducting coil under the frequency, increasing of inductance is better in place of the capacitance. In the absence of enlarging the conducting coil by using a dark substrate, the inducting coil can grow, but a wide substrate occupies large area and increases the machines. It shows that the total formation inductance increases because of the extra plane inducting coil. So, high-inducting coil FSS as peak layer is mandatory to use in place of a traditional patch that acts as a conducting coil plane. Between suspended line and ground, the EBG-FSS array is placed, to form a sandwich-like design. Between two waveguides, the removed line is placed during simulation and the evaluation of S parameters is shown in Fig. 2.

3 Presentation of Receiver Than EBG-FSS and PEC

The normal alloy and regular EBG-FSS designs work to reveal the effectiveness of EBG-FSS. The receiver is placed with same height and dimensions, and value

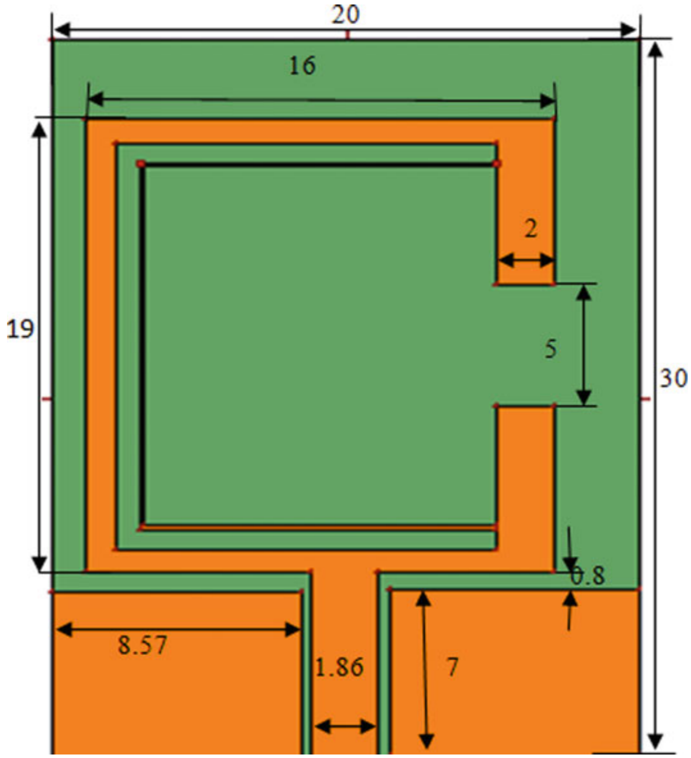


Fig. 1 Geometry of antenna

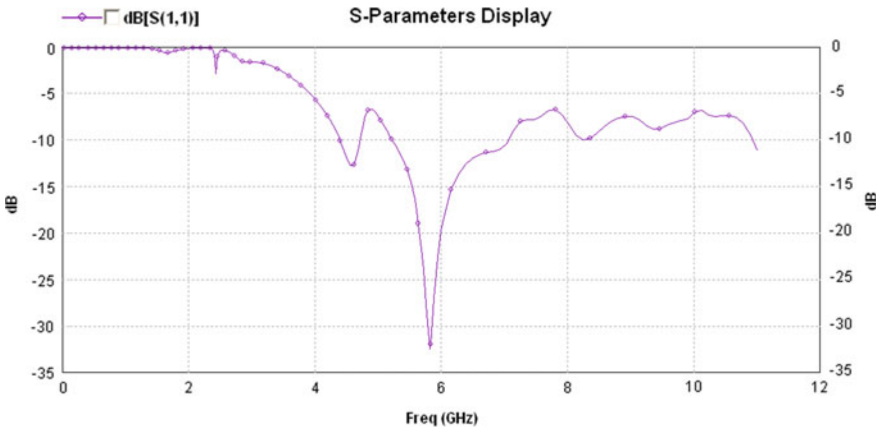


Fig. 2 *S*-parameter display

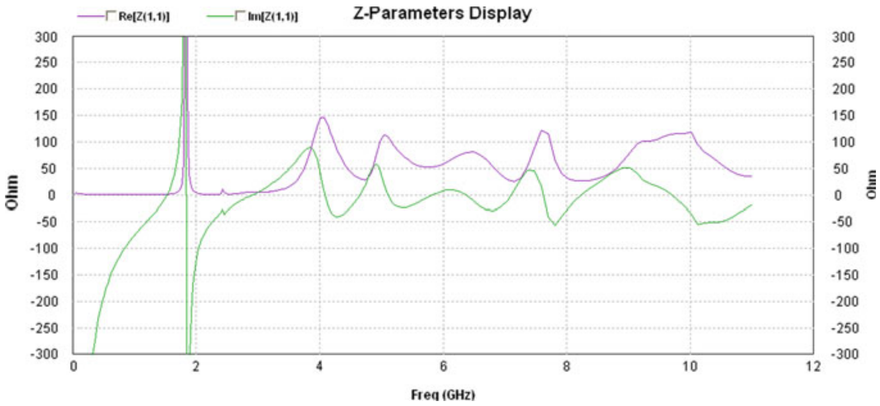


Fig. 3 *Z*-parameter

of antenna is more than −10 dB which is almost linear. Furthermore, to get the better reflection coefficient at 2.4 GHz, EBG-FSS is added to antenna. Because of the periodic action of the EBG-FSS, less impedance matching greater than −10 dB indicates the overall antenna with PEC. The receiver design of EBG-FSS shows good impedance matching with less than −10 dB at 2.4 GHz. So, EBG looks like a good option for an apparel low-profile receiver.

Action of an Antenna on Body

In many requisitions, an apparel receiver is anticipated to damage or obey the anthropoid tissue planes through surgery, frequency and features which manage below bending. Therefore, before using this antenna on human tissue, make sure that antenna is perfect to use on human body. These are taken depending on anthropoid lower limb, upper limb and breast sizes. Sticky tape is used for fixing the antenna on the pol. This leads to bending along the *y*-axis. To decrease the diameter, there must be a slight shift on higher frequency (Fig. 3).

Figure 4 deals with measuring the antenna efficiencies accepting the simulated efficiency results. The proposed theory of multi-input and multi-output system has good efficiency over isolation and simulation. Here, MIMO antenna is a wireless communication technology. It is used at source and destination. The two dimentional radiation pattern of proposed antenna performing well in various angles and the elevation (Fig. 4) and azimuth (Fig. 5) pattern shows maximum gain of 1 dBi at resonant frequency. The efficiency of the antenna at resonant frequency is 75% as shown in Fig. 7.

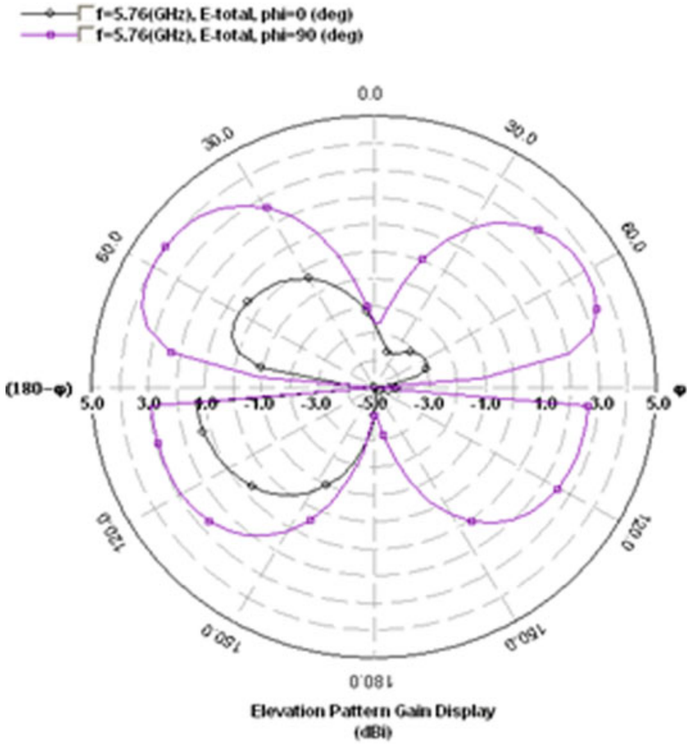


Fig. 4 Elevation pattern

4 Conclusion

A medical body area network application using CPW-fed antenna with 2.4 GHz has been presented. Based on fabric materials, the proposed antenna was designed which is mixed with our daily fabrics. In free space, the antenna was studied and compared to the normal antenna SAR reduced by more than 95%. Therefore, for future MBAN applications, this CPW antenna is applicable.

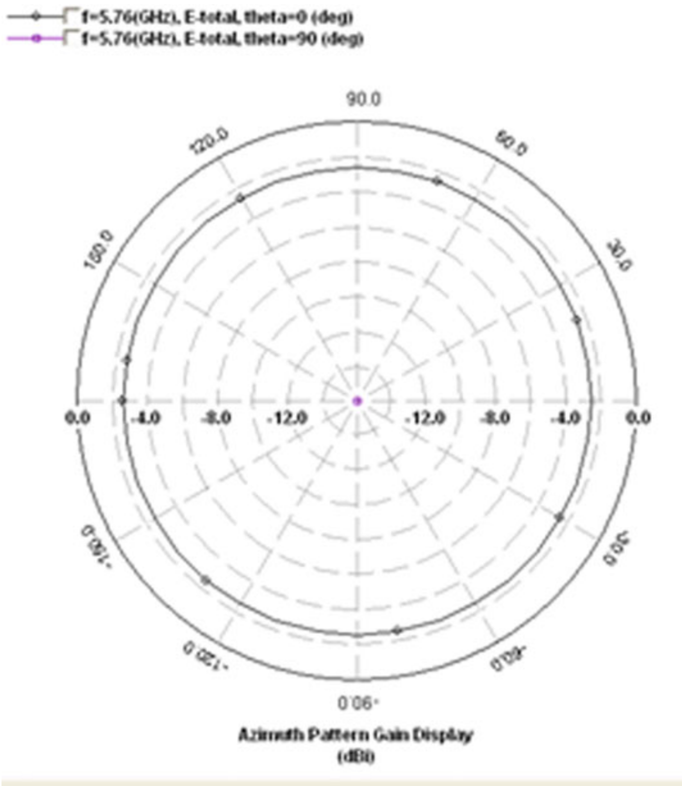


Fig. 5 Azimuth pattern

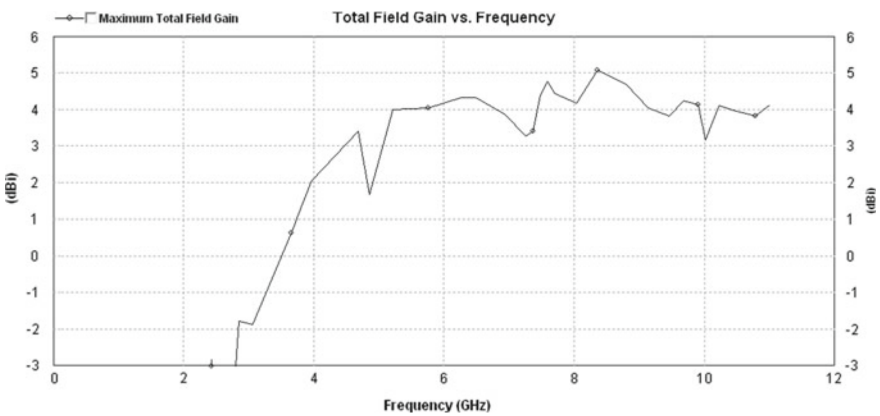


Fig. 6 Field gain versus frequency

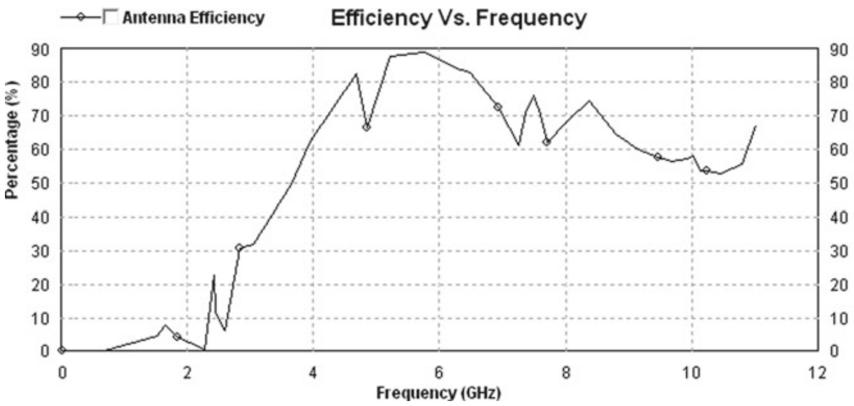


Fig. 7 Efficiency versus frequency

References

1. Ashyap AYI, Abidin ZZ, Dahlan SH (2018) Highly efficient wearable CPW antenna enabled by EBG-FSS structure for medical body area network applications. IEEE Access 6
2. Wang J, Wong H, Ji Z, Wu Y (2019) Broadband CPW-fed aperture coupled meta surface antenna. IEEE Antennas Wirel Propag Lett 18
3. Niamien MAC (2019) Spectral domain-based input impudence prediction of a CPW-fed notch antenna. IEEE Antenna Wirel Propag Lett 18
4. Ketavath KN, Gopi D, Rani SS (2019) In-vitro test of miniaturized CPW-fed implantable conformal patch antenna at ISM band for biomedical applications. IEEE Access 7
5. Kumar SA, Thangavelu S (2015) CPW fed implantable Z-monopole antennas for ISM band biomedical applications. Int J Micro Wire Tech 7:529–533
6. Karthikeya GS, Abegaonkar MP, Koul S (2019) CPW fed wideband corner bent antenna for 5G mobile terminals. IEEE Access 7
7. Yang Y, Zhao Z, Ding X, Nie Z, Liu Q-H (2019) Compact UWB slot antenna utilizing traveling-wave mode based on slot line transitions. IEEE Trans Antennas Propag 67
8. Li YJ, Lu ZY, Yang LS (2019) CPW-fed slot antenna for medical wearable applications. IEEE Access 7
9. Cheng Y-F, Ding X, Xu X, Zhong X, Liao C (2019) Design and analysis of a bow-tie slot-coupled wideband metasurface antenna. IEEE Antennas Wirel Propag Lett 18
10. Ashok Kumar S, Shanmuganantham T (2018 Oct) Design and performance of textile antenna for wearable applications. Trans Electr Electron Mater 19(5):352–355
11. Jun SY, Izquierdo BS, Parker EA (2019) Liquid sensor/detector using an EBG structure. IEEE Trans Antennas Propag 67

A Reduced Planar Omnidirectional MIMO Antenna for Pattern Phase Variety

P. Manaswini, S. Ashok Kumar and T. Shanmuganantham

Abstract A reduced planar omnidirectional MIMO antenna accompanied by pattern phase variety is to be discussed in this paper. Because of the restricted bandwidth of the perpendicular-polarized element, the shape of usual MIMO antennas with two polarizations has extremely occurred in two-dimensional integration. Hence, for planar integration as for omnidirectional MIMO antenna, a two-slot parallel-polarized antenna is introduced. This antenna is made up of folded dipole array with four elements, and to feed the two channels in the company of shared elements, a close packed feed network having a pair of 90° continuous phases is hired, so this integration of antenna can be done in a sheet of substance with 0.8 mm ultra-low profile. Even though the polarization and radiation diagram of both the channels located adjacent to each other are similar to one another, as a consequence of unrelated pattern phase, an advanced isolation is yet achieved. So, a possible solution for planar integration of the system is done by the proposal of dual-channel antenna.

Keywords MIMO system · Pattern phase variety · Multi-band

1 Introduction

Nowadays, the channel capacity polarization of two antennas are more extensively recruited in MIMO systems. Besides, omnidirectional pattern antennas received important attentions in some of the wireless communication systems like WLANs and base stations. Hence, the dual-polarized antennas with omnidirectional patterns

P. Manaswini (✉) · S. Ashok Kumar
Jyothishmathi Institute of Technological Sciences, Karimnagar, India
e-mail: manaswinipuranam@gmail.com

S. Ashok Kumar
e-mail: ashokape@gmail.com

T. Shanmuganantham
Pondicherry University, Puducherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_33

are been focused by many researchers [1]. Some wideband or multi-band with polarization of two antennas is presented with biconical antenna being perpendicular polarization and the circular dipole array being parallel polarization. AMC reflector is added to decrease the outline of the structure [2, 3]. This antenna is used for perpendicular polarization, while the cross bow-tie dipole is used for parallel polarization. These horizontal and vertical slots are merged for the VP together with HP. A novel sabre-like shape is introduced as declining monopole for VP and the denseness of double-polarized antenna in the company of booming cavity for parallel polarization [4–6]. The aforementioned antennas are all three-dimensional (3D) constructions which are effective, and it is very difficult to integrate in planar platform. Some planar omnidirectional MIMO antennas with dual polarization are presented to satisfy the needs of the two-dimensional integration [7]. Although, the proposed antennas are shows narrow bandwidth to the conventional antennas which shows low profile omnidirectional element with perpendicular polarization. To get a solution, a planar two-channel parallel-polarized antenna is presented.

For a reason that both the channels are horizontal polarized, bandwidth of introduced antenna is limitless. The huge separation in the middle of two co-polarization channels is attained by a pattern phase variance approach [8]. Over conventional systems, polarization variance, radiation pattern variance, and spatial diversity procedures are utilized for the increasing of port isolations in the middle of various channels. Over polarization variance, two rectangular emissions are used for decoupling in the middle of two channels. In spatial variance, parasitic formations, slot printing, EBG structures, and neutralization lines are put in the application to decrease reciprocal mixing in the company of close elements [9].

In radiation pattern diversity, for decoupling in the middle of multiple channels, three dissimilar ways are present. In the first way, the utilization of distinct types of patch antenna gives rise to broadside and cone-shaped propagation patterns [10]. The following technique may register in a single radiator, so the structure of this antenna is very simple. The second way uses contrasting radiator to create different beam-pointing patterns [11]. In the third process, usage of unlike feed phases for the design of antenna arrays in different radiation patterns. All the pattern diversity techniques which are shown above introduced are based on amplitude variance of radiation patterns, but they cannot generate two separated omnidirectional radiation courses.

2 Fundamental Research

The geometrical measurements of an antenna the 19.3 mm, 17.4 mm breadth and length, respectively, and the thickness of the antenna is 1 mm. The proposed antenna is easy to fabricate with the relative permitivity of substrate is 1.36 and loss tangent is 0.002. There are three slots radiation element placed on top layer of antenna. The bottom layer is act as ground layer and covered with the part of copper metal. The

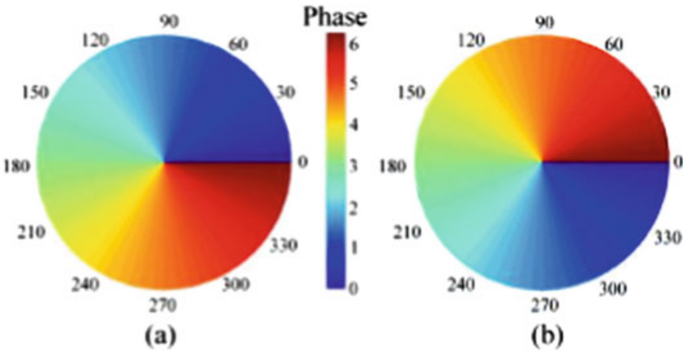


Fig. 1 Pictorial representation of radiation pattern in azimuth plane. **a** Enlarging phase distribution, i.e., +1 mode, **b** declining phase distribution, i.e., −1 mode

proposed MIMO antenna is connected with inner conductor consists of feed point and ground is connected to back of antenna. It is layered antenna.

2.1 Analysis of Pattern Phase Variance

In a MIMO technique, transmission capability may increase if and only if antenna components are related to each other very delicately; thus, the aim of the MIMO antenna is to lower the connection in the middle of every channel [12]. The ECC is a remarkable variable to estimate the relationship in the middle of two channels, and it may have been studied before composite radiated outlying area.

The two divergent phase distributions of far field in the company of omnidirectional diffraction model in azimuth plane have been accompanied in Fig. 1. In Fig. 1a, period of the radiation pattern enlarges against $0 - 2\pi$ in the company of enlargement of φ, consider it as +1. From Fig. 1b, angle decreases against $2\pi - 0$ in the company of enlargement of φ and is considered as −1. Make a note that optical phenomenon and radiation beam of particular two modes are close to one another. For numerically rating of the variance presentation of pattern phase variance technique, ECC in the middle of the dual modes is determined by considering only composite electric field in azimuth plane. Hence, by pattern phase variance procedure, the registration of pattern phase variance is done.

2.2 Antenna Formation

Systematic scheme to reach parallel polarization radiation pattern in all directions is engaging the dipole array which gives rise to a zero-order loop mode accompanied by in-angular currents, which is similar to emission of a magnetic current. Although the design aspect of zero order loop is not changed in azimuth plane, so it is not suitable for the design aspect of various process. For MIMO medium 1, the elevated angle is left-handed 90° developing phase which is explained as +1, and the phase of antenna which is far away from the field is same as shown in Fig. 1a. While for MIMO medium 2, elevated angle is right-handed 90° progressive which is explained as −1, same as Fig. 1b. However, the two mediums share the similar dipole array of four elements, and a high separation with design aspect variation process can be attained.

2.3 Antenna Evolution

As, before mentioned 90° developed phase, a straight dipole array with four elements does not function as represented due to coupling in the middle of neighboring elements. If port 1 and port 3 are elevated using same amplitude with opposite period while port 2 and port 4 are contested, the electric fields of two elevated dipoles are opposite to one another. The correlative coupling in the middle of the nearest elements will decline the active *S*-parameters of the components. Although separations in the middle of dipole elements are stronger when compared to 15 dB, *S*-parameters of ports 1 and 2 are completely dissimilar with harshly declines of port 2 and port 4.

3 Plan of Antenna

3.1 Antenna Layout

The ultimate design for suggested two-channeled two-dimensional folded dipole array antennas is presented in Fig. 2, and dimensions are given in Table 1. One thick feed point has been stamped on upper side (orange) with thickness of 0.8 mm with $\varepsilon_r = 4.4$, $\tan\delta = 0.02$, whereas ground plane as for feed point including two-dimensional folded dipole array is stamped above the rear side (green) of FR-4 substrate. That being of complete antenna may add in one sheet of FR-4 board. A feed point has been created of a 3-dB connector, a crosswalk along with four microstrip lines. The resistivity match of dipoles can be tuned by Lf1 and Lf3 dimensions. Ground plane's size should be small enough to keep away from attacking the omnidirectional radiation pattern.

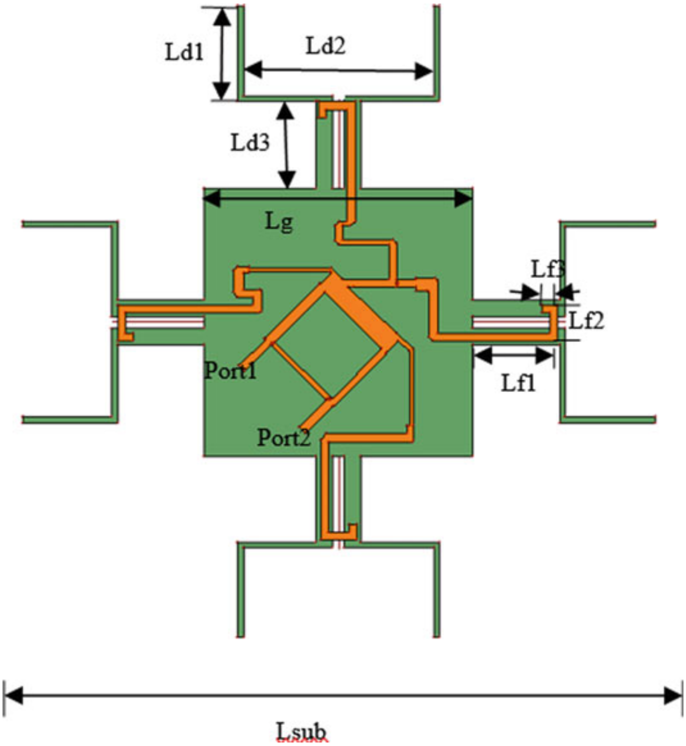


Fig. 2 Geometry of wearable antenna

Table 1 Antenna parameters in mm

Lg	Lf1	Lf2	Lf3
48	17	6.5	2.8
Ld1	Ld2	Ld3	Lsub
17	36	15.5	120

3.2 *Integrated Feed Point*

To keep away from influencing radiation pattern as for suggested antenna, a compressed feed point in the company of two sets of 90° continuous phases is schemed in the midpoint of the dipole array to be anti-discriminated. The presented feed network is having two input ports which are port 1 and port 2, and we have four output ports which are port 3 to port 6. Record that integral lines of port 3 and port 5 are opposite to one another. When feeding through port 1, the phases on output port 3 to port 6 are 0°, 90°, 180°, 270° and for port 2 are 90°, 0°, 270°, 180°. So, orthogonal phase excitations with two sets can exist engaged with pattern phase variance, have been gained. The feed point with features of amplitude, a good impedance match, and suggested production in the middle of two input ports are acquired.

3.3 Parameter Analysis

Size of Lg is a significant framework to modify the omnidirectional radiation pattern. The verified changes in gain at 5.8 GHz with Lg varied are estimated. For both port 1 and port 2, two least gain changes are observed with optimization of Lg to get 48 mm, and gain decreases quickly by increasing of Lg. So, for getting a fine radiation pattern, the size of ground plane should reduce for an optimal dimension, that is limited to design of two-channeled 90° continuous phase feed plane. As given previously, distance in the element is with great importance to dynamic resistivity matching. A weak matching with dynamic resistivity reflects energy against dipole elements, which leads to bad separation in the middle of two channels. The separation in the middle of two channels with element distance Ld3 changed. The separation deteriorates quickly accompanied by distance of the element which is different, and an optimal separation of the antenna is gained with Ld3 = 15.5 mm.

4 Results and Discussion

Isolation and reflection coefficient in the middle of the two channels is calculated by using IE3D simulator. The reflection coefficient S11 and S22 are exhibiting lower return loss above −10 dB across 2.6–5.8 GHz as shown in Fig. 3. It shows broader bandwidth around 3.2 GHz between the resonant frequency. The obtained separations have a good regularity, and evaluated highest separation is 31.7 dB at frequency of 3.1 GHz. Moreover, from the *s*-parameter display, it is observed that return loss is above −20 dB and frequency is in between 2.6 to 5.8. The proposed antenna shows perfect impedance of 50 ohm matching at resonant frequency of 3.1 GHz as shown in Fig. 4. Width of the T-structured component is 5 (it should be constant). And the height of the T-structured element can be changed. By increasing the U-shaped element height, the gain is decreased. So, it is better to keep the U-shaped element height as constant.

4.1 Antenna Performance

When antenna is attached to the vicinity of patient tissue, then resonance frequence is shifted to greater frequency and impedance matching is destroyed. The presented antenna shows 2D pattern of radiation pattern of both elevation and azimuthal planes are shown in Figs. 5 and 6. It radiates maximum gain of 6 dBi between operating

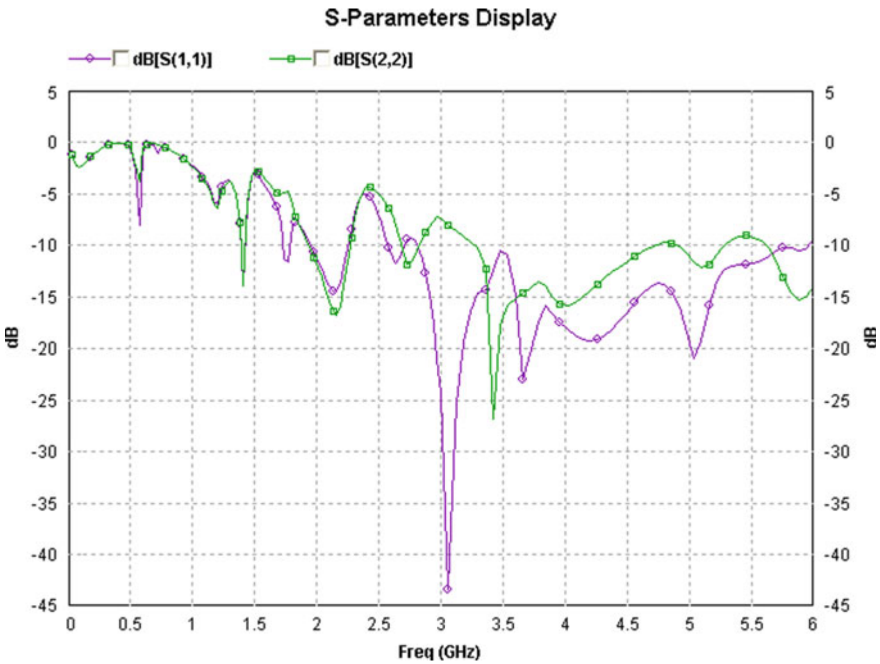


Fig. 3 *S*-parameter display

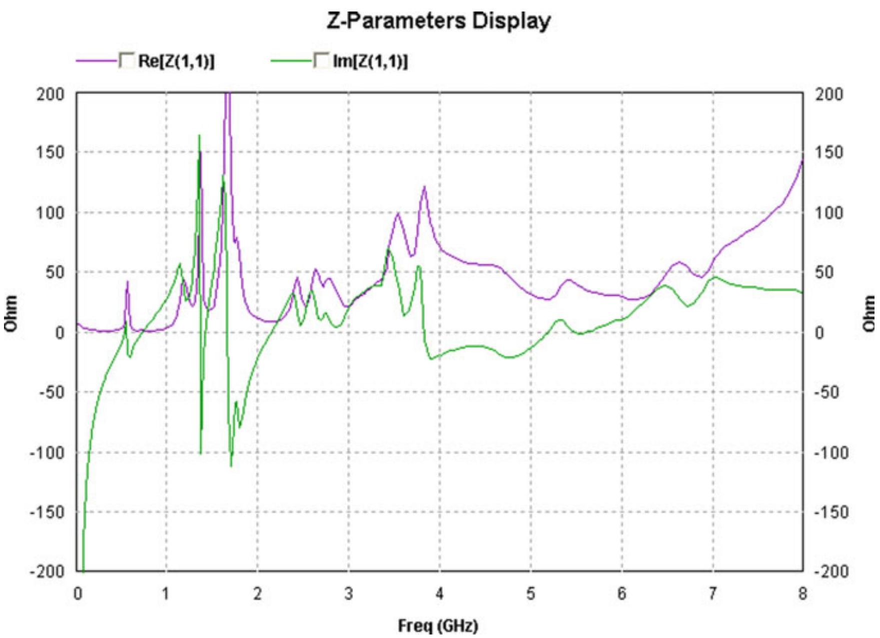


Fig. 4 *Z* parameter

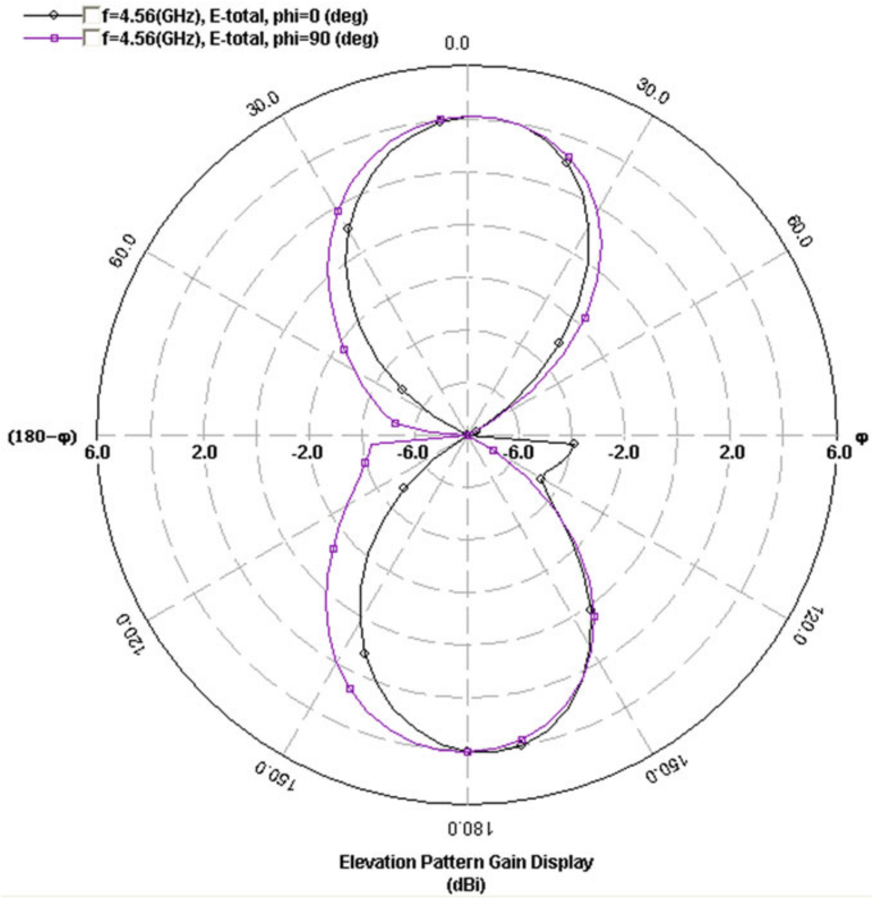


Fig. 5 Elevation pattern

frequency, and it has shown in Fig. 7. If we mounted antenna to the human body, then the resonance frequency is shifted to the greater frequency and impedance matching is decayed but finally reaches the aim of the antenna gain ISM band. The simulated result and measured result relatively agree, so it leads to the fabrication error. There will be the difference between measured and stimulated results.

5 Conclusion

A reduced planar omnidirectional MIMO antenna designed for pattern phase variety is to be discussed in this paper. This work deals with two HP channels of planar MIMO antenna. For disconnection in the middle of two channels compared with

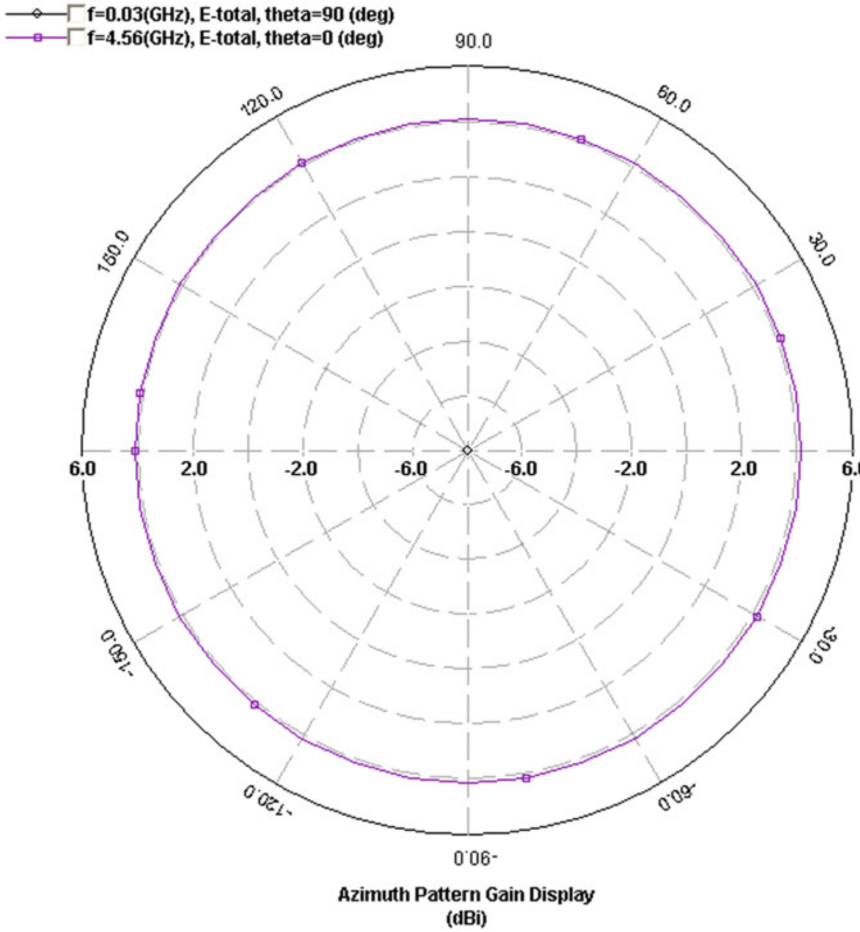


Fig. 6 Azimuth pattern

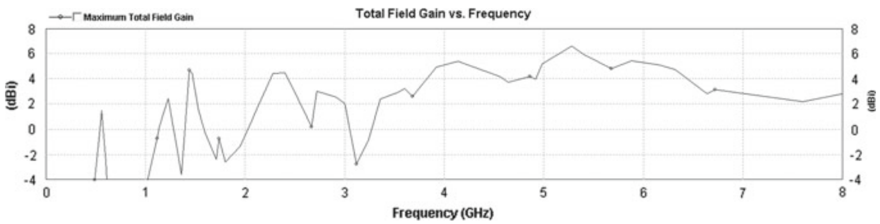


Fig. 7 Field gain versus frequency

the similar polarizations and radiation patterns, a radiation pattern phase variance approach has been used. The obtained bandwidth impedance for both channels is bigger when compared to 66.6%, and separated bandwidth among the two channels is 32.7%. The calculated and gained ECCs in the middle of both the mediums are less compared to 0.05 to other side of the separated band that displays a fine variance execution.

References

1. Sanad M, Hassan N (2019) Orthogonally polarized MIMO LTE/5G terminal antennas for handsets and IoT applications. In: IEEE radio and wireless symposium (RWS)
2. Liu C, Xia X-G, Li Y, Gao X, Zhang H, Omnidirectional quasi-orthogonal space-time block coded massive MIMO systems. IEEE Commun Lett, pp 1–1
3. Eltrass AS, Elborae NA (2019) New design of UWB-MIMO antenna with enhanced isolation and dual band rejection for WiMAX and WLAN systems. IET Microwaves Antennas Propag 13:683–691
4. Guo W, Lu A-A, Meng X, Gao X, Ma N, Broad coverage precoding design for massive MIMO with manifold optimization. IEEE Trans Commun 67:2792–2806
5. Kumar SA, Thangavelu S (2015) CPW fed implantable Z-monopole antennas for ISM band biomedical applications. Int J Micro Wire Tech 7:529–533
6. Hasturkoglu S, Almarashli M, Lindenmeier S (2019) A compact wideband terrestrial MIMO-antenna set for 4G, 5G, WLAN and V2X and evaluation of its LTE-performance in an urban region. In: 13th European conference on antennas and propagation (EuCAP)
7. Shuhrawardy M, Chowdhury MHM, Azim R (2019) A four-element compact wideband MIMO antenna for 5G applications. In: International conference of electrical, computer and communication engineering (ECCE)
8. Liu L, Liu C, Li Z, Yin X, Chen ZN (2019) Slit-slot line and its application to low cross-polarization slot antenna and mutual coupling suppressed tripolarized MIMO antenna. IEEE Trans Antennas Propag 67:4–16
9. Ashok Kumar S, Shanmuganantham T (2018 Oct) Design and performance of textile antenna for wearable applications. Trans Electr Electron Mater 19(5):352–355
10. Raleigh GG, Cioffi JM (1998) Spatio-temporal coding for wireless communication. IEEE Trans Commun 46:357–366
11. Dong Y, Choi J, Itoh T (2018) Vivaldi antenna with pattern diversity for 0.7 to 2.7 GHz cellular band applications. IEEE Antennas Wirel Propag Lett 17(2):247–250
12. Wu J, Yang S, Chen Y, Qu SW, Nie Z (2017) A low profile dualpolarized wideband omnidirectional antenna based on AMC reflector. IEEE Trans Antennas Propag 65(1):368–374

Design of Slot Antenna in Medical Wearable Applications

Tharalitha Reddy, S. Ashok Kumar, and T. Shanmuganantham

Abstract In this paper, a novel slot antenna has simple structure and small size designed for medical wearable applications. Good insulation is used as the substrates for flexible antenna, and it can be proposing for the medical applications and military purposes. When relative permittivity values are going to be deceased to low range, this is to activate performance of the antenna. The flexibility nature of the structure is the interaction between the gaps which make the antenna performing the radiation effectively. This antenna is health monitoring devices for human and animals. The proposed antenna consists of slot radiation elements fed by a coplanar waveguide (CPW) structure and a floating ground. The impedance matching bandwidth of the antenna is from 3.2 to 3.6 GHz, which can cover the 3.4 GHz band and perform good radiation characteristics.

Keywords Slot antenna · Wearable applications · Microstrip · Flexible substrate

1 Introduction

Flexible antennas are of lightweighted antenna which would stand for mechanical strain up to a certain extent. Electronic devices usually involve in communication element for transferring the data source. Communication will occur through element associated on the body to the external device. So, it is known as biological monitoring system [1]. Since the antenna is placed on human body, the back radiation should be minimized to keep the SAR value in an acceptable level which was set by FCC [2]. When the antenna is being worn, it functions as per the specifically design it

T. Reddy (✉) · S. Ashok Kumar
Jyothishmathi Institute of Technological Sciences, Karimnagar, India
e-mail: tharalithareddy@gmail.com

S. Ashok Kumar
e-mail: ashokape@gmail.com

T. Shanmuganantham
Pondicherry University, Puducherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_34

meant. Examples include smart watches, glasses, bracelets. The antenna unit communicates with the person in the hospital which is connected to phone. Microstrip patch antennas have less cost, easy to construct, and integration is simple with feed network; it has enormous application in wireless communication. Due to biocompatibility, when antenna is placed to the clothes, it will be comfortable to the patients and will monitor the patient data [2, 3]. Nowadays, the processors are performing 3.8 GHz or high-speed radiations with multi-layers. The rectangular antenna has broad bandwidth and with maximum gain. This paper deals with recent trends of antenna, and different criterial considerations that have to be knowledged include smart material composition and connectivity methods and analysis. These are very essential to design wearable antenna. The flexible antenna can be bend as for the body of the patient for the comfortable of the patient.

An antenna plays essential role in wearable devices. As the antenna is very close to the tissue of the human, so it has various curvatures, and so, it should have bending nature to decrease the mismatch and losses caused by the shape of body [4–6] Electromagnetic wave propagation to human body is reduced by the ferrite material, and height increment is needed. Electromagnetic band gap (EBG) and artificial magnetic conductor (AMC) plane would be used in SAR deduction [7, 8]. These structure is need specially design and double layer structures and will complex the manufacturing procedure. The multi-layered structure and sometimes a size of their factors are greater than antenna element which enlarges its surface of total antennas structure by using reflector element effect. Patient stress is alleviated when the antenna is worn, due to the high frequency leads to the small size of the antenna. CPW-fed antenna has the floating ground structure which leads to the ISM applications [9]. As the wireless equipment in the hospitals is using 2.4 GHz, so we are selecting the 3.4 GHz band frequency to avoid the interaction. The conductive parts which surround the slots and form the floating ground are made from conductive cloth.

2 Antenna Design

The geometrical measurements of an antenna are 19.3 mm, 17.4 mm breadth and length, respectively, and the thickness of the antenna is 1 mm as shown in Fig. 2, and dimensions are given in Table 1. Antenna has low cost and easy to fabricate, with the relative permittivity 1.36 and tan angle 0.02. There are three slots radiation element placed on top layer of antenna. Substrate bottom is covered with the part of the metallic layer, the floating ground is designed, and this not connected to radiation element. The antenna is fed by coplanar waveguide (CPW), and the feeding point is connected to the inner conductor, and the ground point is linked to the outer conductor of the coaxial cable as shown in Figs. 1 and 3.

The main part of radiation in antenna is slot-1 and slot-3. The compact of antenna is increased by the slot-3. In Fig. 3, the shorten of the total length of slot-1 and slot-3 causes the resonance frequency shifts to higher frequency. The (W2) vertical slot band gap is decided by the CPW feeding structure. In the above figure, the change of

Table 1 Antenna parameters in mm

L1	19.3	W1	17.4	L5	4.5	W5	4.2
L2	7.1	W2	0.5	L6	7.1	W6	2.2
L3	8.0	W3	1.0	L7	15	W7	2.1
L4	14.0	W4	0.8	L8	1.5	W8	5.7

Fig. 1 Overall view of antenna

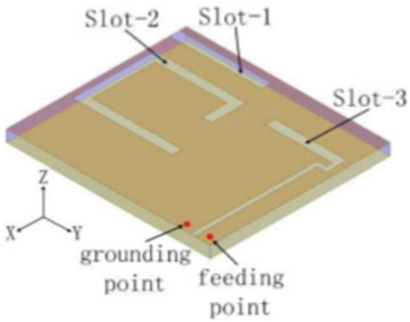


Fig. 2 Geometry of wearable antenna

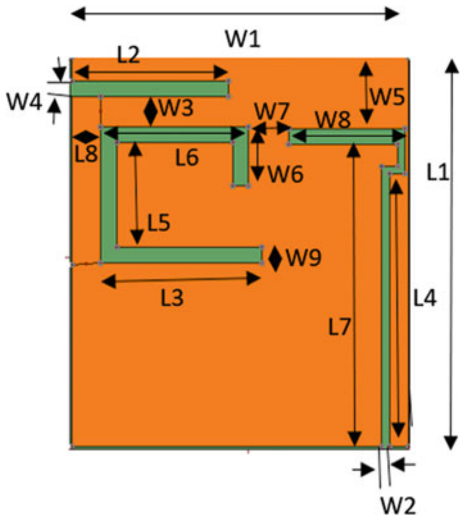
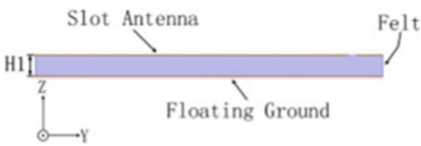


Fig. 3 Pictorial representation of side view of antenna



the slots width will cause the shift of resonance frequency, but the bandwidth is unaffected. Quality factor and the bandwidth are inversely proportional, and the radius of minimum sphere enclosed the antenna related to quality factor, and this cannot change by changing the W9. The direction of tapped currents changes according to the distance between the second and third slots that is W7. The radiating element operates like two antennas with equal radiation characteristics. So the total radiation is strengthened. By this, we can see that the forward and backward radiation is increased. As resonance frequency is not affected by changing distance in between second and third slots, impedance matching will be affected to some extent.

3 Results and Discussion

The performance of wearable antenna is influenced by the human body's high relative dielectric constant. The proposed antenna operates on human body at a distance of 7 mm; the distance between tissue and the antenna is denoted as the k0. It can also perform as per the requirement and safely.

Moreover, from the s-parameter display, it is observed that return loss is above −20 dB, and frequency is in between 3.2 and 3.6. By using few specifications, we can know the operating assumption of MIMO antenna. Width of the T-structured component is 5(it should be constant). And the height of the T-structured element can be changed. By increasing the U-shaped element height, the gain is decreased. So, it is better to keep the U-shaped element height as constant.

Obviously, the Antenna 2 covers the band of 3.4 GHz at the time of re-optimization, and the distance between four antennas is same, by which they are fabricated and calculated. The stabilized and reproduced *S*-parameters are arranged, and they are displayed in Fig. 4. By doing the optimization again, the region of the antenna can be varied comfortably. *Z* parameter of the proposed MIMO shows perfect impedance matching at resonant frequency and the current distribution as exhibited in Fig. 5.

The above-mentioned figure deals to measure the antenna efficiencies accepting the simulated efficiency results. The proposed theory of multi-input and multi-output system has good efficiency over isolation and simulation. Here, MIMO antenna is a wireless communication technology [10, 11]. It is used at source and destination. The two dimensional radiation pattern of the proposed antenna performs good radiation in various angles and elevation and azimuth patterns are exhibited in Fig. 6 and 7. The antenna gain can be plotted in Fig. 9, it shows maximum of 4 dBi at resonant frequency. The efficiency of the antenna at resonant frequency is 75% (Fig. 8).

Antenna Performance

When antenna is attached on the vicinity of patient tissue, then resonance frequence is shifted to greater frequency, and impedance matching is destroyed. Also, the extent of impact and the radiation pattern is affected by the distance of tissue. However, the safety of this antenna can still be proved. If we mounted antenna to the human

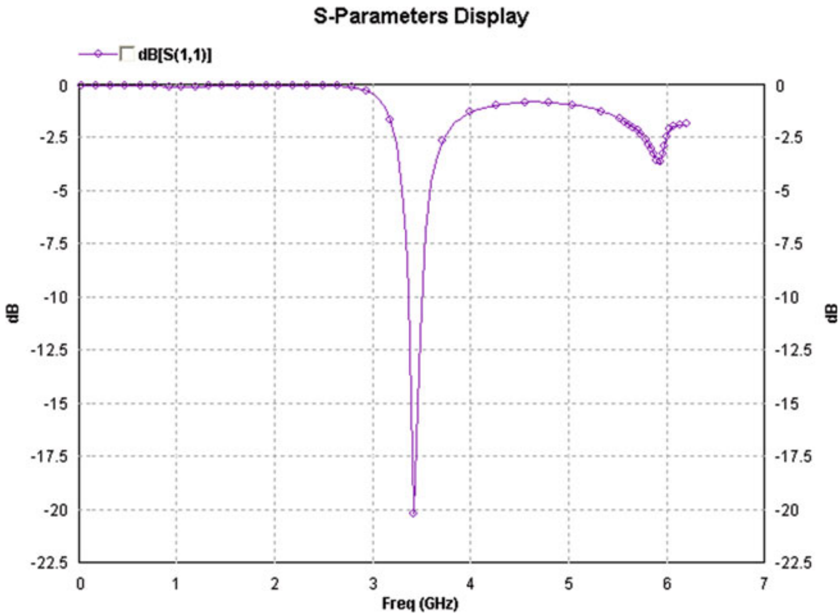


Fig. 4 *S*-parameter display

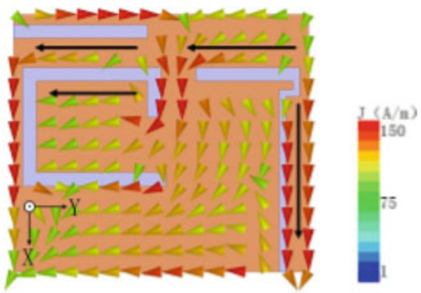


Fig. 5 Current distribution

body, then the resonance frequency is shited to the greater frequency, and impedance matching is decayed but finally reaches aim of the antenna gain ISM band. The simulated result and measured result relatively agree, so it leads to the fabrication error. There will be the difference between measured and stimulated results.

BSAR: The quantitative method for measuring the radiation effect on the human tissue (SAR). It can be calculated by the formula [12].

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

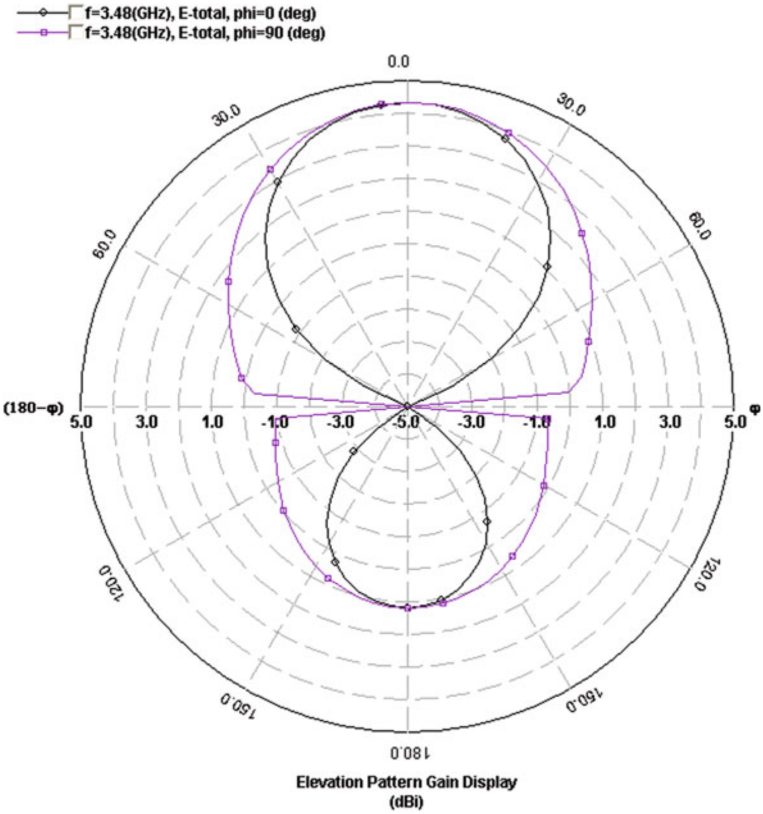


Fig. 6 Elevation pattern

where

σ conductivity
E Electric field intensity.

Bending Performance

During bending movement like forward or down, the compatible planar antenna may be operate from planar to bending configurations. In this work, the proposed antenna is mounted on the human body, and bending performances as shown in Fig. 10. The performance of the wearable antenna should not be variable. The bending does not affect the bandwidth as discussed earlier, but, impedance matching is deteriorated. The antenna can perform the radiation at 3.4 GHz frequency.

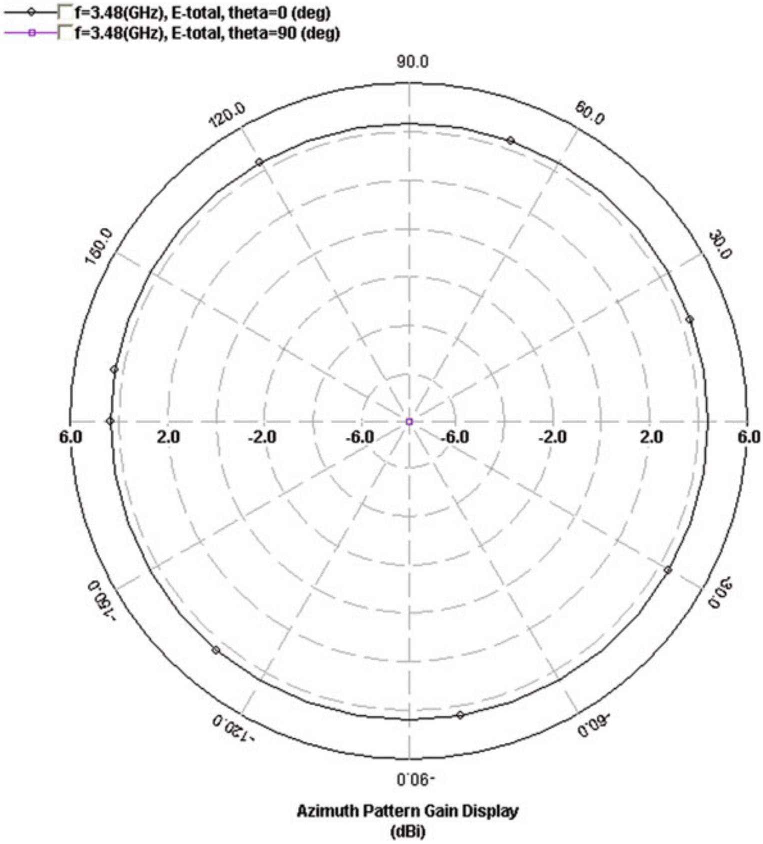


Fig. 7 Azimuth pattern

Fig. 8 Antenna placed to body of human

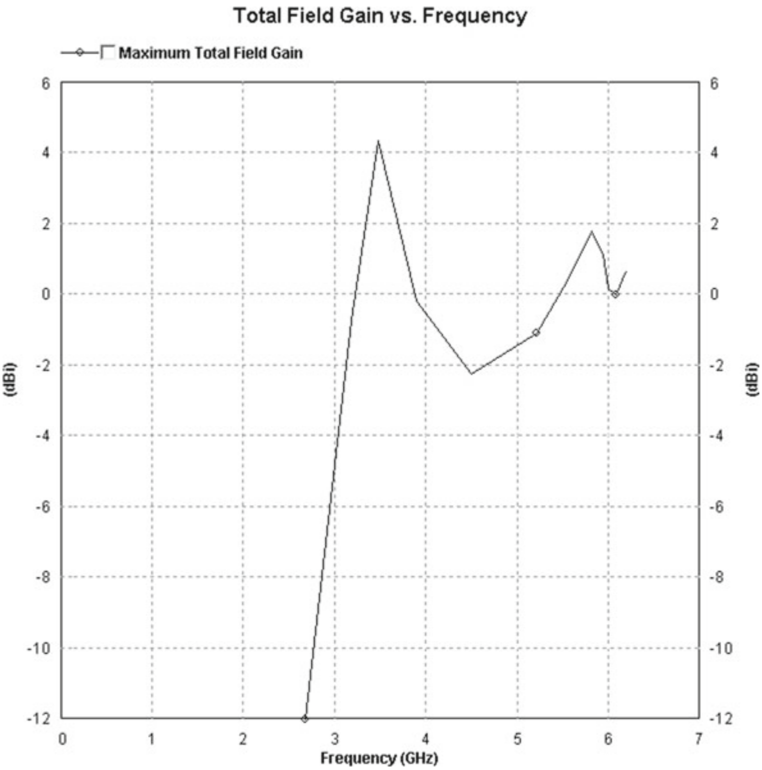
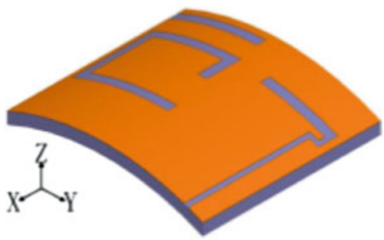


Fig. 9 Field gain versus frequency

Fig. 10 Bending nature of antenna



4 Conclusion

This paper discusses about the design of the medical wearable antenna. A compact, low-profile, compatible slot antenna for wearable applications is proposed. Good impedance matching can be obtained when the antenna is placed on outside of body. It is wireless device. The proposed antenna shows good performances when it is in bent mode and that transmission can perform effectively good during the change of posture.

References

1. Li YJ, Lu ZY, Yang LS (2019) CPW-fed slot antenna for medical wearable application. IEEE 7:4210–42112
2. Sabban A (2019) Small new wearable antenna for IOT, medical and sport application. In: European conference on antennas and propagation (EUCAD), pp 2164–2974
3. Alemaryeem A, Noghaniah S (2019) On-body low-profile texitle antenna with artificial magnetic condutors. In: 13th European conference on antenna and propagation (EUCAD), vol 67, pp 3649–3656
4. Arif A, Zubair M, Ali M, Khan MU, Mehmood MQ (2019) A compact, low-profile fractal antenna for wearable on-body WBAN applications. IEEE Antennas Wirel Propag Lett 18:981–985
5. Kumar SA, Thangavelu S (2015) CPW fed implantable Z-monopole antennas for ISM band biomedical applications. Int J Micro Wire Tech 7:529–533
6. Gao G, Hu B, Wang S, Yang C (2019) Wearable circular ring slot antenna with EBG structure for wireless body area network. IEEE Antenna Wirel Propag Lett 17:434–437
7. Ruaro A, Thaysen J, Jakobsen KB (2016) Wearable shell antenna for 2.4 GHz hearing instruments. IEEE Trans Antennas Propag 64:2127–2135
8. Ashok Kumar S, Shanmuganantham T (2018 Oct) Design and performance of textile antenna for wearable applications. Trans Electr Electron Mater 19(5):352–355
9. Mendes C, Peixeiro C (2017) A dual-mode single-bond wearable microstrip antenna for body area networks. IEEE Antennas Wirel Propag Lett 16:3055–3058
10. Yan S, Vandenbosch GAE (2016) Radiation pattern reconfigurable wearable antenna based on metamaterial structure. IEEE Antennas Wirel Propag Lett 15:1555–1718
11. Ashyap AYI, Abibin ZZ, Dahlal SH (2017) Compact and low profile textile EBG-based antenna for wearable medical applications. IEEE Antennas Wirel Propag Lett 16:2550–2553
12. Li G, Gao G, Bao J, Yi B, A watch strap antenna for the applications of wearable systems. IEEE Access 5:10332–10338

A Comparative Study on Data Cleaning Approaches in Sentiment Analysis

H. Mohamed Zakir and S. Vinila Jinny

Abstract Sentiment analysis has become an important opinion mining technique; in recent years, it becomes one of the most interesting fields in artificial intelligence. Pre-processing is considered as a significant stage in sentiment analysis, but it is not given much attention in the literature or models. The data which are collected from different sources might contain redundant and duplicates; it needs to undergo some detection process for any occurrence of redundancy in the datasets. This paper reviews, analyzes, and compares different data cleaning algorithms such as DySNI, PSNM, and brushing for identifying redundancy in the datasets. Further, it analyzed the effects of general data cleaning methods to enhance accuracy when it is applied to different classifiers. The result reveals that the DySNI algorithm gives the highest accuracy and the brushing algorithm (BAA-DD) helps to reduce the dataset size to a greater extent. Further, applying negation replacement and acronym expansion techniques helps to enhance the accuracy level.

Keywords Pre-processing · Data cleaning algorithms · Data redundancy · Data quality

1 Introduction

With the rapid development of social networking media, more and more people share their feelings, opinions, and suggestions with their friends or even strangers on their bought items in social networking applications or e-commerce systems [1]. Sentiment analysis (SA) is the process of obtaining writer's feelings as positive, negative, and neutral by reading enormous data generated in social media sites, review portals, website, etc. As an important research branch in NLP, it helps the data analysts of large organizations to measure public opinion regarding a product, services, or about

H. Mohamed Zakir (✉) · S. Vinila Jinny
Noorul Islam Centre for Higher Education, Kumaracoil, India
e-mail: hmohamedzakir@gmail.com

S. Vinila Jinny
e-mail: vinijini@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_35

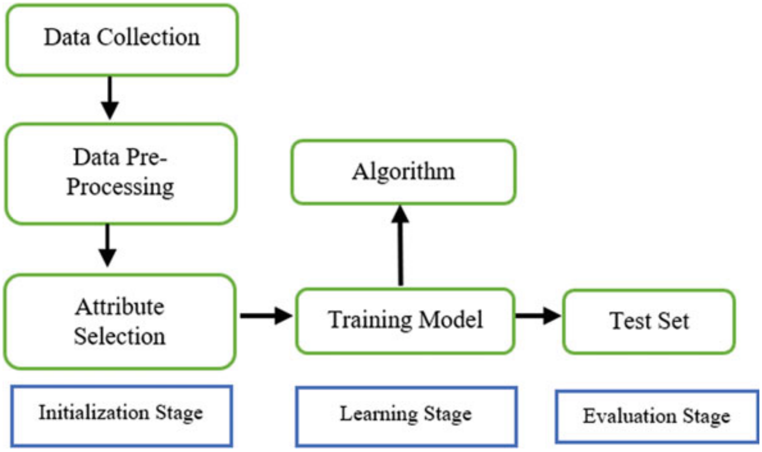


Fig. 1 Training a classifier for sentiment analysis

the current issues discussed in the social media [2]. Pre-processing contains a series of methods such as data cleaning, feature extraction, and selection to produce quality data for the next phase. The Fig. 1 shows the basic steps for training a classifier for sentiment analysis.

The data can have many irrelevant and missing parts when it is integrated from different sources. Duplicates in data can occur due to several reasons such as data maintenance, manual entries, and error in devices, and data cleaning is performed to overcome this. From an organizational perspective, data cleaning is considered very crucial in data processing and decisions might be inappropriate if the data elements used seem to be not suitable, incomplete, and inaccurate [3]. This paper analyzed different data cleaning algorithms such as PSNM, DySNI, and brushing to identify the redundancy in the datasets. Further, the impact on accuracy is discussed, when the data cleaning methods such as removing URLs, negation replacement, reverting repeated letters, stop words and numbers removal, and acronym expansion are applied on sentimental classifiers. Each approach has been evaluated using different datasets and the performance is noted. The remaining section of the paper is structured as follows. The related work is discussed in Sect. 2. Section 3 presents the experimental metrics used to assess the three algorithms. The experimental results are described in Sect. 4 and discussions in Sect. 5. The conclusions are made in Sect. 6.

2 Related Work

2.1 Dynamic Sorted Neighborhood Indexing for Real-Time Entity Resolution (DySNI)

Banda Ramadan et al. [4] proposed a tree-based technique to query records from the dynamic database with reduced query times based on the sorted neighborhood indexing method. In this work, the authors emphasize that data cleaning cannot be applied directly on all datasets, doing so may lead to loss of information on some real-world applications. An accuracy of 86.23% is observed on one large real-world database and two synthetic datasets. The average query comparison time is reduced between 20 and 70%. This technique can be applied to investigate both static and adaptive window approaches in entity resolution (ER). More memory is required which increases the cost to improve the query time.

2.2 Brushing—An Algorithm for Data Deduplication (BAA-DD)

Prasun Dutta et al. [5] proposed a brushing algorithm to resolve deduplication in individual files. This work overcomes the space-related and the time complexity problem. In this work, hashing and bloom filters are the main two methods used which are combined as a single process. The algorithm has been evaluated on their database and the file size is reduced to 73.21%. This algorithm can be used to remove the duplicates from the text files, and it does not support other file types like audio, video, etc.

2.3 Progressive Sorted Neighborhood Method (PSNM)

Srinivasa Reddy et al. [6] proposed a method to identify duplicates and null values in real-time datasets using progressive sorted neighborhood method (PSNM). It is a sorted neighborhood method, which works based on predefined sorted key value. The distance of two records is calculated and rankings are assigned based on sorter order. The resulted (unique) data can be accessed in different file conversion formats based on the user types. An accuracy of 70.56% is achieved in detecting duplicates on World Ocean Database (NODC). The memory restrictions cause the datasets to split, in order, to fit into the memory.

2.4 Comparison Study on Pre-processing Methods (CSPM)

Jianqiang et al. [7] explored the effects of different pre-processing methods such as removing URLs, negation replacement, reverting repeated letters, stop words and numbers removal, and acronym expansion when it is applied on different classifiers. The effects are assessed by monitoring the variations in classifiers when measured with respect to accuracy and average $F1$-measure. The methods have experimented on five datasets; the accuracy reaches 6.85% after acronym expansion in the n-grams model and 8.23% after negation replacement in SVM classifier. The URL in the tweets is removed by assuming that it does not contain useful information, but this conclusion is not consistent (Bao et al. [8]).

2.5 Deep Recurrent Convolution Neural Network (DRCNN)

Zhang et al. [9] proposed this paper to address the challenges in the existing approaches of subjectivity classification. He introduces a deep recurrent convolution neural network (DRCNN) model with multi-features combination for subjectivity classification of microblog text. The pre-processing steps include removal of stop words and URL links, replacing negations and emoticons, acronyms expansion, and tokenization using the tweet NLP. This model is experimented on six Twitter datasets. The highest accuracy of 88.92% is achieved on the Stanford Twitter Sentiment Test dataset, and the minimum accuracy is 82.12% on the SemEval2014 dataset.

2.6 Glove-Deep Convolution Neural Network (G-DCNN)

Zhao Jianqiang et al. [10] proposed this paper to obtain the sentiment features, by introducing word embeddings method obtained from unsupervised learning. The pre-processing steps performed in this work include replacing negations and emoticons, acronym expansion, removal of non-English characters, URLs, numbers, and stop words. The highest accuracy of 87.62% is achieved on standard Twitter Sentiment Test dataset. The advantage of this model is 10-fold cross-validation which is applied on all five datasets. The URL links are removed from the tweet without parsing the destination page.

3 Experiment Evaluation

In this area, we present the experimental assessment of three algorithms which are DySNI, PSNM, and brushing. Also, the evaluation metrics, dataset utilized, and setting up configurations for the experiments are discussed. The algorithm has been evaluated using MATLAB 10.0 and the whole experiment is conducted in Windows 10 Home 64-bit Operating System with Intel® Core™ i7-8550U CPU @ 1.80 GHz 1.99 GHz processor and 8.00 GB of RAM.

3.1 Evaluation Metrics

The main parameter for analyzing an algorithm to calculate the correctness of the system is mainly decided by the number of duplicates detected. The correct identification of duplicates is commonly regarded as true positive (TP). In another case, if a non-duplicate record is detected as a duplicate, then it is considered as false positive (FP). Records that are duplicates but not detected as duplicates are considered as false negatives (FN). The duplication identification measurements have been utilized to assess the efficiency of various duplication identification techniques in terms of recall, precision, and *F*-score.

- ***Recall*** measures the proportion of accurately detected duplicates compared to all true duplicates.
- ***Precision*** is also called as a positive indicator value; again, it measures the proportion of correctly detected duplicates compared to all other declared duplicates.
- ***F-score*** measure captures this trade-off by merging recall and precision through a harmonic mean.

In the information retrieval (IR) field, precision in combination with recall is widely used as a measure for precision–recall visualization graph [11, 12]. Table 1 represents the mathematical formula to calculate recall, precision, and *F*-score [13].

The second component in the evaluation of any duplication identification approach is the time required to get an outcome when the data are being cleaned and moved to the data warehouse. The amount of time that the algorithm takes to execute a function is called as processing time. So, in this work, the execution time of the algorithm is viewed as significant to scale the algorithm's efficiency.

Table 1 Quality measurement formula

Quality measure	Formula
Precision	TP/(TP + FP)
Recall	TP/(TP + FN)
F-score	2 * (Precision * Recall)/(Precision + Recall)

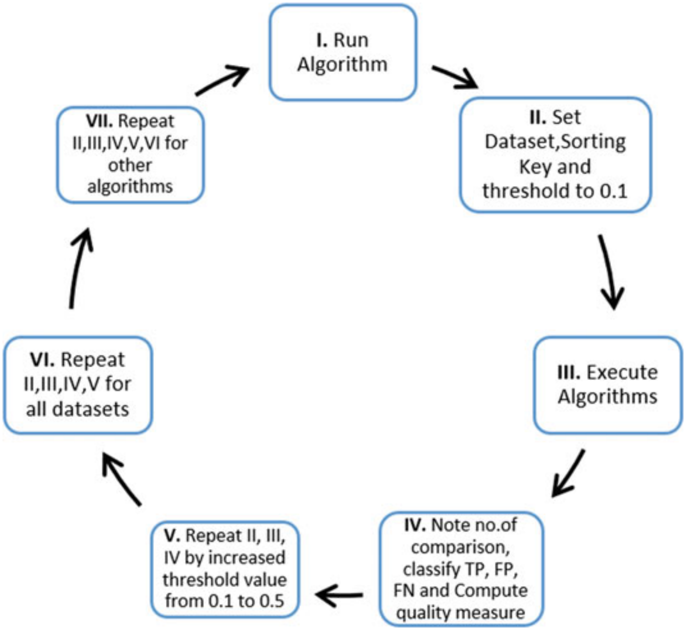


Fig. 2 Algorithm execution steps

3.2 Datasets Description

Restaurant datasets are used to evaluate the three algorithms: DySNI, PSNM, and brushing. It is used by many researchers to experiment with their algorithms [14, 15]. It consists of 864 records including 112 duplicates with different fields like "name, address, telephone, city and type".

3.3 Steps for Algorithm Execution

The experiments are carried out using the steps shown in Fig. 2.

4 Experimental Results

Table 2 shows the results of different metrics for brushing algorithm with restaurant dataset.

Table 3 shows the results of different metrics for DySNI algorithm with restaurant dataset.

Table 2 Results of brushing algorithm

Threshold	No. of comparisons	TP	FP	FN	Precision	Recall	*F*-score
0.1	431	111	339	0	0.24	1	0.39
0.2	431	111	339	0	0.24	1	0.39
0.3	431	111	339	0	0.24	1	0.39
0.4	431	111	329	0	0.25	1	0.40
0.5	431	111	320	0	0.25	1	0.40
0.6	461	111	290	0	0.27	1	0.43
0.7	697	111	53	0	0.67	1	0.80
0.8	925	103	2	8	0.98	0.92	**0.95**

Table 3 Results of DySNI algorithm

Threshold	No. of comparisons	TP	FP	FN	Precision	Recall	*F*-score
0.1	8723	111	750	0	0.12	1	0.22
0.2	8723	111	750	0	0.12	1	0.22
0.3	8723	111	750	0	0.12	1	0.22
0.4	8723	111	750	0	0.12	1	0.22
0.5	8723	111	750	0	0.12	1	0.22
0.6	461	111	290	0	0.27	1	0.43
0.7	697	111	53	0	0.67	1	0.80
0.8	925	103	2	8	0.98	0.92	**0.95**

Table 4 shows the results of different metrics for PSNM algorithm with restaurant dataset.

Table 4 Results of PSNM algorithm

Threshold	No. of comparisons	TP	FP	FN	Precision	Recall	*F*-score
0.1	424	111	310	0	0.26	1	0.41
0.2	453	111	303	0	0.26	1	0.42
0.3	783	111	230	0	0.32	1	0.49
0.4	1510	94	162	17	0.36	0.84	0.51
0.5	2289	93	56	18	0.62	0.83	0.71
0.6	2598	87	12	24	0.87	0.78	0.82
0.7	2831	73	2	40	0.97	0.64	0.77
0.8	2959	57	0	54	1	0.51	0.67

5 Discussion

Each data cleaning approach has experimented on different datasets, and the results are mentioned in Table 5.

From the results shown in Fig. 3, the accuracy value of the DySNI method is competitive to the other two methods DRCNN and Glove-DCNN. In DySNI, the data cleaning process's time is reduced up to 70% when compared to the other two methods. Moreover, it is capable of processing the dynamic datasets, where the other methods do not support new tweets included in the database. Hence, it is evident that the DySNI algorithm performs better out of others on all datasets in terms of supporting dynamic datasets, execution time, and accuracy. The brushing algorithm reduced the size of the dataset to 73.21%. Also, from CSPM, DRCNN, and G-DCNN, it is observed that negation replacement and acronym expansion can be given more importance than others to enhance accuracy further. The tweets in the URL need to be processed rather than omitting to grab the important information. Tables 6 and 7 show the overall summary, and the merits and demerits of each algorithm respectively.

Table 5 Accuracy analysis

Method name	Accuracy (%)	Database name
DySNI	86.23	NC, OZ-x, Febrl
BAA-DD	73.21	Own database
PSNM	70.56	Meteorology dataset
CSPM	8.23	STS-test, SemEval2014, STS-Gold, SS-Twitter, SE-Twitter
DRCNN	88.92	SE-Twitter, SS-Twitter, SemEval2014, STS-Test, SemEval2015-Test, SemEval2016
G-DCNN	87.62	STSTd, SE2014, STSGD, SED, SSTD

Fig. 3 Accuracy analysis chart

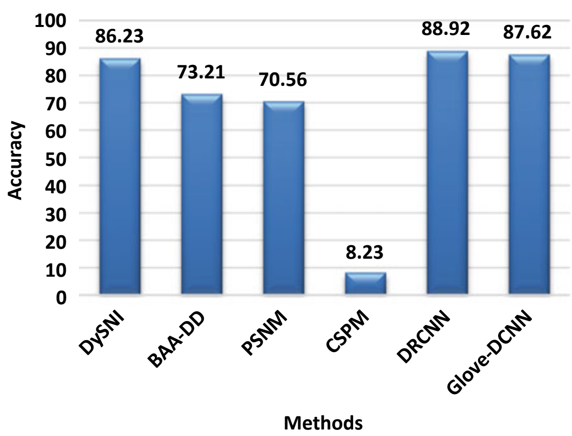


Table 6 Overview of data cleaning approaches

Methods	Publication	Year	Data cleaning steps
DySNI	ACM	2015	Entity resolution
BAA-DD	Springer India	2016	Removing space and similar entities
PSNM	Springer Nature	2017	Identifying duplicates and null values
CSPM	IEEE	2017	Negation replacement, acronym expansion, removal of URLs, stop words, and numbers
DRCNN	IEEE	2018	Replacing negations and emoticons, acronym expansion, removing stop words, non-English characters, URLs, and numbers
G-DCNN	IEEE	2019	Replacing negations and emoticons, acronym expansion, removing stop words, non-English characters, URLs, and numbers

Table 7 Merits and Demerits of each data cleaning approaches

Methods	Merits	Demerits
DySNI	It can be used to investigate both static and adaptive window approaches	More memory is required which increases cost to improve the query time
BAA-DD	It is a lightweight algorithm	It won't support files of other formats like audio, video, etc.
PSNM	The resulted unique data can be converted to different file formats with user type restrictions	Minimal datasets are used for conducting the experiments
CSPM	Acronyms and slangs were expanded to its original form	The URL in tweets are removed during pre-processing, which may reduce accuracy
DRCNN	Dimensionality and data sparsity problem is reduced in this method	Required more execution time
G-DCNN	10-fold cross-validation is applied on all five datasets	The URL links are removed from the tweet during pre-processing without parsing the destination page

6 Conclusion

This work aims to find the best data cleaning approaches by experimentally comparing three relevant algorithms such as PSNM, DySNI, and brushing, and also the impact on accuracy when the general data cleaning methods such as removing URLs, negation replacement, reverting repeated letters, stop words and numbers removal, and acronym expansion are applied on sentimental classifiers. Based on the experimental results, the DySNI algorithm performs better by reducing the execution time to 70% and by supporting the dynamic datasets. The recommended data cleaning methods which can be further applied are negation replacement and acronym expansion to improve the accuracy level further. The brushing algorithm also performs better by reducing the file size by 73.21%. Thus, the DySNI and brushing algorithms can be combined to achieve a better accuracy in reduced execution time, along with applying negation replacement and acronym expansion. Hence, we strongly believe that this hybrid approach will be helpful for researchers to produce quality data in the pre-processing phase, which is very crucial for achieving a better classification accuracy. The end application of this research is that it can be used as a pre-processing step (i.e., data cleaning) in e-commerce applications at the time of decision making. Our future work is to apply this hybrid model in different domain datasets such as product reviews from Amazon and geopolitical datasets.

References

1. Zhang S, Zhong H (2019) Mining users trust from E-commerce reviews based on sentiment similarity analysis. IEEE Access 7:13523–13535. https://doi.org/10.1109/ACCESS.2019.2893601
2. Li L, Wu Y, Zhang Y, Zhao T (2019) Time + User dual attention based sentiment prediction for multiple social network texts with time series. IEEE Access 7:17644–17653. https://doi.org/10.1109/ACCESS.2019.2895897
3. Alenazi SR, Ahmad K (2017) An efficient algorithm for data cleansing. J Theor Appl Inf Technol 95:6183–6191
4. Ramadan B, Christen P, Liang H (2014) Dynamic sorted neighborhood indexing for real-time entity resolution. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence, Lecture Notes Bioinformatics), vol 8506 LNCS, 1–12. https://doi.org/10.1007/978-3-319-08608-8_1
5. Datta P, Pattnaik P, Sahu RK (2016) Brushing-An algorithm for data deduplication. Information systems design and intelligent applications: Proceedings of Third International Conference INDIA, Vol 1 Adv. Intell. Syst. Comput. 433. https://doi.org/10.1007/978-81-322-2755-7_23
6. Reddy SL, Vighneshwar SP, Ravikiran B (2017) PSNM: An Algorithm for Detecting Duplicates in Oceanographic Data. Computer Communication, Networking and Internet Security Vol 5, 149–156. https://doi.org/10.1007/978-981-10-3226-4_29
7. Jianqiang Z, Xiaolin G (2017) Comparison research on text pre-processing methods on twitter sentiment analysis. IEEE Access 5:2870–2879. https://doi.org/10.1109/ACCESS.2017.2672677
8. Bao Y, Quan C, Wang L, Ren F (2014) The role of pre-processing in twitter sentiment analysis. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence, Lecture Notes Bioinformatics), vol 8589 LNAI, pp 615–624. https://doi.org/10.1007/978-3-319-09339-0_62
9. Zhang X, Huang S, Zhao J, Du X, He F (2019) Exploring deep recurrent convolution neural networks for subjectivity classification. IEEE Access 7:347–357. https://doi.org/10.1109/ACCESS.2018.2885362
10. Jianqiang Z, Xiaolin G, Xuejun Z (2018) Deep convolution neural networks for twitter sentiment analysis. IEEE Access 6:23253–23260. https://doi.org/10.1109/ACCESS.2017.2776930
11. Yan S, Lee D, Kan MY, Giles LC (2007) Adaptive sorted neighborhood methods for efficient record linkage. In: Proceedings of the ACM international conference digital libraries, 185–194. https://doi.org/10.1145/1255175.1255213
12. Thesis AM (2012) A comparative study of duplicate record detection techniques
13. Skandar A, Rehman M, Anjum M (2015) An efficient duplication record detection algorithm for data cleansing. Int J Comput Appl 127:28–37. https://doi.org/10.5120/ijca2015906401

14. Tejada S (2002) Copyright 2002 Sheila Tejada. Memory
15. Rehman M, Esichaikul V (2009) Duplicate record detection for database cleansing. In: 2009 2nd international conference machine vision, ICMV 2009, pp 333–338. https://doi.org/10.1109/ICMV.2009.43

Sybil Attack in VANET Operating in an Urban Environment: An Overview

Nitha C. Velayudhan and A. Anitha

Abstract Networking technology in the present scenario has the latest invention called ad hoc networks. VANET is a type of ad hoc network where vehicles communicate with each other in a network. From the last few decades, VANET is becoming popular where relocation, mobility, and converge of location by the vehicles were impossible to achieve by using a wired network. VANET provides the growth of intelligent transport systems by providing safety and comfort consideration for both drivers and passengers. VANET is an open-access network that may prone to various types of attacks. This paper is an overview of the study of various attacks in VANET focusing on Sybil attack in VANET which operates under the urban environment. A comparison of different methods of detection technique for Sybil attack is also included.

Keywords Vehicular ad hoc network · Onboard unit (OBU) · Application unit (AU) · Roadside unit (RSU) · Sybil attack

1 Introduction

The road crashes are the serious issues of injuries, property damages, and cause of death. The overall report by the World Health Organization about road safety in 2018 stated that the number of annual road traffic death rate has reached 1.35 million. On this account, our study focuses on the avoidance of the dangerous situations of accidents caused by human error by providing a fast and accurate warning to the driver half a second before the collision. In this context from the past few decades, VANET is becoming an important part of the research.

N. C. Velayudhan (✉) · A. Anitha
Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Kumarakovil, Tamil Nadu, India
e-mail: nithacvelayudhan@gmail.com

A. Anitha
e-mail: anidathi@yahoo.co.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_36

VANET is a subgroup of ad hoc where the nodes are vehicles that communicate by the means of North American Standard dedicated short-range communication (DSRC) that utilizes the IEEE 802.11p standard in the vehicular domain. The vehicle in the VANET communicates without the existence of fixed infrastructure.

The safety-related application is the major objective of VANET. When a road accident is detected, a vehicle can continuously broadcast the information of critical situations to the approaching nearby vehicles, and the other reason for researches in VANET is to sell more vehicles by providing innovative software and even vehicles without drivers. The currently used methods rely on turning vehicles in a VANET into a wireless router or a node, which in turn creates a wide range of networked vehicles and when a vehicle leaves the signal range of network, then that particular vehicle is dropped from the network enabling another vehicle to join the network [1].

This technology can be used for safety-related applications where police and emergency vehicles can communicate about the critical situation. As the VANET provides a wide range of driving and road safety also makes provision for increasing traffic efficiency, it is becoming popular in automotive companies like General Motors, Toyota, BMW, and Ford. These companies have started promoting VANET by installing the components necessary required for the communication between the vehicles. Most of the countries have made VANET mandatory part of traffic safety, and there was a drastic drop in the ratio of road accidents.

VANET as a part of the ad hoc network is equipped with wireless processing capability devices providing specific features of high processing power, large storage capacity, and sufficient amount of energy backup by working over the battery of vehicles. VANET also detects the time and position of the vehicle. These features enormously distinguish VANET from other ad hoc networks.

VANET operating in urban environment has enumerable attractive services, namely security of the vehicle, toll payment which can be made automatic, traffic management, Internet access, navigation, and location-based services for finding the nearest propellant stations, travel cottage or restaurants, etc., but the traffic behavior and network connectivity of vehicles moving in urban environment are very challenging task; besides, there are also countless attacks that affect the urban VANET as the vehicles are anonymous and short time connectivity to the network. Moreover, the location of the vehicle must be kept confidential. This paper reviews the works done on a variant and most deleterious type of attack called Sybil attack effecting VANET while operating in the social and urban environment.

This paper is organized as follows: Sect. 2 discusses the overview of urban VANET. Section 3 has analyzed the security requirement and applications of urban VANET. The classification of attackers and attacks in urban VANET is included in Sect. 4. The survey on Sybil attack in urban VANET is included in Sect. 5. Section 6 is the conclusion and describes the future works.

2 Overview of Urban VANET

Various researchers took a deliberate effort in studying many types of threats in VANET, but the first person to coin the term about the Sybil attack was Douceur [2], who described the attacker as a malicious node entering into the network by taking the identity of multiple nodes and behaving like a normal node in the network. The Sybil attack causes huddles in the network by changing the traffic-related information, resource allocation; in turn, it can lead to traffic jams and accidents.

The architecture of urban VANET is similar to an ordinary VANET which includes:

- **Ad hoc environment**: Onboard unit (OBU) and application unit (AU) constitute the ad hoc environment. All vehicles participating in urban VANET should have an OBU for communication. The application unit (AU) works behind OBU executing the necessary programs required by the OBU to communicate.
- **Infrastructure environment**: consists of roadside unit (RSU) which is installed at roadsides or parking slots and acts as an access network for providing on-road Internet connectivity.

The communication in urban VANET is:

- **V2V communication**: The communication between vehicles (OBUs) and vehicles (OBUs) is called **V2V** communication. It enables intelligent vehicular communication capability.
- **V2I communication**: The communication between infrastructures (RSU) installed at the roadside or parking slot with vehicles (OBUs) is called **V2I** communication.
- **Hybrid communication**: is the combination of **V2V** and **V2I**.

The set of three types of communications use dedicated short-range communication (DRDC) [3]; the Federal Communication Commission (FCC) has allocated a band of 75 MHz in 5.9 GHz [4]. Figure 1 shows the architecture of VANET.

There are two most prominent technologies endowed for efficient communication in urban VANET:

- **IEEE 802.16 (Wireless MAN/WiMAX)**: designed to impart wireless multimedia application over 30 miles
- **IEEE 802.11p (WAVE)**: In a licensed ITS band of 5.9 GHz, the technology enables V2V and V2I communication in VANET.

There are denumerable applications of urban VANET which are subdivided into two categories:

- **Security-based application**: collision avoidance, traffic analysis, interactive driving
- **User-based application**: entertainment, Internet connectivity on road, providing information about restaurant, and fuel pump.

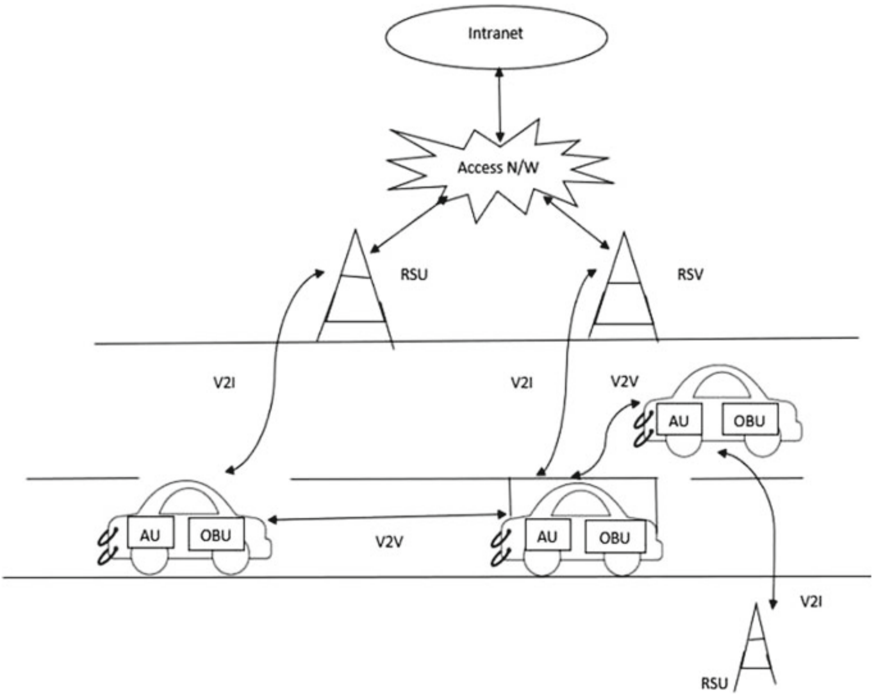


Fig. 1 VANET architecture

Despite various applications, there are various security challenges and requirements faced by an urban VANET during communication:

- Consistency of data
- The rapid change in mobility of the vehicles
- Error tolerance
- Latency control
- Key management
- Location privacy preservation of vehicles.

3 Security Requirements and Applications of Urban VANET

It is evident that any malign behavior of the vehicle in the VANET, such as modification of messages, could be fatal to other vehicles. Therefore, the most important feature of VANET is security. The requirement of security in urban VANET includes:

- **Message authentication and integrity**: Message must be protected from any modification, and the recipient of a message must verify the authentication of the sender.
- **Message non-repudiation**: The sender of a message should not deny the authenticity of sending a message.
- **Entity authentication**: The receiver should validate the sender of the message and should add the evidence of the present condition or the aliveness of the sender.
- **Access control authorization**: establishes what each node is permitted to do in VANET.
- **Message confidentiality**: The content of a message should be kept confidential from unauthorized nodes.
- **Privacy and anonymity**: Conditional privacy is provided in urban VANET. Only the authorities will be given the access to identify the sender of the message in the case of a dispute such as a crime or accident case investigation.
- **Availability**: The applications running on VANET and VANET network should remain functional even after the removal of the faulty node from the network.
- **Liability identification**: Users of vehicles are legally answerable for their intentional wrongdoing or accidental actions that cause harm or disturb the proper functioning of other nodes or transportation systems. In this case, the authorities will be given access to identify the sender of the message.

Depending on the purpose the VANET are used, we can summarize the existing applications of VANET as:

- **Safety applications**: VANET has become a noticeable feature of the transport system; the major intention of VANET is the life saving from road accidents. Safety application of VANET includes intersection collision avoidance by providing a warning about the vehicles and pedestrian crossing violation of traffic rules, providing public safety event-driven messages of nearby emergency vehicles, and it also alerts the drives about sign extension kept on the road.
- **Comfort/entertainment applications**: This type of application is a non-safety application providing a pleasant journey and traffic efficiency. The drivers and passengers are given information about a traffic jam ahead, nearest ATM, restaurants, petrol pumps, and the prices of fuel. Furthermore, the vehicles are connected to an infrastructure (RSU) the passenger can play online games and can avail Internet access throughout the journey.

4 Types of Attackers and Classification of Attackers in Urban VANET

It is important to note that the VANET is a decentralized network and has unrestricted access to all kinds of vehicles; owing to that, the VANET may prone to various types of attacks. Attackers attacking may alter or deny data while transmitting; attackers may obtain unauthorized privileges and insert false data into the network, and they may also change the information in the network, perform traffic network analysis and by using malicious software attackers can also disrupt the network. Therefore, the nodes (vehicles) in the VANET have to guarantee the security measures of privacy, integrity, accessibility, and verification of the node during the transit of data.

The categorization of attackers in VANET is [4]:

Based on membership of the attacker

- **Internal attackers**: are authorized authorities that perform the malicious activity for their personal benefit, and this is a more serious type of attacker than an external attacker; the internal attacker may disturb the network.
- **External attackers**: are intruders who enter into the network by impersonation or by some other kinds of attack. The external attacker can create congestion in the network, which can lead to the unavailability of the network and services; they can also send false routing information.

Based on the activity of the attacker

- **Passive attackers**: are the attackers who silently observe the network for the traffic analysis and release the message content to other nodes without altering the data.
- **Active attacker**: are the attackers that modify the packet or generate new packet carrying malicious information to damage the network.

Based on the intention of the attacker

- **Rational attackers**: are the attackers that attempt to take their personal advantage from the attacks.
- **Malicious attackers**: are the attackers not gaining any personal advantage; their main motive is to create hurdles in the proper functioning of the network.

In VANET, as there is no fixed infrastructure and decentralized administration, there are various types of attacks that are prominent in urban VANET. The classifications of attack are:

- **Monitoring attacks**: The attacker in a monitoring attack is a passive attacker with malicious intention, the attacker silently monitors the network and creates huddles in the network. Impersonation and session attack falls into the category of monitoring attack.
- **Social Attacks**: The attacker sends unethical messages to the driver creating an emotional deception in the driver, which creates driving confusion.
- **Timing attacks**: attack which delays the arrival of packets in the network by making the alteration in the time slot of messages which is sent from a sender.

- **Application attack**: Attackers huddle the application running on the VANET. Eavesdropping and bogus information are examples of application attacks.
- **Network attack**: The attacker attacks the whole network; it is the most serious attack. DOS attack and Sybil attack are the most prominent attacks under this category.

5 Survey of Sybil Attack in Urban VANET

Most of the safety application of VANET requires cooperation between the vehicles. The communicating vehicles should be legitimate enough to transmit genuine information to the other vehicles participating in VANET. However, Sybil attack creates artificial and forbidden ids to claim access to the network. The nodes in such type of attack are called Sybil nodes. These nodes smash the network and create confusion in the network which in turn leads to numerous accidents.

The Sybil attack was first introduced and described by Douceur [2]; in his paper, the Sybil attack was detected by suggesting a resource testing approach; this technique was based on the assumption that each vehicle will have a limited number of computation resources, but this method was not suitable for VANET, as the attacker in VANET environment can have strong computing resources than the legitimate node. Newsome [5] improved the approach proposed by Douceur by testing radio resources nodes, and it was also based on assumption that vehicular node can only have one wireless communication module and each can occupy only one channel resource at a time. But, this technique was hazards to implement.

Park et al. [6] provided a timestamp series approach to detect the Sybil attack. In this approach, whenever RSU detects a vehicle in its vicinity, RSU assigns a timestamp to the vehicle. The Sybil attack was detected on the basis of similarly timestamp series. But, this approach has some limitation that when the distance between two RSU is more than 100 km, the detection becomes time consuming, and this detection works if Sybil attack stays under only one RSU.

Hao et al. [7] used the GPS position of the vehicle to detect the Sybil attack. Each vehicle periodically broadcasted its GPS position within the network to detect Sybil attack, but GPS spoofing was necessary for this approach.

Lee et al. [8] discussed the DTSA technique for Sybil attack. This was based on using the session key based on the certificate to verify the vehicle's identities. In their approach, each vehicle verifies the others in its networks to make sure it is not Sybil node, but this technique of verification requires a high-performance capability and places a high load on the network.

Tang et al. [9] use central authority to prevent Sybil attack; this method proposes a ticket for authenticating neighborhood vehicles which then is verified and validated by a central authority, but it imposes an overhead on the central authority.

Studer et al. [10] proposed the use of a regional authority. The authority generates a temporary anonymous certified key (TACK) and maintains a group signature for security and privacy properties of a vehicle participating in the VANET. The authority can detect the misbehaving OBU of the vehicle in the VANET and nullify the participation of malicious node in the VANET, but the TACK has a very short lifetime and updating should be done when the lifetime of a vehicle in a region gets ended up or when the vehicle joins a new geographical region; this leads to delay in likability.

Haas and Hu [11] in his paper compare security routing protocol IECDSA with TESLA in the new simulator; the security routing protocol IECDSA detects and removes Sybil attack in the VANET. But, this protocol will not detect and isolate location-based security attacks in the VANET; it is limited only to detect both identity and location used for the simulation and hybrid type of protocol-based security attacks in the network.

Wireless LAN Working Group [12] proposed that the vehicles in VANET exchange information about the speed, location, direction, and emergency information through basic safety messages (BSMs), but this data can be compromised by Sybil attack; it can lead to the unavailability GPS, and it can even lead to bugs in software; therefore, it is necessary to provide localization mechanism.

Zhou et al. [20] proposed hashing pseudonyms to detect the Sybil attack. This scheme will not disclose the identity vehicle but the scheme inert overhead on a central authority, hence not suitable for highly dynamic vehicular domain.

Xiao et al. [13] referred to a technique POEST in his paper. This technique relies on RSSI-based localization mechanism using stationary RSU to localize vehicle, but it suffers due to high inaccuracy and mobility.

Garip et al. [14]proposed a mechanism called INTERLOCK to find the location of vehicle when it is not available because of noisy or unavailability of GPS; it effectively localizes a node to collect the RSSI value, and it uses vehicle on the map to localize another vehicle. Therefore, it is also extremely robust to the change in rates of mobility in the network. This mechanism calculates an estimated point and evaluates the success rate, but if this point is closer to another vehicle, then the vehicle being localized fails.

Grover et al. [15] presented the Sybil attack detection by finding the similarity in neighboring vehicles, but vehicle densities affect the scheme and the detection performance could be a challenge.

Many researchers have been carried out to detect and prevent Sybil attack in an ordinary VANET, but only a few types of research have been carried out in urban VANET where it needed to keep the identity of vehicles to be kept confidential.

Chen et al. [16] presented a novel and efficient scheme to detect Sybil attack with limited infrastructure in urban VANET; here, each node accomplished Sybil attack detection independently by comparing the difference in neighboring node's digital signature, but this scheme can be accomplished only under normal condition.

Chang et al. [17] proposed footprint method for the VANET operating in an urban environment where the identity of the vehicles is not considered and geographical information of the vehicles was confidential. Sybil detection was conducted online. Besides the advantage, it requires an infrastructure of RSUs and a trusted authority (TA).

Hamed et al. [18] proposed a scheme for the routine communications among the nodes and RSUs to detect the attackers, but there may be the situation when RSU gets compromised to malicious nodes.

6 Conclusion

In this paper, at the beginning, we presented the overview of an urban VANET; the security requirements and applications of urban VANET are also manifested. We have also introduced various types of attackers and classification of various attacks prominent in urban VANET, followed by a detailed review of some major technique available today for the detection of Sybil attack in an ordinary VANET succeed by few types of research which have been carried out in urban VANET.

7 Future Work

VANET, because of its diverse application, has eventually involved in Indian government projects of the National Highways Authority of India (NHAI) [19] for the electronic toll collection (ETC) systems covering the nation. This system will constitute radio frequency identification (RFID) by installing an onboard unit (OBU) on a vehicle and a stationary roadside unit (RSU) at the toll plaza.

The U.S. Department of Transportation's National Highway Traffic Safety Administration (NHTSA) uses VANET to support safety applications and communication between vehicle and stationery roadside units (RSU) to reduce collisions.

Even though there are significant applications of VANET and much research has been made during the last decade in the field of Sybil detection in VANET but there are a lot of challenges before the research community as there are only a few researches which have been carried out in urban VANET. The algorithms which are fast, simple, and accurate are required to detect Sybil attack in urban VANET where the identity of vehicles should be kept confidential.

References

1. Nzouonta J, Rajgure N, Wang G, Borcea C (2009) VANET routing on city roads using real-time vehicular traffic information. IEEE Trans Veh Technol 58(7):3609–3626
2. Douceur JR (2002) The Sybil attack. In: Proceeding of the 1st international workshop on peer-to-peer systems
3. Kenney JB (2011) Dedicated short-range communications (DSRC) standards in the United States. Proc IEEE 99(7):1162–1182
4. Uzcategui R, Acosta-Marum G (2009) WAVE: a tutorial. IEEE Commun Mag 47(5):126–133

5. Newsome J, Shi E, Song D, Perring A (2004) The Sybil attack in sensor network: analysis a defense. In: Proceedings of the third international symposium on information processing in sensor networks. ACM Press, New York, NY, USA, pp 259–268
6. Park S, Aslam B, Turgut D, Zou CC (2009) Defense against Sybil attack in vehicular ad hoc network based on roadside unit support. In: Military communications conference, 18–21 Oct 2009. IEEE, pp 1–7
7. Hao Y, Tang J, Cheng Y (2011) Cooperative Sybil attack detection for position based applications in privacy preserved VANETs. In: Global telecommunication conference (GLOBECOM 2011). IEEE
8. Lee, BK, Jeong E, Jung I (2013) A DTSA (detection technique against a Sybil attack) protocol using SKC (session key based certificate) on VANET. Int J Secur Appl 7(3):1–10
9. Tang L, Hong X, Bradford PG (2008) Privacy-preserving secure relative localization in vehicular networks. Secur Commun Netw 1:195–204
10. Studer A, Shi E, Bai F, Perrig A (2009) Tacking together efficient authentication, revocation, and privacy in VANETs. In: SECON'09
11. Haas JJ, Hu Y-C (2007) Real-world VANET security protocol performance, pp 1–7
12. wireless LAN Working Group. Wireless access in vehicular environments. IEEE Standards, July2010
13. B. Xiao, B. Yu, and C. Gao. Detection and localization of Sybil nodes in vanets.InDIWANS'06,2006.
14. Garip MT, Kim PH, Reiher P, Gerla M (2017) INTERLOC: an interference-aware RSSI-based localization and Sybil attack detection mechanism for vehicular ad hoc networks. In: 14th CCNC, pp 1–6
15. Grover J, Laxmi V, Gaur MS (2014) Sybil attack detection in VANET using neighboring vehicles. Int J Secur Netw 9(4):222–233
16. Chen C, Wang X, Han W, Zang B (2009) A robust detection of Sybil attack in urban VANETs. In: 2009 29th IEEE international conference on distributed computing workshops
17. Chang S, Qi Y et al (2012) Footprint: detecting Sybil attacks in urban vehicular networks. IEEE Trans Parallel Distrib Syst 23(6):1103–1114
18. Hamed H, Keshavarz-Haddad A, Haghighi SG (2018) Sybil attack detection in urban VANETs based on RSU support. In: Iranian conference on electrical engineering (ICEE). https://doi.org/10.1109/icee.2018.8472629
19. http://www.nhai.org/
20. Zhou T, Choudhury R R, Ning P, Chakrabarty K (2011) P2DAP—Sybil attacks detection in vehicular ad hoc networks. IEEE J Sel Areas Commun 29(3):582–594

Double-Band Coplanar Antenna for GSM and UWB Applications

Dhivya Raj and C. V. Anil Kumar

Abstract A compact ($25 \times 30 \times 1.6\ \text{mm}^3$) coplanar antenna suitable to operate on both ultra-wideband (UWB) and global system for mobile communications (GSM) is presented. A coplanar waveguide (CPW)-fed antenna for 3.1–10.6 GHz (UWB) along with 1.7–1.8 GHz (GSM) will be useful for high-speed data transfer applications in embedded communication systems. Vital characteristics like radiation pattern, return loss, and bandwidth of the proposed antenna are investigated.

Keywords Coplanar antenna · Dual band · UWB · GSM · Bluetooth

1 Introduction

Frequency range of 3.1–10.6 GHz has been allotted for commercial communication applications on February 14, 2002, by the US Federal Communication Commission (FCC) [1]. Since then, antennas operating in ultra-wideband (UWB) spectrum have found wide applications due to its less power and high data rate features. The radiating element that can radiate/receive in both UWB and GSM bands is much more useful in the design of compact communication systems [2–4]. The global system for mobile communications (GSM) ranging from 1.7 to 1.8 GHz assists voice calls and short message service (SMS). Also, ultra-wideband (UWB) technology varying from 3.1 to 10.6 GHz can provide even higher data transfer rates of the order of 100 Mbits/s [5–7].

D. Raj (✉)
Department of ECE, College of Engineering Chengannur, Chengannur, Kerala, India
e-mail: dhivyaraj@ceconline.edu

C. V. Anil Kumar
Department of ECE, College of Engineering Karunagappally, Karunagappally, Kerala, India
e-mail: cvanil@ceconline.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_37

Table 1 Literature survey

Multiband antenna	Antenna dimensions ($H \times W \times T$)	Wireless bands
Proposed antenna	$30 \times 25 \times 1.6$	UWB, Bluetooth
Antenna in Ref. [8]	$30 \times 25 \times 1.6$	UWB
Antenna in Ref. [9]	$25 \times 24 \times 1.6$	UWB, GSM, GPS, Bluetooth
Antenna in Ref. [10]	$30 \times 30 \times 1.6$	UWB, GSM, WCDMA, WLAN
Antenna in Ref. [11]	$18 \times 30 \times 1.6$	Bluetooth, UWB
Antenna in Ref. [12]	$33 \times 28 \times 1.6$	WLAN, WiMAX, HiperLAN (5.25 GHz)
Antenna in Ref. [13]	$10 \times 34.5 \times 0.8$	LTE/WWAN
Antenna in Ref. [14]	$10 \times 6 \times 0.4$	2.4/5 GHz—WLAN

2 Literature Survey

A wide survey on different multiband antennas was conducted and is summarized in Table 1.

The antenna in Ref. [8] was developed on FR4 substrate (thickness = 1.6 mm, loss tangent = 0.02, and dielectric constant = 4.4). Various multiband printed antenna geometries have been studied with different ways to gain impedance matching and bandwidth enhancement in the literature for multiband applications [9–14].

3 Methodology

The simulation analysis is completed with high-frequency structure simulation (HFSS) Ansoft's software package.

4 Antenna Design

When upper part of ground plane is extended as a horizontal metallic strip from ground plane in the rectangular slot of UWB antenna in [8], a frequency of 2.2 GHz is obtained within the lower frequency range as in Fig. 1. When the horizontal strip from ground plane is once more extended and hung vertically down, the resonant frequency shifted more lower to 1.8 GHz. The thickness of horizontal metallic strip in the slot from ground plane (wn) and thickness of hanging strips (hn) are 0.3 mm. The optimum value obtained for length of the hanging strip, Lhp, is 5.2 mm. The optimum point *P* at which ground plane is extended as a hanging metallic strip from

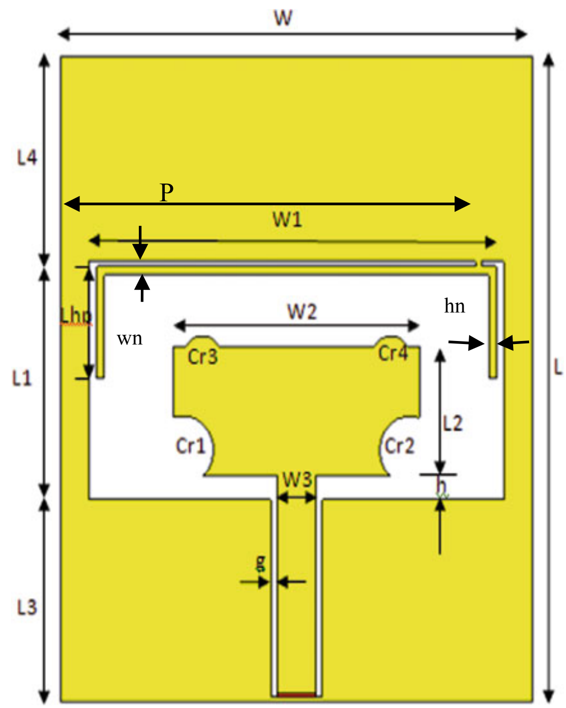


Fig. 1 Geometry of the proposed GSM/UWB antenna

ground plane into the rectangular slot is obtained as 22 mm of total width of antenna and is shown as *P* in Fig. 1.

The antenna in Ref. [8] and the one developed in this paper were developed on FR4 substrate (thickness = 1.6 mm, loss tangent = 0.02, and dielectric constant = 4.4); see Fig. 1 for dimensions of the antenna.

Size of the antenna is 30 mm × 25 mm. Microstrip feed width is 2 mm for a characteristic impedance of 50 Ω. The optimization of antenna dimensions is done by parametric analysis. So obtained dimensional values are given in Table 2. This antenna for GSM/UWB operation can radiate at 1.8 GHz with a bandwidth of 60 MHz and at UWB of 2.7–10.9 GHz with a bandwidth of 8.2 GHz. Figure 2 shows the simulation result of the one proposed.

Table 2 Dimensions of the proposed GSM/UWB antenna (unit mm)

W	L	L1	L3	L4	Cr1	Position of Cr1
25	30	11.1	9.4	9.5	1.6	(6.5, 11.7, 1.6)
W1	W2	Lhp	g	hn	Cr2	Position of Cr2
22	13	5.2	0.35	0.3	1.6	(18.5, 11.7, 1.6)
h	*P*	wn	Cr4	Position of Cr4	Cr3	Position of Cr3
1.1	22	0.3	01	(17.5, 16, 1.6)	01	(7.5, 16, 1.6)

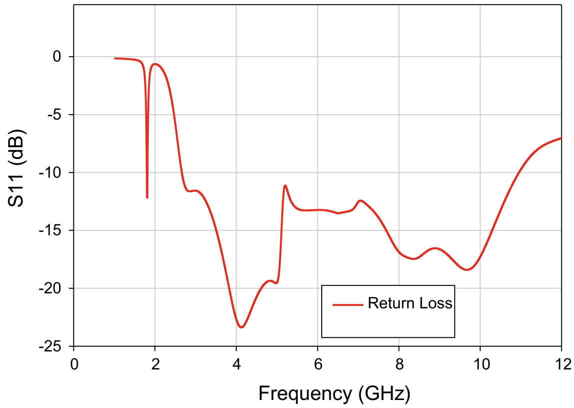


Fig. 2 Return loss of the proposed GSM/UWB antenna

5 Simulation Results

Various simulation results are discussed under return loss, parametric analysis, and radiation pattern in the following Sects. 5.1–5.3.

5.1 *Return Loss*

At resonant frequency of 1.8 GHz and 2.7–10.9 GHz, the return loss of GSM band is nearly as good as −13 dB as shown in Fig. 2. It also exhibits a better impedance matching from 2.7 to 10.9 GHz.

The current distribution pattern at 1.8 and 7.5 GHz illustrates that the hanging metallic strip from ground plane is accountable for the resonance at lower frequency of GSM band; see Fig. 3a, b, respectively.

5.2 *Parametric Analysis*

The thickness of horizontal metallic strip from ground plane (wn), thickness of hanging strips (hn), and the point at which ground plane extends as a hanging metallic strip from ground plane into the rectangular slot (*P*) are identified as critical parameters in the geometry.

The effect of Lhp, length of metallic strip hung from ground plane, is such that return loss of GSM band and UWB band is better maintained when Lhp = 5.2 mm; see Fig. 4. If Lhp is increased beyond 5.2 mm, say to 6.2 mm, the bandwidth (BW) of UWB is reduced. Also, if Lhp is decreased than 5.2 mm, say to 4.2 mm, the lower band got shifted away from GSM band though UWB band was still maintained.

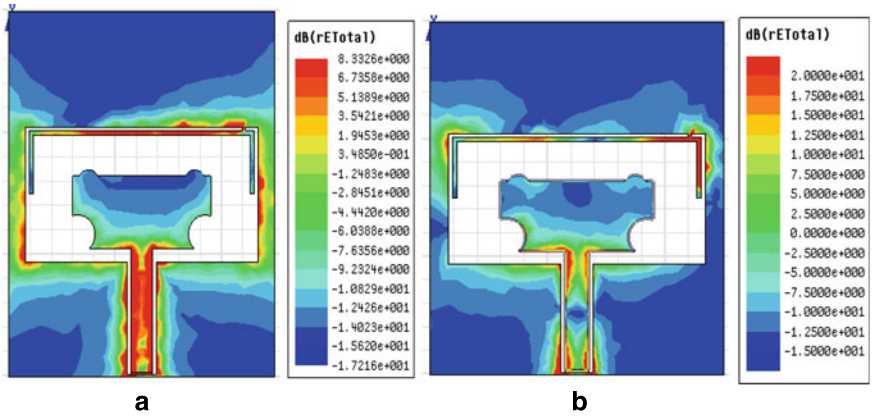


Fig. 3 **a** Surface current distribution at 1.8 GHz. **b** Surface current distribution at 7.5 GHz

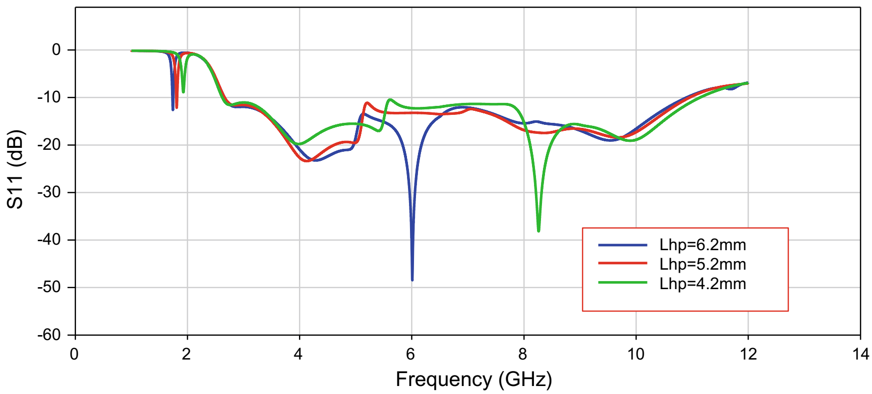


Fig. 4 Effect of Lhp on resonance

Hence, an optimum value for getting better resonance at GSM and better BW at UWB is found when Lhp = 5.2 mm.

The influence of the point, *P*, at which ground plane starts extending as hanging metallic strip into rectangular slot is such that bandwidth of UWB band degrades significantly for even small variations from the optimum value of *P* = 22 mm; see Fig. 5. When *P* is lower than the optimized value, say at 21.5 mm, and when *P* is greater than 22 mm, say at 22.5 mm, the total bandwidth of UWB band is reduced though resonance at GSM band is maintained. Hence, position of *P* is found critical in getting both dual bands of GSM and UWB.

Influence of width of the hanging metallic strip at horizontal section, wn, and vertical section, hn, is shown in Figs. 6 and 7, respectively. The optimum values obtained are wn = 0.3 mm and hn = 0.3 mm because at these values, GSM band had better resonance compared to other values of wn and hn while UWB band performed

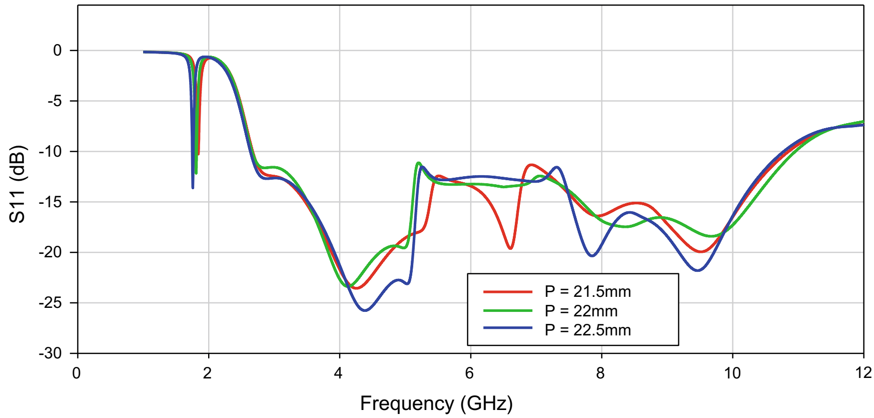


Fig. 5 Effect of *P* on resonance

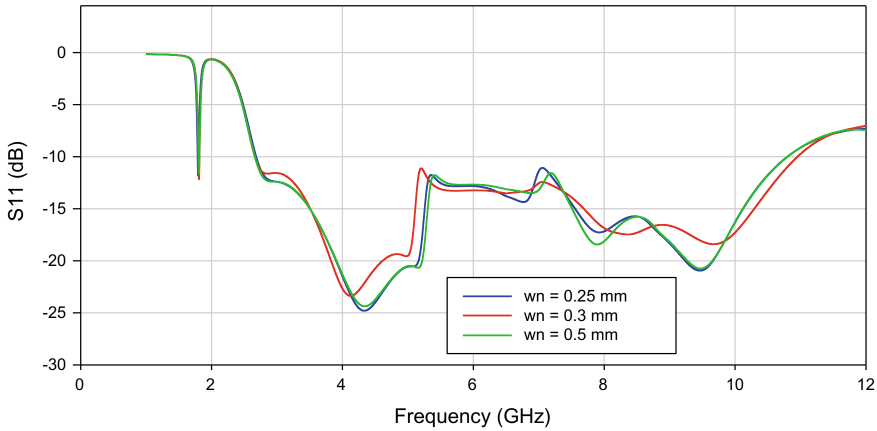


Fig. 6 Effect of wn on resonance

equally well in all other values of wn and hn; see Figs. 6 and 7. Hence, these are crucial parameters in designing the proposed GSM/UWB antenna.

5.3 Radiation Pattern

Radiation characteristics of the proposed antenna are analyzed. The two-dimensional radiation pattern in elevation and azimuth planes at the operating frequency of 1.8 GHz is shown in Fig. 8a, b, respectively. Good far-field radiation properties are achieved. The symmetrical structure resulted in an almost omnidirectional pattern in the *xz* plane.

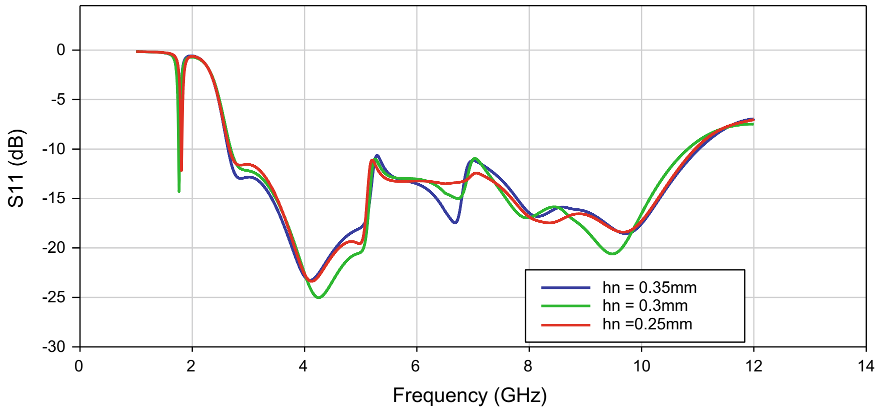


Fig. 7 Effect of hn on resonance

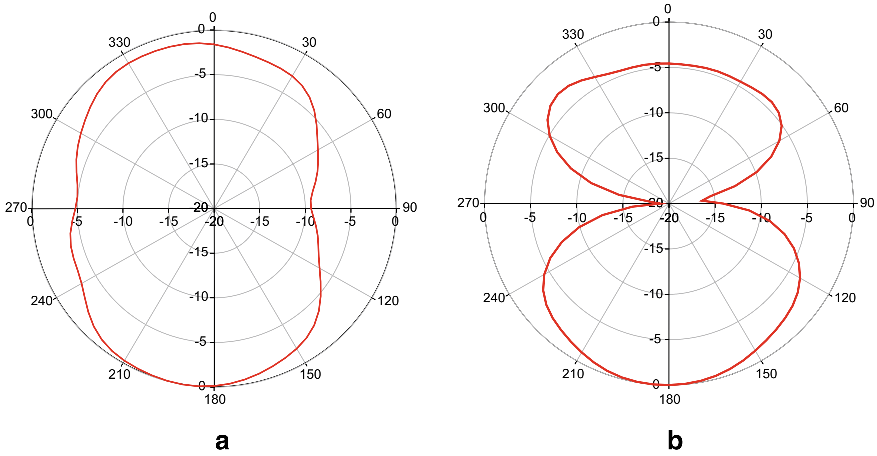


Fig. 8 **a** 2D radiation pattern at H-plane. **b** 2D radiation pattern at E-plane

6 Experimental Results

Fabrication of the antenna was done by photolithographic process, and the measurements were taken. The fabricated antenna and the measured return loss are shown in Figs. 9 and 10, respectively. An unexpected variation in UWB happened in experimental result (see Fig. 10). The poor return loss that happened must be due to variations in antenna dimensions during fabrication. The simulated and measured results at lower band are in good agreement between the simulated and measured results as shown in Fig. 10, and hence, the simulation result is validated.

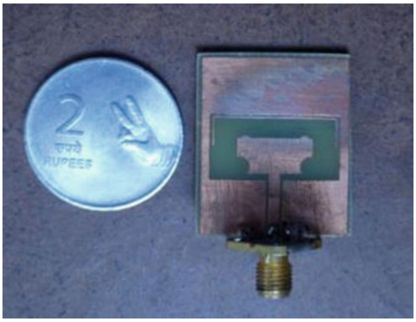


Fig. 9 Top view of fabricated antenna

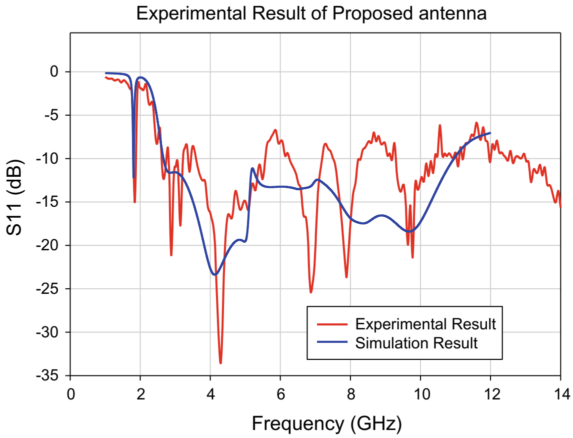


Fig. 10 Measured and simulated results

7 Conclusion

The paper presented here for a novel dual-band, cost-effective, and low-profile coplanar antenna is suitable for applications in GSM band at 1.8 GHz and UWB (3.1–10.6 GHz). Extensive parametric analysis has been conducted and critical parameters reported. The optimized structure resonates at 1.8 GHz with a bandwidth of 60 MHz starting from 1.77 to 1.83 GHz which almost covers the GSM band at 1.8 GHz. The structure can resonate at individual frequencies not throughout UWB. The band is extended due to multiple resonances in the whole UWB range of 2.7–10.9 GHz. The radiation properties are also promising for further investigations to develop compact multiband antenna suitable for wireless applications.

Acknowledgements The authors wish to thank the research scholars and faculty members of Centre for Research in Electromagnetics and Antennas, Cochin University of Science and Technology, for their help and support in antenna fabrication and experimental results.

References

1. FCC (2003) First report and order, revision of the part 15 commission's rules regarding ultra-wideband transmission systems, 98–153
2. Chen ZN (2007) Antennas for portable devices, 2nd edn. Wiley, New York
3. http://www.bluetooth.org. Last accessed 9 Dec 2018
4. Bensky A (2000) Short-range wireless communications. LLH Technology Publishing
5. Balanis CA (2016) Antenna theory: analysis and design, 3rd edn. Wiley, Hoboken
6. Schantz HG (2003) A brief history of ultra wideband antennas. In: IEEE UWBST conference
7. Hui ZHU (2008) Basic research on ultra wideband antenna and indoor positioning systems. Dissertation, Waseda University
8. Anil Kumar CV, Damodaran S, Paul B (2015) A printed antenna with circular cut for bandwidth enhancement. In: IEEE TENCON (2015)
9. Luo Y, Xu L, Xin Z, He S (2013) A compact CPW-fed UWB antenna with GSM, GPS, Bluetooth and dual notch bands applications. Prog Electromagnet Res C 35:205–219
10. Li G, Zhai H, Li T, Ma X, Liang C (2013) Design of a compact UWB antenna integrated with GSM/WCDMA/WLAN bands. Prog Electromagnet Res 136:409–419
11. Xiong L, Gao P (2012) Dual band planar monopole antenna for Bluetooth and UWB applications with WiMAX and WLAN band-notched. Prog Electromagnet Res Lett 28:183–194
12. Anil Kumar CV, Varghese B, Paul B (2013) A novel printed antenna for WLAN applications at 2.45 GHz. IJSER 4(8)
13. Wong K-L, Chen M-T (2013) Small-size LTE/WWAN printed loop antenna with an inductively coupled branch strip for bandwidth enhancement in the tablet computer. IEEE Antennas Wirel Propag Lett 61:6144–6151
14. Chien H-Y, Sim C-Y-D, Lee C-H (2013) Dual-band meander monopole antenna for WLAN operation in laptop computer. IEEE Antennas Wirel Propag Lett 12:694–697

Matrix Adaptor for Instrument Interface Interchangeability in ATE

Maria George, R. Sethunadh, S. Athuladevi, and B. Valsa

Abstract Two critical aspects of automation of test systems are the ability to operate different test instruments through a controller and their seamless integration. Instrument interchangeability is a key factor in test automation as it can make automation easier and overcome the obsolescence of instruments, thereby extending the overall lifetime of an automatic test equipment (ATE). This becomes more important in fields such as aerospace and nuclear systems where the test systems and test programs have operational lifetimes covering several generations of test instrumentation. Across the generations, a feature that changes in the instrumentation side is the instrument interface. Various instrument interfaces are developed such as RS232, GPIB and in recent times USB, Ethernet and so on. This many a time necessitates the need to ensure compatibility with the interface available at the controller side. A novel product called matrix adaptor is proposed here to overcome this limitation and guarantee an interface independent communication between the controller and test equipment. The proposed product can be used to adapt any interface on the test instrument side to any other on the controller side, and novelty of the idea is that the same product can be used for achieving more than one type of interface conversion at the same time. This work elaborates on the design considerations for such a product and gives the implementation details of a proto model with USB, RS232 and GPIB interfaces.

Keywords Automatic test equipment (ATE) · Instrument interface · GPIB · USB · RS232

M. George (✉) · R. Sethunadh · S. Athuladevi · B. Valsa
Vikram Sarabhai Space Centre (VSSC), Indian Space Research Organisation (ISRO), Trivandrum, Kerala, India
e-mail: mgtvm29@gmail.com

R. Sethunadh
e-mail: r_sethunadh@vssc.gov.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_38

1 Introduction

The increase in the complexity of systems being tested triggered the necessity of automatic testing since it allows quick and error free detection and analysis of faults. Test automation took off in the 1950s as digital logic and computers became available. Today, with the advancements in technology, computers became faster and powerful with increased memory and more capable software, making it the focal point of an automated test system. The term automatic test equipment (ATE) encompasses all phases of computer controlled testing and is based on the integration of instruments, computers and software [1]. The continuing challenge in the field of test automation is to improve ATE performance and increase the level of automation, thereby making the system more user-friendly.

The ATE performance is enhanced when a hybrid system is implemented where the strengths of many types of instruments can be combined, including legacy equipment and specialized new generation instruments. The performance of ATE is always closely linked to the performance of the instrument interface architecture used in the system [2]. Test and measurement instruments used in ATE are generally provided with one or more of the standard interfaces, such as General Purpose Interface Bus (GPIB), RS232, RS485, Universal Serial Bus (USB), Ethernet, Bluetooth, etc. The controllers also must have one or more of these interfaces. The automation of test and measurement is accomplished by programming the instruments using the controller through these standard interfaces. Among these interfaces, RS232 and GPIB are common in traditional instruments while USB and Ethernet are popular in modern instruments [2]. The general purpose PCs and laptops are nowadays very commonly used as controller in ATEs. The above interfaces except USB are not included as a default interface in new generation PC or laptop. Additional interface communication cards are hence required to be plugged into the controller for using these interfaces. Thus, in many cases, it is seen that the test and measurement instrument and the controller have no common interface. In this scenario, the interfaces that the user has to address are multiple in numbers. This raises the need for a single adaptor which can handle different interfaces and solve the problem of incompatibility at the instrument-controller interface as illustrated in Fig. 1.

The necessity of a single adaptor is further heightened by the role it plays for reducing complexity in instrument interchangeability, thus aiding test automation and life extension of the ATE. Instrument interchangeability is the ability to replace a given instrument with an alternative instrument of sufficient capability, but of a different model, class, design or manufacturer [3]. Interface is one main aspect of difference among the various instruments. This issue has to be addressed both at the software level as well as at the hardware level. At the software level, certain standards like virtual instrument software architecture (VISA) [4], interchangeable virtual instrument (IVI) [5] drivers, etc., exist to manage this. However, this has to be supplemented at the hardware level by having the corresponding interface at the controller side. VISA is a technology which helps to adapt to any interface via software. However, this requires the controller to have the corresponding instrument

Fig. 1 Concept of matrix adaptor

interface as a built-in option or as an additional add-on card. Another available solution is the use of interface converters at the controller side. Such solutions are highly dependent on the kind of interface used in the system. Hence, to cater to controllers and instruments from different generations, multiple such converters are required which adds cost and complexity of the test system. This necessitates a single adaptor which can be used irrespective of the kind of interface so that the system developer need not bother about the type of interfaces available at the controller side or instrument side. The proposed matrix adaptor is such a system which acts as a common interface to connect the test equipment to the controller. It is more useful in ATEs developed with modular instrumentation, i.e., ATE in which the instruments can be added, deleted or replaced. This may result in an ATE with instrumentation across various generations and hence different kinds of interfaces. In such scenario, matrix adaptor guarantees instrument interface interchangeability, thus enabling seamless integration of instruments in ATE from different generations.

The paper is organized as follows. Section 2 describes the system model. Section 3 details the system implementation, both hardware and software aspects. Section 4 presents the results and discussion, and Section 5 is the conclusion.

2 System Model

The proposed system is conceived as a matrix adaptor which can adapt any of the standard instrument interfaces such as GPIB, RS232, USB, Ethernet and Bluetooth to any other in the set. These interfaces differ mainly in three aspects, viz physical connector, electrical signal levels and communication protocol. Thus, an adaptor for interchange of interfaces should satisfy the physical layer compatibility for which the following three requirements have to be met:

- Interface connectivity: Different interfaces have different hardware connection for communication. For example, RS232 interfaces are usually terminated as a 25 pin or a 9 pin D-subconnector, USB interface is terminated as a 4 pin receptacle, GPIB as a 24-pin microribbon connector, Ethernet as an 8 pin RJ 45 connector, and Bluetooth devices have a built-in short-range radio transmitter. Hence, for proper mating of the different interfaces, a proper physical connector is required.
- Electrical signal level compatibility: The electrical signal levels vary for different interface standards. For example, RS232 standard specifies the use of negative, bipolar logic (± 3 to ± 15 V), whereas GPIB uses TTL logic levels with negative true logic. USB protocol specifies different signaling schemes at different speeds of transmission. At low speed, 1s and 0s are transmitted by sending a 5 V on one or the other of the two signal wires. For high speed lines, differential transmission is used. The standard Ethernet uses a baseband signal with Manchester encoding. Bluetooth being a wireless technology converts bits to signals by using frequency shift keying with Gaussian bandwidth filtering (GFSK). Thus, the proposed adaptor should be able to compensate for the variations in signal levels and signaling schemes of different interfaces.
- Protocol conversion: The communication protocol varies for different interfaces. The different schemes used for synchronization and supported data rates in these interfaces also differ greatly. For example, RS232 has a serial communication protocol, whereas GPIB has a parallel communication protocol. Unlike RS232 where the format of data being sent is not defined, USB is made up of several layers of protocols with packet data transfer. These interfaces are used for connecting peripheral devices; on the other hand, Ethernet is an interface for networking. It uses the carrier sense-multiple access/collision detection (CSMA/CD) to control access to the sharing medium. Bluetooth is different from all the above as it is a wireless protocol and the access method is time-division multiple access (TDMA). GPIB and RS232 are interfaces designed specifically for instrument control applications. On the other hand, USB and Bluetooth have become popular as interfaces for computer peripherals and Ethernet for networking applications. In spite of all these differences, the adaptor should be able to convert from one protocol to another without affecting the data integrity.

In the proposed design, interface connectivity is achieved by providing corresponding ports, electrical signal level compatibility by the usage of respective transceivers and protocol conversion using embedded software in a microcontroller. Figure 2 shows the structure of the proposed matrix adaptor.

3 System Implementation

The proposed adaptor is one step ahead of the available interface converters by virtue of its ‘any to any’ configuration. Hence, the challenges in its design were

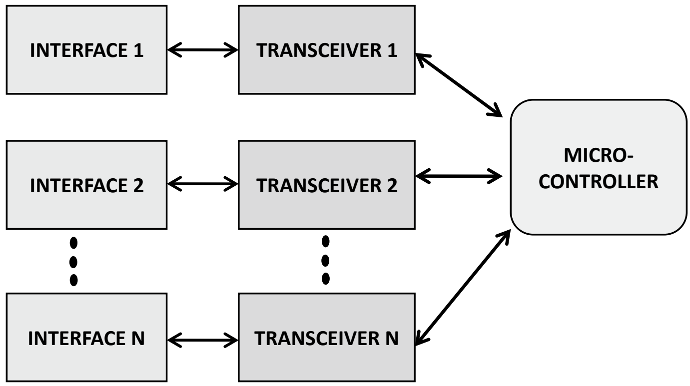


Fig. 2 Structure of the proposed matrix adaptor

many. During the implementation, the different design considerations which have been made are as follows:

- Selection of microcontroller: The microcontroller forms the core of the system. A careful selection of the microcontroller is an important aspect for the efficient realization of the system. There are certain off the shelf microcontrollers which have in built USB interface such as Cypress 68013A, Atmel AT90S and there are certain system on chip (SoC) devices like Microsemi SmartFusion which have inbuilt Ethernet interfaces. The usage of such microcontrollers makes it easier to implement the interface part as the interface transceiver is embedded within the controller. Also certain controllers such as Microchip ENC28J60-a standalone Ethernet controller with SPI interface, Silicon labs CP2200—single chip Ethernet controller, Texas Instruments LMX9838—Bluetooth serial port module with UART interface, etc., can be used along with microcontroller to increase the number of interfaces that are supported. In the proposed system, Cypress 68013A-based design has been implemented.
- Buffer size: A good amount of onboard storage is required to manage the difference in data rates of different protocols. RS232 protocol supports data rate of up to 20 kbps only while GPIB supports a data rate of up to 8 Mbps. Depending on the version, USB supports 1.5, 12 Mbps or 480 Mbps. Standard Ethernet has a data rate of 10 Mbps, whereas the latest version is 10 Gigabit Ethernet, which operates at 10 Gbps. The current data rate of Bluetooth is 1 Mbps with 2.4 GHz bandwidth. Thus, to have a seamless and error free protocol conversion and communication, a considerable buffer size is required. In our case, this was managed by double buffering techniques and usage of external memories.
- Synchronization: This is an important aspect of any kind of communication. Parallel communication protocols such as GPIB takes care of this by means of three handshaking signals. RS232 protocol specifies hardware handshaking using Data Terminal Ready/Clear to Send (DTR/CTS) signal lines or software handshaking by Xoff/Xon character. USB protocol defines a sync field in every packet to ensure

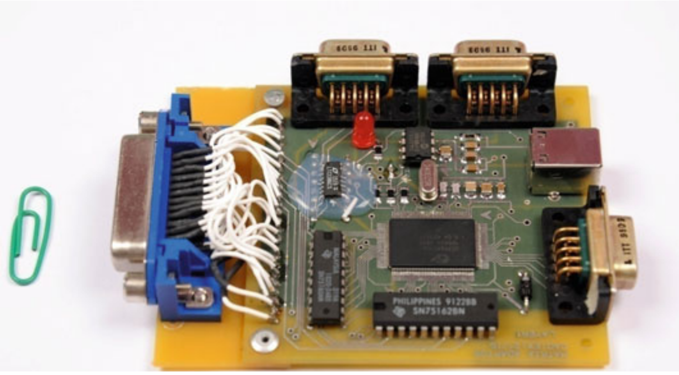


Fig. 3 Proto model of matrix adaptor

synchronization of the receiver clock with that of the transmitter. In Ethernet standard, the first field of the Ethernet frame 'Preamble' contains seven bytes which enable the receiving system to synchronize its clock if it is out of specification. In Bluetooth, this is managed by using synchronization bits in the 72-bit field 'access code.' The adaptor should maintain the synchronization irrespective of the kind of interface.

- Ease of use: An ideal system should have a plug-and-play ease of use. In the case of matrix adaptor, this largely depends upon the mode of identification of the required protocol conversion. It can be implemented by a simple manual selection by means of a switch or a software selection using a graphical user interface in the host controller. An alternate and ideal choice is the automatic identification of the required conversion based on the status of the live connections. This is implemented by incorporating status polling or interrupt routines. The simultaneous use of different conversions can also be achieved by maintaining exclusive buffers for each conversion.
- Overall size and power consumption: The size of the adaptor forms a critical part in determining its suitability for an ATE. More compact the adaptor can be, more suitable it becomes to form part of an ATE. Also, the power management is an important aspect to be considered. When connected to USB, it can be configured to be bus powered and otherwise it can be battery powered.

3.1 Implementation Details

In the sections below, the details of implementation of a proto model of matrix adaptor which has the most commonly found three interfaces—USB, RS232 and GPIB is explained.

Hardware Implementation: As shown in Fig. 3, the implementation achieves interface connectivity by giving appropriate physical hardware connectors—Type B receptacle for USB, 9 pin D-sub connector for RS232 and 24-pin microribbon connector for GPIB.

The circuit schematic is as shown in Fig. 4: The main component selected for the proto model of matrix adaptor is Cypress CY7C68013A-128AXC microcontroller. The primary reason for choosing this microcontroller is that it has a built-in USB interface. The other major components are RS232 transceiver IC LTC1386CS, GPIB transceiver ICs SN75160 and SN75162 for maintaining electrical signal compatibility during protocol conversion and 24 MHz 20 pF crystal for clocking. The main power is derived either through USB or external power source at 5 V DC, and a voltage divider circuit with LT series regulator LT1117IST is used to generate 3.3 V for supply to microcontroller as well as RS232 transceiver.

Software Implementation: The primary element in software implementation is programming of microcontroller which has been done in Keil μVision integrated development environment. A technical challenge here is to implement protocol conversions such that more than one interface conversion can happen at the same time. The means for achieving this is twofold: firstly, by providing independent buffers for each conversion, and secondly, by proper management of conversion identification. To implement independent buffers, this work makes use of independent endpoints. USB protocol defines registers known as endpoints which are indirectly accessed by the device drivers for data exchange [6]. In this implementation, four endpoints were used—One for communication from USB to RS232, one for RS232 to USB, one for USB to GPIB and one for GPIB to USB data transfer. The USB specification defines four transfer types: control, interrupt, isochronous and bulk. Bulk transfer has been used in this implementation.

The next challenge of proper management of conversion identification has been implemented by choosing different identification mechanism for each conversion. RS232 protocol being a serial communication protocol, its identification was implemented in the microcontroller as interrupt-based and GPIB protocol being a parallel communication protocol, polling-based identification was implemented.

The RS232 interface was implemented using the UART port of the microcontroller along with RS232 transceiver IC LTC1386CS. Configuration of the UART port in terms of baud rate, no. of stop bits, no. of parity bits, etc., was made in software.

GPIB interface was implemented using the standard input/output port of the microcontroller along with GPIB transceiver ICs SN75160 and SN75162. GPIB protocol defines three signals for handshaking—Not Ready for Data (NRFD), Not Data Accepted (NDAC) and Data Valid (DAV). GPIB protocol defines talker as the instrument which can send data over the bus and listener as the instrument which can accept data from the bus. During data transfer, ‘DAV’ is controlled by the ‘Talker’ and ‘NRFD’ and ‘NDAC’ are controlled by the ‘Listener.’ Additionally, before communication, the address has to be properly set, and in this work, it is assumed that the first byte sent is the GPIB address. So the first byte is sent as address and the remaining as data. To achieve proper protocol conversion, the data that is available in one buffer has to be transferred to the data lines of GPIB by properly managing

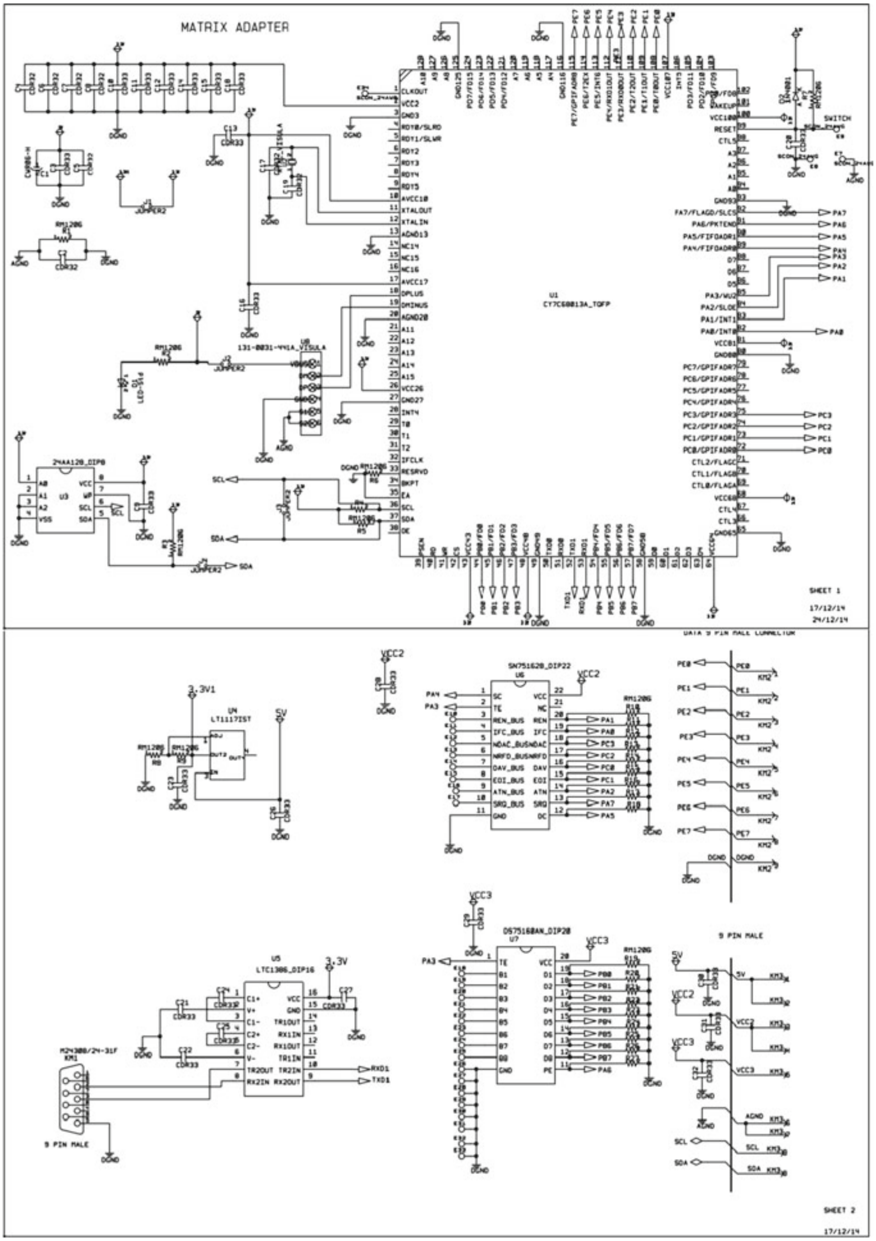


Fig. 4 Circuit Schematic of the proposed matrix adaptor

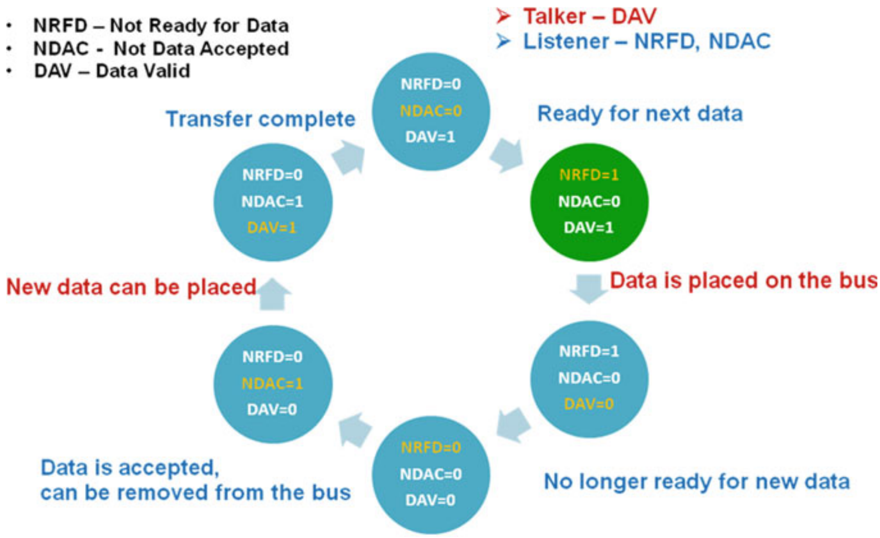


Fig. 5 Implementation of GPIB protocol

the pins corresponding to the above-mentioned signals. The implementation of the handshaking protocol for one byte of data transfer is depicted in Fig. 5.

4 Results and Discussion

The following test setup was made to verify the functionality of the product. A laptop was chosen as the controller, and a power supply with RS232 interface (TDK Lamda) and a function generator (Yokogawa FG110) as well as a digital multimeter (Keithley 2000) with GPIB interface were connected as instruments. In the current test setup, the GPIB addresses chosen were 11 and 12 for function generator and multimeter, respectively. Due to lack of RS232, GPIB interfaces in the laptop, there is no way to connect it directly, and matrix adaptor effectively poses a single solution. The test setup is as shown in Fig. 6. The total current drawn from USB was less than 200 mA which is well within 500 mA, the maximum current can be drawn from USB.

A graphical user interface (GUI) was created in LabVIEW to control the matrix adaptor as shown in Fig. 7. This can be used for the control of the matrix adaptor in other applications also. The interface has provision to enter the commands, separately for RS232 and GPIB, provide GPIB address in case of GPIB instruments as well as provision to display reply if any. Through this GUI, a sequence of commands was given through USB interface of the laptop to enable the various instrument functions. Sample results of the test are given in Table 1 and are found to be satisfactory.

The product was also found to be useful under the following cases:

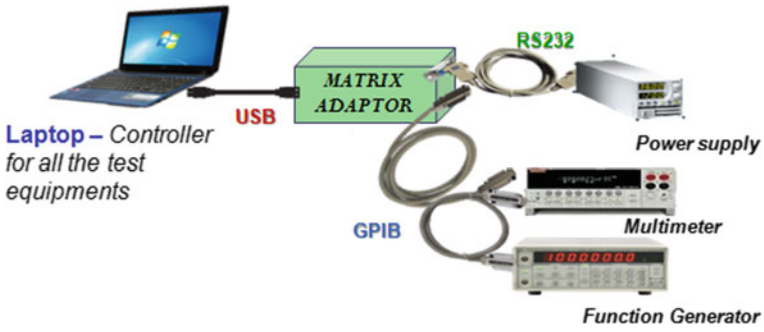


Fig. 6 Test setup with the proposed matrix adaptor

Fig. 7 GUI for the proposed matrix adaptor

Table 1 Sample results of functional verification test of matrix adaptor

S. no.	Test parameters entered in GUI			Result
	RS232 command	GPIB address	GPIB command	
1	:RMT2 :VOL3.0	12	:FUNC 'VOLT:AC'	Power supply is in remote mode and has voltage set to 3 V Multimeter is in remote mode and function has switched to AC voltage mode
2	:OUT1	11	:FUNC SIN:FREQ 100Hz:AMPL 1 V:OFFS 0.2V:OUTP ON;	Power supply has turned ON Function generator is in remote mode and sine wave is selected with frequency 100 Hz, amplitude 1 V and offset 0.2 V
3	:OUT0	12	:FUNC?	Power supply has turned OFF Multimeter has given the reply corresponding to AC voltage mode

- To overcome the obsolescence of equipments:
  In general, any test set up consists of different test equipments. However, as time passes, one or more of the equipment may become obsolete. The search for similar equipment need not result in equipment with the same interface. Matrix adaptor helps in this scenario by seamlessly interfacing the new interface to the system. Thus, it is useful in reducing the dependency on one particular brand or type of equipment with respect to the type of instrument interface.
- To make the ATE compact and portable:
  ATEs from earlier generations are based on PCs and control equipments through GPIB add-on cards and RS232 ports. New-generation laptops can be used to make the system compact and portable, but they lack RS232 ports and provision for additional add-on cards. Using the matrix adaptor, laptop was inducted as the controller, thus reducing size and increasing portability of the system without making any change in instrument interface requirements.
- To control multiple instruments through a single controller:
  The use of matrix adaptor helps in increasing the number of instruments that can be controlled. For example, if a controller has one USB port, then ideally only one instrument with a USB interface can be controlled. However, with the proto

model of the matrix adaptor, a total of 15 equipments (one instrument via RS232 and 14 instruments via GPIB) could be controlled.

5 Conclusions

Here, a novel system is proposed which can be used to develop and maintain an ATE. It transparently recognizes instrument interface standards, thereby helping in the seamless integration of the controller of an ATE with different generations of instruments. Higher the number of interfaces that can be handled, more versatile the adaptor is. The proposed system is implemented in a modular way, so that more types of interfaces can be added to the matrix configuration. The system can be further extended to increase the number of interfaces that can be simultaneously interlinked, thus expanding the scope of application.

References

1. Automated testing. IEEE Aerosp Electron Syst Mag 15(10):125–130 (2000)
2. Sethunadh R (2005) Using new instrument interface standards to increase automatic test system performance. In: Proceedings of IEEE AUTOTESTCON, 2005, pp 677–682
3. Ramachandran N, Oblad RP, Neag IA, Tyler DF (2000) The role of a signal interface in supporting instrument interchangeability. In: Proceedings of IEEE AUTOTESTCON, 2000, pp 403–416
4. Web site: www.vxipnp.org, VISA Specifications
5. Spoelder HJW (1999) Virtual instrumentation and virtual environment. IEEE Instrum Meas Mag 2:14–19
6. de Almeida Pereira Zuquim LD et al (2001) An embedded converter from RS232 to Universal Serial Bus. In: Proceedings of 14th symposium on integrated circuits and system design, pp 91–96

Performance Analysis of Security Algorithms

E. Soundararajan, Nikhil Kumar, V. Sivasankar, and S. Rajeswari

Abstract Computers have become more and more potent over the years, and consequently, it has become more comfortable and easier to break encryption and hashing algorithms. Hence, it is essential to study the performance of these algorithms and analyze them, and possibly come up with better algorithms based on the knowledge we have gained from the analysis. In this paper, a comprehensive literature study of the conventional encryption and hashing algorithms is done, followed by a practical comparison of the time-efficiency and CPU usage of these algorithms, two important performance evaluation parameters.

Keywords Hashing · Encryption · Running time · CPU usage

1 Introduction

Security is the primary concern of any communication over the network. As the number of devices connected through Internet increases, there is an increasing demand for enhanced security without much compromise on performance overheads. The primary purpose of security of information is to achieve confidentiality and integrity. Hashing and encryption algorithms play an active role in reducing security threats.

E. Soundararajan (✉) · V. Sivasankar · S. Rajeswari
Indira Gandhi Centre for Atomic Research, Kalpakkam, India
e-mail: sound@igcar.gov.in

V. Sivasankar
e-mail: sivav@igcar.gov.in

S. Rajeswari
e-mail: raj@igcar.gov.in

N. Kumar
Birla Institute of Technology and Science, Pilani, India

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_39

1.1 Hashing

Hashing is the process of converting a message or a file into a string of fixed length, through a one-way process called a hashing algorithm. Hashes cannot be reversed in a reasonable time frame, which is why hashing is termed a one-way process.

Cryptographic hash algorithms must satisfy three core properties, as follows:

- Pre-image resistance: Given a hash value h, it should not be computationally feasible to generate a message m such that hash (m) = h.
- Second pre-image resistance: Given a message m1 with hash h, it should not be computationally feasible to generate another message m1 = m2 such that hash (m2) = h.
- Collision resistance: It should be computationally infeasible to generate two messages m1 and m2 which have the same hash h.

1.2 Encryption

Encryption, like hashing, is a method of scrambling data (the plaintext) to make it unreadable (the ciphertext). However, the major difference is that encryption is a two-way process, and it can be reversed. Encryption is followed by secure transfer of the ciphertext and decryption (unscrambling of data).

Symmetric key algorithms These algorithms use the same key for both encryption and decryption. Note that in this kind of algorithm, the key needs to be transported across the network from the sender to the receiver, making it vulnerable to attacks. But on average, encryption and decryption happen faster for symmetric algorithms as compared to asymmetric algorithms.

Asymmetric key algorithms These algorithms use two keys—a public key and a private key. If A needs to send a message to B, he encrypts it using B's public key. The only key which can decrypt a message encrypted by B's public key is B's private key, thus ensuring that only B can decrypt the message that was sent for him. Note that in this algorithm, the key need not be transported across the network, making it more secure than symmetric key algorithms. However, asymmetric key algorithms are more expensive in terms of time for encryption and decryption.

2 Structure of Hashing and Encryption Algorithms

Most hashing and encryption algorithms follow a particular set structure. All the hashing algorithms considered in this paper support the Merkle–Damgård construction, while among the encryption algorithms considered, Camellia, CAST5 and DES follow the Feistel structure.

2.1 Merkle–Damgård Construction [3]

The steps involved in this template are:

1. The message is padded until it is of a certain length.
2. The modified message is padded again with a binary representation of the original message length, until the message is of a length (in bits) equal to a multiple of a certain length n bits. This is called Merkle–Damgård strengthening.
3. The message is divided into blocks of size n bits.
4. A buffer of size equal to the length of the final hash value is initialized.
5. Compression functions are used to modify the buffer in each round. The compression functions take in inputs of a certain word length w, so the n-bit block is divided into n/w words of length w. Each word is processed using the compression functions, and the buffer is updated.
6. Once all words in a block are processed, the same process is repeated for the remaining blocks, and the buffer being updated in every iteration.

Merkle–Damgård construction for MD4, MD5, SHA1 and SHA256 As listed above, all the steps of the Merkle–Damgård construction can be seen in the block diagram below. First, the message is padded in the pattern 100…0.0 until it is of length (512n—64) bits. Second, the modified message is padded with a 64-bit representation of the length of the original message (as two 32-bit words in little endian format, with the lower-order bit first). Third, a buffer is initialized, where the initial buffer is provided as part of the algorithm. Then the message is divided into 512-bit blocks, and each block is further subdivided into 16 32-bit words. Each word is processed along with the buffer using the compression function to update the buffer. This is done for every word in a block, and every block in the message, until the last block is reached and the final hash (buffer) is generated (Fig. 1).

Merkle–Damgård construction for SHA512 The procedure is very similar for SHA512 as well, except here, the message is broken down into 1024-bit blocks, and each word is 64 bits long.

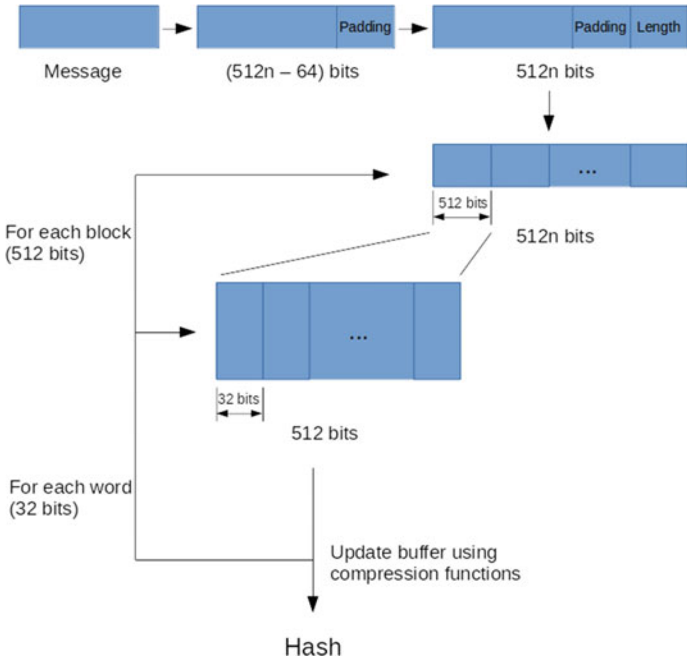


Fig. 1 Merkle–Damgård construction for MD4, MD5, SHA1 and SHA256

2.2 Feistel Structure [8]

Just as most hashing algorithms follow the Merkle–Damgård construction, many encryption algorithms follow a template called the Feistel structure.

Before delving into the specifics of the Feistel structure, let us first discuss the types of ciphers.

1. Block ciphers: These ciphers process plaintext block-by-block. Their input is a block of bits of fixed length, and each block is processed individually. Many block ciphers follow the Feistel structure. Of the algorithms considered, all except RC4 are block ciphers [9].
2. Stream ciphers: These ciphers process plaintext bit-by-bit, which is why their working appears to be 'continuous,' hence stream. Stream ciphers do not follow the Feistel structure. RC4 is a stream cipher.

The Feistel structure consists of processing the message in multiple rounds. The output of one round goes as input to the subsequent round. Each round consists of the following steps [8]:

1. Break the input into halves—left and right.
2. Keep the left half as is, and apply a function to the right half.

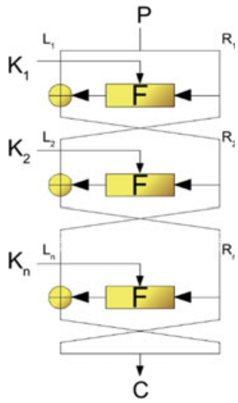


Fig. 2 Feistel structure (*source* https://upload.wikimedia.org/wikipedia/commons/thumb/d/d2/Feistel.png/220px-Feistel.png)

3. XOR the left half and the result of the function.
4. Swap the left and right halves (Fig. 2).

3 Results

The algorithms were run on machine with i5 processor having 3.20 GHz clock speed, 4 GB DDR3 RAM and 500 GB SATA HDD on a Linux platform. The algorithms were compared on running time and CPU usage for different file sizes. OpenSSL library functions were used to implement both the encryption and hashing algorithms in C.

3.1 Hashing

Running time. Running time of each algorithm was measured for files containing words, each file being of a different size. The time was measured using the clock() function available in the C time.h library. Running time was measured for MD4, MD5, SHA1, SHA256 and SHA512, some of the most common cryptographic hash algorithms (Figs. 3 and 4).

MD4: $y = 0.019x + 0.012$
MD5: $y = 0.022x + 0.010$
SHA1: $y = 0.022x + 0.016$
SHA256: $y = 0.032x + 0.014$
SHA512: $y = 0.043x + 0.016$

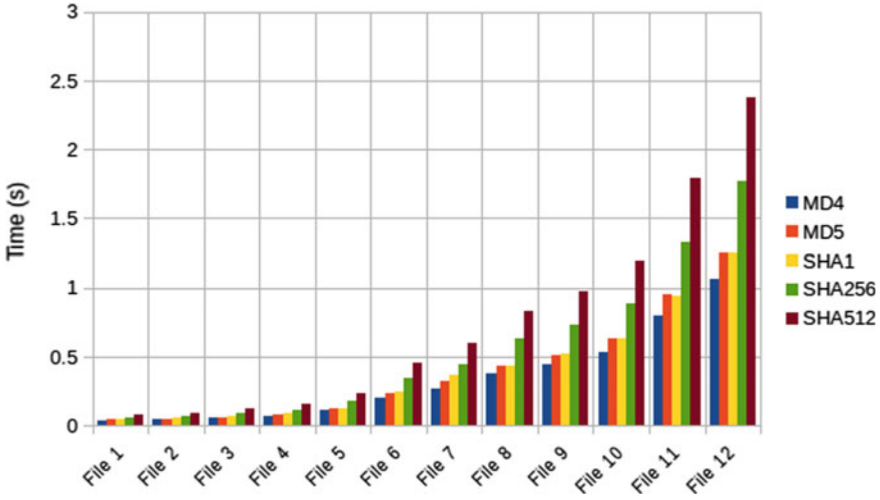


Fig. 3 Comparison of time taken for hashing

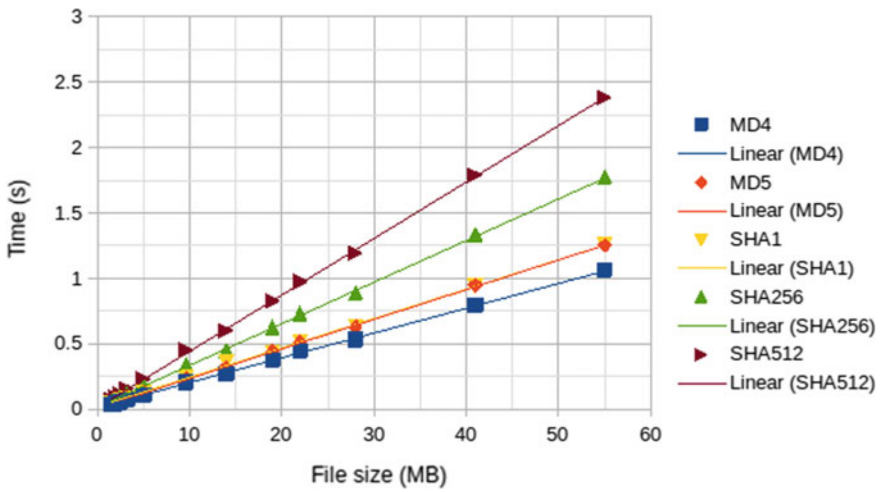


Fig. 4 Time versus file size regression lines

The regression lines have the following equations:

We observe that the variation of time for hashing with the size of the file is very clearly linear. This is consistent with the fact that the algorithms have complexity that vary as $\theta(n)$ 10, 11, where 'n' is the size of the file. Note that the slope of each graph is in the order MD4 < MD5 < SHA1 < SHA256 < SHA512. This is indicative of the fact that the SHA2 (SHA256 and SHA512) algorithms take more time on average than the MD4, MD5 and SHA1 algorithms.

Table 1 Number of rounds per block for hashing algorithms

Algorithm	Number of rounds
MD4	48
MD5	64
SHA1	80
SHA256	64
SHA512	80

These patterns are observed because the running time depends on two factors, namely

1. The number of rounds of operation
2. The complexity of the functions used (Table 1).

The running time varies as

$$\text{MD4} < \text{MD5} < \text{SHA1} < \text{SHA256} < \text{SHA512}$$

However, the number of rounds per block of these algorithms does not exactly follow this pattern. Sure, the correspondence is clear for MD4, MD5 and SHA1, where both running time and number of rounds vary as MD4 < MD5 < SHA1, i.e., running time and number of rounds have a positive correlation, as expected, but this is not the case with SHA256 and SHA512. SHA256 has a higher running time than SHA1 though it has fewer rounds. SHA512 has a significantly higher running time than SHA1 though both have the same number of rounds.

The above observation is because both SHA256 and SHA512 use more functions in a single round. For instance, SHA1 uses four functions for t-values from 0 to 79, whereas SHA256 uses six functions for t-values from 0 to 63. SHA512 uses six similar functions as SHA256, but for t-values from 0 to 79, which explains why SHA256 < SHA512. Now, it is clear why we observe the pattern SHA1 < SHA256 < SHA512.

Another explanation is that the hash values of SHA1, SHA256 and SHA512 are 160, 256 and 512 bits long, respectively. This means the size of the buffer is in the order SHA1 < SHA256 < SHA512. More the number of words in a buffer, the more the number of assignment operations we need to perform to change the value of each word in the buffer, which takes more time. This too corresponds to the observed order SHA1 < SHA256 < SHA512.

CPU usage It was observed that CPU usage was independent of file size. The top command in Linux was used to measure the CPU usage for each of the algorithms (Table 2).

There is a simple reason as to why CPU usage is independent of file size. While running the algorithm, it was observed that the CPU usage at every instant was almost constant. In other words, the CPU is capable of performing a certain number of hashes per second, which is reflected in the CPU usage percentage. The CPU

Table 2 CPU usage for hashing algorithms

Algorithm	Average CPU usage (%)
MD4	55.9
MD5	55.1
SHA1	52.2
SHA256	53.8
SHA512	52.4

'sees' a certain number of words per second and hashes them. The reason why CPU usage is independent of file size in context of hashing algorithms is that the CPU is bothered about the number of hashes per second (i.e., at a point in time) and the usage is independent on how many hashes have already been performed/how many more hashes are to be performed, and so is independent of file size.

3.2 Encryption and Decryption

Running time Running time of each algorithm was measured for files containing words, each file being of a different size using the clock() function of C. In each case, it was observed that encryption times were longer than decryption times. This is as expected, as AES, Camellia, CAST5 and DES have all been run in ciphertext feedback (CFB) mode. CFB mode encryption always occurs linearly, one block after the other (Fig. 5).

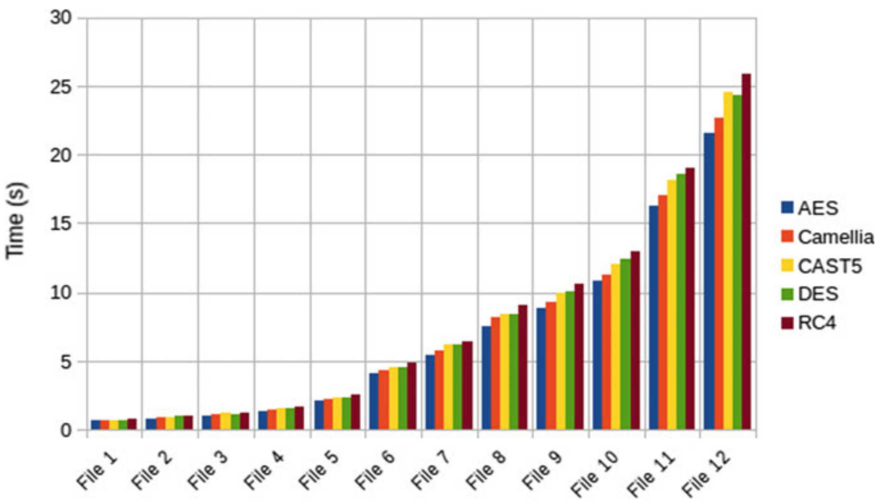


Fig. 5 Comparison of time taken by encryption algorithms

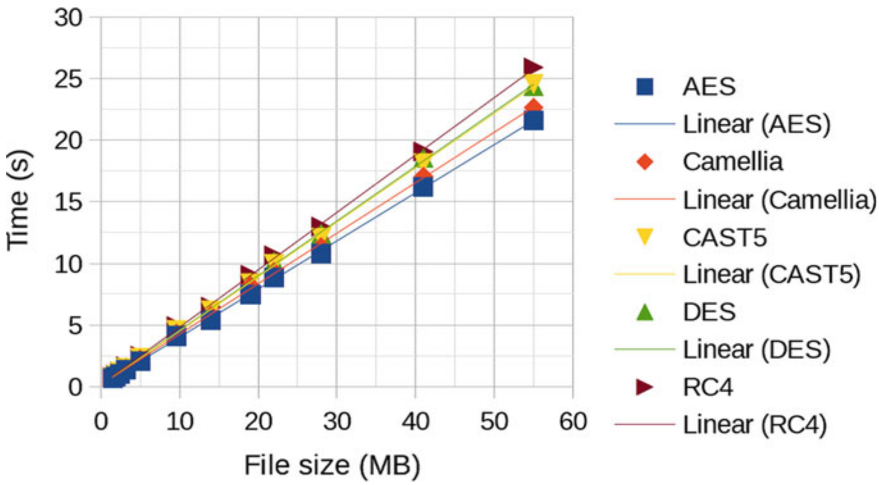


Fig. 6 Running time versus file size regression lines for encryption algorithms

However, CFB mode decryption can be done in parallel, i.e., multiple blocks can be decrypted simultaneously [12]. Thus, decryption takes shorter time than encryption.

In most cases, we observe that running time varies as

$$\text{AES} < \text{Camellia} < \text{CAST5} < \text{DES} < \text{RC4}$$

However, in some cases, running time varies as

$$\text{AES} < \text{Camellia} < \text{DES} < \text{CAST5} < \text{RC4}$$

So, let us take the general pattern to be (Fig. 6)

$$\text{AES} < \text{Camellia} < \text{CAST5} < \text{DES} < \text{RC4}$$

The regression lines have the following equations:

$$\boxed{\begin{array}{l} \text{AES: } y = 0.0390x + 0.111 \\ \text{Camellia: } y = 0.410x + 0.153 \\ \text{CAST5: } y = 0.441x + 0.094 \\ \text{DES: } y = 0.443x + 0.127 \\ \text{RC4: } y = 0.466x + 0.152 \end{array}}$$

CPU usage Again, we find that CPU usage is independent of file size. While running the algorithm, it was observed that the CPU usage at every instant was almost constant. In other words, the CPU is capable of performing a certain number of

Table 3 CPU usage for encryption algorithms

Algorithm	CPU usage (%)
AES	55.72
Camellia	56.88
CAST5	56.98
DES	57.13
RC4	55.83

encryptions per second, which is reflected in the CPU usage percentage. The CPU 'sees' a certain number of words per second and encrypts them. The reason why CPU usage is independent of file size in context of encryption algorithms is that the CPU is bothered about the number of encryptions per second (i.e., at a point in time) and the usage is independent on how many encryptions have already been performed/how many more encryptions are to be performed, and so is independent of file size (Table 3).

4 Conclusions and Recommendations

It is important that the suitable hashing and encryption algorithms to be deployed. The choice of the algorithms depends on level of security required, performance and the available computing infrastructure.

4.1 Recommendations for Hashing Algorithms

Following are some recommendations based on the data and analysis above:

1. If speed is the primary concern, and security is not as important, it is preferable to use MD5 or SHA1. MD4 is not recommended as both MD4 and MD5 take similar running time, but MD4 provides less security. The decision between MD5 and SHA1 can be made based on how much one wants to compromise speed for security (SHA1 is more secure, but marginally slower than MD5).
2. If security and speed are both equally important, it is preferable to use SHA256 or SHA512. Therefore, SHA256 and SHA512 are best suited for message authentication and digital signatures. SHA3 is another standard that is coming up and may be suitable for use in the future. SHA256 and SHA512 are still considered relatively fast (but slower than MD4, MD5 or SHA1) hashing algorithms [13], and so are unsuitable for hashing passwords. Though the results of this report may suggest that SHA256 and SHA512 are slow, there are still slower algorithms that are considered safer for password hashing, such as bcrypt, scrypt and PBKDF2 [13].

4.2 Recommendations for Encryption Algorithms

Following are some recommendations based on the data and analysis above:

1. RC4 is the slowest in terms of running time. Also, several vulnerabilities have been found in RC4, and it is no longer considered safe [14]. Therefore, RC4 is not recommended to be used at all.
2. Among the other ciphers, the order of security [15] is CAST5 < DES < Camellia < AES. The running time is in the order AES < CAST5 ≈ DES < Camellia and CPU usage is in the order AES < RC4 < Camellia < CAST5 < DES. AES scores the best in security, running time and CPU usage. Therefore, AES is by far the best among the algorithms considered. This is consistent with the fact that AES is considered the best and most respected cipher in industry today. Thus, all parameters considered, it is best to use AES.

References

1. AlAhmad MA, Alshaikhli IF (2013) Broad view of cryptographic hash functions. Int J Comput Sci Issues (IJCSI) 10(4):239. https://pdfs.semanticscholar.org/a5ea/6caf93c2d4b74b30ff095cb7e1b91c19a085.pdf
2. Ramachandran J (2002) Designing security architecture solutions. Wiley, New York
3. Tiwari H (2017) Merkle-Damgård construction method and alternatives: a review. J Inf Organ Sci 41(2):283–304. https://pdfs.semanticscholar.org/7abf/5eb51960c26dfe5fedbd0404c97ab4df5277.pdf?_ga=2.221023384.1428183823.1560855221-910554265.1559028851
4. Rivest R (1992) The MD4 message-digest algorithm (No. RFC 1320). https://tools.ietf.org/rfc/rfc1186.txt
5. Rivest R (1992) The MD5 message-digest algorithm (No. RFC 1321). https://www.ietf.org/rfc/rfc1321.txt
6. Eastlake D 3rd, Jones P (2001) US secure hash algorithm 1 (SHA1) (No. RFC 3174). https://www.ietf.org/rfc/rfc3174.txt
7. Eastlake D 3rd, Hansen T (2006) US secure hash algorithms (SHA and HMAC-SHA) (No. RFC 4634). https://tools.ietf.org/html/rfc4634
8. Kanthety SS (2017) Network security—feistel structure (Video file). https://www.youtube.com/watch?v=drI2shandyk
9. Wikipedia Contributors (2019) Feistel cipher. In: Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Feistel_cipher&oldid=877632002
10. Rachmawati D, Tarigan JT, Ginting ABC (2018) A comparative study of Message Digest 5 (MD5) and SHA256 algorithm. J Phys Conf Ser 978(1):012116. https://pdfs.semanticscholar.org/0bd7/dfda1dfca713044dac2c16d23124968e9c1b.pdf?_ga=2.210161429.1428183823.1560855221-910554265.1559028851 (IOP Publishing)
11. StackOverflow Contributors (2017) Space & time complexity of SHA-2. https://stackoverflow.com/questions/47557136/space-time-complexity-of-sha-2
12. StackExchange Contributors (2015) Is encryption time greater than decryption time? https://crypto.stackexchange.com/questions/25465/is-encryption-time-greater-than-decryption-time
13. lucaskauffman (2013) About secure password hashing (Blog post). https://security.blogoverflow.com/2013/09/about-secure-password-hashing/

14. Weerasinghe TDB, Disanayake C (2018) A research study: usage of RC4 stream cipher in SSL configurations of web servers used by Sri Lankan Financial Institutes. Int J Cyber Secur Digit Forensics 7(2):111–119. https://www.researchgate.net/publication/323118887_A_research_study_Usage_of_RC4_stream_cipher_in_SSL_configurations_of_web_servers_used_by_Sri_Lankan_Financial_Institutes
15. Nkiama M, Herve MS, Herve S. Comparative study Of AES, Blowfish, CAST-128 and DES encryption algorithm. Structure 4(8):4. https://pdfs.semanticscholar.org/7af1/ac803d2ba5a2a9b419eb41974811c8fcf558.pdf

Quality Grading of the Fruits and Vegetables Using Image Processing Techniques and Machine Learning: A Review

M. K. Prem Kumar and A. Parkavi

Abstract In this paper, the study on fruits and vegetables grading system by using image processing techniques with various machine learning algorithms is presented. In our study, we observed that the images of fruits are used for analyzing about fruits. The parameters like shape, size and color of the samples are extracted for grading the fruits' quality. Some of the research has done on fruit disease classification. Different techniques like color-based segmentation, artificial neural networks and different classifiers can be used to classify the grade or diseases of the sample. This review work is to provide a study of some machine learning techniques and morphological features with color-based grading for fruits and vegetables. The authors have proposed a hybrid system of fruit grading and disease detection system.

Keywords Fruit grading · Vegetable grading · Fruit disease · Feature extraction · Fruit classification · Computer vision · Morphological · Texture

1 Introduction

1.1 A Subsection Sample

The major income of India is from agriculture, if the overall production is good, then the income will also be good. The developing requirement to supply continuous high-quality food products in a fixed timeframe is promoting the automated grading of the agricultural products. This can be done by using computer vision for quality inspection and evaluation purposes.

In the past, the fruit or plant diseases were identified by the experienced farmers. The agricultural experts identify plant diseases by naked eye observation. Based on the observations and suggestions, farmers were deciding the method and medicines for the treatment of the diseased plant and fruits.

M. K. Prem Kumar (✉) · A. Parkavi
M.S. Ramaiah Institute of Technology, Bengaluru 560054, India
e-mail: pkumarmk@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_40

As now, there are new diseases in which a farmer may fail to decide the type of the disease. So, the farmers need to consult the agricultural experts. For the visit, it may cost more to the farmers. This method is slow and hard and getting trust worthy experts also difficult.

The physical nature of agriculture product's like shape, color shading and size, etc., cannot be assessed online by the conventional strategies. Improvements in image processing, software and hardware technology have made easier to distinguish the fruits' and vegetables' quality by using vision detecting technology. High-quality yield will increase the income of the farmers.

This paper discusses various methods and techniques involved in the field of image processing to grade various types of fruits and vegetables and classify the diseases of the fruit. The possible improvements in each of the research work have been identified and specified at the end of the paper.

2 Literature Survey

Mishra et al. [1] proposed a system to distinguish the fruits as good and bad based on their quality. They made use of preprocessing techniques, segmentation techniques, feature extraction, training and matching. The preprocessing steps involve (1) input image (2) background subtraction (3) convert RGB to gray (4) convert gray image to binary image (5) filter the image using MATLAB. They used k-means clustering algorithm for classification and achieved with 80% accuracy.

Prince et al. [2] proposed a system that the captured image is processed by considering the physical features like size, color and shape of fruit samples extracted. Artificial neural network (ANN) is used to train and test the system. Their proposed system used neural network to detect shape, size and color of fruit. The results obtained are promising with the combination of these physical features. The preprocessing steps involve acquisition and image conversion: (1) input image (2) background subtraction (3) RGB image to gray image (4) gray image to binary image (5) filtering. In their model, median filter is used for smoothening of the image, Canny edge detector is used for segmentation process and ANN is used to classify the image. The clustered values are considered more for good and bad fruits.

Telang and Shirsath [3] proposed system for detecting fruit size and grading using processing flow, color detection, detection of the edge, fruit size detecting algorithm and grading the fruit size. In this model, RGB color model is used for color detection and Canny method is used for segmentation which differs than the others as it can handle noise, likely to detect actual weak edges. Standard size detecting algorithm works on fruits' symmetry which finds and measures two attributes: the center coordinates of fruit and axis in the image. Size is graded by detecting the diameter of an apple. This algorithm will categorize the quality of an apple into two grades, big and small, based on the size.

Khadabadi et al. [4] discussed various methods and techniques involved in the field of image processing to detect diseases in various types of vegetables. The disease and their symptoms on vegetables varies in color, size and shape based on the cause. Machine vision techniques are used in this model to the solve problems based on extraction of morphological features and analysis which include features like size, shape, color and texture of the surface.

Jadhav and Patil [5] proposed a classification and grading system. They designed a model to join three procedures, for example, feature extraction, sorting as per color shade and evaluating the grade as per size. Here, grading is categorized into four groups, based on the color and size: red small, green small, red big and green big. Grading based on size is determined by the diameter of an apple which is based on two attributes: (1) identifying the center coordinates of fruit (2) finding axis.

Patil and Gaikwad [6] proposed a model for recognition of a particular vegetable. The input image is split into frames, and the same is differentiated into cubes which represent the features. These features are converted into the data set. Having these values, certain frame is categorized into one of the subsets of the input image, and at the conclusion, the vegetables are identified based on the isolation of the object. They have used computer vision, image processing and convolutional neural network (CNN) to implement system. The CNN avoids the complex preprocessing steps and receives the original image as input. They have implemented a model to classify the vegetables with short training time and resulted 99% of high accuracy.

Pandey et al. [7] discussed on different classification techniques and their comparison based on their merits and demerits. They gave study of machine learning, color-based grading algorithms and their components which are useful for automatic fruit grading system. Their model has mainly concentrated on color classification.

Bhoumik [8] proposed an automated approach for fruit grading. First, they classified the group of the fruit by using the speeded up robust features (SURF), and then after classification, they detected the grade of the fruit (1) good (2) medium (3) bad using the color features and area of the fruit. In this model, minimum distance classifier is used for the classification. The average accuracy for fruit detection is 87.48% and grading is 78.9%. Debasmita found that the SURF method is robust and fast for comparison of images and gets local similarity invariant representation. The main interest of the SURF approach is using box filters and computing the operators fast. Real-time applications like object recognition use box filters. Debasmita obtained an accuracy of 87.48% for fruit detection, and for fruit grading, the accuracy is 78.9% in their model.

Raja Sekar et al. [9] done study on the analysis of different fruit disease detection techniques. They observed that the morphological features, color and texture exhibit highest accuracy rate in disease detection. Hue, Saturation, Intensity (HIS) is a color model used for grading as it is near to human perception. Support vector machine (SVM) resulted high accuracy, but adaptive neuro-fuzzy interference system (ANFIS) results the best out of other compared techniques, and fuzzy model is easy to implement but results low accuracy rate.

Seema et al. [10] reviewed a collection of papers highlighting their advantages and disadvantages, listed and showed the comparison between the methodologies. Research work concluded that SVM exhibits high accuracy.

Arakeri and Lakshmana [11] proposed a quality evaluation system for tomato. The system overcomes the drawbacks of manual techniques used in the sorting process of fruit. This system includes two steps: (1) fruit handling (2) image processing. The fruit handling step is for image acquisition. The acquired images are further analyzed using image processing techniques. This will determine defect and ripeness of the tomato. Image processing steps include segmentation, feature extraction and selection of color statistics, color texture. Defective or non-defective of tomato is indicated by implementing n-input, h-hidden and 1-output neuron in three-layer feed-forward neural network. Here, the backpropagation procedure supervises the neural network and adjust the weights. Experiments are carried out on several images of the tomatoes. The software classified the tomato with the accuracy rate for defective or non-defective is 100% and for ripe or unripe is 96.47%, respectively.

Al Ohali [12] proposed a grading system for date fruits using RGB images of the dates. The physical and external visible features are extracted from these images. The extracted features are further classified into three categories (grades—1, 2 and 3) as defined by experts. The performance of the BPNN classifier has studied, and accuracy of the system had tested on selected date fruit samples. The system resulted with 80% of accuracy in sorting of dates. Preprocessing module steps carry a binarization threshold using image intensity histogram. Feature definition and extraction include flabbiness. The estimation of the flabbiness is done by distributing the intensity of the color in gray-level image. It is observed that the image of the least flabby date fruit is darker than flabbier date fruit. Size of the fruit is considered to decide the quality of the dates. Bigger in the size is the good quality. The size is calculated by the coverage area of the image. Shape irregularity is also considered for quality measure. Irregular shapes are also considered to grade better quality. More wrinkles and more edges estimate more intensity. The image is binarized to determine the intensity, and Sobel operator is used to extract the edges. For classification, backpropagation is used along with neural network which is BPNN. BPNN contains an input layer, one output layer and at least one intermediate hidden layer which contribute to the maximum delay.

Ganatra and Patel [13] discussed the factors which cause the plant diseases and various techniques used for detecting the disease, for segmentation of the affected part to classify the disease. They made the comparative analysis on various segmentation techniques, feature extraction techniques and different classifiers with their benefits and drawbacks.

Dubey and Jalal [14] proposed the disease detection and classification system for apple fruit. K-means-based image segmentation is preferred to detect the region of the infected part. They used multi-support vector machine for the classification of the disease. For the feature extraction experiment, they made use of local binary pattern (LBP), complete local binary pattern (CLBP), color coherence vector (CCV), global color histogram (GCH). They observed that the result was similar with small variations. They concluded that CLBP feature resulted with more accuracy on different diseased apples, resulting average accuracy 93.14%.

Table 1 Lists the fruit samples and the techniques used in different stages and purposes with accuracy level

Sample	Attributes	Method	Accuracy	Reference
Tomato	Color	Multilayer neural network	Defective/non-defective: 100%, ripe/unripe: 96.47%	Arakeri and Lakshmana [11]
Dates	Flabbiness, shape, size, intensity, defect, RGB	BPNN	80%	Al Ohali [12]
Apple Orange Pineapple	Color (for detection)	RGB color space	89.95% 91.45% 81.06%	Bhoumik [8]
Lemon	Color, shape, size (defect segmentation) RGB, k-means	ANN	Prototype success	Prince et al. [2]
Apple	Color, size	Standard formula to find diameter (for size)	–	Telang and Shirsath [3]
Apple	Color, texture HSV, CLBP	Multi-SVM	93.14%	Dubey and Jalal [14]
Vegetables	Color	CNN	99%	Patil and Gaikwad [6]

3 Comparative Analysis

See Table 1.

4 Proposed System

There is a need of hybrid system which will be useful in both grading system in post-harvesting and disease detection system in pre-harvesting. The factors that cause diseases in agriculture products can be classified into two categories: living and non-living agents. Living agents are virus, bacteria, fungus and insects, while non-living agents are excess moisture, insufficient light, less nutrients and pollution [13].

The database consists of data which describes the diseases and mapped with the possible treatment.

Steps:

1. User will send the digital image of the fruit/vegetable sample to the application server through smart phone.

2. The system will process the image in series of steps, preprocessing, segmentation, extraction and classification.
   Acquisition: Images can be obtained by digital camera or any other media.
   Preprocessing: Obtained images need to eliminate unwanted area, removing noise, smoothening the image, converting the image to gray/binary image.
   Feature Extraction: It can obtain the features like color, shape, size and texture to reduce the resources to describe the data set, so that classification can be done using less data.
   Classification: Classifier will classify the image based on the morphological features extracted in the previous step. This helps in recognizing the disease. Here in the proposed system, convolutional neural network is used for classification.
3. Map the disease and possible treatment in map table which is stored in database.
4. Send the disease and possible treatment as reply to the user's smart phone.

The proposed system can classify the disease and map the disease with the possible treatment in database.

Method: CNN is a method of deep learning and a class of neural network used in Image recognition by specialized way of processing on grid of on pixels. The main building blocks of CNN involve layers such as input layer, an output layer and intermediate hidden layers that include multiple convolutional layers, pooling layers and finally fully connected layer (Fig. 1).

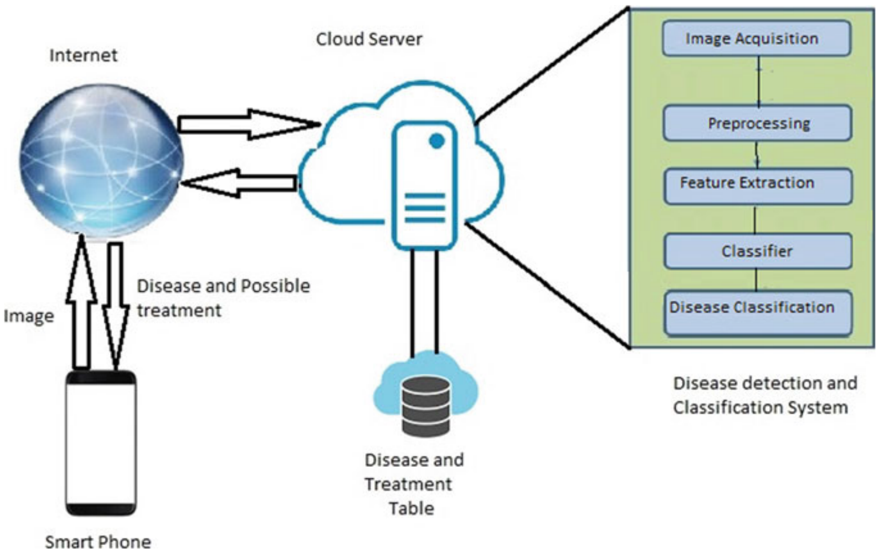


Fig. 1 Model shows how the image undergoes each step, and treatment is returned to the corresponding disease

The advantage of CNN is less number of neurons and distributed weights, pre-processing step can be customized and image can be directly provided as input. The series of multiple hidden layers convolve with a multiplication or other dot product as shown in (see Fig. 2).

CNN reduces processing requirements by using multilayer perceptron. The efficiency of the image processing increases when the operation done on resolution level. A CNN can successfully capture the spatial and temporal dependencies in an image through relevant filters.

CNN will result more accuracy compare to SVM as the pattern of fruit disease will change in shape and color intensity and results in irregularities. SVM results good in sorting the given input and classify the category. But CNN can analyze the image by processing the data with grid-like topology and recognize the patterns on the input object.

The advantage in using conventional neural network are the number of neurons in input layer will be significantly less, computationally efficient and accuracy is more.

Convolution layer will extract features from the given input image, and it holds the relationship between pixels by learning image features using the small frames of input data. This layer computes the dot product between matrix of learnable parameter and kernel. CNN compares the images piece by piece and gets a better similarity than whole image matching. The result of one layer is convolved to subsequent layer as input (Fig. 3).

Rectified Linear Unit (ReLU layer) is an activation function which activates a node if the input is above a value. The output is zero if the input is negative, but when the input rises above a certain threshold, it has a linear relationship with the dependent value. This function will remove the negative value and replace zero in the corresponding matrix of feature as shown in (see Fig. 4).

Pooling Layer will reduce the size of the input. It will scale down the input by keeping important information. This helps in solving overfitting and improving the

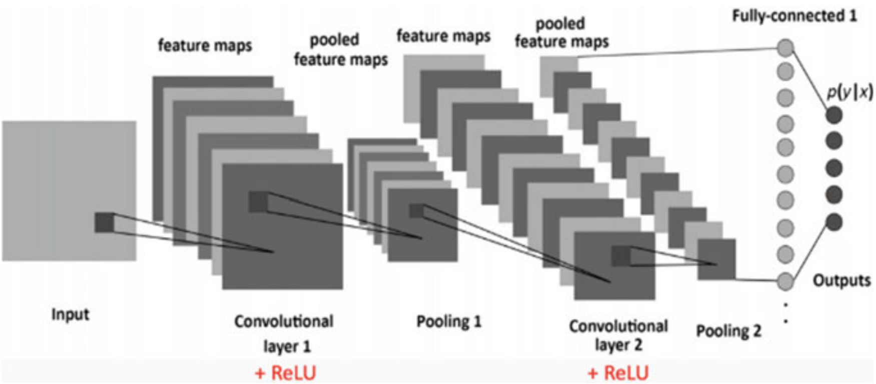


Fig. 2 A model of CNN with two convolutional layers, two pooling layers and a fully connected layer which classifies the image into one of the categories in the final layer

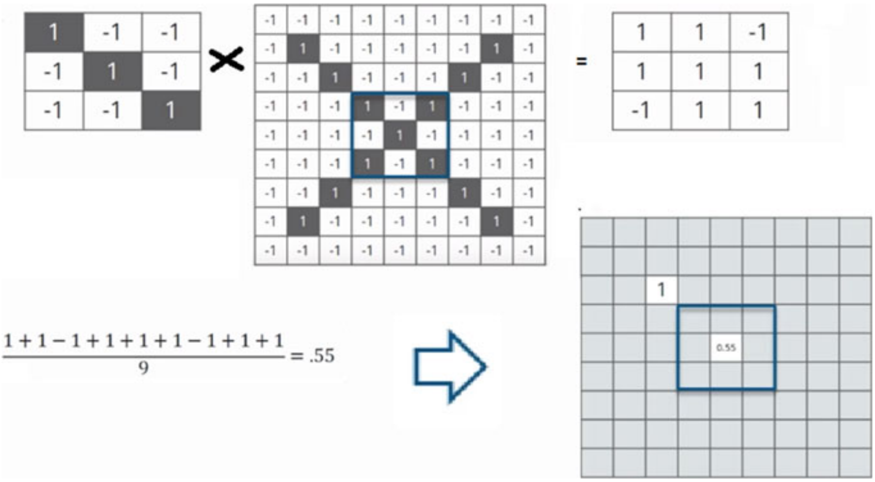


Fig. 3 How the original image segment is calculated with selected filter/feature

Fig. 4 ReLU function is half rectified. $f(x)$ is zero when x is less than zero and $f(x)$ is equal to x when x is above or equal to zero

performance. Below image shows how pooling will summarize the features (see Fig. 5).

Fully Connected Layer: The last layer is similar to regular neural network. It is fully connected to its preceding neurons which receives and computes the class scores from preceding layers and results one-dimensional array of size equal to the number of classes. Now, these learned vectors are used to compare with input image. The high values are summed up and divided by the total number of high values. This value is now compared with total number of high values in the vectors. The pattern is matched with the vector which results high compared to other vectors.

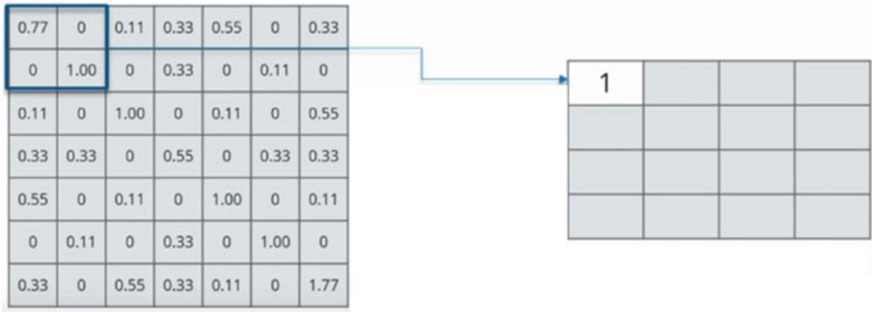


Fig. 5 Shrinking the image from 2 * 2 to 1, further considered to single vector after passed through fully connected layer

5 Conclusion

Research shows that different analytics methods will generate different accuracy level on different fruits/vegetable. Combination of the color scheme, segmentation technique and classification methods also results in variation of accuracy.

The proposed method is a theoretical idea and needs to be implemented. At the end, the technology should reach to the end users in simple way to solve their problems. More research has done on post-harvest period which will be an advantage for the high-end retailers. There is a need to combine the grading system and disease detection system into one application which can be used in both pre-harvesting and post-harvesting where the farmer can get the advantage by identifying the disease of the fruit earlier and reduce the loss in the yield.

References

1. Mishra A, Asthana P, Khanna P (2014) The quality identification of fruits in image processing using MATLAB. Int J Res Eng Technol 3(10). NCCOTII 2014
2. Prince R, Sathish H, Sai Kumar R, Karthik V, Jeya Anusuya S, Vijay Vardhan Reddy SP (2018) Identification of quality index of fruit/vegetable using image processing. IJARIIT 4(2)
3. Telang SS, Shirsath SM (2017) Fruit quality management using image processing. In: International conference on ideas, impact and innovation in mechanical engineering (ICIIIME 2017), vol 5, no 6, pp 727–733
4. Khadabadi GC, Rajpurohit VS, Kumar A, Nargund VB (2015) Disease detection in vegetables using image processing techniques: a review. Int J Emerg Technol Comput Sci Electron (IJETCSE) 14(2)
5. Jadhav RS, Patil SS (2013) A fruit quality management system based on image processing. IOSR J Electron Commun Eng (IOSR-JECE) 8(6):1–5
6. Patil O, Gaikwad V (2018) Classification of vegetables using TensorFlow. Int J Res Appl Sci Eng Technol (IJRASET) 6(IV)
7. Pandey R, Naik S, Marfatia R (2013) Image processing and machine learning for automated fruit grading system: a technical review. Int J Comput Appl 81(16). ISSN:0975-8887

8. Bhoumik D (2018) Fruit grading system using computer vision techniques. Int J Res Technol Stud 5(5)
9. Raja Sekar L, Ambika N, Divya V, Kowsalya T (2018) Fruit classification system using computer vision: a review. Int J Trend Res Dev (IJTRD) 5(1)
10. Seema, Kumar A, Gill GS (2015) Automatic fruit grading and classification system using computer vision: a review. In: Second international conference on advances in computing and communication engineering
11. Arakeri MP, Lakshmana (2016) Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry. In: 7th international conference on communication, computing and virtualization 2016. Procedia Comput Sci 79:426–433
12. Al Ohali Y (2011) Computer vision based date fruit grading system—design and implementation. J King Saud Univ Comput Inf Sci 23:29–36
13. Ganatra N, Patel A (2018) A survey on diseases detection and classification of agriculture products using image processing and machine learning. Int J Comput Appl 180(13)
14. Dubey SR, Jalal AS (2012) Detection and classification of apple fruit diseases using complete local binary patterns. In: Third international conference on computer and communication technology

Exploring Various Aspects of Gabor Filter in Classifying Facial Expression

Seetha Parameswaran, Murali Parameswaran, Shelbi Joseph, and Daleesha M. Viswanathan

Abstract Facial expression detection is a well-studied domain in which facial features are extracted and then classified into six common expressions. One of the most common techniques used for extracting features is the Gabor filter. In literature, for extracting the features, the combined magnitude and phase values of the Gabor filter are used. This paper is exploring the performance of methods using the combined filtering method, using magnitude alone and using phase alone in the domain of facial expression detection. It is observed that considering phase values with the support vector machine classifier yielded an additional 8% accuracy when compared to combined methods.

Keywords Facial expression recognition · Gabor wavelets · Gabor magnitude · Gabor phase · Feature extraction · Support vector machine · Classification

1 Introduction

Facial expressions convey approximately 55% of messages in human communication [20]. Human expressions are very complex, involving multiple muscular contractions, resulting in changes in facial texture. Though there are multiple expressions, based on prominent facial muscle movements, expressions are usually classified into the six predefined emotions, namely happiness, sadness, disgust, anger, surprise, and

S. Parameswaran (✉) · S. Joseph · D. M. Viswanathan
Cochin University of Science and Technology, Cochin, India
e-mail: seethap@cusat.ac.in

S. Joseph
e-mail: shelbi@cusat.ac.in

D. M. Viswanathan
e-mail: daleesha@cusat.ac.in

M. Parameswaran
Ramrao Adik Institute of Technology, Navi Mumbai, India
e-mail: murali.p@rait.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_41

fear [7]. Figure 1 illustrates the basic expressions commonly used in research related to the automatic facial expression recognition (FER) problem.

A typical automatic facial expression recognition system involves detecting the face, extracting the facial expression features, and recognition and classification of expressions as shown in Fig. 2. Several robust algorithms have already been suggested in the literature [9, 10, 22, 27, 28, 33] to identify a face from an image and detect and classify the expression.

This paper is concentrating on a holistic or appearance-based approach. A variety of methods like principal component analysis (PCA) [4], discrete cosine transformation (DCT) [13], histogram of gradients (HoG) [8], independent component analysis (ICA) [2], Gabor filter [5, 16, 19, 32], hidden Markov model (HMM) [29], and local binary patterns (LBP) [1, 23, 24, 34] were applied to face regions to extract the changes in facial appearance [26]. This paper is restricted to analyzing the effectiveness of the use of a Gabor filter for the extraction of features for FER.

The Gabor filter-based technique involves generating the signal values by convolving the input image with the complex filter. The values generated are in a complex plane, having both magnitude and phase components. In the literature, these complex values, representatives of facial features, are directly passed to the next stage of FER. In Oppenheim and Lim [21], it was suggested that using phase values of Fourier signals can improve feature extraction. In this paper, phase values determined by the imaginary plane from Gabor filters are used for extracting features as part of FER.

Fig. 1 Facial expressions in CK data set [14]

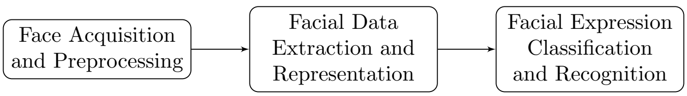


Fig. 2 Automatic facial expression recognition systems [28]

The rest of the paper is organized as follows. A brief survey of related literature work is presented in Section 2. Section 3 details the mathematical aspects of the Gabor filter and associated FER framework. Section 4 discusses the result, and the summarization of the paper is presented in Section 5.

2 Literature Review

Detecting facial changes due to facial expression is key in facial expression recognition. The second step of Fig. 1 plays an important role in identifying the facial changes that occur due to expression. These changes are dynamic in nature. Added to this, the issues that accompany the image acquisition processes like illumination variance and pose variance add-on to the challenges of FER.

This work is using a holistic or appearance-based approach for extracting features. The other approaches that can be used for feature extraction are a geometric approach based on fiducial points or regions on faces and hybrid method based on both appearance and geometry of the face [9].

The more recent research shows that a deep convolutional network can also be applied to give a better performance of the FER system.

A comparison of various methods in the literature is shown in Table 1.

To classify the expressions, Kotsia et al. [15] used the Gabor filter and support vector machine (SVM), where surprise showed an accuracy of 99.67% and anger resulted in 82% accuracy.

Shan et al. [24] used both local binary pattern and Gabor filter for feature extraction. These extracted features were classified using SVM with linear kernel, poly-

Table 1 Comparison of results using Gabor filter bank

Reference	Feature extraction method	Database	Expression best captured	Expression worst captured
Zhang et al. [33]	Gabor, MLP	JAFFE	Surprise	Fear
Kotsia et al. [15]	Gabor, SVM	CK	Surprise, happy	Anger
Shan et al. [24]	Gabor, SVM	CK	Surprise, disgust	Fear, sad
Hsieh and Jiang [12]	Gabor, LoG, SVM	CK	Happy	Fear
Gu et al. [11]	Gabor, FLD, KNN	CK	Surprise, disgust	Anger, fear
Sun and Yu [26]	Gabor, LBP, PCA, LDA, SVM	CK+	Surprise, happy	Anger, fear
Zhao et al. [35]	LBPTOP, Gabor, CLM	CK+	Happy, disgust	Anger, fear
Chengeta and Viriri [3]	Gabor filtered 3D LBP	CK+	Disgust	Anger

nomial kernel, and kernel with radial basis function (RBF). The author observed that the RBF kernel showed a better result in classification. In this paper, the above conclusion that RBF kernel in SVM classifier is used to classify the expressions.

In [11], local blocks were extracted from cropped images, and the Gabor filter was applied on each block. These responses were encoded and fed to the classifier.

In [26], Gabor features and LBP features were fused together, and the SVM-RBF classifier was used.

Zhao et al. [35] use a fusion of spatiotemporal LBP motion feature and multi-orientation Gabor results for classification on the CK+ [18] database.

Gabor filtered 3D LBP variants comprising six intersection points and three mean orthogonal planes were used in [3].

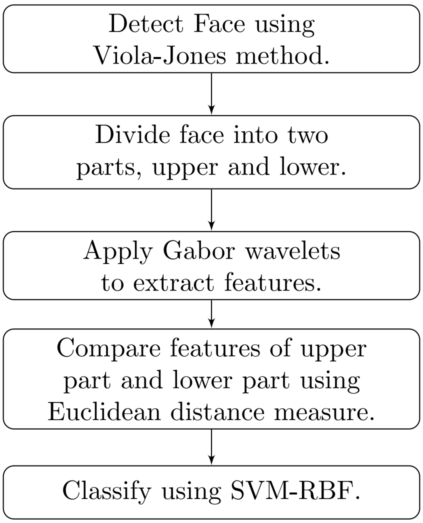
It is observed from Table 1 that the expressions surprise, happy, and disgust have better classification performance, whereas anger and fear suffer most of the time.

3 Methodology

In this work, the Viola–Jones method [30] is used to detect faces. The detected face is extracted using image cropping. The cropped image is resized to 240 × 320 pixels. The images representing neutral and peak pose are selected. Then these images are divided into two, upper, and lower halves. The Gabor filter is applied to the above sub-images.

The Euclidean distance is computed between the upper and lower halves, which becomes the extracted feature vector. The SVM classifier is trained using this feature vector. This workflow is represented in Fig. 3.

Fig. 3 Methodology used in this paper



3.1 Face Detection

The facial expression starts from the neutral expression and evolves to the peak expression. Videos with an only frontal view of the face are considered here. The frames containing the neutral expression and peak expression are extracted for further processing. From each frame, the face is detected using the Viola–Jones object detection framework that detects faces approximately 15 times faster than any previous approach [30]. Viola–Jones method is based on Haar wavelets and the AdaBoost learning algorithm.

3.2 Preprocessing

The detected face is represented as a normalized grayscale image and cropped to a rectangle of size 240 $\times$ 320 pixels. An elliptical mask is applied to the cropped face as shown in Fig. 4.

The masked faces are then partitioned. The partitioning is performed horizontally along the elliptical x-axis into upper face region and lower face region as shown in Fig. 5.

3.3 Gabor Wavelet Feature Representation

To represent the facial expressions, the Gabor wavelet method is used, which is used for textural analysis [31].

To compute the Gabor feature vector, convolve the input image and Gabor filter bank [17, 25]. The Gabor filter bank acts as a bandpass filter for spatial frequency signals.

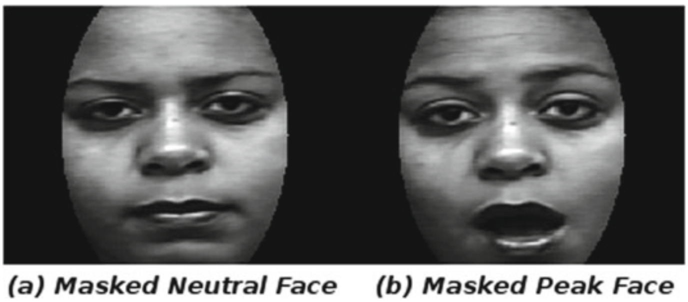


Fig. 4 Masked faces of two expressions: neutral and surprise

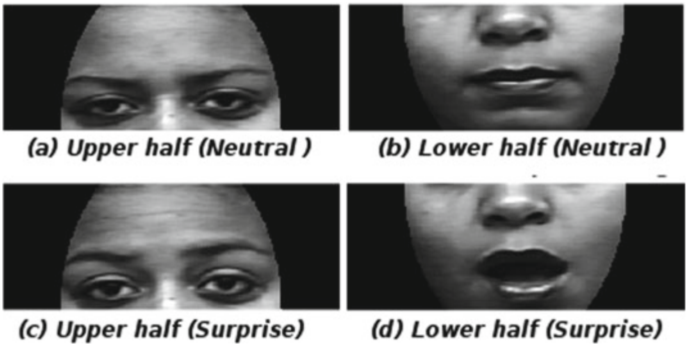


Fig. 5 Upper half and lower half of the masked face of two expressions: neutral and surprise

The 2-D Gabor filter $G(x, y, f, \theta)$ is defined as modulation of a sinusoidal wave, by a Gaussian kernel as in Eq. 1 [6].

$$G(x, y, f, \theta) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{\frac{(x_1^2+y_1^2)}{2\sigma^2}} \exp^{(2\pi f x_1)} \tag{1}$$

where

$$x_1 = x \cos\theta + y \sin\theta$$
$$y_1 = -x \sin\theta + y \cos\theta$$

and

f is the central frequency of the sine wave,
θ is the Gabor filter orientation,
σ is the standard deviation of the Gaussian envelope.

Facial expression cropped, grayscale image $I(x, y)$ is convolved with the Gabor filter $G(x, y, f, \theta)$ to obtain the feature vector as in Eq. 2 [6].

$$G_{u,v}(x, y) = I(x, y) * G(x, y, f, \theta) \tag{2}$$

The convolution output $G_{u,v}(x, y)$ is a complex number that can be decomposed into magnitude and phase components.

The magnitude component of the convolution output can be obtained as in Eq. 3 [6].

$$M_{u,v}(x, y) = \mathrm{Re}[G_{u,v}(x, y)] \tag{3}$$

The phase component of the convolution output can be computed as in Eq. 4 [6]. The phase components of Gabor array for five wavelengths and eight orientations are shown in Fig. 6.

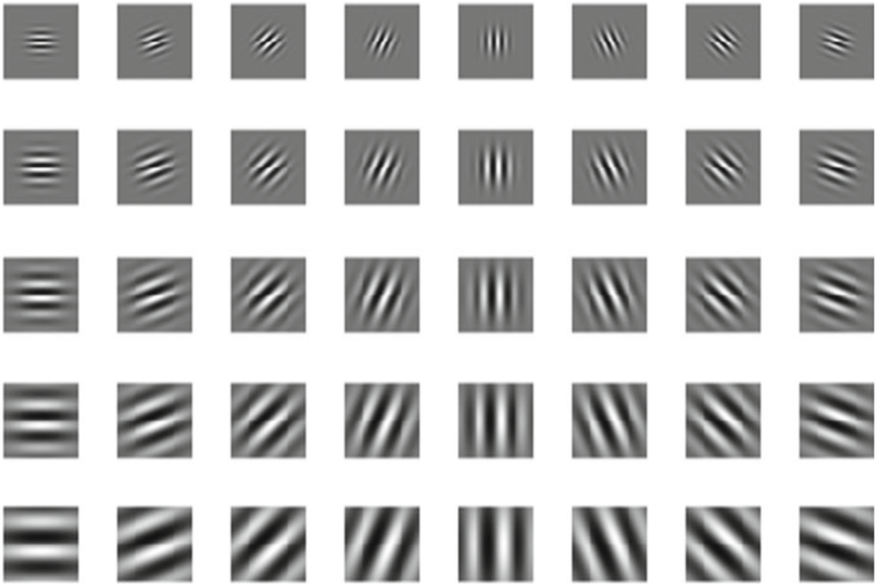


Fig. 6 Phase components of Gabor array for five wavelengths and eight orientations

$$P_{u,v}(x, y) = \mathrm{Im}[G_{u,v}(x, y)] \tag{4}$$

3.4 Support Vector Machine (SVM)

Support vector machine, the supervised machine learning algorithm, gives high classification accuracy for small data set. SVM shows good generalization on data set that is difficult to separate linearly. This paper is focusing on the feature extraction, hence general SVM with RBF kernel is used as the classifier.

4 Results and Analysis

This paper uses the Cohn-Kanade (CK) database [14] for conducting the experiment. In the first experiment, the Gabor filter bank with five wavelengths and eight orientations, in the complex form as seen in most of the literature, is applied (Fig. 7).

The extracted features are classified using SVM radial basis kernel (RBF), after optimizing the hyper-parameter C value, for higher accurate results. The data are partitioned in the 70:30 ratio to train and test the classifier. The experiments are conducted using 30% test data from the CK database. Figure 8 shows the confusion matrix.

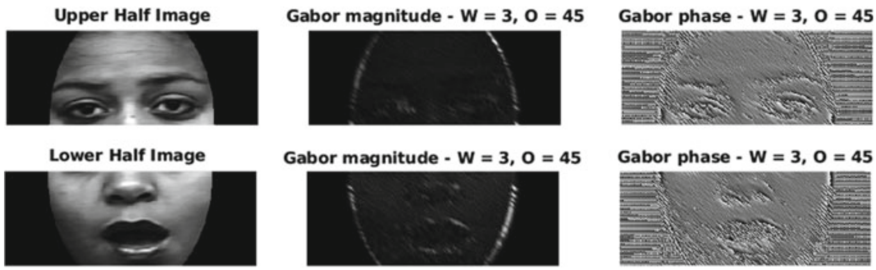


Fig. 7 Masked halves after convolving with Gabor filter with wavelength = 3 and orientation = 45

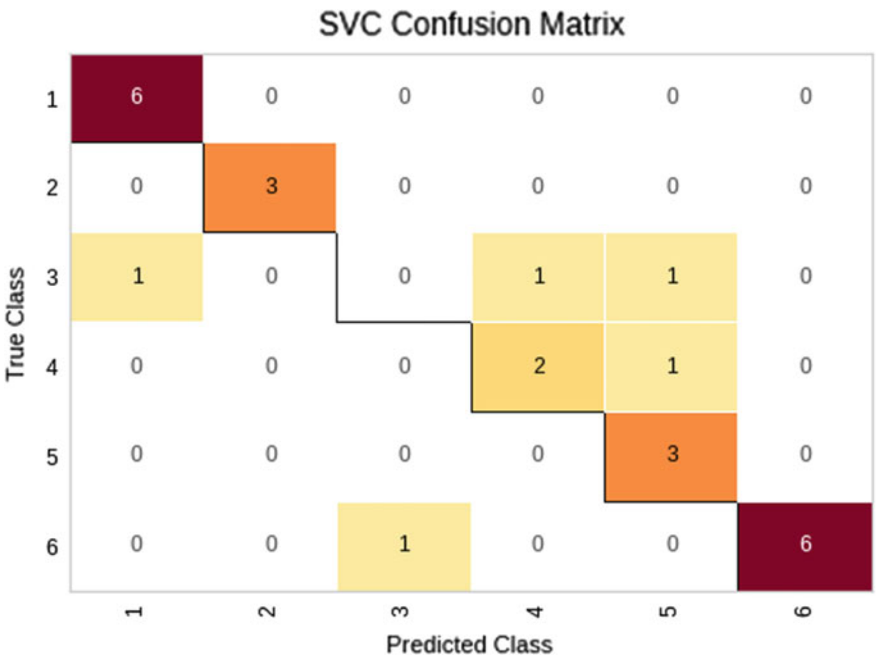


Fig. 8 Confusion matrix for SVM-RBF classification using Gabor filter of complex form, five wavelengths, and eight orientations

In the second experiment, the magnitude components of the Gabor bank with five wavelengths and eight orientations are convoluted with the face images to extract the real features. The above images are as shown in Fig. 7.

The feature vectors obtained after convolving with the magnitude component of Gabor filter bank are classified using SVM using the radial basis kernel (RBF). The confusion matrix of the above classification is shown in Fig. 9.

In the third experiment, the phase components of the Gabor bank with five wavelengths and eight orientations are convolved with the face images to extract the features. The feature vectors obtained after convolving with the phase component

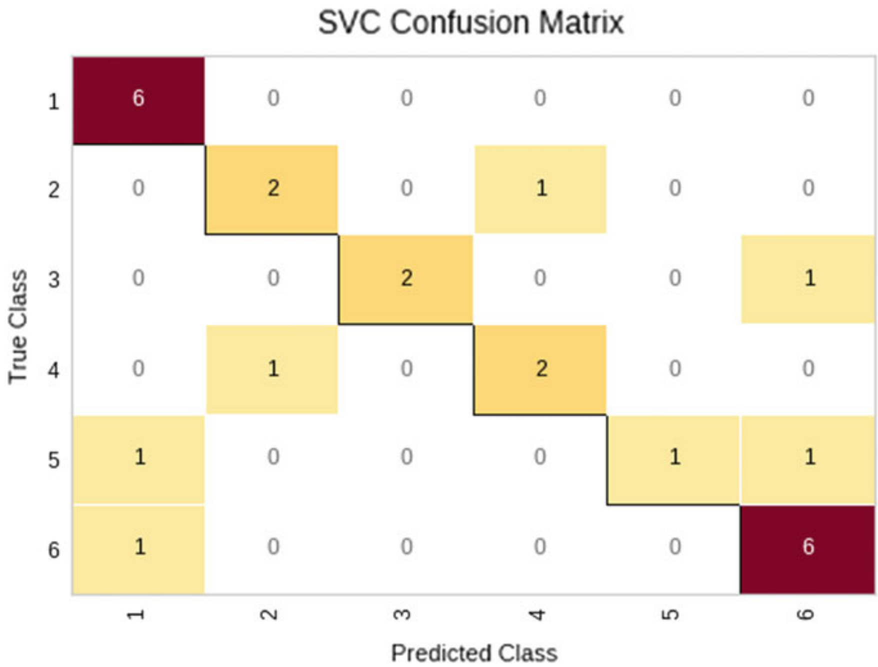


Fig. 9 Confusion matrix for SVM-RBF classification using magnitude component of Gabor filter, five wavelengths, and eight orientations

of the Gabor filter bank are classified using SVM-RBF. The confusion matrix of the above classification is shown in Fig. 10.

The classification results are shown in Table 2. For five wavelengths and eight orientations, feature vectors extracted after using Gabor filter bank in the phase component gives 88.0% accuracy, whereas the accuracy of 80.0 and 76.0% is obtained when the magnitude components of Gabor filter and the complex components are used, respectively.

The locally estimated Gabor wavelet coefficients are robust to illumination change and affine transformations. It is also observed that using the Gabor filter bank phase components increases the classification accuracy considerably.

The phase components are antisymmetric, giving larger responses for edges, which in turn produces better results for facial texture information. The low-level features like pixel intensity and gradients are better captured when using phase components. In other words, the edges and facial wrinkles have more pixel intensity which gives significant results when using phase components of the Gabor filter bank.

Similar to the above, three more experiments are conducted. The complex form, which includes the magnitude and phase parts of the Gabor filter bank using eight wavelengths and thirteen orientations, is convolved with the face images to extract the features. Table 3 lists the results.

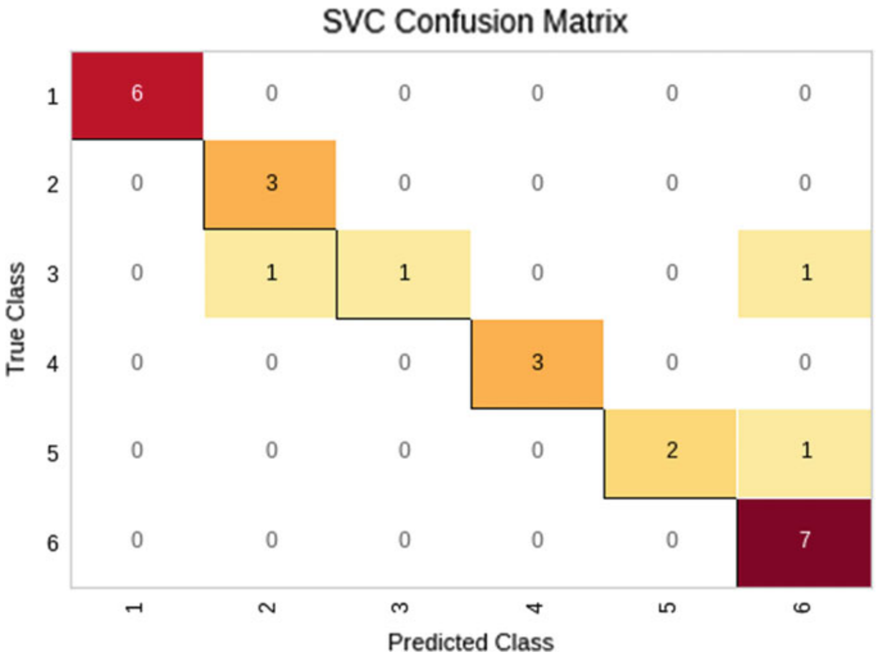


Fig. 10 Confusion matrix for SVM-RBF classification using phase component of Gabor filter, five wavelengths, and eight orientations

Table 2 Experimental results on CK data set using five wavelengths and eight orientations

Gabor component	Filter bank with five wavelengths and eight orientations		
	Accuracy (%)	Precision	Recall
Complex components	80.0	0.69	0.75
Magnitude components	76.0	0.81	0.70
Phase components	88.0	0.92	0.83

Table 3 Experimental results on CK data set using eight wavelengths and thirteen orientations

Gabor component	Filter bank with eight wavelengths and thirteen orientations		
	Accuracy (%)	Precision	Recall
Complex components	79.16	0.80	0.72
Magnitude components	87.5	0.92	0.83
Phase components	91.0	0.94	0.89

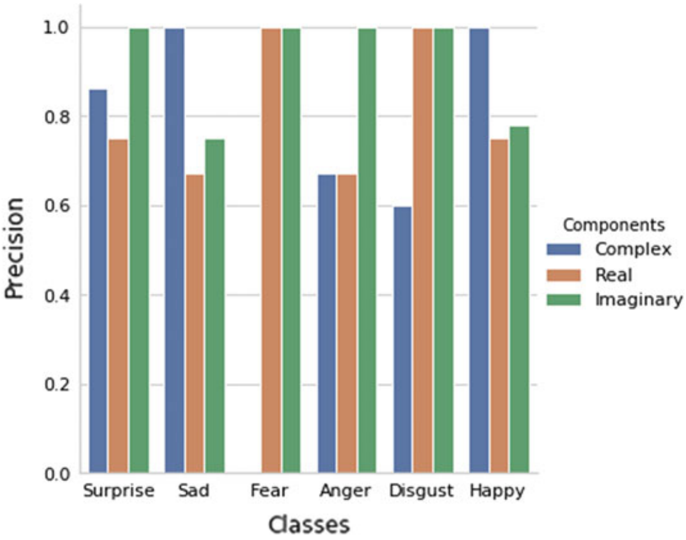


Fig. 11 Precision values for six expressions, using five wavelengths and eight orientations filter bank

Another observation is that there is a significant increase in the accuracy values when eight wavelengths and thirteen orientations are used for feature extraction as shown in Table 3. This increase can be due to extracting more features at a finer granularity when the size of the filter bank is increased.

A comparison of precision and recall for each expression when using five wavelengths and eight orientations is given in Figs. 11 and 12.

From precision Fig. 11 and recall Fig. 12 graphs, the conclusions of Table 4 can be drawn.

It is observed from Table 4 that when using Gabor filter bank, expressions like fear and anger are suffering the most, while 100% accuracy is obtained for surprise expression. Further, it can be observed that when using phase components, more expressions, surprise, sadness, anger, and happiness can be detected at 100% accuracy.

5 Conclusion

In this paper, we explore the impact of the Gabor filter bank for FER. The SVM classifier was used on the Cohn-Kanade data set. The comparative evaluation of three variants of Gabor filter—feature vectors obtained after convolving with Gabor filter bank, feature vectors obtained after convolving with the real component of Gabor

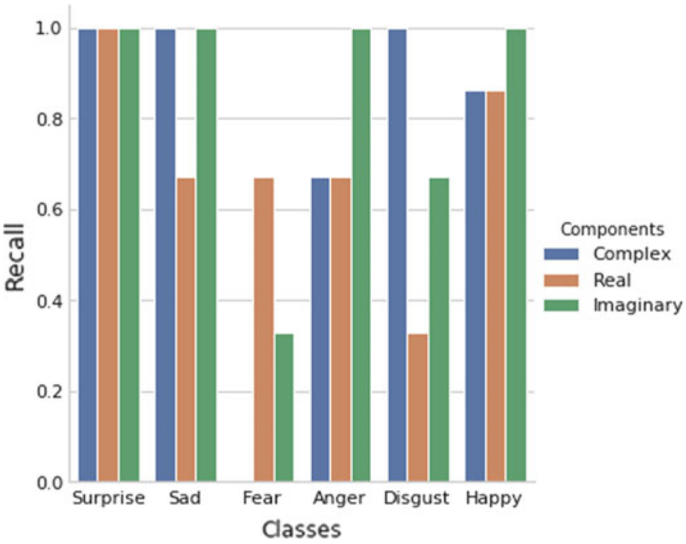


Fig. 12 Recall values for six expressions, using five wavelengths and eight orientations filter bank

Table 4 Experimental results on CK data set

Gabor component	Expression best captured	Expression worst captured
Complex components	Surprise, sad, disgust	Anger, happy, fear
Real components	Surprise	Anger, sad, happy, fear, disgust
Imaginary components	Surprise, anger, sad, happy	Fear, disgust

filter bank, and feature vectors obtained after convolving with imaginary Gabor filter bank—was performed.

When only the phase component of the Gabor filter bank of five wavelengths and eight orientations was used, the performance improved by 8% for the CK data set. Similarly by using phase component of the Gabor filter bank of eight wavelengths and thirteen orientations, the performance improved by 11% for the CK data set. It is hence established in this paper that for the domain of FER, the use of the phase component of the Gabor filter improves the accuracy results of facial recognition when used in conjunction with the SVM classifier. By excluding the magnitude component of the Gabor filter, though we have improved feature extraction, there are unaddressed challenges. There are some emotions for which the classifier resulted in a high false-positive rate. Excluding magnitude, components did not result in any change in the classification of these outliers.

References

1. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 12:2037–2041
2. Buciu I, Pitas I et al (2003) ICA and Gabor representation for facial expression recognition. In: 2003 international conference on image processing, 2003. ICIP 2003. Proceedings, vol 2. IEEE, p II-855
3. Chengeta K, Viriri S (2018) Video sequencing based facial expression detection with 3d local binary pattern variants. In: 5th international conference on computer science, engineering and information technology
4. Dailey MN, Cottrell GW (1999) PCA = Gabor for expression recognition. UCSD CSE TR CS-629
5. Daugman JG (1988) Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression. IEEE Trans Acoust Speech Signal Process 36(7):1169–1179
6. Derpanis KG (2007) Gabor filters
7. Ekman P, Friesen WV (1976) Measuring facial movement. Environ Psychol Nonverbal Behav 1(1):56–75
8. Fan X, Tjahjadi T (2015) A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences. Pattern Recogn 48(11):3407–3416
9. Fasel B, Luettin J (2003) Automatic facial expression analysis: a survey. Pattern Recogn 36(1):259–275
10. Fellenz W, Taylor J, Tsapatsoulis N, Kollias S (1999) Comparing template-based, feature-based and supervised classification of facial expressions from static images. Comput Intell Appl 19(9):9
11. Gu W, Xiang C, Venkatesh Y, Huang D, Lin H (2012) Facial expression recognition using radial encoding of local Gabor features and classifier synthesis. Pattern Recogn 45(1):80–91
12. Hsieh CC, Jiang MK (2011) A facial expression classification system based on active shape model and support vector machine. In: 2011 international symposium on computer science and society. IEEE, pp 311–314
13. Jiang B, Yang GS, Zhang HL (2008) Comparative study of dimension reduction and recognition algorithms of DCT and 2DPCA. In: 2008 international conference on machine learning and cybernetics, vol 1. IEEE, pp 407–410
14. Kanade T, Cohn JF, Tian Y (2000) Comprehensive database for facial expression analysis. In: Proceedings fourth IEEE international conference on automatic face and gesture recognition (Cat. No. PR00580). IEEE, pp 46–53
15. Kotsia I, Buciu I, Pitas I (2008) An analysis of facial expression recognition under partial facial image occlusion. Image Vis Comput 26(7):1052–1067
16. Lee TS (1996) Image representation using 2d Gabor wavelets. IEEE Trans Pattern Anal Mach Intell 10:959–971
17. Liu C, Wechsler H (2002) Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans Image Process 11(4):467–476
18. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: 2010 IEEE computer society conference on computer vision and pattern recognition—workshops. IEEE, pp 94–101
19. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with Gabor wavelets. In: Third IEEE international conference on automatic face and gesture recognition, 1998. Proceedings. IEEE, pp 200–205
20. Mehrabian A (2017) Nonverbal communication. Routledge, New York
21. Oppenheim AV, Lim JS (1981) The importance of phase in signals. Proc IEEE 69(5):529–541
22. Pietikäinen M, Hadid A (2005) Texture features in facial image analysis. In: Advances in biometric person authentication. Springer, Cham, pp 1–8

23. Shan C, Gong S, McOwan PW (2005) Robust facial expression recognition using local binary patterns. In: IEEE international conference on image processing, 2005. ICIP 2005, vol 2. IEEE, p II-370
24. Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: a comprehensive study. Image Vis Comput 27(6):803–816
25. Struc V, Pavesic N (2010) From Gabor magnitude to Gabor phase features: tackling the problem of face recognition under severe illumination changes. In: Face recognition. IntechOpen
26. Sun Y, Yu J (2017) Facial expression recognition by fusing Gabor and local binary pattern features. In: International conference on multimedia modeling. Springer, pp 209–220
27. Tian YI, Kanade T, Cohn JF (2001) Recognizing action units for facial expression analysis. IEEE Trans Pattern Anal Mach Intell 23(2):97–115
28. Tian YL, Kanade T, Cohn JF (2005) Facial expression analysis. In: Handbook of face recognition. Springer, Cham, pp 247–275
29. Valstar MF, Pantic M (2007) Combined support vector machines and hidden Markov models for modeling facial action temporal dynamics. In: International workshop on human-computer interaction. Springer, pp 118–127
30. Viola P, Jones M et al (2001) Rapid object detection using a boosted cascade of simple features. CVPR 1(1):511–518
31. Viswanathan DM, Idicula SM (2015) SVM based recognition of facial expressions used in indian sign language. Int J Image Process (IJIP) 9(1):32
32. Wu T, Bartlett MS, Movellan JR (2010) Facial expression recognition using Gabor motion energy filters. In: 2010 IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW). IEEE, pp 42–47
33. Zhang Z, Lyons M, Schuster M, Akamatsu S (1998) Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. In: Proceedings third IEEE international conference on automatic face and gesture recognition. IEEE, pp 454–459
34. Zhao G, Pietikainen M (2007) Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans Pattern Anal Mach Intell 29(6):915–928
35. Zhao L, Wang Z, Zhang G (2017) Facial expression recognition from video sequences based on spatial-temporal motion local binary pattern and Gabor multiorientation fusion histogram. Math Probl Eng 2017

Voting-Based Ensemble of Unsupervised Outlier Detectors

Roy Thomas and J. E. Judith

Abstract Datasets may contain small sets of data objects whose characteristics are not in accordance with the mainstream characteristics of the data objects in a dataset. These data objects, which are not noise, may contain valuable information and are called outliers. Outlier detection is a topic of research in many fields like detecting malwares in cyber security, finding fake financial transactions, identifying defects in industrial products, detecting abnormality in health data, etc. Researchers have developed several application methods for detecting outliers and a few generic methods. These methods can be grouped into unsupervised methods, supervised methods and semi-supervised methods based on the readiness of class labels. We, in this paper, present the performance of three outlier detection algorithms using the realworld datasets. The algorithms used are one-class SVM, elliptic envelope and local outlier factor. In order to improve the performance, all these algorithms were selected and ensemble based on voting mechanism. The influence of dimensionality reduction on the proposed ensemble method has also been studied. Experiments using publicly available datasets show that the proposed technique outperforms individual outlier detectors.

Keywords Data mining · Dimensionality reduction · Ensemble · Outlier detection · Unsupervised

1 Introduction

Data mining is a field in computer science that is intended to find relevant and crucial information from large collections of datasets most of which are unstructured. Considering the volume, the data in the world is increasing tremendously at every moment and it becomes very difficult to extract the desired information from the datasets. Datasets are identified by the mainstream characteristics of the dataset

R. Thomas (✉) · J. E. Judith
Noorul Islam Centre for Higher Education, Kumaracoil, India
e-mail: roygptc@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_42

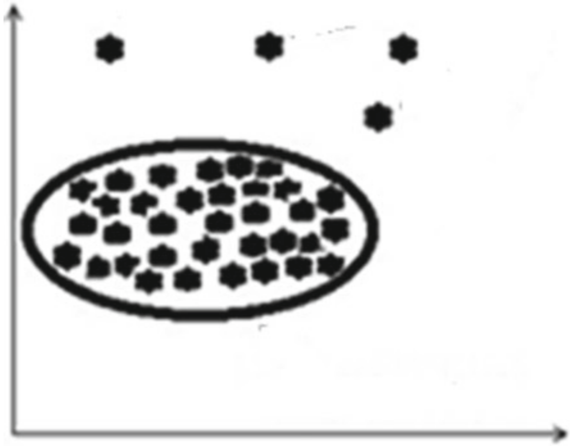


Fig. 1 Objects inside the ellipse are normal data and outside the ellipse are outliers

attributes, even though these datasets may contain data objects whose characteristics are very much different from the characteristics of the majority data objects in the dataset. These hidden data objects may contain valuable information and hence cannot be treated as noise. They are called outliers as their properties are not conformed to the balance data objects in the dataset. Outlier detection plays a vital role in many branches of study such as statistics and data mining. Detection of hidden and important information from large datasets has been a research field with diverse application areas for the past few decades.

Detection of hidden and important information has become an essential research field in many branches of statistics and computer science such as data mining, machine learning, information theory and spectral theory. Outlier detection is aimed at finding infrequent data objects containing valuable information and is not conformable with the majority of data objects in the dataset. Figure 1 shows the difference between normal data objects and outliers. Most of the normal objects are clustered together as they are having similar properties. An outlier is a data object that is distant from most of the remaining data objects in the dataset and is not clustered with the majority data objects. Outlier detection is defined as the discovery of data objects that are dissimilar, distant or uneven with respect to the majority of data objects in the dataset. Outlier detection provides critical information in a variety of domains such as military and aeronautical fields.

The applications of outlier detection are essential in many areas. Outlier detection methods can be applied to find fake financial transactions, unauthorized insurance claim, unauthorized computer network access, to find fake credit card transactions in banking, abnormality in medical and public health data, damages in industrial products, inadequacies in image processing, conflicts in Web applications, irregularities in robot behavior, variances in astronomical data, discrepancies in census data, etc.

Outliers can be categorized into point outliers, collective outliers and contextual outliers depending on the nature of outliers. An individual data object that differs considerably from other data objects in its set is called a point outlier or global outlier. For example, the features of the credit card transactions like the type of items purchased, their quantity, amount spent, etc., by the fraud shall be very much different from the normal purchase features of the credit card transaction by the authenticated card owner. The credit card transactions that do not follow the regular pattern can be treated as an outlier. Collective outlier is a subgroup of data objects which together differs from the entire dataset. The distinct data objects in a collective outlier are not outliers individually. For example, heavy traffic for a limited time is a normal incidence in a metropolitan city. However, if it lasts for two or three days, it becomes a collective outlier. Contextual outlier is a data object that differs considerably from the dataset based on a selected situation and is a normal data object with reference to some other situation. In contextual outliers, the data objects are defined using two attributes—a context attribute and a behavior attribute. Rain in summer is a contextual outlier. Here, rain is the behavioral attribute and summer is the contextual attribute.

2 Related Work

Outlier detection methods have been a research subject in statistics and data mining from decades before, and now, it is extended to many fields such as deep learning and information theory. Different techniques have been developed to find outliers in various application domains. Hodge and Austin [4] observed outlier detection approaches as extracted mainly from three fields of computing—statistical, neural networks and machine learning. The classification outlier detection approaches done by Hodge and Austin into three categories—Type-I, Type-II and Type-III—are analogous to unsupervised, supervised and semi-supervised approaches, respectively. Patcha and Park [8] accumulated the methods into statistical-based, classification-based, clustering-based and nearest neighbor-based groups. Chandola et al. [1] added two more groups of outlier detection methods to these four groups called spectral methods and information theoretic methods. Han et al. presented methods for detecting collective outliers and contextual outliers in addition to global outliers. They also presented the methods for finding outliers in subspaces and high-dimensional data.

3 Outlier Detection Methods

The categorization of outlier detection methods can be done in different ways. One of these categorizations is based on the availability of class labels. Depending on the availability of class labels, outlier detection methods are grouped into supervised, semi-supervised and unsupervised methods [2].

3.1 Supervised Outlier Detection

Supervised approaches are suitable when the data objects are labeled and can be separated into normal objects and outliers based on the class label. This is a classification-based technique where the labels are used to make the model of the normal class or outlier class. Comparing to the existence of normal objects, the occurrences of outliers are very infrequent, and classification techniques which are able to handle highly imbalanced sets are needed to separate the data objects into normal objects, and outliers are needed in a supervised scenario.

3.2 Unsupervised Outlier Detection

Supervised outlier detection methods cannot be used when the datasets are not labeled. Unsupervised methods are used to find outliers in an unlabeled dataset by assigning each object an outlier score which indicates its degree of outlierness. These scores indicate how much different is a data object from other data objects in the dataset. Unsupervised methods do not use a training dataset, and clustering methods are normally used to find outliers, which are data objects that belong to neither of the clusters or to sparse clusters.

3.3 Semi-supervised Outlier Detection

Semi-supervised methods are used in situations when labels are available for only a small part of normal objects or outliers [5]. The labeled objects, either normal or outlier, are used in the training phase to obtain a model of the normal objects or outliers. This is similar to a binary classification problem in which only one set of objects are labeled. This label information is used to divide the data objects into two sets in which one set contains all the normal objects and the other set contains the outliers.

Researchers have developed a number of algorithms to find outliers using supervised, semi-supervised or unsupervised techniques. Some of them are given in the following sections.

3.4 Local Outlier Factor

Local outlier factor (LOF) algorithm is an efficient method to find outlier detection in moderately high-dimensional datasets [9]. The algorithm computes the outlier score called local outlier factor which reflects outlier degree of the object. It computes the deviation in the local density of an object with respect to its neighbors. Objects whose local density is much lower than its neighbors are treated as outliers. Usually, the LOF score of an object is calculated by comparing its local density with the average local density of its k-nearest neighbors. The algorithm needs the value k for k-nearest neighbors and the threshold value for the outlier score as input parameters.

3.5 Elliptic Envelope

Outlier detection using elliptic envelope assumes that the normal data objects form a known distribution, like Gaussian distribution [6]. Data objects are classified into normal data objects and outliers based on this assumption in which the normal objects occur in high probability region of the distribution and outliers occur in low probability regions or do not follow this distribution. Elliptic envelope technique fits a robust covariance estimate to the data objects and fits an ellipse to the central data points. After finding the central data point, a distance measure is used to detect outlier degree of the data object. The method needs the percentage of outliers in the dataset as hyper parameter, which is not easy to estimate.

3.6 One-Class Support Vector Machine

One-class support vector machine (SVM) is a classification method when the dataset contains only one class [7]. In this, the support vector machine model is trained to gather the properties of normal data objects only. The goal is to predict whether a data object belong to this class or not. This method can be used to determine the presence of outliers in the dataset. A data object is considered as normal object or outlier depending on whether or not it belongs to the class.

4 Proposed Method

The proposed method is a voting-based ensemble of three existing methods. The technique involves the following steps.

Step 1. Obtain the predicted outlier values of the datasets using three different algorithms—one-class SVM, elliptic envelope and local outlier factor.
Step 2. Outlier is detected by aggregating the results of the individual predictors in the ensemble through voting mechanism.
Step 3. Detect the outlier after reducing the dimensionality of the datasets using principle component analysis and compare the results with the results before reducing the dimensionality.

Table 1 Characteristics of datasets used for experiment

Dataset	#instances	#attributes	Attribute type	#classes	#normal	#outlier
Iris	150	4	Real	3	50	5
Breast cancer	569	32	Real	2	300	30

5 Datasets

Datasets from the UCI machine learning repository are used for our experiments. These datasets are publicly available for experiments. The datasets used are 'iris' dataset and 'breast cancer' dataset. The reason for taking these two datasets is one dataset contains only four attributes, whereas the other contains 32 attributes. This helps to find out the influence of the dimensionality reduction in high-dimensional as well as low-dimensional datasets. The description of the datasets is given in Table 1.

6 Experiments

Experiments using publicly available datasets were conducted in an Intel core i3-based laptop using Python. Separate experiments were conducted for each outlier detection algorithm using different datasets. Only two classes of data objects were used for the experiment. Samples for experiments were taken from the dataset randomly in which about 90% of data belongs to the normal class and the remaining small set as outliers. The algorithms used for detecting outliers were one-class SVM, elliptic envelope and local outlier factor. The results obtained from the individual algorithms were compared with the actual dataset and evaluated the performance of each algorithm.

6.1 Performance of Individual Detectors

Figure 2 shows the scatter plot of the 'iris' data sample used for detecting outliers using different algorithms. The same sample is used for all t three algorithms and also for our proposed ensemble detector. The results obtained from the individual detectors using the 'iris' dataset are shown in Figs. 3, 4 and 5. The evaluation measures used for comparing the performance of individual algorithms are precision, recall and F1-score. The values of these measures obtained from the individual detectors using iris dataset and breast cancer dataset are shown in Tables 2 and 3, respectively.

All the predictors were able to detect the outliers in the iris dataset, but they wrongly classified some of the normal objects as outliers. Elliptic envelope predictor showed a better precision and F1-score. However, the algorithm took more time to complete. Result analysis of different algorithms shows that the performance of

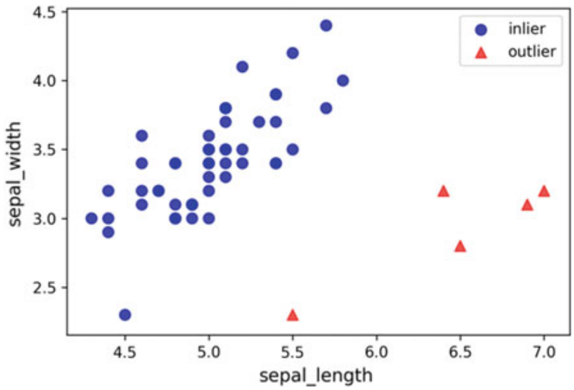


Fig. 2 Scatter plot of iris dataset

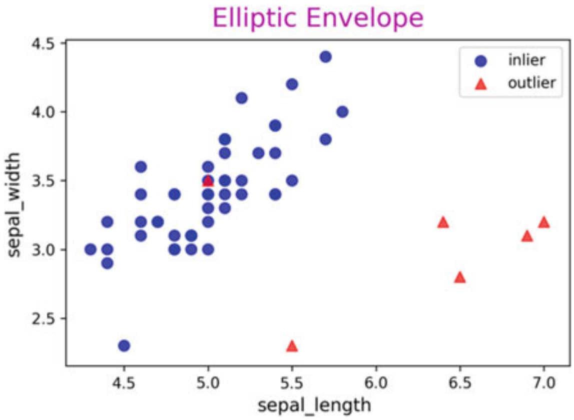


Fig. 3 Scatter plot of the results from elliptic envelope

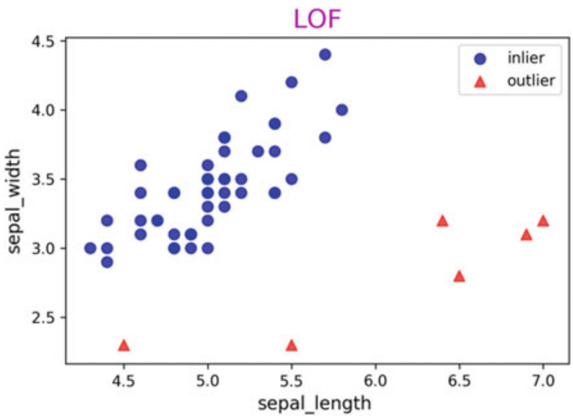


Fig. 4 Scatter plot of the results from LOF

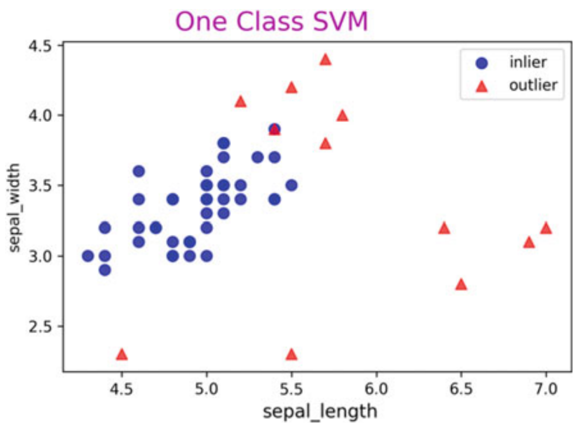


Fig. 5 Scatter plot of the results from one-class SVM

Table 2 Evaluation result of individual predictors using iris dataset

Predictor	Precision	Recall	F1-score	Duration
Elliptic envelope	0.833	1.000	0.909	0.065
One-class SVM	0.417	1.000	0.588	0.003
LOF	0.833	1.000	0.909	0.003

Table 3 Evaluation result of individual predictors using breast cancer dataset

Predictor	Precision	Recall	F1-score	Duration
Elliptic envelope	0.576	0.633	0.603	0.112
One-class SVM	0.485	0.533	0.508	0.005
LOF	0.485	0.537	0.507	0.006

an algorithm depends on the nature of the dataset. Hence, the performance can be improved by combining the results of various algorithms together. The following subsection shows the method used for combining the outputs of the predictors and also the results obtained from the proposed method.

6.2 Performance of Proposed Ensemble Detector

In the proposed ensemble detector, the outlier values are calculated by aggregating the results of the individual predictors in the ensemble through voting mechanism. Here, an object is considered as outlier if majority of the individual algorithms predicted that object as outlier; otherwise, it is considered as a normal object. The results obtained by the proposed ensemble detector are shown in Table 4. The precision

Table 4 Evaluation result of the proposed ensemble detector

Dataset	Precision	Recall	F1-score	Duration
Iris	1.000	1.000	1.000	0.071
Breast cancer	0.739	0.567	0.642	0.123

and F1-score of the proposed ensemble detector are higher than all other detectors. The proposed ensemble detector is a better method as it outperforms the individual detectors.

6.3 Effect of Dimensionality Reduction

Principal component analysis (PCA) is a popular method for reducing the dimensions of datasets [3]. Dimensionality reduction plays an important role in data mining as it finds the relationship among the attributes and helps in data visualization. Here, the performance of the outlier detection algorithms is evaluated after applying dimensionality reduction in the dataset. All three algorithms as well as the proposed ensemble method are used for analyzing the influence of dimensionality reduction in outlier detection techniques. Both the datasets, iris and breast cancer are used for this experiment. Experiments were performed by reducing to different dimensions, and the results obtained after reducing the dimensions are shown in Tables 5 and 6. The effect is more clear from the breast cancer dataset as the original dataset contained 32 attributes and reduced to much lower dimensions. Dimensionality reduction decreased the execution time of the outlier detection algorithms for both the datasets. The performance is improved for the breast cancer dataset, while there is little effect on the performance for iris dataset.

Table 5 Effect of dimensionality reduction on predictors (iris dataset)

Dimensionality	Predictor	Precision	Recall	F1-score	Duration
Dimension reduced from 4 to 3	Elliptic envelope	0.833	1.000	0.909	0.037
	One-class SVM	0.455	1.000	0.625	0.001
	LOF	0.833	1.000	0.909	0.001
	Proposed ensemble	0.833	1.000	0.909	0.039
Dimension reduced from 4 to 2	Elliptic envelope	0.833	1.000	0.909	0.036
	One-class SVM	0.455	1.000	0.625	0.001
	LOF	0.833	1.000	0.909	0.001
	Proposed ensemble	0.833	1.000	0.909	0.038

Table 6 Effect of dimensionality reduction on predictors (breast cancer dataset)

Dimensionality Reduction	Predictor	Precision	Recall	F1-score	Duration
Dimension reduced from 32 to 16	Elliptic envelope	0.636	0.700	0.667	0.074
	One-class SVM	0.471	0.533	0.500	0.003
	LOF	0.485	0.533	0.508	0.004
	Proposed ensemble	0.800	0.667	0.727	0.081
Dimension reduced from 32 to 10	Elliptic envelope	0.697	0.767	0.730	0.054
	One-class SVM	0.500	0.533	0.516	0.002
	LOF	0.485	0.533	0.508	0.003
	Proposed ensemble	0.815	0.733	0.772	0.059

7 Conclusion

In this paper, we studied three different techniques for detecting outliers and analyzed the performance of these detectors individually and also on our proposed ensemble technique using two publicly available datasets, namely iris and breast cancer. We also studied the influence of dimensionality reduction in detecting outliers using our proposed ensemble detector and also with the individual outlier detectors. We have found that our proposed ensemble-based detector outperforms individual outlier detectors, and we were able to achieve higher precision and F1-score than the scores of individual detectors. After applying dimensionality reduction, it is found that even though there is not much difference in the evaluation score for low-dimensional dataset, there is a reduction in execution time of the algorithms. The execution time as well as the performance of the outlier detection algorithms has improved much better for higher-dimensional data as a result of dimensionality reduction. The observations showed that dimensionality reduction has an important role in high-dimensional data, and it has little effect on low dimensional data.

References

1. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv 41(3):1–58
2. Han J, Kamber M, Pei J (2012) Data mining: concepts and techniques. Morgan Kaufmann, Massachusetts, US
3. Hauberg S, Feragen A, Enficiaud R, Black MJ (2016) Scalable robust principal component analysis using Grassmann averages. IEEE Trans Pattern Anal Mach Intell 38(11):2298–2311. https://doi.org/10.1109/tpami.2015.2511743
4. HodgeVJ Austin J (2004) A survey of outlier detection methodologies. Artif Intel Rev 22(2):85–126
5. Ienco D, Pensa RG, Meo R (2017) A semi-supervised approach to the detection and characterization of outliers in categorical data. IEEE Trans Neural Networks Learn Syst 28(5):1017–1029. https://doi.org/10.1109/tnnls.2016.2526063

6. Kumar R, Kundu PP, Phoha VV (2018) Continuous authentication using one-class classifiers and their fusion. In: IEEE 4th international conference on identity, security, and behavior analysis (ISBA). https://doi.org/10.1109/isba.2018.8311467
7. Mũnoz-Marí J, Bovolo F, Gómez-Chova L, Bruzzone L, Camp-Valls G (2010) Semisupervised one-class support vector machines for classification of remote sensing data. IEEE Trans Geosci Remote Sens 48(8):3188–3197. https://doi.org/10.1109/tgrs.2010.2045764
8. Patcha A, Park J-M (2007) An overview of anomaly detection techniques: existing solutions and latest technological trends. In: Computer networks
9. Radovanovic M, Nanopoulos A, Ivanovic M (2015) Reverse nearest neighbors in unsupervised distance-based outlier detection. IEEE Trans Knowl Data Eng 27(5):1369–1382. https://doi.org/10.1109/tkde.2014.2365790

Double-Band Coplanar Antenna for ISM and UWB Applications

Dhivya Raj and C. V. Anil Kumar

Abstract A compact 25 × 30 × 1.6 mm^3 coplanar antenna suitable to operate on both ultra-wideband (UWB) and industrial, scientific and medical bands (ISM) is presented. A coplanar waveguide (CPW)-fed antenna for 3.1–10.6 GHz (UWB) along with 2.4–2.48 GHz (ISM) will be useful for high-speed data transfer applications in embedded communication systems. The proposed antenna is fabricated on FR-4 epoxy substrate, and the experimental results are reported. Vital characteristics like radiation pattern, return loss and bandwidth have been investigated.

Keywords Coplanar antenna · Dual band · UWB · ISM · Bluetooth

1 Introduction

Frequency band of 3.1–10.6 GHz was allocated for commercial communication applications by the US Federal Communication Commission (FCC) on February 14, 2002. Since then antennas operating in ultra-wideband (UWB) spectrum have found wide use due to its low power and high data rate features [1]. The radiating element that can radiate/receive in both UWB and ISM bands is much more useful in the design of compact communication systems [2–4]. The industrial, scientific and medical bands (ISM) ranging from 2.4 to 2.485 GHz range are used widely in today's Bluetooth connectivity devices. At the same time, ultra-wideband (UWB) technology ranging from 3.1 to 10.6 GHz can provide even higher data transfer rates of the order of 100 Mbits/s [5–7].

D. Raj (✉)
Department of ECE, College of Engineering Chengannur, Chengannur, Kerala, India
e-mail: dhivyaraj@ceconline.edu

C. V. Anil Kumar
Department of ECE, College of Engineering Karunagappally, Karunagappally, Kerala, India
e-mail: cvanil@ceconline.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_43

Table 1 Literature survey

Multiband antenna	Antenna dimensions (H × W × T)	Wireless bands
Proposed antenna	30 × 25 × 1.6	UWB, Bluetooth
Antenna in Ref. [11]	30 × 25 × 1.6	UWB
Antenna in Ref. [8]	25 × 24 × 1.6	UWB, GSM, GPS, Bluetooth
Antenna in Ref. [9]	30 × 30 × 1.6	UWB, GSM, WCDMA, WLAN
Antenna in Ref. [10]	18 × 30 × 1.6	Bluetooth, UWB
Antenna in Ref. [12]	33 × 28 × 1.6	WLAN, WiMAX, HiperLAN (5.25 GHz)
Antenna in Ref. [13]	10 × 34.5 × 0.8	LTE/WWAN
Antenna in Ref. [14]	10 × 6 × 0.4	2.4/5 GHz-WLAN

2 Literature Survey

A wide survey on different multiband antennas was conducted and is summarized in Table 1.

The antenna in ref [11] was developed on FR-4 substrate (thickness = 1.6 mm, loss tangent = 0.02 and dielectric constant = 4.4). Various multiband printed antenna geometries have been studied with different ways to gain impedance matching and bandwidth enhancement in the literature for multiband applications [8–10, 12–14].

3 Methodology

The simulation analysis is completed with high-frequency structure simulation (HFSS) Ansoft software package.

4 Antenna Design

When upper part of ground plane is extended as a horizontal metallic strip from ground plane in the rectangular slot of UWB antenna in [11], a frequency of 2.2 GHz is obtained within the lower frequency range as in Fig. 1. When the horizontal strip from ground plane is once more extended and hung vertically down, the resonant

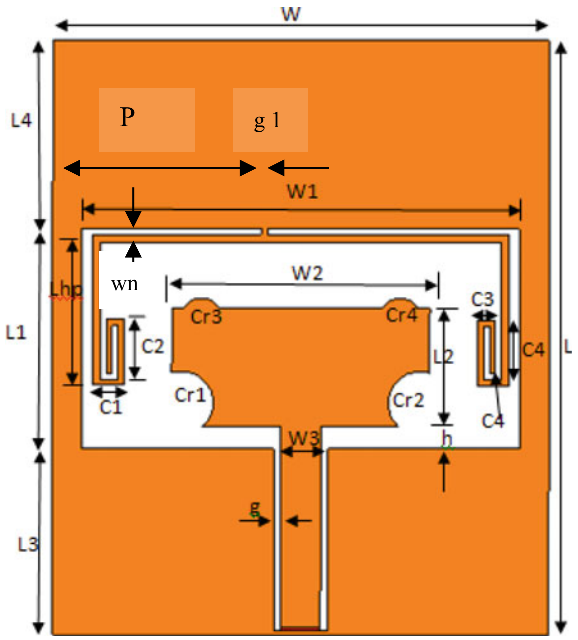


Fig. 1 Geometry of the proposed antenna

frequency shifted more lower to 1.8 GHz. The thickness of horizontal metallic strip in the slot from ground plane (wn) and thickness of hanging strips (hn) are 0.3 mm. The optimum value obtained for length of the hanging strip Lhp is 7.6 mm. The optimum point *P* at which ground plane is extended as a hanging metallic strip from ground plane into the rectangular slot is obtained as 10.5 mm of total width of antenna and is shown as *P* in Fig. 1. The width of the gap where the ground plane starts extending into the slot is $g_1 = 0.35$ mm. When hanging metallic strips were coiled at their ends, 2.45 GHz—ISM band was introduced along with UWB range of frequency, see Fig. 1.

The antenna in ref [11] as well as the one developed in this paper was developed on FR-4 substrate (thickness = 1.6 mm, loss tangent = 0.02 and dielectric constant = 4.4), see Fig. 1 for dimensions of the antenna.

Size of the antenna is 30 mm × 25 mm. Microstrip feed width is 2 mm for a characteristic impedance of 50 Ω. The optimization of antenna dimensions is done by parametric analysis. So obtained dimensional values are given in Table 2.

5 Simulation Results

Various simulation results are discussed under return loss, parametric analysis and radiation pattern in the following Sects. 5.1, 5.2 and 5.3.

Table 2 Optimized parameters of the proposed antenna (Unit: mm)

W	L	L_1	L_2	L_3	L_4	Lhp
25	30	11.1	06	9.4	9.5	7.6
W_1	W_2	W_3	$\mathrm{Cr}_1 = \mathrm{Cr}_2$	$\mathrm{Cr}_3 = \mathrm{Cr}_4$	g	g_1
22	13	02	1.6	01	0.35	0.5
h	P	Wn	hn	C_1	$C_2 = C_4$	C_3
1.1	10.5	0.3	0.3	1.2	2.7	0.6

5.1 Return Loss

At resonant frequency of 2.45 and 3.1–10.6 GHz, the return loss of ISM band is nearly as good as −15 dB and −17 dB, respectively, as shown in Fig. 2. It also exhibits a better impedance matching from 2.43 to 2.48 GHz and from 2.9 to 10.6 GHz.

The current distribution pattern at 2.45 and 7.5 GHz illustrates that the hanging metallic strip from ground plane into the rectangular slot and coil strips at its end are accountable for the resonance at lower frequency of ISM band, see Fig. 3a, b, respectively.

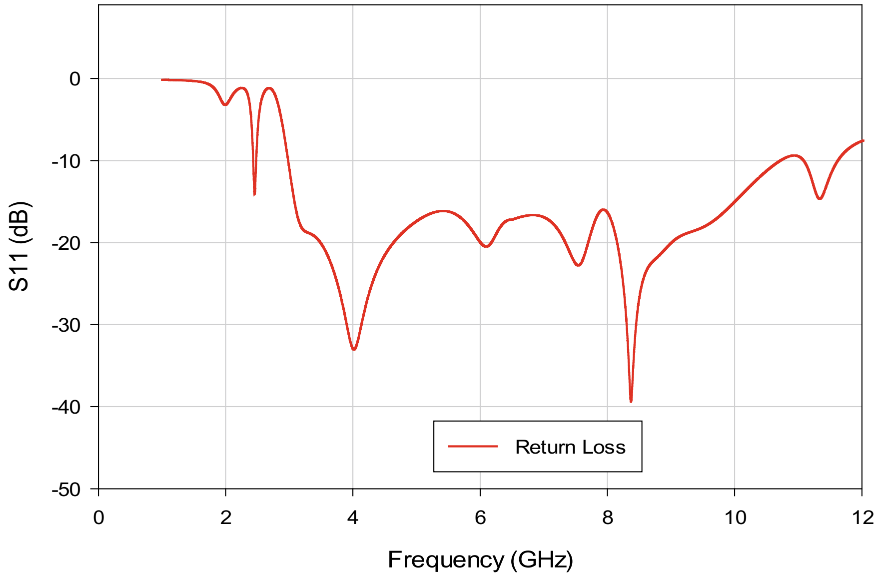


Fig. 2 Return loss of the proposed ISM/UWB antenna

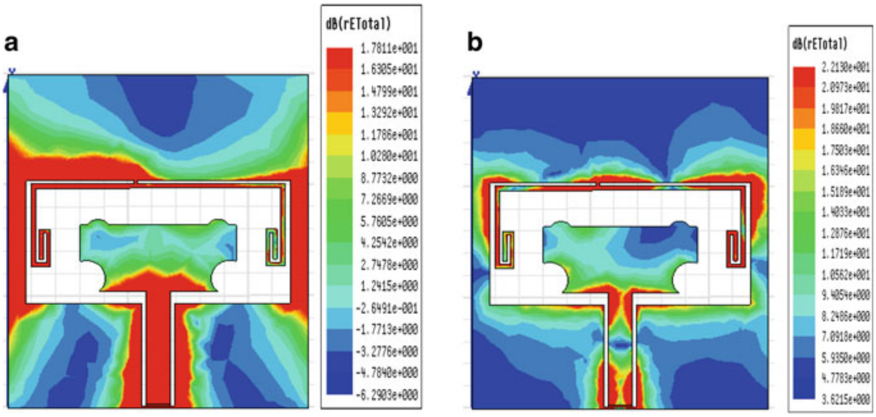


Fig. 3 **a** Surface current distribution at 2.45 GHz. **b** Surface current distribution at 7.5 GHz

5.2 *Parametric Analysis*

Through an extensive parametric analysis, Lhp, P, hn, wn and coil strip lengths C_1, C_2, C_3, C_4 are identified as critical parameters in the geometry. The impact of length of hanging metallic strip, Lhp from ground plane is such that return loss of ISM band and UWB band is better maintained when Lhp = 7.6 mm, see Fig. 4. The influence of P is also equally important as shown in Fig. 5. Influence of hn and wn is shown in Fig. 6 and coil strip length effect is shown in Fig. 7. These parameters have significant role in contributing lower bands along with UWB.

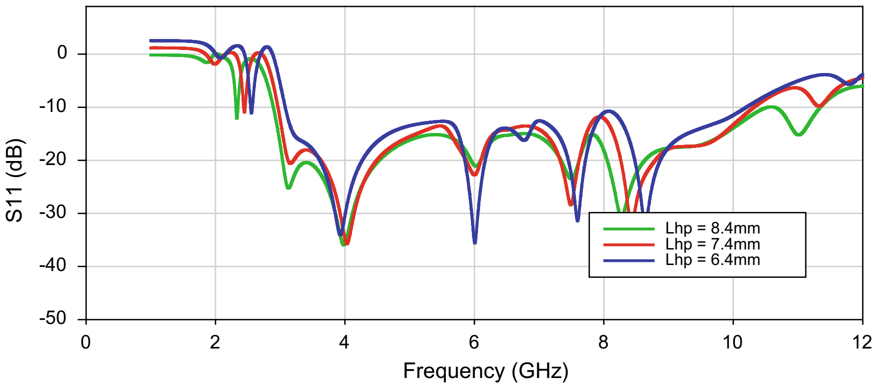


Fig. 4 Effect of Lhp on resonance

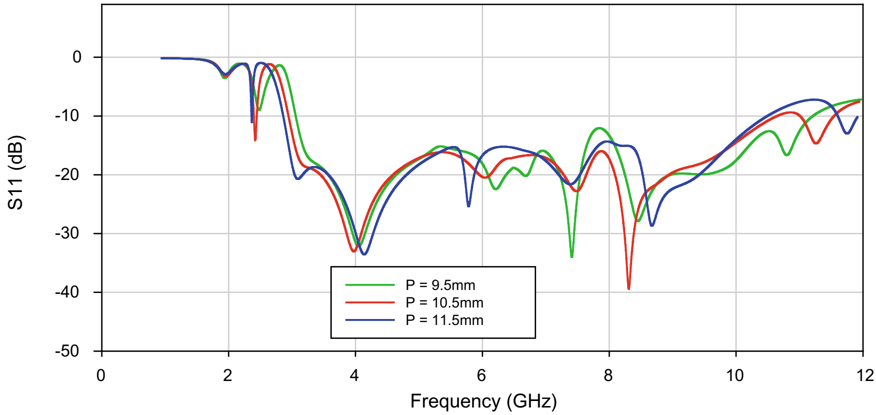


Fig. 5 Effect of *P* on resonance

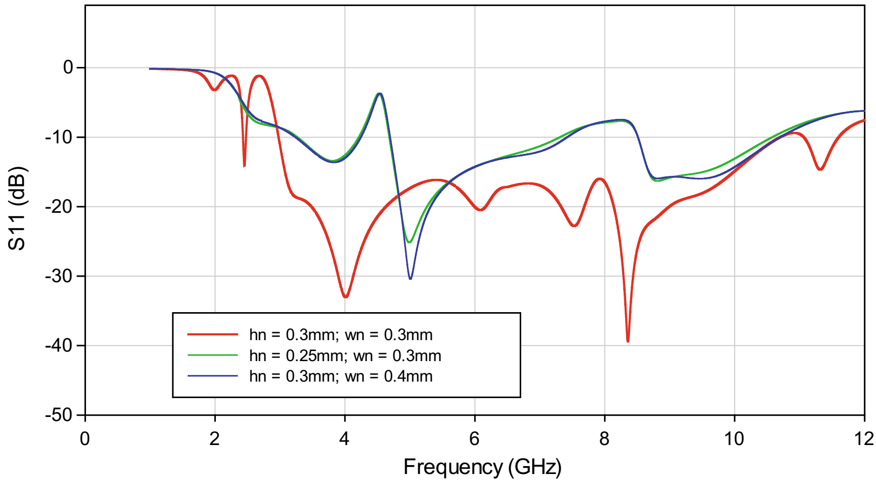


Fig. 6 Effect of hn and wn on resonance

5.3 Radiation Pattern

Analysis of radiation features of the proposed ISM/UWB antenna is done. The 2D radiation pattern in H-plane and E-planes at the operating frequency of 2.45 GHz is shown in Fig. 8a, b, respectively. Good far-field radiation properties are achieved. An almost omnidirectional pattern is obtained in the xz plane due to the symmetrical structure of the antenna.

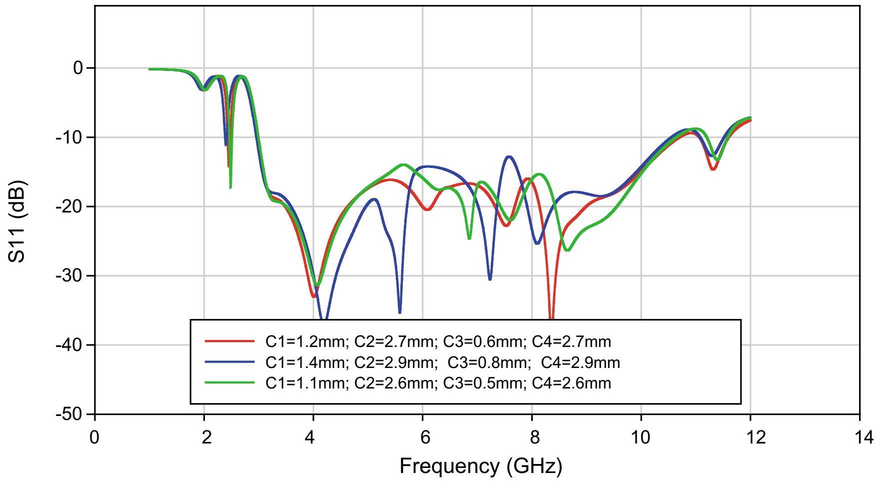


Fig. 7 Effect of coil strips on resonance

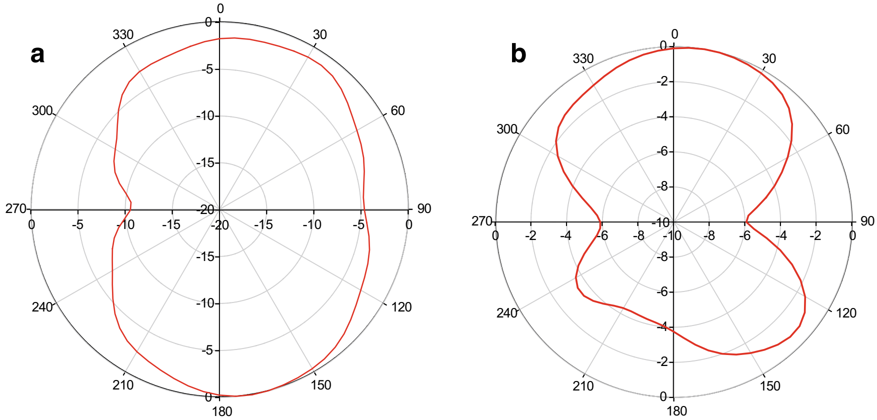


Fig. 8 **a** 2D radiation pattern at H-plane. **b** 2D radiation pattern at E-plane

6 Experimental Results

Fabrication of the antenna was done by photolithographic process, and the measurements were taken. The fabricated antenna and the experimentally measured return loss are shown in Figs. 9 and 10, respectively. An unexpected band at 2 GHz is found in addition to the designed ones in the experimental result (see Fig. 10). The simulation result obtained is in good agreement between the simulated and measured results as shown in Fig. 10 and hence validated the simulation result.

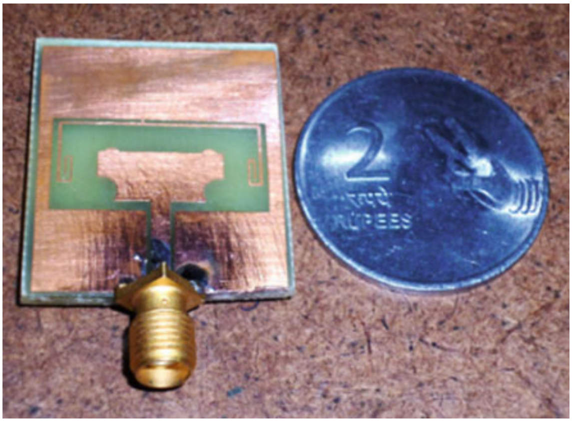


Fig. 9 Top view of fabricated antenna

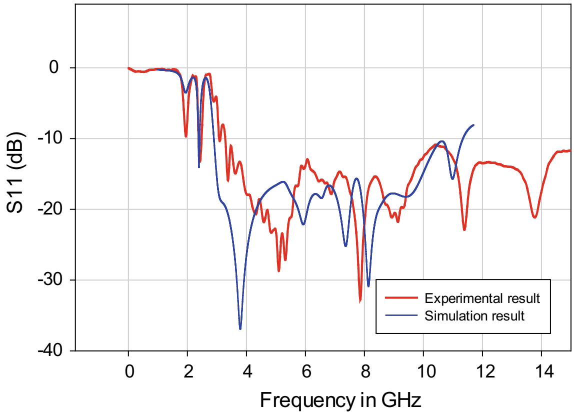


Fig. 10 Experimental result of the proposed antenna

7 Conclusion

The paper presented here for a novel dual-band cost-effective and low-profile coplanar antenna is suitable for applications in ISM band at 2.45 GHZ and UWB (3.1–10.6 GHz). Extensive parametric analysis has been conducted and critical parameters identified. The optimized structure resonates at 2.45 GHz with a bandwidth of 50 MHz starting from 2.43 to 2.48 GHz which fully covers the ISM band at 2.45 GHz. The structure also resonates in the whole UWB range of 3.1–10.6 GHz. An additional band is also seen at 2 GHz in the experimental result which could be further improved to obtain a triple-band antenna. The radiation properties are also promising for further investigations to develop compact multiband antenna suitable for wireless applications.

Acknowledgements The authors wish to thank the research scholars and faculty members of Centre for Research in Electromagnetics and Antennas, Cochin University of Science and Technology, for their help and support in antenna fabrication and experimental results.

References

1. FCC (2003) Revision of the part 15 commission's rules regarding ultra-wideband transmission systems. First report and order, pp 98–153
2. Chen ZN (2007) Antennas for portable devices, 2nd edn. Wiley
3. http://www.bluetooth.org. Last Accessed 9 Dec 2018
4. Bensky A (2000) Short-range wireless communications. LLH Technology Publishing
5. Balanis CA. Antenna theory: analysis and design, 3rd edn
6. Schantz HG (2003) A brief history of ultra wideband antennas. In: IEEE UWBST conference
7. Hui Z (2008) Basic research on ultra wideband antenna and indoor positioning systems. Dissertation, Waseda University
8. Lin KC, Lin CH, Lin Y-C (2013) Simple printed multiband antenna with novel parasitic element for multistandard mobile phone applications. IEEE Trans Antennas Propag 61(1):488–491
9. Li G, Zhai H, Li T, Ma X, Liang C (2013) Design of a compact UWB antenna integrated with GSM/WCDMA/WLAN bands. Prog Electromagnet Res 136:409–419
10. Xiong L, Gao P (2012) Dual band planar monopole antenna for bluetooth and UWB applications with WiMAX and WLAN band-notched. Prog Electromagnet Res Lett 28:183–194
11. Anilkumar CV, Damodaran S, Paul B (2015) A printed antenna with circular cut for bandwidth enhancement. In: IEEE TENCON
12. Anilkumar CV, Varghese B, Paul B (2013) A novel printed antenna for WLAN applications at 2.45 GHz. Inter J Sci Engg Res (IJSER) 4(8)
13. Wong K-L, Chen M-T (2013) Small-size LTE/WWAN printed loop antenna with an inductively coupled branch strip for bandwidth enhancement in the tablet computer. IEEE Antennas Wirel Propag Lett 61(12):614–6151
14. Chien H-Y, Sim C-Y-D, Lee C-H (2013) Dual-band meander monopole antenna for WLAN operation in laptop computer. IEEE Antennas Wirel Propag Lett 12:694–697

A Modified Partitioning Around Medoids Clustering-Based Cluster Head Selection Scheme for Data Offload in Mobile Cloud Sensor Network

S. Jeen Shene and W. R. Sam Emmanuel

Abstract Mobile cloud-deployed mobile sensing networks are a growing area of technology where attaining energy utilization is a challenging task during data transmission from mobile sensor devices to the cellular base station. Data offload can address drawbacks like network delay, poor performance, and high-energy consumption. Such a setup requires an efficient scheme that focuses on energy efficiency in a better way reducing the faster death of nodes. In this paper, an energy-aware approach named modified partitioning around medoids with cluster head selection (MPAM-CHS) is proposed, that aims for better clustering of mobile devices and the fairer selection of group head to minimize the energy utilization of the nodes. The proposed scheme consists of four phases like initialization, clustering, cluster head formation, and transmission phase. Initially, the nodes are randomly deployed in the network field and then clustering is performed on them using a modified PAM algorithm to determine the actual cluster points for partitioning the nodes into small groups. Next, the cluster head (CH) or the group head is determined based on the criteria such as residual energy, signal-to-noise ratio (SNR), path loss, and average path loss between the sensor and the sink. Finally, the sensed information collected from the nodes is offloaded to the group head, aggregated, and then sent to the sink. The experimental analysis shows that the proposed algorithm has a significant gain in energy consumption in terms of network utilization and lifetime metrics.

Keywords Energy efficiency · Mobile cloud · Mobile sensing · Node clustering · Cluster head · Network lifetime · Threshold distance

---

S. Jeen Shene (✉) · W. R. Sam Emmanuel
Department of Computer Science and Research Centre, Nesamony Memorial Christian College, Marthandam, Affiliated to Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India
e-mail: jeenshene@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_44

1 Introduction

Mobile smartphone devices are nowadays gaining more popularity and increasingly embedded with various sensors and several other computing capabilities. These sensors are responsible for acquiring a large amount of sensing data and provide useful information with high precision data to the users [1]. The recent version of mobile devices is equipped with several sensors such as ambient sensor, accelerometer, Wi-Fi, and GPS [2]. This wide range of capabilities tends to apply in various domains include environmental monitoring, healthcare, transportation, social networks, smart homes, etc [3]. The application ranges from the individual personal monitoring (such as personal health care) to the large-scale criteria such as military field and traffic area monitoring. These capabilities provide great attention and enabled innovation in mobile sensing technologies [4]. To meet the user requirements, several features have to be improved for future changes in developing computing technologies [5, 6].

Mobile cloud computing is a promising approach that offers several services in terms of the median of mobile as an interface to the users. One of the advantages is the offloading strategy that allows data to be stored in cloud and to be returned to the users on demand [7]. The major task of the mobile sensing is to enhance the performance while maintaining low-energy consumption during transmission of data. When Wi-Fi is not available, the mobile node utilizes an additional energy for transmitting the sensed data. Thus, each of the mobile sensors losses its energy and soon causes more dead nodes in the network. Even though offloading offers a greater advantage in computation capability, energy depletion is a greater challenge in mobile sensing systems [8].

For reduced energy utilization, clustered communication is followed, that allows dividing of the sensor nodes into small groups and further sending forward the information to the sink after aggregation. This type of transmission usually involves two stages including a setup sage and a steady-state stage. In the setup stage, partitioning of nodes is performed, and in steady state, data is then sent forward to the base station [9, 10]. In order to enhance the efficiency, a head is selected for each cluster that is responsible to perform compression and aggregation of data. This makes the cluster to remove the redundancy of the data and to transmit the information to the sink in an efficient manner. Thus, the cluster head (CH) selection plays a significant role in reducing the energy usage of the mobile nodes in the sensing field. For this purpose, an efficient clustering followed by CH estimation is performed and the threshold probability computed changes the group head on a regular time interval basis. The performance is estimated in terms of network lifetime duration and network utilization.

2 Related Work

Several models have been proposed earlier for the CH selection that utilizes the energy as low as possible. However, the existing models have certain limitations due to the selection of low efficient node as a group head that causes high-energy consumption, network delay, additional overhead, and the earlier death of nodes.

Khan et al. [11] presented a fuzzy-TOPSIS method for the cluster head selection based on the multi-criteria decision making to maximize the lifespan in the mobile wireless sensor network. Bhatti et al. [12] proposed a novel clustering approach using fuzzy c-means method (FCM) and the group head formation based on the location, fusion center, energy, and signal-to-noise ratio (SNR) measure. These are considered to estimate the CH selection based on minimum energy consumption and reduced energy utilization of neighboring nodes. Loomba et al. [13] proposed two algorithms that aim to optimize the energy utilization of CH selection for data transmitted between the mobile sensing devices and the cellular base station. The first algorithm modulates the nearest neighbors based on the distance factor criterion. The second model adds an additional forgetting factor, to ensure not to use the already obtained group head for set intervals of time.

Sarkar et al. [14] developed a firefly algorithm with cyclic randomization (FCR) that optimally selects the CH to reduce the energy usage. But the FCM approach suffers from the high complexity for determining the node to become CH. Darabkh et al. [15] discussed two methods, namely centralized density- and threshold-based CH replacement, combined with adaptive data distribution methods. In this, base station (BS) estimates the remaining energy of the network and decides whether a node to serve as CH or not.

Kaswan et al. [16] discussed a multi-objective particle swarm optimization scheme that estimates the path for the sink based on the Pareto dominance approach. This algorithm determines the rendezvous points for each node and designs a path for the mobile sink (MS). Darabkh et al. [17] proposed a modified threshold-based replacement of CH that aims to enhance the functionality of the network based on the threshold energy level. Here, the threshold energy is modified to estimate the actual number of clustering operations carried out by the node. Hong et al. [18] proposed a threshold-based low-energy adaptive clustering hierarchy method that alters the CH based on the threshold-based residual energy level in the sensor nodes. In this, the present node continues to act as a CH even after the specific interval of time since they maintain the energy above threshold limit.

Darabkh et al. [19] proposed an energy and density-based clustering routing protocol for data collection in WSN. The purpose of this method is to distribute the load evenly between the available sensor nodes to balance the energy consumption in the network. Therefore, a new generic network model is introduced to partition the network into identical dimensions and sub-layers. Finally, a relaying algorithm is used in cluster heads to be aware of sensor nodes located in the layer ahead to the base station to pick up the relay node according to the highest priority weight, resulting in good network lifetime and energy consumption.

Zafar et al. [20] proposed a clustering-based routing approach that enables us to improve the energy efficiency and scalability based on three-layer techniques of mobility-aware centralized clustering algorithm (MCCA) and mobility-aware hybrid clustering algorithm (MHCA). This MCCA algorithm employs centralized layers of clustering hierarchy, and MHCA employs gridding at the upper layer and distributed at the lower layers.

In the previous approaches, CH consumes more energy due to the continuous monitoring of the network field and heavier task performed in both the setup and steady-state phases. This causes faster node death and degradation in the performance of the system. To overcome such issues in the mobile sensing transmission system, the proposed work on MPAM approach partitions the sensors based on the medoid points located in the central portion of the clusters, which in turn selects a fairer CH to improve network utilization and network lifetime, ultimately able to reduce the power consumed by the nodes during data transmission.

3 Modified Partition Around Medoid-Based Cluster Head Selection (MPAM-CHS) for Data Offload

3.1 Clustering Model

The proposed work considers a clustering model of B $\times$ B sample area as shown in Fig. 1, where the application requests for sensed data from mobile devices equipped with sensors. The cluster head aggregates data compresses it and offloads it to the nearest base station then on to the cloud server where the intended applications are hosted. This work assumes certain technology of communication between the nodes to CH, CH to base station (BS) and BS to the cloud server as shown in Table 1.

3.2 Clustering Phases

The proposed work involves the use of a MPAM-CHS algorithm that performs clustering and cluster selection for effective transmission in mobile cloud sensing environment. The proposed methodology has four phases: initialization/node setup phase, clustering/cluster setup, CH selection, and transmission phases which are shown in Fig. 2.

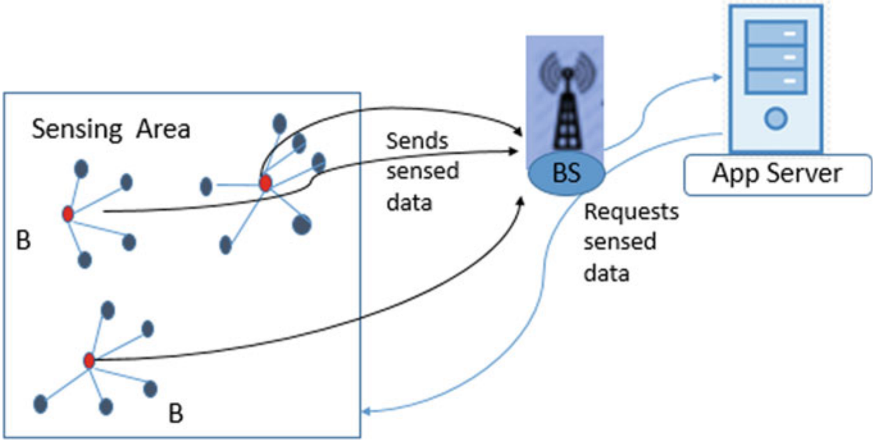


Fig. 1 Clustered architecture in a B*B area

Table 1 Communication versus technology used in the model

Communication	Technology
Devices to cluster head	Wi-Fi direct transmission
Cluster head to base station	3G transmission
Base station to cluster head	3G transmission

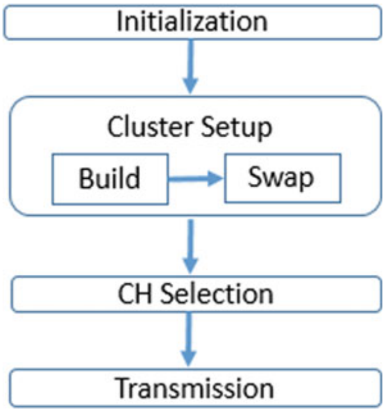


Fig. 2 Cluster head selection process

3.2.1 Node Setup Phase

Consider the number of mobile sensor nodes, S, that monitors the network field and transfers the aggregated information to the mobile sink, M. Initially, the nodes are placed randomly on the network field. After this deployment, the sink broadcast the hello packets that define the information about its location information.

3.2.2 Clustering/Cluster Setup by Modified PAM Algorithm

After the initialization phase, the sensors placed in the network field are divided into small groups called clusters. The number of clusters is fixed, and it is set using initial settings in the beginning. For example, if the cluster size is set as two, then two clusters are formed and the size of the each cluster is depending on the number of nodes generated randomly within the cluster. There is no restriction on the maximum number of nodes per cluster. The algorithm chooses some nodes and groups them to form a cluster. This algorithm consists of two phases: build phase that estimates and identifies the initial center of the cluster and swap phase that improves the clustering by swapping (replacing) the medoid by non-medoid point, thereby finding the best medoid point for clustering. This clustering algorithm uses the k-representative objects s_d called medoids instead of mean measure which has the total deviation (T_v) as objective; the dissimilarities obtained with respect to the point $x_i \in c_i$ to the medoid s_d of its cluster are given by Eq. (1), where d_m represents the Manhattan distance.

$$T_v = \sum_{j=1}^{k} \sum d_m(x_i, s_d) \tag{1}$$

Build Phase

In this phase, initial set of nodes is selected to determine the good clustering with the minimum deviation factor. The main objective of this algorithm is to decrease the dissimilarity of the other nodes in the specified cluster set. For this purpose, the algorithm determines the medoid in the place of the mean point that is actually located in the center of the cluster. For the distance measure, Manhattan distance function is used and the smallest distance is calculated by Eq. (2).

$$T_{\mathrm{vi}} = T_{\mathrm{vi}} + d(x_0, x_i) \tag{2}$$

Initially, the k set of representative points is chosen arbitrarily to find the initial clusters. The first iteration obtains the initial medoid point based on the smallest distance sum and assigns the node to the nearest medoid point. The second iteration estimates the other medoid points, adds those medoid points, and estimates the distance measure (β) between two points (x_0, x_i), based on Eq. (3).

$$\beta = d(x_0, x_i) - \min_{f \in s_1, \ldots, s_i} d(x_0, f) \tag{3}$$

Then, the distance of each node to the representative point is computed and selection of set of clusters with minimum deviation factor is performed.

Swap Phase

In this phase, non-medoid points are allowed to replace the medoid points and the benefit is estimated to enhance the quality of clustering. If the exchange can assure improvement, then the previous medoid is exchanged to the current medoid point. Here, nodes are assigned to the new representative point and the sum of distance of all the nodes to the current medoid point is calculated. If the obtained result is equal to the current distance, then the algorithm stops else it updates the old medoid to the new medoid point in the cluster.

The work considers the distance between nodes inside a cluster in the build phase and the distance from a node to medoid point in swap phase. These two phases of processing facilitate gaining improved medoid points and the formation of effective sensor node clusters in the network.

3.2.3 Cluster Head Formation

After clustering, the corresponding cluster head is estimated that collects the sensed information from the member nodes and transfers it to the sink. The selection process uses four parameters such as residual energy, SNR, average path loss between the node and the other nodes, and the path loss between the sensor node and the base station. For the cluster (l), the CH is chosen based on the objective function (ψ) and is given in Eq. (4),

$$\psi_i^l = \left(\frac{R_i^l \, \gamma_i^l}{\alpha \, P_i^l + (1 - \beta) P_{\mathrm{bs}}^l} \right) \tag{4}$$

where R_i^l is the residual energy (i.e., remaining energy left later sensing and transmission), γ_i^l is the SNR node, P_i^l is the average path loss between the ith sensor and other member nodes in the l cluster, P_{bs}^l denotes the path loss between the ith sensor and BS, and α denotes the weight assigned of P_{bs}^l and P_i^l The average path loss (P_i^l) is calculated by Eq. (5),

$$P_i^q = \frac{\sum_{j=1}^{M_l} P_{i,j}^l}{M_l} \tag{5}$$

where M_l defines the total number of sensors in the cluster l, $P_{i,j}^l$ is the path loss of jth sensor and its cluster member node i and is given by Eq. (6),

$$P_{i,j}^l = 10r \log_{10} D_{i,j}^l \tag{6}$$

where $D_{i,j}^l$ defines the distance between the ith and jth sensors, r is the path loss exponent. Similarly, the path loss P_{bs}^q between the ith mobile sensor and BS is given by Eq. (7),

$$P_{\text{bs}}^q = 10r \log_{10} D_{\text{bs}}^q \tag{7}$$

where D_{bs}^q is the distance function estimated between the base station and the cluster node.

The Modified PAM Algorithm

Based on the estimated value in Eq. (4), the node with maximum objective measures is determined as the head for the specific cluster set. This work involves cluster setup and cluster head selection using MPAM-CHS algorithm. The flow of steps involved in the process is depicted by Algorithm 1, which takes set of input nodes to be clustered as input and returns k group of clusters with efficient CH node identified as output. It starts with initializing the set of nodes in the network and proceeds through clustering and cluster head selection to end with transmission between the member nodes and CH. The cluster setup randomly picks medoids, considers the Manhattan distance between two nodes, computes deviation factor, and looks for reduced deviation factor while including nodes in clusters around the chosen medoid. It continuously improves the selection and proceeds by swapping with identified best medoid. The CH selection involves identifying the best cluster head based on SNR and path loss values of participating nodes.

Algorithm.1: MPAM-CHS algorithm

Input: Set of input nodes to be clustered

Output: clustering of nodes into k groups and efficient CH node

1.Initilaization:

Initilaize the set of nodes in the network

2.Cluster setup using Modified PAM:

Randomly choose the initial set of reprersentative points as medoids

Estimate the distance between two data points based on the Manhattan distance measure

for each medoid //***BUILD phase***

Initialize the total deviation factor and medoids as (∞,null)

for each set of representative point ***do***,

initialize the deviation factor as 0,

estimate the T_v for the every other node and choose the smallest distance

for other medoids do

for the representative points other than medoids, set total deviation factor as 0.

for every nodes other nodes and medoids points ***do***

compute β from equation (3)

determine the best reduction in deviation factor

return

repeat

swap the medoid and non-medoid points // ***SWAP phase***

estimate the benefit of new medoid based on the distance measure

if the best improvement is made, then swapping is performed

return (total deviation, medoids, cluster)

3.CH formation

calculate SNR, residual energy of the nodes in the network

Calculate the path loss from equation (5) and (7)

Estimate the best CH based on the equation (4)

4 .exit

3.2.4 Transmission Phase

The current cluster head is responsible to coordinate the transmission, and the mobile sensed data transmissions happen between the cluster head and the other nodes accordingly. From the selected CH, the sensed information collected from their corresponding member nodes is aggregated and dispatched to the base station during specific time intervals.

4 Results

The network simulation of the proposed approach in energy-aware CH selection is implemented using MATLAB platform. Couple of metrics such as network utilization (N_U) and network lifetime (N_L) are used for the evaluation of the proposed method. The analysis with obtained results involves a comparative study of proposed technique with a few other existing methods on the chosen metrics as shown in Table 2.

While seeing through all the network dimensions, there is a consistent increase of both network lifetime and utilization across dimensions for the proposed technique. It proves that the proposed technique outperforms the existing methods in terms of network lifetime and utilization, thereby implicates energy saving and it is the better approach.

Figure 3 shows the network utilization versus the network dimension in comparing the proposed MPAM-CHS method with the existing techniques. Here, T-LEACH shows the least performance due to the faster node death as threshold factor. The performance of the MT-CHR, C-DTB-CHR-ADD is fairly better relatively than T-LEACH due to cluster management that causes power reduction. However, the proposed method's plot lies on top achieving a better performance than the existing methods under the study due to the adoption of robust clustering and efficient cluster head selection strategies.

Table 2 Comparison of network lifetime and utilization metrics with existing approaches at various dimensions

Network dimensions	T-LEACH [18]		MT-CHR [17]		C-DTB-CHR-ADD [15]		Proposed	
	N_L (rounds)	N_U (%)	N_L (rounds)	N_U (%)	N_L (rounds)	N_U (%)	N_L (rounds)	N_U (%)
100×100	800	73	900	77	1400	79	2488	83
200×200	700	67	800	72	1280	74	2480	81
300×300	620	61	700	66	1150	69	2468	75
400×400	500	56	580	61	950	64	2450	66

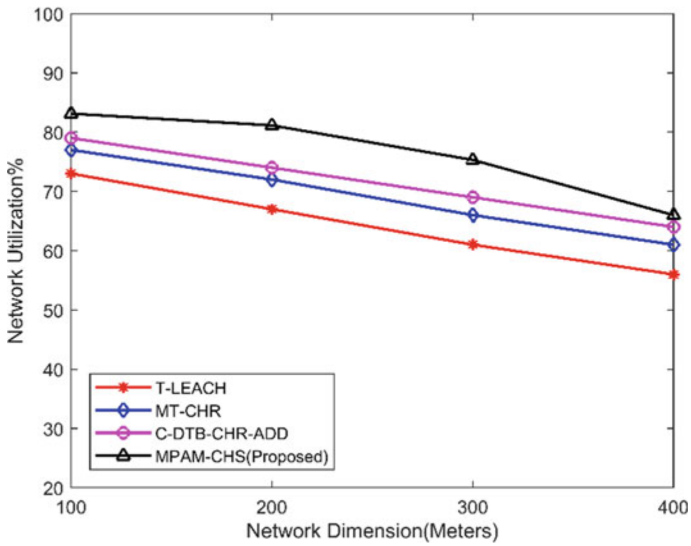


Fig. 3 Network utilization versus network dimension

Figure 4 indicates the comparison of the proposed method with the existing approaches in terms of network lifetime measure. Here, T-LEACH shows the least network lifetime where MT-CHR is slightly better than T-LEACH technique. This is due to the threshold probability to select the node at every node, which leads to

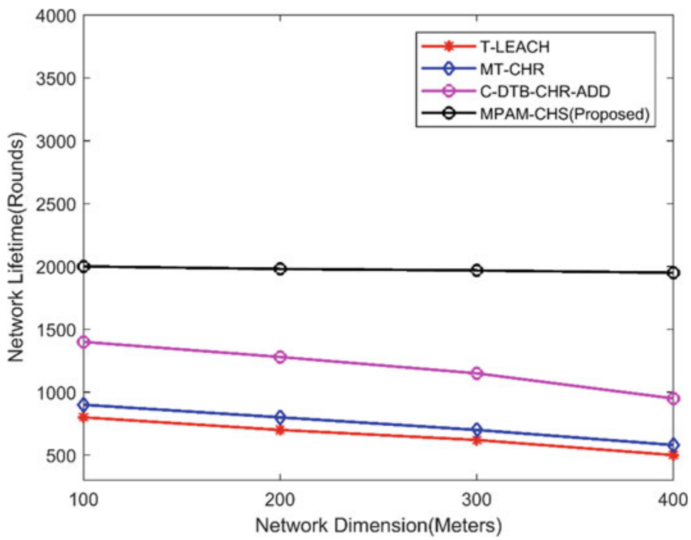


Fig. 4 Network lifetime versus network dimension

faster node death. C-DTB-CHR-ADD shows improved result than the T-LEACH and MT-CHR because it interconnects the minimum distance and balances the load in transmission process. However, the proposed method stands out of all the other methods studied and compared due to the energy-saving strategy of cluster setup and CH selection in the network, thereby enhancing the network lifetime.

5 Conclusion

This paper presents a new energy-efficient clustering scheme with the cluster head selection approach that transmits the sensed information collected from the nodes to the BS, then offloads data on to a cloud application server, in a mobile cloud-deployed environment of mobile sensor network. The proposed scheme performs clustering based on the modified PAM approach that uses the medoid points that determine the centrally located point, thereby minimizing the distances between the nodes to the labeled center in the specific cluster set. Then, cluster head is selected based on the four estimated parameters. The performance is evaluated under two metrics, and the results show that our proposed method achieves a significant improvement in energy efficiency than the existing approaches. This technique can be further enriched by introducing more parameters in selecting an efficient CH, and study can be extended involving more metrics to evaluate the performance of the model.

References

1. Rachuri KK, Musolesi M, Mascolo C, Rentfrow PJ, Longworth C, Aucinas A (2010) EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. In: Proceedings of the 12th ACM international conference on ubiquitous computing. ACM, pp 281–290. https://doi.org/10.1145/1864349.1864393
2. Zhang X, Yang Z, Sun W, Liu Y, Tang S, Xing K, Mao X (2015) Incentives for mobile crowd sensing: a survey. IEEE Commun Surv Tutorials 18(1):54–67. https://doi.org/10.1109/COMST.2015.2415528
3. Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. IEEE Commun Mag 48(9):140–150
4. Khan WZ, Xiang Y, Aalsalem MY, Arshad Q (2012) Mobile phone sensing systems: a survey. IEEE Commun Surv Tutorials 15(1):402–427. https://doi.org/10.1109/SURV.2012.031412.00077
5. Kumar K, Lu YH (2010) Cloud computing for mobile users: can offloading computation save energy? Computer 43:51–56. https://doi.org/10.1109/MC.2010.98
6. Haghighi V, Moayedian NS (2018) An offloading strategy in mobile cloud computing considering energy and delay constraints. IEEE Access 6:11849–11861. https://doi.org/10.1109/ACCESS.2018.2808411
7. Othman M, Madani SA, Khan SU (2013) A survey of mobile cloud computing application models. IEEE Commun Surv Tutorials 16(1):393–413. https://doi.org/10.1109/SURV.2013.062613.00160

8. Liu X, Yang Q, Luo J, Ding B, Zhang S (2018) An energy-aware offloading framework for edge-augmented mobile RFID systems. IEEE Internet Things J. https://doi.org/10.1109/jiot.2018.2881295
9. Chidean MI, Morgado E, Sanromán-Junquera M, Ramiro-Bargueno J, Ramos J, Caamano AJ (2016) Energy efficiency and quality of data reconstruction through data-coupled clustering for self-organized large-scale WSNs. IEEE Sens J 16(12):5010–5020. https://doi.org/10.1109/JSEN.2016.2551466
10. Souza ÉL, Pazzi RW, Nakamura EF (2015) A prediction-based clustering algorithm for tracking targets in quantized areas for wireless sensor networks. Wirel Netw 21(7):2263–2278. https://doi.org/10.1007/s11276-015-0914-3
11. Khan BM, Bilal R, Young R (2018) Fuzzy-TOPSIS based cluster head selection in mobile wireless sensor networks. J Electr Syst Inf Technol 5(3):928–943. https://doi.org/10.1016/j.jesit.2016.12.004
12. Bhatti D, Saeed N, Nam H (2016) Fuzzy c-means clustering and energy efficient cluster head selection for cooperative sensor network. Sensors 16(9):1459. https://doi.org/10.3390/s16091459
13. Loomba R, de Frein R, Jennings B (2010) Selecting energy efficient cluster-head trajectories for collaborative mobile sensing. In: 2015 IEEE global communications conference (GLOBECOM). IEEE, pp 1–7. https://doi.org/10.1109/glocom.2015.7417727
14. Sarkar A, Murugan TS (2019) Cluster head selection for energy efficient and delay-less routing in wireless sensor network. Wirel Netw 25(1):303–320. https://doi.org/10.1007/s11276-017-1558-2
15. Darabkh KA, Wala'a S, Al-Zubi RT, Alnabelsi SH (2017) C-DTB-CHR: centralized density- and threshold-based cluster head replacement protocols for wireless sensor networks. J Supercomputing 73(12):5332–5353. https://doi.org/10.1007/s11227-017-2089-4
16. Kaswan A, Singh V, Jana PK (2018) A multi-objective and PSO based energy efficient path design for mobile sink in wireless sensor networks. Pervasive Mobile Comput 46:122–136. https://doi.org/10.1016/j.pmcj.2018.02.003
17. Darabkh KA, Wala'a S, Hawa M, Saifan R (2018) MT-CHR: a modified threshold-based cluster head replacement protocol for wireless sensor networks. Comput Electr Engg 72:926–938. https://doi.org/10.1016/j.compeleceng.2018.01.032
18. Hong J, Kook J, Lee S, Kwon D, Yi S (2009) T-LEACH: the method of threshold-based cluster head replacement for wireless sensor networks. Inf Syst Front 11(5):513. https://doi.org/10.1007/s10796-008-9121-4
19. Darabkh KA, Odetallah SM, Al-qudah Z, Ala'F K, Shurman MM (2019) Energy-aware and density-based clustering and relaying protocol (EA-DB-CRP) for gathering data in wireless sensor networks. Appl Soft Comput 80:154–166
20. Zafar S, Bashir A, Chaudhry SA (2019) Mobility-aware hierarchical clustering in mobile wireless sensor networks. IEEE Access 1(7):20394–20403

Design of Dual-Band Compact Planar of MIMO Antenna with Pattern Diversity Characteristics

Vineesha Alladi, S. Ashok Kumar, and T. Shanmuganantham

Abstract The multi-input multi-output (MIMO) terminal characteristics are used for the high examined element of dual-band MIMO antenna with the pattern diversity that has been proposed. The MIMO terminal antenna is having two similar monopole components. By calculating the envelope related to each other coefficient (ECC) and radiation patterns, the MIMO terminal pattern is checked. This antenna exhibits good efficiency and gain. The outcomes obtained for the designed antenna are very perfect.

Keywords Pattern diversity · MIMO system · Compact self-isolated antenna

1 Introduction

Multi-input multi-output (MIMO) technology has gained a significant attentiveness in the ancient decade due to their property of excess data rate with good reliability and range. The MIMO technology uses number of antennas at transmitter and receiver sides, and the antenna meets the expectation of the modern wireless communication system, where the data is increasing exponentially combining the multi-standard communication system [1]. The several numbers of bands in MIMO order will inherently be much better. Various double-band multi-input and multi-output receivers are to be reported with the dissimilar decoupling circuits and to defeat the coupling phenomenon in the existed bands [2, 3]. The MIMO antenna configuration has a common ground plane. Moreover, orthogonal position of antenna elements helps in attaining good diversity. Mostly, separate intrinsic characteristics are needed for each band [4, 5]. While printing a circuit board, it consists of complex dissociate

V. Alladi (✉) · S. Ashok Kumar
Jyothishmathi Institute of Technological Sciences, Karimnagar, India
e-mail: vineeshrao.alladi98@gmail.com

S. Ashok Kumar
e-mail: ashokape@gmail.com

T. Shanmuganantham
Pondicherry University, Puducherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_45

circuits created or occupied more space in between antenna elements. The physical properties of the MIMO terminal like size and shape result in the difficulty levels to attain portability of the devices [6]. Therefore, it is very difficult to design a double-band MIMO antenna terminal which has good efficiency and polarization diversity characteristics.

2 Configuration of MIMO Antenna

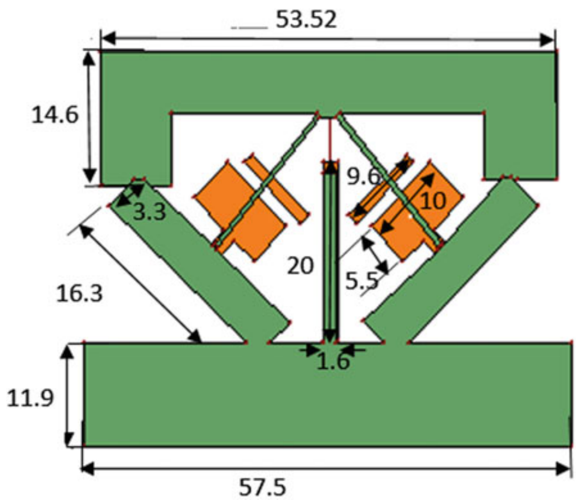
The structure is simulated and arranged in a 3-D electromagnetic (EM) and that can be done by using Zeland IE3D software. The structure is designed on a Neltec substrate, which has a top surface of 1.524, constant dielectric is 3.4, and loss tangent is 0.002. Therefore, the first element of a presented MIMO antenna is a redesigned version of already known antenna and is reproduced for MIMO application. Geometry of the presented antenna is shown in Fig. 1.

2.1 Analysis of MIMO Antenna

For better understanding the design pattern of the radiating element, the antenna structures are shaped as well as simulated by using Zeland IE3D software. In this design, two similar-shaped objects are attached to the half surface plane of the MIMO terminal to attain a double-band resonance.

The development of a separation can also be checked by comparing the reproduce plane of the present distribution MIMO receiver condition. In the MIMO receiver,

Fig. 1 Single-MIMO antenna configuration



almost the combining is done because of the CGP state which supplies an easy way to combine. MIMO antenna is moved outward by 1 mm. To invite a next resonance, further a metal rob is attached to the ground plane side.

2.2 Pattern Diversity Performance Analysis

The diversity of the presented antenna design has 3D radiation patterns of the MIMO terminal, and further, it helps in visualization of the pattern diversity of MIMO antenna prototype. MIMO antenna is distinguished by the envelope correlation coefficient (ECC) and calculated by using radiation pattern. ECC has less than 0.05 (ECC < 0.05) which preserves the good pattern diversity of a double-band MIMO antenna [7]. The MIMO antenna structure supports pattern diversity by the elements arranged in the orthogonal location. To envision of the pattern diversity, the 3D radiation patterns to be provided. The used two different types of patterns are giving ECC values which are very small in radiation simulation. The efficiency of isolation may trigger by applying the addition of thin metallic strips with equal distance to the elements.

3 Performance Validation

The model structure is established on a Neltec substrate, and the s-parameter can be corrected with dual-port vector network analyzer (VNA). As compared with the efficiency, results of simulation with the measured values slightly vary and that can be negligible. The measured results can acceptable with the simulated results.

3.1 Parametric Analysis of MIMO Antenna

The analysis of the main parameters can be PM1 and PM2. The parameters are proposed to examine the difference in the MIMO antenna. During the process, either PM1 or PM2 is changed, by keeping the other dimensions constant. The parameters may increase or decrease symmetrically about the mean position. The parameters PM1 and PM2 are at the front plane and ground plane. They are varied from 8.6 to 10.4 mm and 38.33 to 47.33 mm. The PM1 is used for achieving the upper band, and PM2 is used for achieving the lower band. The *s*-parameters are shown in Fig. 2.

The channel capacity standard with 3.4 GHz band is superior to 34 bits/s/Hz. The MIMO system *s*-parameter is influenced by specification t: the impedance matching and resonance frequency are affected (where T-structured component will reduce the impedance matching in antenna). The resonance frequency can be varied by stubs' length. Increasing the values of p and q will get the resultant band that is a small bandwidth. So, antenna size is reduced. Gain potential of 4×4 antenna configuration

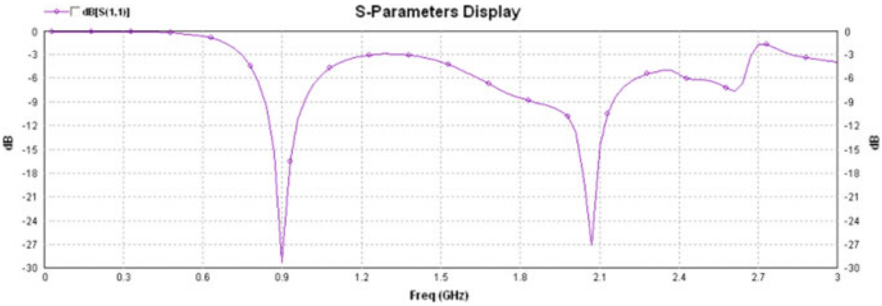


Fig. 2 *S*-parameter display

is good. The height of the antenna should be constant. By changing the values of length, we will get the gain in positive standard.

Moreover, from the *s*-parameter display it is observed that return loss is above −20 dB and frequency is in-between 900 and 2000 MHz. By using few specifications, we can know the operating assumption of MIMO antenna. Width of the *T*-structured component is 5 (it should be constant), and the height of the *T*-structured element can be changed. By increasing the *U*-shaped element height, the gain is decreased. So, it is better to keep the *U*-shaped element height as constant.

Obviously, the Antenna 2 covers the dual-band at the time of re-optimization and the distance between four antennas is the same, by which they are fabricated and calculated. The stabilized and reproduced *S*-parameters are arranged, and they are shown in Fig. 2. By doing the optimization again, the region of the antenna can be varied comfortably. *Z*-parameter of proposed MIMO shows the perfect impedance matching at resonant frequency as shown in Fig. 3.

The mentioned figure deals to measure the antenna efficiencies accepting the simulated efficiency results. The proposed theory of multi-input and multi-output system has good efficiency over isolation and simulation. Here MIMO antenna is a wireless communication technology. It is used at source and destination. The two

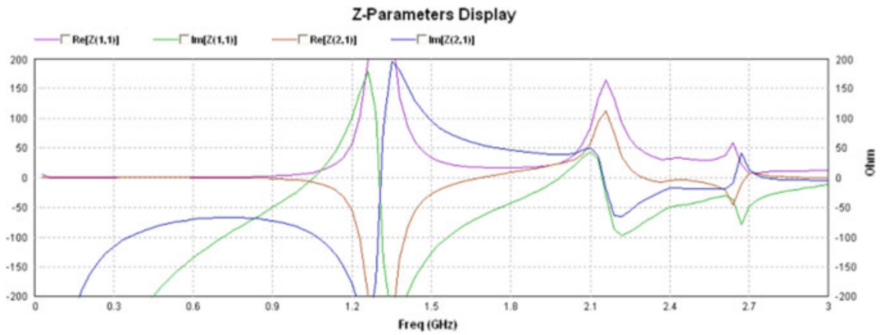
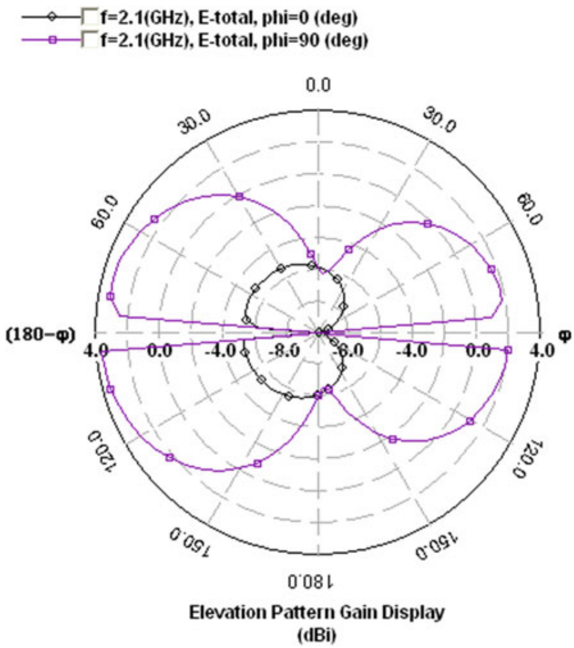


Fig. 3 *Z*-parameter

Fig. 4 Elevation pattern



dimensional radiation pattern of proposed antenna performs well in various angles and maximum gain of the above designed antenna shows 1dBi at resonant frequency as shown in Figs. 4 and 5. The efficiency of the antenna at resonant frequency is 75%.

Figures 6 and 7 show the directivity and gain between 4.5 and 5 dBi. MIMO-based system price is too high when compared to other antenna system, because MIMO-based systems contain both hardware and advanced software specification. A router with 4 × 4 MIMO has four antennas, and it can communicate on at once. If several 4 × 4 MIMO devices connected to that router, they would all maintain a connection of four data streams at the same time. Devices with more antennas are generally more expensive because of its hardware. Modern flagship phones generally have 2 × 2 MIMO. That extra wireless hardware will use a bit of extra power, so 2 × 2 MIMO might reduce battery life a tiny bit compared to 4 × 4 MIMO. But, the faster wireless speed and improved signal strength are always good. From figures, we can observe the gain value as a positive, where the gain of the single antenna is low when compared to the gain of 4 MIMO antennas.

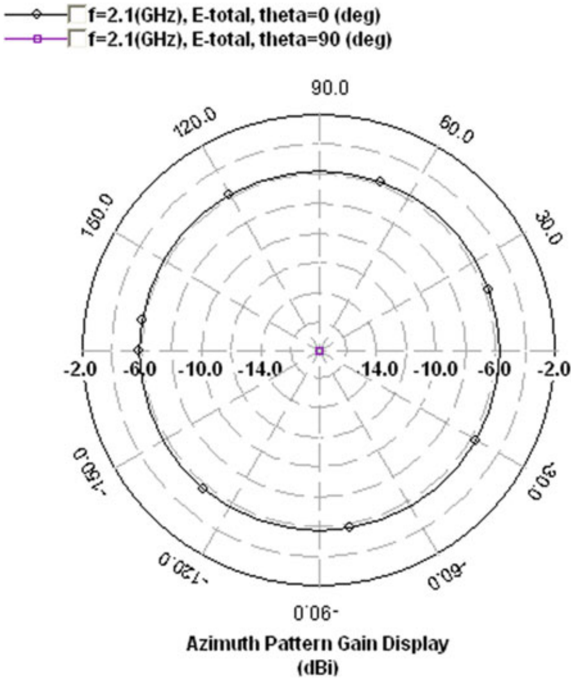


Fig. 5 Azimuth pattern

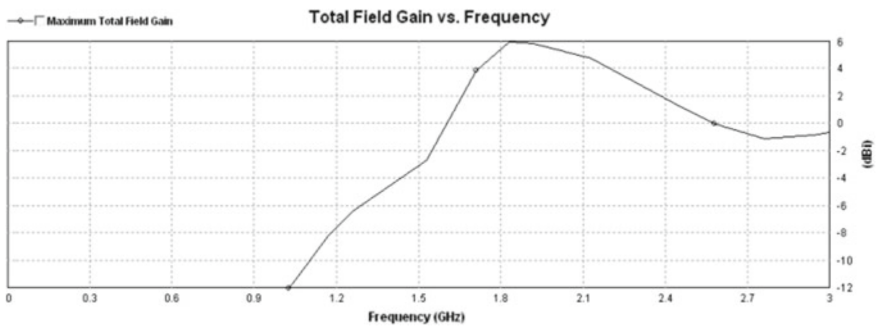


Fig. 6 Field gain versus frequency

4 Conclusion

MIMO antenna of the planar dual-band is proposed and verified. The MIMO antenna characteristics have high isolation and well-patterned diversity. The antenna elements are orthogonally located which outcomes in pattern diversity of the MIMO terminal. High isolation can gain from both the bands by the installation of metallic strips.

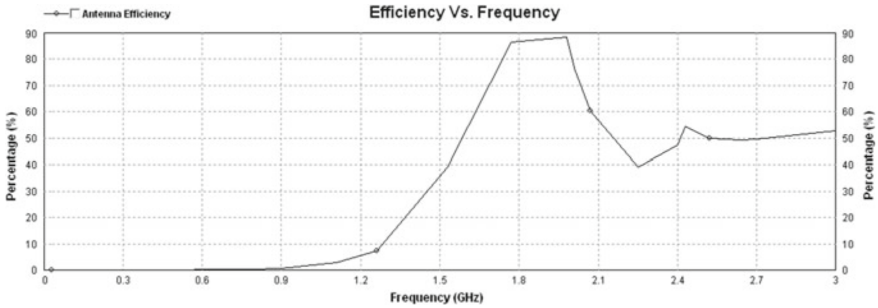


Fig. 7 Efficiency versus frequency

All the calculated outcome shows a good deal with the simulated characteristics. The above structure indicates that the presented antenna is superior for diversity applications.

References

1. Wang SM, Hwang LT, Lee CJ, Hsu CY, Chang FS (2015) MIMO antenna design with built-in decoupling mechanism for WLAN dual-band applications. Electron Lett 51(13):966–968
2. Zhang S, Pedersen GF (2016) Mutual coupling reduction for UWB MIMO antennas with a wideband neutralization line. IEEE Antennas Wirel Propag Lett 15:166–169
3. Kumar SA, Thangavelu S (2015) CPW fed implantable Z-monopole antennas for ISM band biomedical applications. Int J Micro Wire Tech 7:529–533
4. Gao Guo-ping, Bin Hu, Wang Shao-fei, Yang Chen (2019) Wearable circular ring slot antenna with EBG structure for wireless body area network. IEEE Antenna Wirel Propag Lett 17:434–437
5. Ashok Kumar S, Shanmuganantham T (2018) Design and performance of textile antenna for wearable applications. Trans Electricals Electron Mater 19(5):352–355
6. Raleigh GG, Cioffi JM (1998) Spatio-temporal coding for wireless communication. IEEE Trans Commun 46:357–366
7. Dong Y, Choi J, Itoh T (2018) Vivaldi antenna with pattern diversity for 0.7 to 2.7 GHz cellular band applications. IEEE Antennas Wirel Propag Lett 17(2):247–250

Energy Harvesting of Traditional Cantilever-Based MEMS Piezoelectric Energy Harvester

Arjunan Nallathambi and T. Shanmuganantham

Abstract Wireless sensor networks are important developments in remote sensing applications. Supplying power to these systems is invincible one, and replacing batteries all the time is inefficient and not an appropriate solution. Piezoelectric materials convert mechanical energy into electrical energy from the external vibration of the environment through its direct piezoelectric effect. In this paper, rectangular with and without hole based unimorph piezoelectric cantilever geometries are proposed using finite element method simulation and analysis developed. The outcomes display that the rectangle with hole piezoelectric structure is having a lower resonant frequency and harvests more energy compared to other structures.

Keywords Piezoelectric energy scavenger · Unimorph cantilever · Piezoelectric effect · MEMS tip mass · Finite element method

1 Introduction

Energy harvester is the method of changing presented moving energy into practical electrical energy over the usage of specific things of transduction device. In wireless technology, it is highlighted as an alternative to the power supply of its sensor nodes. The cantilever beam can consume extreme deformation when compared to other structures, and also its resonant frequency is lesser. The model of energy collecting with the cantilever constructions can be expected as blends of the seismic mass on a mechanism and under vibrations, its net effort can effort a dashpot which translates the mechanical energy into electrical energy [1–4]. Later sensor connection in remote places and reliable process in wireless sensor networks is an excited task. Likewise,

A. Nallathambi
Marri Laxman Redddy Institute of Technology and Management, Hyderabad 500043, India
e-mail: nallathambiarjunan@gmail.com

T. Shanmuganantham (✉)
Departments of Electronics Engineering, Pondicherry University, Puducherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_46

continuous supply of capacity is essential for process in an actual manner, and battery backup is not an appropriate and suitable solution. Hence, power harvesting is a subject for these devices which can be simply scavenged from the obtainable atmospheres of close environment. Therefore, ZnO materials are a devoted one to produce energy due to its benefits such as outstanding linear input–output connection, large dynamic range and small size easy incorporation with electronics [4]. The energy harvested of the power harvester is inversely proportional to the resonant frequency; hence, it is a determining aspect of any energy harvesters' constructions. Its outcome on the power harvesting output is expected one [1, 2]. While using these models as fluid flow path and flow rate sensor, the sensitivity is dependent on the Reynolds number which describes the fluid flow design and the applied force due to fluid flow that is connected to the strain created in the piezoelectric layer, and therefore, the sensitivity is also diverse [5–16].

Recently, altered unimorph cantilever with altered cross-sections is working to produce additional extent of energy. Also outdated cantilever beams, through changing the cross-section such as RCRC, RCTAC, TACRC, TRCTAC and TACTAC altered beams constructions, were used for energy collecting and also showed that the TRCTAC can produce extra voltage related to others [3–7]. In this work, two different traditional models such as rectangular and rectangular with hole models with silicon tip mass at free end are designed. The analytical method of these energy harvesters is presented and is simulated in the (FE) finite element tool-based software tool to compare its energy harvesting capability.

2 Design Material for Piezoelectric Sensor

Here, two different types of phenomena are used, for example, direct and indirect piezoelectricity. In conventional piezoelectricity, a voltage is produced as strain is created in the piezoelectric materials, and an example of indirect piezoelectricity, a voltage is generated due to the existence of electric field. There are two categories of piezoelectric materials such as piezoceramics like lead zirconate titanate (PZT) and piezopolymers such as polyvinylidene fluoride (PVDF). While piezoelectric materials are used for voltage detection and deformation of the design. The electrical and mechanical performance of these materials can be shown by two basic equations [2].

$$D = d_t T + e^E E \tag{1}$$

$$S = s^E T + d_t E \tag{2}$$

where D—electric displacement, E—electric field, S—mechanical strain, T—applied mechanical stress, e^E—permittivity matrix at continuous mechanical strain,

s^E—matrix of elasticity under environments of constant electric field and d_i—piezo-electric coefficient matrix. In this proposed model of the cantilever, it is used in two altered modes as 31 modes and 33 modes. In 33 mode, energy is produced in three directions, corresponding to the way of functional power which is named as compressive energy.

In 31 mode, power is produced in one direction, perpendicular to the direction of useful power which is named as transverse mode. In this work, the 31 modes will be used for power generation purpose.

$$S = sT \tag{3}$$

where s—is the compliance vector of silicon substrate, S—is the mechanical strain vector and T is the mechanical stress vector.

3 Model of Piezoelectric Energy Scavenger

Williams and Yates [2] established the standard model constructed on the kinetic energy presented in Fig. 1. It is a second-order dynamic structure which associated the input vibrations $y(t)$ to the output displacement $z(t)$.

By using the dynamic equation of the D'Alembert's law, it is given as:

$$m\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} + b\frac{\mathrm{d}z}{\mathrm{d}t} + kz = -m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} \tag{4}$$

where m—mass, b—damping coefficient and k—spring constant. While there is damping of the structure, because of the damper, there is remaining transfer of mechanical power into electrical power. For sinusoidal excitation as

$$y(t) = Y \sin \omega t$$

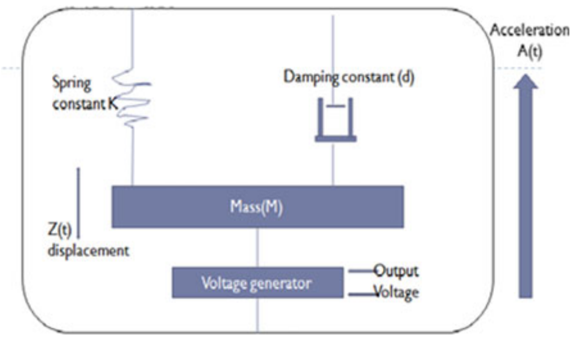


Fig. 1 Energy harvester generic model

The generated power is considered as

$$p(w) = \frac{m\zeta_T Y^2 \left(\frac{w}{w_r}\right)^3 \omega^3}{\left(1 - \left(\frac{\omega}{\omega_r}\right)^2\right)^2 + \left(2\zeta_r \frac{\omega}{\omega_r}\right)^2} \tag{5}$$

By resonant frequency, the power harvested from the model is maximum and is defined as

$$P = \frac{ma^2}{8\omega_r}.Q = \frac{F.a}{8\omega_r}Q \tag{6}$$

where $a = Y\omega^2$, Q is the quality factor and acceleration functionality of the piezoelectric material correspondingly. Since Eq. 6, it has been detected that the energy harvested is proportional to the power applied and quality factor of the typical and likewise inversely proportional to the resonant frequency the proposed cantilever beam. Thus, the lesser resonant frequency of the cantilever is desirable for higher energy construction.

4 Model of Piezoelectric Energy Scavenger

See Figs. 2 and 3.

These two proposed cantilever models are having the piezoelectric layer as of the PZT (lead zirconate titanate) and the substrate thin film as of the silicon bulk material. The mechanical properties of materials and all these models of tip mass geometry and silicon bulk materials are planned in Table 1. The dimensions of length, width

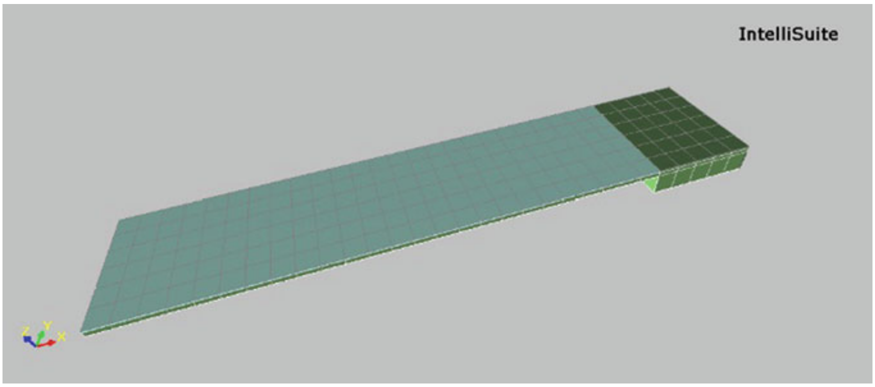


Fig. 2 Proposed rectangular tip mass cantilever

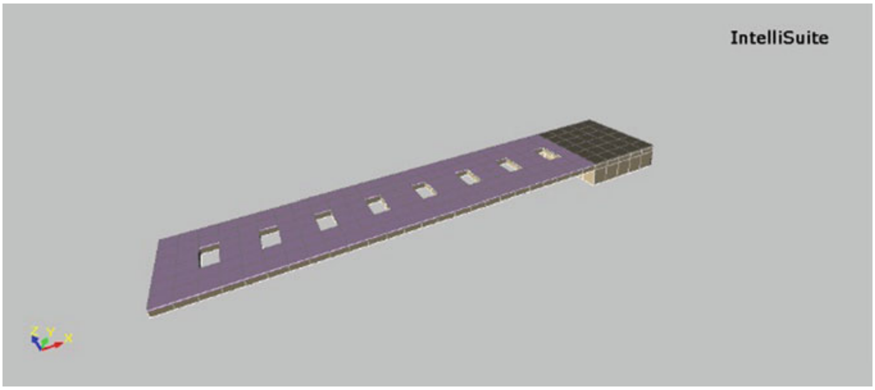


Fig. 3 Proposed rectangular with holes tip mass cantilever

Table 1 Properties of material

Material	Young's modulus (Gpa)	Poisson's ratio	Density (gm/cm^3)	Coefficient of piezoelectric material (d311, d322, d333)
Silicon	170	0.26	2.3	–
ZnO	10,000	0.226	7.5	−0.0001, −0.0001, 0.0003

and thickness of the cantilever are planned in Table 2. Then, the cantilever with the hole of the complete models is diverse to lower the resonant frequency. With respect to this geometry, these models are designed (Fig. 4).

Table 2 Dimensionsof cantilever and tip mass

Design parameters	Material	Dimensions (μm)
Substrate length, *l*	Silicon	250
Substrate width, *w*	Silicon	250
Substrate thickness, *t*	Silicon	2
Piezoelectric Length, *l*	ZnO	200
Piezoelectric width, *w*	ZnO	200
Piezoelectric thickness, *t*	ZnO	1
Length of tip mass Si, *l*	Silicon	50
Width of tip mass Si, *w*	Silicon	60
Thickness of tip mass Si, *t*	Silicon	10

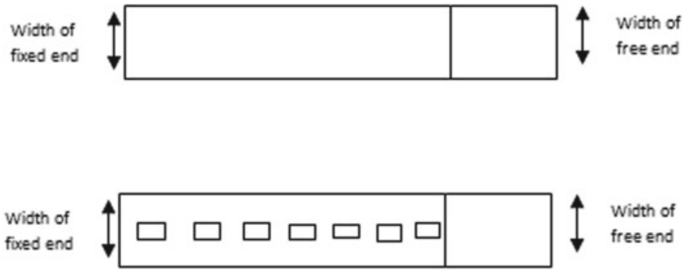


Fig. 4 Schematic diagrams of proposed models

5 Mathematical Modeling of Proposed Structure

5.1 Resonant Frequency of an Unimorph Cantilever Beam

Piezoelectric cantilever beam has divided into three types such as parallel configurations, unimorph and bimorph series. While beam has introduced for the silicon substrate layer. Formerly, it is defined as a unimorph cantilever beam.

The resonant frequency of nth mode of the cantilever without proof mass is [9, 10]

$$f_n = \frac{v_n^2}{2\pi l^2}\sqrt{\frac{EI}{m'}} \tag{7}$$

In expressions of bending, modulus per unit width is defined as

$$f_n = \frac{v_n^2}{2\pi l^2}\sqrt{\frac{D_p}{m}} \tag{8}$$

D_p is the role of the young's moduli of two materials as E_p (PZT) as young's modulus of piezoelectric material, E_s (Si) as young's modulus of silicon substrate and their thickness as t_p and t_s,

where $D_p = \frac{E_p^2 t_p^4 + E_s^2 t_s^4 + 2E_p E_s t_p t_s (2t_p^2 + 2t_s^2 + 3t_p t_s)}{12(E_p t_p + E_s t_s)}$, then the mass per unit area m is intended from the thickness and densities as ρ_s (silicon substrate) and ρ_p (piezoelectric layer) is given by

$$m = \rho_p t_p + \rho_s t_s$$

The resonant frequency with silicon tip mass is given by

$$f_n = \frac{v_n'^2}{2\pi}\sqrt{\frac{K}{m_e + \Delta m}} \tag{9}$$

where $\nu_n'^2 = \nu_n^2\sqrt{\frac{0.236}{3}}$ and the effective cantilever mass is given by $m_e = 0.236mwl$, K is the effective spring constant and Δm is the mass of tip mass of the cantilever beam and is given by

$$K = \frac{3D_p w}{l^3}$$

For distributed tip mass at the end of the proposed cantilever beam, the center of the tip mass is $\frac{l_m}{2}$ distance from the free end. The effective spring constant is given by

$$K' = K\left(\frac{l}{l - \frac{l_m}{2}}\right)^3$$

The resonant frequency of the unimorph cantilever with tip mass at the end of the beam is given by

$$f_n' = \frac{\nu_n'^2}{2\pi}\sqrt{\frac{0.236D_p w}{\left(l - \frac{l_m}{2}\right)^3(0.236mwl + \Delta m)}} \quad (10)$$

5.2 Energy Relations of the Unimorph Piezoelectric Microcantilever

The charge accumulated in the piezoelectric layer is given as [11]

$$Q = \frac{-3d_{31}s_s s_p t_s(t_s + t_p)l^2 F}{X} \quad (11)$$

where stiffness constant for silicon and piezoelectric layer is related as

$$s_s = \frac{1}{E_s} s_p = \frac{1}{E_p}$$

The voltage generated on the piezoelectric layer of the unimorph cantilever is defined as

$$V = \frac{-3d_{31}s_s s_p t_s t_p(t_s + t_p)lF}{\varepsilon_{33}^T wX\left(1 + \left(\frac{3s_p^2 s_s t_p t_s^2(t_p+t_s)^2}{s_h X} - 1\right)K_{31}^2\right)} \quad (12)$$

And the potential energy generated in the system is defined as

$$U = \frac{-9d_{31}^2 s_s s_p^2 t_s^2 t_p (t_s + t_p) l^3 F^2}{\varepsilon_{33}^T w X^2 \left(1 + \left(\frac{3s_p^2 s_s t_p t_s^2 (t_p + t_s)^2}{s_h X} - 1\right) K_{31}^2\right)} \quad (13)$$

where

$s_h = s_s t_p + s_p t_s$
$X = s_s^2 t_p^4 + 4 s_s s_p t_s t_p^3 + 6 s_s s_p t_s^2 t_p^2 + 4 s_s s_p t_p t_s^3 + s_p t_s^4$
K_{31} is the coupling coefficient of the piezoelectric materials.

6 Results and Discussions

Since the cantilever is fixed at one end, the mechanical relations are completed by setting the vertical aspects of one adjacent of the cantilever, whereas all further surfaces are left forced and the electrical relations are completed by fluctuating the piezoelectric thin film and by preparation the substrate of thick layers to couple the voltage generated. Giving to the essential boundary conditions, l these two samples are considered and simulated in the FEM-based IntelliSuite software (Figs. 5, 6 and 7).

A. Analysis of Resonant Frequency

Energy harvesters of piezoelectric cantilever are very capable when these are determined at the resonant frequency, and the voltage generation will be extreme. It is detected that the rectangular with hole cantilever model consuming the lowermost resonant frequency. Therefore, energy producers will be extra related to other models.

B. Displacement Versus Load

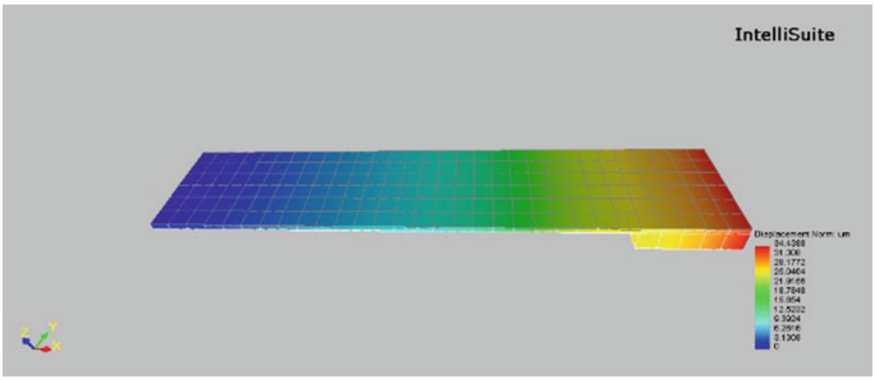


Fig. 5 Simulation results of displacement analysis

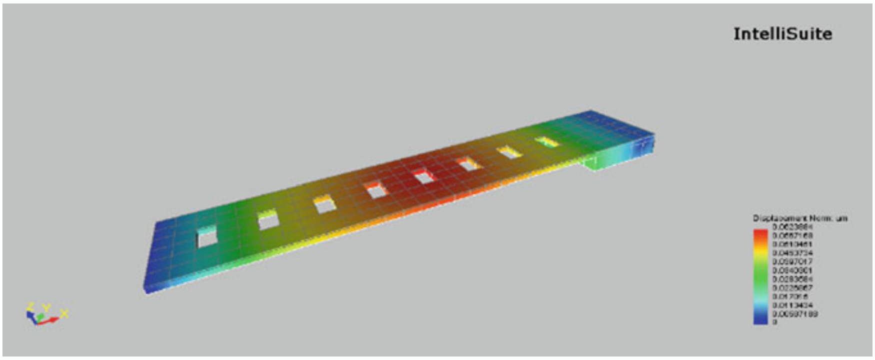


Fig. 6 Simulation results of displacement analysis

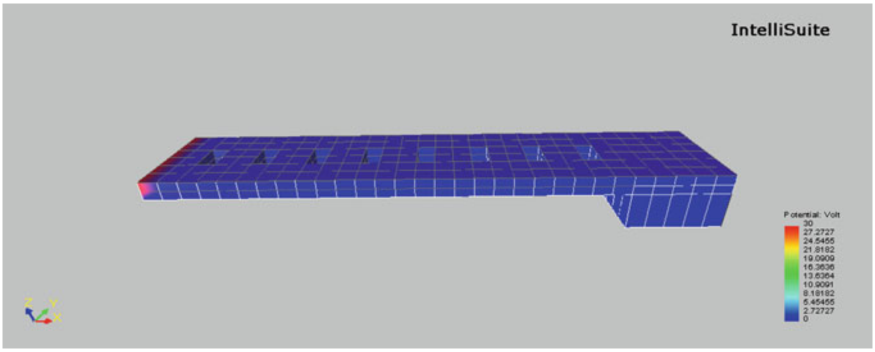


Fig. 7 Simulation results of voltage analysis

Subsequently, the piezoelectric thin film due to strain harvests extra charges on its outward. Hence, determinaion of displacement can be found using corresponding pressure weight. This analysis has been performed (Fig. 8).

It is observed that the displacement produced on the rectangle with hole piezoelectric surface layer is low as nearly as 0.028E−6 compared to other models.

C. Generation of Voltage Versus Load Applied

With esteem to obtainability of load in the nearby shaking sources, two different geometries of the model are being experienced altered pressure range from 1e−5 Mega Pascal to 1e−1 Mega Pascal [9], and likewise, it is detected that the rectangle with hole cantilever beam is generating more voltage related to others in organized the cases of the amount of load applied on it. For example, the load rises, and the strain bent in the rectangle with hole beam is more which in turn harvests moment of inertia (Fig. 9).

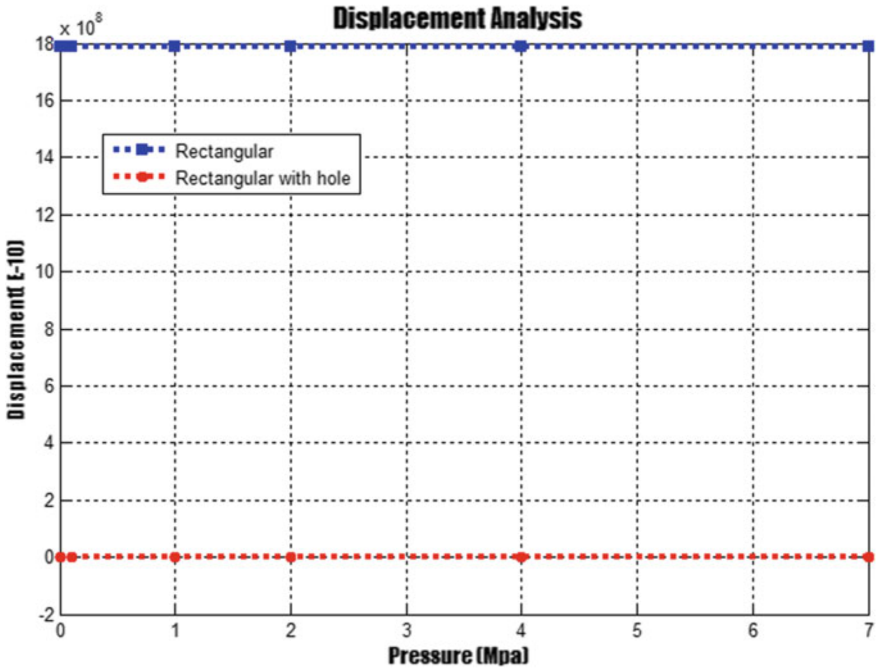


Fig. 8 Displacement of simulated energy scavengers

From all these analyses, it is observed that among all cases, the rectangular with hole beam with a bulk at the free end having all the characteristic to harvest the most amount of energy from all surrounding vibrations (Table 3).

7 Conclusions

The MEMS energy scavenging traditional cantilever with a tip mass is intended to use a 3D builder of IntelliSuite software. The dimensions of the tip mass and same piezoelectric material such as PZT are evaluated by setting the essential boundary conditions. In this proposed designed geometry was used for the analysis of displacement, pressure, voltage for the harvesters. Therefore, it is lastly determined that the MEMS reversed rectangular with hole harvests extra energy and strain in all the surroundings of obtainability of nearby load with evaluation of more respective piezoelectric cantilever models. Other, these proposals likewise will be used for traditional of repeatedly tuned energy harvesters.

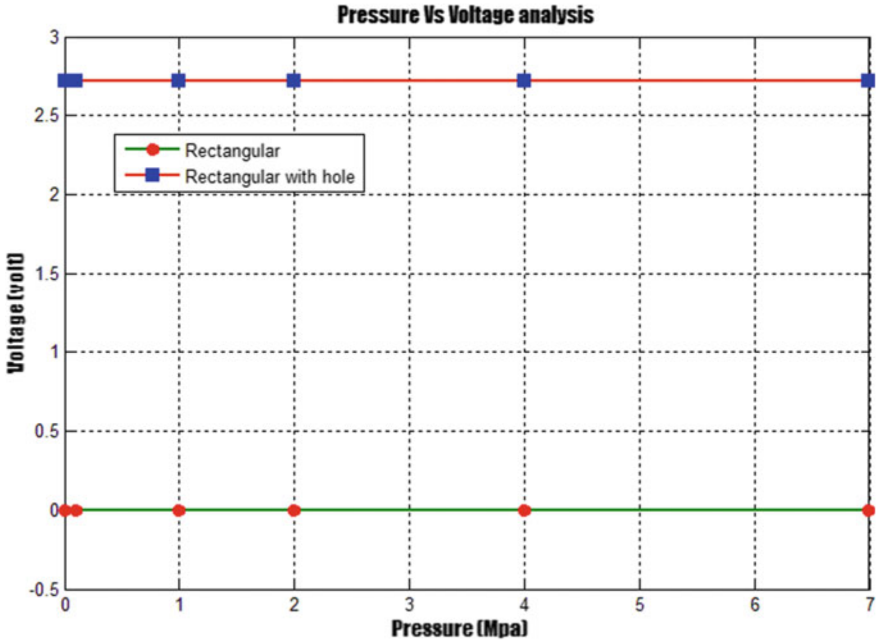


Fig. 9 Energy produced versus load applied

Table 3 Displacement and energy produced versus load applied

Pressure load (MPa)	Displacement	
	Rectangular model 1	Rectangular with hole model 2
1e−1	0.239	0.0287
Pressure load (MPa)	Energy production	
	Rectangular model 1	Rectangular with hole model 2
1e−1	−0.0496 V	2.728 V

Acknowledgements The authors would like to acknowledge National Program on Micro and Smart Systems (NPMaSS), Government of India, for providing the MEMS software facilities at Department of Electronics Engineering, Pondicherry University, Pondicherry.

References

1. Ali W, Ibrahim S (2012) Power analysis for piezoelectric energy harvester. Energy Power Eng 4(6):496–505
2. Sunithamani S, Lakshmi P, Eba Flora E (2012)Simulation and optimization of a MEMS piezoelectric energy harvester with nontraditional geometries. In Comsol conference, Bangalore
3. Alpurerturk, Inman DL (2011) Piezoelectric energy harvesting. Wiley Ltd., publication
4. Chen ZS, Yang YM, Dong GQ (2009) Analytical and experimental study of vibration energy harvesting behaviours of piezoelectric cantilevers with different geometries. In: IEEE conference of sustainable power generation and supply, SUPERGEN, pp 1–6
5. Yang K, Li Z, Chen D (2012) Design and fabrication of a novel T shaped Piezoelectric Zno canitilever sensor. J Act Passive Electron Compon 2012, Article ID 834961
6. Gogoi UJ, Shanmuganantham T (2014) Sensitivity enhancement of MEMS fluid flow rate and flow direction sensor. In: IEEE international conference on smart structures and smart systems, Chennai
7. Matova SP et al (2013) Effect of length/width ratio of tapered beams on the performance of piezoelectric energy harvesters. Smart Mater Struct 22:075015
8. Liu H et al (2012) A new S shaped MEMS PZT cantilever for energy harvesting from low vibrations frequency below 30 Hz. J Microsyst Technol 18(4):497–506
9. Blevins RD (2001) Formulas for natural frequency and mode shape. Krieger Publishing Company, Malabar, p 158
10. Yi JW, Shih WY, Shih W-H (2002) Effect of length, width, and mode on the mass detection sensitivity of piezoelectric unimorph cantilevers. J Appl Phys 91:1680–1686
11. He Q, Yan Z, Fei LK, Song B (2012) Modelling and analysis of piezoelelctric vibrating generator of cantilever. Appl Mech Mater 148–149, 1327–1330
12. Ajitsaria J, Choe SY, Shen D, Kim DJ (2007) Modeling and analysis of a bimorph piezoelectric cantilever beam for voltage generation. Smart Mater Struct 16(2):447
13. AzizollahGanji B, Taybi M (2014) Modelling of resonance frequency of MEMS corrugated diaphragm for capacitive acoustic sensors(technical Note). IJE Trans C Aspects 27(12):1850–1854
14. Gogoi UJ, Shanmuganantham T (2015) Energy harvesting of cantilever with silicon tip mass based MEMS energy scavengers. In IEEE international conference on electronics and communication systems, ICECS 2015, Coimbatore, India, 26–27 Feb 2015
15. Nallathambi A, Shanmuganantham T (2018) Study and analysis of MEMS based flow rate and path sensor system. In: 2018 IEEE international conference on circuits and digital enterprises technology (ICCSDET 2018), Saintgits, Kottayam, Kerala
16. Gogoi UJ, Shanmuganantham (2015) Energy harvesting of cantilever with silicon tip mass based MEMS energy scavengers. In: IEEE international conference on electronics and communication systems, ICECS 2015, Coimbatore, India, 26–27th Feb

Challenges and Impacts of RFID Technology in a Research Library

V. Sivasankar, E. Soundararajan, and S. Rajeswari

Abstract Libraries are constantly looking for a solution to deliver fast and efficient information services and workflows, which brings a high level of satisfaction for both patrons and library staff. The library needs to constantly evolve with rapid changes in the field of information technology. As a result, digital library services are taking over the traditional library services. In this context, the radio frequency identification technology (RFID) is used for smooth circulation management in the library besides ensuring the security of the books. Also, the application of RFID in the library field will save the time of patrons and library staff. RFID technology helps libraries to improve user satisfaction with a self-service experience and increases staff efficiencies by augmenting several value-added services. This paper highlights the comprehensive application of RFID technology in the library domain with a case study of migrating RFID infrastructure implemented at the Scientific Information Resource Division (SIRD), Indira Gandhi Centre for Atomic Research (IGCAR) at Kalpakkam. The various challenges faced while implementing the technology and the impact of RFID technology in the library are highlighted.

Keywords RFID · Barcode · ILMS · Digital library

1 Introduction

The fifth law of library science by Dr. S.R. Ranganathan states that 'library is a growing organism,' indicates libraries are changing continuously [1]. Scientific Information Resource Division (SIRD) at Indira Gandhi Centre for Atomic Research

V. Sivasankar (✉) · E. Soundararajan · S. Rajeswari
Indira Gandhi Centre for Atomic Research, Kalpakkam, India
e-mail: sivav@igcar.gov.in

E. Soundararajan
e-mail: sound@igcar.gov.in

S. Rajeswari
e-mail: raj@igcar.gov.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_47

(IGCAR) is constantly adapting to state-of-the-art advanced digital library infrastructure in all aspects including its circulation system. It has implemented barcode technology in the year 2002 and later migrated to RFID technology in 2006 for delivering fast and efficient workflows.

RFID technology solutions are receiving much attention for meeting varied applications ranging from industrial automation, pharmaceutical/health care, airline baggage, airplane parts, animal tagging, passports, auto tires access control, supply chain, and library transaction management. It can be used in all areas of automatic data capture, allowing contactless identification of objects.

RFID is based on radio frequency and microchip technology which uses radio waves for transmission between the tag and the reader [2]. Recently, SIRD has commissioned the latest RFID-based integrated library management system with new RFID book tags, tagging stations, handheld readers, electronic article surveillance (EAS) gates, self-check-in/check-out kiosks, etc. This paper highlights the challenges and impacts of RIFD technology in a research library.

2 Literature Survey

Gupta et al. examined the theoretical and practical aspects of the implementation and use of RFID technology in libraries, including the benefits, pros, cons, challenges, and problems related to RFID application [3]. Yusof et al. formalized a formal innovation-decision framework is used to analyze the literature on the use of RFID in libraries [4]. Shamshul bahri et al. developed a process model of emerging IT implementation in a library. The model is divided into three phases: before, during, and after the installation. It consists of ten activities, starting from soliciting requirements until enhancing the RFID system [5]. John Chellaiah et al. proposed the best practice model through the insights gained by the people involved in the RFID implementation [6].

3 History of Research Library at IGCAR

SIRD caters to the information resource needs of scientists and engineers of IGCAR. SIRD was started in the year 1972 with a few thousand books and subscriptions to a dozen journal titles. Over the period, the holdings of SIRD increased to about 60,000 books, 220,000 reports, and 800 periodicals.

The advancements in information technology in both hardware and software are being utilized for providing services to patrons with enhanced quality. The barcode system was initially used for data acquisition for library transactions. Now with the advent of technology, the scope of collection management, circulation, and security management of the documents migrated to RFID-based systems. RFID is an identification technology and offers more functionality including EAS, high-speed inventory management, and shelf management.

4 RFID-Based Integrated Library Management System at IGCAR

The radio frequency identification (RFID)-based integrated library management system (ILMS) which was commissioned in SIRD during the year 2011 had become obsolete and started showing symptoms of failure. Moreover, the RFID book tags which were pasted on books during the year 2006 are more than a decade old and have crossed their ideal retention life of 10 years. Hence, it was proposed to enhance SIRD with state-of-the-art RFID infrastructure and open-source integrated library management system (ILMS).

The proposed system is a comprehensive solution that included open-source Koha integrated library management software (ILMS), RFID tags with enhanced memory, lifetime, and latest RFID gadgets with value-added features. The scope of the work also includes migration of the complete data from the existing LibSys to Koha, retagging of all SIRD holdings, and performing a stock verification.

5 RFID-Based Library Schema

The layers of the RFID-based Library schema are shown in Fig. 1. The schema comprises five layers, namely the physical layer, an interface layer, the transportation layer, the processing layer, and the application layer. The proposed schema is an altered architecture from Ustundag et al. [7]. The functionality of each layer is described in detail in the following subsections.

Fig. 1 Layers of RFID-based library schema

5.1 Physical Layer

The physical layer involves the attachment of RFID tags to books, back volumes, multimedia cases, and identification/membership cards of the patrons. The commonly used tag in libraries is a high-frequency (HF) passive RFID tag. It is used to store the unique accession number of the book/item. The patron identification card is also an RFID card that holds a library membership number information besides the identification number of the patron.

5.2 Interface Layer

This is the layer that interacts with the physical layer; it captures real-time data for the system. The RFID writer is used to write the patron's library membership number to the MIFARE employee card. The handheld RFID reader, self-check-in/check-out terminals, electronic article surveillance gate are equipped with fixed antennas to read accession number from the RFID tags fixed in the library books/items. Handheld readers are used for shelving, inventory, management, and locating the books.

5.3 Communication Layer

The Standard Interchange Protocol version-2 (SIP2 protocol) is used for communication between library management software and self-service circulation terminals. In SIP2, requests to perform operations are sent over a connection, and responses are sent in return [8]. There is no middleware used for communication between RFID gadgets and library management system software.

5.4 Processing Modules

In this layer, the data captured through RFID readers are stored in the relational database systems such as MySQL or PostgreSQL. Structured Query Language (SQL) queries are used to process the data with the help of integrated processing modules such as check-in/check-out, inventory, authorizing, book record, user record, and reporting.

5.5 *Application Layer*

This is the topmost layer, where the system administrator or the patrons directly interacts with the integrated management system. It is a user-friendly interface for administrators to process the library management modules. The application layer coordinates to integrate the processing modules and manage business logic with the system applications and the library information management system.

6 Functionality and Components of RFID-Based ILMS at SIRD

The business process of RFID-based ILMS is shown in Fig. 2. It consists of six core components such as user management, book tagging, self-check-in/check-out, electronic article surveillance, inventory and shelf management, and ILMS administration. The ILMS components are interconnected through the local area network (LAN) with the Koha open-source library management software. Each component is described in the following sections:

- User management console

The employee identification card is programmed with the patron's library membership number using an RFID writer, and a new patron profile is created in the Koha ILMS. A dedicated work station embedded with an in-house developed application is used to write a membership number on the patron cards.

- Book tagging station

RFID tags have integrated microchips with an inbuilt aluminum antenna used for communication. Library staff can tag RFID labels with necessary information using

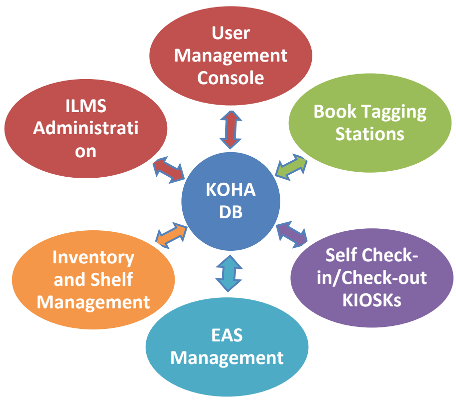


Fig. 2 Workflow of RFID-based ILMS

Fig. 3 Book tagging station and KIOSK

staff connect conversion software; transactions at staff desk with the ILMS using staff connect circulation software. The book tagging station attached to staff workstation and kiosk is shown in Fig. 3.

- Self-check-in/check-out KIOSKs

KIOSKs are used for self-check-in/check-out of books by the users, and it is based on touchscreen technology with a mini PC and thermal slip printer, as shown in Fig. 3. The bibliotheca quick-connect self-check software is used to do the circulation process. KIOSKs are fixed with separate RFID readers to read both book tags and patron ID cards, which are based on ISO 15693 and ISO 14443 standards, respectively.

- Electronic Article Surveillance (EAS) Gate Management

The EAS gates detect the theft bits of RFID tags in any orientation and give the audio and visual alarm as shown in Fig. 4. It fully supports tags based on ISO 15693, ISO 18000-3-A standards. It has the transmitting power up to 4 W maximum. RFID has an operating frequency of 13.56 MHz. Besides the surveillance, these gates can also record the number of people visiting the library through the infrared (IR) sensors installed within them.

Fig. 4 EAS and inventory management

- Inventory and shelve management

RFID handheld reader is used for library collection management including shelf rectification, re-shelving, sorting, searching, weeding, and exception-finding. It enables the staff to perform the inventory and stock verification quickly as shown in Fig. 4. Bibliotheca staff connect data software is used to do the above activities.

- ILMS Administration

The administration module enables library administrators to manage circulation records, setup library policies, perform inventory checking, etc., and also, it provides various management and statistical reports to the management.

7 RFID Tagging and Commissioning at SIRD

SIRD uses a passive tag based on ISO 15693 Standard with 13.56 MHz frequency. This is a vicinity tag with a detection range of up to few centimeters. These tags are affixed in the inside rear cover of the books as shown in Fig. 5.

Each tag has 2.5 K memory with a built-in antitheft detection bit. The accession and call numbers of the book are stored in the tags. The accession number is a primary key used to retrieve the data stored in the KOHA database.

The programming and affixing of RFID tags are a voluminous job for about 60,000 books. A strategic plan was formulated to complete the job in less than a month. The RFID tagging process was divided into three phases, namely pre-RFID programming, RFID programming, and post-RFID programming as shown in Fig. 6. In pre-RFID programming process, the books are de-shelved from the rack. The paper packet from the back cover, which is used to hold the manual issue card, is removed and cleaned. RFID programming process includes deactivation of old tag, pasting of new RFID tag, programming the accession number to the new tag, and pasting of security label. This process also includes checking of correctness of programmed data against accession number and title of the item. The final process is post-RFID programming process, which includes fixing of quality control sticker, the Universal Decimal Classification (UDC) on books and shelving the items back to the rack.

Fig. 5 RFID tagging

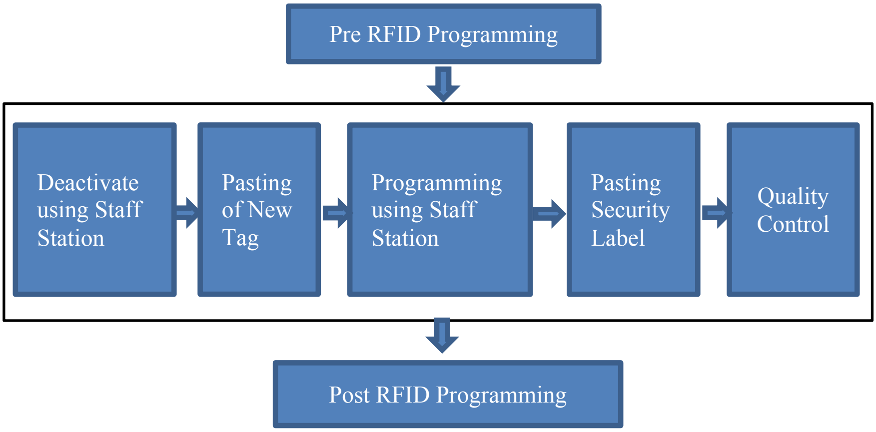


Fig. 6 RFID tagging process

8 Challenges Faced While Implementing RFID

The RFID tags can store several data fields about the document. As in the case of any data entry process, care must be taken to ensure the correctness of the bibliographic data.

8.1 Tagging Issues

The initial investment for RFID is high due to the infrastructure, tags, associated management software. The investment also varies with the collection size.

Using the library management software, the document data are uploaded into tags. Tags may fail due to several reasons such as high electrical resistance likely to develop between the chip and the antenna, a crack in the microchip, and failure in antenna due to metal fatigue. Library items are often bent, twisted, impacted, subjected to pressure, and as a consequence, the tags attached to them experience mechanical stress. Mechanical stress is one of the prime reasons for failure. To overcome this, RFID tags can be pasted with the lamination to ensure durability. Environmental exposure can also cause deleterious effects on tags, producing corrosion, cracking, or other damages.

RFID tag fails gradually with a modest or a severe reduction in the read range. The tag can sometimes suddenly become unresponsive with no warning. Tags may also fail due to less frequency of usage. To resolve this problem, periodical activation of the tags for their seamless functioning has to be done. There are also issues faced while re-fixing the used tags. For instance, re-binding of the damaged books requires the old tag to be reaffixed, and mostly, this will lead to detection problems or failure.

Fig. 7 RFID tag fixing locations

Hence in such cases, fresh tags are to be placed. To enhance the data retention period of tags, it is required to provide proper preservation strategies such as adequate air-conditioning, periodical activation, and proper handling of tags.

8.2 Circulation and Inventory Issues

The detect range for the reader used at times may not cover all the books placed at a time and thus leads to non-detection. In such cases, it is required to put back the non-detected books for check-in/check-out. This requires special supervision by the counter-personnel. During the inventory, in case of a non-detection of the books by RFID stock verification device, there is a need to do a multi-pass scan through the rack. Placing two tags close together can lead to interference of signals and affects the read capability of the scanner.

To overcome this, the tags are strategically placed in SIRD at different locations inside the rear cover page, such that no tags are fixed at the same orientation and location for every consecutive three books as shown in Fig. 7.

8.3 Security Issues

The RFID tags used in libraries do not have restrictions in reading the information, such as patron details, title, author, or last check-in/check-out time. In rewritable tags, a preventive measure should be taken to avert adversaries erasing the tag data, changing the security status of tags with security bits. In SIRD, the encryption methodology has been implemented for protecting the unauthorized accessing of tags data. RFID security pedestals need to be integrated with the existing security components such as closed-circuit television (CCTV) and logging systems to zero down the book loss. Also, special software is integrated with RFID gate to monitor unauthorized check-outs of books through pop-up messages to the librarian's desktops.

9 The Impact of RFID Technology at SIRD

User satisfaction levels can be measured in terms of quality, availability, and efficiency of services. For efficient library operation, the following functions are important.

9.1 Stock Verification and Shelf Management Analysis

One of the most important requirements for the library is to carry out stock verification periodically. For libraries with large collections, it is a voluminous task and consumes intensive work and time. Earlier barcode-based stock verification had helped SIRD to improve efficiency and reduced the time considerably. It required every book to be manually taken out of the rack during the verification process. Also, the library transactions were suspended during stock verification. In the case of RFID, the verification process is carried out with the help of a handheld inventory reader. It is rapidly moved across a shelf of books to read the unique identification information attached to them. As a result, stock verification can be carried out at a much faster rate with less workforce in SIRD.

SIRD has 73 racks with 30 shelves each. Each shelf can hold for about 25–30 books approximately. The time taken for inventory exercise carried out for the collections using barcode was 170 h (with approximately 10 s per book) and using RFID was 34 h (with approximately 2 s per book). This implies that the average scanning time in inventory management is reduced by 80% by the RFID technology, as shown in Fig. 8. Similarly, as shown in the same figure, the man-days required for the inventory exercise are found to be 22 days and 5 days for barcode and RFID, respectively. The RFID-based technology was also used to identify misplaced books, reordering of the shelf, etc.

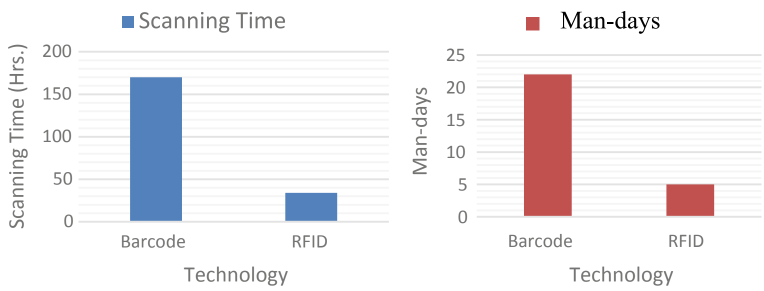


Fig. 8 Barcode versus RFID: inventory scanning time and man-days

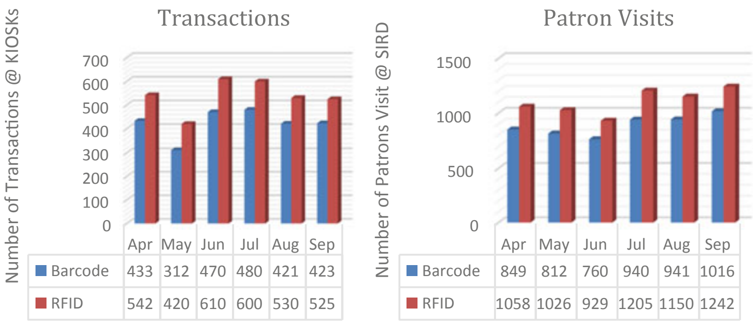


Fig. 9 Barcode versus RFID transactions and patrons' visits

9.2 Counter-Transaction Productivity Analysis

Productivity analysis can be defined as enhancement in the circulation process, resulting in the improvement in the user satisfaction level with the reduced requirement of the workforce. Barcode-based circulation necessitates the staff to be present and also takes more of the time of the user. This method also necessitates security monitoring to be done manually. RFID technology provides better self-service capabilities by reducing the amount of time required for circulation operations. One way or the other, patrons spend less time waiting in check-out lines. Besides, technologically advanced services enhance the satisfaction levels of the patrons toward the library. Not only does self-service reduce check-out wait times and also reduces the requirement of staff for manning the counters [9].

The SIRD can achieve enhanced user satisfaction and reduce the workload at the circulation counter through RFID technology. After self-check-in and check-out stations in SIRD were made accessible to the users, the satisfaction level increased dramatically. The average number of transactions during six months with barcode technology and with RFID technology is compared and found to be 20% (approximately) increase with RFID as shown in Fig. 9. Similarly, as shown in the same figure, there is also a significant increase (more than 19%) in patrons' visits after RFID implementation.

9.3 RFID-Based Member Cards

ISO 14443 employee ID card of Indira Gandhi Centre for Atomic Research has provision for storing the library information in specific sectors allocated for this purpose. The need for a separate RFID card has been avoided by integrating the library member details into the employee ID card. A USB-based in-house RFID reader has been developed to read the library membership information and communicate it

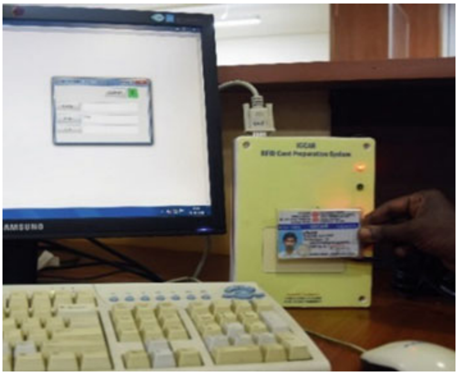


Fig. 10 Employee ID tagging

to the library self-check-in and check-out software. The programming of library membership number with employee ID card is shown in Fig. 10.

The integration of member's information to employee ID card increased the user satisfaction level as it eliminates the need to carry a separate library membership card.

9.4 Generating Management Information Reports

The use of RFID technology opens up many advantages to the library community, and one such is the ability to generate various management reports. The books that have been referred or read by the users can be scanned and stored into the database using the handheld RIFD reader before it is re-shelved. This data becomes a base for knowledge management initiatives in the library. The book usage detection on carrels using RFID handheld devices is shown in Fig. 11.

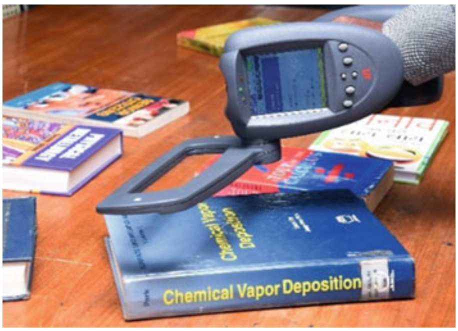


Fig. 11 Book usage detection

Table 1 Book usage detection on carrels

Top five books accessed at carrels during 09/02/2019–09/06/2019 week			
S. no.	Titles	Authors	Access count
1	Signal processing for radiation detectors	Nakhostin (Mohammad)	9
2	Pulsed laser deposition of thin films	Chrisey (D. B.)	8
3	Fundamentals of 5G mobile networks	Rodriguez (Jonathan) ED	8
4	Introduction to nuclear engineering	Lamarsh (John R.)	7
5	Let us C	Kanetkar (Y. P.)	7

The data captured through the RFID device is imported into the database, and it is analyzed further using SQL queries. A sample report on book usage on carrels during a week at SIRD is listed in Table 1.

10 Summary and Conclusion

RFID is boon for many areas of application including library as it provides several advantages due to technological advancements. SIRD at IGCAR, Kalpakkam, has effectively implemented the RFID technology for providing efficient services to the users. The RFID-based library management system commissioned at IGCAR increased the speed of the circulation system with little or no manual intervention of bookkeeping. The methodology, issues, and impacts related to implementing RFID enabled library were discussed. The integration of employee ID cards with RFID and the customized reports on usage statistics are the unique features of this project. IGCAR is one of the libraries in India to integrate RFID technologies with open-source Koha ILMS using SIP-2 protocol without any commercial middleware.

References

1. Barner K (2011) The library is a growing organism: Ranganathan's fifth law of library science and the academic library in the digital era. Libr Philos Pract 2011
2. Finkenzeller K (2010) RFID handbook: fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication. Wiley. https://doi.org/10.1002/9780470665121
3. Gupta P, Madhusudhan M (2017) RFID technology in libraries: a review of literature of Indian perspective. DESIDOC J Libr Inf Technol 37:58–63
4. Yusof MK, Saman Y (2016) The adoption and implementation of RFID: a literature survey. LIBRES Libr Inf Sci Res Electron J 26:31
5. Bahri S, Ibrahim A (2013) RFID in libraries: a case study on implementation. Libr Hi Tech News 30:21–26

6. Chelliah J, Sood S, Scholfield S (2015) Realising the strategic value of RFID in academic libraries: a case study of the University of Technology Sydney. Aust Libr J 64:113–127
7. Ustundag A, Kilinc MS (2010) Design and development of RFID based library information system. In: WMSCI 2010—The 14th world multi-conference on systemics, cybernetics and informatics, proceedings, vol 2, pp 234–238
8. 3M Standard Interchange Protocol (3M, 2006)
9. Haley CK, Jacobsen LA, Robkin S (2007) Radio frequency identification handbook for librarians

An Efficient Multiauthority Attribute-Based Encryption Technique for Storing Personal Health Record by Compressing the Attributes

F. Sammy and S. Maria Celestin Vigila

Abstract Personal health record system allows patient to share their health information in a remote location. Security of the data is an important criterion to be considered while outing the facts in the remote servers. Encryption is the common technique to provide the security for the data before outsourcing. This paper focus on multiauthority attribute-based encryption (MAABE) approaches, by compressing the least value attributes. In all the attribute-based encryption, the attributes are used to encode and decode the data. As the quantity of attributes used increases, the calculation time will be increased for the performance. So, in order to increase the system efficiency, least value attribute details should be compressed before encoding and decoding the information. Also, a binary tree structure is used to administer the attributes in various authority levels of the algorithm. This binary tree structure helps to decrease the encoding and decoding timing of the algorithm.

Keywords Multiauthority attribute-based encryption · Efficiency · Compression

1 Introduction

The growth of public websites and their numerous amounts of private data are stockpiled over the websites wants to be protected. Keys were used in the past cryptosystems which are required to be common so that another party can use the same key for encoding or decoding the data. But, in the novel cryptography algorithms, attributes of the customers are worn to encode and decode the records. It has been familiar in grounds such as social network and cloud storage, etc. Correlated to the past encryption algorithms, ABE allows better control over the access structure and make easier to develop an elastic access control structure.

F. Sammy (✉) · S. M. C. Vigila
Department of Information Technology, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu 629180, India
e-mail: fvr.sammy@gmail.com

S. M. C. Vigila
e-mail: celesleon@yahoo.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_48

Vigila and Muneeswaran have proposed a key generation algorithm using elliptic curve. The system is providing security against cryptanalytic attacks, demographic, interrelated, key receptive and computational attacks [1].

In recent years, researchers concentrate their research work in developing the algorithms for securing the details of the users which are private, with the help of their own user's attributes. This yields to the generation of different forms of algorithms based on fuzzy logic, ciphertext, key policy and role-based by using the attributes of the users. All these algorithms mentioned above are generated, in aspects to offer the safekeeping of the valuable details of information, when they are stored in the third-party server location [2].

The next stage of development results in providing an access control structure along with the algorithm, by providing access rights to view the data by a particular user using the logic gate called AND [3]. The next stage yields to the development of MAABE; here, more than one authority is used to control the data. This is an effective algorithm, but having a problem of collusion attack. To solve this problem, researchers introduce a technique of allocating universal identifier (UID) to all users. This can be achieved by allocating a Central Authority (CA), (or) more than one CAs, (or) without any CA [4, 5]. All these algorithms were developed to provide the security of the data. Additional related work will be cited in the upcoming sections.

Grounded on overhead research, the context recommends an algorithm based on multiple authority arrangement to explain the tricky of authorization in the scheme using AND gate which suits in real-world application correctly, e.g., Aadhaar Card in India. The Aadhaar Card is used for multiple purposes like banking, insurance, etc. The authority for banking is different from insurance. The keys are generated based on the attributes needed for their purposes. That is, when we are using banking services, the key is generated based on the attribute needed for banking. In the same way, the insurance company generates the key based on their attributes. Only the needed attributes are used to generate the Key. This effectually decreases the computational cost and reduces the ciphertext's length by ignoring the least value attribute information. In order to increase the system efficiency, least value attribute information wants to be compressed.

Furthermore, to provide higher security for the system, we are using polycentric model. The paper is systematized as shown below. The starting section discusses about the development process of ABE. The second section briefly concentrated on various related works with their comparison. The third section proposes the system model with a proposed algorithm. At the ending section, the completed work is highlighted.

2 Related Work

The rapid growth of internet and their increasing use of online services, there is a lack of single authority control system. Nowadays, all the services are based on multiauthority control systems. Nowadays, scholars are begun to pay attention on

Table 1 Analysis of the related works

Research papers	Access structure	Measurement of ciphertext	Multiauthority arrangement
[7]	Based on tree	Non-persistent	Solo CA
[9]	Based on AND gate	Persistent	No
[10]	Based on AND gate	Non-persistent	Multiple CAs
Proposed scheme	Based on AND gate	Non-persistent	Multiple CAs

multiple approvals techniques. In [6], this algorithm uses more than one CA that result in increasing the processing time when compared to solo CA. To overcome this problem, latter design an algorithm by using the complex group to reduce their processing time. In [7], threshold sharing of secret key was introduced in the algorithm using attributes. In [8, 9], they introduce the concept of trusted authority (TA), who is the ultimate control of the system. The TA is the only sole, having the privilege to create the next-level authorities. So, the TA should be a trusted person; any failure in the TA results in the collapse of the entire system. In [10], giving to the preceding, MAABE become the more effective algorithm by providing more elastic access structure with the logic of using the AND gate which helps in reducing the processing time and to rise the effectiveness. The comparative analysis of the related research will be equated, as in Table 1.

Nowadays, the majority of the algorithms are designed by means of the gate to provide better access framework and use CA to control the different authorities.

3 System Model

The core area of this planned system model is to offer safe personal health record (PHR) access and well-organized encryption and decryption. The vital focus is to split the arrangement into numerous security provinces (SPs) specifically, personal provinces and public provinces. Personal provinces are related with family members or close friends. The public provinces are related with users based on professional roles. The data owner is responsible for providing the access privileges for both the provinces. This is shown in Fig. 1.

These algorithms are preformed throughout the processes of storing the PHR files: key setup and key generation, encryption, decryption and revocation of user attributes. Here, the algorithm is similar to Li et al. [11], except the concept of selecting the attributes in the binary tree structure for effective encryption and decryption based on polycentric model, i.e., try to select the most common attributes from the different provinces. Here, an access construction $A \in N$ (N is the attributes set represented for global) is recognized by the users.

In the current research, a binary recursive algorithm traversal is defined; therefore, two child nodes are there in each interior node of the binary tree:

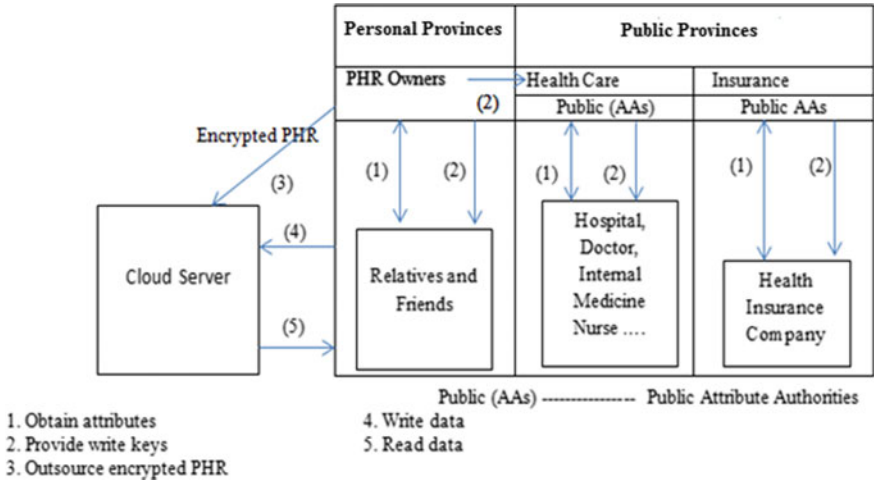


Fig. 1 Proposed framework for PHR

```
Function Traversal
Begin
Attributes atk,j and atk,j+1
If Child Nodes
then
return values arej and j  + 1;
End If
If Non- Child Nodes
atk,j and atk,j+1 = 1;
then
    return value 1;
else if atk,j and atk,j+1 both are equal to 0
then
return value 1;
else
return 0;
End If
End Function
```

This information is passed to the key generation algorithm to perform better encryption and decryption in order to reduce computational time.

4 Conclusion

The proposed technique increases system efficiency, by compressing the least value attribute information for the encoding and decoding of the data. In order to administer the attributes in each AA, binary tree is used to represent the child nodes. Binary tree representation helps to decrease encoding and decoding timing and length of the ciphertext.

References

1. Vigila SMC, Muneeswaran K (2011) Elliptic curve based key generation for symmetric encryption. In: IEEE International conference on signal processing, communication, computing and networking technologies, pp 824–829. https://doi.org/10.1109/icsccn.2011.6024664
2. Sahai A, Waters B (2005) Fuzzy identity-based encryption. In: Annual international conference on theory applications and cryptographic techniques, Aarhus, Denmark, Springer, pp 457–473. https://doi.org/10.1007/11426639_27
3. Cheung L, Newport C (2007) Provably secure ciphertext policy ABE. In: Proceedings of the 14th ACM conference on computer and communications security, New York, USA, pp 456–465. https://doi.org/10.1145/1315245.1315302
4. Li W, Xue K, Hong J (2015) TMACS: a robust and verifiable threshold multi authority access control system in public cloud storage. IEEE Trans Parallel Distrib Syst 27(5):1484–1496. https://doi.org/10.1109/tpds.2015.2448095
5. Chase M, Chow SSM (2009) Improving privacy and security in multiauthority attribute-based encryption. In: Proceedings of the 16th ACM conferences on computer and communications security, New York, USA, pp 121–130. https://doi.org/10.1145/1653662.1653678
6. Ning J, Dong X, Cao Z-F, Wei L (2015) White-box traceable ciphertext-policy attribute-based encryption supporting flexible attributes. IEEE Trans Inf Forensics Secur 10(6):1274–1288. https://doi.org/10.1109/tdsc.2016.2608343
7. Luo E, Liu Q, Wang G (2016) Hierarchical multiauthority and attribute-based encryption friend discovery scheme in mobile social networks. IEEE Commun Lett 20(9):1772–1775. https://doi.org/10.1109/lcomm.2016.2584614
8. Zhou Z, Huang D, Wang Z (2013) Efficient privacy-preserving cipher text policy attribute based-encryption and broadcast encryption. IEEE Trans Comput 64(1):126–138. https://doi.org/10.5815/ijcnis.2016.10.07
9. Guo F, Susilo W, Wong D, Varadharajan V (2014) CP-ABE with constant-size keys for lightweight devices. IEEE Trans Inf Forensics Secur 9(5):763–771. https://doi.org/10.1109/tifs.2014.2309858
10. Han J, Susilo W, Mu Y, Yan J (2012) Privacy-preserving decentralized key-policy attribute-based encryption. IEEE Trans Parallel Distrib Syst 23(11):2150–2162. https://doi.org/10.1109/tpds.2012.50
11. Li M, Yu S, Zheng Y, Ren K, Lou W (2013) Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE Trans Parallel Distrib Syst 24(1):131–143. https://doi.org/10.1109/tpds.2012.97

BrainNET: A Deep Learning Network for Brain Tumor Detection and Classification

Aditya Raj, Abhishek Anil, P. L. Deepa, H. Aravind Sarma, and R. Naveen Chandran

Abstract The increased use of technology had an impact to the overall wellbeing. Health experts have increasingly taken advantage of the benefits of these technologies thus generating a scalable improvement in the area of health care. Because of this, there is paradigm shift from manual monitoring toward more accurate virtual monitoring with minimum percentage of error in the area of health care. Advances in artificial intelligence (AI) led to exciting solutions with good accuracy for medical imaging and is a key method for future applications in health care. Brain tumor detection is an important task in medical image processing. Early detection of brain tumors plays an important role in improving treatment possibilities and thus increasing the survival rate of the patients. Manual detection of the brain tumors for cancer diagnosis from a large amount of MRI images generated in clinical routines is a difficult and time consuming task due to the complexity and variance of tumors and medical data. So, there is a need for automatic brain tumor detection from brain MRI images. With the help of deep learning networks, we can automate the detection process. For that, we have proposed a new network known as BrainNET which reads the MRI images coming from the MRI machine, and then, it detects as well as classifies the brain tumor if present.

A. Raj (✉)
QBurst, Technopark, Thiruvananthapuram, India
e-mail: adityaraj1797@gmail.com

A. Anil
Infosys, Mysore, India
e-mail: abhishekanil14@gmail.com

P. L. Deepa · R. Naveen Chandran
Mar Baselios College of Engineering and Technology, Thiruvananthapuram, India
e-mail: deepa.pl@mbcet.ac.in

R. Naveen Chandran
e-mail: naveenchandran97@gmail.com

H. Aravind Sarma
Vinvish Technopark, Thiruvananthapuram, India
e-mail: aravindsarmah@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_49

Keywords Deep neural network · Brain tumor detection · Brain tumor classification · Deep network designer

1 Introduction

The tumors are swollen mass formed in some parts of human body caused by an abnormal growth of cells. The occurrence of tumor in brain led to brain tumors. Brain tumor occurs when cells get abnormally multiplied within the brain. Early detection of brain tumor is of utmost importance. The tumors can be either of malignant (cancerous) or benign (non-cancerous). Malignant tumors will spread without any control and affect other cells, but benign tumor will not spread to the nearby tissues and hence less risky. Brain tumors are one of the deadly cancers which are of different types and grades. The main diagnosis methods are neurological examination, magnetic resonance imaging (MRI) scan, computer tomography (CT) scan, spinal tap and biopsy. Among these techniques, basically MRI scan is used for detecting the brain tumor. MRI is preferred than CT scan because the human brain is filled with soft tissues and MRI is focused on soft tissue, but whereas CT is focused on hard tissues. The accuracy of detection depends on the radiologist. If he fails to do it accurately, it will affect the patient's life. So, here we are introducing deep learning techniques for the efficient detection of the brain tumors and classify them into different classes.

Heba et al. [1] used deep learning technology for brain tumor classification. They have incorporated the features extracted using DWT and PCA with the deep learning classifier for classifying tumors into four classes, e.g., normal, glioblastoma, sarcoma and metastatic bronchogenic carcinoma tumors. Varuna Shree and Kumar [2] used gray level cooccurrence matrix (GLCM) and DWT of brain for getting better performance. Sobhaninia et al. [3] conducted studies about different angles of MRI and used different networks for segmentation. They obtained a die score of 0.73 in single network and 0.79 in multiple network. Chinmayi et al. [5] discussed a method to segment and classify the MRI brain image using Bhattacharyya coefficient into two classes, normal and abnormal. Bahadure et al. [6] used BWT and SVM classifiers to obtain an accuracy of 96.5%, specificity of 94.2% and sensitivity of 97.2% for detecting normal and abnormal tissues from brain MRI. Isın et al. [7] done with the automatic segmentation of MRI images using deep learning methods. Antony et al. [8] done with the efficient classification of brain MRI images into normal and abnormal cells using deep CNN. Lina Chato et al. [4] used machine learning techniques for the prediction of glioma brain tumor. They have used histogram and a pretrained CNN model known as Alexnet for this purpose. The maximum classification accuracy obtained is 68.5%.

In [9], brain MRI images are classified into three types: normal, malignant and benign using a probabilistic neural network (PNN). Here, principal component analysis is used for image compression, and K-means clustering algorithm is used to segment out the affected area of the brain. In this paper on comparing PNN and

CNN, PNN is said to have more advantages which is due to the fact that PNN learns from the training data instantaneously. Here, MATLB 2013 is used for simulation. In [10], Alexnet model is used for classifying tumors into different types as a base model along with Faster R-CNN algorithm. Here, transfer learning is being used during the training phase. The proposed system can predict the type of tumor correctly with good accuracy. Here, the convolutional layers output of alexnet is taken as the convolutional feature map which is given as input to the Region Proposal Network and ROI pooling layer. Here, Python language was used for implementation.

In [10], we have discussed about detection of brain tumor by applying transfer learning to already existing pretained networks such as Alexnet, VGG16 and VGG 19. For doing the detection with pretrained networks, there is a need of resizing the MRI images because the pretrained network's input image sizes are 224×224 and 227×227. But MRI images are usually of size 512×512. While resizing it, there are chances of data lose. To overcome this limitation, we proposed a new network known as BrainNET which can process 512×512 images. Also, we have proposed four classification networks which can classify the input brain MRI tumor images into three classes: glioma, meningioma and pituitary tumor.

In this paper, we are proposing a new DL network for the detection of brain tumor from MRI images. The existing best classification networks are retrained using the technique of transfer learning for converting it to classify the tumor into three different classes. Also, we are proposing a new DL network for classification of brain tumor data into three different classes. The organization of the paper is as follows. Section 2 describes the proposed networks followed by the network implementation in Sect. 3. Section 4 describes the details of graphical user interfaces followed by results and discussion in Sect. 5. We are concluding in Sect. 6.

2 Proposed System

Automated brain tumor detection and classification system will help the doctor as well as radiologist to diagnose the tumor area correctly and more accurately. This paper deals with the detection and classification of brain tumor from MRI. The block schematic of the proposed system is shown in Fig. 1. At first, the input MRI images are collected to form the database. Pre-processing step includes the resizing of images for pretrained network. The pretrained models are created by using the method known as transfer learning. The dataset is divided into training data and testing data. The deep learning network is trained using the training dataset for detection if it is the detection network. Then, the query can be given as input, and the network will check for tumor area. If it is present, the image is given to post-processing step for segmenting out the tumor area. For classification, the classification network is trained using the classification dataset and tested using a query containing the tumor.

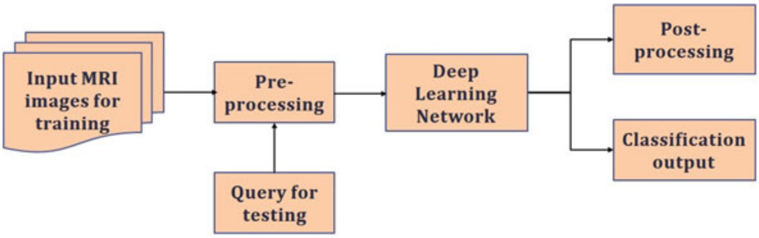


Fig. 1 Block schematic of the proposed system

2.1 Proposed Detection Network: BrainNET

BrainNET is a type of convolutional neural network (CNN), designed for effective detection and classification of brain tumors from MRI data having an input size same as that of MRI image size that is 512 × 512. Here, we are proposing two different networks: one for detection and second for classification. Classification network will be active only if brain tumor is detected by the detection network. BrainNET detection network shown in Fig. 2 consists of 22 layer stacks in which 17 are convolutional layers and remaining 5 are fully connected layers. Each convolutional layer stack consists of a Relu layer, and some contains a max pooling layer. The network has a total of 56 layers and can detect the tumor images more accurately. The filter size used is 3 × 3 having a single stride thus helps to increase the accuracy and efficiency. By using 3 × 3 filter size, fine tuning of MRI image is possible so that the minute tumors can be detected. The input layer of the network consists of 786,432 (512 × 512 × 3) neurons, and at last, the output is detected using the softmax classifier which has two neurons, and it classifies the MRI image as normal or tumor image.

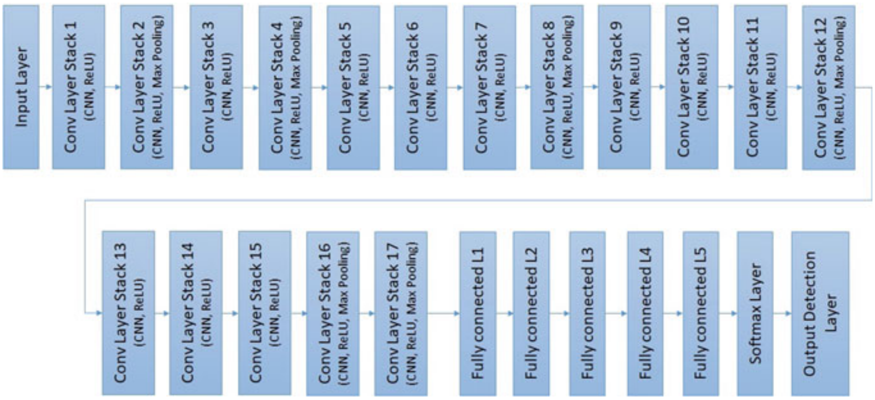


Fig. 2 Architecture of BrainNET detection network

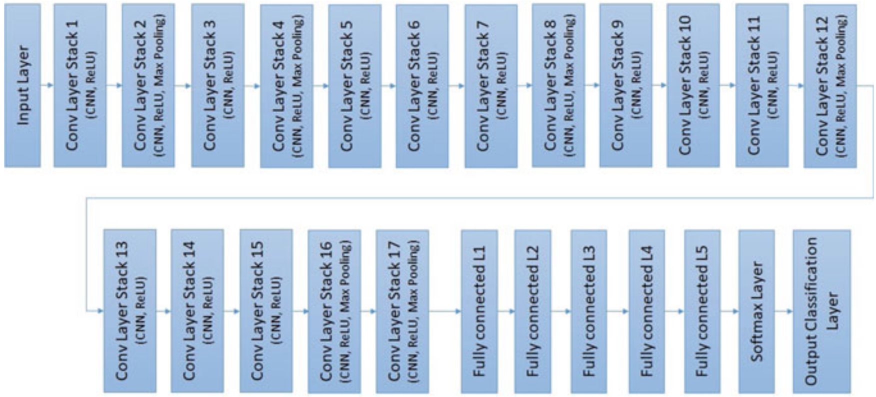


Fig. 3 Architecture of BrainNET classification network

2.2 *Proposed Classification Networks*

Alexnet [11] and VGGnets [12] are the some of the best classifiers available in the literature. So, we have used these pretrained models for creating different classification networks. Last fully connected layer of each of these models is reconfigured and retrained by the process of transfer learning, and the trained models are obtained. We have done this for Alexnet, VGG16 and VGG19 for getting three different classification models. As the input size is fixed, we need to resize the input image size according to the size of the input layer, and thus, data lose may occur. So, we have proposed a new classification framework similar to BrainNET which can directly handle the input data of size $512 \times 512 \times 3$.

BrainNET classification architecture shown in Fig. 3 is similar to BrainNET detection architecture having 22 layer stack followed by Relu and max pooling. Here also, input layer of the network consists of 786, 432 ($512 \times 512 \times 3$) neurons. The network classifies the data into mainly three classes they are glioma, meningioma and pituitary tumor using softmax classifier. Here, softmax classifiers are used for the classification of the MRI data. Softmax layer consists of three neurons which are used to classify the data as the above-mentioned classes.

3 Network Implementation

BrainNET is implemented using the Deep Learning Toolbox in MATLAB2018b. Deep Learning Toolbox in MATLAB is incorporated with a deep network designer which can be defined as "a point-and-click tool for creating and modifying deep neural networks." Using this app, we created a new network by drag and drop new layers according to our requirement, and thereby, new network is created and can

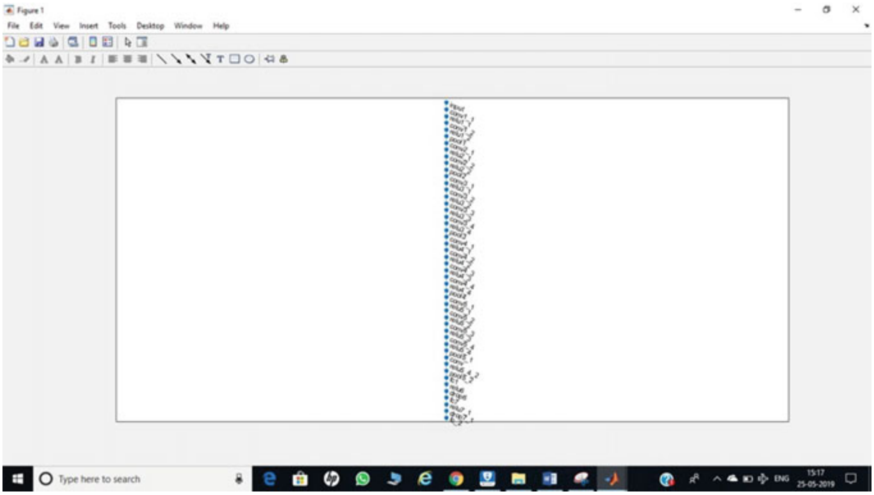


Fig. 4 Layer graph of BrainNET

analyze the network for checking whether our connections are correct. If there is no error, then the layer graph will be generated after that network training is done. Layer graph is shown in Fig. 4. Deep learning network can also be created without the help of the deep network designer. This can be done by MATLAB programming. The layers can be implemented using the MATLAB functions. After creating the network structure either by using the network designer or by using functions, save the layered structure in ".mat" file format for future use. The network can be trained by using the following functions:

1. trainingOptions(solver Name): This function creates a set of training options according to the solver specified by the solver Name for training the network using the next function. The possible values of the solver Name are "sgdm," "adam" and "rmsprob". The other main input arguments are "Initial Learn Rate" which represents the initial learning rate used for training the network; "Max Epochs" which represents the maximum number of epochs used for training and "Mini Batch Size" which represents the size of the mini batch used for each training iteration.
2. trainNetwork(ds, layers, options): This function is used for training the network. It will return the trained classification network which can be saved as a ".mat" file for future purposes. "ds" is the training dataset; "layers" is the layer structure developed previously, and "options" is the set of training options created. After training, the network can be tested by using the function "classify( )" which can find out the class with which the input data belongs to. After testing the network with sufficient amount of test data, we can assess the performance by creating the confusion matrix.

Datasets which are available which can be given as the inputs to the input layer are Oasis dataset, Pre- and Post-tumor dataset and types of tumor dataset. The Oasis dataset contains over 2000 MR sessions which include T1w, T2w, FLAIR, ASL, SWI, time of flight, resting-state BOLD and DTI sequences. Many of the MR Sessions are accompanied by volumetric segmentation files produced through free surfer processing. In this work, MRI images of brain tumor patients have been used. Pre- and post-tumor dataset available in the Internet consists of MRI images of patients having tumor and the MRI images of patients after removing the tumor. From the above-said databases, we have consolidated a total of 23,444 brain MRI images which consist of 16,944 tumor data and 6500 normal data. Among these, 80% (18,755) is used for training and 20% (4689) for testing. Types of tumor dataset are available in the Mathworks site. It contains MRI images of three classes: glioma, meningioma and pituitary tumor. From this dataset, we have used a total of 3064 images which consist of 1426 glioma images, 708 meningioma images and 930 pituitary tumor images for classification purpose. Among these, 80% (2451) is used for training and 20% (613) for testing.

4 Graphical User Interface of the Proposed System

We have created a graphical user interface (GUI) for making it easier for anyone to operate this system. The GUI was created in MATLAB 2018b. The start button is used to display the image to be detected and classified. The button allocated for BrainNET detection network is used to output the result as normal/tumor and also display the segmented tumor area. The button allocated for the BrainNET classification network is used to output the classified MRI image into one among glioma/meningioma/pituitary tumor.

We have created a GUI for dealing with MRI video too. This can read the MRI video data from the output of the MRI instrument if interfaced properly and can detect the slices having tumor data correctly. For example, if we are inputting a video data, it will first play the video data and predict from which slice to which slice the tumor is present.

5 Result and Discussion

5.1 *BrainNET Detection Network*

A total of 4689 brain MRI images were taken for testing the BrainNET detection network. Out of it, 3389 were tumor images and 1300 were normal images. After training, the new network is saved as a .mat file, and testing is done with the testing data. From the obtained output, a confusion matrix was created using the above data

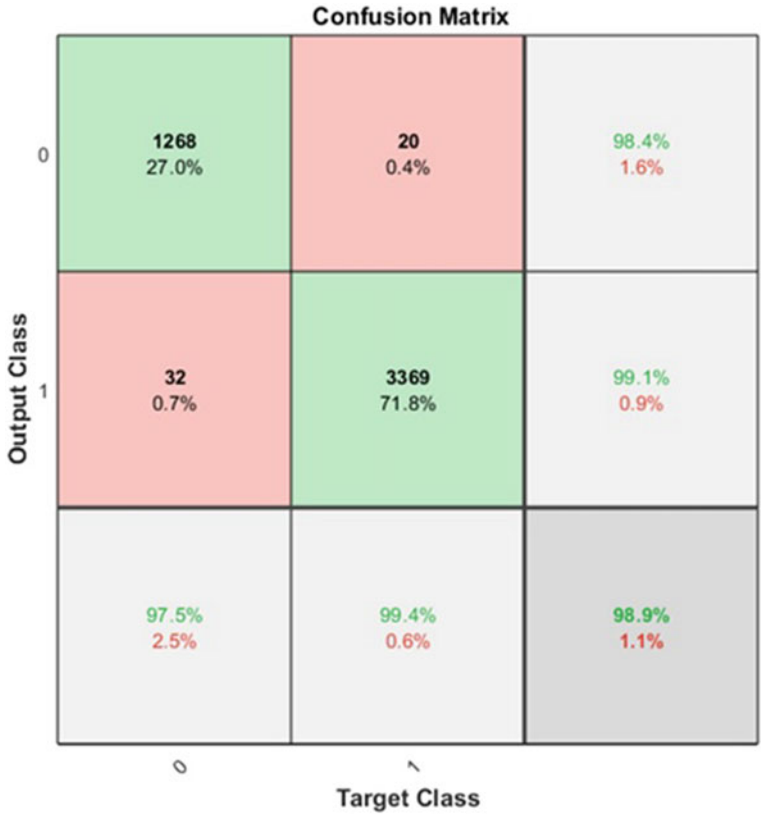


Fig. 5 Confusion matrix of BrainNET detection

which determines various parameters of the network such as sensitivity, specificity, error rate and accuracy. The network correctly detected 3299 tumor images out of 3369 images and 1268 normal images out of 1300 images. The sensitivity of the network was calculated to be 99.1%, specificity as 98.4%, error rate as 1.1%, and the accuracy of the network was obtained as 98.9%. This is shown in the confusion matrix in Fig. 5.

5.2 *Classification Networks*

The classification networks are also trained with the classification dataset, and the trained models are saved as .mat files. Then, the network was tested with the testing dataset. The trained model will classify the data into three tumor types: glioma, meningioma and pituitary tumor. A total of 613 tumor MRI images were taken for testing the classification networks. Out of these, 285 were glioma, 142 were meningioma and 186 were pituitary tumor images. Among these 613, 80% is used for training and remaining 20% for testing.

For Alexnet, VGG16 and VGG19, we resized the input data according to the size of the input layer. The accuracies obtained are 90.05%, 94.45% and 95.11%, respectively.

5.3 BrainNET Classification Network

BrainNET classification network is also trained with the same data without resizing, and the confusion matrix obtained after testing is shown in Fig. 6. 281 out of 285 glioma images, 142 out of 142 meningioma images and 184 out of 186 pituitary tumor images were correctly classified, and therefore, the network accuracy is calculated to be 99.0%.

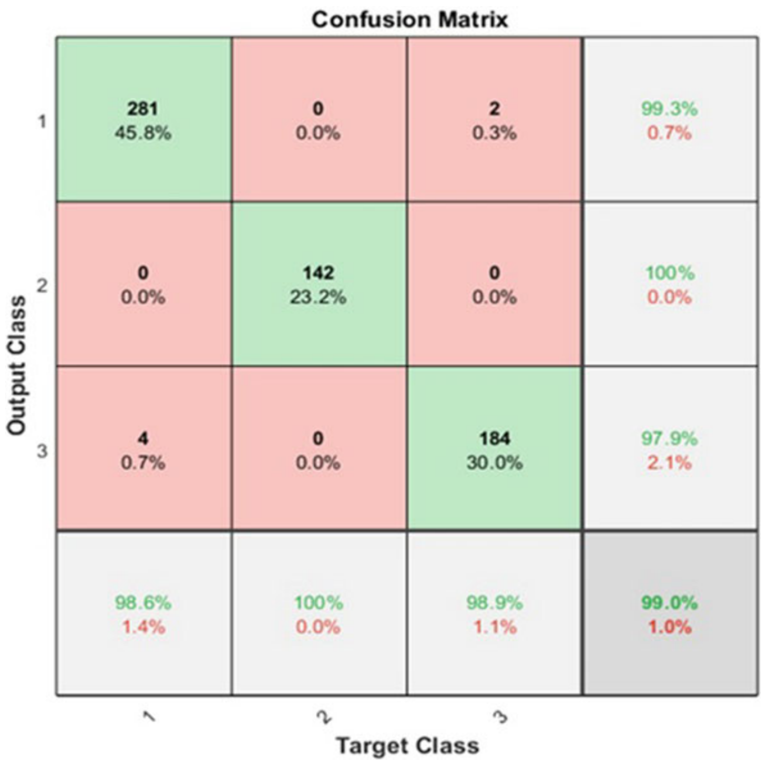


Fig. 6 Confusion matrix of BrainNET classification

6 Segmentation

After detecting whether the given MRI data contains tumor or not, BrainNET detection network will also segment out the tumor area as shown in Fig. 8. The first image represents the given MRI, and the second image shows the tumor area segmented out form the tumor image. The third image shows the tumor area inside the MRI image.

We have checked the GUI by giving a glioma image as shown in Fig. 7. The segmented output is shown in Fig. 8. The second GUI for video is tested with an MRI video consists of 52 slices. From the output shown in Fig. 9, it is evident that tumor is present from the 11th slide to 32th slide. This will facilitate the doctor to evaluate the approximate size of the tumor present in human brain.

We have done a total of four detection networks for comparing BrainNET with the existing best classifiers like Alexnet and VGGnet. Their results can be consolidated as shown in Table 1.

Fig. 7 GUI of BrainNET detection and classification

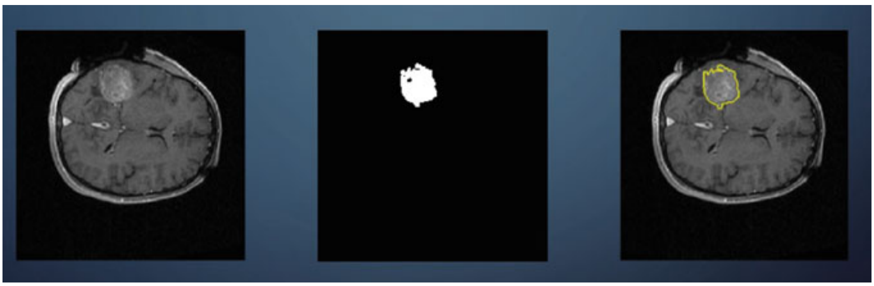


Fig. 8 Segmented output

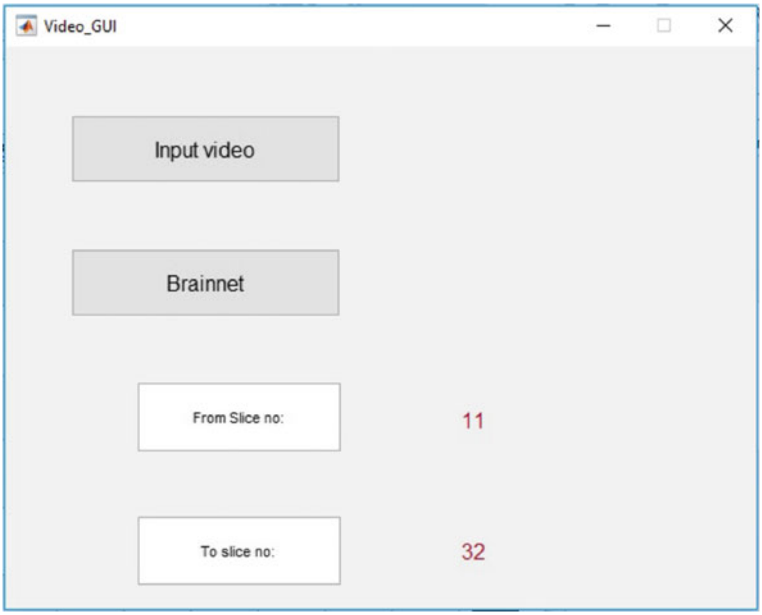


Fig. 9 Segmented output

Table 1 Performance of detection networks

Detection networks	No. of layers	Training execution time—one time process (HH:MM:SS)	Testing execution time for single image (s)	Correctly detected out of 4689	Specificity (%)	Sensitivity (%)	Detection accuracy (%)
Alexnet	8 stacks (25 layers)	11:38:24	7.59	4203	84.1	91.5	89.64
VGG16	16 stacks (41 layers)	21:32:56	19.07	4371	89.0	94.8	93.22
VGG19	19 stacks (47 layers)	24:45:15	20.23	4491	93.0	96.8	95.78
BrainNET (proposed network)	22 stacks (56 layers)	36:11:37	26.29	4637	98.4	99.1	98.89

Also, we have done a total of four classification networks for comparing BrainNET with the existing best classifiers like Alexnet and VGGnet. Their results can be consolidated as shown in Table 2.

Table 2 Performance of classification networks

Classification networks	No. of layers	Training execution time—one time process (HH:MM:SS)	Testing execution time for single image (s)	Correctly detected out of 613	Classification accuracy (%)
Alexnet	8 stacks (25 layers)	03:38:25	5.02	552	90.05
VGG16	16 stacks (41 layers)	12:36:46	10.47	579	94.45
VGG19	19 stacks (47 layers)	18:43:35	18.19	583	95.11
BrainNET (proposed network)	22 stacks (56 layers)	24:19:17	23.21	607	99.02

7 Conclusion

In this paper, we have proposed five new networks: one for detection of brain tumor detection and four for brain tumor classification. The proposed detection network can effectively detect brain tumors from MRI images as compared to the state of the art methods. Among the four classification networks, three are generated by the method of transfer learning, and the fourth is a completely new network. All the networks are done in MATLAB2018b, and the layer graphs obtained are saved as .mat files so that anyone can use this network in any of the platform. By the method of transfer learning, this can be used for any two class and three class classification purposes. MRI data are taken from the standard datasets available in the literature. Image resizing is not required for the new networks and therefore no data loss. The detection and classification networks are giving good accuracies compared to the existing methods. We have created graphical user interfaces (GUI) for detecting and classifying brain tumor from image as well as video. It provides the user an easier way to access our system. This system can be used by the radiologist for fixing his findings in an accurate manner. Here, we have considered only three classes of brain tumors as they are the most commonly occurring ones. In future, we can increase the number of classes for making it more user friendly and convenient for effective use in medical field.

References

1. Heba M, El-Dahshanb El-Sayed A, El-Horbaty El-Sayed M, Salem Abdel-Badeeh M (2018) Classification using deep learning neural networks for brain tumors. J Future Comput Inform 3(1):68–71

2. Varuna Shree N, Kumar TNR (2018) Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. J Brain Inform 5(1):23–30
3. Sobhaninia Z, Rezaei S, Noroozi A, Ahmadi M, Zarrabi H, Karimi N, Emami A, Samavi S (2018) Brain tumor segmentation using deep learning by type specific sorting of images. In: IEEE international conference on computer vision and pattern recognition (CVPR), Sept 2018
4. Chato SL, Latifi S (2017) Machine learning and deep learning techniques to predict overall survival of brain tumor patients using MRI images. International conference of bioinformatics and bioengineering, Oct 2017
5. Chinmayi P, Akilandeswari L, Prabu Kumar M, Muralibabu K (2017) An efficient deep learning neural network based brain tumor detection system. Int J Pure Appl Math 117(17):151–160
6. Bahadure NB, Ray AK, Thethi HP (2017) Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int J Biomed Imaging. Article ID 9749108
7. Isın A, Direkoglu C, Sah M (2016) Review of MRI-based brain tumor image segmentation using deep learning methods. J Procedia Comput Sci 102:317–324
8. Antony A, Ancy Brigit MA, Fathima KA, Raju D, Binish MC (2017) Brain tumor detection and classification in MRI images. Int J Innov Res Sci Eng Technol 6(5)
9. Lavanyadevi R, Machakowsalya M, Nivethitha J, Niranjil Kumar A (2017) Brain tumor classification and segmentation in MRI images using PNN. In: IEEE international conference on electrical, instrumentation and communication engineering (ICEICE), Apr 2017
10. Anil A, Raj A, Aravind Sarma H, Naveen Chandran R, Deepa PL (2019) Brain tumor detection from brain MRI using deep learning. In: 9th international conference on recent engineering and technology (ICRET), Aug 2019, vol 3, no 2, pp 451–457. https://doi.org/10.29027/IJIRASE.v3.i2.2019
11. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. In: IEEE International conference on computer vision and pattern recognition (CVPR), Sept 2014
12. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: ACM 25th international conference on neural information processing systems (NIPS'12), vol 1, Dec 2012, pp 1097–1105
13. Ezhilarasi R, Varalakshmi P (2018) Tumor detection in the brain using faster R-CNN. In: IEEE international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), Aug 2018, pp 388–392

High-Speed Inversion Using x^{4^n} Units

M. Kalaiarasi, V. R. Venkatasubramani, A. Christina Grace, and S. Rajaram

Abstract Inversion is a significant operation in ECC processors and is the most complex and time-consuming operation among operations like addition, subtraction and multiplication. Thus, proposing an algorithm and designing its architecture to compute inverse with minimum number of clock cycles are mandatory. In this brief, high-speed inversion of NIST recommended pentanomial GF(2^{163}), based on traditional Itoh–Tsujii inversion algorithm (ITIA) is proposed. This proposed inversion algorithm is then implemented on FPGA Virtex-5 platforms to analyze its performance. This design minimizes the latency and, thereby, improves speed. The developed high-speed Itoh–Tsujii inversion algorithm HS-ITIA computes inversions in 18 clock cycles with maximum clock frequency 64.5 MHz, which thereby yields a rise in performance by 56%.

Keywords Field-programmable gate array (FPGA) · Elliptic curve cryptography (ECC) · High-speed Itoh–Tsujii inversion algorithm (HS-ITIA)

1 Introduction

Inversion is traditionally carried out by two methods: extended Euclidean algorithm (EEA) and ITA using Fermat's little theorem (FLT). The implementation of the extended Euclidean algorithm is complex due to its excess computing time on the operands. Thus, inversion architectures based on FLT are more efficient. The Itoh-Tsujii inversion algorithm is based on FLT and was designed using square blocks [1] which was found to be advantageous over existing sequential and recursive type computations. In order to improve the performance, it was reformulated using quad blocks. The quad Itoh–Tsujii algorithm uses FPGA resources better and requires short addition chains than the existing squarer circuits [2]. A theoretical model for any

M. Kalaiarasi (✉) · V. R. Venkatasubramani · A. Christina Grace · S. Rajaram
Thiagarajar College of Engineering, Madurai, India
e-mail: kalaibhuvan26@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_50

Galois field and k-input LUT-based FPGA ($k > 3$) [3] aided the hardware designer to quickly opt the ideal design parameters. The inversion algorithm is reformulated to generalize the primitive one so that the universal algorithm for all binary fields is achieved [4].

In this work, the high-speed Itoh–Tsujii inversion algorithm (HS-ITA) using 2^n blocks [5] is reformulated using 4^n blocks to minimize the clock cycles for computation. Thus, the conventionally used 2^n units are modified into non-iterative logic using 4^n blocks. This architecture is compatible with other ECC operations, and the design is carried out in FPGA Virtex-5. Unlike the existing works, the proposed design seems to show an optimal decrease in latency without compensating its performance.

The rest of the document is arranged such that Sect. 2 presents the existing high-speed Itoh–Tsujii inversion algorithm (HS-ITIA) using 2^n blocks. Section 3 delivers the suggested hardware architecture using 4^n blocks followed by the proposed algorithm in Sect. 4. Section 5 deals with the control word development and the calculation of clock cycles to compute inversion. Section 6 details the hardware implementation part. Section 7 provides the experimental results and Sect. 8 concludes the paper.

2 Existing HS-ITA Using (2^n) Blocks

Inversion of a variable x using Fermat's little theorem is given by

$$x^{-1} = x^{2^m-2} = x^{2\left(2^{m-1}-1\right)} \tag{1}$$

Thus, it requires $m - 2$ squaring and $m - 1$ multiplications. In order to simplify, inversion is computed by the following conditions.

$$2^s - 1 = \begin{cases} \left(2^{\frac{s}{2}} - 1\right)\left(2^{\frac{s}{2}} + 1\right), \text{when } s \text{ is odd} \\ 2\left(2^{\frac{(s-1)}{2}} - 1\right)\left(2^{\frac{(s-1)}{2}} + 1\right) + 1, \text{when } s \text{ is even} \end{cases} \tag{2}$$

$$\begin{aligned} x^{-1} &= x^{2^{163}-2} = x^{2\left(2^{162}-1\right)} = x^{2\left(2^{81}+1\right)} \\ &\times\{2(2^{40}+1)(2^{20}+1)(2^{10}+1)(2^5+1)[2(2^2+1)(2+1)+1]+1\} \end{aligned}$$

This includes seven 2^n blocks ($2^{81}, 2^{40}, 2^{20}, 2^{10}, 2^5, 2^2, 2^1$). Due to the fact that all 2^n values are pre-computed, the value of inverse is computed within 19 clock cycles.

3 Proposed Modular Inversion for GF(2^m) Using 4^n Blocks

The purpose of this document is to reformulate the high-speed ITIA using 2^n blocks, with 4^n blocks.

$$x^{-1} = x^{2^{163}-2} = x^{2(2^{162}-1)} \tag{3}$$

$$= x^{2(4^{81}-1)} \tag{4}$$

This is then processed using the following conditions:

$$(4^s - 1) = \begin{cases} \left(4^{\frac{s}{2}} - 1\right)\left(4^{\frac{s}{2}} + 1\right), \text{ when } s \text{ is odd} \\ 4\left(4^{\frac{(s-1)}{2}} - 1\right)\left(4^{\frac{(s-1)}{2}} + 1\right) + 3, \text{ when } s \text{ is even} \end{cases} \tag{5}$$

Inversion using 4^n blocks over NIST recommended field GF(2^{163}) is given by:

$$x^{-1} = x^{2\{4(4^{40}+1)(4^{20}+1)(4^{10}+1)(4^{5}+1)[4(4^{2}+1)(4+1)+3]+3\}} \tag{6}$$

It requires a pre-computation register to store the cube value, and the inverse calculation is complete after the final squaring. This uses only six 4^n blocks to obtain the final inverse value. Since $x^{(4^2)}$, $x^{(4^5)}$, $x^{(4^{10})}$, $x^{(4^{20})}$, $x^{(4^{40})}$ are computed in advance, the clock cycles depend only on the multiplications. Therefore, it requires 18 clock cycles for the implementation of inversion over GF(2^{163}).

4 Algorithm for HS-ITA Using 4^n Units

Input requirement: x
Ensure: $x - 1$. In this proposed high-speed Itoh–Tsujii inversion algorithm (HS-ITA), there are two fundamental operations that streamlined the algorithm. First, the operation is to calculate $x^{((4^n+1)a)}$ when the input is x^a, which can be achieved as follows (Table 1):

$$\begin{aligned} &z \leftarrow (x^a)^{4^n},\ 4^n \text{ unit in GF}(2^m) \\ &y \leftarrow x^a \,.\, z,\ \ \text{multiplication unit in GF}(2^m) \end{aligned} \tag{7}$$

The second operation is to calculate $x^{(2a+1)}$ when the input is x^a and x, which can be achieved by

$$\begin{aligned} &\text{z} \leftarrow (x^a)^4,\ \text{quad unit in GF } 2^m \\ &y \leftarrow x.z,\ \ \ \text{multiplication unit in GF}(2^m) \end{aligned} \tag{8}$$

Table 1 Algorithm for HS-ITA using 4^n units

1	z < − x*x	x^2
2	y < − z*x	x^3 (p register)
3	z < − y^4	x^{12}
4	y < − y*z	x^{15}
5	z < − y^{4^2}	x^{240}
6	y < − z*y	x^{255}
7	z < − y^4	x^{1020}
8	y < − z*p	x^{1023}
9	z < − y^{4^5}	$x^{1044480}$
10	y < − y*z	$x^{1045503}$
11	z < − $y^{4^{10}}$	$x^{1.096*10^{12}}$
12	y < − y*z	$x^{1.096*10^{12}}$
13	z < − $y^{4^{20}}$	$x^{1.205*10^{24}}$
14	y < − y*z	$x^{1.205*10^{24}}$
15	z < − $y^{4^{40}}$	$x^{1.457*10^{48}}$
16	y < − y*z	$x^{1.457*10^{48}}$
17	z < − y^4	$x^{5.828*10^{48}}$
18	z < − final squaring	$x^{1.165*10^{49}}$

5 Control Word Development for Inverse Computation over GF(2^{163})

Each 4^n unit is composed of 4^p units connected in cascade. We are therefore proposing a reduced architecture composed of 4^n units. This streamlined version is based on the reality that, instead of the logical operation, complexity depends on the cascade connection (Table 2).

6 Hardware Implementation of High-Speed ITA

The circuit is designed with two 4:1 multiplexers, a 2:1 multiplexer and a 8:1 multiplexer. The HS-ITA architecture has an input x, fed to the 4:1 multiplexer. Initially, the input is sent to the squarer using 8:1 multiplexer. At the next clock cycle, the squared output is given to the multiplier to compute the cube value, which is then stored in a register. This pre-computed value is used whenever needed, from the register. For

Table 2 Control word format for GF(2^{163})

Steps	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	Q_{OUT}	M_{OUT}	Output
$\alpha_1(a)$	0	0	0	0	0	d	d	d	1	d	x^2
$\alpha_2(a)$	d	d	d	d	d	0	0	0	d	1	x^3
$\alpha_3(a)$	0	1	0	0	1	d	d	d	1	d	x^{12}
$\alpha_4(a)$	d	d	d	d	d	0	1	0	d	1	x^{15}
$\alpha_5(a)$	0	1	0	1	0	d	d	d	1	d	x^{240}
$\alpha_6(a)$	d	d	d	d	d	0	0	1	d	1	x^{255}
$\alpha_7(a)$	0	1	0	0	1	d	d	d	1	d	x^{1020}
$\alpha_8(a)$	d	d	d	d	d	0	1	0	d	1	x^{1023}
$\alpha_9(a)$	0	1	0	1	1	d	d	d	1	d	$x^{1044480}$
$\alpha_{10}(a)$	d	d	d	d	d	0	0	1	d	1	$x^{1045503}$
$\alpha_{11}(a)$	0	1	1	0	0	d	d	d	1	d	$x^{1.09*10^{12}}$
$\alpha_{12}(a)$	d	d	d	d	d	0	0	1	d	1	$x^{1.09*10^{12}}$
$\alpha_{13}(a)$	0	1	1	0	1	d	d	d	1	d	$x^{1.25*10^{24}}$
$\alpha_{14}(a)$	d	d	d	d	d	0	0	1	d	1	$x^{1.25*10^{24}}$
$\alpha_{15}(a)$	0	1	1	1	0	d	d	d	1	d	$x^{1.45*10^{48}}$
$\alpha_{16}(a)$	d	d	d	d	d	0	0	1	d	1	$x^{1.45*10^{48}}$
$\alpha_{17}(a)$	0	1	0	0	1	d	d	d	1	d	$x^{5.8*10^{48}}$
$\alpha_{18}(a)$	1	0	0	0	0	d	d	d	1	d	$x^{1.16*10^{49}}$

d Do not care

Table 3 Comparison results of modular inverse

Algorithm	Device	$F_{\max}$ (MHz)	Cycles	Latency (ns)	LUTs	Performance	Field
[3] ITA	Virtex-5	177.6	20	107	21,278	439	163
[8]	Virtex-4	73.64	30	407	12,260	214	163
[7]	Virtex-4	87.71	25	285	15,602	225	163
Proposed HS-ITA	Virtex-5	65.43	18	111	23,512	397	163
[6] HS-ITA	Virtex-5	96.43	49	103	7,739	254	163

NIST recommended field over GF(2^{163}), the control word format is developed which clearly proves that the inverse is computed in 18 clock cycles, which is lesser than the existing quad ITA which is shown in Table 3 (Fig. 1).

7 Experimental Results

Performance is defined to determine the efficiency of the design using the number of clock cycles and LUTs.

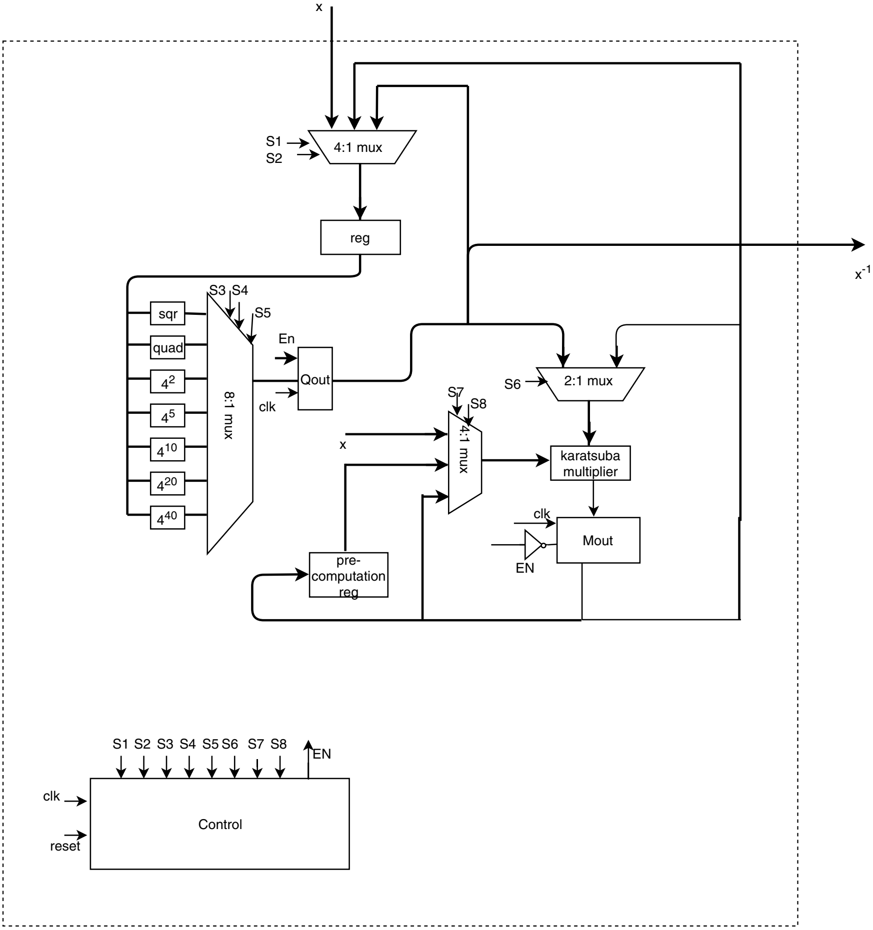


Fig. 1 Block diagram for $GF(2^{163})$ using high-speed ITA

$$\text{Performance} = \frac{1}{\text{LUTs} \times \text{Clockycles} \times \text{Period}} \tag{9}$$

The result of the proposed algorithm with other existing works [3, 6–8] is shown in Table 3.

The proposed HS-ITA are implemented on Virtex-5 and compared with the existing implementations in the most popularly used $GF(2^{163})$. The proposed HS-ITA not only increases the overall clock frequency, but also improves the overall performance. It takes only 18 clock cycles when pre-determined $4n$ blocks are used. The number of LUTs is found to be 23,512 with a maximum clock frequency of 65.43 MHz.

8 Conclusion

This brief recommends the hardware implementation of inversion over GF(2^{163}) using 4^n blocks on FPGA platform for analyzing different parameters. This architecture reduces two clock cycle when compared with the squarer block. Unlike the available works, the overall performance is found to be improved as a result of decrease in clock cycle.

References

1. Itoh T, Tsujii S (1988) A fast algorithm for computing multiplicative inverses in $GF(2^m)$ using normal bases. Inf Comput 78(3):171–177
2. Rebeiro C, Roy SS, Reddy DS, Mukhopadhyay D (2011) Revisiting the Itoh-Tsujii inversion algorithm for FPGA platforms. IEEE Trans Very Large Scale Integr Syst 19(8):1508–1512
3. Roy SS, Rebeiro C, Mukhopadhyay D (2011) Theoretical modeling of the Itoh-Tsujii inversion algorithm for enhanced performance on k-LUT based FPGAs. In: Design, automation & test in Europe conference & exhibition, 2011. IEEE, pp 1–6
4. Hu J, Guo W, Wei J, Cheung RCC (2015) Fast and generic inversion architectures over $GF(2^m)$ using modified Itoh Tsujii algorithms. IEEE Trans. Circuits Syst. II Express Briefs 62(4):367–371
5. Li J, Li Z, Xue C, Zhang J, Gao W, Cao S (2018) A fast modular inversion FPGA implementation over $GF(2^m)$ using modified x^{2^n} unit. In: Proceedings of IEEE international symposium on circuits and systems (ISCAS), May 2018, pp 1–5
6. Rashidi B, Farashahi RR, Sayedi SM (2017) High-performance and high-speed implementation of polynomial basis Itoh Tsujii inversion algorithm over $GF(2^m)$. IET Inf Secur 11(2):66–77
7. Li L, Li S (2017) Fast inversion in $GF(2^m)$ with polynomial basis using optimal addition chains. In: Proceedings of IEEE international symposium on circuits and systems, May 2017, pp 1–4
8. Sinha RS, Chester R, Debdeep M (2012) Generalized high speed Itoh-Tsujii multiplicative inversion architecture for FPGAs. Integr VLSI J 45(3):307–315

Design of CPW-Fed Slot Antenna for 5G Mobile Applications

B. Soumya, S. Ashok Kumar, and T. Shanmuganantham

Abstract In this paper, CPW-fed slotted flexible antenna has been presented for 5G mobile applications. The element of the bent antenna topology is analyzed to reduce the physical footprint, and a simple inspection with the terminals of a mobile is operated in the frequency band of 4.1 GHz. Ratio in the middle of the front and back of the bent antenna shows a high specific mapping up to the rate of a ratio of 1 dB. A compact wideband reflector is reported with a direct aspect ±90° and transmission <−25 dB. TThe broadband reflector is combined with a region twister over an antenna with 0.135λ long distance from the emitting design. The region twisted receiver is utilized in 4.1 GHz band combined with reflector, with a ratio in the middle of the front and back greater than 14 dB over the frequency and an onward gain of 5.5 dBi.

Keywords Slot antenna · CPW · 5G applications

1 Introduction

Due to massive explosion in broader bandwidth applications of cellular technology, the hardware design for 5G requires to support the mm Wave frequency. The bent antenna capacity is for simplicity in spectral overcrowding in sub-6 GHz bands. So, resetting to top carrier frequencies can decrease the structure and the hardware ecosystem development to cooperative with the presented loops [1–3]. The provocation with resonant frequency band diffusion is reduced, and destruction is in high path reduction, and diffusion in losses is noticed. For the recounting of these, antennas

B. Soumya (✉) · S. Ashok Kumar
Jyothishmathi Institute of Technological Sciences, Karimnagar, India
e-mail: bhykanisowmya0410@gmail.com

S. Ashok Kumar
e-mail: ashokape@gmail.com

T. Shanmuganantham
Pondicherry University, Pondicherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_51

allocated in communication system require to reduce its effects. When the beamwidth would be low, then reducing analysis is known as high gain. It is a trade-off in between coverage area and gain, which requires to be examined for the structure and allocation [4].

The required aspects of a 5G mobile distraction antenna involve in a design and small solid footprint, high impedance bandwidth to bear the further spectrum at 4 GHz, steady patterns across the whole band by certain low appropriation rate, huge gain to obtain pattern on the mobile terminal and polarization diversity to bear applications in many directions [5]. Different arrangements of antenna for 5G mobile extreme have been analyzed corresponding to several film designs so absorbing comparatively a huge region, with 9% resistive bandwidth. Even though the receiver component is presented in 10% bandwidth, the propagation designs may not be acceptable for combined with the mobile extreme [6–8]. The suggested antenna has a bandwidth of greater than 35% along a gain of 5.5 dBi, but the structure is two dimensional. The design presented and offers 20% bandwidth with a SIW feed, hence increasing the complexity of the design [8]. It should be observed that most of the described designs are two dimensional and structured with microstrip feed.

2 Antenna Geometry

The CPW-fed slit receiver is shown in Fig. 1. It is designed on Nelco NY9220 substance among efficiency of 2.2 and height of mm. A relative permittivity is selected to keep the ground plane to a small as opposed to a longer dielectric steady substance. Electrically, the thin substance was selected to continue on the cross-pole diffusion to a lowest level [9]. A 2.65 mm broad co-planar waveguide line along a space of

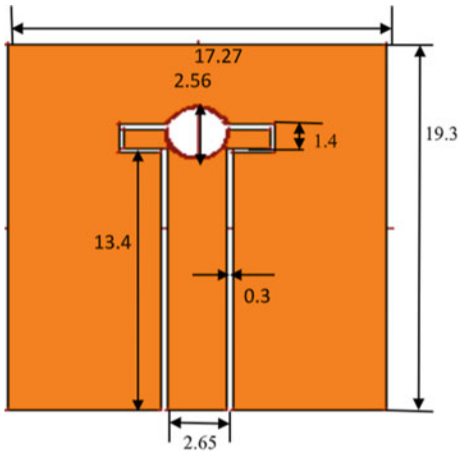


Fig. 1 Geometry of an antenna

0.2 mm is chosen for the supply line considering broad to detect lines massive than emitted divination at the frequency of delight direct guide across the mode antenna.

The geometry of slot antenna is shown in Fig. 1, and the radiation opening was bear with a clearance way from the rectangular ground and hence the reaction in dual beam with design in the broadside. The splitting opening was preserving with a detailed way from the rectangular ground to make a sensible ray on the two sides in the XY region. The extension in the middle of the feed and the pattern are stationary at 1λ to make well propagation design assessments with the end connector. It is taken from Fig. 1 that the propagation from a standard antenna is to be strengthened to approach the utilizer. The presented topology occupies lower capacity (volume) and constructs the plane design.

3 Results and Discussion

The designing, meshing, and simulation work of this proposed antenna structure are simulated using IE3D simulation software. The antenna is designed with dimensions and modified to get better results. The exploratory outcomes are demonstrated as follows. Reflection coefficient indicates the loss of power in a signal, which is returned/reflected because of a discontinuity in a fiber. The discontinuity may be a terminating load or by inserting a device in the transmission line.

It achieves return loss −19 dB at frequency of 4.1 GHz. The construction and observation of T-slot antenna are executed utilizing guide designs IE3D test system. The adjusted antenna plan along the simulations has been carried out, and its generation was affirmed. This antenna effectively operates at the resonant frequency of 4.0–4.2 GHz with the arrival loss return −19 dB. It is evident that the result uncovers full recurrence and transmission capacity of an antenna with a sensible exactness. It shows return loss of −19 dB as appeared in Fig. 2. The recreated VSWR quality of T-shape antenna demonstrates 1:2 VSWR proportion at 4.1 GHz. The operating frequency of the antenna lies in the range of 4–4.2 GHz having return loss −19 dB. It can be observed that the result shows full recurrence and transmission capacity of an antenna with a greater accuracy (Fig. 3).

Figures 4 and 5 show the radiation pattern of the slot antenna at both elevation and azimuth pattern. Therefore, for 5G applications, the proposed antenna is a perfect solution. The proposed antenna exhibits a maximum gain of 5.5 dBi at 4.1 GHz as exhibited in Fig. 6. The proposed antenna shows maximum radiation efficiency of 80% as shown in Fig. 7.

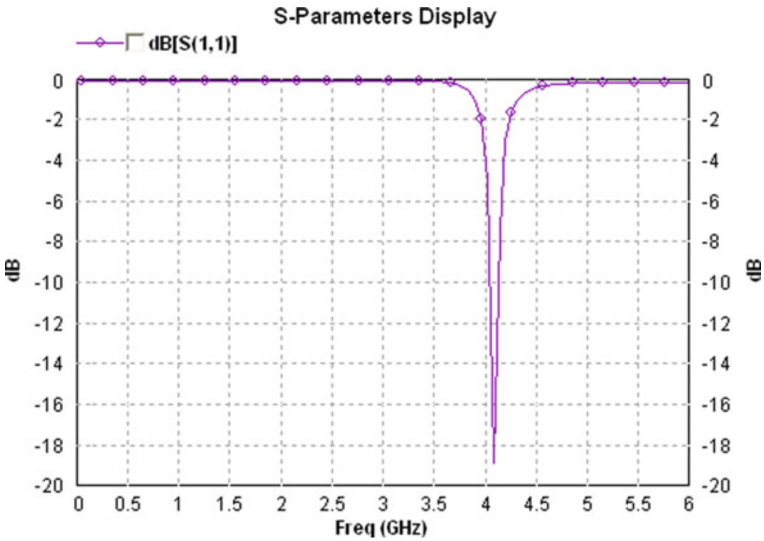


Fig. 2 S-parameter

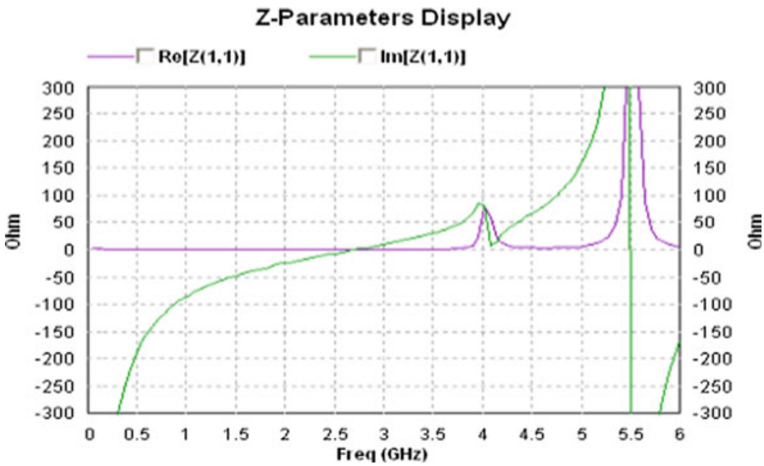


Fig. 3 Z-parameters display

Fig. 4 E-plane

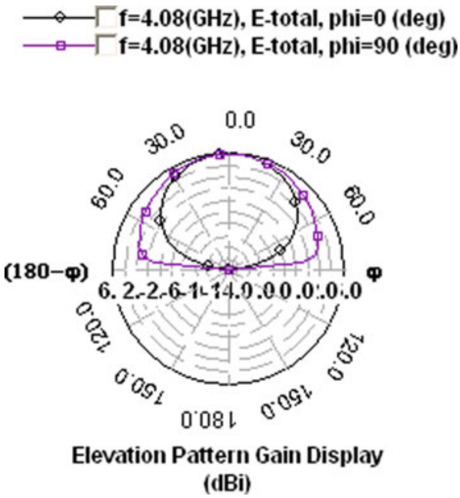
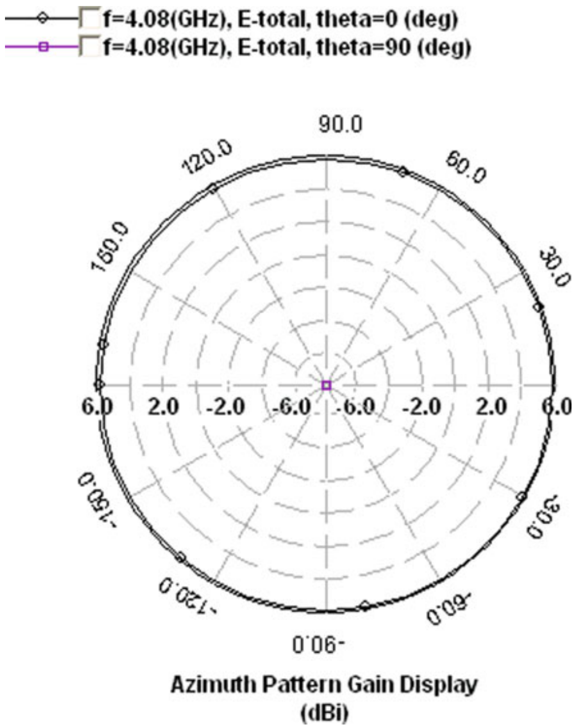


Fig. 5 H-plane pattern



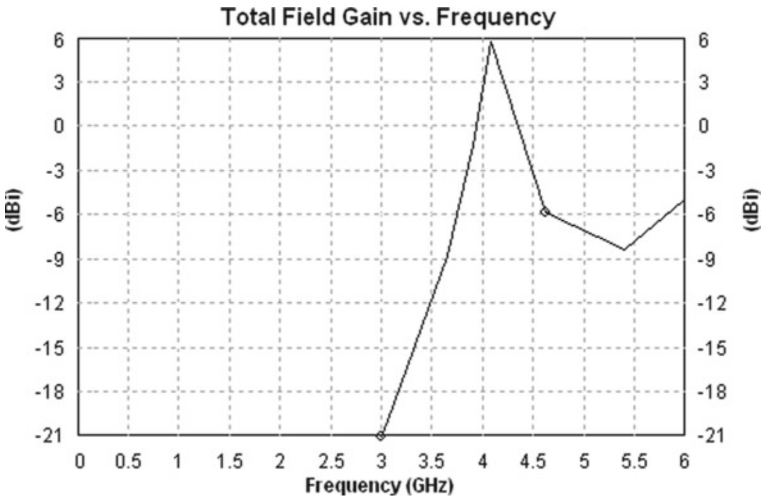


Fig. 6 Gain characteristics

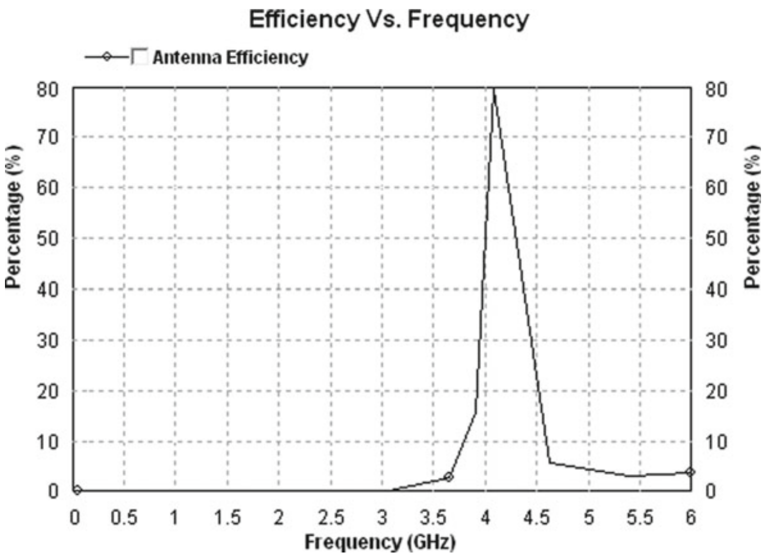


Fig. 7 Antenna efficiency

4 Conclusion

A wide impedance bandwidth of CPW fed antenna has been designed for 5G mobile applications. The corner twisted topology of an antenna has been investigated, due to high dielectric constant, the proposed antenna shows very good results at resonant frequency. The impedance bandwidth is greater when compared with 20% in the company of stable radiation design along with gain of 5.5 dBi over the conventional antennas.

References

1. Jun SY, Izquierdo BS, Parker EA (2019) Liquid sensor/detector using an EBG structure. IEEE Trans Antennas Propag 67(5):3366–3373
2. Prasad H, Jaleel Akhtar M (2017) Flexible substrate based CPW-fed antenna for WiMAX application. In: IEEE applied electromagnetic conference (AEMC), pp 1–2
3. Wang J, Wong H, Ji Z, Wu Y (2019) Broadband CPW-fed aperture coupled metasurface antenna. IEEE Antennas Wirel Propag Lett 18(3):517–520
4. Karthikeya GS, Abegaonkar MP, Koul SK (2019) CPW fed wideband corner bent antenna for 5G mobile terminals. IEEE Access 7:10967–10975
5. Zhou Y, Liu T, Tang F, Tinashe M (2019) An unlinkable authentication scheme for distributed IoT application. IEEE Access 7:14757–14766
6. Park E, Kim T, Han Y (2019) An efficient time-shifted random access scheme for cellular-based IoT networks. IEEE Commun Lett 23(3):522–525
7. Li YJ, Lu ZY, Yang LS (2019) CPW-fed slot antenna for medical wearable applications. IEEE Access 7:42107–42112
8. Ashok Kumar S, Shanmuganantham T (2018) Design and performance of textile antenna for wearable applications. Trans Electr Electron Mater 19(5):352–355
9. Cheng Y-F, Ding X, Xu X, Zhong X, Liao C (2019) Design and analysis of a bow-tie slot-coupled wideband metasurface antenna. IEEE Antennas Wirel Propag Lett 18(7):1342–1346

Energy-Conserving Cluster Method with Distance Criteria for Cognitive Radio Networks

M. S. Sumi and R. S. Ganesh

Abstract Due to development in wireless communication, advanced technology is necessary to meet the growing demands in this field. Since the number of users increases, spectrum resources have to be utilized in a planned and effective manner. Cognitive Radio paves the way for proper spectrum utilization. Balancing energy consumption in Cognitive Radio can be obtained by clustering methods. In this paper, cluster-based cooperative sensing is analysed. Energy-conserving cluster method with distance criteria is proposed, where cluster heads are selected based on their distance from primary user and fusion center, and cluster members are grouped to the nearest cluster head, thereby reducing reporting energy brings conservation in overall energy. Also a relay assistance approach is proposed where relay users placed between cluster head and fusion center assist cluster head in times of low energy. An energy consumption analysis is made to compare the proposed method with conventional methods. Simulations are performed using MATLAB R2016a software, and it is observed that the proposed method conserves energy and improves detection.

Keywords Cluster · Signal-to-noise ratio · Distance · Cooperative spectrum sensing · Fusion center · Energy

1 Introduction

As a result of the great advancement in technology in the area of wireless communications and mobile devices, more and more devices are brought into market day by day at an affordable price. Not only the price of the commodity is becoming reasonable, but also the technology is made to reach different people at wide and remote destinations to a greater extent. By the year 2021, the number of mobile devices utilized may exceed 28.3 billion [1]. As the number of devices utilized every day increases,

M. S. Sumi (✉) · R. S. Ganesh
Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu 629180, India
e-mail: sumirhn@gmail.com

R. S. Ganesh
e-mail: r_s_ganesh@rediffmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_52

the need for spectrum resources to be provided to combat the growth also increases. But the available spectrum is limited [2]. Also it has been observed that as per the Federal Communications Commission (FCC), spectrum is not completely utilized by their respective incumbent users at all durations and locations [3]. And that in addition to scarcity, spectrum also remains underutilized during most of the time durations. Hence, to meet the spectrum demands of new users and also to balance the spectrum underutilization problem, Cognitive Radio (CR) has been suggested for consideration. It is derived from the Latin word "Cognoscere", which means to know [4]. CR is defined as an intelligent radio whose transceiver automatically detects available free channels in wireless spectrum and instantly occupies them. The users who are licensed to make use of these spectrum resources are called primary users (PUs) or incumbent users. The main function of CR is to determine the availability of unused spectrum band which is also termed as white spaces or spectrum holes. The users that are permitted to temporarily make use of the vacant spectrum bands are called secondary users (SUs) or unlicensed users. Spectrum has to be monitored continuously, so that it enables the SUs to vacate the occupied frequency band and shift to another free channel once the incumbent user reappears in its band without creating any interference problem to PUs [5]. For proper operation of CR network, four functions are present, which are spectrum sensing, spectrum sharing, spectrum decision and mobility [4]. Spectrum sensing is the important functionality in which the SU senses the frequency band to check for the availability of PU signal. Also to avoid interference due to reappearance of PU back into its channel, sensing has to be performed continuously by the CR network.

The performance of sensing by a single SU is greatly affected due to problems like fading, shadowing and noise uncertainty conditions [6]. Therefore, to overcome this concern, cooperative spectrum sensing (CSS) is being considered, where SUs sense the channel simultaneously and either report their respective decisions regarding PU availability to a common fusion center (FC) or share among themselves. The former method is centralized type of CSS, where the local results from all SUs are combined by the FC using certain fusion rules, while the latter is distributed type, where each individual SU combines the received results to make the final decision. The fusion rules play a vital role in proper estimation of the global decision regarding PU status from the transmitted local results. They can be either hard or soft fusion rules [7]. In the case of hard fusion, each SU makes its local decision as a single bit indicating the presence or absence of PU as 1 or 0, respectively, and forwards its one-bit decision to FC. At the FC, single-bit results received are combined using methods like AND, OR, Majority fusion, etc.… to obtain the final decision [8]. In soft fusion, each SU transmits the entire energy signal estimated to FC, where based on the soft fusion method employed such as equal gain combining, selection combining and maximum ratio combining. The received signals are combined together to obtain the final test statistics and global decisions are made [9, 10].

Even though CSS improves the performance of sensing in cognitive networks, cognitive users undergo issues of increased energy consumption. Energy consumption of a network increases due to the increase in the number of SUs especially as sensing and reporting energies. Also a trade-off exists between spectrum efficiency

(SE) and energy efficiency (EE). Hence, to overcome these problems and improve the EE of the network, some advanced methods like clustering, censoring and relaying are being proposed in literature [11]. This paper explains the analysis of clustering in CR networks, and an energy-saving cluster method using distance criteria is being proposed to improve the energy conservation of the SUs employed in the CR networks. The rest of this paper is given as follows. Section 2 gives a description about the related works about cluster-based CSS in literature. Section 3 explains a simple cluster model enabling energy detection-based CSS which explains the analysis of clustering in CSS, the proposed method of EECMD is explained in Sect. 4 along with the energy consumption analysis, and also SE-EE trade-off is discussed. Simulations and discussions are given in Sect. 5. Finally, the paper is concluded with future work in Sect. 6.

2 Related Works

One of the advanced methods of improving spectrum and energy efficiencies in CR network is clustering. Different methods of cluster formation and cluster head (CH) selection are proposed and analysed in literature. A cluster-based CSS is proposed in [12] over fading reporting channels, where sensing performance is enhanced by decreasing the reporting errors through user selection diversity. The user having largest reporting gain is chosen as CH. A cluster method with centralized LEACH protocol is proposed in [13], where a reporting strategy is used to reduce the average number of reporting users as only CR with decision 1 is alone forwarded to CH. The adjacent SUs with similar signal-to-noise ratio (SNR) are grouped to form cluster in [14], and the best sensing node is selected as CH for a dynamic sensing interval. Here, frequency division-based parallel reporting mechanism is followed. A reputation-based hierarchical CSS is proposed in [15], where the reputation is a weight assigned to cluster decisions at FC based on history. Detection performance and robustness are improved. Still low SNR condition is not discussed. Equal gain combining method of fusion is used at cluster level and modified majority rule at FC. CH is selected in [16] based on reporting channel gain, distance from FC and energy levels of CRs. Sensing efficiency is improved due to the multiple hops employed. Still a trade-off persists between the energy consumed and sensing delay. As per the size of CR networks, the number of hops is optimized.

A multilevel hierarchical cluster method with double fusion is proposed in [17] to improve sensing performance and overhead. An iterative algorithm is used to obtain the suboptimal number of levels. Fusion rule and threshold are optimized. A three-stage cluster method is introduced in [18] to reduce sensing overhead, where during pruning stage, CR with improper sensing results is avoided and most reliable sensing user is chosen as CH during selecting stage. Following this, cluster is constructed with respect to correlation in sensing results. In [19], the cluster method depends on a multi-objective genetic algorithm to reduce power consumption and improve energy gain. CH selection depends on the user's distance from FC and energy level.

A hierarchy with soft–hard fusion is proposed in [20], where soft fusion is followed at the lower and hard fusion at higher levels. Also the number of levels and users at each level is optimized. Clustering is based on fuzzy c-means method in [21]. Spatial correlation is also employed here, where uncorrelated SUs alone report sensing results to CH, thereby saving energy.

In the merged clustering proposed in [22], if all SUs involve in clustering, they are grouped based on their distance. If only subset of users is alone considered, cluster formation depends on SNR of the users with respect to PU. The clustering algorithm in [23] depends on geographical position, channels available and database statistics, improving throughput and network stability. The resulting large-sized clusters reduce re-clustering. Reinforcement learning (RL) method is followed in [24] to determine cluster size, which considers the number of PUs and channels in the network and its activity levels. Due to larger clusters, re-clustering and cluster overhead are reduced. Cluster formation in [25] depends on a weight factor shared by SUs with its single-hop neighbours. This weight is estimated considering the speed of the SU, number of channels available and interference to PU. CHs are chosen among the maximum weighted users.

3 System Model

In this work, cooperative sensing in CR network is performed with clustering methodology. Consider the system model for CR network with clustering as in Fig. 1. Here, the CR network is centralized and consists of a single PU sensed by N SUs which are grouped into K clusters. Each cluster has a CH for reporting its respective decision regarding PU status to a common FC. The FC makes the final decision regarding

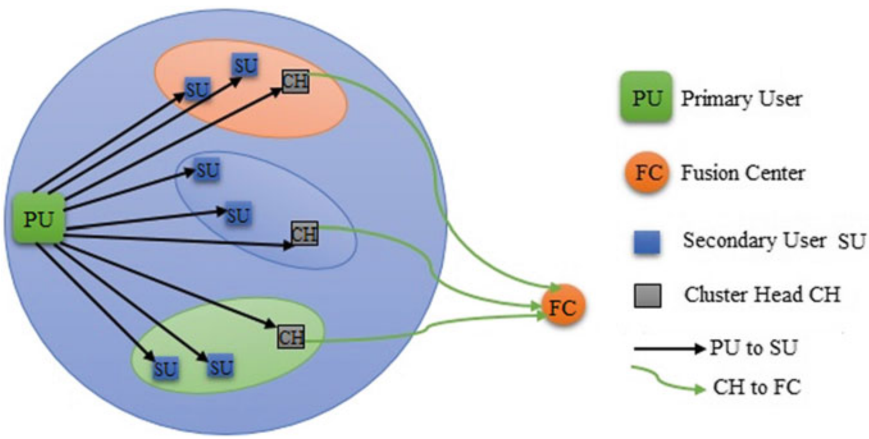


Fig. 1 System model representing clustering in CR network

channel availability to the clusters through their respective CHs. Two stages of fusion are followed here, one at the cluster level and another at the FC.

The sensing process is carried out by SUs for L samples by means of non-coherent energy detection method which is considered to be simple, as it does not need any prior information regarding PU signal [26]. In this method, the sensing SU receives the PU signal which is converted into test statistics, which can be compared with an estimated threshold value as in hard fusion or as such forwarded to FC as in soft fusion where the received test statistics will be combined and compared with the threshold determined. The two important parameters declaring the performance of sensing are global probabilities of false alarm (Q_{fa}) and detection (Q_{d}). Lower the value of Q_{fa} enables more channel availability for SUs, but at the same time may result in missed detections. Thus, there exists a trade-off between Q_{fa} and Q_{d}.

3.1 Cluster-Based Cooperative Sensing with SNR Criteria

The performance of CSS is mostly affected by factors such as sensing accuracy, fusion techniques employed, reliability of reporting channel and network overhead. Clustering method is an efficient approach to overcome the performance degradation in CR networks due to reporting channel uncertainty. Also overhead of control channel is addressed for larger networks with larger number of SUs [16]. The important phases in a cluster-based CSS are cluster formation, PU channel sensing, results forwarding and decision fusion. This cluster formation phase includes both CH selection and grouping of SUs into relevant clusters. Based on the methodology adopted in the above-said sub-phases, the performance of the CR network is improved in terms of throughput, energy efficiency and network reliability. Cluster formation as mentioned in Sect. 2 is mainly based on features such as channel gain, residual energy, SNR, sensing features, distance between SUs as well as from FC, available channels and geographical position [13, 14, 16, 19, 22]. Here, an energy-efficient cluster scheme for CSS is analysed and simulated, where SNR of the SU with respect to the received PU signal [16] and distance from FC are the criteria considered for cluster formation and CH selection. Initially, the optimal number of clusters K required is determined as in [27]. It is given as

$$K = N\frac{\ln\left(\left(1-\frac{Q_{\mathrm{fa}}}{N}\right)\left(1-0.5\mathrm{e}^{\frac{1}{2M\Upsilon1^{2}}\operatorname{erfc}^{-1}\left(\frac{2Q_{\mathrm{fa}}}{N}\right)}\right)\right)}{\ln\left(\frac{\mathrm{e}^{0.5\left(\operatorname{erfc}^{-1}\left(\frac{2Q_{\mathrm{fa}}}{N^{0.25}}\right)\right)^{2}}}{3.5\sqrt{3.14\,N}\operatorname{erfc}^{-1}\left(\frac{2Q_{\mathrm{fa}}}{N^{0.25}}\right)}+\frac{Q_{\mathrm{fa}}}{N^{0.25}}\right)} \tag{1}$$

where M is the time bandwidth product and $\Upsilon1$ is the average SNR value of the network with respect to received PU signal. Based on the optimal number of clusters obtained, K CHs are chosen for the CR network. The distance of each SU from FC is estimated, and K SUs with the smallest distance from the FC are selected as CH. Later among the non-CH users, SUs having similar SNR to any CH group into clusters with their respective CHs. Energy detection-based spectrum sensing is performed by all members of each cluster, and the results are reported to their respective CH. Certain fusion rules (hard or soft) are followed at the CH to obtain the detection results, which are then forwarded to the FC. Now the second stage of combining occurs at the FC employing either hard or soft fusion methods to obtain the global detection result. Flow chart given in Fig. 2 explains the process of cluster-based CSS with SNR criteria.

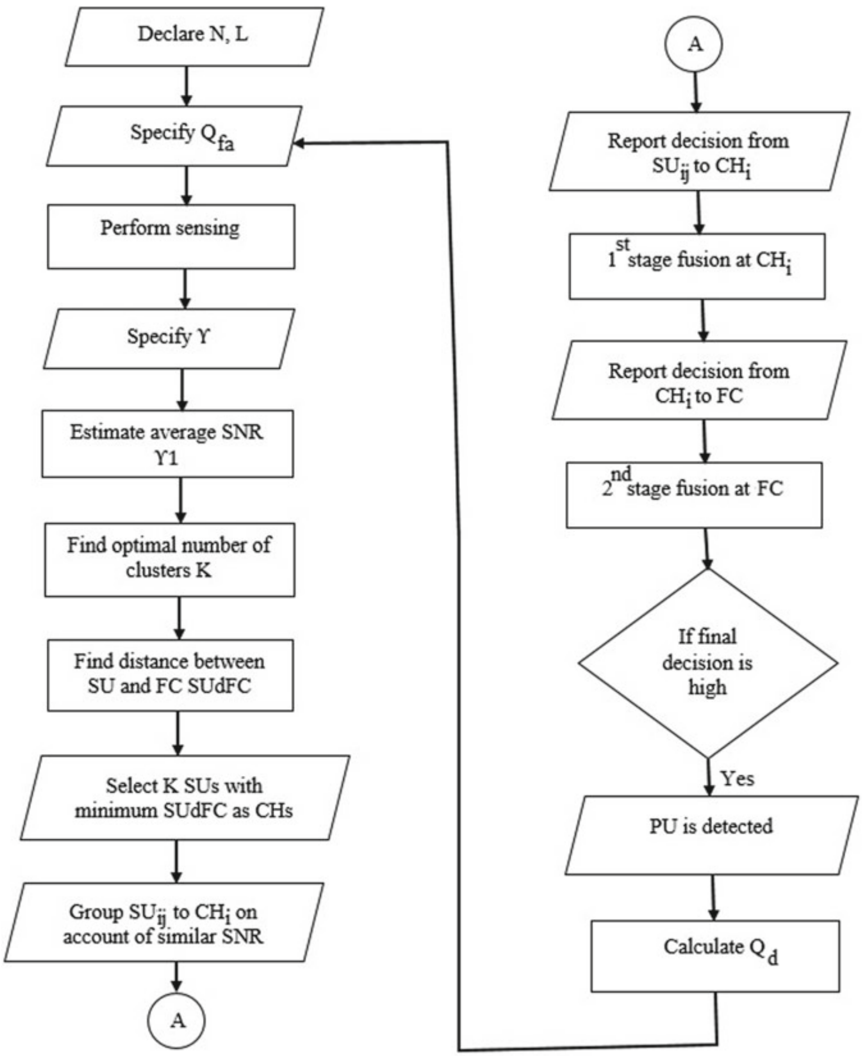


Fig. 2 Flow chart representing cluster-based CSS with SNR criteria

Since only the CHs closer to FC are forwarding the results to FC, it paves the way for reduction in reporting energy and hence energy can be saved. Since the cluster members are grouped with respect to similar SNR criteria, local decisions of members become similar, thereby improving the sensing reliability of the CR network.

4 Proposed Energy-Conserving Cluster Method with Distance Criteria (EECMD)

An energy-conserving cluster method with distance criteria (EECMD) is proposed. Here, the CHs are chosen in such a way that they are far away from PU as well as nearer to FC. As a result, reporting energies of CHs are minimized. As the CHs are far from PUs, their sensing results may get deteriorated due to distance and hence may affect the reliability if sensing. Hence, they are relieved from participating in sensing. As the number of sensing users reduces, sensing energy decreases. Initially after determining the optimal number of clusters K using (1), the distance of each SU with respect to both PU (SUdPU) and FC (SUdFC) is estimated. Then, K SUs which have maximum SUdPU and minimum SUdFC are selected as CHs. Now for grouping the cluster members, the distance of each non-CH users from the CHs is estimated and the SUs are added to the cluster with the nearest CH. This grouping is limited at each cluster when the maximum number of cluster members (CMs) for each cluster is reached. Even though distance criteria is a popular method for cluster head selection in literature [22], here the SUs are also grouped into cluster on the basis of their distances with respect to selected CHs. And the CHs selected are not only nearer to FC but also far away from PU. Also relay assistance has been introduced to the required CHs. The system model given in Fig. 3 represents the proposed model of energy-conserving cluster method with distance criteria (EECMD) for CR networks.

The CMs of each cluster report their decisions to their respective CH with minimum energies as their reporting distances with respect to CH are reduced. As a result in this method of clustering, both sensing and reporting energies of CMs and CHs are reduced, thereby paving the way for energy efficiency. In addition to bringing energy conservation to the CR network by the proposed clustering method, single-user energies can be conserved by introducing relay nodes for reporting. Relay nodes employed here do not involve in channel sensing, and hence, their sensing energies are not exhausted. They can be placed in the network to assist the CHs for reporting to FC. As a result, overall energy consumption of network is not improved. Still an alternate reporting path is established between CHs and FC with the help of relay users, so that the CHs can be prevented from draining out of energy. As a result, network stability is not affected. Also as the CMs and CHs report to their respective destinations at minimum distance, the reporting signal quality is not affected and hence reliability is improved. Flow chart in Fig. 4 explains the process of EECMD for CR networks. Hence, in this EECMD method, by reducing the number of sensing

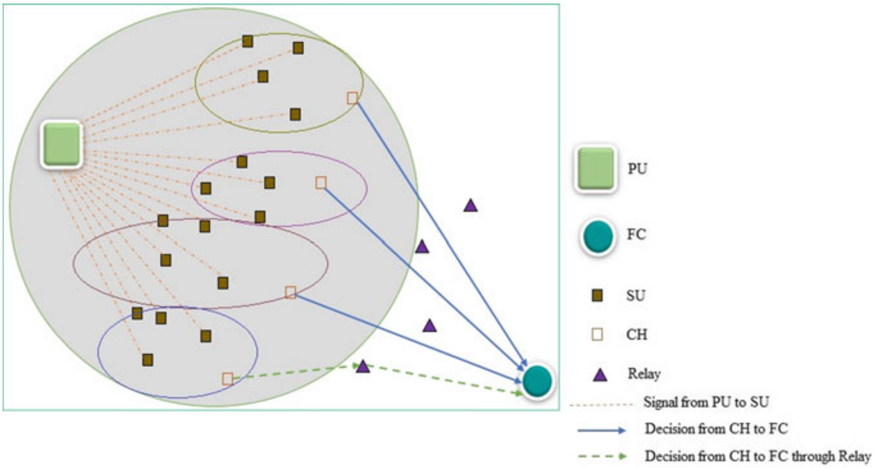


Fig. 3 System model for energy-conserving cluster method with distance criteria (EECMD)

users, sensing energy is conserved. Also by reducing the reporting distance to both CH and FC, reporting energies are reduced, thereby saving energy.

4.1 Energy Consumption Analysis

The CSS process in CR network is carried out for a time frame of period T_F in seconds for a data rate R, in bits per second and is given as in [28].

$$T_F = T_s + T_r + T_t \tag{2}$$

where T_s is the total sensing time which is assumed to be fixed and is a common sensing period for all SUs to perform sensing. T_r is the reporting time, shared among all reporting users equally using, and is given as NT_u, where T_u is the reporting time for one-bit result (hard fusion). T_t is the remaining time period available for data transmission for the SUs, once an idle channel is sensed. As the sensing period is fixed, the remaining time period available is shared for reporting and data transmission. The sensing and reporting/transmission powers are represented as P_s and P_t, respectively. Also as the number of sensing users is reduced, energy consumption reduces, paving the way for energy efficiency. Since the reporting time is reduced in the proposed method due to decrease in reporting distance, more time is available for data transmission which improves throughput.

P_{rsui} is the reporting power for ith SU to CH. SUdCH is the distance between SU and CH. P_{rchj} is the reporting power for jth CH to FC, and CHdFC is the distance between CH and FC distance. The number of sensing users is N_s which is given as

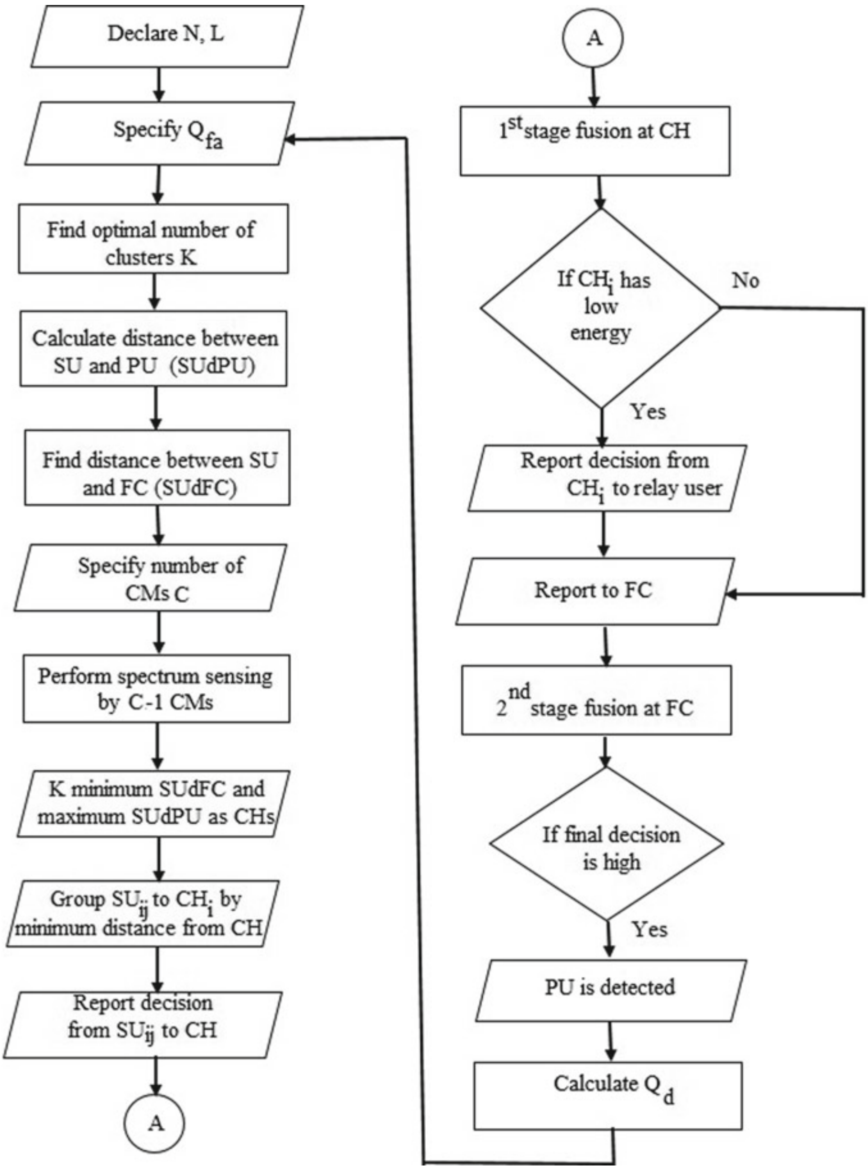


Fig. 4 Flow chart representing energy-conserving cluster method with distance criteria (EECMD)

$$N_s = N - K \tag{3}$$

Sensing energy of CMs is given as

$$E_s = N_s T_s P_s \tag{4}$$

Reporting energy of CMs is given as

$$E_r = \sum_{j=1}^{N_s} T_u P_{\mathrm{rsu}i} \tag{5}$$

Reporting energy of CHs is given as

$$E_{\mathrm{rc}} = \sum_{j=1}^{K} T_u P_{\mathrm{rch}j} \tag{6}$$

Total sensing and reporting energies for a CR network following EECMD method of clustering are given as

$$E_t = E_s + E_r + E_{\mathrm{rc}} \tag{7}$$

4.2 Relay Assistance in EECMD

The individual energy levels of a single CH can be conserved by introducing relay assistance in EECMD. These relay users (RUs) are not sensing users. If each sensing user is observed to be in a condition of low energy, the nearest available RU can be selected and an alternative reporting path can be established between the low-energy CH and relay to report their respective cluster decisions to FC. This approach is not intended to bring overall energy conservation in the entire system, but supports individual reporting users to conserve energy. Algorithm 1 given below describes the process of relay assistance in EECMD

Algorithm 1

1. Specify energy values of CHs
2. Determine threshold energy value
3. Identify CHs lower than threshold energy value
4. Specify positions of available relay users RU
5. Identify the nearest RU for lower energy CH
6. Establish reporting part between CH and FC through selected RU
7. Compare the energy consumed in reporting to FC by CH with and without relay assistance.

4.3 Spectrum-Efficiency (SE) and Energy-Efficiency (EE) Trade-off

While advancing wireless devices' technology towards 5G and 6G [29], both spectrum and energy efficiencies are to be essentially addressed. Maximizing SE requires the increase in achievable data at a frequency band, and EE represents the number of bits successfully transmitted to the total energy used [30]. Optimizing any one efficiency degrades the other resulting in a trade-off [31]. Increasing the number of cooperating users improves SE as spectrum usage is increased, but more number of sensing users also result in increase in energy consumption. Hence, a trade-off occurs. Hence, optimizing the number of sensing users stabilizes the trade-off, thereby improving both efficiencies [32].

5 Simulations and Discussions

In this section, the detection performance of cognitive network with cluster-based CSS is simulated and results obtained are analysed. Energy detection is employed at all sensing users, and hard fusion methods are used at both levels of the combining to achieve the channel status. A 100×100 region with N CR users and a single PU and centralized FC are considered at location (50, 200). Simulations are carried out for the cluster methods by varying the number of SUs for different ranges of SNR values and compared with the conventional non-cluster CSS methods. MATLAB R2016a software is being using to perform the simulations.

5.1 Simulation of Cluster-Based CSS with SNR Criteria

For 20 SUs, the optimal number of clusters using (1) is obtained as 5. The distance of each SU from FC is determined, and five nearest SUs are selected as CHs. Their SNR values are observed, and the non-CH members are grouped to CH having similar SNR values. For different ranges of SNR, detection results are obtained and compared. Figures 5 and 6 represent the comparison of cluster-based CSS with conventional CSS for 20 SUs using AND and Majority fusion methods for SNR range of SUs lying between -20 and -1. From the results obtained, it is that there is an increase in detection probability for the simulated cluster-based CSS method with SNR criteria than the conventional CSS methods for both AND and Majority fusion methods.

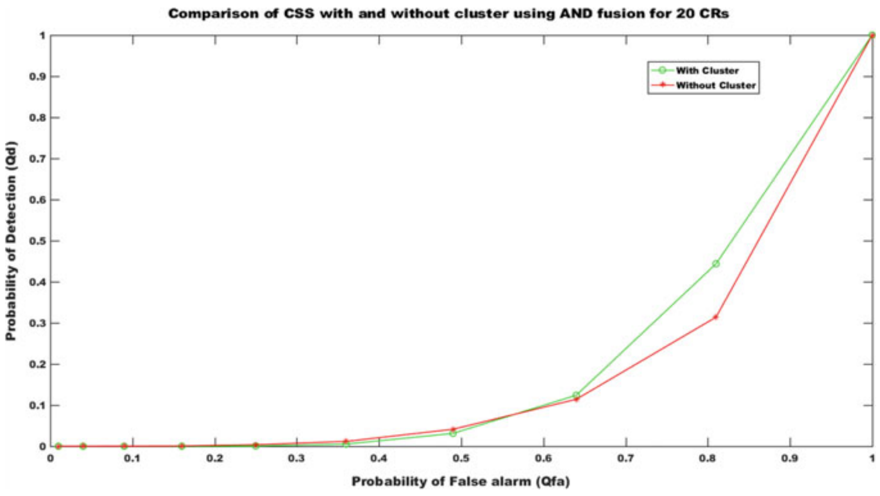


Fig. 5 Comparison of cluster-based CSS with conventional CSS for 20 CRs using AND fusion

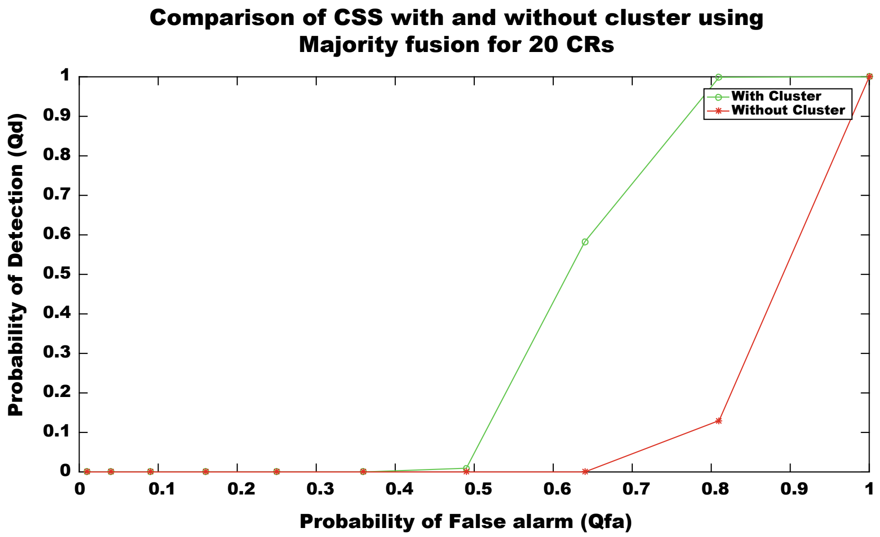


Fig. 6 Comparison of cluster-based CSS with conventional CSS for 20 CRs using Majority fusion

5.2 Simulation of Energy-Conserving Cluster Method with Distance Criteria

After obtaining the optimal number of clusters for both 10 and 20 SUs, the distance of each SU from the FC and PU is estimated. The optimal number of SUs with maximum distance from PU and minimum distance from FC is formed as CHs and the SUs which are at a nearer distance to any CH link with it to form clusters, simultaneously considering the maximum number of non-CH members for each cluster. Table 1 given below gives a description of the EECMD method parameters and detection probability compared with conventional CSS and cluster-based CSS with SNR criteria for 20 SUs. Also Fig. 7 represents cluster formation using EECMD.

Table 1 Comparison of detection probability of EECMD with conventional CSS

Method	No. of sensing users	No. of reporting users	No. of clusters	Detection probability
Conventional CSS	20	20	–	0.0000 0.0000 0.0001 0.0007 0.0030 0.0138 0.0389 0.1203 0.3164 1.0000
Cluster-based CSS with SNR criteria	20	15 CM 5 CH	5	0.0000 0.0000 0.0000 0.0000 0.0006 0.0052 0.0283 0.1322 0.4303 1.0000
EECMD	15	15 CM 5 CH	5	0.0000 0.0009 0.0071 0.0352 0.1089 0.2409 0.4136 0.6059 0.8553 1.0000

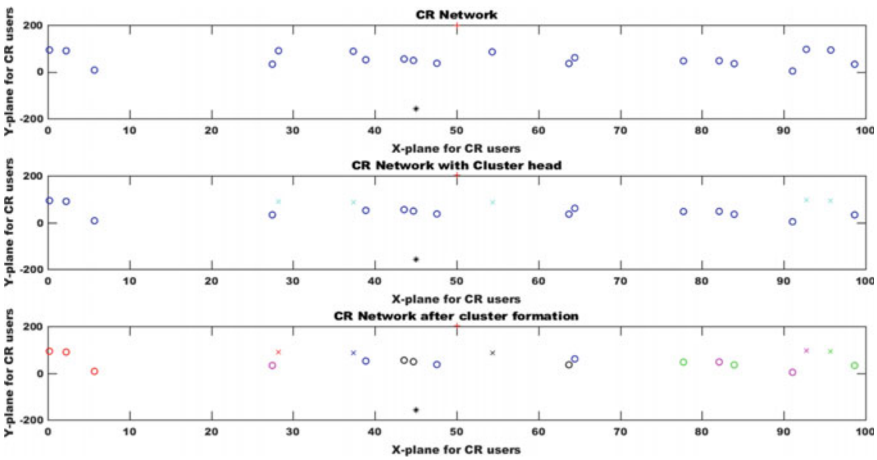


Fig. 7 Cluster formation using EECMD for 20 users

Table 2 Cluster formation using EECMD

Cluster	Distance of CH from FC	Distance between CH and non-CH members
1	102.0335	42.7534, 69.2251, 47.8595
2	102.4650	14.8893, 78.2206, 59.7109
3	103.1979	78.8642, 13.5575, 41.8000
4	106.4482	43.5709, 57.5248, 92.2822
5	108.3185	42.8215, 51.4172, 68.2374

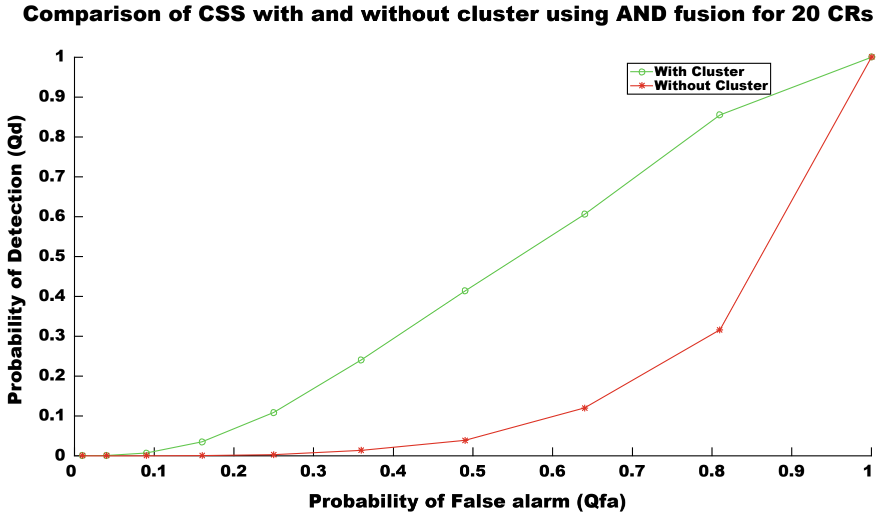


Fig. 8 Comparison of EECMD with conventional CSS for 20 CRs using AND fusion

For 20 SUs, five clusters are formed as mentioned in Table 2. Figures 8 and 9 represent the comparison of detection probability of EECMD process with conventional CSS for 20 and 10 SUs, respectively, using AND fusion method.

From the tables and above figures, it is observed that the proposed method brings increase in detection performance compared with the conventional method and cluster method with SNR criteria.

5.3 Energy Consumption Analysis of Simulated Methods

With reference to the values assumed to parameters in Table 3 given, the amount of energy consumed by the different methods of clustering is compared for a CR network with 20 SUs. Also Table 4 provides the comparison of energy consumption for the proposed EECMD method with the cluster method with SNR criteria.

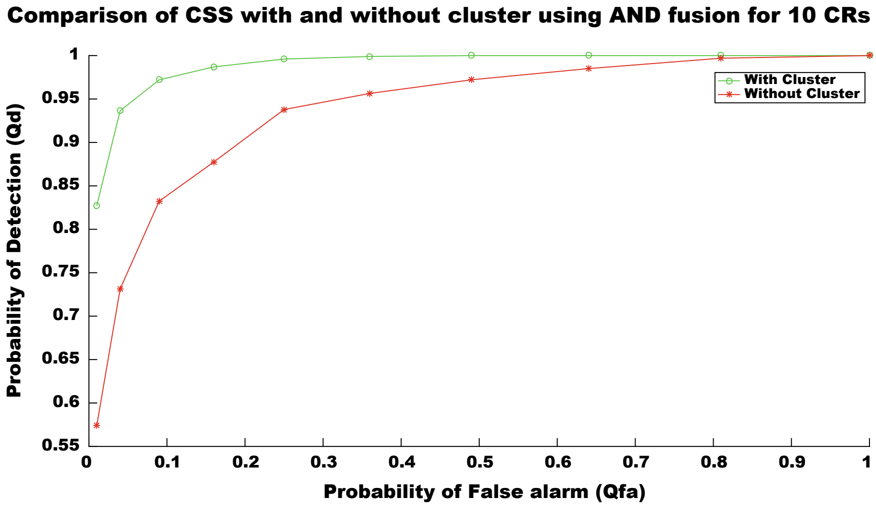


Fig. 9 Comparison of EECMD with conventional CSS for 10 CRs using AND fusion

Table 3 Parameters and values for energy consumption estimation

Parameters	Values
T_s (in s)	0.001
T_u (in s)	0.0001
P_s (in W)	0.010
P_{rsui} (in W)	SUdCH/1000
P_{rchj} (in W)	CHdFC/1000
N	20
K	5
N_s	15

Table 4 Comparison of energy consumption analysis

Cluster formation method	Sensing energy (mJ)	Reporting energy CM to CH (mJ)	Reporting energy CH to FC (mJ)	Total energy (mJ)
Cluster-based CSS with SNR criteria	0.2	0.080842	0.05908	0.34035
Energy-conserving cluster method with distance criteria (EECMD)	0.15	0.078657	0.05908	0.28816

Based on the values from table [33], energy consumed for channel sensing as well as reporting results by both CMs and CHs are estimated, and their combined energy consumption is obtained using (3–7). It is observed from Table 4 that the proposed method has low-sensing energy consumption as the CHs are not participating in sensing. And also the reporting energy is reduced due to reduction in reporting distance between the CHs and FC as well as between CMs and CHs. Hence, the proposed method brings overall energy conservation.

Even though various clustering methods improve the efficiency of sensing in CR [16, 23, 24], the proposed method brings reduction in energy as cluster formation as well as CH selection brings energy conservation, without affecting the reliability of sensing.

5.4 Relay Assistance and Energy Consumption Analysis

The energy values of the CHs are specified as $E_{\mathrm{CH}i} = [900, 575, 850, 780, 825]$, and the energy threshold is found to be 786. Energy values of CH 2 and 4 are observed to be lower than the threshold value and can therefore be supported with relay assistance. The relay users are assumed to be positioned at [(35, 180), (36,182), (37,184), (38,186), (39,188)]. After estimating the distance between CHs and RUs as well as RUs and FC, it is observed that CH2 can be assisted with RU1 and CH4 can be assisted with RU2. As a result, both the low-energy cluster heads can save energy as their respective reporting distances are minimized.

6 Conclusions

In this work, cluster-based method for cooperative spectrum sensing is analysed, and the different methodologies followed for cluster formation and cluster head selection are observed. A cluster-based CSS with SNR criteria is simulated and compared with the conventional method of CSS. From the simulation, it is observed that clustering in CSS not only saves energy, but also improves detection performance of the CR network. An energy-conserving cluster method with distance criteria (EECMD) is proposed, where the CHs are selected based on their distance from PU and FC. Also the non-CH members of each cluster are formed considering their distance with CHs. Here, the CHs in all clusters are not participating in sensing. As a result, both sensing and reporting energies are reduced. Relay users can be introduced in the network to support the low-energy CHs from getting drained out of energy. Cluster heads 2 and 4 are noted to have low energy and are assisted with relay users 1 and 2, respectively, thereby their respective energy consumption can be minimized. This work can be further modified considering that the number of cluster members as not constant. Also the relay assistance can be further improved in such a way that both single-user energy and entire network energy can be conserved. Optimization of the number of sensing users per cluster for addressing SE-EE trade-off can be obtained.

References

1. Gupta MS, Kumar K (2019) Progression on spectrum sensing for cognitive radio networks: a survey, classification, challenges and future research issues. J Netw Comput Appl 143:47–76
2. Akyildiz IF, Lee WY, Vuran MC, Mohanty S (2008) A survey on spectrum management in cognitive radio networks. IEEE Commun Mag 46(4):40–48
3. Salahdine F (2017) Spectrum sensing techniques for cognitive radio networks. arXiv preprint arXiv:1710.02668
4. Garg R, Saluja N (2016) Spectrum sensing in cognitive radio: components and methodologies. In: World congress on engineering and computer science, San Francisco, USA
5. Alias DM (2016) Cognitive radio networks: a survey. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 1981–1986
6. Akyildiz IF, Lo BF, Balakrishnan R (2011) Cooperative spectrum sensing in cognitive radio networks: a survey. Phys Commun 4(1):40–62
7. Teguig D, Scheers B, Le Nir V (2012) Data fusion schemes for cooperative spectrum sensing in cognitive radio networks. In: 2012 military communications and information systems conference (MCC). IEEE, pp 1–7
8. Hossain MS, Abdullah MI, Hossain MA (2012) Hard combination data fusion for cooperative spectrum sensing in cognitive radio. Int J Electr Comput Engg 2(6):811
9. Qin Q, Zhimin Z, Caili G (2009) A study of data fusion and decision algorithms based on cooperative spectrum sensing. In: 2009 sixth international conference on fuzzy systems and knowledge discovery, vol 1. IEEE, pp 76–80
10. Fu Y, He Z, Yang F (2017) A simple quantization-based multibit cooperative spectrum sensing for cognitive radio networks. In: 2017 14th international computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, pp 220–223
11. Sumi MS, Ganesh RS (2017) Performance enhancing techniques in cognitive radio networks. In: 2017 IEEE international conference on circuits and systems (ICCS). IEEE, pp 172–178
12. Sun C, Zhang W, Letaief KB (2007) Cluster-based cooperative spectrum sensing in cognitive radio systems. In: 2007 IEEE international conference on communications. IEEE, pp 2511–2515
13. Kozal AS, Merabti M, Bouhafs F (2012) Energy efficient clustering approach for cooperative spectrum sensing in cognitive radio networks. In: The 13th annual post graduate symposium on the convergence of telecommunications, networking and broadcasting (PGNet)
14. Nguyen-Thanh N, Koo I (2013) A cluster-based selective cooperative spectrum sensing scheme in cognitive radio. EURASIP J Wire Commun Network 2013(1):176
15. Huifang C, Lei X, Xiong N (2014) Reputation-based hierarchically cooperative spectrum sensing scheme in cognitive radio networks. China Commun 11(1):12–25
16. Kozal AS, Merabti M, Bouhafs F (2014) Spectrum sensing-energy tradeoff in multi-hop cluster based cooperative cognitive radio networks. In: 2014 IEEE conference on computer communications workshops (INFOCOM WKSHPS). IEEE, pp 765–770
17. Awin FA, Abdel-Raheem E, Ahmadi M (2015) Optimization of multi-level hierarchical cluster-based spectrum sensing structure in cognitive radio networks. Digit Sig Process 36:15–25
18. Jiao Y, Yin P, Joe I (2016) Clustering scheme for cooperative spectrum sensing in cognitive radio networks. IET Commun 10(13):1590–1595
19. Salah I, Saad W, Shokair M, Elkordy M (2016) Minimizing energy of cluster-based Cooperative spectrum sensing in CRN using multi objective genetic algorithm. In: 2016 12th international computer engineering conference (ICENCO). IEEE, pp 178–183

20. Salout N, Awin FA, Alqawasmeh AF, Abdel-Raheem E (2017) Hierarchical cluster-based cooperative spectrum sensing in cognitive radio employing soft-hard combination. In: 2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE). IEEE, pp 1–4
21. Bhatti DMS, Shaikh B, Zaidi SIH (2017) Fuzzy c-means and spatial correlation based clustering for cooperative spectrum sensing. In: 2017 international conference on information and communication technology convergence (ICTC). IEEE, pp 486–491
22. Cichoń K, Kliks A, Bogucka H (2017) Energy-efficient cooperative spectrum sensing with a merged clustering measure. In: 2017 IEEE 13th international conference on wireless and mobile computing, networking and communications (WiMob). IEEE, pp 38–43
23. Li X, Xiao Q, Zhang Y (2017) Clustering algorithm for cognitive radio. Boletín Técnico 55(12):491–497
24. Javed Z, Yau KLA, Mohamad H, Ramli N, Qadir J, Ni Q (2017) RL-Budget: a learning-based cluster size adjustment scheme for cognitive radio networks. IEEE Access 6:1055–1072
25. Kumar S, Singh AK (2018) A localized algorithm for clustering in cognitive radio networks. J King Saud Univ Comput Inf Sci
26. Pandya P, Durvesh A, Parekh N (2015) Energy detection based spectrum sensing for cognitive radio network. In: 2015 fifth international conference on communication systems and network technologies. IEEE, pp 201–206
27. Guo C, Peng T, Xu S, Wang H, Wang W (2009) Cooperative spectrum sensing with cluster-based architecture in cognitive radio networks. In: VTC spring 2009-IEEE 69th vehicular technology conference. IEEE, pp 1–5
28. Sarijari MA, Abdullah MS, Janssen GJ, Van der Veen AJ (2015) On achieving network throughput demand in cognitive radio-based home area networks. EURASIP J Wire Commun Network 2015(1):221
29. Tariq F, Khandaker M, Wong KK, Imran M, Bennis M, Debbah M (2019) A speculative study on 6G. arXiv preprint arXiv:1902.06700
30. Kim SW, Cui C, Neihart N (2015) Optimum sensing bandwidth for energy-efficient cognitive radio communications. In: 2015 IEEE international conference on ubiquitous wireless broadband (ICUWB). IEEE, pp 1–5
31. Hu H, Zhang H, Liang YC (2015) On the spectrum-and energy-efficiency tradeoff in cognitive radio networks. IEEE Trans Commun 64(2):490–501
32. Hsu CC, Chang JM, Chou ZT, Abichar Z (2013) Optimizing spectrum-energy efficiency in downlink cellular networks. IEEE Trans Mob Comput 13(9):2100–2112
33. Althunibat S, Granelli F (2014) Energy efficiency analysis of soft and hard cooperative spectrum sensing schemes in cognitive radio networks. In: 2014 IEEE 79th vehicular technology conference (VTC Spring). IEEE, pp 1–5

Quality Enhancement of Low Bit Rate Speech Coder with Nonlinear Prediction

Ancy S. Anselam, Sakuntala S. Pillai, and K. G. Sreeni

Abstract Toll quality speech codec design with a low bit rate is really a challenging task in modern communication because of the drastic increase in end-users in social networks. Most of the low bit rate speech codecs are based on linear prediction. The code-excited linear prediction codec (CELP) gives good quality decoded speech at a lower bit rate of 4.8 Kbps. But, it neglects the natural nonlinear effects present in speech production process. So, some adaptive techniques are to be used to make the system nonlinear to perform better than linear prediction speech codecs. An adaptive technique with nonlinear prediction of speech, based on truncated Volterra series, is used to generate the nonlinear prediction coefficients. The generated nonlinear prediction coefficients are implemented in G723.1 CELP codec to introduce code-excited nonlinear prediction (CENLP) codec. Advancements in the performance are evaluated using subjective and objective quality measures and compared with the normal G723.1 CELP codec.

Keywords CELP · Nonlinear long-term prediction · Pitch period · Prediction gain · Volterra series

1 Introduction

To accelerate the efficient transmission and compact storage capabilities, some systems need low bitrate and low delay speech coders with very good quality. Commonly

A. S. Anselam (✉) · S. S. Pillai
Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, India
e-mail: ancy.anselam@mbcet.ac.in

S. S. Pillai
e-mail: sakuntala.pillai@gmail.com

K. G. Sreeni
Department of Electronics and Communication Engineering, College of Engineering Trivandrum, Thiruvananthapuram, India
e-mail: sreenikg79@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_53

Table 1 Values of quality measures and delay of existing LPC-based speech codecs

Linear prediction codecs parameters	MELP 2.4 Kbps	CS-ACELP 8 Kbps	LD_CELP 16 Kbps	CELP 4.8 Kbps
Delay (ms)	122.5	10	0.625	37.5
MOS score	2.9	3.98	3.95	3.15

used speech coding algorithms are based on linear prediction. In this, the current sample is predicted using future or past speech samples. Various speech coding algorithms with linear prediction are linear prediction coder [1], code-excited linear prediction (CELP) [2], low delay CELP [3], mixed excitation linear prediction (MELP) [4] and conjugate structure—Algebraic CELP [5]. Among these speech coders, the CELP 4.8 gives the moderate quality and delay. The trade-off between the quality measure (MOS score) and delay of LPC-based speech codecs is shown in Table 1. It shows that to achieve toll-quality decoded speech with low bitrate and low delay alterations in CELP 4.8 are essential.

During the development of speech production models, it is assumed that, a power source (i.e. lungs) produces a puff of air (i.e. sound) and the surroundings (i.e. vocal tract) shapes or filters the same before it reaches the lip end [6]. Linear prediction model works with an assumption that the lungs and the vocal tract are independent, which allows the separation of the source and filter in the model (i.e. source-filter model). It removes the interdependence of the excitation source and the vocal tract. But, this assumption will not work well for all speech processing applications. Therefore, adaptive procedures that consider the nonlinear generation process (dependency lungs and vocal tract) of the speech signal need to be considered, which may outperform the linear techniques, enabling better performance speech processing applications.

A variety of procedures dealing with nonlinear speech prediction are reported in the literature. Among them, the most widely used technique is nonlinear adaptive filtering using Volterra series model [7]. This deals with the short-term and long-term prediction (LTP) of speech based on truncated Volterra series. In this work, the nonlinear characteristics of speech signal are modelled using the Volterra series for short-term and long-term signals [6]. Then, the nonlinear predictor coefficients are obtained using the adaptive techniques. The adaptive technique used is 'recursive least squares (RLS) algorithm' which offers fast convergence and good error performance.

2 Code-Excited Linear Prediction Speech Codec (CELP)

CELP is viewed as the upgraded version of LPC model, in which codebooks used are vector quantization based. It follows analysis-by-synthesis approach [2]. The formant synthesis filter function is:

$$H_f(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{i=1}^{M} a_i z^{-i}} \tag{1}$$

$A(z)$—the system function of analysis filter and M—order of the filter. The amount of noise in the reconstructed speech is adjusted by using a weighing filter. The system function used for the implementation of perceptual weighing filter is:

$$W(z) = \frac{A(z)}{A(z/\gamma)} = \frac{1 + \sum_{i=1}^{M} a_i z^{-i}}{1 + \sum_{i=1}^{M} a_i \gamma^i z^{-i}} \tag{2}$$

where γ—constant with values [0, 1]. If it is sampled at 8000 samples per second, γ is between 0.8 and 0.9. Best matching vector is selected from codebook by using MMSE criterion [8].

Figure 1 shows CELP encoder design flow. The speech in PCM format is given to the encoder [2]. Linear prediction is done in two different levels—short-term and long-term. To obtain formant information, short-term prediction is done on speech frames and long-term prediction is done on sub-frames to get pitch and intensity. Finally, CELP bitstream formed by encoding the linear prediction coefficients (LPCs), long-term LP parameters, gain and codebook excitation index.

Figure 2 shows the design flow of CELP decoder. From the received bitstream, parameters are decoded and extracted. The decoder takes the code vector from the codebook corresponding to the index received. The obtained code vector is scaled by the matching gain. The decoder filters the scaled code vector by pitch synthesis and formant synthesis filters with the help of received parameters. Then, the synthesized speech is passed through the post-filter to improve the perceptual quality.

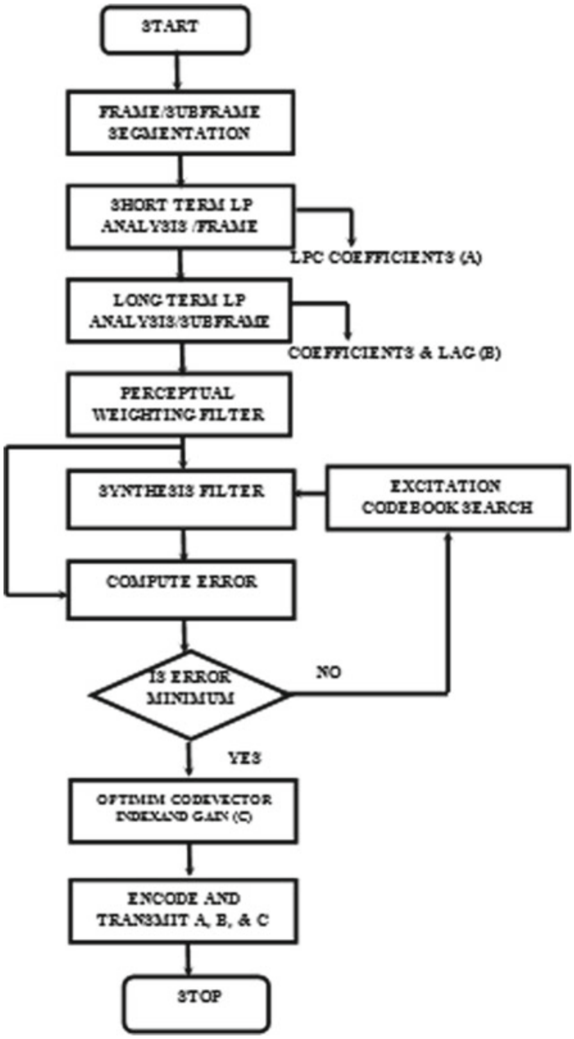
3 Estimation of Nonlinear Prediction Coefficients Using Volterra Series

For the approximation of nonlinear behaviour in systems, the Volterra series model is used similar to Taylor series. It has the ability to capture 'memory' effects, and this makes the model different from Taylor series. The input–output relation of discrete-time Volterra series with infinite memory is as follows [6, 7].

$$\begin{aligned} y(n) = h_0 &+ \sum_{k=0}^{\infty} h_1 \,.\, s(n-k) + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} h_2(i,j) \,.\, s(n-i) \,.\, s(n-j) \\ &+ \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} h_2(i,j,k) \,.\, s(n-i) \,.\, s(n-j) \,.\, s(n-k) + \cdots \end{aligned} \tag{3}$$

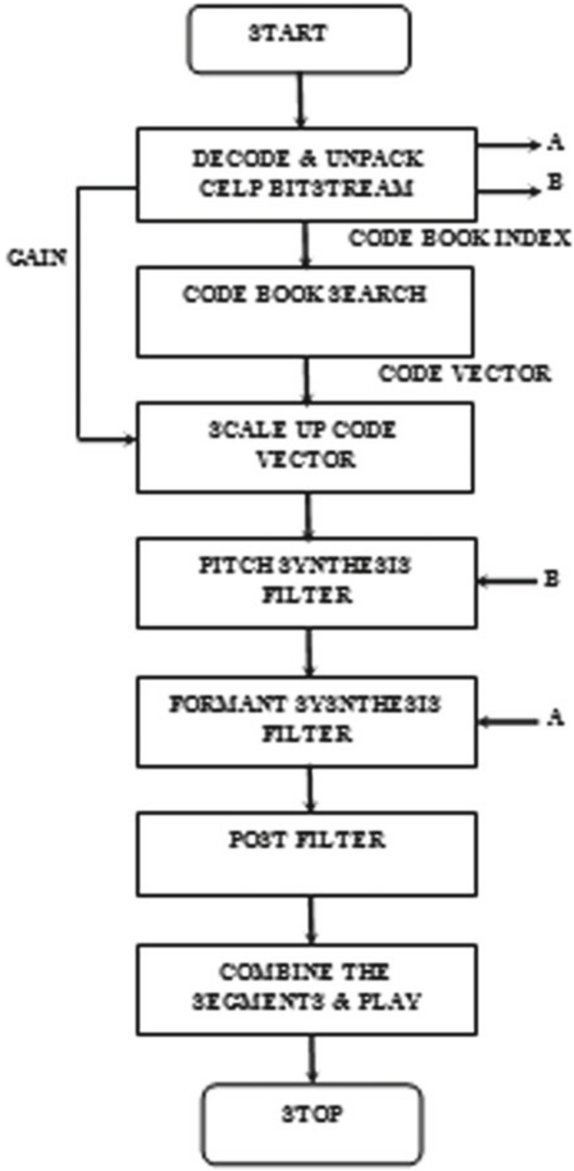
where $s(n)$—input signal, $y(n)$—output of the model, h_0—bias coefficient, h_1—linear coefficients, h_2—quadratic coefficients, h_3—cubic coefficients.

Fig. 1 CELP encoder



Here, the output is expressed as the linear combination of nonlinear functions of the input signal. So, the adaptation algorithms valid in linear case can be extended to Volterra series. Both LMS and RLS are used to identify unknown parameters. With the order of nonlinearity, the number coefficients of increases exponentially resulting in a significant increase in system complexity. So, the Volterra series is truncated to low linearity orders.

Fig. 2 CELP decoder



3.1 Short-Term Speech Prediction: Second-Order Volterra Model

Volterra series model-based nonlinear predictor estimates the current signal value as a linear combination of past signal values and linear combinations of products of past signal values. The second-order Volterra model treats the system with the first- and second-order kernels only and the predicted signal is given by

$$\hat{s}(n) = h_0 + \sum_{k=1}^{P} h_1(k)s(n-k) + \sum_{i=1}^{M}\sum_{j=1}^{M} h_2(i, j)s(n-i)s(n-j) \tag{4}$$

where $\hat{s}(n)$ is the estimate of $s(n)$. This model consists of two parts: a linear part of prediction order P with coefficients $h_1(k)$ and a quadratic part of prediction order M with coefficients $h_2(i, j)$. With the coefficient symmetry, i.e. $h_2(i, j) = h_2(j, k)$, the overall number of coefficients for the second-order Volterra predictor is

$$n_c = P + \frac{M(M+1)}{2} \tag{5}$$

The system model of a second-order Volterra-based nonlinear predictor with prediction order $P = 2$ is given in Fig. 3.

Prediction error for the filter model is,

$$\begin{aligned} e(n) = s(n) - \hat{s}(n) = s(n) - h_0 - \sum_{k=0}^{P} h_1(k)s(n-k) \\ - \sum_{i=0}^{M}\sum_{j=0}^{M} h_2(i, j)s(n-i)s(n-j) \end{aligned} \tag{6}$$

The filter coefficients are computed by minimizing the criterion function based on RLS algorithm,

$$J(n) = \sum_{k=0}^{n} \rho^{n-k} \,.\, \mathrm{e}^2(k) \tag{7}$$

where $e(n)$—prediction error and ρ—forgetting factor ($0 \ll \rho \leq 1$).

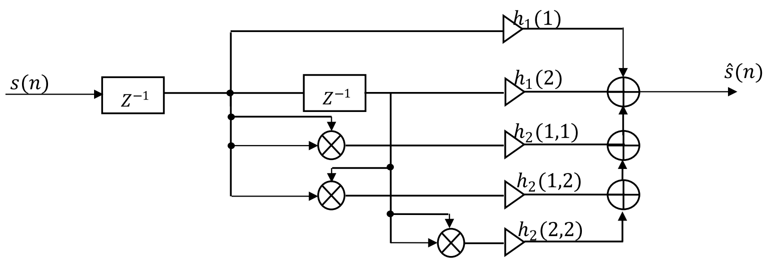


Fig. 3 Second-order Volterra filter model

3.2 *Long-Term Speech Prediction Using Volterra Model*

Only the correlation between nearby samples is eliminated by short-term linear prediction. To model voice signal adequately, the prediction order must be high enough to include at least one pitch period. Due to large delay and increased complexity, this is not acceptable in most of the practical applications. Hence, it is really important to design long-term predictor (LTP) that is capable of removing far-sample redundancies due to the presence of a pitch excitation. The standard solution to this problem in linear prediction is the use of a model with two linear predictors, short-term and long-term, connected in cascade as shown in Fig. 4. This is used in many of the linear prediction speech coders [2].

Second-Order Long-Term Volterra Prediction To perform the nonlinear long-term prediction of speech signal, a short-term linear predictor and a long-term second-order Volterra predictor are connected in cascade as shown in Fig. 5. This filter model predicts the current speech sample from a past sample which is one or more pitch periods away.

The predicted sample from the above filter model is

$$\hat{e}_s(n) = h_1 \,.\, e_s(n-T) + h_2 \,.\, e_s^2(n-T) \tag{8}$$

where T—pitch period and h_1 and h_2—long-term prediction coefficients. For a given T, the coefficients are computed by minimizing the sum of squared error.

Sum squared error is given as,

$$J = \sum_n \left(e_s(n) - \hat{e}_s(n)\right)^2 \tag{9}$$

The long-term prediction coefficients obtained are,

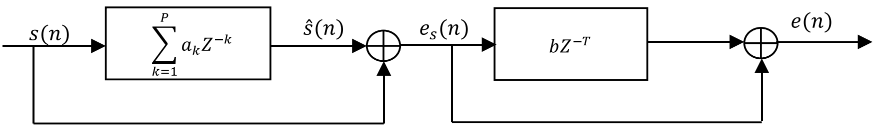


Fig. 4 Short-term and long-term linear predictors in cascade

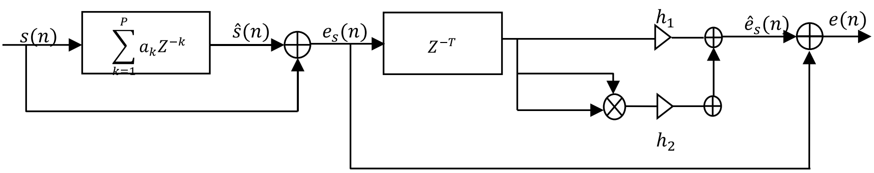


Fig. 5 Short-term predictor and second-order long-term Volterra predictors in cascade

$$h_1 = \frac{\sum_n e_s^4(n-T) \sum_n e_s(n) \,.\, e_s(n-T) - \sum_n e_s^3(n-T) \sum_n e_s(n) \,.\, e_s^2(n-T)}{\sum_n e_s^4(n-T) \sum_n e_s^2(n-T) - \left(\sum_n e_s^3(n-T)\right)^2} \tag{10}$$

$$h_2 = \frac{\sum_n e_s^2(n-T) \,.\, \sum_n e_s(n) \,.\, e_s^2(n-T) - \sum_n e_s^3(n-T) \,.\, \sum_n e_s(n) \,.\, e_s(n-T)}{\sum_n e_s^4(n-T) \sum_n e_s^2(n-T) - \left(\sum_n e_s^3(n-T)\right)^2} \tag{11}$$

4 Code-Excited Nonlinear Prediction Speech Coder (CENLP)

The CENLP–4.8 coder uses the analysis-by-synthesis method of encoding the signal, i.e. the encoder synthesizes the signal (without any channel errors). The CENLP encoder and decoder flow diagrams are shown in Fig. 6.

Short-term linear prediction: Estimation of the formant structure (the period of the formants is less than 1 ms) and the nonlinear prediction coefficients of different orders like 2, 4, 10. These nonlinear coefficients are used to find the error signal by the adaptive technique, which enhance the quality of speech signal.

Long-term linear prediction: Estimation of the pitch and intensity of the speech signal (the period of these parameters is in the order of 3–15 ms). The degree of nonlinearity is maximum up to 3, since the complexity in structure makes it computationally difficult to find the coefficients. Here the long-term predictions up to orders 2 and 3 are done to found the coefficients, and thus, the error signal is determined and the signal is reconstructed for the quality checking purpose.

The nonlinear prediction coefficients (NLPCs) are given as input to the prediction-error filter which performs inverse filtering, i.e. the formant information provided by the NLPCs is removed from the speech sub-frame signals, giving the resultant signal from which long-term parameters like pitch and intensity can be estimated (by estimating the frequency and amplitude of the signal). The pitch and intensity are the long-term parameters that are sent for encoding to be transmitted. The input frame is also given as input to a perceptually weighted filter. The filter weights are determined based on a perceptual compression algorithm. The short-term prediction output (i.e. NLPCs) and the long-term prediction output (pitch and intensity) are given as input to formant synthesis filter and pitch synthesis filter, respectively.

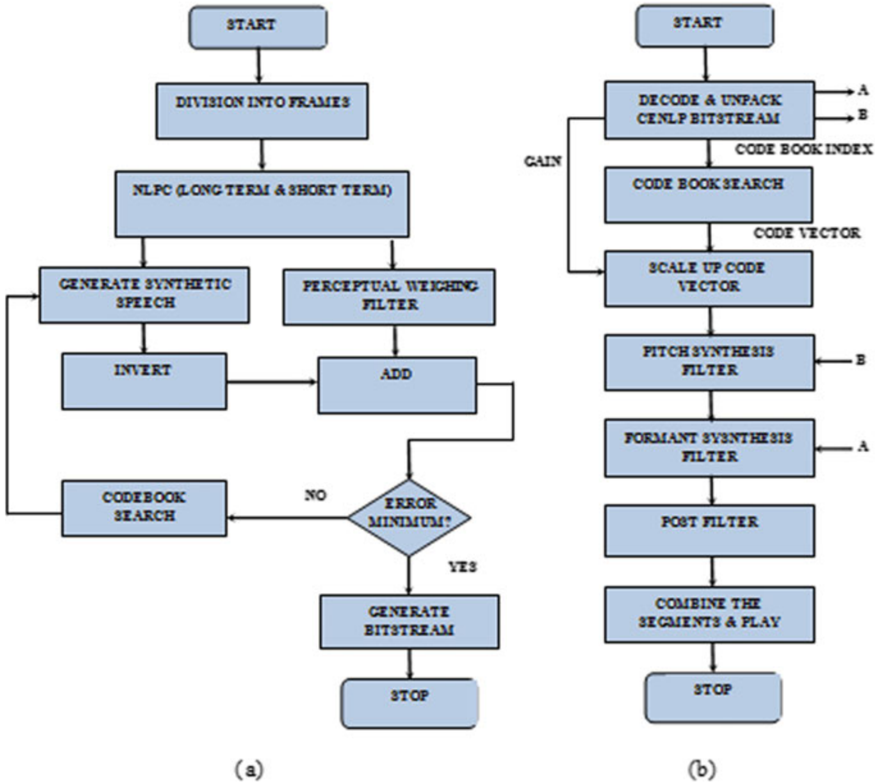


Fig. 6 **a** Design flow of CENLP Encoder **b** Design flow of CENLP Decoder

4.1 Implementation Details of CENLP Coder with Bitrate 4.8 Kbps

CENLP codec is implemented in MATLAB for a bit rate 4.8 Kbps. Speech sampled at 8 kHz is given to the encoder. The speech is divided into frames of 240 samples each (30 ms). The number of frames decides the number of times the encoder must run. Then it is subdivided into sub-frames of 60 samples each (75 ms).

The frames are given to the second-order Volterra short-term nonlinear predictor to compute the nonlinear prediction coefficients. The filer used to have the linear prediction order $P = 4$ and nonlinear prediction order $M = 3$. The output of this block is an array of coefficients 1, $h_1(1)$, $h_1(2)$, $h_1(3)$, $h_1(4)$, $h_2(1{,}1)$, $h_2(1{,}2)$, $h_2(1{,}3)$, $h_2(2{,}2)$, $h_2(2{,}3)$ and $h_2(3{,}3)$. These coefficients are nonlinear prediction coefficients (NLPCs).

The predicted speech sample is given by the equation,

Table 2 Variation in prediction gain

Type	Prediction gain in dB
Short-term LPC ($P = 10$)	12.5
Short-term LPC and long-term linear prediction	14.8
Short-term LPC and Volterra long-term prediction	17.6

$$\hat{s}(n) = 1 + \sum_{k=1}^{4} h_1(k)s(n-k) + \sum_{i=1}^{3}\sum_{j=1}^{3} h_2(i,j)s(n-i)s(n-j) \tag{12}$$

To find the periodicity of voiced speech frames, a long-term predictor combined with a short-term predictor is used. The long-term prediction is done at the sub-frame level. For this, a cascade of linear short-term predictor and a second-order Volterra long-term nonlinear predictor is used. Short-term prediction error appears as the input long-term predictor. Through the squared error minimization procedure, long-term prediction coefficients (LTPCs) are computed. Then, a search procedure is done within a range $T_{\min} \le T \le T_{\max}$ to determine the pitch period (T) [6]. Typically, $T_{\min} = 20$ ms and $T_{\max} = 140$ ms.

Sub-frame level nonlinear predictions are done under different conditions, and variation in prediction gain is listed in Table 2.

Parameters considered for the encoding of CENLP 4.8 codec are as follows.

Number of samples/s	8000
Frame duration	60 ms
Frame length	240 samples
Number of nonlinear prediction coefficients (NLPCs)	10/frame ($P = 4$ and $M = 3$)
Pitch filter coefficients (LTPCs)	2/sub-frame
Number of gain values	4/frame
Pitch filter delays	4/frame
Codebook size	512 code vectors (index length–9).

These parameters are encoded as per the following bit allocation strategy.

NLPC coefficients—$6 \times 10 = 60$, gain—$2 \times 4 = 8$, Pitch filter coefficients—$3 \times 8 = 24$, delay—$4 \times 4 = 16$ and codebook index $= 4 \times 9 = 36$.

5 Experimental Results

Both the nonlinear and linear predictions were implemented in the G723.1 CELP codec, and the results were observed. The reconstructed signals had very good quality

Table 3 MOS score CELP and CENLP

Language	File type	MOS score	
		CELP	CENLP
Malayalam	Mal_female.wav	3.68	3.56
	Mal_male.wav	3.45	3.85
	Mal_Conv.wav	2.85	3.55
English	Eng_female.wav	4.34	4.12
	Eng_male.wav	4.78	4.55
	Eng_Conv.wav	3.80	3.98
Hindi	Hin_female.wav	4.12	4.35
	Hin_male.wav	3.45	4.26
	Hin_Conv.wav	3.64	3.75
Tamil	Tam_female.wav	3.45	3.75
	Tam_male.wav	4.10	4.00
	Tam_Conv.wav	3.89	3.95
Kids	Hin_female.wav	3.40	3.89
	Hin_male.wav	2.98	3.52
Average		3.694	3.92

compared to the original speech signal. The subjective testing and objective testing were performed on the reconstructed speech signals.

5.1 Subjective Testing

Subjective evaluation of in CELP and CENLP codecs was done by conducting informal listening test (MOS testing) of reconstructed speech [11]. It was conducted using the speech.wav files of the database created for Indian dialects. The results of the MOS testing of coder outputs for several speech signals are included in Table 3. The MOS score obtained from reconstructed speech files for CENLP codec is 3.92, which approaches toll quality.

5.2 Objective Testing

Objective distortion measures of codecs were computed from the frequency and time characteristics of original and reconstructed signals. Objective parameters used for the analysis are segmental signal-to-noise ratio (SNRseg), log likelihood ratio (LLR), weighted spectral slope (WSS), and perceptual evaluation speech quality (PESQ). Values of these parameters are computed using MATLAB routines [7, 11].

Values of objective quality measures of CELP and CENLP codecs for different Indian dialects are listed in Table 4. Figure 7 shows the change in average values of quality measures among the above-mentioned codecs. Experimental results show that PESQ and SNRseg values are very good and high for CENLP. These measures give the enhancement in the quality of reconstructed signal with the use of proposed code-excited nonlinear prediction speech coder. LLR and MSE show low score for the proposed system. The agreement between the spectral magnitudes of original and decoded speech is indicated by LLR and its low value corresponds to close agreement.

6 Conclusion

In this modern era, the efficient transmission of audio and video signals at low bitrate has a greater requirement. CELP 4.8 is a good quality linear prediction speech coder with moderate bitrate and delay. In this work, a new speech codec is proposed by implementing nonlinear prediction in CELP coder to improve the quality by maintaining the bitrate and low delay. The proposed codec CENLP is validated through the subjective and objective quality analysis procedures. Processing delay is also measured for both CELP and CENLP algorithms. Both the algorithms have almost the same delay because of equal number of coefficients. Experimental results clearly say that, in CENLP there is an increase in the prediction gain and thereby an increase in quality that approaches toll quality. Further improvement in the speech quality can be achieved by doing codebook optimization in CENLP.

References

1. Makhoul J (1975) Linear prediction: a tutorial review. Proc IEEE 63:561–580
2. Schroeder M, Atal B (1985) Code-excited linear prediction (CELP): high-quality speech at very low bit rates. In: IEEE international conference on acoustics, speech, and signal processing, vol 10, pp 937–940
3. Chen J-H, Cox RV, Lin YC, Jayant N, Melchner MJ (1992) A low-delay CELP coder for the CCITT 16 Kb/s speech coding standard. IEEE J Sel Areas Commun 10(5):830–849
4. McCree A, Truong K, George EB, Barnwell TP, Viswanathan V (1996) A 2.4 Kb/s MELP coder candidate for U. S. Federal Standard. In: IEEE international conference on acoustics, speech and signal processing, vol 1, pp 200–203
5. ITU-I Rec. G.729 (1996) Coding of speech at 8 kbps using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)
6. Despotović V, Perić Z (2013) Design of nonlinear predictors for adaptive predictive coding of speech signals. In: 2013 21st Telecommunications forum (TELFOR), Serbia, Belgrade, 26–28 Nov 2013
7. Despotovic V, Goertz N, Peric Z (2012) Nonlinear long-term prediction of speech based on truncated Volterra series. IEEE Trans Audio Speech Lang Process 20(3):1069–1073
8. Chu WC (2003) Speech coding algorithms: foundation and evolution of standardized coders. Wiley

Table 4 Objective quality measures of CELP and CENLP codecs

Language	Objective measures											
	CELP						CENLP					
	LLR	SNRseg	WSS	PSNR	PESQ	MSE	LLR	SNRseg	WSS	PSNR	PESQ	MSE
Malayalam	0.65	1.11	44.75	125.67	2.27	0.028	0.58	1.32	60.43	132.45	3.25	0.013
English	0.67	1.04	62.83	117.87	2.42	0.044	0.58	1.01	80.67	125.32	3.12	0.0.22
Tamil	0.55	1.21	68.34	132.55	2.45	0.034	0.57	1.49	89.32	139.27	2.89	0.012
Hindi	0.69	1.15	61.12	131.41	2.41	0.051	0.61	1.43	68.77	139.23	3.21	0.041
Kids	0.53	1.23	60.32	127.45	2.97	0.052	0.44	1.57	76.54	131.66	3.31	0.040
Average	0.62	1.15	59.47	126.99	2.50	0.042	0.56	1.36	75.15	133.59	3.16	0.026

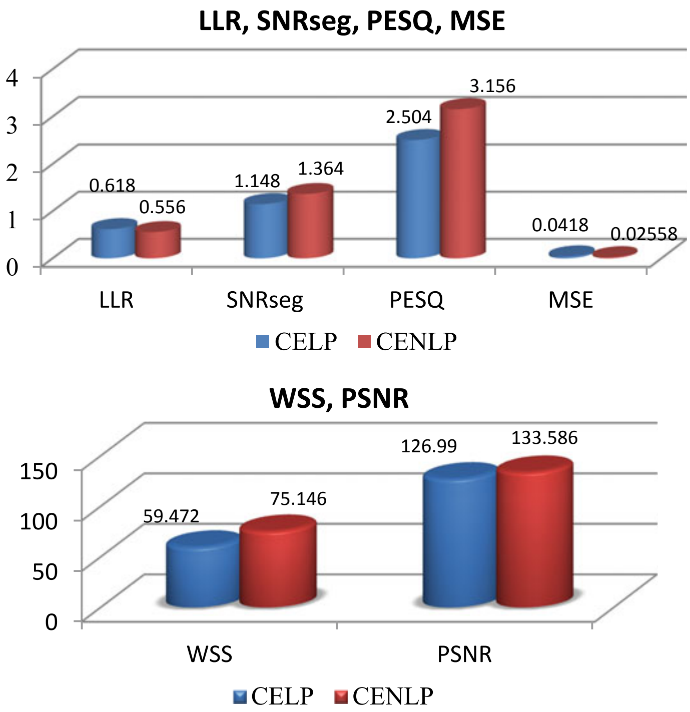


Fig. 7 Average objective quality measures of CELP and CENLP: LLR, SNRseg, PESQ, MSE, WSS and PSNR

9. Despotovic V, Goertz N, Peric Z (2012) Low-order Volterra longterm predictors. In: Proceedings 10th ITG symposium on speech communication, Braunschweig, Germany, Sept 2012, pp 26–28
10. Despotović V, Görtz N, Perić Z (2012) Improved non-linear long-term predictors based on Volterra filters. Vienna University of Technology, Institute of Telecommunications, Gußhausstr. 25–29, 1040 Vienna
11. Anselam AS, Pillai SS (2017) Optimization of code excited linear prediction speech coder with PSVQ-genetic codebook. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET)

Comparative Analysis of FCFS and SJF for Multimedia Process Scheduling

R. Magdalene and D. Sridharan

Abstract Multimedia data processing has been gaining popularity due to its increasing demand in today's scenario. In this paper, we discuss the various scheduling strategies that have been employed for multimedia processes. We study the performance of traditional scheduling algorithms such as First Come First Servce (FCFS) and Shortest Job First (SJF) for multimedia tasks and analyse the variation in their performance under different scenarios such as different sets of data and by varying the number of processes. Experimental results prove that the scheduling schemes vary in their performances under different situations.

Keywords Multimedia process · Scheduling · Operating systems · CPU scheduling

1 Introduction

Multimedia data processing involves handling multiple media files such as audio, video, text, and graphics—all at the same time. When there are a multiple number of processes, only one process can be executed at a time and the other processes need to wait in the queue to be executed by the CPU. To overcome this constraint, multiprogramming is introduced. CPU scheduling [1–6] helps in managing which ready process in the queue should be assigned to the CPU. There are a wide number of scheduling algorithms in existence to schedule the processes such first come first serve (FCFS) [5], shortest job first (SJF) [5–7], round robin scheduling [5, 8] and priority-based scheduling algorithms [7, 9–11] such as earliest deadline first (EDF) [8, 12, 13], and least slack time first (LSF) [14]. In the traditional FCFS algorithm, the process that arrives first in the queue is executed first. In other words, the scheduling is performed as per the arrival time of each process. In SJF algorithm, the shortest

R. Magdalene (✉) · D. Sridharan
Anna University, Chennai, India
e-mail: magdalene.rc@gmail.com

D. Sridharan
e-mail: sridhar@annauniv.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_54

job is executed first pushing the longest job to be executed last. The priority-based algorithms EDF sorts the job as per the deadline of each task and thus the task that has the latest deadline is executed first and in LSF, the one that has the least slack time is executed first.

The rest of the paper is organized as follows: The second section explains multimedia task scheduling while the third section discusses the research work carried out on multimedia processes. The experimental results of the FCFS and SJF applied to multimedia processes are shown in Sect. 3.1. Section 4 gives a summary and conclusion to the paper followed by the references.

2 Multimedia Task Scheduling

Multimedia specific tasks are quite challenging to work with as shown by most of the recent research works [1, 15–19]. The increase in multimedia data is a major reason for the emergence of bigdata. Recent techniques such as cloud-based [15, 18–20] scheduling and even genetic algorithms [4, 21, 22] have been employed in order to find an optimized scheduling scheme.

2.1 Multimedia Task

Multimedia is a content that comprises of elements such as text, audio, video, graphics, and so on. A multimedia process can thus be defined as the process that deals in combination with all the elements mentioned above. As explained in Fig. 1, the figure explains the basic process flow of multimedia task scheduling.

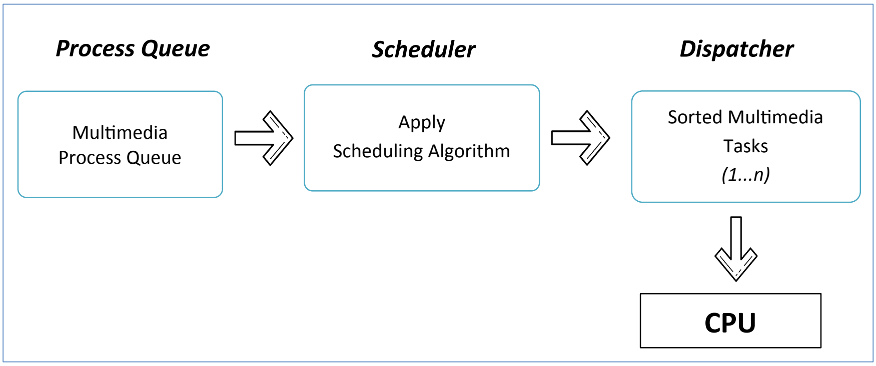


Fig. 1 Multimedia task scheduling flow

2.2 Scheduling Criteria

To determine the performance of various scheduling algorithms, we use different types of assessment factors. These factors should either be minimized or maximized accordingly. The following are some of the common scheduling criteria, (1) maximizing the CPU utilization [1] (2) fair allocation of the CPU [13] (3) Maximum throughput [23] (4) Minimum waiting time [4] (5) Minimum turnaround time [4, 7] (6) Minimum response time (7) Minimizing the starvation of CPU [3] (8) Minimizing the missing deadlines [22] (9) Balanced utilization (10) Priority [9, 10, 24] Every scheduling algorithm outperforms the others in one or more scheduling criteria mentioned above. However, determining an ideal scheduling scheme depends on the type of application and the type of processes that are handled by the computing systems. Hence, for multimedia type of applications [4, 12, 18, 23, 25–28], the scheduling algorithm must satisfy the requirements of multimedia-related demands.

2.3 Basic Definitions [29]

- **Completion time—**The time taken for the execution to be over from the time the process arrived at the queue
- **Burst time**—The total execution time of each process is referred to as burst time
- **Turnaround time—**The time taken for a process from the time of submission till it is executed, thus turnaround time is calculated using Eq. (1)

$$\lambda^t = \sum_{i=1}^{n} \tau^i - \sigma^i \tag{1}$$

where

τ^i is the completion time of the process
σ^i is the arrival time of the process

- **Waiting time**—The time each process had to wait in the queue before the execution is the waiting time. Mathematically, this is calculated by subtracting the burst time from the turnaround time, thus waiting time is given by the Formula (2)

$$\omega^t = \sum_{i=1}^{n} \lambda^i - \beta^i \tag{2}$$

where

λ^i is the turnaround time of the process
β^i is the burst time of the process

- **Slack Time—**The time left after the task has been completed if the task is processed now.
- **Response time**—The time taken for the process to receive its first response after the request is submitted.
- **Load Average**—The load is referred to the number of the processes that are ready in the queue waiting to be executed; the load average is calculated by finding the average number of processes that are usually there in the ready queue.
- **Throughput**—The number of processes executed in each period. This can range from one process per 10 s or one process per hour depending on the specific processes

$$T = n/t \tag{3}$$

where

n is the number of process
t is defined as per unit time

- **CPU Utilization**—In an ideal case, we expect the CPU to be utilized 100% but in a real-time scenario, the CPU utilization ranges between 40 and 60%. The CPU utilization rate can be calculated by using the Formula (4)

$$\delta^r = \frac{\zeta^t}{\zeta^t + I^t} \tag{4}$$

where

ζ^t is the busy time of the CPU
I^t is the idle time of the CPU.

3 Comparative Study of FCFS and SJF for Multimedia Task Scheduling

3.1 Experimental Results and Discussion

For this experiment, we have taken different synthetic datasets for obtaining test results to analyze the performance of the algorithms in different instances. The figurative representation of the results is shown in Figs. 2 and 3. For every multimedia process $\mathrm{M}_{\mathrm{P}}^{\mathrm{i}}$, the arrival time, A^{T}, and burst time, B^{T} are dynamic, for example, in every execution, the arrival time and burst time are randomly generated and the algorithm is evaluated for the performance at different instances.

The figurative results explained in Figs. 2 and 3 are the performance analysis of the FCFS and SJF algorithms for multimedia process scheduling. From the observations

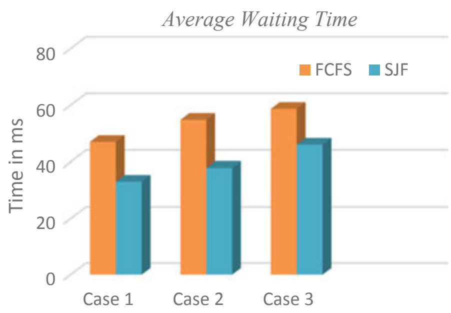


Fig. 2 Comparison of average waiting time for FCFS and SJF

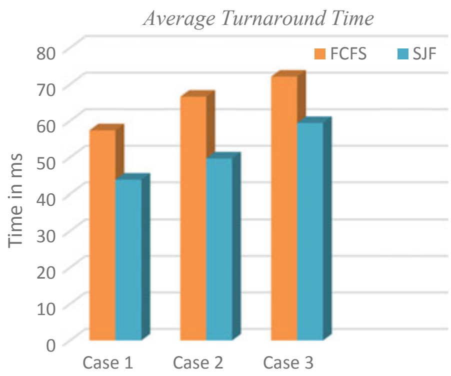


Fig. 3 Comparison of average turnaround time for FCFS and SJF

Table 1 Case 1

M_P id	A_T	B_T
1	7	5
2	8	14
3	6	4
4	6	7
5	3	7
6	4	19
7	3	16
8	6	13
9	9	6
10	7	13

performed through the simulation study, we learn that in case 1, case 2, and case 3 as per the dataset listed in Tables 1, 2 and 3 containing different sets of multimedia processes, SJF performs better when compared to FCFS at all the instances.

Table 2 Case 2

M_P id	A_T	B_T
1	6	3
2	310	4
3	1	18
4	0	18
5	0	15
6	8	7
7	5	2
8	2	1
9	7	15
10	6	17

Table 3 Case 3

M_P id	A_T	B_T
1	3	17
2	2	17
3	9	13
4	0	16
5	7	11
6	3	3
7	1	12
8	5	8
9	9	17
10	2	20

4 Conclusion

A comparative study of FCFS and SJF for multimedia process scheduling has been carried out. The results show that for multimedia tasks scheduling, the SJF performs better compared to FCFS while applying the algorithms to three different datasets. A set of random arrival time and burst time is generated and the performances of each algorithm under different datasets were computed. With many traditional data scheduling techniques still in existence, experimentations are being carried out to find out the algorithm that works best for multimedia data.

References

1. Hu W, Ma T, Wang Y, Xu F, Reiss J (2018) TDCS: a new scheduling framework for real-time multimedia OS. Int J Parallel Emergent Distrib Syst 1–16

2. Yau DKY, Lam SS (1997) Adaptive rate-controlled scheduling for multimedia applications. In: Proceedings of the fourth ACM international conference on multimedia, vol 5, no 4, pp 129–140
3. Muhammad A, Abdullahi SE, Junaidu SB (2015) Optimization of priority based CPU scheduling algorithms to minimize starvation of processes using an efficiency factor. Int J Comput Appl 132(11):24–32
4. Neshat M, Sargolzaei M, Najaran A, Adeli A (2012) The new method of adaptive CPU scheduling using fonseca and fleming's genetic algorithm. J Theor Appl Inf Technol 37(1):1–16
5. Siahaan APU (2016) Comparison analysis of CPU scheduling: FCFS, SJF and round robin. Int J Eng Dev Res 4(3):124–131
6. Baiceanu V, Cowan C, McNamee D, Pu C, Walpole J (1996) Multimedia applications require adaptive CPU scheduling. IEEE RTSS Work Resour Alloc Prob Multimed Syst. 1–6
7. Khan MR, Kakhani G (2015) Analysis of priority scheduling algorithm on the basis of FCFS & SJF for similar priority jobs. Int J Comput Sci Mob Comput 4(9):324–331
8. Saraswat PK, Gupta P (2006) Design and implementation of a process scheduler simulator and an improved process scheduling algorithm for Multimedia Operating Systems. In: Proceedings—2006 14th international conference on Advanced Computer and Communications (ADCOM 2006), pp 513–517
9. Dehbi Y, Mikou N (2008) Priority assignment for multimedia packets scheduling in MANET. In: SITIS 2008—Proceedings of 4th international conference on signal ımage technology and ınternet based systems, pp 32–37
10. Chen H, Xia J (2009) A real-time task scheduling algorithm based on dynamic priority. In: Proceedings—2009 international conference on embedded software and systems (ICESS 2009), no 1, pp 431–436
11. Zhenxia Y, Fang M, Sheng S (2008) Scheduling algorithm based on task priority in heterogeneous computing environment. In: Proceeding international conference on computer science and ınformation technology (ICCSIT 2008), pp 12–16
12. Lim TM, Lee BS, Yeo CK (2007) Quantum-based earliest deadline first scheduling for multiservices. IEEE Trans Multimed 9(1):157–168
13. Steinmetz R (1995) Multimedia file systems survey: approaches for continuous media disk scheduling. Comput Commun 18(3):133–144
14. Hwang M, Choi D, Kim P (2011) Least slack time rate first: an efficient scheduling algorithm for pervasive computing environment. J Univers Comput Sci 17(6):912–925
15. Liu Y, Li C, Luo Y, Shao Y, Zhang J (2018) Scheduling multimedia services in cloud computing environment. Enterp Inf Syst 12(2):218–235
16. Xu J, Ota K, Dong M (2018) Real-time awareness scheduling for multimedia big data oriented in-memory computing. IEEE Internet Things J. 5(5):3464–3473
17. Bok K, Hwang J, Lim J, Kim Y, Yoo J (2017) An efficient mapreduce scheduling scheme for processing large multimedia data. Multimed Tools Appl 76(16):17273–17296
18. Alonso Á, Aguado I, Salvachúa J, Rodríguez P (2017) A methodology for designing and evaluating cloud scheduling strategies in distributed videoconferencing systems. IEEE Trans Multimed 19(10):2282–2292
19. Meriam E, Tabbane N (2017) A survey on cloud computing scheduling algorithms. In: Proceedings—2016 global summit on computer and information technology GSCIT 2016, pp 42–47
20. Singh S, Negi S, Verma SK, Panar N (2018) Comparative study of existing data scheduling approaches and role of cloud in VANET environment. Procedia Comput Sci 125:925–934
21. Li X, Gao L, Shao X, Zhang C, Wang C (2010) Mathematical modeling and evolutionary algorithm-based approach for integrated process planning and scheduling. Comput Oper Res 37(4):656–667
22. Yoo M, Gen M (2006) Multimedia task scheduling using proportion-based genetic algorithm. IEEJ Trans Electron Inf Syst 126(3):347–352
23. Xing M, He J, Cai L (2016) Maximum-utility scheduling for multimedia transmission in drive-thru internet. IEEE Trans Veh Technol 65(4):2649–2658

24. Li C, Zhu L, Liu Y, Luo Y (2017) Resource scheduling approach for multimedia cloud content management. J Supercomput 73(12):5150–5172
25. Kweon H, Do Y, Lee J, Ahn B (2007) An efficient power-aware scheduling algorithm in real time system. In: IEEE pacific rim conference on communications, computers and signal processing—proceedings, pp 350–353
26. Steinmetz R (1995) Analyzing the multimedia operating system. IEEE Multimed 2(1):68–84
27. Taboada I, Liberal F, Fajardo JO, Ayesta U (2013) QoE-aware optimization of multimedia flow scheduling. Comput Commun 36(15–16):1629–1638
28. Thinagar AS, Jenkins L (2003) Integrated scheduling of hard real-time and multimedia tasks. In: IEEE Reg 10 Annual international conference proceedings/TENCON, 4:1618–1622
29. Silberschatz A, Galvin PB, Gagne G. Operating system concepts, 9th edn. Wiley

Face Recognition Using Improved Co-HOG Features

C. H. Hima Bindu and K. Manjunatha Chari

Abstract Face recognition is one of the most sought-after biometric technologies in the field of machine learning and computer vision in recent years. Histogram of oriented gradients (HOG) descriptor was originally developed for human detection and recently, it is also being applied to face recognition. However, when compared with other successful feature descriptors such as SIFT, LBP, Gabor, and so on, there is still considerable research space on the application of HOG features for face recognition. Co-HOG, a variant of HOG, uses a pair of gradient orientations as its basic building block, unlike HOG which uses a single gradient. Using the pair of orientations, the co-occurrence matrix is computed and histograms are calculated. However, in Co-HOG, gradient direction alone is considered and magnitude is ignored. It is believed that gradient magnitude also carries significant information about features. In this paper, we develop a face recognition system that utilizes Co-HOG features with embedded gradient magnitude information. The experimentation is done on ORL face dataset and it is observed that the proposed model is better than other existing methods with a maximum accuracy value of 97%.

Keywords Gradient orientation · HOG · Co-HOG · Face recognition

1 Introduction

Nowadays, biometrics is replacing a majority of conventional identification systems in various security-related applications. Face recognition, one of the biometric technologies, is being widely used for this purpose and considered to be the fastest-growing technology. There are a plethora of face recognition techniques existing in the computer vision and pattern recognition literature. But the idea of developing

C. H. Hima Bindu (✉)
GITAM (Deemed to be University), Rudraram, Sangareddy, Telangana, India
e-mail: himabindu.chelluri@gmail.com

K. Manjunatha Chari
ECE Department, GITAM (Deemed to be University), Rudraram, Sangareddy, Telangana, India
e-mail: hodece_hydcampus@gitam.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_55

a robust face recognition system that addresses various challenging scenarios like illumination variation, occlusion viewpoint variation, etc., is still an open problem. The face recognition method, broadly, consists of two steps: (i) feature extraction and (ii) classification. Feature extraction is a process of effectively describing a large set of data accurately using a minimum amount of information. In the pattern recognition literature, various feature extraction methods have been widely used. Besides holistic methods such as PCA and LDA, local descriptor methods are also gaining immense popularity recently. An ideal descriptor should have inter-class variance in maximum and intra-class variance in minimum indicating the requirement of the descriptor to be robust to varying illumination, slight deformations, image quality degradation, and so on.

To enhance the overall accuracy of the face recognition system, one needs to improve the quality of two aspects of the system. They are features and classifier. The basic features of an image can be extracted using two mathematical approaches. One is differential geometry and the other one is scale-space theory [1], where the former approach is based on local anisotropic variations of pixel intensities and the latter is based on multi scale representation of an image. Among the various differential feature extraction methods, gradient-based methods are being used widely as they effectively describe important features of an image. A few examples of gradient-based edge features are edge orientation histogram [2], histograms of oriented gradients (HOG) [3], co-occurrence HOG (Co-HOG) [4], shapelets [5], etc.

HOG allocates the gradient magnitudes of a portion of image patch into different orientation bins to form a feature and is successfully applied in many research fields specifically in pedestrian detection. Co-HOG, proposed by Watanabe et al., is an extension of HOG that embodies additional information about the spatial relationship between gradient orientations. Both HOG and Co-HOG show good performance in various range of object detection/recognition applications. However, in Co-HOG, the gradient information is not completely used, i.e., gradient direction alone is considered and magnitude is ignored in the process of framing the descriptor. It is believed that gradient magnitude also carries an optimum amount of information that can contribute while describing features. To address this shortcoming, Pang et al., proposed an improved version of Co-HOG by using the complete gradient information that considers magnitude in addition to gradient orientation in their work that deals with human detection in a video/image. In this paper, we intend to use Co-HOG features into a face recognition system that also takes into account magnitude details of the gradient information and hence improves the recognition rate [6].

2 Related Work

Recognizing a face, particularly, becomes challenging under uncontrolled environments due to the presence of big distinctions in illumination, pose, accessories and changes due to aging, etc. [7]. In this scenario, a powerful feature extraction method always plays a major factor to tackle these issues. Hence, feature extraction is an important issue in pattern recognition and computer vision.

HOG is a histogram-based descriptor that works on the gradient orientation of pixels of an image. It is primarily used for object detection/recognition and was applied firstly for pedestrian detection. It captures face contour information and also has the advantage of being insensitive to light variations and small offset. It is closely related to the SIFT technique. However, SIFT is intended to run at specific points called interest points whereas HOG runs over a dense grid. Firstly, using vertical and horizontal filters, the gradient of the image is computed. Then, the gradient image is divided into multiple cells *C,* each of size $p \times q$ pixels. For each cell, a histogram of b bins is computed and then, all the histograms of all cells are concatenated to form a feature vector of size $C \times b$. HOG is very simple and effectively describes local features. However, one area of concern about HOG is that it overlooks information pertaining to spatial and structure. Researchers have proposed different methods to improve HOG. Albiol et al. used HOG features in combination with EBGM for face recognition where EBGM is used to localize facial landmarks and HOG is used to calculate descriptor of these landmarks [7]. In [8], in order to overcome errors due to occlusion, pose, illumination changes, Deniz et al., extracted HOG features using a regular grid. Also, the descriptors are extracted at different scales of the image to improve accuracy. Yi et al. [9] proposed global sampling (GHOG) for character recognition, yet another improved variant of the HOG which focuses on better modeling of character structures and hence better performance.

3 Background of Co-HOG

Co-HOG is another derivative of HOG that has evolved with an objective to improve the feature extraction. This technique is widely used in object detection and recognition due to its high accuracy and less false-positive rates. In Co-HOG, a pair of oriented gradients form the fundamental element, unlike HOG in which the fundamental element is, a single gradient orientation. A co-occurrence matrix is formed by combining a pair of neighboring gradient orientations at a given offset in the image and the pairing of orientations is done over a small region of the image [4]. This combination of gradient orientations at an offset (Fig. 1) offers more accurate features of the image and hence contributes to better image classification. The elements of the co-occurrence matrix for a given image of gradient orientations of size $n \times m$ are formulated as given in Eq. (1):

$$M_{\Delta x,\Delta y}(i,j) = \sum_{x=0}^{n-1}\sum_{y=0}^{m-1} \begin{cases} 1 \text{ if } I(x,y) = i \text{ and } I(x+\Delta x, y+\Delta y) = j \\ 0 \text{ otherwise} \end{cases} \quad (1)$$

where I is the GO image, i and j are GOs, and $\Delta x, \Delta y$ are offsets in vertical and horizontal directions. The equation for orientations is given as:

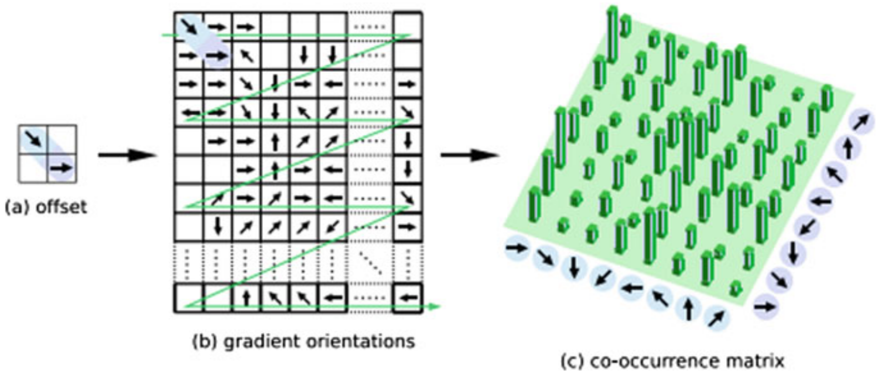


Fig. 1 Co-occurrence matrix of gradient orientations [4]

$$\theta = \tan^{-1}\frac{v}{h} \tag{2}$$

where v and h are gradients in vertical and horizontal directions, respectively. The orientations are divided into eight divisions in the range of $(0, 2\pi)$ at an interval of $\frac{\pi}{4}$. Thereby leading to a co-occurrence matrix of size $8 \times 8 = 64$ elements. The total Co-HOG descriptor features amounts to $m \times n \times d^2$ where m is the number of tiled regions, n is the number of offsets, and d is the number of gradient orientation bins. Finally, the Co-HOG descriptor is determined by concatenating all co-occurrence matrices whose size is calculated as $2 \times 2 \times 6 \times 8^2 = 1536$.

Tian et al. [10] applied the co-occurrence of HOG (Co-HOG) features for scene text recognition to capture the spatial distribution of neighboring orientation pairs rather than only a single gradient orientation. The Co-HOG method improves HOG significantly in scene character recognition. Watanabe et al. used Co-HOG features successfully for pedestrian detection using gradient orientations [4]. However, in their work, the magnitudes of the gradient orientations of the features were seldom used and discarded.

4 Proposed Method

In this paper, we propose a face recognition system (Fig. 2) based on Co-HOG features. Initially, in the preprocessing stage, the face image is resized and converted into a gradient image using a Sobel filter.

The image is then split into non-overlapping blocks of size 3×6. For each block, Co-HOG features are computed and subsequently, features of all blocks are concatenated to obtain the final feature vector of the image as shown in Fig. 3. Co-HOG features are derived from HOG which is based on a pair of gradient orientations against HOG which is based on a single gradient. After obtaining features, an SVM

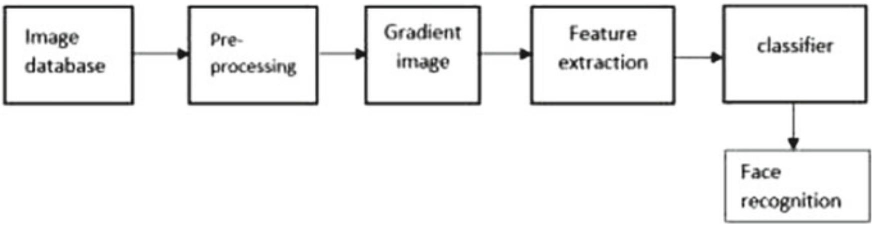


Fig. 2 Proposed approach for Co-HOG features

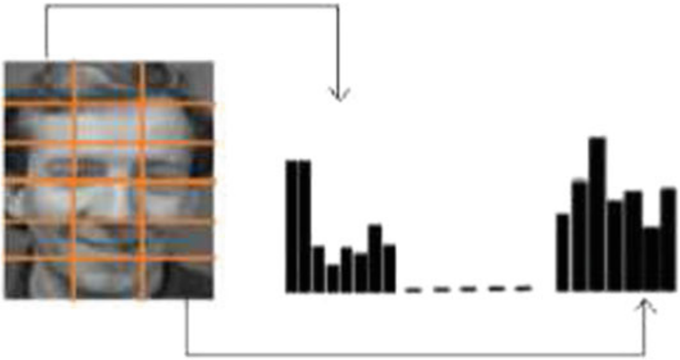


Fig. 3 Feature extraction representation

classifier [15] is used for classification purpose. In the proposed method of face recognition, we intend to use complete information of gradients, i.e., both magnitude and orientation by encoding magnitude information into Co-HOG features and thereby it is expected to extract better features and hence better feature description.

4.1 Feature Extraction of the Proposed Method

To encode the gradient magnitude into Co-HOG feature, firstly, the gradients are decomposed into eight gradient planes corresponding to eight orientations. For a given pixel (x, y), let the intensity be $f(x, y)$. Let the gradient at this pixel is given by $g(x, y) = (g(x, y), \theta(x, y))'$ where g is the magnitude and θ is the orientation. According to the force decomposition theory, magnitude g is decomposed into g_1 and g_2 between two angles θ_1 and θ_2 as:

$$g_1 = \frac{g_x \sin\theta_2 + g_y \cos\theta_2}{\cos\theta_1 \sin\theta_2 - \sin\theta_1 \cos\theta_2} \tag{3}$$

$$g_2 = \frac{g_x \sin\theta_1 + g_y \cos\theta_1}{\cos\theta_2 \sin\theta_1 - \sin\theta_2 \cos\theta_1} \tag{4}$$

where

$$g_x = g\cos\theta = g_1\cos\theta_1 + g_2\cos\theta_2 \tag{5}$$

$$g_y = g\sin\theta = g_1\sin\theta_1 + g_2\sin\theta_2 \tag{6}$$

where g_1 and g_2 are gradient magnitudes along θ_1 and θ_2 respectively.

In the next step, gradient smoothing is performed using a convolution mask to avoid the effect of unavoidable noise on the gradient. After decomposition and smoothing, gradient matrices $G_{\theta_i,\Delta x,\Delta y}$ and $G_{\theta_j,\Delta x,\Delta y}$ which are in the direction of θ_i and θ_j, respectively, are combined to form the co-occurrence gradient matrix given by, $G_{\theta_i,\theta_j,\Delta x,\Delta y}$. The co-occurrence gradient matrix is further smoothed and illumination normalized.

5 Experimental Results

We evaluate the proposed method on ORL datasets [14]. This face dataset has images of 40 distinct persons with ten different expressions per person with the size of each image being 92 × 112. The metrics used to evaluate the performance of the proposed face recognition system are false acceptance rate (FAR), false recognition rate (FRR), and accuracy. The metrics are then compared with the existing models, namely SIFT [11], KSIFT [12] and WHOG [13]. The performance of the proposed face recognition system is investigated for an offset value of 2. The classifier is trained with training face images for measuring the performance of the proposed model. For experimentation, the entire face database is separated into two: a training set and a testing set. The tests are conducted at five different instances of varying percentages of training data. At each instance, the remaining database forms the testing data. For example, at a given instance, if training face data is 60% of the total face database, then the remaining 40% forms testing face data. The training data values are 50, 60, 70, 80, and 90% of the face database. Figure 4 shows the experimentation results graphically for accuracy, FAR, and FRR and the same are tabulated in Table 1a–c. The accuracy of SIFT ranges from 90 to 97% as compared to WHOG whose accuracy ranges from 93 to 95%. KSIFT has a maximum recognition accuracy of 98.5%. On the other hand, Co-HOG also exhibits relatively good performance with a high accuracy of 98% for higher percentages of training data. The FRR of SIFT ranges from 0.07 to 0.12 and for WHOG, the values range from 0.06 to 0.08. Co-HOG has a low FRR of 0.03 at higher percentages of training data. Similarly, for FAR, Co-HOG exhibits equally good/better values when compared with other methods.

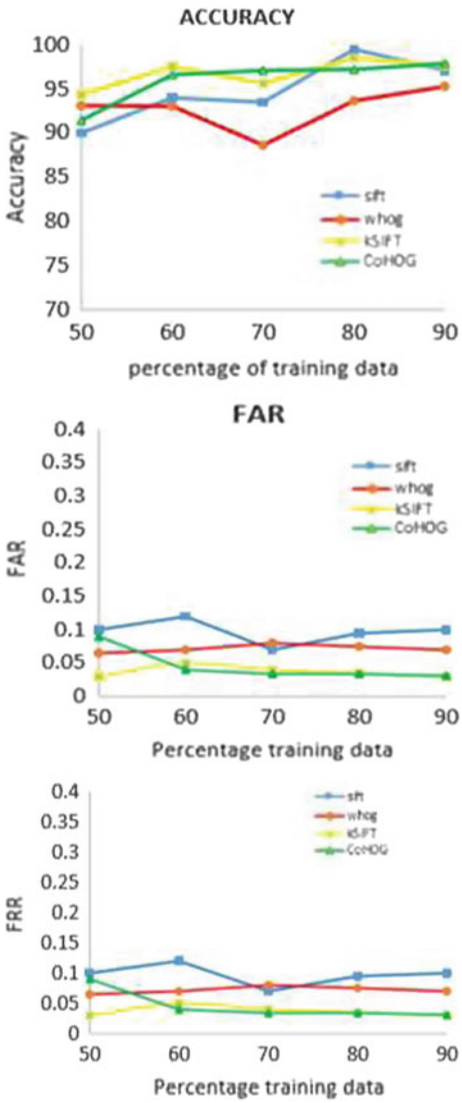


Fig. 4 **a**–**c** Experimental graphs for accuracy, FAR and FRR, respectively

6 Conclusions

In the proposed work, we applied Co-HOG technique for feature extraction in the face recognition problem. Co-HOG is a gradient-based method in which only gradient orientations were utilized discarding magnitudes of the gradients. Because of the discriminative nature of the magnitudes of the gradients during feature extraction, we intend to use them in the proposed approach of face recognition using Co-HOG

Table 1 **a** Accuracy, **b** false acceptance rate (FAR), **c** false recognition rate (FRR)

Training data (%)	WHOG	SIFT	Co-HOG	KSIFT
(a)				
50	93.11	90.02	91.46	94.35
60	93.03	94	96.61	97.58
70	88.64	98.51	97.12	95.62
80	93.65	99.5	97.23	98.53
90	95.28	97.65	97.89	97.9
(b)				
50	0.06	0.1	0.09	0.03
60	0.07	0.12	0.04	0.05
70	0.08	0.07	0.034	0.04
80	0.075	0.095	0.034	0.035
90	0.7	0.1	0.031	0.03
(c)				
50	0.06	0.1	0.09	0.03
60	0.07	0.12	0.04	0.05
70	0.08	0.07	0.034	0.04
80	0.075	0.095	0.034	0.035
90	0.7	0.1	0.031	0.03

features. The experimentation is done on ORL face dataset and it is observed that the proposed model is better than other methods with a maximum accuracy value of 97%.

References

1. Tizon X (2004) Algorithms for the analysis of 3D magnetic resonance angiography images. 316
2. Lopez-Molina C, Baets BD, Bustince H (2011) Generating fuzzy edge images from gradient magnitudes. Comput. Vis. Image Und. 115:1571–1580
3. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition (CVPR 2005), vol 1, pp. 886–893). IEEE
4. Watanabe T, Ito S, Yokoi K (2009) Co-occurrence histograms of oriented gradients for pedestrian detection. Lect Notes Comput Sci 5414:37–47
5. Sabzmeydani P, Mori G (2007) Detecting pedestrians by learning shapelet features. In: CVPR 2007, IEEE conference on computer vision and pattern recognition, pp 1–8
6. Pang Y, Yan H, Yuan Y, Wang K (2011) Robust CoHOG feature extraction in human-centered image/video management system. IEEE Trans Syst Man Cybern Part B (Cybern) 42(2): 458–468
7. Albiol A, Monzo D, Martin A, Sastre J, Albiol A (2008) Face recognition using HOG–EBGM. Pattern Recogn Lett 29(10):1537–1543

8. Déniz O, Bueno G, Salido J, De la Torre F (2011) Face recognition using histograms of oriented gradients. Pattern Recogn Lett 32(12):1598–1603
9. Yi C, Yang X, Tian Y (2013) Feature representations for scene text character recognition: a comparative study. In: 2013 12th International conference on document analysis and recognition, pp 907–911. IEEE
10. Tian S, Lu S, Su B, Tan CL (2013) Scene text recognition using co-occurrence of histogram of oriented gradients. In: 2013 12th International conference on document analysis and recognition (ICDAR), pp 912–916. IEEE
11. Wang H, Jiani Hu, Deng W (2017) Face feature extraction: a complete review. IEEE Access 6:6001–6039
12. Bindu H, Manjunathachari K (2018) Hybrid feature descriptor and probabilistic neuro-fuzzy system for face recognition. Sensor Review
13. Singh AK, Shukla VP, Tiwari S, Biradar SR (2015) Wavelet based histogram of oriented gradients feature descriptors for classification of partially occluded objects. Int J Intell Syst Appl 7(3):54
14. The database of faces (2002) AT&T Laboratories Cambridge, [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
15. Lin W-H, Wang P, Tsai C-F (2016) Face recognition using support vector model classifier for user authentication. Electron Commer Res Appl 18:71–82

Non-destructive Testing for Cracks in Concrete

Deepali Koppad and Nirmala Paramanandham

Abstract Non-destructive testing (NDT) is the process of analyzing the materials, components, structures, etc. without causing damage to it. In this paper, a NDT technique is proposed for detecting the cracks in the concrete surfaces. Initial results are obtained using a neural network for concrete crack detection. A convolution neural network model has been developed and trained using both positive (crack images) and negative (non-crack) images. In this work, a database consisting of 40,000 images is used. The model is trained with 36,000 images, 4000 for validation and 4000 for testing. To evaluate the effectiveness of model, accuracy, recall, precision and F1 score parameters are calculated.

Keywords Concrete crack detection · Non-destructive testing · Deep learning · CNN

1 Introduction

Over a period of time, given any material, such as concrete, iron, steel and plastic will develop a crack. Every material has applications in different areas. For example, concrete is mostly used in dams, bridges, swimming pools, homes, streets, plain cement tiles, lamp posts, drain covers and so on. Steel is used in buildings, ships, automobiles, etc. All these structures tend to develop cracks while being used either due to mishandling or over a period of time. It is important that these cracks are to be detected as soon as they are developed before they cause any severe damage.

Crack detection is the method of finding the crack in either at the surface or subsurface level using any of the processing techniques. Crack can be detected using destructive and non-destructive testing (NDT). Non-destructive testing (NDT) is a method of analyzing materials, components, etc. without causing damage to it. The

D. Koppad (✉) · N. Paramanandham (✉)
Ramaiah Institute of Technology, M S R Nagar, MSRIT Post, Bangalore 560054, India
e-mail: deepali.koppad@gmail.com

N. Paramanandham
e-mail: nirmalavp.ece@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_56

advantage of NDT is that after the testing is performed, the material or component can still be in use. NDT forms a significant part in the quality checking process of any manufacturing industry which aims to make high-quality products fulfilling the customer requirements. NDT techniques are progressively more used to ensure safety and reliability of components in operational units. For safety and reliability, NDT methods primarily aim at detecting and characterizing faults.

Crack detection of different materials is a major research area. This paper addresses the issue of crack detection in concrete material. Various techniques, either manual or automatic, have been developed over the years to detect cracks. With advances in technology, it is recommended to have an automatic crack detection mechanism without affecting the serviceability of the structure.

One such method has been attempted in this paper to detect cracks in concrete materials automatically. The idea is to first take images using a camera of cracks present in concrete. Using these images, a database is developed. Each image is pre-processed to get required information. These preprocessed images are automatically processed for indicating if the image captured has a crack or not. General block diagram is given in Fig. 1.

The most commonly used technique for automatic detection is that of artificial neural networks (ANN). ANNs are being widely used recently in all applications to automate things and also for faster response. The advancement over ANN is the convolution neural network (CNN). These have been extensively used in medical applications [1], banking services [2], forecasting [3], military [4], etc. They are also being employed in NDT [5].

Note: Interested readers are requested to read [5, 6] for more information on deep learning and also in-depth knowledge on ANN, CNN, RNN and various other techniques available.

This paper makes an attempt to apply CNN for NDT to detect cracks in concrete materials. The rest of the paper is organized as: Sect. 2 discusses the related work in this area. Section 3 explains in brief the CNN model and the methodology developed.

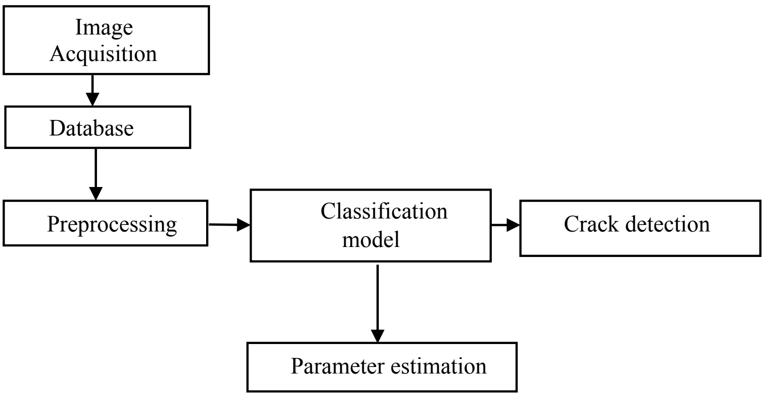


Fig. 1 General block diagram

The results of the methodology are explained in Sect. 4. Finally, Sect. 5 concludes the paper and also proposes the future work in this area.

2 Related Work

This section briefly describes the work done in crack detection using various techniques.

The paper [7] discusses the existing methods for crack detection, crack image capturing using different techniques. The paper also outlines the research challenges and achievements in the field of crack detection. The authors describe the image processing techniques such as camera based, IR based, ultrasonic based, laser based and others. Using these techniques, the crack images are captured, and then, the crack detection process is performed. The crack detection processes are also discussed in this paper.

The authors in [8] proposed a crack detection network based on image segmentation network. They trained the Conditional Similarity Network (CSN) model using the normal crack dataset and also two integrated datasets. The CSN network was tested under different conditions such as hue, white noise and brightness. The advantage of this paper is that the crack images captured are not scaled down to lower resolution like the common practice in deep learning applications. But, the image dataset used is small with only 242 crack images. Also, since the image sizes are processed in their original form, the processing time is also more. The results of precision and recall under various brightness conditions are mentioned in the paper.

In [9], the authors use a CNN technique to perform crack detection. They have used AlexNet with modifications for their application. Their dataset consists of 1250 images which were cropped into 60,000 smaller images of 256 $\times$ 256 pixel resolutions. This paper conducted the study using Caffe on a CPU system with a GPU installed on it. They achieved a validation accuracy of 99.06% at 13,450th epoch while training the model for 15,000 epochs.

NDT methods are utilized to verify the quality and structural integrity of manufactured parts. It is essential for the inspection of discontinuities and breakdown without destructing and damaging the part. In situ monitoring and inspection [10], using NDT technologies during the printing process afford both time and cost savings as the process can be stopped if a failure is found.

In this paper [11, 12], the authors proposed a crack detection technique for concrete tunnel surfaces. The technique used convolutional layers of CNNs for selecting the actual features. They achieved better results in terms of lower implementation time when compared to traditional techniques. The results obtained in the paper indicate that the technique used is better.

3 Methodology

In order to perform crack detection, a CNN model is used. This includes four steps, namely building a database, training the model, validating the model and finally testing the model.

3.1 Database

An existing database available at [13] has been used for training, validating and testing the CNN model. The database contains concrete images having cracks collected from METU Campus Buildings. These images are divided into two categories, positive (with crack) and negative (without crack). Each category has 20,000 images with a total of 40,000 images with 227 × 227 pixels with RGB channels. The dataset is generated from 458 high-resolution images (4032 × 3024 pixel). No data augmentation in terms of random rotation or flipping is applied.

This dataset is directly fed into the CNN model, and classification is performed. CNN model is described as follows.

3.2 CNN Model

This subsection explains the CNN architecture. The basic architecture used in this paper is the AlexNet architecture [13].

It consists of 5 convolution layers, 3 max pool layers, 3 fully connected layers and finally an output layer with a softmax function. All these layers are briefly mentioned in the following sections. Rectified Linear Unit (ReLU) is also used as an activation step in the model.

AlexNet architecture: This is a powerful and extensively used CNN model for image classification. It extracts useful information about the images and also learns new features in each layer. This architecture performs image classification where the input is an image which may be one among 1000 classes. The output of this architecture is a 1000 vector output. The number of classes in this paper is only 2, and hence, the output layer is modified to have only 2 classes (crack and without crack).

Convolutional layer: This is the core building block of a CNN model. The main function of this layer is to reduce the size of the image using kernels/filters but still maintaining the important features in order to identify the image. The idea behind reducing the image size is to ease up the processing and also enhance the speed. In this paper, 5 convolution layers are used. Details of the CNN model are mentioned in Table 1.

Table 1 CNN model

Layer	Filter size	Stride	Activation	Num_Output
Conv 1	11 × 11	4	ReLU	96
MaxPool 1	3 × 3	2	–	96
Conv 2	5 × 5	1	ReLU	256
MaxPool 2	3 × 3	2	–	256
Conv 3	3 × 3	1	ReLU	384
Conv 4	3 × 3	1	ReLU	384
Conv 5	3 × 3	1	ReLU	256
MaxPool 2	3 × 3	2	–	256
FC1	–	–	ReLU	4096
FC2	–	–	ReLU	4096
FC3	–	–	ReLU	4096
Output-Softmax	–	–	–	2

Max pooling layer: This layer is also used for reducing the size of the convolved image. This layer aids in reducing the computational power that is required to process the data by reducing the dimension of the image. It still extracts the main features and thus maintains the quality of the image for effective training.

There are two main types of pooling: max pooling and average pooling. In this paper, the max pooling method is used. This method returns the maximum value from the portion of the image covered by the kernel associated with the pooling layer.

ReLU activation function: A neural network has many activation functions such as sigmoid, tanh, Rectified Linear Activation function (ReLU) and others. In this paper, ReLU is used as an activation function since it is commonly used in CNN models. This function returns the input as it is if the input is greater than 0.0, else it returns a 0 for any negative inputs. In short, it can be represented as $f(x) = \max(0, x)$ [14].

Fully connected layer: The input to this layer is from the previous convolution/pooling layer. This layer uses the results from the previous layer to classify the image into a particular class. In this paper, 3 fully connected layers are used.

Softmax layer: Since, in this paper only binary classification (crack and non-crack) is performed, either softmax or sigmoid function can be used. In this paper, the softmax function is used by setting the output to have only 2 classes

4 Results

The first experiment was carried out on an Intel Core-i5 with 8 GB of memory. The training could be done only for 4000 samples, and then, the system gave an out of memory error. Hence, the training was carried out on an Intel Core-i5 with an

Table 2 Confusion matrix

	Predicted No	Predicted Yes
Actual No	TN	FP
Actual Yes	FN	TP

NVIDIA TITAN X (PASCAL) GPU with 12 GB memory. The dataset with 40,000 images was split into 3: training, validation and testing. 80% of the data was used for training. The remaining 20% was split equally for validation and testing. The model was trained for 2000 epochs with a batch size of 1000. The model took 2.5 h for training and validating using 32,000 and 4000 images, respectively. At 1729 epoch, a validation accuracy of 99.5% was achieved.

After the model was fit using the training and validating samples, the performance of the model was checked using the test samples, i.e., the remaining 4000 samples kept aside for testing. These samples were never used in training or validation of the model. In these samples, the images with crack and without crack are categorized as true positives (TP) and true negatives (TN), respectively, otherwise they are categorized as false positive (FP) and false negative (FN).

The main parameter to evaluate the effectiveness of the model is to obtain the confusion matrix. This is a kind of a table that indicates predictions made by the model and how right/wrong are those predictions. The confusion matrix for the CNN model proposed in this paper is as shown in Table 2.

Based on the four terms, TP, TN, FP and FN, the following parameters were monitored and results obtained

Accuracy: indicates out of all the predictions (TP, TN, FP, FN) what percentage of the fraction is predicted correctly. Formula for accuracy is given in (1)

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{1}$$

Precision: is a term that indicates how precise the model is out of the true predicted positive values, how many are actually positive. Formula is given in (2)

$$\text{Precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})} \tag{2}$$

The TP and FP together form the true predicted positive values.

Recall: The TP and FN together form the actual positive values. Recall indicates out of the actual positive values, how many TP the model is able to identify. Formula for the same is given in (3)

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \tag{3}$$

Table 3 Parameters evaluating the effectiveness of the CNN model

Parameters	Values (%)
Accuracy	99.73
Precision	99.60
Recall	99.85
F1 score	99.73

F1 Score: indicates the accuracy of the model including precision and recall. A good F1 score, i.e., close to 1, indicates that there are low FP and FN. This indicates that there are less false alarms. Formula for the same is given in (4)

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{4}$$

The actual results obtained using the 4000 test samples, for all the parameters mentioned above, are given in Table 3.

Confusion Matrix

N = 4000	Predicted No	Predicted Yes
Actual No	TN = 1979	FP = 8
Actual Yes	FN = 3	TP = 2010

As can be seen from the confusion matrix, only 11 out of the 4000 images have been misclassified as either FN or FP. The rest of the images have been identified as correct (either TN or TP) as indicated with a test accuracy of 99.73%.

5 Conclusions and Future Work

This paper presents the initial results obtained by using a neural network for concrete crack detection. The standard AlexNet has been modified as per the requirement, and a CNN model was built. This model is trained using an existing database consisting of 40,000 images (20,000 for with crack and 20,000 without cracks). The training was done using 80% of the images, and the remaining 20% was used for validating and testing the model. The model was trained for 2.5 h and acquired a validation accuracy of 99.5% at the 1729th epoch (while the model was run for a total of 2000 epochs). The model was obtained more than 99.6% in all evaluation parameters such as accuracy, precision, recall and F1 score as indicated in Results section.

This model was trained only for concrete images. As a future scope, the authors intend to improve the model by using images of steel, iron and other materials. Also, in the images of the database, there is no disturbance such as noise or blurring. The authors will be extending this work to include such disturbances and monitor the performance of the model.

References

1. Essam H, Houssein, Moataz Kilany (2017) ECG signals classification: a review. Int J Intell Eng Inform 5(4)
2. Leo M, Sharma S, Maddulety K (2019) Machine learning in banking risk management: a literature review. Risks 7(29)
3. Zhao Y, Li J, Yu L (2017) A deep learning ensemble approach for crude oil price forecasting. Energy Econ 66:9–16. ISSN: 0140-9883
4. Sonesson L, Boffard K, Lundberg L, Rydmark M, Karlgren K (2018) The potential of blended learning in education and training for advanced civilian and military trauma care. Injury 49(1): 93–96. ISSN: 1879-0267
5. Hou W, Wei Y, Guo J, Jin Y, Zhu C (2018) Automatic detection of welding defects using deep neural network. J Phys Conf Ser 933
6. Goodfellow I, Bengio Y, Courville A (2017) Deep learning. The MIT Press, Cambridge, MA
7. Mohan A, Poobal S (2018) Crack detection using image processing: a critical review and analysis. Alexandria Eng J 57(2):787–798
8. Lee D, Kim J, Lee D (2019) Robust concrete crack detection using deep learning-based semantic segmentation. Int J Aeronaut Space Sci 20:287–299
9. Li S, Zhao X (2019) Image-Based concrete crack detection using convolutional neural network and exhaustive search technique. Hindawi Adv Civ Eng. Article ID 6520620, 12 p
10. Lu QY, Wong CH (2018) Additive manufacturing process monitoring and control by non-destructive testing techniques: challenges and in-process monitoring. Virtual and Physical Prototyping 13(2):39–48. https://doi.org/10.1080/17452759.2017.1351201
11. Protopapadakis E, Voulodimos A, Doulamis A, Doulamis ND, Stathaki T (2018) Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing. Appl Intell J 49:2793–2806
12. Özgenel ÇF, Sorguç AG (2018) Performance comparison of pretrained convolutional neural networks on crack detection in buildings. ISARC 2018, Berlin
13. Krizhevsky Alex, Sutskever Ilya, Hinton Geoffrey E (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
14. Nair V, Hinton GD (2010) Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning, pp 807–714. Haifa, Israel, June 2010

On Computer-Aided Diagnosis of Breast Cancers Using Nuclear Atypia Scoring

Soorya Shaji, M. Sreeraj, Jestin Joy, and Alphonsa Kuriakose

Abstract Computer-aided systems are gaining interest in every field over the world. The medical field has also enhanced to a great extent with the use of computer-aided diagnosis. Among the different diseases in the era, breast cancer has shown a rapid hike in the number of deaths. Scoring of nuclear atypia is an efficient method for the prognosis of breast cancer. The biopsy samples taken from the suspicious tissues are analysed under microscope by the pathologists and are graded. But manual grading highly depends on the pathologists and can cause variation in the results. Hence, the requirement of computer-aided systems for grading has increased. Many studies related to nuclear atypia scoring have taken place in the literature based on different algorithms and classifiers. This paper gives an overview of the different studies in the literature, related to nuclear atypia scoring. Various techniques are used for nuclear atypia scoring. Multifarious image processing techniques are used for this. The aim of this study is to analyse these techniques and their results and know the most efficient one from them. Our analysis shows that promising results are achieved by machine learning techniques. Scores obtained using these techniques are comparable to manual grading.

Keywords Medical imaging · Nuclear atypia scoring · Machine learning

S. Shaji (✉)
Department of MCA, CUCEK, Kuttanad, Kerala, India
e-mail: soorya12340@gmail.com

M. Sreeraj
Department of Computer Science, Sree Ayyappa College, Eramallikara, India
e-mail: sreerajtkzy@gmail.com

J. Joy
Department of Computer Science and Engineering, FISAT, Angamaly, Kerala, India

A. Kuriakose
Department of Computer Science and Engineering, VJCET, Vazhakulam, Kerala, India

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_57

1 Introduction

Cancer is one of the most common death-causing diseases in the current era. The disease involves the abnormal growth of cells which divides itself, spreading to other parts of the body. Hence, this growth should be controlled or else it results in death. Some kinds of cancer have rapid growth, whereas others divide at a slow rate. Different types of cancer include breast cancer, prostate cancer, lung cancer, leukaemia, etc. Cancer can be caused due to many reasons. Causes of cancer include consumption of alcohol, smoking, increased body weight, age, the problem with immune systems or even genetics. According to the recent study by the American Cancer Society [26] in 2018, about 1.7 million people were expected to diagnose with cancer, and death of 609,640 was estimated. Since these numbers are huge, the problem has to be solved. Many studies are taking place for the prevention and detection of this disease. Cancer occurs in different stages. It varies from stage 0 to stage 4. Stage defines the size and spread of the tumour. However, the determination of stage is important because it is based on the stage that the treatments and prognosis are determined. If breast cancer is diagnosed in the initial stage, there is a high possibility for a cure. That is, early detection has a high influence on recovery.

Among the different types of cancers, the most common kind of cancer is breast cancer. It is mostly seen in women. According to the study by the World Health Organization (WHO), about 627,000 women in the world die from breast cancer. The cancer is formed in the lobules or ducts of the breast. The symptoms for breast cancer include a lump or mass in the breast, thickness, redness, scaling, swelling, distortion, and more. The main two types of breast cancer are ductal carcinoma and lobular carcinoma. They can be invasive or non-invasive. An invasive carcinoma spreads fast to other tissues. On the other hand, non-invasive carcinoma does not spread from its original tissue. Age, gender, family history, pregnancy, hormones, body weight, etc., can be the reason behind the cause of breast cancer. But the exact reason for breast cancer is still undefined. However, once a patient is diagnosed with cancer, the only solution is to take the treatment as soon as possible. The common treatment involves the lumpectomy and mastectomy. Lumpectomy is the surgical removal of the tumour and the surrounding tissues. On the other hand, a mastectomy involves surgical removal of the breast. Other than the surgery, chemotherapies, radiation therapies, hormone therapies, and other medications are also taken by the patient.

Early detection followed by appropriate treatment can help to increase the rate of survival from the disease. Even though there are many technologies for the detection of cancer, a biopsy is used to finalize malignancy. The system focuses on nuclear atypia scoring which is based on the Nottingham grading system. The biopsy samples are graded based on nuclear atypia.

The manual process of grading requires a lot of human effort, and it can vary according to the pathologist. Therefore, computer-aided systems are now gaining interest. Many studies have emerged based on computer-aided systems for breast cancer grading. The grading mainly involves pre-processing, segmentation, feature selection, and classification of the images. Different algorithms are used for each of the stages for obtaining an accurate result.

Figure 1 shows a comparison of stages in the manual and computerized grading system. The manual grading is depicted in Fig. 1a. It begins with the sample collection followed by the preparation of slides. The slides are stained to distinguish the cells under the microscope. These slides are then observed under the microscope by the pathologists and are graded. Figure 1b defines the automated grading. Training and testing are the two phases involved. Each phase comprises pre-processing, segmentation, and feature extraction. The data is trained first to generate a model, and this model is used for testing and grading.

Mammography is one of the most common technologies used for breast cancer detection. It can be used for detecting the disease at an early stage. But the problem was that its results were not perfect. It resulted in many false positives and false negatives [11, 19, 27]. Another technique is ultrasound imaging, which assesses the morphology, orientation, and internal structure of the tumour. It is usually done after mammography. If mammography generates a positive result, the ultrasound imaging is used to detect the size, location, and shape of the tumour. The problems with the machines and other objects result in noise and complicate the process. Magnetic resonance imaging (MRI) is also widely used for detection. From the image, the traces of cancer are identified, if any. An MRI detects cancer more efficiently, compared to mammography and ultrasound [17]. But MRI is too much sensitive and results with many false positives. It also lacks specificity, and it is costly. However, people with

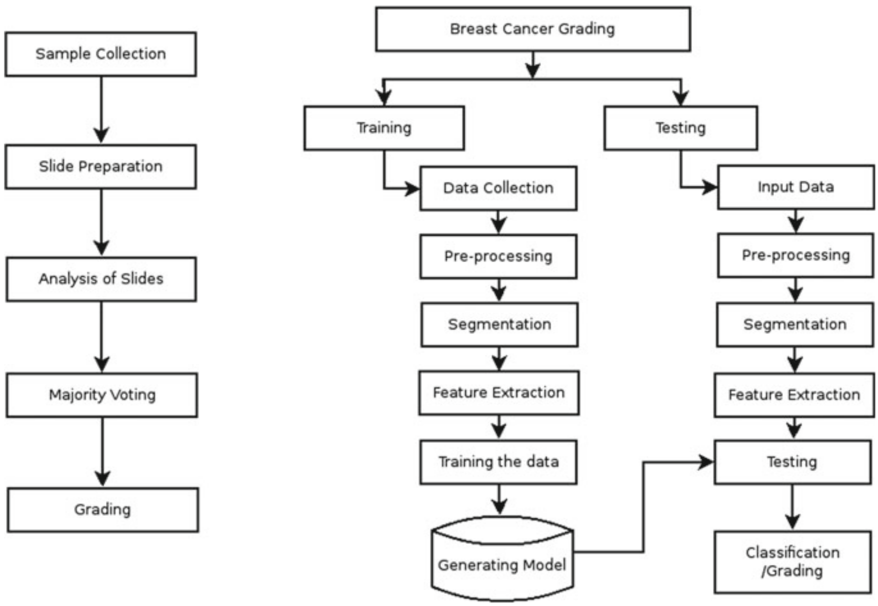


Fig. 1 **a** Manual grading; **b** Automated grading

a high chance of the disease are recommended to take an annual MRI [26]. Even though there are different methods for breast cancer detection, the final verification has to be done with the help of breast biopsy. In a biopsy, samples of tissues or cells are extracted from the suspicious region, and then a pathologist analyses them through the microscope to check malignancy. The analysis of the biopsy sample is a lengthy process, and there are many methods for it [25]. The most common prognostic determinants for malignancy in a biopsy are lymph node (LN) status, tumour size, and histological grade. Among the three, the most accepted determinant is histological grading which is simple and less expensive. It defines the morphological features of the tumour cells and provides high accuracy results. Hence, the use of histological grading in breast cancer has gained wide acceptance.

But the manual process of breast cancer grading (BCG) [29] requires a large amount of time and effort. It requires the pathologist to examine thousands of slides via a microscope to check malignancy. The analysis is to be carried out by more than one pathologist for the confirmation of the result, and it requires a great amount of human labour. A comparison of manual and automatic grading is given in [27] by F. Schnorrenberg. The accuracy of manual process is affected by the inter–intra observer variation, and the individual nucleus is counted and classified based on the stain intensity. On the other hand, the automated system provides the facility to assess many areas and gives a better result.

As a result, image processing and computer-aided diagnosis (CAD) have received much attention.

Image processing is based on processes of digital images. It is used for classification, feature extraction, pattern recognition, and so on. CAD also contributes to the early detection of breast cancer. It is the use of computer technologies to create technical designs or patterns. Lots of studies have taken place in the literature related to image processing [1, 12, 18, 28] systems in mammography. The mammography images are analysed to identify the mass and calcifications. The same can be used to automate histological grading. Tutac et al. [29] suggest a solution to the automatic breast cancer grading using the Nottingham grading system.

Nottingham grading system [9] is a widely used breast cancer grading system. It analyses the invasive adenocarcinomas based on three components: (1) tubule formation (2) mitotic count, and (3) nuclear atypia (NA). Tubule formation gives the degree of glandular differentiation in invasive ductal carcinoma. Mitotic count refers to the rate of spreading of cancer cells. Nuclear atypia refers to the difference in the structure of cell nuclei compared to the original cells. There are many studies that make use of the Nottingham grading system for breast cancer grading [6, 24]. The three components in NGS are given a score of 1, 2, or 3. At the end, the three grades are added together, providing a minimum score of 3 ($1 + 1 + 1$) and a maximum score of 9 ($3 + 3 + 3$). These scores are then converted to histological grades. Besides the combined use of the three components, each of them can be used individually for breast cancer detection. In tubule formation, the cells form a tubular structure. It contains a clear white blob, called lumina, surrounded by tubular structures. The smaller the number of tubular structures, the higher will be the grade. In [23], Nguyen et al. define the segmentation of the tubule structure for grading. The system identifies all the nuclei and lumens from the images and then classifies the malignant ones using a random forest classifier. On the other hand, Ko et al. in [16] suggest a method for tubule detection. It initially detects ROIs at the lesion areas. Then, it segments the gland and lumen. Later, the gland features, lumen features, and texture features are extracted and are then extracted using SVM classifier.

Whereas mitotic count relies on the rate of cell division. Mitosis is the process in which a cell is broken down to identical daughter cells. The counts of these daughter cells are then taken. They are found less in regions of tubule formation. Nateghi et al. [22] describe the use of a CAD system for mitosis detection. The paper suggests the use of a maximum likelihood algorithm to estimate the parameter of mitosis. Candidates are detected using this algorithm. The process is then followed by feature extraction and classification. The classification is done using the SVM classifier. Logambal and Saravanan in [20] and Amitha et al. in [2] use the SVM classifier for mitosis detection.

2 Methodology

There are many studies in the literature based on tubular formation and mitotic count, whereas the studies based on the third component, nuclear atypia (NA), are found less in the literature. Nuclear atypia differentiates the nuclei based on its size and shape. The nuclei are compared with original cells to determine the variation. Compared to these two methods, nuclear atypia (NA) gives an identical result for grading. In nuclear atypia scoring (NAS), the score of nuclear atypia is used for histological grading. The size, shape, and the chromatin distribution of the nucleus are analysed and are given a score accordingly. A score of 1 indicates low nuclear atypia (highly differentiated) and 3 indicate a strong nuclear atypia (rarely differentiated).

As specified before, automated grading can be done based on CAD. Chekkoury et al. in [4] suggest a system using CAD for malignancy detection. It differentiates the H&E stained biopsy samples based on malignancy. The slides are stained with haematoxylin and eosin (H&E) for microscopic evaluation. The method involves three stages: (1) image processing, (2) feature extraction, and (3) classification. The systems extract morphological, textual, and topological features for processing and are then classified to benign or malign using SVM classifier.

All these studies are carried out on Whole-Slide Histopathological Images. There are many technologies and algorithms of CAD and image processing in the literature, which are used in combination to grade the histological images. Some methods segment the image into regions for further processing. Images with high resolution are needed to be segmented for efficient processing [5, 21]. At the same time, images are also processed without segmentation.

Cosatto et al. in [5] describe methods of analysing the histological micrographs for grading of nuclear pleomorphism. The process is initiated with the colour separation process. Then, the nuclei that are larger than the normal threshold are detected and features are extracted using Hough transform. Then, an active contour model (ACM) is used for the purpose of segmentation of images. The outlines of the nuclei are then extracted for further classification. The shape features and texture features are taken into account. The system is then trained using support vector machine (SVM) classifier.

Lu et al. in [21] segmented the images into patches, after the process of stain normalization and separation. Textual and morphological features are then extracted, and histograms for each feature are constructed. Later features are extracted from each of the segment, and the support vector machine (SVM) classifier is used for the task of classification. It inferred that textual features are likely to contribute more classification process accuracy. Use of deep learning for nuclear pleomorphism was discussed in [31]. It focuses on two tasks in which one is the nuclear atypia scoring. It makes use of multi-resolution convolutional network (MR-CN) for processing. The method resulted in an accuracy of 79.87%.

The idea of a novel image level descriptors is presented in [15]. The images are segmented into regions. The region covariance (RC) descriptors are calculated for each region and are combined to a single form using the geodesic mean of region covariance (RC) or gmRC. This symmetric positive definite (SPD) matrix is derived and is plotted on a Riemannian manifold. For the purpose of classification, they used different machine learning classifiers like decision tree (DT) classifier, RUSBoost, kNN, linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA). The result was derived with an accuracy of 83.3% under G-kNN classifier.

As defined in [15], the idea of SPD matrices and Riemannian manifold is used in [8]. But it performs sparse coding and dictionary learning on the SPD matrices for the task of classification. The experiment resulted in a higher accuracy of 87.92%. Apart from this, Dalle et al. [7] suggested a method that automatically selects and segments the critical cells in place of selecting all the cells and thus reducing the expense. It adapts the idea of region of interest (ROI) for cell selection and classification. It resulted in a minimum error rate of 7.8%. Whereas Faridi et al. [10] perform manual grading of nuclear atypia, after efficient pre-processing. While Dalle et al. [7] dealt with the critical cells, Faridi et al. in [10] suggest detecting the cancerous cells and deriving its exact outline. The segmentation is done in two ways: by detecting the centroid and cell boundary of the nuclei. The paper resulted in an accuracy of 87%.

3 Dataset

Most of the studies related to nuclear atypia scoring are done on the MITOS-ATYPIA-14 dataset [8, 10, 15, 21, 31]. This was a challenge conducted by the International Conference on Pattern Recognition (ICPR). It was a contest using breast cancer histological images. The contest had two parts: mitotic count and nuclear pleomorphism. The dataset contains haematoxylin and eosin (H&E) stained slides. These are scanned using two slide scanners. They are Hamamatsu Nanozoomer 2.0-HT and the Aperio Scanscope XT. Images of 10×, 20×, and 40× magnification are taken.

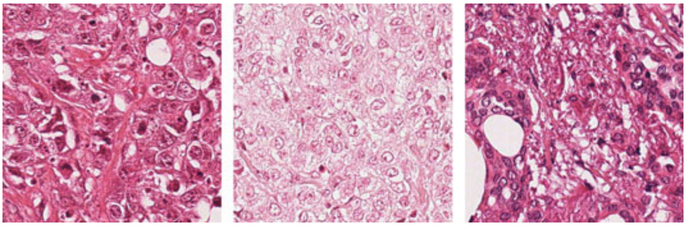


Fig. 2 Sample dataset images with NAS = 1, NAS = 2 and NAS = 3

The dataset is divided into a training set and testing set. 80% is set aside for training and the rest for testing. Figure 2 shows the sample images in the dataset with nuclear atypia score (NAS) score 1, 2, and 3.

4 Taxonomy

Figure 3 depicts the taxonomy on different techniques used in different stages of the grading. Pre-processing, segmentation, feature selection, and classification are the important stages in the computer-aided diagnosis [3]. Different methods can be adopted for each stage, and they are used in different combinations to provide better result.

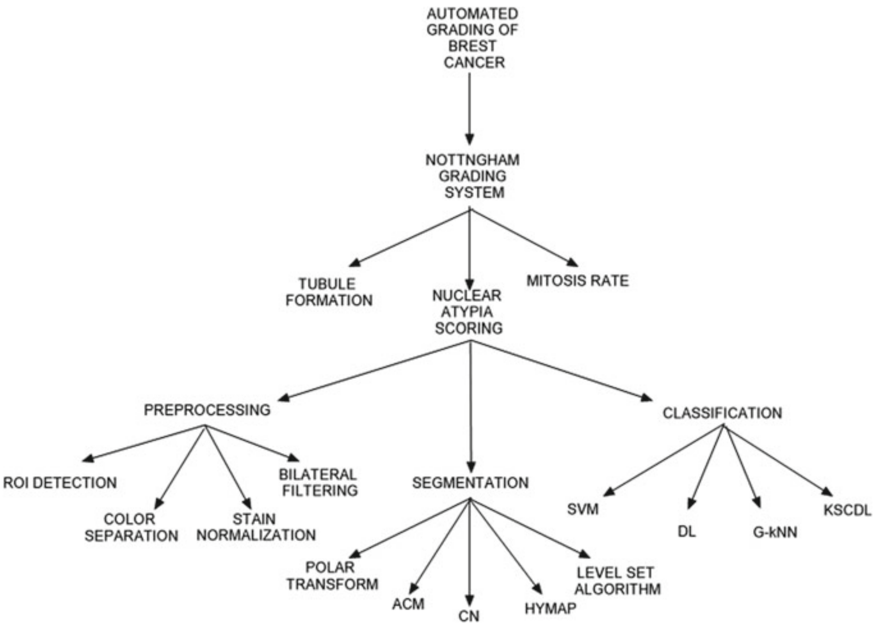


Fig. 3 Taxonomy

4.1 Pre-processing

Pre-processing aims at the removal of unwanted distortions and highlighting the required features. It performs noise reduction and background removal. There are many pre-processing methods defined in the literature. It is performed prior to segmentation and detection. Different pre-processing techniques used for image processing. According to Irshad et al. [13], pre-processing methods are illumination normalization, colour normalization, noise reduction, image smoothing, and ROI detection. By illumination normalization, it means to correct the illumination of the image using white shading correction or by deriving its pattern from a set of images. Colour normalization techniques include histogram or quantile normalization. It is mainly focused on stain removal because the samples we take are stained with haematoxylin and eosin (H&E) dyes. Noise reduction aims at minimizing unwanted noise and artefacts. It uses morphology for this purpose. Region Of interest (ROI) plays simultaneously with noise reduction. Techniques like thresholding and clustering are used for this purpose. Cossato et al. [5] and Dalle et al. [7] used pre-processing for ROI detection. On the other hand, Khan et al. [15] and Das et al. [8] used stain normalization. Bilateral filtering was used by Faridi et al. [10], which is the same as image smoothing or noise reduction.

4.2 Segmentation

Segmentation is required to divide the images into smaller regions. The images can be of high resolution, and sometimes it is required to break it down to patches. The traditional image processing techniques for segmentation are specified the in paper by Irshad et al. [13]. They are thresholding, morphology, region growing, watershed, ACM, K-means clustering, probabilistic models, and graph cuts. Thresholding refers to the process of converting an image to a binary form. It assigns each pixel a value of 0 or 1 depending on a threshold value. Morphology processes the image as geometrical structures based on the morphological features. Region growing is a segmentation method which initially identifies the seed point and then classifies its neighbouring pixels. Watershed is also a segmentation method which begins from a particular pixel. It mainly aims at dividing the foreground and background image. Veta et al. [30] suggest the use of watershed for segmentation. It initially processes the image using region of interest (ROI) and is then segmented. Active contour model (ACM) is used to obtain the contour of objects using the gradient information. It is also called deformable models. It is also used to identify the nuclei boundary in nuclei segmentation. Jing et al. [14] suggest the use of a hybrid ACM for segmentation. The K-means clustering divides the image into k clusters iteratively. The probabilistic model is an extension of K-means clustering. And finally, graph cut considers the image as a weighted undirected graph, and a certain set of algorithms are used for segmentation.

Table 1 Segmentation

Authors	Methods	Parameters	Accuracy (%)
Khan et al. [15]	HyMap	Tissue morphology	87
Dalle et al. [7]	Polar transform segmentation	Polar space, centre of cell nuclei (origin), radius, angle	84.1
Farirdi et al. [10]	Level set algorithm	Identify points at which the image brightness changes	87

Nguyen et al. in [23] and Cosatto et al. in [5] use an ACM for segmentation. On the other hand, Dalle et al. [7] segment the image using polar transform segmentation. It segments on the basis of polar coordinates. Xu et al. [31] use convolution network, and Khan et al. [15] use HyMap. HyMap stands for a hybrid magnitude-phase approach. Among all these, ACM was found to have a better result. Table 1 summarizes some of the techniques used for segmentation.

4.3 Feature Selection

Feature selection involves the extraction of cytological features, morphological features, texture features, topological features, and more [3]. Cytology features comprise size and shape. Morphological features detect shape and margin. Texture features indicate the pattern of nuclei tissue. Similarly, there are many features. Table 2 lists the different types of features and their description. Figure 4 shows the different features that can be extracted from the image in the MITOS-ATYPIA1-14 dataset. They include the nuclei size, anisonucleosis, nucleoli size, chromatin density, nuclei contour, and membrane thickness.

Table 2 Feature selection

Feature type	Features
Cytological features	Increased nuclear size, pleomorphism, lack of differentiation, mitosis
Morphological features	Size and shape of the nucleus
Texture features	The pattern of the inside of the tissue
Topological features	Geometric features (points, lines, edges, space, etc.)

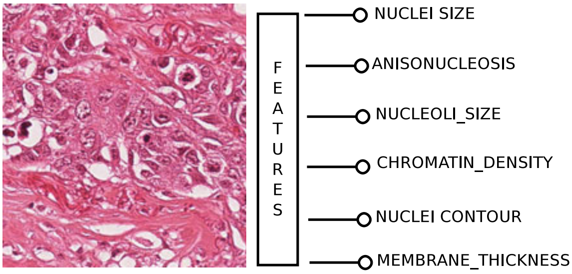


Fig. 4 Features

4.4 Classification

Features extracted are given as input for classification. The most commonly used classification is the SVM classifier. SVM with ACM resulted in high accuracy of 89% [5]. Deep learning has also been widely used in classification. Studies infer that deep learning has a better impact on classification [31]. Other algorithms used

Table 3 Techniques

Name	Pre-processing	Segmentation	Feature set	Classification	Results
Grading nuclear pleomorphism on histological micrographs [5]	ROI	Active contour	Morphological	SVM	89%
Nuclear pleomorphism scoring by selective cell nuclei detection [7]	ROI	Polar transform segmentation	Morphological	SVM	82.2%
Deep learning for histopathological image analysis: towards computerized diagnosis on cancers [31]	–	Convolutional network	Texture morphological	CNN	79.87%
Automated image analysis of nuclear atypia in high-power field histopathological image [21]	Stain normalization, colour separation	Seed detection, local threshold, and morphology operations	Texture morphological	SVM	74.87%
A global covariance descriptor for nuclear atypia scoring in breast histopathology images [15]	Stain normalization	HyMaP	Morphological	Geodesic K-nearest neighbour (G-Knn)	83.29%
Sparse representation over learned dictionaries on the Riemannian manifold for automated grading of nuclear pleomorphism in breast cancer [8]	Stain normalization, colour separation	-	Morphological	Kernel-based sparse coding and dictionary learning	87.92%
An automatic system for cell nuclei pleomorphism segmentation in histopathological images of breast cancer [10]	Bilateral filtering	Level set algorithm	Morphological	Manual classification	87%

were G-kNN [15] and KSCDL [7]. Table 3 lists out the different techniques used in different stages in different literature studies.

5 Conclusion

In this paper, we reviewed the different techniques used in literature for nuclear atypia scoring. The studies made use of different techniques for better results. However, the CAD systems has highly influenced the field of breast cancer detection. Many methodologies and techniques have been suggested in the literature for efficient grading using CAD systems. In the whole process, the analysis of biopsy sample is a crucial step in grading of cancer. It is based on this grading that further procedures are determined. Each stage in processing is important, and hence, there are related studies for each.

References

1. Abdallah YM, Elgak S, Zain H, Rafiq M, Ebaid EA, Elnaema AA (2018) Breast cancer detection using image enhancement and segmentation algorithms. Biomed Res 29(20):3732–3736
2. Amitha H, Selvamani I, Dhas DAS (2017) Development of computer aided system for detection and classification of mitosis using SVM. In: 2017 international conference on inventive computing and informatics (ICICI). IEEE, pp 954–958
3. Aswathy M, Jagannath M (2017) Detection of breast cancer on digital histopathology images: present status and future possibilities. Inform Med Unlocked 8:74–79
4. Chekkoury A, Khurd P, Ni J, Bahlmann C, Kamen A, Patel A, Grady L, Singh M, Groher M, Navab N, et al (2012) Automated malignancy detection in breast histopathological images. In: Medical imaging 2012: computer-aided diagnosis, vol 8315. International Society for Optics and Photonics, p 831515
5. Cosatto E, Miller M, Graf HP, Meyer JS (2008) Grading nuclear pleomorphism on histological micrographs. In: 2008 19th international conference on pattern recognition. IEEE, pp 1–4
6. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC (2008) Automatic breast cancer grading of histopathological images. In: 2008 30th annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 3052–3055
7. Dalle JR, Li H, Huang CH, Leow WK, Racoceanu D, Putti TC (2009) Nuclear pleomorphism scoring by selective cell nuclei detection. In: WACV
8. Das A, Nair MS, Peter SD (2018) Sparse representation over learned dictionaries on the Riemannian manifold for automated grading of nuclear pleomorphism in breast cancer. IEEE Trans Image Process 28(3):1248–1260
9. Elston CW, Ellis IO (2002) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991; 19; 403–410: author commentary. Histopathology 41(3a):151
10. Faridi P, Danyali H, Helfroush MS, Jahromi MA (2016) An automatic system for cell nuclei pleomorphism segmentation in histopathological images of breast cancer. In: 2016 IEEE signal processing in medicine and biology symposium (SPMB). IEEE, pp 1–5
11. Guo R, Lu G, Qin B, Fei B (2018) Ultrasound imaging technologies for breast cancer detection and management: a review. Ultrasound Med Biol 44(1):37–70

12. Guzmán-Cabrera R, Guzmán-Sepúlveda J, Torres-Cisneros M, May-Arrioja D, Ruiz-Pinales J, Ibarra-Manzano O, Aviña-Cervantes G, Parada AG (2013) Digital image processing technique for breast cancer detection. Int J Thermophys 34(8–9):1519–1531
13. Irshad H, Veillard A, Roux L, Racoceanu D (2013) Methods for nuclei detection, segmentation, and classification in digital histopathology: a review-current status and future potential. IEEE Rev Biomed Eng 7:97–114
14. Jing J, Wan T, Cao J, Qin Z (2016) An improved hybrid active contour model for nuclear segmentation on breast cancer histopathology. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI). IEEE, pp 1155–1158
15. Khan AM, Sirinukunwattana K, Rajpoot N (2015) A global covariance descriptor for nuclear atypia scoring in breast histopathology images. IEEE J Biomed Health Inform 19(5):1637–1647
16. Ko CC, Cheng CY, Lin CH (2016) A computer-aided grading system of breast carcinoma: scoring of tubule formation. In: 2016 30th international conference on advanced information networking and applications workshops (WAINA). IEEE, pp 918–923
17. Lehman CD, Schnall MD (2005) Imaging in breast cancer: magnetic resonance imaging. Breast Cancer Res 7(5):215
18. Li Y, Chen H, Cao L, Ma J (2016) A survey of computer-aided detection of breast cancer with mammography. J Health Med Inf 4(7)
19. Løberg M, Lousdal ML, Bretthauer M, Kalager M (2015) Benefits and harms of mammography screening. Breast Cancer Res 17(1):63
20. Logambal G, Saravanan V (2015) Cancer diagnosis using automatic mitotic cell detection and segmentation in histopathological images. In: 2015 global conference on communication technologies (GCCT). IEEE, pp 128–132
21. Lu C, Ji M, Ma Z, Mandal M (2015) Automated image analysis of nuclear atypia in high-power field histopathological image. J Microsc 258(3):233–240
22. Nateghi R, Danyali H, Sadegh Helfroush M, Tashk A (2014) Intelligent cad system for automatic detection of mitotic cells from breast cancer histology slide images based on teaching-learning-based optimization. Comput Biol J 2014
23. Nguyen K, Barnes M, Srinivas C, Chefd'Hotel C (2015) Automatic glandular and tubule region segmentation in histological grading of breast cancer. In: Medical imaging 2015: digital pathology, vol 9420. International Society for Optics and Photonics, p 94200G
24. Petushi S, Garcia FU, Haber MM, Katsinis C, Tozeren A (2006) Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. BMC Med Imaging 6(1):14
25. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR et al (2010) Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res 12(4):207
26. Society AC (2017) Breast cancer facts & figures 2017–2018
27. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y (2009) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Trans inf Technol Biomed 13(2):236–251
28. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y (2009b) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Trans Inf Technol Biomed 13(2):236–251
29. Tutac AE, Racoceanu D, Putti T, Xiong W, Leow WK, Cretu V (2008) Knowledge-guided semantic indexing of breast cancer histopathology images. In: 2008 international conference on biomedical engineering and informatics, vol 2. IEEE, pp 107–112
30. Veta M, Van Diest PJ, Kornegoor R, Huisman A, Viergever MA, Pluim JP (2013) Automatic nuclei segmentation in H&E stained breast cancer histopathology images. PLoS ONE 8(7):e70221
31. Xu J, Zhou C, Lang B, Liu Q (2017) Deep learning for histopathological image analysis: towards computerized diagnosis on cancers. In: Deep learning and convolutional neural networks for medical image computing. Springer, Cham, pp 73–95

Using Images for Real-Time Violence Detection in the Edge

Aniruddha Pandey, T. Samarth, and S. Raghuram

Abstract Surveillance cameras have become commonplace in urban scenarios. These cameras are currently proving to be a valuable aid in apprehending criminals after the camera feed is analyzed. Their utility can be further enhanced if detection capabilities are provided, via additional hardware for realizing specific functionality. Currently, Deep Neural Networks are the most popular set of tools for image classification and have also been successfully adapted for other modalities such as video, speech and ECG signals. However, many of these Deep Nets are designed for very complex multi-class problems and hence do not allow for real-time functionality on an embedded platform, due to the large number of parameters involved during the classification task. In this work, we use a modified version of an existing dataset that can be easily trained with published networks to give reasonably good detection accuracy. It is shown via experiments that even smaller nets perform as well as state-of-the-art networks and provide the same level of detection accuracy as the more complex networks. Hence, the simpler networks can be used for a detection task by downloading it to an embedded system and allowing for classification to happen in real time.

Keywords Deep Neural Networks · Classification · Video analytics · Security

1 Deep Neural Networks and Violence Detection

Deep Neural Networks (DNNs) are currently the state of the art for image classification tasks. Even though the first of such networks was proposed as early as 1998 [1], hardware resources and data availability have only recently allowed for their full potential to be realized, starting with the seminal paper in 2012 [2]. Currently, DNNs are considered to be the most accurate model for image classification tasks.

A. Pandey (✉) · T. Samarth · S. Raghuram
M.S. Ramaiah Institute of Technology, Bengaluru 560054, India
e-mail: anipandey1998@gmail.com

S. Raghuram
e-mail: raghuram@msrit.edu

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_58

Since these types of models do not require any feature engineering, they have been rapidly adapted for other modalities such as video analytics, natural language processing, speech recognition, speech synthesis, various classification and prediction tasks related to healthcare technologies. In this work, we are primarily interested in a particular application of edge computing—violence detection in surveillance video camera feeds using DNNs. In the next section, we briefly describe the work done related to DNNs and violence detection.

1.1 Violence Detection Using Deep Neural Networks

The domain of violence detection falls under the more general topic of anomaly detection in surveillance video cameras. The anomalies that have been investigated in recent literature include events such as a fire in a crowded locality, large scale violence, for example, in sports arenas, individuals engage in physical violence, etc. In the rest of this section, we discuss some recent works using DNNs for anomaly detection.

The authors in [3] have created a dataset suitable for anomaly detection and also proposed a DNN that is capable of classifying the videos with reasonable levels of accuracy. Creation and availability of quality datasets related to violence detection are non-trivial due to (1) lack of access to surveillance cameras and (2) relatively sparse occurrences of anomalies, making the data heavily skewed toward normalcy. Due to these problems with existing datasets, [3] is among the first efforts to manually collected videos from YouTube, etc. to create a suitable dataset of 1900 videos, of which exactly half contain a clear view of an anomaly. The set of anomalies collected includes classes such as assault, burglary, accidents and fighting. The classifier itself divides each video into temporal segments, from which features are extracted. These extracted features are then fed to a convolutional neural network for the classification task. Using their dataset and the classifier designed, they have obtained an AUC of about 75.

One of the most important characteristics of DNNs is their capability to learn the features via the data, i.e., requirement of hand-crafted features are not required for the classification task. The authors in [4] have proposed a Deep Net called Appearance and Motion Deep Net or AMDN which uses stacked denoising autoencoders to learn separately features corresponding to motion and appearance. A data fusion scheme is also proposed to combine the features before final classification. From the results generated, similar work using convolutional autoencoders obtains an AUC of 80, while this work achieves an AUC of 87.

A seminal work on activity recognition in videos using DNNs, the authors of [5] propose the concept of two-stream ConvNet architecture for classification. One stream is used to recognize spatial clues in a static image, while the other stream is trained on optical flows generated from consecutive images to recognize the movement of the objects between frames. Individual ConvNet scores are then combined using late fusion. On the UCF-101 dataset with about 100 classes, the model achieves a maximum accuracy of 87% when using an SVM for late fusion.

More recently, DNNs with memory capabilities have become popular for video classification due to the temporal nature of the data. The authors of [6] have used a certain type of Recurrent Neural Network called a Long Short-Term Memory (LSTM) network for unsupervised learning of video representations. Unsupervised learning is gaining importance due to the large amount of data being generated "in the wild" and an absence of resources to label these data. The actual architecture consists of multiple LSTMs: one LSTM functions as an encoder of input frames. Based on the functionality (e.g., future prediction, reproduction), one or more LSTMs are used to act as the decoder. This model too achieves around 75% accuracy on the UCF-101 dataset.

1.2 Edge Computing and Violence Detection

Edge computing [7] is primarily motivated by industrial applications in domains such as IoT. In relation to this work, edge computing becomes extremely relevant due to the high volume of data being continuously generated in surveillance video cameras in an urban setting. Hence, it becomes impractical due to bandwidth limitations to continuously transmit these video to central servers for real-time analysis. In this situation, it becomes essential to implement an edge-based system that is capable of performing analytics at the source of the data and then only send periodic signals about the activities until an event of interest is detected, which will again trigger a signal to be sent with high priority to manual monitoring stations. The overall flow for an example of a detection system implemented on the edge is shown in Fig. 1.

Objectives

Video analytics is a challenging task due to the structure of the data. Further, due to the generality of the classification tasks attempted in the literature sources (e.g., the UCF-101 dataset has 100 classes) the architecture of the DNN becomes complex. Due to this reason, an embedded hardware platform is not capable enough to support real-time classification tasks for these DNN architectures. Further, focusing on typical street-level anomalies related to violence narrows the number of classes down to less than ten classes, which have to be detected in real time. To overcome the above-mentioned bottlenecks, in this work we have focused on two activities:

- Using individual images from surveillance video camera feeds to detect the anomalous activities belonging to a certain class of violent physical actions
- Design of a DNN architecture that is capable of classifying the data, but is also lightweight enough to be downloaded onto an embedded platform.

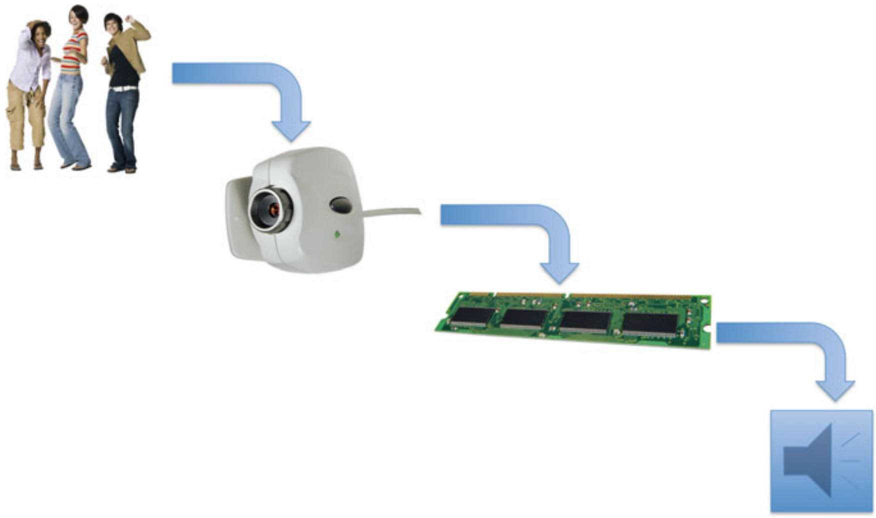


Fig. 1 Sample system for edge activity detection

2 Images for Violence Detection on the Edge

During the last decade, there have been many video datasets like UCF-101, YouTube 1 M dataset, etc. More recently, video datasets have also been made available for surveillance video cameras. It is well known that DNNs requires significantly large amounts of training data for the features to be learnt. To satisfy this requirement, the authors of [3] have published the UCF-Crime Dataset, a video dataset that contains 128 h of video comprising of 13 classes of anomalies. Further, the dataset is also generated from surveillance video cameras, and hence, a model trained on this dataset can be used with little to no modification for any other surveillance scenario.

However, a video dataset will require a complex DNN architecture that will overload the resources of an embedded platform. For example, the SSD GoogLeNet requires about 350 MB of memory for its deployed version and can process only 0.39 Frames/second on a 1.2 GHz Raspberry Pi 3 board with 1 GB of memory [8]. Since the type of violence that we are interested in (e.g., physical violence between individuals) itself will be only a few seconds in length, processing power in embedded systems is a serious bottleneck when DNNs are designed for video architectures.

In this work, we have created images from the video dataset published in [3]:

- The video dataset is converted into an image dataset from the frames of the individual videos. Image datasets are more amenable to classification with relatively smaller DNN architectures. Hence, real-time classification capabilities can be realized with models deployed on embedded systems.

- The number of classes is reduced to exactly two—i.e., whether violence is present in the image or not. Other classes in the original UCF-Crime Dataset are not used, e.g., abuse, arson, etc. since these are not the activities of interest for our work.

2.1 Reduced UCF-Crime Images

The images obtained from the UCF-Crime Dataset video files contain only the actions depicting physical violence between two individuals. Other violence modalities such as arson, robbery and abuse are not included to reduce the complexity of the model. The images in the dataset fall under two classes: direct physical contact of a violent nature between two individuals, and multiple individuals involved in non-violent interactions such as talking, walking and hugging. Examples of the former and latter classes of images are shown in Figs. 2 and 3, respectively. Choosing these two classes is important to reduce the classification errors—i.e., the cases where non-violent interactions are classified by the model as violent could be numerous since many actions involve physical contact between two individuals. Hence, the model should learn to discriminate between violent and non-violent types of contact, by including such examples with the correct labels. As part of the training images, we have looked for and chosen many such images where there is a high chance of being classified as violent, but involves only physical contact as regular interaction. The

Fig. 2 Examples of violence types from the UCF-Crime dataset

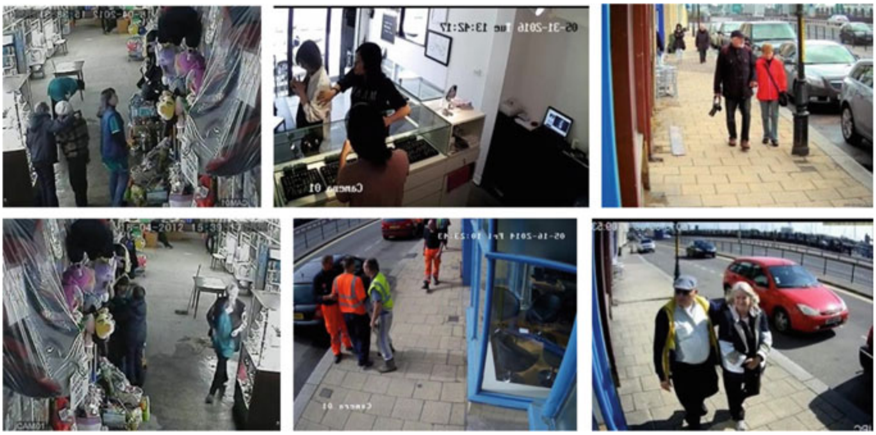


Fig. 3 Normal videos from the UCF-Crime dataset

data contains images to train the network to reduce the instance of false positives, i.e., images classify as normal with no violence, but still have some physical contact such as a handshake, etc. The total number of images classified under the violence category is 1171, while the total number of images classified as normal is 3603. In the next section, we present the results of classification, using various DNN architectures published in the literature.

3 Experimental Results

Rather than design new DNN architectures, we apply many of the already published and popular architectures for our dataset. A key question that we are looking to answer with these experiments is the suitability of smaller architectures for identifying the two classes in our image data. The Keras framework with Tensorflow background was used in realizing the neural networks. We have not reported the training time as we are only interested in the accuracy attained. Further, for ResNet-50 and GoogLeNet we have applied transfer learning [9] and trained only the top two layers, using the pretrained weights from the built-in model in Keras (Figs. 4, 5 and 6).

We have chosen the following three architectures:

- FCN: As a base case, we have chosen a simple neural network consisting of one hidden layer, a fully connected layer with 128 neurons, followed by a Softmax layer with two outputs (we call this architecture FCN in the rest of this paper). A batch size of 50 images was used for training the network in 10 epochs.
- ResNet-50 [10]: The top two layers of the network were only trained with the remaining weights being the same as in the pretrained model. Here too, we used a batch size of 50 images and 8 epochs for the training. The pretrained network

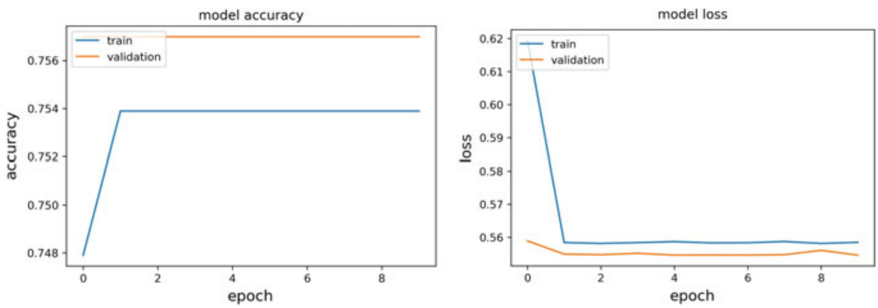


Fig. 4 Training and accuracy for FNC

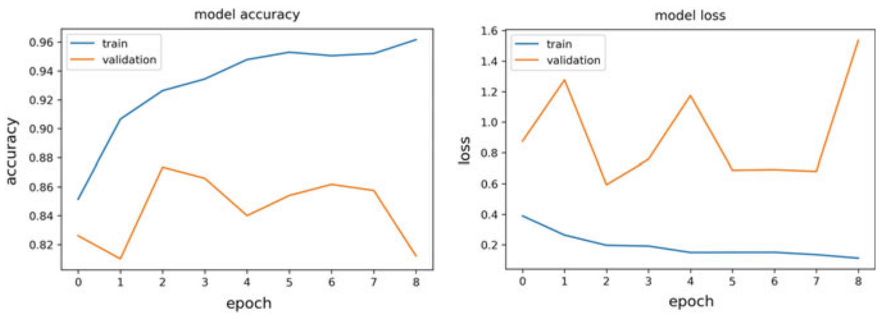


Fig. 5 Training accuracy for GoogLeNet

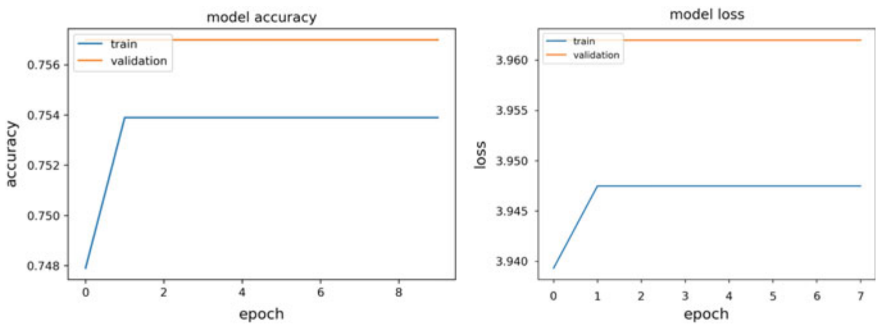


Fig. 6 Training accuracy for ResNet-50

has a total of 50 layers and is trained on the ImageNet dataset to recognize 1000 different categories of objects.

- GoogLeNet [11]: Here too, the top two layers of the network were only trained, with same batch sizes and epochs as the previous two cases. The pretrained weights are obtained as in the previous case by training on the ImageNet dataset.

Table 1 Comparison of accuracy

Network	Number of layers/Number of layers trained	Total trainable parameters (in millions)	Accuracy (%)
FCN	2/2	0.2	75.7
ResNet-50	50/2	25.4	75.4
GoogLeNet	22/2	6.7	81.9

From Table 1, it is clear that even a relatively small network (FCN in this case) can give comparable accuracy to much more deeper sophisticated networks such as ResNet-50 and GoogLeNet. Further, the accuracies are similar to many of the architectures presented in the literature.

4 Conclusions and Future Directions

An experimental comparison of various networks is given for classification of images into violent and normal physical interaction. The study shows that for images with and without physical violence, small networks are also capable in performing a classification task for this 2-class problem. Further, since images are used for the classification, networks architectures can also be relatively simple. These networks can be used to extract features for larger networks, e.g., a 3D convolutional network or a LSTM network, in building true video classification architectures for real-time detection using embedded hardware.

References

1. LeCun Y et al (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
2. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst
3. Sultani W, Chen C, Shah M (2018) Real-world anomaly detection in surveillance videos. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 6479–6488
4. Xu D et al (2017) Detecting anomalous events in videos by learning deep representations of appearance and motion. Comput Vis Image Underst 156:117–127
5. Simonyan K, Zisserman A (2014) Two-stream convolutional networks for action recognition in videos. Adv Neural Inform Process Syst
6. Srivastava N, Mansimov E, Salakhudinov R (2015) Unsupervised learning of video representations using lstms. Int Conf Mach Learn
7. Satyanarayanan Mahadev (2017) The emergence of edge computing. Computer 50(1):30–39
8. Seyed YN, Chen Y, Song S, Xu R, Choi B-Y, Faughnan TR (2017) Smart surveillance as an edge network service: from Harr-Cascade, SVM to a Lightweight CNN, https://arxiv.org/pdf/1805.00331.pdf
9. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press

10. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
11. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich, A. (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9

Diagnosis of Epileptic Seizure a Neurological Disorder by Implementation of Discrete Wavelet Transform Using Electroencephalography

Sanjay Shamrao Pawar and Sangeeta Rajendra Chougule

Abstract Abnormality and presence of neurological brain disorder such as epileptic seizure is diagnosed by analyzing electroencephalography signals accurately. The acquired brain signals are analyzed in time–frequency domains by using wavelet for accurate diagnosis. The online standard EEG database signal is preprocessed to remove power noise and most important eye blink artifact using independent component analysis. Daubechies wavelet is implemented, and decomposition of frequency is carried out up to eight levels. The exact sub-band of frequencies are extracted from band of frequencies which are called as delta band, theta band, alpha band, beta band and gamma band from lower to higher. Suitable features such as Lacunarity, Fluctuation Index, Energy and Entropy, Kolmogorov Entropy, Kurtosis and Skewness are extracted and classified using K-Nearest Neighbor, Support Vector Machine and Probabilistic Neural Network. Performance analysis is carried out by measuring specificity, sensitivity, accuracy and true predictive value. Implementation of independent component analysis to the affected channels in preprocessing block to remove eye blink artifact, decomposition of brain signal using Daubechies wavelet up to eight levels provides accurate diagnosis of epileptic seizure. The proposed seizure detection system provides high metric of performance parameters.

Keywords Daubechies wavelet transform · Electroencephalography · Independent component analysis · Kolmogorov entropy · Lacunarity · Probabilistic Neural Network

S. S. Pawar (✉)
Shivaji University, Kolhapur, India
e-mail: sanjaypawar832@gmail.com

Bharati Vidyapeeth's College of Engineering, Kolhapur, India

S. R. Chougule
College of Engineering, Kolhapur Institutes of Technology, Kolhapur, India
e-mail: shivsangeeta.chougule@rediffmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_59

1 Introduction

The human brain function represents the status of whole body, and central nervous system (CNS) consists of neurons. The neuron transmits information on response to simulate, which is called as action potential, whose voltage is in between −60 and 10 mv, and the action potential remains for 5–10 Ms. An electroencephalogram is a measurement of brain signal using scalp electrodes, and the study of EEG waves is used in detection of brain disorders which are called as neurological disorders and also to find abnormalities present in human body. The EEG signals are used for detection of epilepsy neurological disorder and testing of given drug effect by clinical expert using visual inspection. The amplitude and frequency changes during abnormalities are difficult to detect with trained eyes also, so there is a need of advance signal processing system to accurately analyze electroencephalogram waveform and identify the seizures accurately. The EEG signals need to be acquired by using 10–20 international standard electrode placement system. The EEG signals are contaminated with internal and external noises which are called as artifacts, and the artifacts may be internal or external and need to be separated out. The required features are to be extracted out from the signal, so that it can be correctly classified the EEG signal as seizure signal and non-seizure signal by our classifier [1].

Detection and classification of epileptic seizures have attracted many researchers, and analysis of acquired EEG signal has moved from time or frequency domain to wavelet (time–frequency) domain. The EEG signal was analysis by wavelet using the features such as standard deviation, largest Lyapunov exponent and correlation dimension [2] and by packet wavelet transform by finding the feature as Combined Seizure Index (CSI) [3]. Short-Time Fourier Transform and power spectrum density were used to analysis signal in time–frequency domain [4]. Implementation of wavelet such as discrete wavelet transform and classifier ant colony was done for early predication of seizure [5]. Energy, variance and fractal dimension features were extracted, and Support Vector Machine is used for classification [6]. Detection of epileptic seizure is done by using Shannon Entropy and Logarithmic Energy Entropy and classified using Support Vector Machine classifier [7]. EEG artifacts were removed, and the features such as mean, variance, standard deviation, kurtosis and Skewness were classified by Support Vector Machine [8].

Accurate seizure detection, predication of pre-ictal stage and localization of seizure still remain to be challenge in EEG signal processing. Hence, our objective is to develop seizure detection system to diagnosis and analysis epilepsy accurately. The development of system will reduce healthcare costs and improve medical diagnosis treatment of patients. The features in system are chosen to achieve high metric of sensitivity, specificity, accuracy, positive predictive value and low false rate. Three classifiers are implemented, and analysis of EEG for seizure detection has been done to provide suitable classifier.

2 EEG Signals and Brain Rhythms

Acquiring brain signals for diagnosis of neurological disorder has become very vital. Electroencephalogram is a recording method of brain waves which consists of number of electrodes, differential amplifier and filters. The electrodes are placed on skull using placement system which is defined internationally and known as 10–20 system. To analyze the EEG signal, sampling and quantization are carried out, and then, the data is stored. The EEG brain wave signals are called as brain rhythms, which changes its amplitude and frequency on the occurrence of abnormality. There are five major brain wave categories with respect to their frequencies from lower to higher and are called as delta, theta, alpha, beta and gamma. The delta waves are from 0.5 to 4 Hz and are generated during deep sleep, and the theta waves are from 4 to 7.5 Hz and are generated during drowsiness and meditation. Alpha waves are from 8 to 13 Hz and are generated during general awareness and without concentration and are round sinusoidal in shape. Beta waves are from 14 to 30 Hz and are generated during walking, thinking and solving problems. The frequencies above 30 Hz and up to 60 Hz are named as gamma waves which are lower in amplitude and are very rarely used for investigation of neurological disorder [1, 9].

The abnormal EEG patterns are used in identification and classification of many neurological diseases out of which one is epileptic seizures. There is a sudden change in frequency due to excessive electrical discharge from brain cells during the onset of epilepsy. The electroencephalogram brain signal measurement technique remains to be effective method to diagnosis and analyze the epileptic seizure [8–10].

3 Methodology

Figure 1 shows seizure detection system, and the EEG signals necessary for our analysis are taken from standard online database, which are recorded and stored using 10–20 international electrode placement system. EEG signals are further preprocessed to remove power noise, artifacts and required band is extracted. Daubechies discrete wavelet transform with eight level of decomposition is implemented on the extracted band of EEG signals. The sub-band frequencies from lower to higher are delta, theta, alpha, beta and gamma waves. Further features such as Lacunarity, Fluctuation Index, Energy and Entropy, Kolmogorov Entropy, Kurtosis and Skewness are extracted from alpha and beta waves which are the band of frequency of our interest. These features vectors are further provided to classifier, and the classifier categorizes the EEG signal as seizure EEG signal and non-seizure EEG signal.

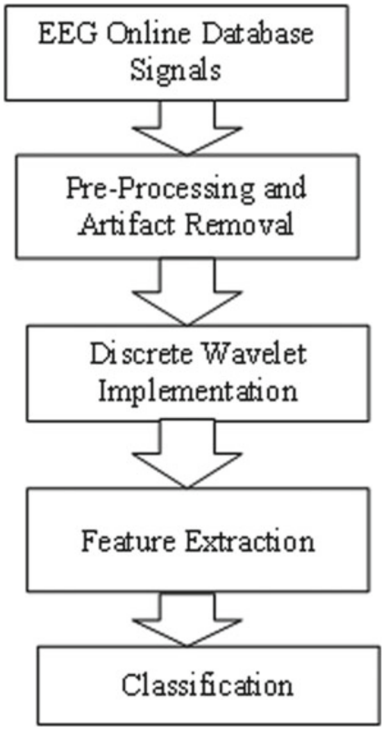


Fig. 1 Seizure detection system

3.1 EEG Database

The online database of EEG signals is collected from children's hospital Boston. The children's in hospital are monitored for several days and hours for their surgical intervention. Electroencephalogram 10–20 international electrode placement system is used to record and store the brain signal. The signals are collected using 23 electrodes, sampling rate of 256 samples per second is used and the EEG signal data is represented with 16 bits of resolution. The data is stored in file whose extension is in European data file (.edf). There are total 23 subjects patients data available out of which six subject's patient's data is analyzed in our experiment. There are two male and four female subjects in the age range of 1.5–22 years, and total 152 EEG signals are used in our experiments [11, 12].

3.2 Preprocessing and Artifact Removal

The EEG signals during recording are mixed with external and internal noises, which are termed as artifacts, and these artifacts affect the performance of seizure detection system. The external artifacts which are not originated from brain, such as signals from mobile phone, human voice, noise generated by movement of tables and chairs or reflection of subject itself during EEG recording. The external artifact can be avoided or minimized by taking precautionary measures during recording of EEG. The neuronal information interest frequency is from 3 to 60 Hz, so the interested band is extracted by implementation of band-pass filters. The power noise of line frequency 50 Hz is removed by implementation of notch filter. The eye blink artifact is the major artifact present in electroencephalographic, which is also called as electrooculograms (EOG). A movement of subject's eyeballs during recording of EEG signal causes a change in the electrical potential, which are picked up by electrodes placed near to it, in our case FP1 and FP2 electrodes are the source of eye blink artifact. The human eyes are dipole in nature which is positively charged from front side and negatively charged from behind. The movement caused by eyeballs or eyelids is easy to detect because of their typical shape, and many algorithms are developed to find and remove eye blink artifacts automatically from acquired EEG signal [13, 14].

The introduced eye blink artifact is removed by implementation of independent component analysis (ICA) to the specified channels. In independent component analysis (ICA), the signals are decomposed into their constituent components, and blind source separation (BSS) method is used for separation of signals. ICA separates independent source signal using the information mixture during recording. ICA is used for separating and removing noise from the EEG signals. ICA is implemented to only those EEG signals which are acquired by the electrodes, whose placements are done near eye and are the source of eye blink artifact. The EEG signal of channel 1, channel 5, channel 9 and channel 13 are passed through ICA as the electrodes FP1 and FP2 are used in plotting these channels. The electrodes FP1 and FP2 placements are near the eyes and can be the source of eye blink artifacts [15].

Figure 2 shows the original EEG signal of one of the channel and the filtered EEG signal after removal of power noise by implementation of notch filter.

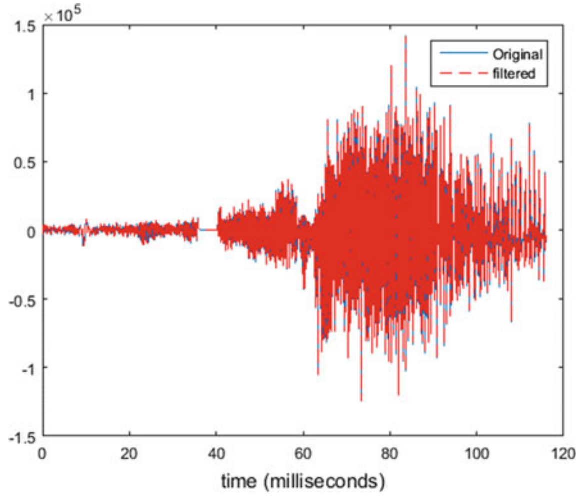


Fig. 2 Filtered EEG signal

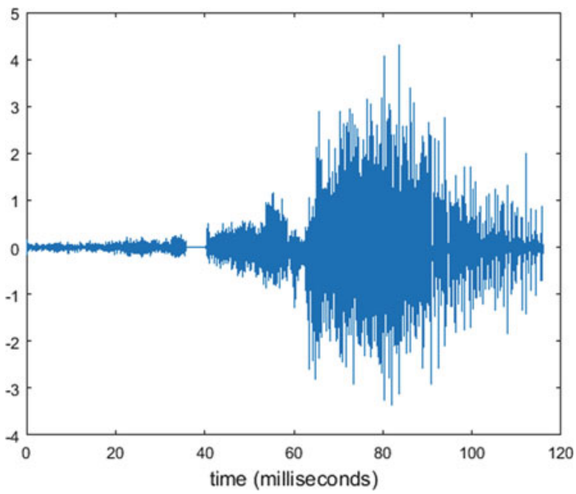


Fig. 3 EEG signal after implementation of ICA

Figure 3 shows the EEG signal of one of the channel when eye blink artifact is removed, and only those channel signals are passed through independent component analysis (ICA) who have consider FP1 and FP2 for plotting signals as these electrodes are placed nearer to eye and are source of eye blink artifact.

3.3 Wavelet Implementation

EEG signals are the non-stationary signals, and the regular time or frequency domain were used to analyze the signal. Compared to the traditional method, wavelet transform (time and frequency analysis) is a powerful and most suitable method for analysis of transient signals. Wavelet transform is the mathematical tool which is used to analyze EEG signal. Wavelet is a waveform which has average value as zero, analysis of signal in wavelet domain means breaking of signal in scaling and shifting. Wavelet transform methods are used to localize the signal component better than time–frequency space.

Daubechies wavelet is found to be most suitable wavelet for analysis of epileptic EEG data. Daubechies discrete wavelet transform with eight level of decomposition has been implemented to separate out delta, theta, alpha, beta and gamma waves in EEG signal. The signal waves which are important for our study and analysis are alpha and beta range frequencies as these waves change its characteristics during occurrence of epileptic seizure and are hence used to derive the feature vector. The implementation of eight-level discrete wavelet transform gives the approximation coefficients and detail coefficients. Figure 4 represents eight levels of decomposition of discrete wavelet transform. The approximation coefficient is CA8, and detail coefficient are named as CD8, CD7, CD6, CD5, CD4, CD3, CD2 and CD1. The

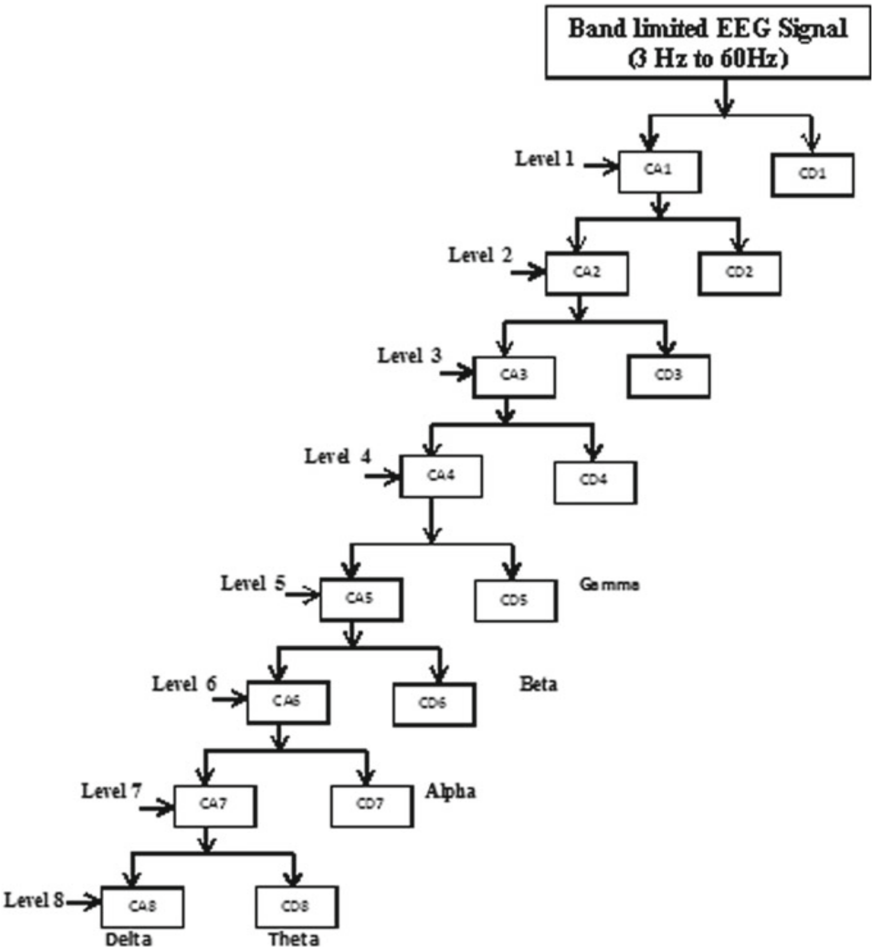


Fig. 4 Discrete wavelet transform decomposition up to eight levels

seizure usually occurs in the frequency band range of 8–30 Hz. These bands of frequency are named as alpha and beta and are important for analysis. The detail coefficients, CD7 denotes alpha frequency band and CD6 denotes beta frequency band, which are further used to derive feature vectors.

3.4 Feature Extraction

Selecting effective features is very important in seizure detection system, and best features need to be selected which represents the characteristics of EEG signals. Feature vector found out should be able to differentiate clearly the given EEG as seizure and non-seizure EEG data. Feature extraction in EEG signal provides more precise and necessary information, which is suitable for further EEG classification. Feature extraction consists of finding hidden information with an objective to find the presence of events in given EEG signal data. These features vectors are important parameters for detection and classification of given biomedical EEG signal processing, it is very important steps in analysis of data [15]. The variations in feature vector distribution of the EEG signals are measured in terms of parameters of Gaussian and their mapped deviation of distribution from standard Gaussian. As alpha and beta frequencies are present in detail coefficients CD7 and CD6, these coefficients are used to extract features. The deviation in Gaussian of the EEG signals can be found out by measuring the values of Lacunarity, Fluctuation Index, Energy and Entropy, Kolmogorov Entropy, Kurtosis and Skewness.

Lacunarity. Lacunarity is a measurement of heterogeneity. A low value of Lacunarity indicates the presence of homogeneity and transitionally invariant in EEG signal, and high value of Lacunarity indicates heterogeneity and not transitionally invariant in EEG signal [18].

Fluctuation Index. Fluctuation Index is the measurement of fluctuation intensity present in EEG signal, and Fluctuation Index of the measured EEG signal during seizures period is usually higher than that of non-seizure periods

Energy and Entropy. Energy of a given signal is measured in terms of sum of squared modulus of the sample values [6]. It is also represented as strength of measured EEG signal. High value of energy indicates the presence of seizure activity in given signal. Entropy is measured as the presence of randomness or uncertainty in the signal, it also reflects the quantity of event in the measured EEG signal and Shannon Entropy is found out as a feature vector [7].

Kolmogorov Entropy. Kolmogorov Entropy is a statistical property which is used to measure irregularity of EEG signal. Kolmogorov Entropy is the entropy of source distribution generating the sequence.

Kurtosis. Kurtosis is the measurement of whether the data is peaked or flat in structure at the mean point, which is related to normal distribution. The data set having high value of kurtosis indicates that there is a presence of peak near the mean point and which declines very rapidly with heavy tails. Data set having low value of kurtosis indicates that there is flat top near the mean point of the signal.

Skewness. Skewness is a measurement of lack of symmetry or presence of symmetry in given data set distribution. Symmetric distribution looks similar to the left as well as to right side from the center point. If the distribution of data set is present more toward right side from the mean point, then the Skewness is termed negative Skewness, and if the distribution of data set is present more toward left side from the mean point, then the Skewness is termed positive Skewness. Skewness is found as zero for symmetric distribution of data.

3.5 Classification

The main objective of classification is to separate out the feature vectors in two or more classes by setting or drawing boundary between classes. Classifier is very essential part for accurate seizure detection system. There are a number of classification techniques developed and implemented in past, and few classification techniques which are mentioned as Support Vector Machines, K-Nearest Neighbor and Probabilistic Neural Network are found to be most accurate and widely used. The above three classifiers are implemented and compared to find the most suitable classifier for seizure detection system. Support Vector Machine (SVM) is a categorized as supervised machine learning algorithm. The EEG classification is done by placing of hyperplanes in a multidimensional space which separates data of different class labels. In two-dimensional data as in our case of EEG signal, a single hyperplane is implemented with largest margin to separate out the data into two groups such as seizure and non-seizure [16–18]. The second classifier is an effective clustering tool which is K-Nearest Neighbor algorithm, and it is an instance-based learning method. K-Nearest Neighbor algorithm uses Euclidean distance metrics function to measure similarity and minimum distance to classify the EEG data. The algorithm divides set of feature vectors in k different clusters [19]. The third classifier is again a type of supervised neural network which is Probabilistic Neural Network (PNN), and it is implemented for classification of EEG data. PNN belongs to the type of Artificial Neural Network and is derived from Bayesian network. PNN has total four layers, first layer is input layer, second layer is pattern layer, third layer is summation layer and the final fourth layer is named as output layer. Advantage of PNN is that it is much faster than above two classifiers but consumes more memory space.

4 Performance Evaluation

The performance evaluation of proposed automatic seizure detection systems is carried out by using the EEG signal acquired from online database of six patients, namely subject1, subject 2, subject 3, subject 4, subject 5 and subject 6. According to gender specification, there are two male and four females with subject age ranges from 1.5 to 22 years, respectively. The EEG data set used in our experiment is collected by 10–20 international standard electrode placement system using 23 electrodes. The signals are plotted using differential bipolar montages to map a signal. A total of 152 EEG signal records with combination of seizures and non-seizures are used for analysis of above six patients. Total 118 record signals are of type non-seizure and 34 record signals are of type seizure. A training set of 51 signals is created for all the three classifiers by using 11 seizure signals and 40 non-seizure signal types. A totally different testing set of EEG signal is created and tested with all three classifiers of 101 EEG signals. Testing set consists of 23 seizure-type signals and 78 non-seizure-type signals as their ground truth value. The 23 differential signals in each file are preprocessed by removing 50 Hz of power noise by implementation of notch filter. An eye blink artifact is removed by implementation of independent component analysis, which is blind source separation method only to those EEG signals where FP1 and FP2 electrodes are considered during mapping of differential signal. Further, the band of interest where seizure may occur is extracted by implementation of band-pass filter which is from 3 to 60 Hz. Discrete wavelet transform with eight level of decomposition is implemented, and various range of frequencies are extracted, alpha and beta band of frequencies which are of our interest where the seizures occur normally. These bands are taken, and the statistical features such as Lacunarity, Fluctuation Index, Energy and Entropy, Kolmogorov Entropy, Kurtosis and Skewness are extracted and provided as input to classifier Support Vector Machines, K-Nearest Neighbor algorithm and Probability Neural Network to classify and find out whether a subject has a seizure or not.

5 Result and Discussion

Proposed seizure detector system performance is characterized or measured in terms of parameters such as specificity, sensitivity, accuracy and true predictive value. The specificity, sensitivity, accuracy and true predictive value of the above three implemented classifiers are calculated and are compared with each other to suggest the best-suited classifier for detecting seizure in EEG signal.

5.1 Sensitivity

It is the ratio of true positives seizures to the total number of seizures labeled by the EEG clinical experts. True positive indicates detected seizure signal by our system which was also identified as seizure by the EEG clinical experts [20].

$$\mathrm{Sensitivity} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$$

5.2 Specificity

It is the ratio of true negatives seizures to the total number of non-seizure labeled by the EEG clinical experts. True negative indicates a signal labeled as non-seizure both by our system well as by the EEG clinical experts [20].

$$\mathrm{Specificity} = \mathrm{TN}/(\mathrm{TN} + \mathrm{FP})$$

5.3 Accuracy

It is the ratio of number of correctly identified seizure and non-seizure signals to the total number of signals.

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

5.4 True Predictive Value

It is the ratio of number of seizure identified by our system as well as labeled by EEG clinical experts.

$$\text{True Predictive Value} = \text{TP}/(\text{TP} + \text{FP})$$

where true positive (TP) is a number of seizures detected by our proposed system and also identified by clinical expert. False negative (FN) is a number of seizures missed by our proposed system but are identified and marked by clinical expert. True negative (TN) is a number of non-seizures identified by our proposed system and also identified and marked by clinical expert, and finally, false positive (FP) is defined as non-seizure signal detected as seizure by our proposed system.

Table 1 shows the comparison between three classifiers used in our system for same EEG data set in terms of specificity, sensitivity, accuracy and true predictive value. With the above features and preprocessing of EEG signal 78 non-seizure and 23 seizure signals, total 101 EEG signals were used for testing. Support Vector Machine classifier classifies 21 true positive signals, 78 true negative signals, 0 false positive signal and 2 false negative signals. K-Nearest Neighbor algorithm classifier classifies 20 true positive signals, 78 true negative signals, 0 false positive signal

Table 1 Comparison between SVM, K-NN and PNN in terms of sensitivity, specificity, accuracy and true predictive value

Classifier	Sensitivity in %	Specificity in %	Accuracy in %	True predictive value in %
Support vector machines	91.30	100	98.02	100
K-Nearest neighbor algorithm	86.96	100	97.03	100
Probability neural network	73.91	100	94.06	100

and 3 false negative signals. Probability Neural Network classifier classifies 17 true positive signals, 78 true negative signals, 0 false positive signal and 6 false negative signals.

6 Conclusion

Automatic seizure detection system is significant for diagnosis of epileptic seizure in prolonged recording of EEG. In this paper, we have proposed automatic seizure detection system for accurate seizure diagnosis. EEG signal is preprocessed, and eye blink artifact is removed by ICA. EEG signal is analyzed in both time and frequency domain by implementation of eight-level discrete wavelet transform. Statistical features such as Lacunarity, Fluctuation index, Energy and Entropy, Kolmogorov Entropy, Kurtosis and Skewness are extracted from alpha and beta waves and are provided as input to classifier Support Vector Machines, K-Nearest Neighbor and Probabilistic Neural Network to identify and classify as seizures and non-seizure. Performance analysis of proposed system is carried out in terms of specificity, sensitivity, accuracy and true predictive value. Experimental results of online standard EEG data sets show that our proposed system achieves a highest accuracy of 98.02%, specificity of 100%, sensitivity of 91.30% and true positive value of 100% for Support Vector Machines and lowest accuracy of 94.06%, specificity of 100%, sensitivity of 73.91% and true positive value of 100% for Probabilistic Neural Network. So the implementation of ICA in preprocessing block to remove eye blink artifact and implementation eight level of decomposition in discrete wavelet transform with Support Vector Machine as classifier in seizure detection system provides accurate diagnosis of epileptic seizure with high metric of performance parameters.

In the future for still better accuracy in detection of epileptic seizure and to find preictal stage before seizure occurrence, intracranial electroencephalography (IEEG) signals can be tested and analyzed with our seizure detection system.

References

1. Pawar S, Chougule SR, Tirmare AH (2017) Diagnosis of epilepsy a neurological disorder using electroencephalogram (EEG). Int J Mod Trends Eng Res 4:144–149. https://doi.org/10.21884/ijmter.2017.4195.u8cbh
2. Ghosh-Dastidar S, Adeli H, Dadmehr N (2007) Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Trans Biomed Eng 54:1545–1551
3. Zandi AS, Javidan M, Dumont GA, Tafreshi R (2010) Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform. IEEE Trans Biomed Eng 57:1639–1651
4. Tzallas AT, Tsipouras MG, Fotiadis DI (2009) Epileptic seizure detection in eegs using time-frequency analysis. IEEE Trans Inf Technol Biomed 13:703–707

5. Salem O, Naseem A, Mehaoua A (2014) Epileptic seizure detection from EEG signal using discrete wavelet transform and ant colony classifier. In: IEEE ICC 2014 selected areas in communication symposium, pp 3530–3534
6. Kumar A, Kolekar MH (2014) Machine learning approach for epileptic seizure detection using wavelet analysis of eeg signals. In: International conference on medical imaging, m-health and emerging communication systems (MedCom), pp 412–416
7. Vasudha H, Puntambekar VP, Raviteja K, Mahalakshmi P (2018) Detection of epileptic seizure using wavelet analysis based shannon entropy and logarithmic energy entropy and support vector machine. Int J Eng Technol 935–939
8. Pawar SS, Chougule SR (2019) Diagnosis and analysis of epileptic seizure neurological disorder using electroencephalography. IOSR J VLSI Signal Process (IOSR-JVSP). 9:23–27. https://doi.org/10.9790/4200-0901012327
9. Pawar SS, Suryawanshi N (2017) Epilepsy seizure detection using wavelet based by artifact reduction. Int J Innovations Eng Res Technol 4:11–18
10. Pawar SS, Narayankar KG (2017) Tsallis entropy based seizure detection. In: Int J Innovations Eng Res Technol 4:86–92
11. CHB-MIT scalp EEG database [Online]. http://physionet.org/physiobank/database/chbmit/
12. Shoeb A (2009) Application of machine learning to epileptic seizure onset detection and treatment. Ph.D. Thesis, Massachusetts Institute of Technology
13. Savelainen A (2010) An introduction to EEG artifacts. Independent Res Projects Appl Math
14. Balbir S, Wagatsuma H (2017) A removal of eye movement and blink artifacts from EEG data using morphological component analysis. Comput Math Methods Med
15. Pawar SS, Chougule SR (2019) Diagnosis and classification of epileptic seizure a neurological disorder using electroencephalography. In: 2nd International conference on intelligent computing, instrumentation and control technologies (ICICICT), pp 973–978
16. Liu Y, Zhou W, Yuan Q, Chen S (2012) Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. IEEE Trans Neural Syst Rehabil Eng vol 20
17. Bhuvaneswari P, Kumar JS (2013) Support vector machine technique for EEG signals. Int J Comput Appl 63
18. Zhou W, Liu Y, Yuan Q, Li X (2013) Epileptic seizure detection using lacunarity and bayesian linear discriminant analysis in intracranial EEG. IEEE Trans Biomed Eng 60:3375–3381
19. Sanei S, Chambers JA (2007) EEG signal processing. centre of digital signal processing cardiff University. Wiley, UK
20. Yan A, Weidong Z, Yuan Q, Yuan S, Wu Q, Zhao X, Wang J (2015) Automatic seizure detection using Stockwell transform and boosting algorithm for long-term EEG. In: Epilepsy & Behavior

Detection and Recognition of License Plate Using CNN and LSTM

Anaya Anson and Tessy Mathew

Abstract Due to the world's rapid economic growth, the motor vehicles' number has significantly increased which lead to the necessity for the security of vehicles. It becomes difficult to manually track the vehicles which are over speeding and to monitor the traffic so as to avoid congestion problems, tracking stolen cars, detecting the unlawful activities that involve a vehicle. Automatic number plate recognition (ANPR) is an image processing technology which allows users to track, identify, and monitor moving vehicles by automatically detecting the license plate from the vehicle. A convolutional neural network (CNN) is used which slides over the entire image so as to detect the characters from an image. After detecting the characters, the non-plate region must be eliminated from plate region. For that a plate/non-plate CNN classifier is used. After detecting the license plate, LSTM is being used for obtaining the feature sequence. Then the characters are decoded, and the license plate is recognized. By combining both CNN and LSTM methods, it gave an accuracy of 85.0%.

Keywords Convolutional neural network · Long short-term memory · Recurrent neural network · Bidirectional recurrent neural network · Connectionist temporal classification

1 Introduction

In developing countries, there is a huge rise in vehicles every day. Due to the unlimited rise of vehicles, it has become difficult to manually track and monitor the vehicles. Automatic number plate recognition (ANPR) [1] will solve this issue as it can be used for the identification of vehicles, controlling theft, providing security to vehicles,

A. Anson (✉) · T. Mathew
Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, India
e-mail: anayaanson0919@gmail.com

T. Mathew
e-mail: tessmathew@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_60

checking the status of vehicles, collecting toll amount automatically, controlling the traffic flow, tracking the vehicles which are over speeding and finding the vehicles which had involved in criminal and unlawful activities. License plate recognition (LPR) automatically reads the license plate of vehicles.

While automatically reading the number plate, the pixels of image are converted to ASCII text. The challenge lies in capturing an image of a fast-moving vehicle as there is a chance to get a blurred image which can reduce the accuracy when recognizing the number plate. In plate localization, it locates the license plate on the vehicle. The dimensions were adjusted to the required size using plate orientation and sizing. Normalization was done which enhances both the contrast and brightness of an image. After license plate has been detected from the image, the characters are then segmented into individual characters and then the license plate is recognized. The characters and positions are checked against country-specific rules. ALPR primarily consists of three steps which are detecting license plate, segmenting each character, and recognizing the characters.

This work provides the following contributions:

In license plate detection and recognition (LPDR) system, the license plates are viewed as strings of text. To detect the presence of characters, a character CNN must be used which slides over the entire image. After detecting the characters, a plate or non-plate CNN classifier was used which eliminate the non-plate regions from the license plate. After detecting the license plate, the characters were recognized. Long short-term memory (LSTM) was used for recognizing the feature sequence [2]. LSTM will give some probability values, and these sequences of probability values must be converted to character string. Connectionist temporal classification (CTC) was used which performs the decoding of character string in the plate [3].

Automatic license plate recognition (ALPR) was used to read the license plate from vehicles. The vehicle image is obtained from a camera and can be used by the police officers for finding the unlawful activities involving vehicles around the world and can also be used for checking whether the vehicle has been registered or not registered.

Automated license plate recognition works for the following scenarios:

- Finding the stolen cars. A list of stolen cars will be prepared, the LPR will compare the passing cars with the list, and a siren will be activated when it finds a match.
- It can be used while parking a car in a busy shopping center. When the car enters the parking area, the license plate of the vehicle will be recognized and is stored. When the car exits the parking area, the license plate will be read and a charge will be issued to the driver according to the duration of parking.
- Another application is access control as an authorized vehicle will be permitted to pass the gate after recognizing the number plate.
- It can be used to track over speeding vehicles. The over speeding vehicles can be found out by finding out the time, and it will take to move from one camera to the next camera.

- It can be used in tolling places. The license plate will be read when it enters the toll lane and a pass card will be presented and will be compared with the pass information.

To detect an object from an image, its general characteristics must be understood and must find the difference between that object and the other objects in that image. For detecting the license plate from a car, its region of interest (ROI) must be found.

The general characteristics of a number plate are:

- The number plates have a higher contrast which makes them to recognize easily.
- The number plates have a standard size same aspect ratio in each country, and its shape is a horizontal rectangle.
- The number plate is usually located at the bottom part and middle position of the car.

The ROI can be obtained by following these characteristics. After obtaining the number plate, the characters in that number plate must be recognized.

The detection of license plate may have to face many environmental challenges like:

(1) *Background*: The background color may be the same as the number plate, and also, it may contain the same patterns like the plate, and it will get baffled with the number plate.
(2) *Interference of objects*: Some interfering objects may occlude the number plate.
(3) *Illumination*: The weather conditions may affect the illumination.
(4) *Shadows*: Objects may cause shadow which affects the detection process.

2 Literature Review

To recognize a license plate from an input image, it needs to perform mainly three steps. This section gives a brief description of the works done by various researchers.

2.1 License Plate Detection

License plate detection will determine the location of license plate, and a bounding box is generated over the detected license plate.

Anagnostopoulos et al. [4] proposed that the region of interest (ROI) can be obtained by sliding a window for detecting the license plate. The disadvantage of this method was that it was not possible to detect the plate when it is dirty or is damaged. There were also illumination problems: During daytime, the illumination changes slowly and can even give incorrect results due to change in climatic conditions or of passing objects like clouds. The next problem is due to the similarity of characters, such as 2 and Z, 1 and I, 5 and S, 0 and O, 8 and B.

Many pre-treatment steps [5] including conversion of RGB to HSV image, morphology filtering, Gaussian blur filter, adaptive threshold; geometric filtering must be performed before applying convolutional neural network (CNN) model, and a bounding box will be formed over the detected license plate.

Background extraction and preprocessing [6] were done to minimize the noise. The detection of license plate from moving vehicles is done by a template matching procedure which learns automatically and will adjust itself to the plate size. License plate was detected by using canny edge detection operator from the vehicle image [7].

A CNN classifier [8] was being used which will slide over the entire image, and the text-like regions are detected. The license plate must be distinguished from other text regions. For that, a plate or a non-plate CNN classifier is used.

A vehicle image is first obtained. Then the localization of license plate is done using 'feature-based license plate localization' method [9]. All background noise was eliminated, and license plate area was preserved.

The detection of license plate first includes preprocessing which includes plate detection, correction, and segmentation. For shadow removal, Bernsen algorithm is used [10].

License plate image of a vehicle is being captured and that image converted to gray image. To highlight the white area in the image, the gray conversion is done. The unwanted noise is being removed by using median filter which makes the background portion of the image much clearer [11].

Saunshi et al. [12] proposed that an image of a car is captured. Preprocessing is performed to enhance the image quality. The image is converted from RGB format to gray scale as to reduce the number of colors. Image noises can be found which are caused due to fault in camera or due to changing weather conditions. To remove noise, bilateral filter is being used. Binarization is performed to restrict the color shades to two colors. The license plate was being extracted by using Sobel's edge detection algorithm or smearing algorithm.

Preprocessing was performed to enhance the input image [13]. License plate localization is done in two steps which are placing a bounded rectangle on the license plate and determining the exact location of the license plate.

The vehicle image captured is changed to a grayscale image. The candidate regions are been detected by using sliding concentric windows (SCW). Filtering and verification of color for candidate region are done by using HSI color model. The license plates which are tilted must be corrected by using the least squares method [14].

Sparse auto-encoder (SAE) [15] can be used for detecting the license plate. Images are taken from a camera, or videos can be obtained, these images are converted into grayscale images. Hough transform [16] is performed to determine the Hough lines, and the input images are fed to the edge detector for detecting the edges. The reason for detecting the edges is to reduce the amount of information, and it filters out the information which is not wanted and maintains the significant information.

The first step is to obtain the vehicle image using camera. The images are obtained in RGB format. The next is the extraction of license plate. This method is being used in Sindh which has the number plate with yellow background with black characters. For that, an algorithm called yellow search is used to search for yellow color pixels. The pixel having a yellow color is set to 1, and the pixel having no yellow color is set to 0. Thus, the image obtained is in white-and-black format. After identifying the ROI, an image filtering is done. The number plate is extracted using smearing algorithm [17].

A Hough transform method is used for detecting the license plate. Hough transform [18] is performed in such a way that the straight lines are detected. The disadvantage was that it consumed time and memory.

Edge-based approaches can be performed for detecting the license plates. The license plates are rectangular in shape with the same aspect ratio. The license plate boundary can be found out by applying color transition.

Sobel filter [19] was been used for detecting the edges, and the license plate can be found out in high-density regions. The candidate regions can be performed by applying horizontal and vertical edge detections. This method mostly uses vertical edge detection; as while using horizontal edge detection, it gave inaccurate results; as it was a little bit puzzled with the car bumper edges.

A color-based approach [20] which detects the license plates using color. The license plate color varies from the background color. In this method, a genetic algorithm is used which will detect those color regions. This method was trained in different conditions and will learn to detect the colors that it should look for on its own.

A method which uses both local Haar-like features and global statistical features [21] was performed. Most of background regions were eliminated by global features, and the local Haar-like features will help to get accurate results with change in conditions.

Matas et al. [22] used a region-based approach. The character-like regions are found out by employing a neural network which classifies the regions in the image. The presence of combinations of character regions will detect the license plate.

Li et al. [23] method will detect the number plate by using a 37-class convolutional neural network (CNN), and the false positives are eliminated by using another CNN. This method is comparatively slow moving, takes 5 s per image, and however, removes the necessity of character segmentation.

A technique is used in which an image saliency is employed [24]. The initial tactic detects the characters solely and then segments the characters based on intensity saliency map. After that, a sliding window is moved on the segmented characters which detect the license plate.

2.2 Character Segmentation

The characters which are present on a license plate will be mapped, and they are segmented out into each image. There are many techniques which can be used and can segment each character after detecting the license plate.

Anagnostopoulos et al. [4] proposed a segmentation technique called sliding concentric windows—SCW. The connected component analysis method is used for obtaining the license plate position and to perform character segmentation.

The segmentation steps include the conversion of RGB image to grayscale image which is done [5]. Then the image is enhanced to get more clarity. Image edge detection is applied. Then contour extraction technique is performed, and boundary boxes of characters are detected.

The segmentation of characters involves contrast extension which makes the image sharpened followed by median filter which was used for removing the unwanted areas which have noise and blob coloring method was performed to determine the character boundaries [7].

Pre-segmentation is not to be used to extract the features [8].

The detected number plate can be segmented into single characters using an 'image scissoring' [9] algorithm. This method is being restricted to some factors like scripts on the license plate, vehicle speed, skew in the image.

During segmentation stage, image gray equilibrium is an important step. For image enhancement, gray statistical approach (GSA) [10] was been used. After the license plate was segmented correctly, character recognition step is performed. Feature extraction is performed by the elastic mesh. This method cannot identify properly when the plate is damaged or contains dirt. This method is suitable for the black characters which are on a white plate, but does not fit for white characters which are on a black plate.

Multi-thresholding [11] is used for image segmentation to distinguish the object from background pixels.

Badr et al. [13] proposed that the characters are segmented by two steps. The first step converts the license plate image to a binary image. Then noise is removed. After removing the noise, filtering is applied such that, no two components are merged. The false characters can be detected and can be rejected.

The candidates are decomposed by position histogram [14].

A block-based image segmentation technique [15] is performed which divides the whole image into several individual blocks. A deep learning model (SAE) was trained for selecting candidate blocks. Finally, an exact extraction of the license plate is done by using block merging and block selection.

Prabhakar et al. [16] proposed that the step after detecting the license plate is segmenting of characters. The character segmentation must be performed in an appropriate manner as in some cases the characters may have been divided improperly or in some cases the two characters may be improperly unified. Adaptive thresholding filter is used before the segmentation phase to distinguish the dark foreground from light background.

Projection-based approaches find the beginning and end locations of characters by applying a vertical projection over the binary image [25]. After that, the characters are segmented by applying a horizontal projection over that image.

A pixel connectivity-based approach [26] is used which labels all the connected pixels on the binary image to perform segmentation. Those pixels which have the same characteristics are grouped into the characters. The disadvantage of this method was that if the characters are broken apart or overlapping of characters exists then it will be impossible to segment the characters.

2.3 License Plate Recognition

License plate recognition firstly segments the characters, and then optical character recognition (OCR) will be performed to recognize each segmented characters. The characters which were segmented earlier were identified.

Anagnostopoulos et al. [4] proposed that the segmented characters were recognized using optical character recognition when it is forwarded to artificial neural network.

The characters are recognized by using a 37-class CNN model consisting of 26 uppercase characters and 10 classes of digits, and non-character classes were also provided [5].

A dynamic displacement method [6] is used to track the detected plate, and the nearest neighbor classifier is used to recognize the characters. The disadvantage was that the recognition of the plate characters becomes difficult in low-quality videos or in bad lighting conditions. The other problem is that the license plate in each frame must be compared with its previous recognized result and must check whether a recognition error is found and if the present must be corrected.

For recognizing the characters, two separate ANNs [7] were used to avoid the complexities that arise when detecting the similar characters in numbers and letters. The disadvantage of this method was that the recognition of characters from the license plate does not give accurate results when the illumination conditions are not good.

The feature sequence was recognized by using bidirectional recurrent neural networks (BRNNs) and long short-term memory (LSTM) [8]. The recognized characters are decoded using connectionist temporal classification (CTC).

Kulkarni et al. [9] proposed that after performing the character segmentation step, the characters will be extracted from an image, and the character enhancement technique is applied to enhance each character. These enhanced characters were then moved to a character recognition module, which recognizes the characters based on feature extraction.

The recognition of characters is performed by artificial neural networks [11].

The characters are recognized by convolutional neural network [12]. This system showed some errors while recognizing the characters from the license plate. The characters were recognized by using feedforward neural network [13]. This system gave correct results for clearer images, but showed incorrect results for blurry and skewed snapshots.

The characters are recognized by using artificial neural network [14]. This method was tested for various points of view, illumination conditions, and in various distances between camera and vehicle, and it worked well for these conditions. But it was restricted to images which are blurred which gave false values.

A deep learning model is used which learns the features from the license plate [15].

Prabhakar et al. [16] proposed that for recognition of characters: calculate the intensity score for each character and then compare with the model matrix, and the one with best score will be chosen as the recognized character. This method does not give accurate results in complex cases and for minor rotation and skew.

Optical character recognition (OCR) [17] algorithm was performed which was used for character recognition. In this method, during the recognition of the license plate, the OCR method provided inaccurate results during the case of misalignment and when it is of different sizes.

A template matching approach [18] in which it checks for the similarity of characters, and the most similar ones are chosen as the template. For each font, it is necessary to have a template. The disadvantage of this method was that it does not work well if the characters are broken. The processing time is much high as it needs to take each pixel and has to compare it with every template.

A learning-based approach [27] which distinguishes each character is based on unique features. It is done based on image density. It uses CNN which will learn the unique features of each character on its own.

Li et al. [23] used a pre-trained 9-layer CNN model. The recognition process was faster as it needs to process only fewer features.

3 Methodology

3.1 Convolutional Neural Network

Convolutional neural networks are artificial neural networks which are used for performing classification of images. The images are grouped together based on their similarity, and then with that it will perform the object recognition. It consists of many convolutional layers which are used to detect features such as lines, edges, and color drops. The early layers in CNN will detect the edges in an image. Then in the next layer, it will detect the patterns based on the previous layer. As it moves deeper into the layers, it may detect more complex features of the image. After it has learned all the characteristics of an image, it will be able to recognize it at any point of time.

The convolutional layers are operating based on 3D tensors which are called as feature map. It consists of height, width, and channel axis called depth. In an RGB image, it will contain channel axis as 3 because it uses three channels which are red, green, and blue, while in a black-and-white image, the channel axis is 1. Convolutional neural network consists of the following parameters:

(1) *Filter*: Filter is also known as feature detector or kernel. It is a matrix which when multiplied element-wise with the input image will give a feature map. The main function of the filter is that it lowers the dimension of the image and thus makes the processing speed faster. While reducing the dimensions of the input image, it will not eliminate the significant information as these unique features are used for identifying an object.
(2) *Pooling*: Pooling is used to detect the object even if there is a change in lightning or angle deviation in an image. The most commonly used pooling is max pooling. In max pooling, when a filter is placed on the feature map, it will take the largest value from it and thus keeping all the significant information's of the image. Max pooling aims to reduce the dimension of the image while maintaining the necessary information. It also helps to eliminate over-fitting problem which is mainly caused while training a CNN with more information.
(3) *Flattening*: After performing the pooling operation, flattening is performed. Flattening will transform the pooled feature map into a single dimension column matrix.
(4) *Full connection*: After performing flattening operation, it is passed to the fully connected layer to perform processing. Fully connected layer is the same as the hidden layer which mainly acts as classifier. The output layer outputs many predicted classes. If an error occurs during prediction, it will backpropagate that error through the whole network for improving the prediction.
(5) *Softmax Function*: The output obtained from the fully connected layers is given to the softmax function, where it will give the probability value of each class.

3.2 Recurrent Neural Network

When a human is thinking, he won't start from the beginning. When one person is reading a book, he would understand each word based on the previous words. This means that he would keep the previous word in his memory without throwing them away. This is similar to an RNN. RNN has a memory, and it moves like a loop. When predicting the next word which occurs in a sentence, it will take the current input and the previous output.

The neural network N will take the current input and previous output and will give the output in Fig. 1.

In the unfolding of recurrent neural network in Fig. 2.

It is the input at the time step t.
It is the hidden state at the time step t.

It acts as the 'memory' of the network. It is calculated by using the previous output and current input,

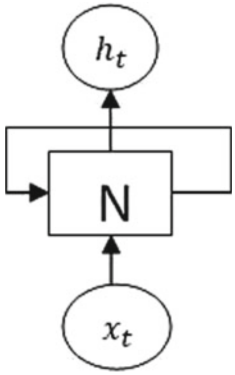


Fig. 1 Recurrent neural network

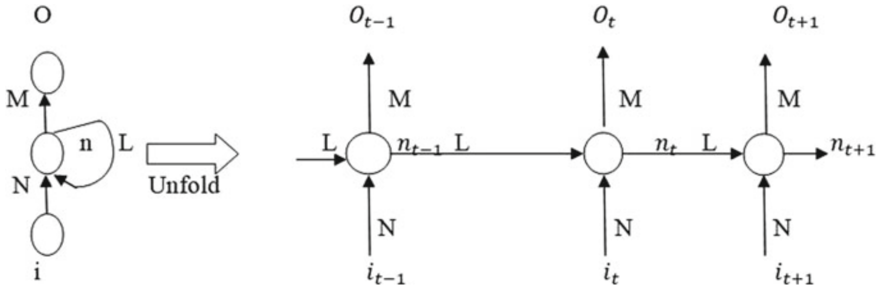


Fig. 2 Unfolding of recurrent neural network

The function f is an activation function which is mainly tanh or ReLU.

It is the output at step t. The next word in a sentence is predicted using the probability values.

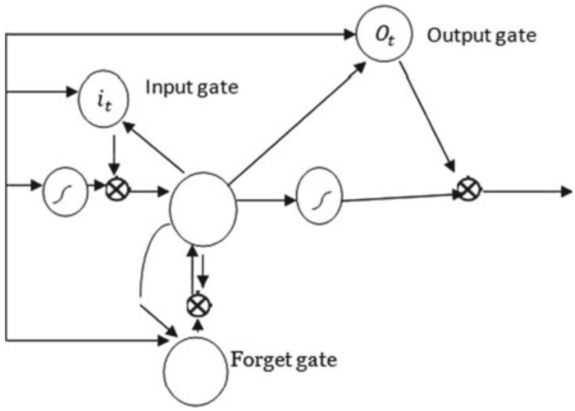
The drawback of RNN is that it cannot remember the past information for a longer time.

3.3 Long Short-Term Memory

Long short-term memory (LSTM) networks can be used which acts as an enhancement of recurrent neural networks. The informations of LSTM are stored in a memory which acts like a computer memory which allows information to read from, write to, and delete from its memory.

Its memory acts like a gated cell, in which the cell can decide whether to store the information or not by assigning importance to that information. The gated cells will pass or reject the information based on the signal they receive and based on that

Fig. 3 Long short-term memory



signals it can decide whether to accept or reject that information. The importance of information can be determined mainly through the weight which has been perceived through the process of iteration by making guesses and by adjusting weights. That means the importance of particular information can be assigned overtime.

LSTM mainly has three gates which are input gate, forget gate, and output gate. The input gate which is used can determine whether the new input can be passed into or not. The forget gate will delete the informations which are not that important. The output gate is used to output the information (Fig. 3).

3.4 Transfer Learning

Training of entire CNN from the beginning is a difficult task. To train a CNN network, it needs a large training dataset.

Transfer learning is done based on these two methods:

(1) *CNN as fixed feature extractor*: In this method, except the last classifier layer, all the other CNN layers are fixed while using the new dataset. The second last layer is a fully connected layer which produces CNN codes for each of the image.
(2) *Fine-tune the CNN*: In this method, instead of changing the last layer, the whole network is retrained. This is done by training again the pre-trained model with the new datasets without changing the weights or its structure. The weights are being fine-tuned. Once the training results meet the required output, the weights are fixed and can use those weights for the next time. The time taken for fine-tuning is large and may also cause over-fitting problem.

While performing transfer learning, two factors can mainly affect the transfer learning task. They are the size of the new dataset and the similarity of the new dataset and the original dataset.

(1) *Small dataset with low similarity*: This situation is the most difficult task as the training set is small, and it can cause over-fitting problem. Even if a new classifier is used, it will not train well as the similarity between the datasets is low.
(2) *Small dataset with high similarity*: In this situation, even if the similarity is high, it won't work well as the dataset is small. As the similarity is high, it can be used on the last fully connected layer while training which can produce better results than in the first situation.
(3) *Large dataset with low similarity*: In this situation, as it is having a large dataset it works well for training. But as the similarity is low, the CNN must be trained from scratch. This option gives more accurate outputs as it is having a large dataset.
(4) *Large dataset with high similarity*: This situation is the most appropriate one. As it is having a large dataset, fine-tuning is much efficient, and as it is having a higher similarity, the features learned from the original dataset can be used for the new dataset also.

4 Proposed System

The input images must be scaled to a standard size before passing to the network. The feature map is being generated for extracting the features from the image when passed to the convolutional layers. The output from convolutional layers is passed as an input to LSTM. The feature map was scanned which moves from left corner to right corner and top to bottom and will generate the feature vectors of feature sequence. The feature map column will correspond to a rectangular part of the input image.

4.1 Car License Plate Detection

The first stage of license plate detection and recognition (LPDR) system is license plate detection. In the first step, the characters are being trained using CNN classifier. The characters are being detected. The detected characters also consist of non-plate regions. These non-plate regions must be eliminated. A plate or a non-plate CNN classifier is used which classifies this license plate and non-license plate. Segmented license plates are used for training. The detection process is done by training a 7-layer CNN to classify character and a non-character. CNN is used which extracts the features from the whole license plate. All the CNN layers are concatenated together. The detected license plate bounding box will be too small or too large so a bounding box refinement must be done. After detecting the license plate, the output of CNN layers is fed to both bidirectional recurrent neural networks (BRNNs)

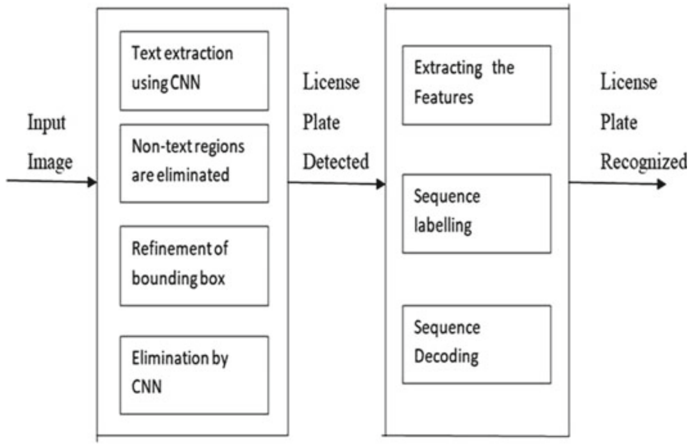


Fig. 4 System architecture

and long short-term memory (LSTM) for recognizing the feature sequence [28]. Bidirectional recurrent neural networks (BRNNs) are used to strengthen the process of feature extraction and selection process which considers both the past and future informations. This method then uses a softmax layer which transforms the BLSTM outputs into different probability values for each individual character classes. Then the output from both LSTM and BRNN is decoded to character string by CTC (Fig. 4).

4.2 Car License Plate Recognition

The second stage of license plate detection and recognition (LPDR) system is license plate recognition (Fig. 5).

(1) *Sequence feature generation*: In license plate recognition, the features are extracted from the cropped license plate image by CNN classifier. After the license plate is detected, the image will be converted to a gray scale and is resized to get the same height as the input height. After that CNN is used for extracting the features from the image.
(2) *Sequence labeling*: After the features are extracted, the sequence labeling was performed. In traditional systems, RNNs were used. The past information can be retrieved from RNN. During sequence recognition process, these past informations are used to get stable recognition results. Recurrent neural networks consist of two inputs as their sources, its current input and the previous output, and based upon that information, it will determine the new data. Recurrent network consists of a feedback loop and has a memory. In each time steps, it takes its current input and previous output, and after performing computations, it will store it in the memory. The output produced will be taken and will be passed to

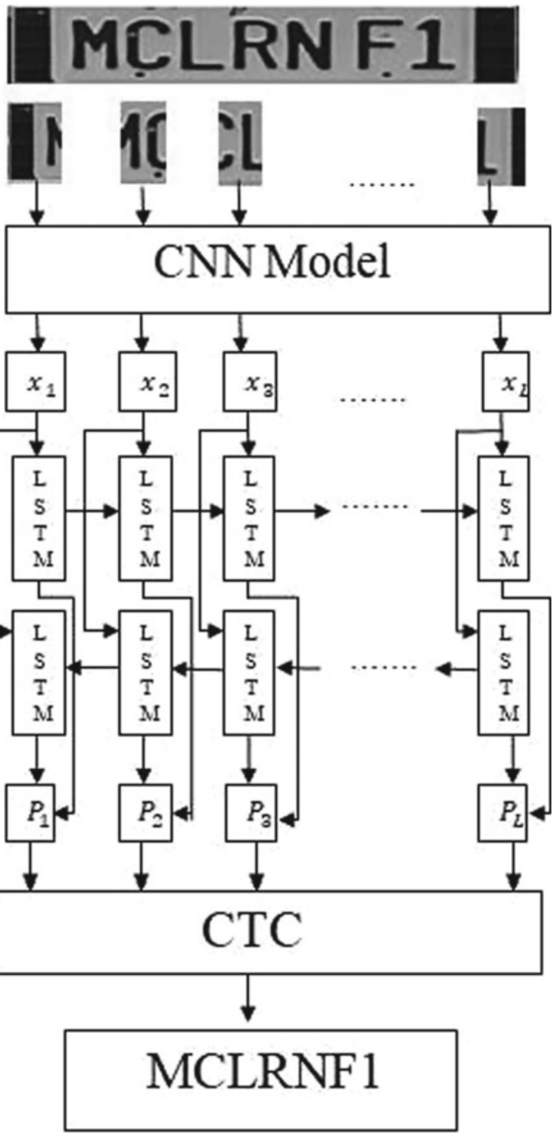


Fig. 5 License plate detection and recognition

the next node. This process will be continued until the sequence of information has been processed. The drawback of RNN is that it cannot remember the past information for a longer time. RNN remembers information only for a short time, and after consuming more information, it cannot recollect the previous information. This problem is called as gradient vanishing problem. This disadvantage is solved by LSTM as it contains memory which can be used to store information for a longer period.

The vanishing gradient problem is solved by LSTM as it will give priorities to information and will reject that information which isn't important. While using LSTM, its training time will be less compared to RNN and the accuracy is high.

(3) *Sequence decoding*: During sequence labeling, it has predicted many probability values. So to transform these sequences of probability values to a character string, connectionist temporal classification (CTC) is used.

5 Experimental Results

The ALPR system has the following steps:

1. Vehicle image capture.
2. Preprocessing.
3. Extraction of number plate.
4. Segmenting of characters.
5. Recognizing the characters.

The initial step of an ALPR is to capture an image of a vehicle. The second step is to obtain the position of the license plate from that image and after that extracting the license plate. The final step is character segmentation. This segmentation process is done using different techniques like histogram analysis, neural network, analysis of colors, and mathematical morphology. Segmentation is done for each single character. Optical character recognition (OCR) [17] can be used which stores each character on a database.

After obtaining an input image, the image is preprocessed; that is, the input image is converted to gray scale as these grayscale images are easily used to perform computation. After converting the input image to grayscale and threshold image, the boundaries of that image are obtained. Characters are being extracted from that image so as to get recognized. After extracting characters, the possible plates must be found out and resizing operations are done on that cropped image. And again apply preprocessing on that license plate. The character recognition is done while training the convolutional neural network. The character prediction is done on while training the LSTM layers.

An input image of car will be chosen. In this method, plates are found first and then the characters within the plate. First the input image is preprocessed which converts the original image to grayscale and threshold image. Then get the list of possible characters from the image which can contain more than 2000 contours which are not even characters. So it is necessary to eliminate non-characters from the list. For that, a condition will be specified. The condition includes calculating width, height, and aspect ratio. The contours that do not satisfy those conditions will be eliminated from the list. Thus, the number of contours gets reduced. Now find the list of matching characters. To find the matching characters, it will compare it to every other contour and then to the next contour compare that to every other contour and so on. The comparing of contours is done based on some conditions such as angle between characters, distance, area, width, and height. The matching characters which are grouped or clustered and are adjacent to each other can be possibly a license plate, and those plates are extracted from the image. After extracting the plates, preprocessing is done on those extracted plates. The characters in the plate are found. Inner overlapping characters are being removed as for example, after thresholding and quantifying the threshold on the letter ‘O’ will get the outer part of ‘O’ and inner part of ‘O’ separately recognized contours. So to avoid that problem, inner overlapping is removed which throws out the inner ‘O’ and leaves only the outer ‘O’. A list of matching characters would be found based on some conditions. The matching characters must be close to each other. The distance would not be considered as a match as it cannot be a part of the same license plate, and from that list, the longest list of potential character will be identified as the actual character. The individual characters from each plate are recognized and are shown with a red bounding rectangle (Fig. 6).

The system was built based on a pre-trained model. It was trained with 47,605 images of 36 classes ranging from 0 to 9 and A–Z and was tested with 1292 images. An input of 64×64 was given in the first convolutional layer. The activation function used in the convolution layer was rectified linear unit (ReLU). A softmax function was used in the final layer to predict the score of the input image. This system uses root mean square optimizer (RMS) with the starting learning rate as 0.001 which gradually decreases by 0.005 after each epoch. This system was being trained for 5 epochs. The accuracy was calculated based on the formula:

$$\text{Accuracy} = (\text{correct} * 100)/\text{Total images}$$

where correct is the number of license plates which were detected correctly. Total images are the total number of license plate images.

The loss was calculated using CTC lambda, and the loss rate was very less (Figs. 7, 8 and 9).

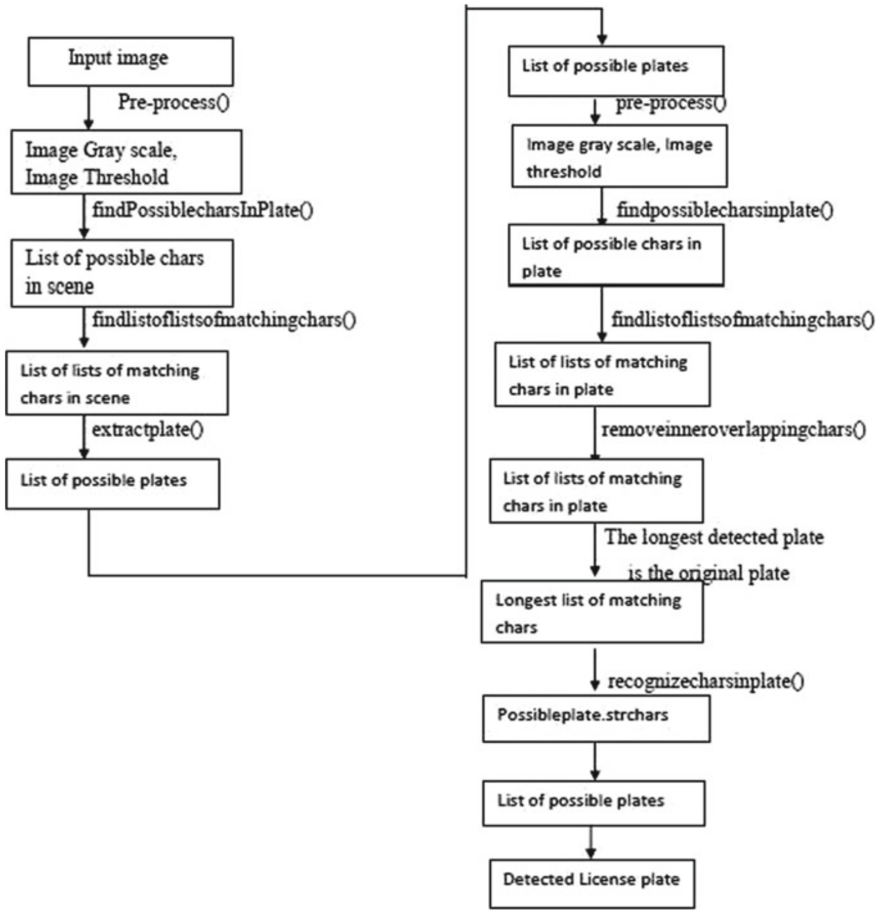


Fig. 6 Sequence followed in proposed system

6 Conclusion

The ALPR system has given a huge rise in demand for detecting the license plate with reliable and high performance in the transportation systems. The license plate localization will find the position of the license plate, and after detecting the license plate, the characters are segmented into individual character. Then the license plate is recognized. In this present work, a pre-trained model was used and had trained it with UK number plates. It was able to detect and recognize the number plate. The only limitation was that it was difficult to recognize the letter I and digit 1. The remaining characters and digits worked well in this system. License plate can vary in shape, color, and style. It can vary in different countries. License plate detection and recognition can also get affected by different illumination conditions. Here the

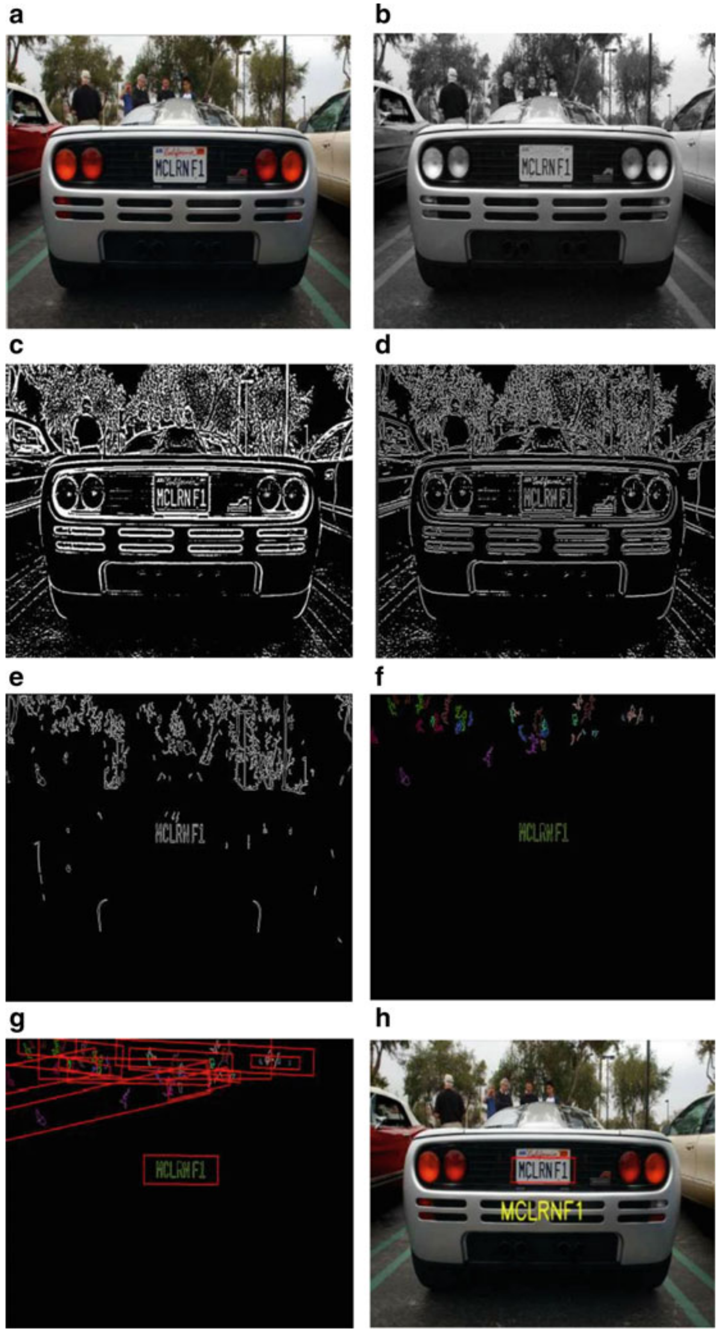


Fig. 7 Steps for license plate detection **a** initial image, **b** grayscale image, **c** threshold image, **d** all contours, **e** filtered contours, **f** contours with matching characters, **g** contours with matching characters are extracted, **h** vehicle number plate [29]

Fig. 8 Other results using this technique [29]

Fig. 9 Prediction output

```
Class:  22
predicted class: 22
Detected character -  M
Concatenated -  M

Class:  12
predicted class: 12
Detected character -  C
Concatenated -  MC

Class:  21
predicted class: 21
Detected character -  L
Concatenated -  MCL

Class:  27
predicted class: 27
Detected character -  R
Concatenated -  MCLR

Class:  23
predicted class: 23
Detected character -  N
Concatenated -  MCLRN

Class:  15
predicted class: 15
Detected character -  F
Concatenated -  MCLRNF

Class:  1
predicted class: 1
Detected character -  1
Concatenated -  MCLRNF1
```

system was trained on the UK number plate. As the future enhancement, this system can be used for Indian number plates. This system detects the number plate when an input image was fed on it which can be extended to videos on real-time applications.

Acknowledgements First and foremost, I express my heartfelt thanks to God the Almighty for showering infinite blessings on me that strengthened me through the successful completion of this milestone. Most importantly, I would like to express my gratitude to my project guide Dr. Tessy Mathew. I am extremely thankful to her for sharing her expertise, sincere and valuable guidance, and encouragement extended to me. Finally yet importantly, I take this opportunity to express my gratitude to all the department faculty members for their help and support. I also thank my parents for their unceasing encouragement, support, and attention. I am also grateful to my friends who supported me throughout this venture.

References

1. Du S, Ibrahim M, Shehata M, Badawy W (2013) Automatic license plate recognition (ALPR): a state-of-the-art review. IEEE Trans Circuits Syst Video Technol 23(2):311–325
2. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
3. Graves A, Liwicki M, Fernandez S (2009) A novel connectionist system for unconstrained handwriting recognition. IEEE Trans Pattern Anal Mach Intell 31(5):855–868
4. Anagnostopoulos CNE, Anagnostopoulos IE, Loumos V, Kayafas E (2006) A license plate-recognition algorithm for intelligent transportation system applications. J IEEE Trans Intell Transp Syst 7(3):377–392
5. Selmi Z, Halima MB, Alimi AM (2017) Deep learning system for automatic license plate detection and recognition. In: 14th IAPR international conference on document analysis and recognition, vol 1, pp 1132–1138
6. Sarfraz MS, Shahzad A, Elah MA (2013) Real-time automatic license plate recognition for CCTV forensic applications. J Real-Time Image Proc 8(3):285–295
7. Kocer HE, Cevik KK (2011) Artificial neural networks based vehicle license plate recognition. Sci Dir 3:1033–1037
8. Li H, Wang P, You M, Shen C (2018) Reading car license plates using deep neural networks. Sci Dir Image Vision Comput 72:14–23
9. Kulkarni P, Khatri A, Banga P, Shah K (2009) Automatic number plate recognition (ANPR) system for indian conditions. In: 19th International conference radioelektronika, IEEE transaction (2009)
10. Wen Y, Yue L, Yan J, Zhou Z (2011) An algorithm for license plate recognition applied to intelligent transportation system. IEEE Trans Intell Transp Syst 12(3):830–845
11. Bhushan B, Singh S, Singla R (2013) License plate recognition system using neural networks and multithresholding technique. Int J Comput Appl 84(5):45–50
12. Saunshi S, Sahani V, Patil J, Yadav A, Rathi S, License plate recognition using convolutional neural network. IOSR J Comput Eng 28–33. e-ISSN: 2278-0661
13. Badr A, Abdelwahab MM, Thabet AM, Abdelsadek AM (2011) Automatic number plate recognition system. annals of the university of craiova. Math Comput Sci 38(1):62-71. ISSN: 1223-6934
14. Deb K, Khan I, Saha A, Jo K-H (2012) An efficient method of vehicle license plate recognition based on sliding concentric windows and artificial neural network. Sci Dir Procedia Technol 4:812–819
15. Yang R, Yin H, Chen X (2015) License plate detection based on sparse auto-encoder. In: 8th International symposium on computational intelligence and design, pp 465–469

16. Prabhakar P, Anupama P, Resmi SR (2014) Automatic vehicle number plate detection and recognition. IEEE Trans Int Conf Control Instrum Commun Comput Technol
17. Qadri MT, Asif M (2009) Automatic number plate recognition system for vehicle identification using optical character recognition. In: Proceedings of the 2009 international conference on education technology and computer, pp 335–338
18. Naeem A, Rasheed A, Ishaq O (2012) Automated number plate recognition using hough lines and template matching. In: Proceedings of the world congress on engineering and computer science, pp 199–203
19. Zheng D, Zhao Y, Wang J (2005) An efficient method of license plate location. Pattern Recogn Lett 26:2431–2438
20. Kim S, Kim DW, Kim HJ (1996) A recognition of vehicle license plate using a genetic algorithm based segmentation. ICIP 2:661–664
21. Zhang H, Jia W, He X, Wu Q (2006) Learning-based license plate detection using global and local features. In Proceedings of the 18th international conference on pattern recognition ICPR'06 Washington, DC, vol 02, pp 1102–1105. IEEE Computer Society
22. Matas J, Zimmermann K (2002) Unconstrained licence plate and text localization and recognition. In: 2005 proceedings intelligent transportation systems, pp 225–230. IEEE
23. Li H, Shen C (2016) Reading car license plates using deep convolutional neural networks and lstms. CoRR, vol abs/1601.05610
24. Lin K-H, Tang H, Huang TS (2010) Robust license plate detection using image saliency. In: 2010 IEEE international conference on image processing, pp 3945–3948. IEEE
25. Kim KK, Kim K, Kim J, Kim HJ (2000) Learning-based approach for license plate recognition. In: Proceedings of the IEEE signal processing society workshop 2:614–623
26. Nukano T, Khalid M (2004) Vehicle license plate character recognition by neural networks. Proc Int Symp Intell Signal Process Commun Syst 771–775
27. Jiao J, Ye Q, Huang Q (2009) A configurable method for multi-style license plate recognition. Pattern Recogn 42(3):358–369
28. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE Trans Signal Process 45
29. Car images. https://github.com/MicrocontrollersAndMore/OpenCV_3_License_Plate_Recognition_Python/tree/master/LicPlateImages

Robust Adaptable Segmentation-Based Copy Move Forgery Detection Method

R. Dhanya and R. KalaiSelvi

Abstract The availability of the most advanced image correction tools along with the latest highly classy capturing devices has made the tampering of the images more easier. Copy move method of tampering happens when a particular section in the digital image is copy pasted on a new section on the same image for guarding or else hiding offensive areas. Digital forgery detection methods on images target the discovery of fake regions or the duplicated parts. Many pre- or post-operations are done on the images by persons who do the tampering. A novel method integrating the block-based and key-point-based methods is suggested in this paper. Initially, it adaptively segments the input host picture into non-overlapping and unequal sized slabs and features and thus mined from each block is harmonized with each further to discover labeled feature points that point out doubted fake regions approximately. In order to make it further precise, an extraction algorithm is proposed that substitutes the feature points by minor superpixels as slabs of feature which later fuses with the adjacent blocks which show like local color features to create that detected forgery sections. Lastly, morphological operations are performed on these fused areas to produce the spotted tampered areas. Experiments show that this method of detection achieves better outputs under several perplexing conditions compared with the latest CMFD algorithms.

Keywords Copy move forgery · Forgery detection · Digital forensics · Labeled feature points · Local color feature · Over-segmentation · Superpixels

1 Introduction

The recent technological advancement in the area image processing software has caused the process of digital image falsification to increase in a rapid mode. The popularity of images, mainly through social media, as a basis of information has

R. Dhanya (✉) · R. KalaiSelvi
Department of Computer Applications, Noorul Islam Centre for Higher Education, Nagercoil, Tamil Nadu, India
e-mail: dhanyamanoj2003@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_61

made the reliability of these images a vital issue. This has made more researchers to concentrate their area of research on digital image altering. One of the popular tampering methods is the copy move forgery, which pastes the regions of a selected image into some other area on the same image. Throughout this copy and paste processes, many pre-processing techniques like scaling, compression, rotation, noise addition, and blurring are applied on the regions to make the tampering more substantial. But, as the copied and the pasted regions forms part of the same image, some significant characteristics like the color character and noise component of the image are well-matched with the residue areas, and hence in this case, those methods adapted on the associated image characteristics are not appropriate.

2 Related Works

In the existing scenario, the detection methodologies of copy move forgery mainly are characterized into block-based and key-point-based.

The present methods that are block-based initially split the input image into overlying, even image blocks, after which the forged areas are acquired by corresponding image pixel slabs or transform coefficients. In [1], Fridrich et al. projected an algorithm with the image segmented into overlying rectangular slabs and the location of the forged region is calculated by matching the quantized discrete cosine transform (DCT) coefficients from every chunk. Principal component analysis (PCA) applied by Popescu and Farid [2] diminished the feature dimensions. In [3], Li et al. applied discrete wavelet transform (DWT) along with singular value decomposition (SVD) for feature extraction. 24 blur-invariant moments were used by Mahdian and Saic [4] as features. In [5], reduced rank approximation extraordinary values were estimated by Kang and Wei for each block. Fourier-Mellin Transform (FMT) was applied by Bayram et al. [6] to attain the features. Average intensities of circles with dissimilar radii were used by Wang et al. [7, 8] around the block center to excerpt the features. In [9], from every block and sub-block, gray average outcomes were used as the features by Lin et al. Information entropy was used as block features by Bravo-Solorio and Nandi [10]. Zernike moments were used by Ryu et al. [11, 12].

An alternative to the methods based on block is the key-point based on where key-points attained from the picture are mined, and entire span of the image is matched with to repel some image alterations encountered while recognizing replicated areas. Scale-invariant feature transform (SIFT) [13–16] was function along the original pictures to excerpt feature points and matched to every other later. A threshold value was set, and at the time when the shift vector value surpassed this threshold, the SIFT feature point sets matching were identified. The speeded-up robust features (SURF) was functional in [17–19], to excerpt features, instead of SIFT. Though the mentioned techniques identified the forged regions very well, they cannot accomplish reasonable results and a persistent high recall rate [20] at the same time. In [21], the authors made an evaluation of the prevailing methods, and from their analysis, they recommended that merging the two methods can distinguish the copy move forgery more effectively.

Majority block-based detection methods practice a related outline besides which they all differ in the diverse feature extraction approaches and they apply to abstract the block features. Even though operative in forgery detection, the methods have three major limitations: (a) the bigger size of the image leads to be computationally expensive as the suspected image needs to be divided into overlapping rectangular blocks. (b) These methods cannot effectively report noteworthy geometrical alterations of the fake areas, and (c) these methods suffer from low recall rate as the blocks are of definite shape. The first two demerits are overcome by the key-point-based methods, but their recall rate was established to be very low.

3 Superpixel Segmentation

Superpixels [22] are a collection of pixels with alike characteristics and normally color-based segmentation. Superpixels provide an appropriate primitive from which local features of an image are figured out. In the image, redundancy is captured and the intricacy of succeeding image processing jobs is greatly reduced. It has been verified progressively beneficial in applications such as body model approximation, skeletonization, depth approximation, object localization, and image segmentation. For making superpixels useful, it has to be fast, easily usable, and yield high-class segmentations. But most of the advanced superpixel approaches do not fulfill the mentioned necessities. Often it suffers from high computational cost, very poor segmentation quality, unpredictable dimension and form or contains numerous uncontrollable constraints. The approach we discuss in this work addresses these concerns and creates high-quality solid, closely even superpixels more efficiently.

4 Proposed Method

The traditional forgery discovery approaches, block and key-point-based, are combined in this projected scheme. Like block-based methods, a blocking method is suggested which adaptively segments the image into non-overlying and asymmetrical slabs. Thereafter, like in key-point methods, the points of features are extracted from distinct block as an alternative of being mined from the whole image. Consequently, the features of the blocks are compared with one another such that the labeled feature points are discovered that almost specify the assumed forged regions. In order to identify forged regions further accurately, we put forward the forged area extraction algorithm that substitutes as feature blocks and the feature points with minor superpixels. Later, to create those tampered regions of the image, feature blocks are formed by combining the neighboring blocks which have identical local color features. And lastly, morphological operation is performed on these combined

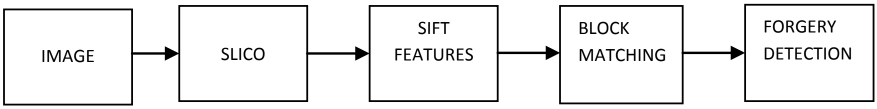


Fig. 1 Proposed system

parts to create those identified fake areas. Here we use SLICO to segment the image. By using this method, the user no longer has to set the parameters of compactness or try diverse values of it. SLICO adaptively selects the compactness parameter for every single superpixel distinctly. It results in the generation of regular shaped superpixels in both textured and non-textured regions alike. And this will not compromise on the computational complexity (Fig. 1).

Scale-invariant feature transform (SIFT) is a four-stage process developed by David Lowe and creates exclusive and extremely expressive features from the image. Features are designed to be invariant to rotation and tough for alterations in scale, noise, illumination, slight changes in viewpoint, etc. Thus, extracted features indicate if there is any relation between spaces within the images. The four stages of the SIFT algorithm given in Lowe's paper are as follows [23].

(1) *Scale-space extrema detection*: Initial step forms the creation of a Gaussian scale-space pyramid for the image. Continually blurred images are produced from the convolution of Gaussian functions which create multiple octaves. Difference between two consecutive images within an octave calculates the difference of Gaussian (DoG). Matching each point in the DoG images and its 26 neighbors, and looking for extreme, the initial set of candidate features are selected.
(2) *Feature localization*: Reduction in the number of features happens in this stage. To locate the exact superpixel, interpolation occurs. Location of the candidate features and low contrast area points or those that are localized along the edges are eliminated.
(3) *Orientation assignment:* The image gradient directions of the pixels in a feature's neighborhood are calculated. This is added to a histogram of orientation with 36 bits. The neighborhood values are Gaussian weighted so that those closer to the center have a better outcome on the resultant orientation. One key orientation is selected for each feature.
(4) *Creating the feature descriptor*: A128-dimensional vector feature descriptor labels the pixel properties of the area in which a feature neighbor. A 4×4 array of 16 histograms is focused upon the feature and rotated so as to match the key orientation that is calculated in the last step. The gradient magnitudes are given a Gaussian weighting, then added to the histograms, and are normalized to create the descriptor.

Once the block features (BF) are obtained, the matched blocks through them are located. In majority prevailing block-based methods, matching procedure yields an exact block pair simply if there occurs several further identical duos in the same conjoint site, supposing them having equal shift vector. This upon surpassing quantified threshold from user, the slabs that matches contribute to the explicit shift vector are recognized to be those areas which may be assumed to have been forged. As the block features are composition of a group of feature points, a new technique is projected to identify those matched blocks. Initially calculated the amount of feature points matched, followed by the adaptive calculation of corresponding matching threshold value. From the outcome, the pairs of block matches are identified. Lastly, matching feature points contained inside the block sets are mined and then characterized so that point of doubted fake section can be traced.

Labeled feature points are mined based on credentials of block differences from the images. The process of labeling feature points is done on the following steps:

In STEP-1, supposing that LPF $= \left\{\left\langle LP_1, \overline{LP_1}\right\rangle, \left\langle LP_2, \overline{LP_2}\right\rangle, \ldots \left\langle LP_n, \overline{LP_n}\right\rangle\right\}$ where $\left\langle LP_i, \overline{LP_\iota}\right\rangle$ signifies feature point set matched, i denotes ith set of labeled feature point, $i = 1, 2, \ldots, n$, and n is the entire feature points number in LPF; the assumed areas SR given by SR $= \left\{\left\langle LS_1, \overline{LS_1}\right\rangle, \left\langle LS_2, \overline{LS_2}\right\rangle, \ldots\ldots.\left\langle LS_n, \overline{LS_n}\right\rangle\right\}$.

The size of the real images determines the primary dimension in the SLICO algorithm S that fragments the original image into smaller superpixels. Through trials, the initial dimension of S is set to 20 in the case of extraordinary resolution images, [image size approx. 3000 × 3000] and to 10 with respect to low-resolution images [image size approx 1500 × 1500].

For every single-doubted region, in STEP-2, $SR_i = \left\{\left\langle LS_i, \overline{LS_\iota}\right\rangle\right\}$ the adjacent blocks demarcated as SR$_i$_neighbour $= \left\langle LS_{i\_\theta}, \overline{LS_{\iota\_\theta}}\right\rangle$ where θ differs by an angle of 45°};

$$F_C\_LS_i = \frac{R(LS_i) + G(LS_i) + B(LS_i)}{3}$$
$$F_C\_\overline{LS_i} = \frac{R\left(\overline{LS_\iota}\right) + G\left(\overline{LS_\iota}\right) + B\left(\overline{LS_\iota}\right)}{3}$$
$$F_C\_LS_{i\_\theta} = \frac{R\left(LS_{i\_\theta}\right) + G\left(LS_{i\_\theta}\right) + B\left(LS_{i\_\theta}\right)}{3}$$
$$F_C\_\overline{LS_{i\_\theta}} = \frac{R\left(\overline{LS_{\iota\_\theta}}\right) + G\left(\overline{LS_{\iota\_\theta}}\right) + B\left(\overline{LS_{\iota\_\theta}}\right)}{3}$$

in which R(), G() and B() give the mean of the RGB constituents of respective slabs. Local color features of neighboring blocks when found similar to the corresponding doubted regions, the local feature will attain a state as follows:

$$\left|F_C\_LS_i - F_C\_LS_{i\_\theta}\right| \le TR_{\text{sim}}$$
$$\left|F_C\_\overline{LS}_i - F_C\_\overline{LS_{i\_\theta}}\right| \le TR_{\text{sim}}$$

The neighboring blocks are fused to the equivalent forged area, where $F_C\_LS_i$ and $F_C\_\overline{LS}_i$ display the color features that are local to the forged area, SR_i.

$F_C\_LS_{i\_\theta}$ and $F_C\_\overline{LS_{\iota\_\theta}}$ are the color features of the adjacent blocks SR$_i$_neighbour, SR_i_neighbour $= \langle LS_{i\_\theta}, \overline{LS_{\iota\_\theta}} \rangle$. TR_{sim} is the threshold value that calculates the correspondence among the local color features. STEP-3 creates a circle with radius associated with the dimensions of real image well-defined the structural element used in the close operation. This process fills those gaps in the fused areas, and in the meantime, the shape of the area is kept unaffected.

Even though the labeled feature points (LFP) give the positions alone of tampered regions extracted, we are yet to trace the exact tampered sections. By keeping in mind that the superpixels are able to fragment the host image, a new method is proposed by substituting LFP with minute superpixels so as to attain the doubted region (SR) and composed of labeled minor superpixels. Additionally, as an improvement to the precision and recall results, the color features of the superpixels which neighbor the doubted regions (SR) are calculated. If the color features found identical to the alleged regions, we combine the adjacent superpixels into the respective suspected regions, and this in conclusion creates the fused regions. To the combined regions, morphological operations are conducted that produces the exact identified copy move forgery regions.

5 Experimental Results

The proposed method is implemented using MATLAB 2016a on Intel Core i3-5005U CPU @ 2.00 GHz and 4 GB RAM. The algorithm is tested on the dataset MICC-F220, where the copy move areas on the images are subjected to various affine transformations like rotation, scaling, and translation.

To analyze the performance of the algorithm, we calculated the Precision, Recall, and F1-measure.

$$p = \frac{T_p}{T_p + F_p} \quad r = \frac{T_p}{T_p + F_n}.$$
$$F1 = \frac{p.r}{p + r}$$

Applied on a pixel basis, *Tp* is the number of forged pixels that are correctly identified; *Fp* is the number of authentic pixels erroneously labeled as forged; *Fn* is the number of forged pixels erroneously labeled as authentic.

Hence, precision is the fraction of pixels identified as tampered that are truly tampered. Recall (or true positive rate) is the fraction of tampered pixels that are correctly classified as such.

The proposed method had given a precision of 97.11%, and a recall value of 1 with F1-measure is equivalent to 0.9853 which is a fairly good value in comparison with the state-of-the art copy move forgery detection methods (Figs. 2, 3, 4, 5, 6 and 7).

Fig. 2 Input image

Fig. 3 SLICO image

6 Conclusion

Copy move operations create the tampered images that are toughest to distinguish. The algorithm fragments the whole image into non-overlying, unequal blocks conferring the size and density of the original host images. For every single image, a

Fig. 4 SIFT features extracted

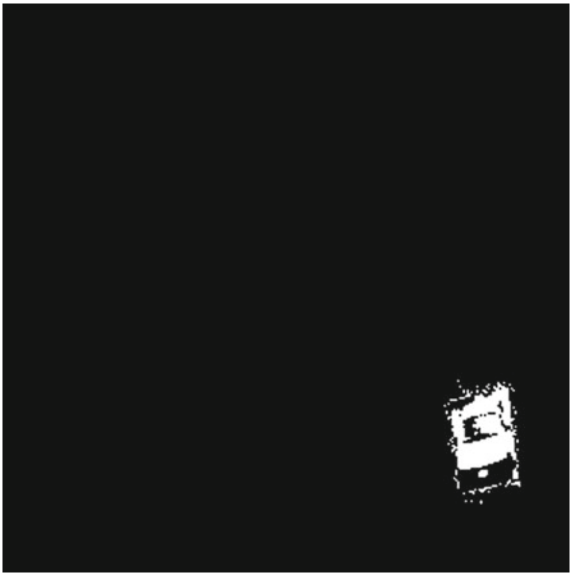


Fig. 5 Labeled feature points

suitable block initial size can be limited to improve the accurateness of the tampering recognition outcomes with reduced computational expenses. Later, from every such block, the feature points are extracted, and they are matched with each other using feature matching algorithm to identify the labeled feature points, that almost specify the supposed tampered areas. Further, the forged regions are identified more accurately

Fig. 6 Merged regions

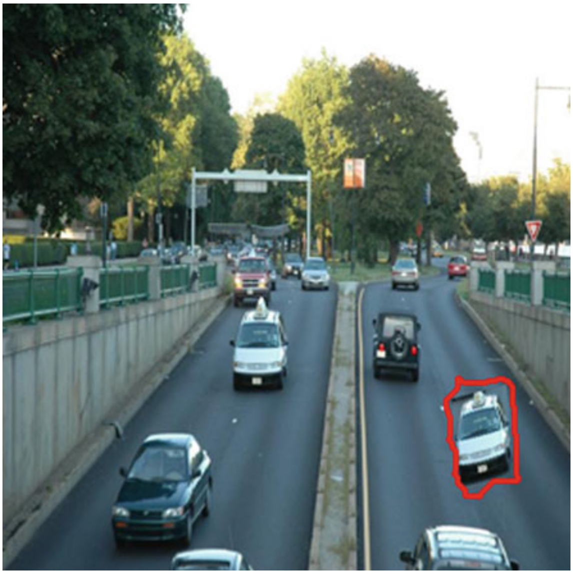


Fig. 7 Detected forged region

by the algorithm through which the minute superpixels substitute labeled feature points with feature blocks, and the neighboring blocks of feature substituted with local color features identical to the block feature are combined to create fused areas. Later, morphological operations are performed that generate the detected tampered areas.

The experimentation shows that the proposed scheme achieves better results than the existing methods. The future work could focus on checking the algorithm under various challenging conditions like geometric transforms and JPEG compression.

References

1. Fridrich J, Soukal BD, Lukáš AJ (2003) Detection of copy-move forgery in digital images. In: Proceedings of digital forensic research workshop
2. Popescu C, Farid H (2004) Exposing digital forgeries by detecting duplicated image regions. Dept Comput Sci Dartmouth College Tech Rep (TR, 2004–515)
3. Li G, Wu, Tu D, Sun S (2007) A sorted neighbourhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. In: IEEE International conference on multimedia and expo, pp 1750–1753
4. Mahdian B, Saic S (2007) Detection of copy–move forgery using a method based on blur moment invariants. Forensic Sci int, vol 171, pp 180–189
5. Kang X, Wei S (2008) Identifying tampered regions using singular value decomposition in digital image forensics. In: 2008 International conference on computer science and software engineering, pp 926–930
6. Bayram S, Sencar HT, Memon N (2009) An efficient and robust method for detecting copy-move forgery. In: 2009 IEEE international conference on acoustics, speech and signal processing ICASSP 2009, pp 1053-1056
7. Wang J, Liu G, Li H, Dai Y, Wang Z (2009) Detection of image region duplication forgery using model with circle block. In: 2009 International conference on multimedia information networking and security, MINES'09. pp 25–29
8. Wang J, Liu G, Zhang Z, Dai Y, Wang Z (2009) Fast and robust forensics for image region-duplication forgery. Acta Automatica Sinica 35:1488–1495
9. Lin H, Wang C, Kao Y (2009) Fast copy-move forgery detection. WSEAS Trans Signal Process 5:188–197
10. Bravo-Solorio S, Nandi AK (2011) Exposing duplicated regions affected by reflection, rotation and scaling. In 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 1880–1883
11. Ryu S, Lee M, Lee H (2010) Detection of copy-rotate-move forgery using zernike moments. In: International workshop on information hiding 2010, pp 51-65
12. Ryu SJ, Kirchner M, Lee MJ, Lee HK (2013) Rotation invariant localization of duplicated image regions based on zernike moments. IEEE Trans Inf Forensics and Secur 8:1355–1370
13. Huang H, Guo W, Zhang Y (2008) Detection of copy-move forgery in digital images using SIFT algorithm. In: 2008 Pacific-Asia workshop computational intelligence and industrial application, PACIIA'08, pp 272–276
14. Pan, XY Lyu S (2010) Region Duplication Detection Using Image Feature Matching. IEEE Trans Inf Forensics Secur 5:857–867
15. Amerini I, Ballan L, Caldelli R, Del Bimbo A, Serra G (2011) A sift-based forensic method for copy–move attack detection and transformation recovery. IEEE Trans Inf Forensics Secur 6:1099–1110
16. Kakar P, Sudha N (2012) Exposing post processed copy–paste forgeries through transform-invariant features. IEEE Trans Inf Forensics Secur 7:1018–1028
17. Bo X, Junwen W, Guangjie L, Yuewei D (2010) Image copy-move forgery detection based on SURF. In: 2010 International Conference on Multimedia Information Networking and Security (MINES), pp 889-892
18. Shivakumar B, Baboo LDSS (2011) Detection of region duplication forgery in digital images using SURF. IJCSI Int J Comput Sci Issues 8

19. Bay H, Tuytelaars T, Van Gool L (2006) SURF: speeded up robust features. In: European conference on computer vision—ECCV 2006, pp 404–417. Springer, Berlin
20. Luo W, Huang J, Qiu (2006) Robust detection of region-duplication forgery in digital image. In: 18th International conference on pattern recognition, ICPR 2006, pp 746–749
21. Christlein V, Reiss C, Jordan J, Angelopouplou E (2012) An evaluation of popular copy-move forgery detection approaches. IEEE Trans Inf Forensics Secur 7(6)
22. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S (2012) SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell 34:2274–2282
23. Lowe DG (1999) Object recognition from local scale-invariant features. In: The proceedings of the seventh IEEE international conference on computer vision, pp 1150-1157

Behavior Tracking in Video Surveillance Applications: A Detailed Study

L. Shana and C. Seldev Christopher

Abstract The problem of video surveillance toward behavior tracking has been well studied. The object tracking in video surveillance and behavior analysis has been used for different problems. Numerous techniques have been presented earlier for the tracking of objects and classification of behaviors. The general approaches use a background model to identify the foreground objects. Based on the foreground objects identified, the object tracking has been performed. Similarly, various approaches consider different features to perform object detection and classification. Each method produces a different result with varying accuracy in behavior tracking. This paper analyses various methods of object tracking and behavior analysis in video surveillance. A detailed survey on the methods of object tracking is performed, and a comparative study is presented in this paper.

Keywords Video surveillance · Object detection · Object tracking · Behavior tracking

1 Introduction

The modern organizations involved in monitoring different process of organization which has been required to make decision and control the operations and other activities of the organizations. To perform such activities, the video surveillance has been used. By using video surveillance, the employees and the activities of the organization can be monitored. It is not possible for the administrator to watch their employees all the time, and they cannot be located in all the places. To make it possible, the video surveillance systems have been emerged. The arrival of video surveillance systems helps various monitoring like animal tracking in forests. In

L. Shana (✉) · C. Seldev Christopher
St. Xavier's Catholic College of Engineering, Chunkankadai, Nagercoil, Kanyakumari, Tamil Nadu, India
e-mail: shanbiji@gmail.com

C. Seldev Christopher
e-mail: cseldev@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_62

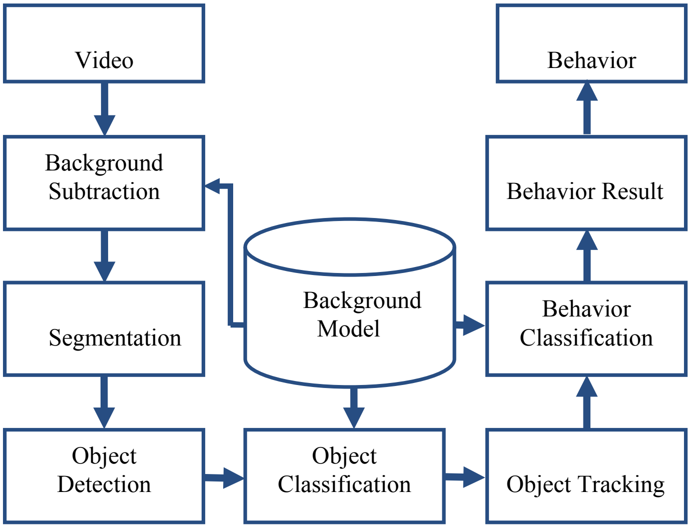


Fig. 1 Architecture of general behavior tracking system

recent days, the world society is more concerned about the growth and safety of wild animals. In order to take senses about the animals living and their movement, it is necessary to monitor them. To perform this, the video surveillance can be used. Similarly, there are a number of applications can be named.

To monitor the human or animal, it is necessary to detect them in the videotapes. The videotapes are the collection of video frames. The video frames are coming with both background and foreground objects. In order to identify the foreground objects, first the background features have to be identified and eliminated.

On the way to remove the underlying features numerous segmentation algorithms are available. Each algorithm uses different underlying models. Likewise, to recognize the foreground objects, dissimilar features are pulled out from the images. The objects sketch, skeleton, shape and texture features are used in identifying the objects. The objects may vary in their size, shape due to the mobility nature, so it is necessary to consider the variant features.

Behavior tracking has become a dominant research in recent days. Behavior tracking has been used in several problems where the dedicated monitoring person cannot be placed. It has been performed by classifying the motion patterns of any object by high-level description. In some situations, it is required to investigate the behaviors of people and to decide whether the detect behaviors are normal or abnormal. For example, the video files from the database can be classified for different actions, according to the behavior tracking algorithms. On the other side, the behavior surveillance can be used in the place of traffic control. Similarly, a number of applications can be named toward behavior tracking. The next section discusses the general working principle of behavior tracking.

Figure 1 shows the behavior tracking system and shows various functional components of the architecture.

1.1 Background Subtraction/Segmentation

The video surveillance tapes have come with both background and foreground features. The foreground features are the objects to be considered. But in order to track the objects, it is necessary to remove the background features. By means of eliminating the underlying features, the foreground features and objects can be recognized.

The background models are used to perform such operations and help to remove the backgrounds. The background model would have different color values and objects and various other features. By keeping track of such features, the background subtraction can be performed. The segmentation is the process of grouping similar features into a specific name. In this case, the segmentation algorithm split the background and foreground features.

1.2 Object Detection/Object Tracking

Object tracking is the method of following the movement of the object from the videotapes available. The objects are identified based on the changing features like shape, texture and sketch. To perform this, the low-level features can be used. However, identifying the varying features like shape and texture plays a vital role in object tracking. But the efficiency of object tracking is highly depending on identifying the object from the sequence of frames. Once the objects and motion of them are detected, the moving objects between different frames are detected. The tracking of features is performed with points and blob.

In [1], a relational model has been presented toward object tracking. The method uses the relationship between the environments and objects to perform tracking. The Bayesian algorithm is used to perform tracking. Similarly, in [2], a web-based multi-object tracking algorithm is presented for image analysis and assessment using a top-down Bayesian formulation that effortlessly integrates state estimation, track management, clutter rejection, occlusion and misdetection handling into a single recursion. In [3], a Bayesian-based multi-label classification algorithm is presented. The method uses the correlation between labels and relationships in object tracking.

In [4], the author presents a novel to detect the passengers in public transport vehicles. The first step is the detection of the human head by a suitable algorithm. The objective of this paper is to recognize the mistrustful behavior of a human in public transport vehicle.

1.3 Object Classification

The shape feature of the moving object would vary. In order to classify the object, the shape feature is more essential. Also, in order to track the behavior of the object, the object has to be classified, according to the shearing of the feature. Similarly, the object classification can be performed based on the motion features. However, the classification algorithm should consider the varying feature like shearing, scaling and so on.

1.4 Object Tracking

The object tracking is performed by considering both foreground and background features. It has to track the moving object by considering the shape features. The object tracking process can be classified as follows:

Region-Based Tracking In this type of approach, the objects can be maintained for their location like the region where it is located. By maintaining the location details of the object between different sequences of frames, the movement of the objects can be tracked.

Contour-Based Tracking In this type of approach, pixel to pixel comparison is performed and such varying locations only considered tracking the objects.

In [5], a contour-based object tracking algorithm is presented. The method tracks the distribution of objects and background distribution. The method performs segmentation based on color values. The background distribution has been used to track the objects.

Feature-Based Tracking Feature-based tracking is based on tracking individual feature points across successive frames in the video. The features are edges, corners, etc. First, the features are found in consecutive image frames. Second, these features are matched between the frames. This method is used to identify and track the different parts of the object. Different algorithms are used to achieve recognition and tracking of objects by pulling out elements, grouping them into higher-level features and then matching the feature between images.

Model-Based Tracking In this type of approach, a specific object model has been used. Specific object has been tracked in each sequence of the image.

Appearance-Based Tracking In [6], the author presents a collaborative model between a pre-trained object. First, the video is divided into frames. Then for each frame, identify the detections and trackers. In this approach, the activities are classified into two classes, and different layers are used to represent them. The abnormal data has been classified efficiently.

In [7], the features are organized hierarchically and the features are classified into a number of classes. The CNN has been used to learn and suggest an area adaptation unit to online adapt the pre-learned features according to the definite target object. Similarly, in [8], an appearance-based approach is presented which uses district frame patch indexing. The frames are clustered using appearance-based linear clustering.

2 Methods of Object Tracking

The methods of object tracking can be classified into several classes according to the feature and method used.

2.1 Shape-Based Object Tracking

In [9], a combined approach has been presented. The method uses both shape and background features for object tracking. The method identifies the shape control points to track the objects. Similarly, in [10], the shape features are used to generate the shape matrix to represent the features. The shape matrix has been generated at different postures, and posture labels have been used to perform classification using the Bayesian classifier.

In [11], the author presents a skeleton-based human action recognition using various poses of human. The method combines kernel features obtained from various pose sequences.

In [12], a traffic sign detection algorithm is presented toward the alert system. Traffic signs are noticed by examining color information, red and blue, contained in the images. Similarly, the shape features are extracted and have been used to classify them. Similarly, to monitor the public transport, an elliptic curve algorithm is presented. The method removes the blur from geometric features to perform classification.

In [13], the shape feature has been used to identify the falls from the human. The method uses the Gaussian mixer model in identifying the falls.

In [14], a combination of vehicle tracking and vehicle detection method is proposed. Kalman filter is used to improve the accuracy of the object detection. First, the video is transformed into frames, which is converted to grayscale images. Three frame differences are the method used to detect the boundary. Algorithm with Sobel matrix is used to detect the edges of all the edges of the grayscale converted image. Kalman filter method is used to track the moving objects.

In [15], the author presents a Histogram of Oriented Gradients to detect the human. The execution time increased by using the salience-windowed images. HOG features are used to detect the human, and k-means algorithm is used to classify the features.

In [16], the Gaussian mixture model is used to identify the object whether it is a human or non-human. In this, HOG algorithm is applied to ROI (Region of Interest) to identify the type of object which is present in that image.

Detecting human beings and recognizing the events in the video plays an important role in computer vision. In [17], cluster-based segmentation is used to detect the human in the video. In this, the input video is separated into a number of frames using frame generation block. Then, the cluster-based segmentation is done. HOG is used for feature extraction. Classification is done by the support vector machine algorithm. The proposed method has an accuracy of 89.59%.

In [18], various regions of a video can be identified by the saliency mapping and spatial saliency mapping. The highly moving object can be detected using temporal mapping, and the regular motion is identified by spatial mapping.

2.2 Occlusion-Based Object Tracking

The objects in the video may be occupied by dust. The dust detection can be used to identify the particles using optical flow can be used to perform tracking.

In [19], the author presents a novel approach to detect and track multiple objects. The result of tracking can be used to perform surveillance in home and business places. The method uses the color values in the cancelation of background features.

3 Methods of Behavior Analysis

In [20], the author delivers a comprehensive review of movement recognition in video surveillance applications. There are many real-time applications including visual surveillance by way of content-based retrieval and human–computer interaction methods.

3.1 Hidden Markov Model-Based Behavior Analysis

In [21], a behavior recognition algorithm by HMM has been presented. The images have been transformed into vectors and sequences. The vector quantization is generated from the image sequences. Based on the vectors, the behavior analysis is performed.

In [22], the author presents group activity recognition in video surveillance systems. The method uses asynchronous hidden Markov model (AHMM) to model the association between people. The method recognizes the group activity and handles both symmetric and asymmetric activities.

A weighted feature fusion scheme is presented in [23]. The method learns the objects in multiple views. Based on the weighted features, the classification is performed. In [24], a human motion recognition algorithm is presented. Vision-based human detection is proposed. Trajectory-based recognition method is used to recognize the presence of the human.

In [25], the author presented HMM-based action classification algorithm. The method uses the skeleton features, and feature vectors are generated. The HMM classifier works on the feature vector to produce results of action.

Human Movement Analysis and its Application to Surveillance Systems [26] presents a sorting arrangement to evaluate the posture of human movements directly from the video sequences. The analyzed movements are converted into a posture sequence in the system. The method is more robust, and the skeleton features have been used to extract context features. The centroid features are used to characterize the shape features.

3.2 Fuzzy-Based Behavior Analysis in Video Surveillance

Numerous investigations on Intelligent Video Surveillance methods for doubtful action detection dangerous review [27], the author present detailed review on different methods of video surveillance. Fuzzy-based method is used for online multiple tracking. Knowledge-based fuzzy inference system with a set of fuzzy if-then rule is used for online tracking. The rules used to determine the fuzzy membership degrees, which can be used to substitute the association probabilities between the objects and the measurements.

In [28], the author presents a different procedure for the detection of road traffic. The system detects the presence of vehicle from the image frames. Hough transform is used to extract the contour lines of the vehicles. Morphological operations are used to remove the noise in the image. The shape of the object regions is increased when the noise is removed. Fuzzy-based integrals are used for the detection of vehicles based on the information gathered.

In [29], it defines a new strategy for detecting abnormal behaviors. Fuzzy logic is being used to examine the video streams coming from the network of surveillance cameras. The camera will be deployed in certain areas of requirements in order to notice irregular behavior. The findings of these behaviors will rise the speed of the response of the security services. This method can be used to have an accurate analysis and detection of events in real time.

A Fuzzy Model for Human Fall Detection in Infrared Video [30] presents an infrared video-based inactivity monitoring. The method uses geometric and kinematic features to track the ROI. The fuzzy algorithm has been used to perform tracking.

In [31], it presents an activity analysis scheme which captures the images and transmits to the system using fuzzy logic. This fuzzy logic method is used for the recognition of human activity in video surveillance applications. The shape features are extracted, and fuzzy values are used to perform inference. The system has been adapted to track the child within specific area.

In [32], the author presents a fuzzy logic and optical correlation-based face recognition method. The application of this method is to monitor the patient in home video surveillance. The algorithm presents a correlation-based subject behavior analysis. The model considers behavior changes in different levels of alarms.

3.3 Support Vector Machine-Based Behavior Analysis

Movement Human Actions Recognition Based on Machine Learning [33] uses optical flow to track the objects which are moving. The flow energy has been used to extract the features of moving objects. The object features are used to generate the convolution neural network, and SVM has been used to perform classification on various behaviors.

Automatic Video Based Surveillance System for Abnormal Behavior Detection [34], in surveillance videos is proposed in this paper. We have targeted to create a system for the recognition of human activity and behavior and extract new information of interest for end users in highly secured indoor surveillance system. The multi-object detection is done by background subtraction with the help of appropriate model and then recognizing the person using HOG feature and SVM classifier.

In [35], the author suggests a novel hybrid optimization of feature selection and support vector machine (SVM) mechanism. Abnormal event detection is done by way of training the SVM model. To reduce the dimensions of multi-feature, an adaptive genetic simulated annealing algorithm (ASAGA) feature selection method is used.

Human Behavior Analysis Based on a New Motion Descriptor [36] generates a descriptor for the moving objects which has been identified using optical flow. The motion filters are used to extract shape and trajectory features. The behaviors are classified using the support vector machine.

Traffic Behavior Recognition using the Pachinko Allocation Model [37] presents for road surveillance which combines SVM and pachinko allocation model (PAM) to perform behavior classification. The method uses the Kalman filter to track the object using the Gaussian mixture models (GMMs). The method extracts the sparse features to track the objects and classify the behaviors in traffic.

In [38], the author presents a hidden Markov model with semi-structure which classifies the activities into low and high levels. Based on the Coxian feature distribution, the method performs classification.

3.4 ANN-Based Behavior Analysis

In [39], an object tracking and behavior classification algorithm is presented. The method captures the number plates of cars and uses artificial neural network for the classification.

Detection of Cattle Using Drones and Convolutional Neural Networks, [40] presents drone tracking algorithm which trains the drone features using the convolution neural network.

Design and Implementation of Behavior Recognition System Based on Convolutional Neural Network [41] presents a CNN-based behavior analysis model which segments the video into images. Second, the method segments the background and extracts the foreground characters. The feature extracted has been trained with CNN to perform behavior classification.

In [42], the author defines a neural network classifier. A data-based modeling neural network such as modified probabilistic neural network (MPNN) is proposed. This neural network algorithm partitions the decision space nonlinearly in order to achieve reliable classification, however, still with acceptable computations.

Convolutional Neural Networks Based Fire Detection in Surveillance Videos [43] presents a fire detection approach in video surveillance. In [44], it proposed a skeletal joint points' configuration captured from Kinect RGB-D camera. Skeletal points have contributed significant features to discriminate different actions. However, some joint points are irrelevant and not moving which acts as noise. This problem will reduce the performance of HAR. Therefore, in this paper a method to identify informative joint point using Shannon's entropy mechanism is mentioned. Then, the features will be an input to the convolution neural network (CNN) classifier.

In [45], the author presents a feed-forward neural network-based behavior analysis algorithm. The method extracts the appearance features and measures the crafted similarity measure. The method uses FFNN toward classification where the method considers the occlusion, rotation and scaling.

In [46], it presents scale and location-based object tracking which uses convolution neural networks. The method uses both temporal and spatial features for classification. The classification is performed by measuring the scale feature based on localization features

3.5 SOM-Based Behavior Analysis

Aquatic Toxic Analysis by Monitoring Fish Behavior Using Computer Vision: A Recent Progress [47] presents a human tracking algorithm for 3D space. The method identifies the abnormal pattern based on the shape features.

Automated Intelligent Surveillance using Human Behavior Analysis in Shopping Malls [48] classifies the human activities toward suspicious event. The methods have been designed to generate alarm on suspicious activity. In Student Behavior Analysis

Using Self-Organizing Map Clustering Technique [49], the method considers the patterns of students for the classification. The classification is performed using self-organizing maps.

Human Action Recognition using Image Processing and Artificial Neural Networks [50] extracts the features of the video and transform the images using two-dimensional discrete cosine transform (2D-DCT). The self-organizing maps have been used to perform classification.

Supervised Self-Organizing Maps in Drug Discovery Robust Behavior with overdetermined Data Sets [51] extracts the nonlinear correlation features for behavior classification. The classification is performed using self-organizing maps.

3.6 KNN-Based Behavior Analysis

In [52], the k-means clustering approach is used to group the patterns of behaviors. Based on the patterns available, the classification is performed.

3.7 PCA-Based Behavior Analysis

In [53], the author defines a vehicle detection method which splits the entire image into a number of blocks. The method measures the gray-level changes, and the principle component analysis has been used to perform classification.

In [54], a semi-automatic framework is presented for the vehicle detection which minimizes propagation curves. The method is capable of handling seed points in road points.

3.8 Other Methods

In [55], attribute-guided features are used for behavior analysis which classifies the activity based on the image features. Skeleton poses are used as a method for the recognition of human actions.

In [56], the author presents a motion detection algorithm which performs segmentation of images and performs action evaluation.

In [57], a knowledge-based algorithm is presented which classifies the user behavior based on the components available at each class. In [58], a Gaussian model has been presented which uses the background feature and motion vectors. The Bayesian classifier is used for behavior classification.

In [59], the author discusses the presentation of moving objects detection algorithm based on spatiotemporal blocks on infrared videos. The procedure decomposes spatiotemporal blocks using dimensionality decrease method to get a compact vector representation of each block and to overpower the effect of noise.

In [60], the author uses different methods such as background subtraction, statistical method and temporal frame differencing for the recognition of moving objects detection. The background subtraction method is used to notice the moving object from the video frames.

In [4], the author presents a threat detection algorithm which uses temporal flow of events in the detection of threat. The method uses information of context to perform classification. In [61], the bicycles in a video have been detected using object tracking. In [62], the contour-based model is presented which consider the robust statistics. The background features are eliminated and apply the affine transform to perform classification.

In [63], the author put forward the characteristics of video data. A campus surveillance video storage (CSVS) system is proposed. The system consists of three phases: client, metadata information and video data. The metadata information and video data are stored in the database. Depending upon the query, the content is pulled out from the database. The key frames are related to the metadata information to create the storage index. The key frame index is used in lookup operations while querying.

In [64], the author presents a target detection algorithm for ships which removes background using the model, and DCT has been used to extract the features. The three main steps including horizon detection, background modeling and background subtraction. The methods are all based on discrete cosine transform (DCT).

In [65], the author presents an automatic human monitoring system which segments the human body, and based on the movement of body parts, the actions and behaviors are classified. Similarly, in [66], the segmentation is performed to remove the background features. With the remaining features, the method performs classification. In [67], an intelligent approach is presented, which tracks the events of system. The region of interest has been identified to generate the knowledge base.

In [66], the author deals with hybrid background modeling modified adaptive Gaussian mixture model (GMM) and three temporal differencing (TTD). This hybrid method can be used in dynamic environment. The background subtraction is the basic step in video surveillance. The background subtraction is developed to separate the foreground and background images. The technique is efficient and robust for the dynamic environment and reaches good accuracy.

In [68], the author presents a novel method for foreground object detection in real-time applications. The displacement between two successive frames is calculated using a correlation-based corner tracker and background generation method termed as ViBE (Visual Background Extractor).

In [69], an invariant feature-based tracking model is presented. The method extracts the temporal features and performs classification using pattern models of CNN.

In [70], the author defines a method that works on the video that is degraded by dust particle. The system will automatically identify whether the video is degraded by dust or not. Dust particles are frequently stuck on the camera. These dust particles are static in the images of a video sequence captured from a dynamic scene is the main problem. The main aim of the system is to remove the noise in the video frames.

In [71], a behavior analysis model for face has been presented. The method classifies the data based on HDx quantization methods.

In [72], the author presented a monitoring algorithm to provide collision-free navigation to the autonomous vehicles.

In [73], a long short-term memory-based object tracking algorithm is presented. The objects are detected using YOLO and trained with CNN. The classification is performed with LSTM association.

In [74], the author presents a motion detection algorithm which limits the computational cost by representing the features as macro-block motion vectors. Similarly, in [75], a semantic approach is presented by extracting the trajectory and human body parts. The method estimates the semantic weights for different actions to perform classification.

In [76, 77], the author develops an efficient adaptive segmentation algorithm which removes the background by estimating the multiple correlation coefficient. The method performs spatial clustering to perform classification. In [78], the author presents rule-based approaches which consider single and multiple pedestrian interactions.

In [79], action recognition is done by using the volumetric features. In this instead of doing background subtraction on the images, over-segmentation of the video is done. Then, complementary nature of shape- and flow-based features for action recognition is performed in this paper.

4 Performance Analysis

The performance of different methods has been evaluated according to the data set Weizmann. The "Weizmann" data set has been used for the evaluation of various algorithms. The data set is the collection of different actions which belongs to nine actors. The videos of the data set are captured in the frontal view. The data set has been used for the evaluation of different algorithms.

Table 1 presents the comparative study on the methods and results of various methods.

5 Conclusion

This paper presented a detailed review on the methods of behavior analysis on video surveillance systems. However, there are a number of methods available and discussed, not all of them are efficient in behavior analysis. The methods have used different features in identifying the objects and tracking them. Apart from other features, the shape features are the only efficient feature which can represent the

Table 1 Comparative analysis on various methods

Method/feature used	Advantage	Disadvantage
HMM/images converted to vectors of sequences	Classification is moderate efficient	Higher time complex
AHMM/relationship between peoples	Low accuracy in classification	Higher false rate
HMM/skeleton features	Good accuracy in classification	Higher time complexity
Fuzzy/contour and shape feature	Moderate accurate	Much false ratio
Fuzzy/geometric, kinematic feature	Moderate accuracy	Time complexity is higher
Fuzzy/shape	Efficient	Higher time complexity
SVM/optical flow movements	Efficient in recognizing human actions	Poor in other actions
SVM/optical flow motion descriptor	Moderate in behavior classification	Produces false ratio with higher time complexity
SVM and directed cyclic graph/sparse features	Produces efficient classification	Higher time complexity
SVM/discrete Coxian distribution	Produces moderate classification accuracy	Higher time complex
CNN/skeletal joint point	Higher irrelevancy in identifying the points.	Higher false ratio
ANN/appearance features	Efficient classification	False classification is noticeable
SOM/toxic	Moderate efficient in classification	Higher false classification
SOM/shape features	Moderate accuracy	Higher time complexity
SOM/shape features	Moderate in classification	Higher false classification

behavior of the person. By considering shape feature with geometric variation like shearing, scaling and invariant features, the performance on behavior analysis can be performed.

References

1. Zhu J (2010) Object tracking in structured environments for video surveillance applications. IEEE Trans Circuits Syst Video Technol 20(2)
2. Kim DY, Vo BN (2019) A labeled random finite set online multi-object tracker for video data. Pattern Recognit 90:377–389
3. Nguyen TT, Nguyen TTT (2019) Multi-label classification via label correlation and first order feature dependence in a data stream. Pattern Recognit 90:35–51

4. Bhargava M, Chen C-C, Ryoo MS, Aggarwal JK (2009) Detection of object abandonment using temporal logic. Springer, Berlin, pp 271–281
5. Allili MS (2009) Effective object tracking by matching object and background models using active contours. In: 16th IEEE international conference on image processing (ICIP)
6. Naiel MA, Ahmad MO, Swamy MNS, Lim J, Yang MH (2017) Online multi-object tracking via robust collaborative model and sample selection. Comput Vis Image Underst 154:94–107
7. Wang L, Liu T, Wang G, Chan KL, Yang Q (2015) Video tracking using learned hierarchical features. IEEE Trans Image Process 24(4):1424–1435
8. Kannappan S, Liu Y, Tiddeman B (2019) DFP-ALC: automatic video summarization using distinct frame patch index and appearance based linear clustering. Pattern Recognit 120:8–16
9. Kim T, Lee S, Paik J (2011) Combined shape and feature-based video analysis and its application to non-rigid object tracking. IET Image Process 5:87–100
10. Kane L, Khanna P (2019) Depth matrix and adaptive Bayes classifier based dynamic hand gesture recognition. Pattern Recognit 120:24–30
11. Saggese A, Strisciuglio N (2019) Learning skeleton representations for human action recognition. Pattern Recognit 118:23–32
12. Paulo CF, Correia PL (2007) Automatic detection and classification of traffic signs. In: IEEE, Eighth international workshop on image analysis for multimedia interactive services, WIAMIS
13. Caroline Rougier C, Meunier J, St-Arnaud A, Rousseau J (2011) Robust video surveillance for fall detection based on human shape deformation. IEEE Trans Circuits System Video Technol 21(5):611–622
14. Rin V, Nuthong C (2019) Front moving vehicle detection and tracking with Kalman filter. In: 4th international conference on computer and communication systems
15. Gajjar V, Gurnani A, Khandhediya Y (2017) Human detection and tracking for video surveillance: a cognitive science. IEEE-2017
16. Banerjee P, Sengupta S, Human motion detection and tracking for video surveillance
17. Seemanthini K, Manjunath SS (2018) Human detection and tracking using HOG for action recognition. Procedia Comput Sci 132:1317–1326
18. Dewan MAA, Hossain MJ, Chae O (2009) Background independent moving object segmentation for video surveillance. IEICE Trans Commun 4633(2):585–598
19. Kim JS, Yeom DH, Joo YH (2011) Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems. IEEE Trans Consum Electron 57(3):1165–1170
20. Vishwakarma S, Agrawal A (2012) A survey on activity recognition and behavior understanding in video surveillance. Vis Comput 29:983–1009
21. Yamato J (2002) Recognizing human behavior using hidden markov models. Video Comput, 90–131
22. Lin W, Sun M-T, Poovendran R, Zhang Z (2010) Group event detection with a varying number of group members for video surveillance. IEEE Trans Circuits System Video Technol 20(8):1057–1067
23. Chaaraoui AA, Padilla Lopez JR (2014) A vision-based system for intelligent monitoring: human behaviour analysis and privacy by context. Sensors 14:8895–8925
24. Vinayak Kale G, Patil VH (2016) A study of vision based human motion recognition and analysis. Int J Ambient Comput Intell 7(2)
25. Farzad A (2015) Recognition & classification of human behavior in intelligent surveillance systems using hidden markov model. Int J Image Graph Signal Process 12:31–38
26. Hsieh J-W, Hsu YT (2008) Video-based human movement analysis and its application to surveillance systems. IEEE Trans Multimedia 10:372–384
27. Mathur G, Bundele M (2016) Research on intelligent video surveillance techniques for suspicious activity detection critical review. In: 2016 international conference on recent advances and innovations in engineering (ICRAIE)
28. Liu X, Liu ZQ (2002) Detection of vehicles from traffic scenes using fuzzy integrals. Pattern Recognit 35(4):967–980

29. Chebi H, Acheli D, Kesraoui M (2017 Apr) Strategy of detecting abnormal behaviors by fuzzy logic. In: 2017 intelligent systems and computer vision (ISCV)
30. Sokolova MV (2013) A fuzzy model for human fall detection in infrared video. J Intell Fuzzy Syst 24(2):215–228
31. Abdelhedi S (2015) Logic based human activity recognition in video surveillance applications. In: Proceedings of the second international Afro-European conference for industrial advancement AECIA 2015
32. Elbouz M, Ayman A, Brosseau C (2011) Fuzzy logic and optical correlation-based face recognition method for patient monitoring application in home video surveillance. Optic Eng Spie 50(6):067003
33. Xu H, Li L, Fang M (2018) Movement human actions recognition based on machine learning. Int J Online Eng (IJOE) 14(04):193
34. Divya J (2013) Automatic video based surveillance system for abnormal behavior detection. Int J Sci Res (IJSR)
35. Miao Y, Song J (2014) Abnormal event detection based on SVM in video surveillance. In: 2014 IEEE workshop on advanced research and technology in industry applications (WARTIA)
36. Huang K, Wang S, Tan T, Maybank SJ (2009) Human behavior analysis based on a new motion descriptor. IEEE Trans Circuits Syst Video Technol 19(12)
37. Huynh T, Banos O (2015) Traffic behavior recognition using the pachinko allocation model. Sensors 15(7):16040–16059
38. Duong TV, Bui HH, Phung DQ, Venkatesh S (2005) Activity recognition and abnormality detection with the switching hidden semi-markov model. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition (CVPR), San Diego, Ca, USA, 20–25, vol 1, 838–845
39. Shareef IR (2016) Design and implementation of smart security system based on artificial neural network. J Eng Appl Sci 11(9)
40. Rivas A, Chamoso P (2018) Detection of cattle using drones and convolutional neural networks. Sensors (Basel). 18(7):2048
41. Yu B (2017) Design and implementation of behavior recognition system based on convolutional neural network
42. Jan T, Piccardi M (2003) Neural network classifiers for automated video surveillance. In: IEEE explore, neural network for signal processing
43. Muhammad K (2018) Convolutional neural networks based fire detection in surveillance videos. IEEE Multimedia Anal, 18174–18183
44. Suriani NS, Ahnad SNFSNF (2018) Human activity recognition based on optimal skeleton joints using convolutional neural network. J Eng Sci Technol Special Issue on ICCSIT 2018, 48–57 (2018)
45. Hu J, Lu J, Tan YP (2016) Deep metric learning for visual tracking. IEEE Trans Circuits Syst Video Technol 26(11):2056–2068
46. Fan J, Xu W, Wu Y, Gong Y (2010) Human tracking using convolutional neural networks. IEEE Trans Neural Netw 21(10):1610–1623
47. Katanyukul T, Ponsawat J (2017) Customer analysis via video analytics: customer detection with multiple cues. Acta Polytech Hung 14(3)
48. Xia C, Fu L (2017) Aquatic toxic analysis by monitoring fish behavior using computer vision: a recent progress. J Toxicol 2018, Article Id 2591924:11 p
49. Alias UF, Ahmad NB (2015 Dec) Student behavior analysis using self-organizing map clustering technique. J Eng Appl Sci 10(23)
50. Chaitra BH, Anupama HS (2013) Human action recognition using image processing and artificial neural networks. Int J Comput Appl 80(9):0975–8887
51. Xiao Y-D, Clauset A, Harris R (2005) Supervised self-organizing maps in drug discovery. 1. Robust behavior with over determined data sets. J Chem Inf Model
52. Lozano A, Manfredi G, Nieddu L (2009) An algorithm for the recognition of levels of congestion in road traffic problems. Math Comput Simul 79:1926–1934

53. Zhou J, Gao D, Zhang D (2007) Moving vehicle detection for automatic traffic monitoring. IEEE Trans Veh Technol 56:51–59
54. Niu X (2006) A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model. J Photogr Remote Sens (ISPRS) 61:170–186
55. Anzalone SM, Xavier J (2019) Quantifying patterns of joint attention during human-robot interactions: an application for autism spectrum disorder assessment. Pattern Recognit 118:42–50
56. Bandera JP, Marfil R (2019) A new paradigm for autonomous human motion description and evaluation: application to the get up & go test use case. Pattern Recognit 18:51–60
57. Georis B, Brémond F, Thonnat M (2007) Real-time control of video surveillance systems with program supervision techniques. Mach Vis Appl 18(3–4):189–205
58. Zhang W, Fang XZ, Yang X (2006) Moving vehicles segmentation based on bayesian framework for Gaussian motion model. Pattern Recognit Lett 27(1):956–967
59. Pokrajac D, Zeljkovic V, Latecki LJ (2005) Spatial-temporal algorithm for moving objects detection in infrared video sequences. IEEE, New York
60. Bojkovic Z, Samcovic A, Turan T (2005) Object detection and tracking in video surveillance systems. Trondheim, Norvegia, 113–116
61. Beran V, Herout A, Reznicek I (2009) Video-based bicycle detection in underground scenarios. In: Proceedings of WSCG'09, 4
62. Araki A, Matsuoka T, Yokoya N, Takemura H (2001) Real-time tracking of multiple moving object contours in a moving camera image sequence. IEICE Trans Inf Syst E83-D(7):1583–1591
63. Ma S, Chen X, Li Z, Yang Y (2018) A retrieval optimized surveillance video storage system for campus application scenarios. J Electr Comput Eng, Article Id 3839104
64. Zhang Y, Li QZ, Zang FN (2017) Ship detection for visual maritime surveillance from non-stationary platforms. Ocean Eng 141:53–63
65. Srinivasan K, Porkumaran K, Sainarayanan G (2014) Background subtraction techniques for human body segmentation in indoor video surveillance. J Sci Ind Res 73(5):342–345
66. Niranjil KA, Sureshkumar C (2015) Background subtraction in dynamic environment based on modified adaptive GMM with TTD for moving object detection. J Electr Eng Technol 10(1):372–378
67. Lim MK, Tang S, Chan CS (2014) Surveillance: intelligent framework for multiple events detection in surveillance videos. Expert Syst Appl 41(10):4704–4715
68. Kryjak T, Komorkiewicz M, Gorgon M (2014) Real-time implementation of foreground object detection from a moving camera using the vibe algorithm. Comput Sci Inf Syst 11(4):1617–1637
69. Kuen J, Lim KM, Lee CP (2015) Self-taught learning of a deep invariant representation for visual tracking via temporal slowness principle. Pattern Recognit 48(10):2964–2982
70. Chen L, Zhu D, Tian J, Liu J (2016) Dust particle detection in traffic surveillance video using motion singularity analysis. Digit Signal Process 58:127–133
71. De La Torre M, Granger E, Sabourin R, Gorodnichy DO (2015) Adaptive skew-sensitive ensembles for face recognition in video surveillance. Pattern Recognit 48(11):3385–3406
72. Savkin AV, Wang C (2017) A framework for safe assisted navigation of semi-autonomous vehicles among moving and steady obstacles. Robotica 35(5):981–1005
73. Jiang M-X, Deng C, Pan ZG (2018) Multi object tracking in videos based on LSTM and deep reinforcement learning. Complexity 2018, Article Id 4695890, 12 p
74. Kiryati N, Riklin T, Ivanchenko Y, Rochel S (2008) Real-time abnormal motion detection in surveillance video. In: IEEE 19th international conference on pattern recognition, 1–4
75. Lao W, Han J, Dewith PHN (2010) Flexible human behavior analysis framework for video surveillance applications. Int J Digit Multimedia Broadcast Article Id 920121:1–9
76. Murali S, Girisha R (2009) Segmentation of motion objects from surveillance video sequences using temporal differencing combined with multiple correlations. Int Adv Video Signal-Based Surveill, 472–477

77. Lian X, Zhang T, Liu Z (2010) A novel method on moving-objects detection based on background subtraction and three frames differencing. In: International conference on measuring technology and mechatronics automation, 252–256
78. Yoshimitsu Y, Naito T, Fujimura K, Kamijo S (2010) Behavior understanding at railway station by association of locational semantics and postures. In: IEEE international conference on systems man and cybernetics, 3033–3038
79. Ke Y, Sukthankar R, Hebert M (2007) Spatio-temporal shape and flow correlation for action recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Minneapolis, MN, USA, 17–22, 1–8

Analysis of Different Substrate Material on Wearable Antenna for ISM Band Applications

S. Bhavani and T. Shanmuganantham

Abstract Recently, wearable antennas are become popular due to the miniaturization of wireless devices. The main advantage of using wearable antennas is that they are part of clothing and easily transmit and receive signals through clothes. These antennas play a vital role in many applications like navigation, tracking, health monitoring, physical training, RFID, medicine, and military. This paper describes the design and development of microstrip patch antenna with different fabric materials as substrates for ISM band applications. The performance characteristics like gain, directivity, return loss, VSWR, and radiation characteristics are simulated and measured by using CST.

Keywords Wearable antenna · ISM band · Microstrip patch antenna · CST

1 Introduction

Portable devices like mobiles and tablets are becoming part of human life due to their advance features. Due to rapid changes in technology, the size and visibility of devices decrease gradually. In the future, human activities are controlled by sensors and devices to monitor various requirements of human including medical conditions. These devices and sensors can communicate with each other and with the outside world also. This type of advancement in communication is possible only by using wearable devices and antennas [1]. To design such antennas, globally unlicensed 2.45 GHz ISM band is suitable.

Wearable antenna can transmit and receive the bio-signals from the human body to device [2]. Wearable antenna should be flexible, hidden, low profile, and it cannot harm to human body. This requires the perfect integration of antenna with daily wearing fabric. The performance of antenna depends on how it should be fit in compact

S. Bhavani (✉) · T. Shanmuganantham
Department of Electronics Engineering, Pondicherry University, Puducherry 605014, India
e-mail: ammav10@gmail.com

T. Shanmuganantham
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_63

area of device. Other than telecommunications, the band that can be internationally reserved for industrial, scientific, and medical applications is ISM band. Authors have strived in designing a wearable textile antennas for biomedical applications like signal monitoring application, telemedicine applications [3], skin cancer detection [4], and for bio-information applications [5].

Microstrip patch antenna is preferred for any wearable application due to easy integration with clothing [6]. Microstrip is a transmission line to transmit microwave frequency signals. It consists of ground layer, substrate layer, and patch can be fabricated on the printed circuit board. Above the ground plane, the substrate is placed to separate the radiating metallic strip from the ground plane [7]. These are also referred as patch antennas. There are many shapes of patches which are available, namely square, rectangular, dipole, circular, elliptical, circular rings, or ring sector [8]. Microstrip is less expensive than existed technology, as well as less in weight and more compact. Due to low fabrication cost, these are more amenable to mass production.

The structure of rectangular patch antenna is shown in Fig. 1.

The shape of patch may be different according to their dimension as dipole, square, rectangular, circular, triangular, circular ring, etc., as shown in Fig. 2 [9]. The second layer is substrate which is middle part and placed on ground of patch antenna. Substrates are used to enhance the antenna radiation capability.

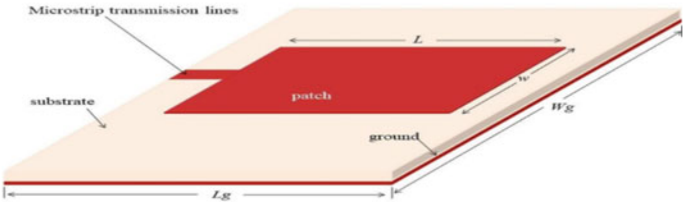


Fig. 1 Rectangular microstrip patch antenna

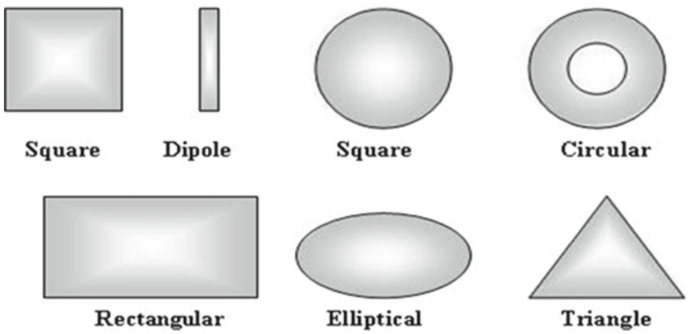


Fig. 2 Common available shapes of microstrip

2 Design and Methodology

The geometry of antenna related to this paper is shown in Fig. 3 and consists of ground plane and rectangular patch with substrate and feeder. The role of feeding is very significant in case of efficient operation of antenna to rectify the antenna impedance matching. There are two types of feed techniques known as contact and non-contact feed methods. Different feed techniques are mentioned in [10].

Choosing of feed point location decides the perfect impedance matching to minimize the return loss of the antenna [8]. The design specifications of antenna are given in Table 1. The dimensions of patch are computed based on microstrip patch radiator. The width of the patch (W) is calculated using the following formula [11].

$$W = \frac{C}{2\mathrm{fr}}\sqrt{\frac{2}{\mathrm{er}+1}} \tag{1}$$

where c is the velocity of EM wave, fr is resonance frequency, and er is the dielectric constant of fabric substrate.

The strip lies between air and dielectric so we have considered effective dielectric constant which is calculated using the following formula:

$$\mathrm{ereff} = \frac{\mathrm{er}+1}{2} + \frac{\mathrm{er}-1}{2} + \left[1 + 12\frac{h}{W}\right] - \frac{1}{2}, \quad w/h > 1 \tag{2}$$

where h is the height of the substrate. The effective length of patch is given by

$$L_{\mathrm{eff}} = \frac{C0}{2\mathrm{fr}}\sqrt{\frac{2}{\mathrm{ereff}}} \tag{3}$$

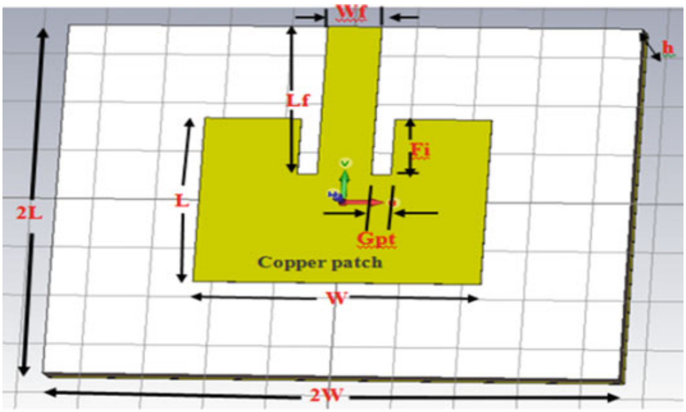


Fig. 3 Microstrip rectangular patch antenna

Table 1 Design specifications of antenna

Substrate material	Dielectric constant εr	Length of patch Lp (mm)	Width of patch Wp (mm)	Loss tangent
RT Duroid 5880	2.2	40.46	48.37	0.0004
GML 1000	3.2	33.7	42.22	0.002
RO4003	3.4	32.71	41.25	0.002
FR-4	4.4	28.81	37.23	0.0025
Jeans	1.7	45.851	52.693	0.078
Lycra	1.5	48.68	54.76	0.0093
Rubber	3	34.81	43.29	
Polyester	3.2	33.7	42.22	0.0045
Taconic TLC	3.2	34.483	43.129	0.002
Bakelite	4.8	27.6	35.93	0.0304
Arlon AD 250	2.5	38.02	46.25	0.0018
Wash cotton	1.51	48.20	54.20	0.0004
Resin	3.5	32.25	40.79	0.0085
Cordura	1.90	42.9	50.2	0.0098
Cotton	1.60	46.2	53.0	0.0040
Quartzel fabric	1.95	42.91	50.41	0.0004
Curtain cotton	1.47	49.1	55.0	0.0004
Teflon	2.1	41.4	49.17	0.0002

Due to the effect of fringing fields, the additional line length is calculated using the following formula

$$\Delta L = 0.412h\frac{(\text{ereff} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\text{ereff} - 0.258)\left(\frac{W}{h} + 0.8\right)} \tag{4}$$

The actual length of patch is given by

$$\text{Lp} = L_{\text{eff}} - 2\Delta L \tag{5}$$

3 Simulations and Results

Microstrip patch antenna is designed and simulated using computer simulation technology (CST) microwave studio software. For antenna design, copper material is used as ground plane and patch. The substrate plays an important role in antenna design

[12]. Wearable antennas must compatible with human clothes to meet the requirements of human being [13, 14]. The antenna is designed with different substrates for observing which substrate is more suitable for wearing purpose. RT Duroid 5880, GML 1000, RO4003, FR-4, Jeans, Lycra, Rubber, Polyester, Taconic TLC, Bakelite, Arlon AD 250, Wash cotton, Resin, Cordura, Cotton, Quartzel fabric, curtain cotton, Teflon substrates are used for analyzing the antennas at ISM band. Simulations are done, and different parameter comparison of different substrates is provided in Table 2.

The main characteristics of antenna are return loss, VSWR, gain, and directivity which are measured for antenna by applying different substrates, as shown in Figs. 4, 5, 6, and 7. Antenna with FR-4 substrate results in less return loss −25.5 dB, and VSWR is 1.11. Antenna with wash cotton fabric results in high directivity 8.9 dB. Antenna with curtain cotton fabric results in high gain 8.04 dB. Antenna with resin material gives the next least values of return loss and VSWR. Antenna with Arlon AD 250 has nearly similar values of gain and directivity.

Directivity decides how much power concentrated by antenna in a given direction. If the directivity is high, then the concentration of antenna is more and beam will travel longer. Generally, directivity is taken with respect to isolated antenna. Antenna with different substrates results in various directivity values that are shown in Fig. 4. From the graph, antenna with wash cotton radiates more compared to others. Antenna with

Table 2 Comparison of various substrates

Substrate material	Directivity (dB)	Return loss (dB)	VSWR	Gain (dB)
RT Duroid 5880	7.84	−10.39	1.866	6.86
GML 1000	7.15	−13.3	1.55	5.67
RO4003	6.98	−12.88	1.587	5.58
FR-4	6.7	−25.5	1.11	3.23
Jeans	8.5	−16.2	1.3	6.4
Lycra	8.84	−18.03	1.31	6.6
Rubber	7.25	−11.1	1.76	6.2
Polyester	5.61	−15.4	1.40	2.21
Taconic TLC	7.19	−13.2	1.553	5.8
Bakelite	6.59	−16.49	1.352	5.26
Arlon AD 250	7.8	−11.6	1.7	7.7
Wash cotton	8.9	−11.3	1.743	7.95
Resin	7.82	−21.0	1.20	4.29
Cordura	8.2	−18.03	1.2	5.78
Cotton	8.6	−12.5	1.6	7.11
Quartzel fabric	8.1	−11.5	1.77	7.1
Curtain cotton	8.8	−11.2	1.75	8.04
Teflon	8.0	−11.4	1.73	7.07

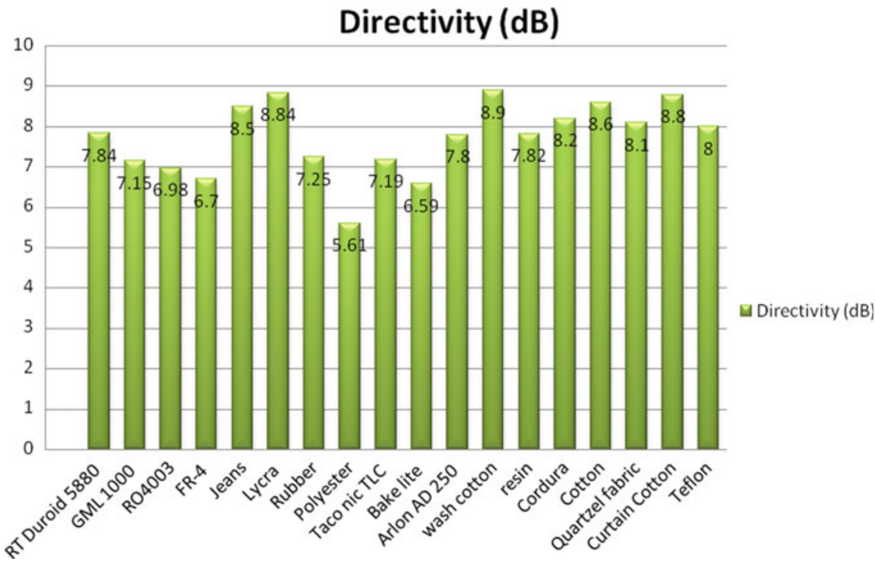


Fig. 4 Directivity comparison graph

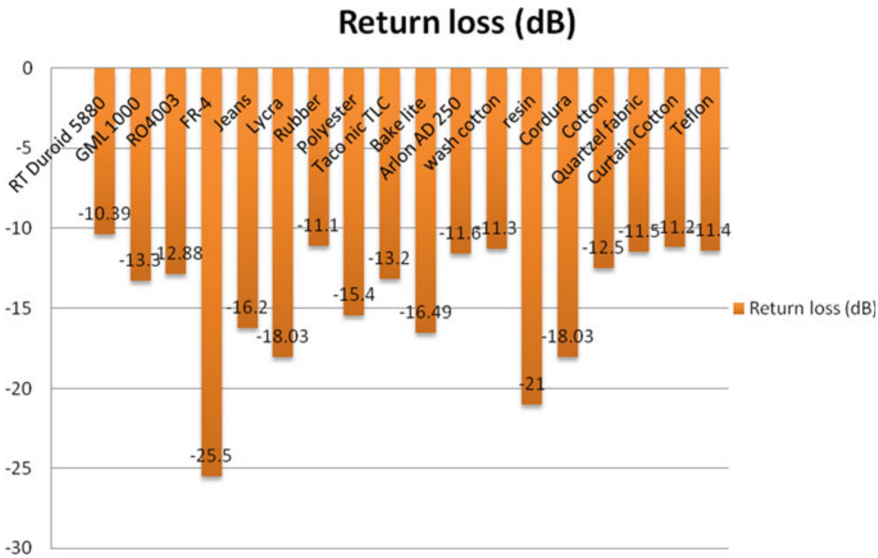


Fig. 5 Return loss comparison graph

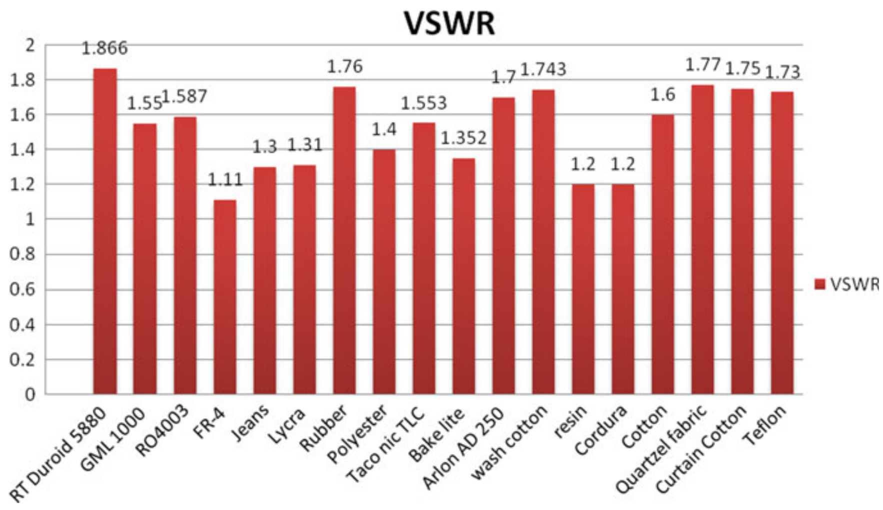


Fig. 6 VSWR comparison graph

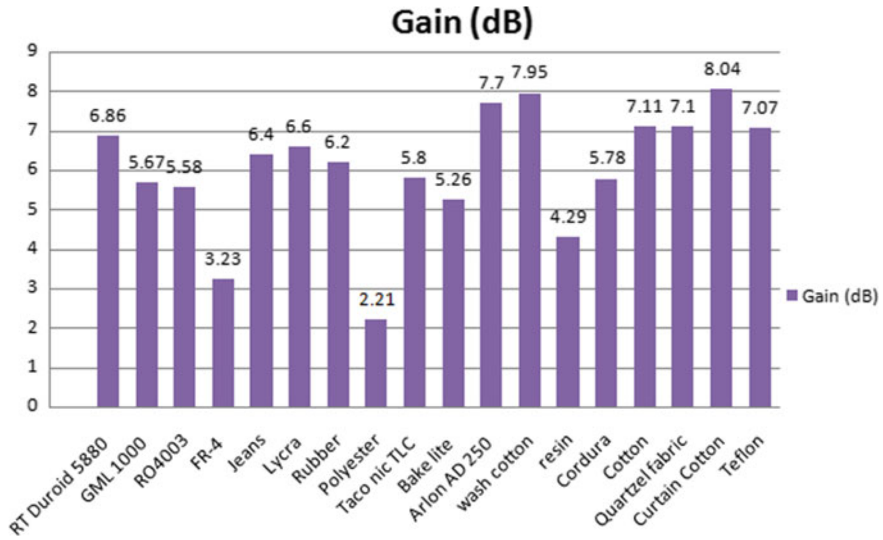


Fig. 7 Gain comparison graph

cotton, Lycra, curtain cotton nearly equally radiates the power in given direction. But antenna with polyester substrate results in poor directivity and indicates less concentration of antenna in desired direction.

Return loss and VSWR are related to measure how much power is reflected toward the antenna due to imperfections or mismatches in the load. If the value of the VSWR is high and indicates high mismatch, then the return loss and VSWR values of antenna

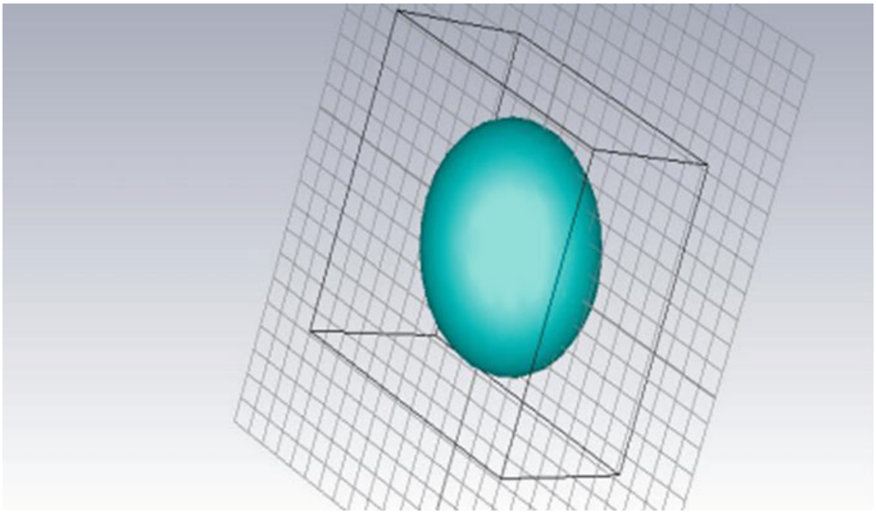


Fig. 8 Head model

using different substrates are compared in Figs. 5 and 6. From the figures, antenna with Fr-4 substrate exhibits low return loss and low VSWR. Antenna with RT Duroid substrate is having more return loss and VSWR, which indicates high mismatching and occurs by using this substrate.

The antenna gain describes the amount of power transmitted in the direction of peak radiation compared to an isotropic source. In case of directivity, losses are neglected but in case of gain calculation losses are considered. Gain of the antenna decides the capability of antenna to convert input power into radio waves or radio waves into electrical signal. The comparison of gain of antenna by using different substrates is shown in Fig. 7. Antenna with wash cotton substrate results in high gain, whereas by using polyester substrate results in low gain.

SAR calculations are done by developing head model and place the designed antenna near to head model. Figure 8 shows the head model of 200 mm in the tool, and Fig. 9 shows the SAR distribution to head model at 2 mm (Table 3).

The specific absorption rate (SAR) value is low for antenna by using RT5880 substrate, and it is in limit specified by FCC standard which is 1.6 W/kg.

4 Conclusion

Design and simulation of various wearable antennas for ISM band have been done using CST Microwave Studio. Wearable antennas can be widely used in many applications like medical, textile, and industrial. Based on the application, substrate material can be varied for each antenna. All the compared antennas are operated in the

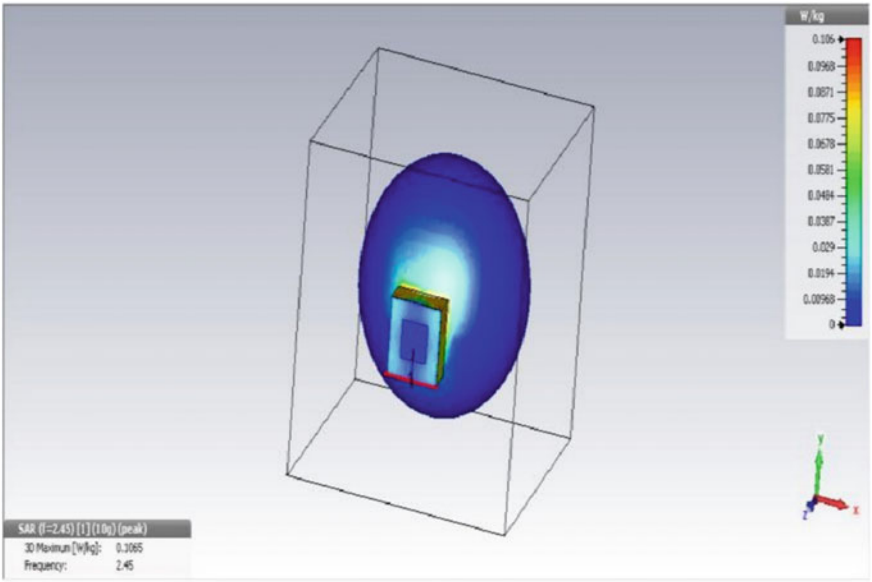


Fig. 9 SAR distribution to head model at 2 mm

Table 3 SAR comparison for different substrates

Substrate	1 g SAR (W/kg)	10 g SAR (W/kg)
FR4	0.010	0.206
RT5880	0.005	0.0784
Cotton	0.0640	0.085
Jeans	0.09	0.04

frequency of 2.45 GHz. From the simulation results by considering return loss and VSWR Fr-4 is preferable and by gain consideration curtain cotton is better among other antennas. In directivity point of view, wash cotton is most suitable. Substrate is selected based on the application and system requirements.

5 Future Work

The presented work is only at basic level of antenna design to get an idea on different substrates, and how they affect the antenna parameters. The future work can be done by designing a wearable antenna with better results and providing good efficiency, better bandwidth, and low SAR value for a specific biomedical application.

References

1. Za'aba AS, Ibrahim SN, Malek NFA, Ramly AM (2017) Development of wearable patch antenna for medical application. In: IEEE regional symposium on micro and nanoelectronics (RSM)
2. Devamalar PMB, Bai VT, Kavya G (2016 Feb) Wearable antenna for bio signal monitoring applications using ISM band. In: Proceedings of international conference on advances in electrical, electronics, information, communication and bioinformatics (AEEICB16), pp 32–34. https://doi.org/10.1109/aeeicb.2016.7538320, 978-1-4673-9745-2 ©2016IEEE
3. Christina GC, Rajeswari A, Lavanya M, Keerthana J, Ilamathi K, Manoranjitha V (2016 Apr) Design and development of wearable antennas for tele-medicine applications. In: Proceedings of international conference on communication and signal processing, pp 2003–2037. 978-1-5090-0396-9/16/$31.00 ©2016 IEEE
4. Kaur NA (2016 Oct) Wearable antennas for skin care detection. In: International conference on next generation computing technologies (NGCT-2016) Dehradun, India, pp 197–201. https://doi.org/10.1109/ngct.2016.7877414, 978-1-5090-32570/16/31.00$ ©2016 IEEE
5. Pei R, Wang J, Leach M, Wang Z, Lee S, Lim EG (2016) Wearable antenna design for bio information. In: Proceedings of IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB), pp 1–4
6. Tanaka M, Jang JH (2003) Wearable microstrip antenna. In: Proceedings of IEEE APS international symposium and URSI North American radio science meeting, Columbus, OH, USA, June 2003
7. Carver KR, Mink JW (1981 Jan) Microstrip antenna technology. IEEE Trans Antennas Propag AP-29(1):21
8. Bakibillah ASM, Hossain MS, Roy IS (2014 Dec) Design of a micro strip patch antenna to minimize return loss for WI-MAX application. Int J Adv Res Comput Commun Eng 3(12):8897–8899
9. Locher I, Klemm M, Kirstein T, Troster G (2006) Design and characterization of purely textile patch antennas. IEEE Trans Adv Packag 29(3):777–778
10. Ndujiuba CU, Oloyede AO (2015) Selecting best feeding techniques for a rectangular patch antenna for an applications. Int J Electromag Appl 5(3):99–107
11. Balanis CA (1997) Antenna theory: analysis and design, 2nd edn. Wiley, Asia Pte Ltd., Singapore, 722–736, 869–870
12. Faiji D, Pagel N, Ngo VL, Seidel H (2010) Effect of substrate on reliability and antenna indicator parameters. In: IEEE conference, 1–4 Nov 2010, Kona HI, pp 1342–1345. ISBN-978-1-4244-8170-5
13. Salomen P, Hurme H (2003) Modeling of a fabric GPS antenna for wearable applications. In: Proceedings of IASTED international conference modeling and simulation, vol 1, pp 18–23
14. Salomen P, Hurme H (2003) A novel fabric WLAN antenna for wearable applications. In: Proceedings of IEEE APS international symposium, vol 2, pp 700–703, June 2003

An Analytical Framework for Comparing Flat and Hierarchical Architectures in Fog Computing Networks

Niyas K. Haneefa, S. Pramod, and R. Manivasakan

Abstract With the emergence of IoT, the number of devices that is getting connected is increasing exponentially, which poses a constraint on the lower latency requirements of processing these tasks. Evolving technologies such as fog computing and edge computing consist of computationally lesser exhaustive power servers, which bring this processing near the edge or onto the devices, thereby reducing the round-trip delay as well as the load on the entire network. In this work, we propose the hierarchical arrangement of servers in the fog computing layer for this processing of data and derive the analytical framework for the same. The proposed architecture works on top-down scheduling policy rule for the incoming packets, defined mathematically in terms of two-dimensional Markov chains. The performance of the proposed architecture is compared with an equivalent flat architecture analytically in terms of the mean sojourn time and the mean computational power and justified using simulation results.

Keywords Fog computing · Hierarchical architecture · Markov analysis · omnet++

1 Introduction

The number of devices that are getting connected is increasing enormously, and it is estimated that around 50 billion devices would be connected around 2020 [1–3]. One of the main requirements for the processing of this data is the low latency

N. K. Haneefa (✉) · S. Pramod · R. Manivasakan
Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600036, India
e-mail: niyaskhaneefa@gmail.com

S. Pramod
e-mail: pramodachar769@gmail.com

R. Manivasakan
e-mail: rmani@ee.iitm.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_64

for task computations. Relying on conventional remote cloud solutions may not be advisable always in this scenario because of the round-trip delay associated with the transmission, especially with the huge traffic the network has to handle. The newly evolved technologies such as fog computing and edge computing address this issue and take the processing capabilities near or onto the edge devices [1]. This eventually reduces the latency associated with the processing of the tasks in comparison with a cloud-centric computation environment. Also, the computational requirements needed for the processing of these tasks may not be enormous, which also justifies the elimination of high capacity cloud computing servers. Rather, the tasks can be processed and computed by the fog or edge servers and later if required can be sent back over to the cloud depending on the need. Fog computing enables low latency, distributed computing at or near to the network edge and overcomes some of the limitations of the centralized cloud-centric architecture mainly at the cost of processing power.

In this work of ours, we mainly compare and discuss analytically and then by simulation the two architectures that could be used for the distribution of servers in a fog layer for computation of tasks. The architectures which we compare include the flat arrangement of servers each with the same service times, whereas the second one corresponds to the servers arranged in a tiered/hierarchical fashion with the servers in each tier having different service times. Not many works have addressed this analytical comparison of flat and hierarchical architectures in a fog computing environment. Josilo et al. discuss a decentralized algorithm to offload the tasks to set of devices or to a cloud [4]. Chekired et al. using probabilistic analysis models claim that the hierarchical deployment of fog servers is better for offloading of tasks when a server in a tier is not able to process it [5]. In this paper, we define a top-down scheduling rule for a hierarchical architecture that can be deployed in a fog computing environment, derive the analytical framework using Markov chains for a (3,1) hierarchical architecture and compare between the two architectures in terms of the mean sojourn time and the mean computational power. Finally, we do a comparison between the two architectures using simulations and verify the results obtained analytically.

2 Network Model and Contributions of Our Work

Figure 1 gives an example for the network model which we consider here. The fog computing layer sits between the cloud layer and the devices layer. In our work, we mainly focus on the analytical derivations and simulations of the processing of tasks by the fog servers in this fog computing layer for two different architectures and then make the comparison between the two.

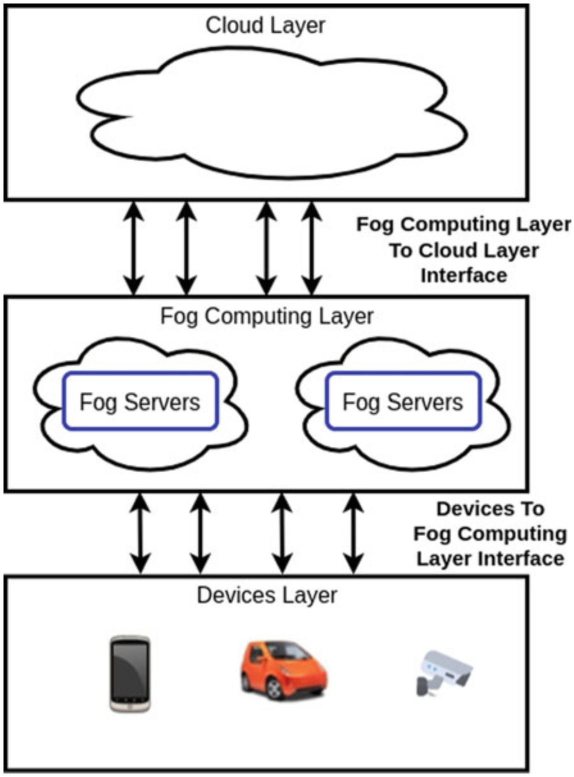


Fig. 1 Example model involving fog architecture

2.1 Contributions

The below points list the contributions of our work with points [1–5] relating to the analytical work and points [6–7] relating to the simulations:

1. We define a top-down scheduling rule for processing the incoming arrival tasks for a hierarchical architecture in a fog computing environment and define a (3,1) two-tiered server arrangement using this rule.
2. Using the above rule, we draw the Markov chain state transition diagram and derive the steady-state stability equations.
3. The steady-state probabilities for each of the defined states are computed from the steady-state stability equations.
4. We define an equivalent flat topology to be compared with the defined (3,1) hierarchical architecture.
5. The hierarchical and flat topologies are compared analytically in terms of the mean sojourn times and the mean computational power.
6. We perform the simulations for the defined flat and hierarchical topologies using omnet++.

7. Finally, we prove that the analytical comparisons we carried out are justifiable with the simulation results, and we conclude by making our final inferences from the results about the two architectures.

3 Analytical Framework and Mathematical Analysis

Below section details the analytical derivations and mathematical analysis for the fog computing layer defined in Sect. 2. We consider and explain the arrangement of fog servers in a hierarchical architecture defined by the top-down scheduling policy followed by the equivalent flat architecture for comparison with the former.

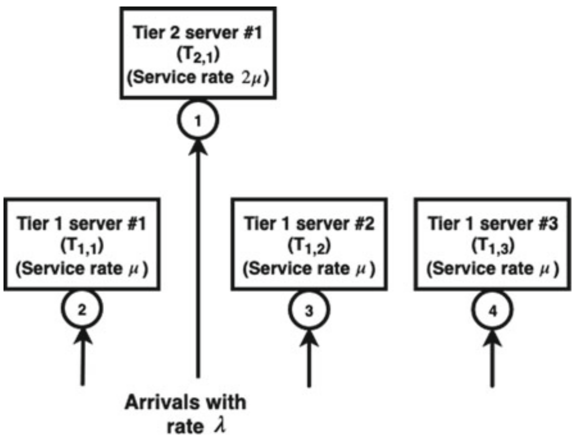
3.1 *Hierarchical Architecture*

Figure 2 gives the hierarchical architecture with three tier-1 servers and one tier-2 server which we have considered in this work, denoted as (3,1).

Definition—Top-Down Scheduling Rule: Let 'S' denote the set of all possible states the servers in the hierarchical architecture can occupy. The term 'state' indicates the occupancy status of a server, i.e. whether it is busy or idle. Let S_k denote the state of the servers in the kth time instant, $S_k \in S$. In the hierarchical architecture we considered as in Fig. 2, $S_k = (i, j)$ where i = number of tier-1 busy servers, $0 \leq i \leq 3$ and j = number of servers tier-2 busy servers, $0 \leq j \leq 1$.
Assumptions:

- The arrival rates denoted by λ are exponentially distributed.
- Tier-1 servers have a service rate of μ and tier-2 servers 2μ.

Fig. 2 Hierarchical Two-Tier Architecture



- Any new job arrivals are checked for the availability of a tier-2 server, if not it is allotted to a tier-1 server if it is available.
- If none of the servers is available, the job is discarded.
- We have assumed a buffer size of zero in our work.

3.2 Analysis

Figure 3 gives the Markov chain state diagram of the hierarchical architecture considered in Fig. 2. The green arrows indicate the incoming arrivals with an arrival rate of λ, while the rest of all coloured arrows indicate the service rates at all the states. The steady-state equations at each of the state nodes for the Markov chain diagram in Fig. 3 are given by (1)–(9).

Steady-State Equations:

$$\text{Stability equation at state (0,0)} \Rightarrow \lambda.\pi_{0,0} = 2\mu.\pi_{0,1} + \mu.\pi_{1,0} \tag{1}$$

$$\text{Stability equation at state (0,1)} \Rightarrow (\lambda + 2\mu)\pi_{0,1} = \lambda.\pi_{0,0} + \mu.\pi_{1,1} \tag{2}$$

$$\text{Stability equation at state (1,1)} \Rightarrow (\lambda + \mu).\pi_{1,1} = \lambda.\pi_{0,1} + 2\mu.\pi_{2,1} \tag{3}$$

$$\text{Stability equation at state (2,1)} \Rightarrow (\lambda + 2\mu)\pi_{2,1} = \lambda.\pi_{1,1} + 3\mu.\pi_{3,1} \tag{4}$$

$$\text{Stability equation at state (3,1)} \Rightarrow (5\mu).\pi_{3,1} = \lambda.\pi_{2,1} + \lambda.\pi_{3,0} \tag{5}$$

$$\text{Stability equation at state (3,0)} \Rightarrow (\lambda + 3\mu)\pi_{3,0} = 2\mu.\pi_{3,1} \tag{6}$$

$$\text{Stability equation at state (2,0)} \Rightarrow (\lambda + 2\mu)\pi_{2,0} = 2\mu.\pi_{2,1} + 3\mu.\pi_{3,0} \tag{7}$$

$$\text{Stability equation at state (1,0)} \Rightarrow (\lambda + \mu)\pi_{1,0} = 2\mu.\pi_{1,1} + 2\mu.\pi_{2,0} \tag{8}$$

$$\text{Also,} \quad \sum_{i,j \in S} \pi_{i,j} = 1 \tag{9}$$

where $\pi_{i,j}$ corresponds to the steady-state probabilities for state i of tier-1 servers and state j for the tier-2 server. Solving (1)–(9) gives the steady-state probabilities of the states as follows:

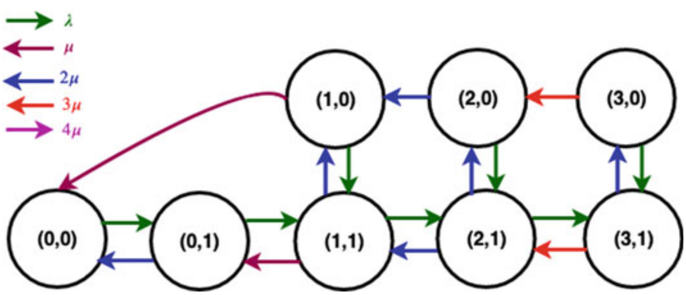


Fig. 3 Markov Chain Diagram

$$\pi_{(0,0)} = \frac{12\left(\lambda^3\mu^3 + 8\lambda^2\mu^4 + 19\lambda\mu^5 + 10\mu^6\right)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{10}$$

$$\pi_{(0,1)} = \frac{6\lambda(\lambda + 2\mu)\left(\lambda^2\mu^2 + 2\lambda\mu^3 + 5\mu^4\right)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{11}$$

$$\pi_{(1,1)} = \frac{6\lambda^2(\lambda + 2\mu)\left(2\lambda\mu^2 + 5\mu^3\right)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{12}$$

$$\pi_{(2,1)} = \frac{3\lambda^3\mu(\lambda + 2\mu)(\lambda + 5\mu)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{13}$$

$$\pi_{(3,1)} = \frac{\lambda^4(\lambda + 2\mu)(\lambda + 3\mu)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{14}$$

$$\pi_{(3,0)} = \frac{2\lambda^4\mu(\lambda + 2\mu)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{15}$$

$$\pi_{(2,0)} = \frac{6\lambda^3\mu^2(2\lambda + 5\mu)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{16}$$

$$\pi_{(1,0)} = \frac{24\lambda^2\left(2\lambda\mu^3 + 5\mu^4\right)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{17}$$

The computed stationary probabilities (10)–(17) are used for comparisons with the equivalent flat architecture in Sect. 3.5.

3.3 Compound Round-Robin Algorithm for ‘K’ Tier Architecture

For a generalized ‘K’ tier hierarchical architecture, the compound round-robin algorithm for the top-down scheduling is detailed in Algorithm 1.

Algorithm 1: Compound round-robin algorithm for K tier hierarchical architecture

---

Result: Return available server id starting from the Kth tier till the first tier.
Initialization: N_1, N_2, ...N_K are the number of available servers in each of the ‘K’ tiers ;

```
for i,Tier-K to Tier-1 do
    for n,0 to N_i do
        if Check nth server available then
            return server id 'n';
            // exit the loop and invoke again on arrival
               of the next arrival
        end
    end
end
return -1 // If no server is free, drop the job.
```

Definition: Let ‘K’ denote the number of tiers in the hierarchical fog architecture with N_i; $i = 0, 1, \ldots K - 1$ as the number of servers in the ith tier.

For a new job arrival, the algorithm checks for any available servers in the topmost tier. If it is available, the job is allotted to that server. If not, the algorithm checks for the next tier from the top. This process repeats until the algorithm finds the first idle server starting from the Kth tier till tier-1. If none is available, the job is discarded.

3.4 *Flat Architecture*

Figure 4 gives the equivalent flat architecture which we have used for comparison with the hierarchical architecture in Fig. 2.

Assumptions:

- The arrival rates denoted by λ are exponentially distributed.
- Only tier-1 servers are present with all the servers having the same service time of μ.
- The states are defined by the notations $S_0 - S_5$ with the subscripts indicating the number of servers occupied at a particular time instant.
- (18) gives the generalized steady-state probability equation at each of the states for a flat architecture.

$$\lambda.\pi_i = (i + 1)\mu.\pi_{i+1} \tag{18}$$

where π_i indicates the steady-state probability at state ‘i’; $i = 0, 1, , \ldots 5$.

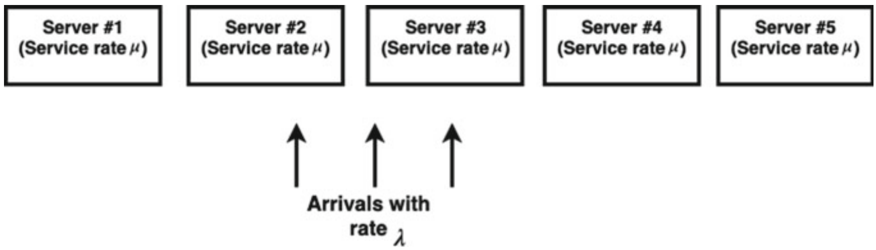


Fig. 4 Flat Architecture equivalent to the hierarchical architecture in Fig. 2

$$\text{Also, } \sum_{i=0}^{5} \pi_i = 1 \tag{19}$$

$$\text{Solving (18) and (19)} \Rightarrow \pi_0 = \frac{1}{26.04\rho^5 + 26.04\rho^4 + 20.83\rho^3 + 12.5\rho^2 + 5\rho + 1} \tag{20}$$

where the utilization factor $\rho = \frac{\lambda}{(5*\mu)}$.

3.5 Comparison Between Flat and Hierarchical Architecture

In this section, we compare the flat and hierarchical architectures in terms of the performance measures of (1) mean sojourn time and (2) mean computational power.

Sojourn Time: Sojourn time is defined as the amount of time a job spends in the queuing system which is equal to the sum of the waiting time the job has to spend in the queue and the service time a server takes to process it. We have considered zero buffer size for the queue in our work.

Mean Sojourn Time—Flat Architecture: We define $E(T_{S,\text{Flat}})$ as the mean sojourn time taken by the flat architecture servers to process 'N' number of tasks. Then,

$$E(T_{S,\text{Flat}}) = \pi_0.0 + \pi_1.\frac{1}{\mu} + \pi_2.\frac{1}{\mu} + \pi_3.\frac{1}{\mu} + \pi_4.\frac{1}{\mu} + \pi_5.\frac{1}{\mu} \tag{21}$$

where π_i = Probability that server/s remain in state 'i'. Defining the utilization factor ρ as $\rho = \frac{\lambda}{m*\mu}$, m = No. of servers

$$E(T_{S,Flat}) = \frac{5\rho + 12.5\rho^2 + 20.83\rho^3 + 26.04\rho^4 + 26.04\rho^5}{\mu * (1 + 5\rho + 12.5\rho^2 + 20.83\rho^3 + 26.04\rho^4 + 26.04\rho^5)} \tag{22}$$

Mean Sojourn Time—Hierarchical Architecture: We define $E(T_{S,\text{Hier}})$ as the mean sojourn time taken by the hierarchical architecture servers to process the same 'N'' number of tasks as in the flat. Then

$$\begin{aligned} E(T_{S,\text{Hier}}) = {} & \pi_{0,0}.0 + \pi_{0,1}.\frac{1}{2\mu} + \pi_{1,1}.\frac{[\frac{1}{\mu} + \frac{1}{2\mu}]}{2} + \pi_{2,1}.\frac{[\frac{1}{\mu} + \frac{1}{2\mu}]}{2} \\ & + \pi_{3,1}.\frac{[\frac{1}{\mu} + \frac{1}{2\mu}]}{2} + \pi_{3,0}.\frac{1}{\mu} + \pi_{2,0}.\frac{1}{\mu} + \pi_{1,0}.\frac{1}{\mu} \end{aligned} \tag{23}$$

$$E(T_{S,\text{Hier}}) = \frac{0.75(\lambda^6 + 10.67\lambda^5\mu + 64.33\lambda^4\mu^2 + 204\lambda^3\mu^3 + 256\lambda^2\mu^4 + 40\lambda\mu^5)}{\lambda^6 + 10\lambda^5\mu + 61\lambda^4\mu^2 + 150\lambda^3\mu^3 + 210\lambda^2\mu^4 + 288\lambda\mu^5 + 120\mu^6} \tag{24}$$

Replacing the utilization factor ρ as $\rho = \frac{\lambda}{m*\mu}$ in (22) and solving (22–24)

$$\Rightarrow E(T_{S,\text{Flat}}) - E(T_{S,\text{Hier}}) > 0 \tag{25}$$

$$\Rightarrow E(T_{S,\text{Flat}}) > E(T_{S,\text{Hier}}) \tag{26}$$

Equation (26) proves analytically that the mean sojourn time of flat is greater than the (3,1) hierarchical architecture we considered here. This implies that for a (3,1) hierarchical architecture with the scheduling rule defined by our top-down allotment policy takes lesser time to finish the same number of jobs as compared to an equivalent flat architecture.

Computational Power: We define computational power as the amount of power (in watts or CPU cycles/ second) a server takes to process a task.
Mean Computational Power—Flat Architecture We define $E(P_F)$ as the mean computational power taken by the servers in a flat architecture to perform 'N' number of tasks. Our assumption is that all the servers in the flat architecture have the same computational capacity. If P_1 denote the power taken by a tier-1 server to execute a task, then

$$\therefore\ E(P_F) = \pi_0.0 + \pi_1.P_1 + \pi_2.2 * P_1 + \pi_3.3 * P_1 + \pi_4.4 * P_1 + \pi_5.5 * P_1 \tag{27}$$

where π_i = Probability that a server remain in state 'i'.

$$\Rightarrow \frac{E(P_F)}{P_1} = \frac{5\rho + 25\rho^2 + 62.5\rho^3 + 104.15\rho^4 + 130.2\rho^5}{(1 + 5\rho + 12.5\rho^2 + 20.83\rho^3 + 26.04\rho^4 + 26.04\rho^5)} \tag{28}$$

Mean Computational Power—Hierarchical Architecture: We define $E(P_H)$ as the mean computational power taken by the hierarchical architecture servers to perform 'N' number of jobs. Our assumption is that the tier-1 servers have a computational capacity of '$P1$', and the tier-2 servers have a computational capacity of '$P2$'. We assume $P_2 = 1.5 * P_1$ and for ease of analysis we assume $P_1 = 1$.

$$\Rightarrow E(P_H) = P_2 * \pi_{0,1} + \pi_{1,1} * (P_1 + P_2) + \pi_{2,1} * (2P_1 + P_2) + \pi_{3,1} * (3P_1 + P_2) + \pi_{1,0} * P_1 + \pi_{2,0} * 2P_1 + \pi_{3,0} * 3P_1 \tag{29}$$

$$\Rightarrow \frac{E(P_H)}{P_1} = 1.5\pi_{0,1} + 2.5\pi_{1,1} + 3.5\pi_{2,1} + 4.5\pi_{3,1} + \pi_{1,0} + 2\pi_{2,0} + 3\pi_{3,0} \tag{30}$$

$$\Rightarrow \frac{E(P_H)}{P_1} = \frac{4.5\left(\lambda^6 + 8.67\lambda^5\mu + 39.\lambda^4\mu^2 + 85.33\lambda^3\mu^3 + 78.\lambda^2\mu^4 + 20.\lambda\mu^5\right)}{\lambda^6 + 10.\lambda^5\mu + 61.\lambda^4\mu^2 + 150.\lambda^3\mu^3 + 210.\lambda^2\mu^4 + 288.\lambda\mu^5 + 120.\mu^6} \tag{31}$$

Replacing $\rho = \frac{\lambda}{m*\mu}$ in (28) and solving for $(E(P_F) - E(P_H))$ from (28–31) gives:

$$E(P_F) - E(P_H) < 0; \qquad \rho < 0.63 \tag{32}$$
$$E(P_F) - E(P_H) > 0; \qquad \rho > 0.63 \tag{33}$$

$$\Rightarrow E(P_F) < E(P_H); \qquad \rho < 0.63 \tag{34}$$
$$E(P_F) > E(P_H); \qquad \rho < 0.63 \tag{35}$$

which proves analytically that the mean computational power of flat is lesser than the equivalent (3,1) hierarchical architecture when the utilization factor $\rho < 0.63$, and for $\rho > 0.63$ the mean computational power of flat is greater than the equivalent (3,1) hierarchical architecture. This implies that for a (3,1) hierarchical architecture with the scheduling rule as defined by our top-down allotment policy takes more computational power for $\rho < 0.63$, but as the incoming traffic increases the hierarchical architecture performs better taking less computational power than the equivalent flat architecture i.e. for $\rho > 0.63$.

4 Simulation Results and Discussions

The simulation platform which we have used for comparing the two architectures is Omnet++ 5.5.1. Simulation experiments were performed for a two-tier (3,1) hierarchical topology and compared with the equivalent flat topology for 10,000 jobs as the input. The arrival rates λ and the service rates μ are exponentially distributed for varying values of $\rho = \frac{\lambda}{m*\mu}$ with 'm' = 5 in our case with the scheduling algorithm (jobs sent to different servers) used as compound round-robin top-down rule. Figures 5 and 6 depict the simulated architectures of flat and hierarchical topologies in Omnet++. For the hierarchical architecture in Fig. 6, the compound round-robin algorithm ensures that only when the tier-2 server is fully occupied, the task goes to a tier-1 server. The buffer size of the queue is assumed to be zero in our work; hence, any packets that arrive when all the servers are fully occupied are discarded.

Figure 7 evaluates the analytical and simulation mean sojourn time curves performance of the flat and the hierarchical architectures of our work. As seen from the figure, the mean sojourn time of the hierarchical architecture utilizing the top-down scheduling rule is lesser for both analytical and simulation curves.

Figure 8 evaluates the analytical and simulation mean computational power curves performance of the flat and the hierarchical architectures of our work. As seen from the curves, the analytical curves of flat and hierarchical architecture crossed for $\rho = 0.63$, which implies that the mean computational power is lesser for the flat architecture in this region and for $\rho > 0.63$, the hierarchical architecture has lesser mean computational power compared to flat. The corresponding simulation

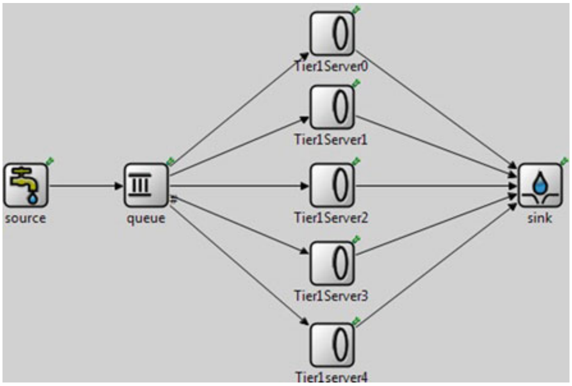


Fig. 5 Simulated flat architecture

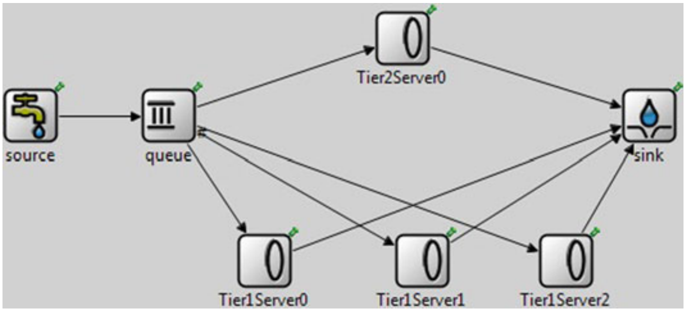


Fig. 6 Simulated hierarchical architecture

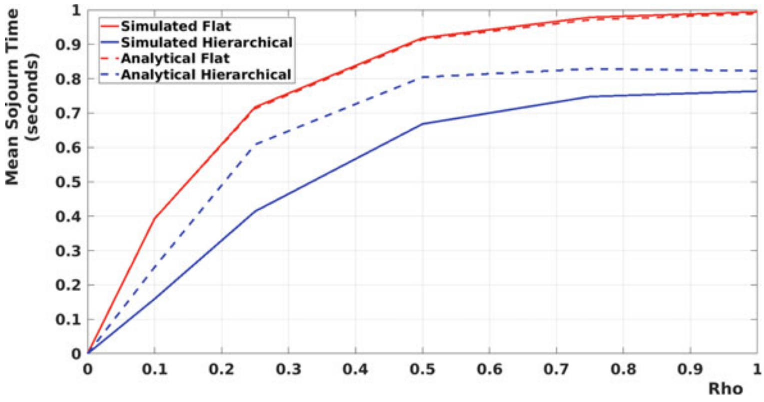


Fig. 7 Mean sojourn time comparison

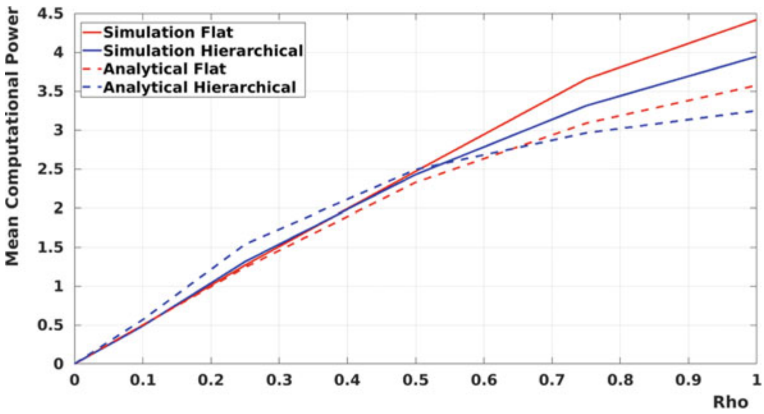


Fig. 8 Mean computational power comparison

curves of flat and hierarchical crossed for $\rho = 0.41$, which implies that the mean computational power of flat is lesser than the hierarchical for $\rho < 0.41$, and it was greater than hierarchical for $\rho > 0.41$ under simulation conditions.

5 Conclusion

Fog computing is an emerging technology that is becoming prominent due to the tremendous increase in the number of devices that is getting connected. The primary usefulness is for low latency low processing requirements eliminating the need for communicating with cloud-centric high processing capability servers. In this work, we propose a hierarchical architecture for processing of the tasks in the fog computing layer. A top-down scheduling rule is formulated for the assignment of the tasks in this architecture, followed by the analytical derivations of a (3,1) hierarchical model using two-dimensional Markov chain state diagrams. The hierarchical architecture we proposed is compared analytically with an equivalent flat topology in terms of the mean sojourn time and the mean computational power performance measures. The hierarchical architecture was better in terms of the mean sojourn time for all values of ρ, and for values of ρ greater than a threshold it was better in terms of the mean computational power also. We simulate the two architectures in Omnet++ and justify the results we obtained for the analytical computations using the simulation curves. Still, there is scope for lots of work by evaluating the blocking probability performance as well as introducing a nonzero buffer size for the queue and then arriving at the best possible architecture which we are currently working on.

References

1. Dawy Z, Saad W, Ghosh A, Andrews JG, Yaacoub E (2017) Toward massive machine type cellular communications. IEEE Wirel Commun 24(1):120–128 Feb.
2. Mozaffari M, Saad W, Bennis M, Debbah M (2016) Unmanned aerial vehicle with underlaid device-to-device communications: performance and tradeoffs. IEEE Trans Wirel Commun 15(6):3949–3963 Jun.
3. Park T, Abuzainab N, Saad W (2016) Learning how to communicate in the Internet of Things: finite resources and heterogeneity. IEEE Access 4:7063–7073 Nov.
4. Jošilo S, Dán G (2019 Feb) Decentralized algorithm for randomized task allocation in fog computing systems. IEEE/ACM Trans Netw 27(1):85–97
5. Chekired DA, Khoukhi L, Mouftah HT (2018 June) Industrial IoT data scheduling based on hierarchical fog computing: a key for enabling smart factory. IEEE Trans Ind Inf 14(10):4590–4602
6. Kleinrock L (1975) Queueing systems, Vol. 1—Theory. Wiley, New York
7. Qamar F, Dimyati K, Hindia MN, Noordin KA, Amiri IS (2019 Apr) A stochastically geometrical poisson point process approach for the future 5G D2D enabled cooperative cellular network. IEEE Access 7:60465–60485

Wideband Spectrum Sensing Using Geolocation Database for Cognitive Radio Networks

G. P. Aswathy, K. Gopakumar, and T. P. Imthias Ahamed

Abstract The cognitive radio technology is considered as an effective method for alleviating the dilemma of spectrum scarcity by creating spectrum access opportunities for the secondary users. Future cognitive devices require advanced sensing techniques for rapid and active identification of spectrum holes over a wideband. Incorporation of prior data from the geolocation database enhances the sensing efficacy, reduces the computation complexity and maximizes the spectrum hole detection capability for cognitive devices. However, unlike the TV white space database, other wideband spectrum databases are not yet available. Therefore, for dynamic and fragile cognitive networks, the geolocation database cannot assist the spectrum sensing process, which necessitated database-independent real-time spectrum sensing. In this paper, the performance of modulated wideband converter-based wideband spectrum sensing is analysed using prior information from geolocation database.

Keywords Wideband spectrum sensing · Cognitive radio · Geolocation database · TV white space

1 Introduction

The rapid development of mobile services along with the Internet of Things (IoT) technology demands more operational spectrum which overwhelms the current government policy of static spectrum allocation. This stimulates the need for advanced

G. P. Aswathy (✉) · K. Gopakumar
Department of ECE, TKM College of Engineering, Kollam, A. P. J. Abdul Kalam Technological University, Thiruvananthapuram, India
e-mail: aswathygp19@gmail.com

K. Gopakumar
e-mail: gopakumar@tkmce.ac.in

T. P. Imthias Ahamed
Department of EEE, TKM College of Engineering, Kollam, A. P. J. Abdul Kalam Technological University, Thiruvananthapuram, India
e-mail: imthiasa@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_65

dynamic spectrum access techniques to alleviate user demands and spectrum supply gap. Many factors led to an increased interest among researchers to develop efficient dynamic access techniques. Commercial broadband operators require additional operational spectrum to counterbalance the increasing demand for wireless broadband services. Meanwhile, the Federal Communication Commission (FCC) recently revealed that the geographical and temporal utilization of the licensed spectrum is very low [1]. Also, there is a rapid development in dynamic spectrum access techniques through geolocation database and other emerging technologies. Underutilization of licensed spectrum along with the dilemma of spectrum shortage promoted the development of innovative cognitive radio technology [2]. This technology plays a major role in providing dynamic spectrum access to secondary users. With this, the secondary users can access the unutilized licensed spectrum when it is vacant.

Spectrum scarcity and the emergence of cognitive radio technology have inspired the government to take crucial steps towards the release of multiple frequency bands for dynamic spectrum access including the 3.5 GHz band Citizens Broadband Radio Service (CBRS) [3], TV White Space (TVWS) [4] and the 5 GHz unlicensed bands [5]. In 2015, the UK Office of Communications (Ofcom) has proclaimed the licence-free regulations for using TVWS [4]. Thus, low-power compact white space devices can be used in TVWS for providing rural broadband and Wi-Fi [6]. This concept can be further extended to other vacant spectral widebands.

Continuous and accurate spectrum sensing is necessary for white space devices to avoid harmful interference to authorized users [7–9]. For this, these devices should either make use of a geolocation database or sense the presence of licensed users to determine the unused spectrum in their vicinity [4]. In TVWS, the current mechanism used by the white space devices to determine the unutilized TV channels is to query the geolocation database [10]. For energy-constrained cognitive devices, spectrum sensing is more expensive in terms of hardware complexity and energy consumption. However, the geolocation database is much easier to implement and requires only simple hardware. The white space devices initially identify their geographical location and query the geolocation database for vacant spectrum. The database system then returns a list of available channels for reuse along with their maximum allowable transmit powers. Meanwhile, there are various drawbacks for geolocation databases. First, it can protect only licensed users, while PMSE (Programme Making and Special Events) users mostly operate in unlicensed basis. Secondly, it uses propagation modelling for determining the vacant spectrum which makes it very conservative in terms of the number of frequency bands it returns for a particular geographical location.

Hybrid sensing in [11, 12] incorporates both spectrum sensing and geolocation database. However, sequential sensing for wide frequency bands increases the sensing delay. To reduce the unaffordable sampling rate requirement, compressive sampling-based sensing scheme is proposed in [13]. However, accuracy and resolution of geolocation database information affect the support recovery process. The recent literature suggests that the procurement of extra a priori spectral data from geolocation database in compressive sensing offer several benefits especially in terms of computational complexity and number of required measurements [14]. Recently, multicoset sampling and sub-space-augmented orthogonal matching pursuit (SA-OMP)-based joint sensing scheme is proposed in [15]. However, multicoset sampling technique has practical limitations which makes it unsuitable for cognitive devices [16].

This paper introduces a less-complex hybrid scheme of modulated wideband converter-based sub-Nyquist wideband spectrum sensing with prior information from geolocation database. The paper is organized as follows. Section 2 introduces the system model. Section 3 describes the block diagram of the proposed scheme. Section 4 describes the joint sensing scheme. Section 5 presents the results. Finally, Sect. 6 gives the concluding remarks.

2 System Model

Assume that the received signal $x(t)$ is given by,

$$x(t) = s(t) + n(t) \tag{1}$$

where $s(t)$ is the primary user signal and $n(t)$ is the additive white Gaussian noise. The primary user signal, $s(t)$ is modelled as analog multiband signal given by,

$$s(t) = \sum_{j=1}^{K} \sqrt{E_j B_j} sinc\left(B_j t\right) \exp\left(2\pi f_j t\right) \tag{2}$$

where E_j is the energy, B_j is the bandwidth and f_j is the frequency associated with each primary user j. In Eq. (1), $n(t)$ is having zero mean and variance σ_n^2, i.e. $n(t) \sim \mathcal{N}\left(0, \sigma_n^2\right)$. Consider that the observed signal $x(t)$ is a continuous-time wideband signal. The frequency-domain representation of $x(t)$ gives a sparse spectrum with most of its energy confined within a limited number of mutually exclusive frequency subbands.

Assume that the wideband spectrum under consideration is uniformly divided into W narrowbands. Let B denote the bandwidth for each segment. These segments are indexed from 0 to $W-1$. Suppose that there are K active bands within the wideband. Let S denote the set of spectral support indices represented as $S = \left[\, S_1 \;\; S_2 \;\ldots\; S_K \,\right]$. The primary aim behind wideband spectrum sensing is to identify the vacant spectrum or equivalently the active channel set S. After recovering S, the vacant spectrum (or white space) is allocated to the secondary users.

3 System Architecture

The system architecture for the proposed spectrum sensing scheme is shown in Fig. 1.

In Fig. 1, the sensor network consisting of numerous sensor nodes performs the real-time sensing of the wideband spectrum. The sensor nodes identify the spectral occupancy status and inform the master white space device (WSD). For TV spectral

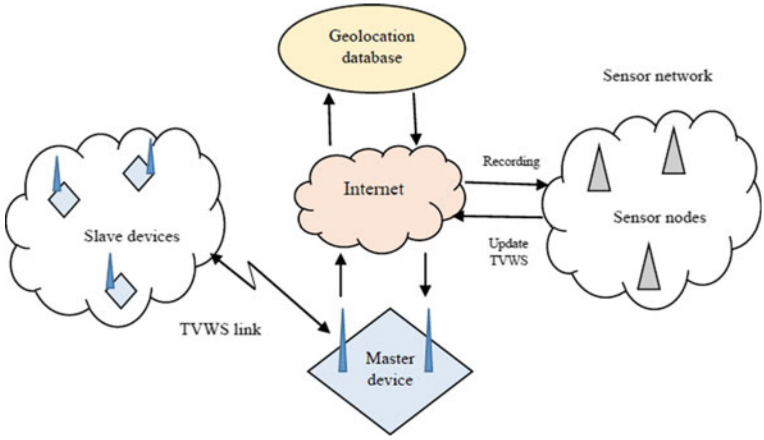


Fig. 1 System architecture for the proposed sensing scheme

bands, the geolocation database is available for reference. Therefore, for TV spectral bands, the sensor nodes perform sensing with prior informtion from the database and informs the master device the spectral occupancy status. Based on the decision made by the master device, the vacant frequency slots are allocated to the slave devices. Meanwhile, the geolocation database is also updated. It is noteworthy that numerous TV spectral bands are heavily used by incumbent users including local radio and TV stations. Therefore, their information available in geolocation database is stable due to the long-run TV broadcasting policy. Even if the prior data from the spectral database is erroneous due to the dynamic nature of spectral occupancy, incorporation of such information only increases the reliability of the spectrum sensing process.

Figure 2 shows the block diagram for the proposed scheme. The wideband signal received by the slave devices is sampled using sub-Nyquist rates, and the spectral indices are identified with prior data from the database. The main objectives of this framework include enhancing the spectrum utilization and enabling communication for white space devices in radio frequency spectral bands. Spectrum utilization can be improved by efficiently using the available spectral holes. For this, spectrum

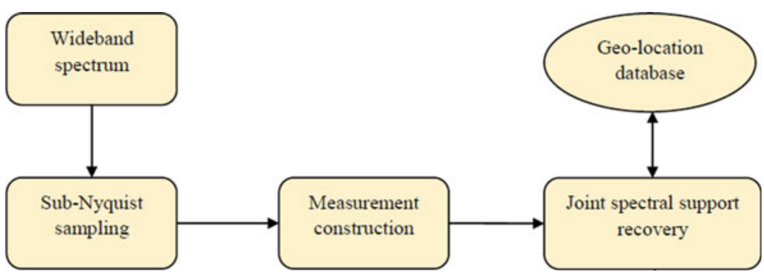


Fig. 2 Block diagram for the proposed scheme

sensing techniques along with a priori knowledge about geolocation database can be used. The WSD's operation in TVWS is controlled using the geolocation database available for TV spectrum. In such scenarios, the WSDs identify their geographical location and search the geolocation database to detect the vacant spectral bands at these locations. However, unlike the TV spectral band geolocation database, other wideband spectrum databases are not yet available. Therefore, for dynamic and fragile cognitive networks, the geolocation database cannot assist the spectrum sensing process, which necessitated database-independent real-time spectrum sensing. Even though spectrum sensing techniques identify the active spectral bands, still there is always a slight probability of harmful interference from incumbent users.

4 Joint-Sparse Recovery Incorporated with Geolocation Database

4.1 Compressed Sensing

Compressed sensing (CS) [17] has gained wide attention in signal processing owing to their successful recovery of samples at sub-Nyquist rates. The success of CS technique for wideband spectrum sensing relies on the frequency-sparse nature of wideband signal. CS methods try to recover the unknown sparse signal $\boldsymbol{X}$ from the partial observation matrix $\boldsymbol{Y}$ by finding solution to an underdetermined system of linear equation as follows:

$$\boldsymbol{Y} = \phi^{\mathrm{T}} \psi \boldsymbol{X} = A\boldsymbol{X} \tag{3}$$

where A is the $m \times d$ sensing matrix with $m < d$, $\psi \in \mathbb{R}$ is the $m \times d$ measurement matrix and $\phi \in \mathbb{C}$ is the $m \times m$ transform matrix. The sensing matrix A is optimally chosen according to some specific constraints so that it provides an almost precise solution to the above set of underdetermined equations defined by (3). One of the main characteristics of the sensing matrix is the restricted isometric property (RIP) which solely determines whether the system of equations is solvable.

4.2 Modulated Wideband Converter

In this work, we adopt compressive modulated wideband converter (MWC) [18] for signal acquisition and sampling. Each sensor node implements a branch of MWC which samples the signal at a sub-Nyquist sampling rate. A low-dimensional sensing matrix is obtained from these low-rate samples. Then, the spectral support recovery is realized using orthogonal matching pursuit (OMP) algorithm. In our proposed approach, the a priori data from the spectral database is also incorporated for support recovery.

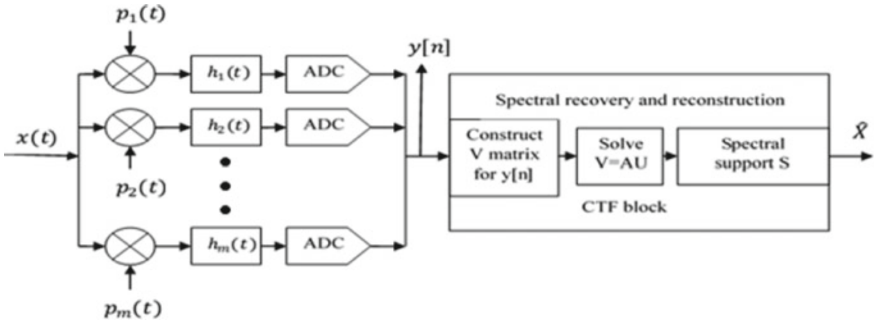


Fig. 3 Modulated wideband converter [19]

Figure 3 depicts a practical MWC [19]. The sensor nodes sense the neighbouring spectrum and receive the signal $x(t)$ which is sampled at a low sampling rate using MWC. It comprises a bank of mixers, low-pass filters (LPFs) and analog-to-digital converters (ADCs) operating at low sampling rates. The MWC consists of m branches which are fed with the same signal $x(t)$. In each mixer 'i', $x(t)$ is mixed with a periodic waveform $p_i(t)$. This waveform should be periodic (period $= T_\mathrm{p}$) and is selected as a piece-wise constant waveform which alternates its value between $\{+1 \text{ or } -1\}$ at equal time intervals, M_s.

$$p_i(t) = p_{il};\, l\frac{T_\mathrm{p}}{M_s} \leq t \leq (l+1)\frac{T_\mathrm{p}}{M_s};\, 0 \leq l \leq M_s - 1;\, p_{il} = \{+1, -1\} \tag{4}$$

The mixing process ensures the aliasing of spectrum from each band to the baseband. In each branch 'i', the LPF, $h_i(t)$ truncates the signal spectrum. The filtered signal is then sampled using a low-rate ADC. The spectral support is recovered using the continuous-to-finite (CTF) block shown in Fig. 3. $y[n]$ represents the output measurement matrix from MWC. The CTF block initially computes the matrix Q as,

$$Q = \sum_{n=-\infty}^{\infty} y[n]y[n]^\mathrm{T} \tag{5}$$

The frame matrix V is constructed from Q matrix using eigenvalue decomposition (EVD),

$$[V, d] = EVD\left[Q, \widehat{K}\right] \tag{6}$$

where $d = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ represent the K principal eigenvalues and V represents their corresponding eigenvectors. The eigenvectors V and the eigenvalue matrix d are jointly used to obtain the lower-dimensional measurement matrix, v.

$$v = V\sqrt{d} \tag{7}$$

Consider the system of linear equations given by,

$$v = Au \tag{8}$$

where A is the sensing matrix [18]. Solve (8) to recover the sparsest vector, $\hat{u}$. The unknown spectral support (S) is identified as,,

$$S = \bigcup_i \text{Supp}\left(\hat{u_i}\right) \tag{9}$$

4.3 Spectral Support Recovery with Prior Information

It is known that v in (8) is row-sparse, i.e. the matrix v has very few rows with non-zero entries. Indices of those rows correspond to the unknown spectral support, S. Recovering sparse solution to (8) is considered as the multiple measurement vector (MMV) problem. Greedy pursuit algorithms [20] and convex relaxation [21] are the most common methods used for solving (8). In this paper, OMP [22, 23] is used for joint spectral support recovery. The geolocation database provides prior channel occupancy status which is further incorporated for joint spectral support recovery.

Let T denotes the prior information from the geolocation database such that $T \subset [0, W-1]$. T is related to the original spectral support S as,

$$S = T \cup \Delta / \Delta_e \tag{9}$$

where $\Delta := S/T$ corresponds to the spectral indices of the newly active channels and $\Delta_e := T/S$ are the newly free channels. Δ and Δ_e can be identified during real-time spectrum sensing. The joint spectral support algorithm is as follows.

Algorithm 1: Joint spectral support recovery using OMP with prior information from geolocation database

Input: Require $\boldsymbol{A}$, $\mathbf{Q}$, $\boldsymbol{T}$
Output: Spectral support, $\boldsymbol{S}$
Procedure:

1. Eigenvalue decomposition of $\mathbf{Q}$ matrix to obtain frame matrix, $\mathbf{V}$
   $[\boldsymbol{V}, \boldsymbol{d}] = \boldsymbol{EVD}\left[\boldsymbol{Q}, \hat{\boldsymbol{K}}\right], \boldsymbol{v} = \boldsymbol{V}\sqrt{\boldsymbol{d}}$
2. Initialize: Count, $\boldsymbol{I} = \mathbf{0}$, $\boldsymbol{S} = \boldsymbol{T} + \mathbf{1}$
   $\boldsymbol{u_0} = \boldsymbol{A}^{\dagger}_{\boldsymbol{T}+\mathbf{1}}\boldsymbol{v}$, residual, $\boldsymbol{r_0} = \boldsymbol{v} - \boldsymbol{A}_{\boldsymbol{T}+\mathbf{1}}\boldsymbol{u_0}$

3. ***while*** $\boldsymbol{I \le \widehat{K} - |T|}$ ***do***
4. Increment $\boldsymbol{I = I + 1}$
5. Find the best spectral support index
   $\boldsymbol{s_t = \mathop{\mathbf{argmax}}_{q} \left\| A_q^H r_{I-1} \right\|_2}$
6. Update the spectral support $\boldsymbol{S}$
   $\boldsymbol{S = S \cup s_t, u_I = A_s^{\dagger} v}$
7. Update the residual
   $\boldsymbol{r_t = v - A_s u_I}$
8. ***end while***
9. ***return S***

Application of OMP reduces the computation costs in low-power resource-constrained sensor nodes. Conventional OMP uses K iterations to recover K-sparse signal. Since T indices are known (obtained from geolocation database), above algorithm requires only $K - T$ iterations. Notably, local spectrum sensing needs to be performed only on a limited number of vacant bands. Thus, computational complexity and recovery delay are thus significantly reduced.

5 Results and Discussions

This section presents the simulation results obtained while analysing the performance of the proposed sub-Nyquist wideband spectrum sensing technique. We investigate the effect of various parameters including the spectral occupancy ratio, signal-to-noise ratio (SNR), number of branches of MWC and the detection capability on the system performance with prior data from the spectral database.

5.1 Simulation Setup

Consider a wideband signal received by the sensor nodes, $x(t) \in \mathcal{F} = [0, 10]$ GHz. The wideband is segmented into $W = 195$ bands, bandwidth of each being $B =$ 50 MHz. Let K be the number of active bands, such that $K \le W$. We choose $K = 10$ for simulation. We use Eq. (2) to model the analog wideband signal. To measure the detection capability of the method, we calculate the success rate which is the fraction of times the active primary channels are correctly being reported as occupied (i.e. $\widehat{S} = S$). The recovered spectral support $\widehat{S}$ is compared against the original active channel set S using 1000 Monte Carlo trials and the success rate is plotted.

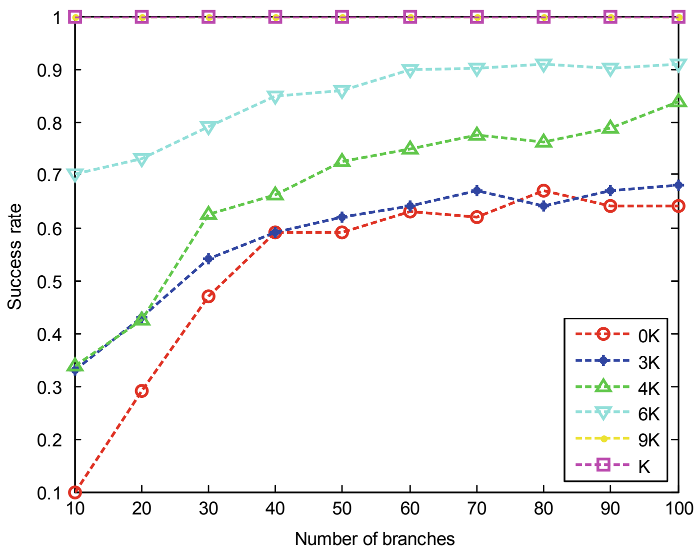


Fig. 4 Success rate versus number of branches under varying prior information from geolocation database

Performance of proposed scheme upon sub-Nyquist sampling ratio and SNR. We evaluate the performance of the proposed spectrum sensing scheme using MWC and OMP. The number of active channels is randomly chosen as $K = |S| = 10$. Thus, $\alpha = 5\%$ is the channel occupancy ratio. The spectral support S and the prior information from the geolocation database T are randomly chosen from $[0, W - 1]$. The size of prior known part from geolocation database, τ, is varied from 0 to K. In Fig. 4, we consider the received SNR as -5 dB, and the number of branches for MWC is varied from 10 to 100. If $\tau = 0K$, there is no a priori knowledge from the spectral database, while the information from geolocation database is completely available if $\tau = K$. For $\tau \geq 0.5K$, less number of branches are required for attaining high detection performance. In Fig. 5, the spectral detection capability is examined by varying SNR from $-$ 10 dB to 20 dB. From Figs. 4 and 5, we observe that the success rate generally increases with increase in the number of branches along with prior information from the geolocation database. As concluding remark, with prior information from geolocation database, less number of branches are required in the joint sensing scheme to achieve higher detection capability.

Average number of iterations.
The number of iterations required for OMP is limited to the sparsity level, K. Conventional OMP uses K iterations to recover K-sparse signal. Since τ indices are known (obtained from geolocation database), OMP requires only $K - \tau$ iterations. Thus, the prior knowledge from the spectral database reduces the required number of OMP iterations to $\widehat{K} - \tau$. This reduces the computational cost and sensing delay significantly.

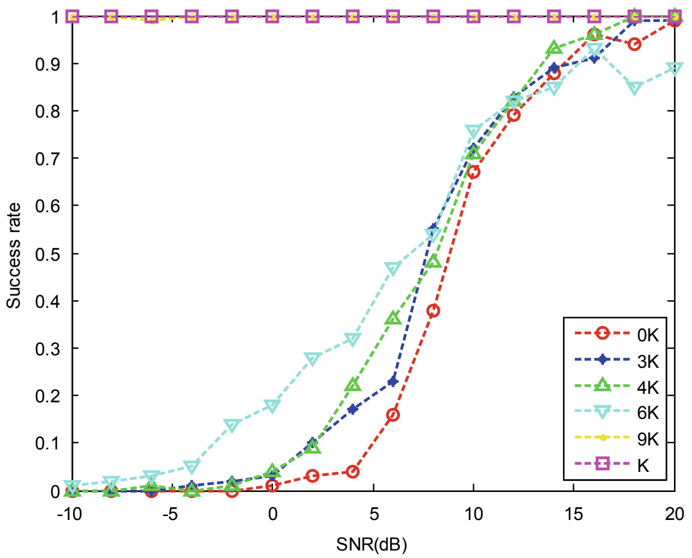


Fig. 5 Success rate versus SNR under varying prior information from geolocation database

Performance of the proposed technique with partially incorrect prior information from geolocation database.
In Figs. 4 and 5, it is assumed that the spectral information stored in the database is completely reliable. But, there are inevitable cases where the stored data in the database is erroneous, especially in a dynamic environment when the channel occupancy status changes with time. In such cases, the actual spectral support can still be recovered with a slight increase in computational cost and number of measurements. Figure 6 shows the success rate versus number of branches of MWC for partially incorrect spectral indices from geolocation database. In Fig. 6, we assume that C out of τ indices in T are correct and W out of τ indices are wrong. The simulation setting is same as that in Figs. 4 and 5 but with different combinations of c and W in T. It is observed that as the prior information from database increases, the proposed sensing scheme recovers the exact spectral support and improves the detection probability.

6 Conclusions

In this work, a less-complex joint sub-Nyquist wideband spectrum sensing technique for white space devices is proposed. MWC and OMP algorithm are used to realize less-complex wideband spectrum sensing. For TV spectral bands, the information available from the geolocation database can assist the spectrum sensing process.

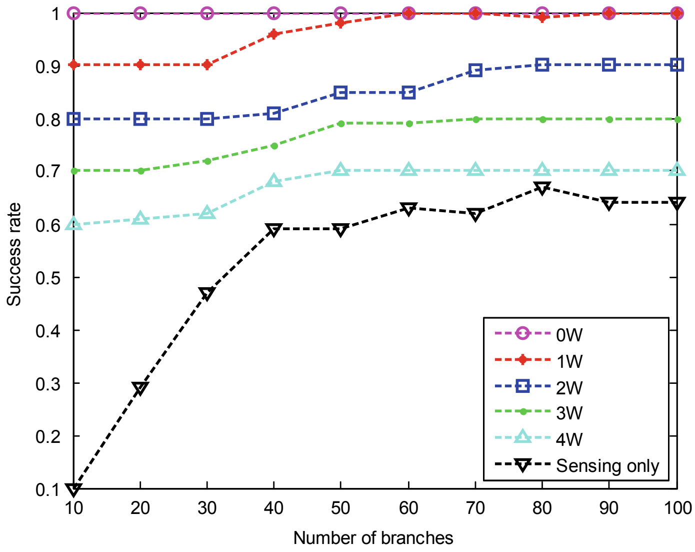


Fig. 6 Success rate versus number of branches for partially incorrect spectral indices from geolocation database

However, for other wideband spectrums, the geolocation database is not yet available, and therefore, geolocation database-independent spectrum sensing is also discussed. The a priori spectral occupancy status from the database reduces the number of measurements, computational cost, sampling rate and the overall sensing delay. Remarkably, it reduces the number of iterations and computation time required for recovering the spectral support. Further, it increased the white space detection performance and improved the spectral utility. This scheme is thus more suitable for low-power compact cognitive devices.

Acknowledgements We would like to acknowledge the Centre for Engineering Research and Development under A. P. J. Abdul Kalam Technological University, Thiruvananthapuram, for the support offered to complete this project.

References

1. FCC (2003) ET Docket No 03-222 Notice of proposed rule making and order
2. Mitola J, Maguire GQ (1999) Cognitive radio: making software radios more personal. IEEE Personal Commun Mag 6(4):13–18
3. Federal Communications Commission (2015) Amendment of the commission’s rules with regard to commercial operations in the 3550–3650 MHz band. Washington, DC, USA, GN Docket no. 12-354
4. Kolodzy P (2002) Spectrum policy task force report. Federal Communications Commission, Washington, DC, USA, ET Docket no. 02-135

5. Federal Communications Commission (2013) In the matter of revision of part 15 of the commission's rules to permit unlicensed national information infrastructure (U-NII) devices in the 5 GHz band: notice of proposed rulemaking. Washington, DC, USA, ET Docket no. 13-49
6. Ofcom (2015) Implementing TV White Spaces. Ofcom, London, U.K. http://stakeholders.ofcom.org.uk/binaries/consultations/white-spacecoexistence/statement/tvws-statement.pdf
7. Ding G, Wang J, Wu Q, Yao Y-D, Song F, Tsiftsis TA (2016) Cellular-base-station assisted device-to-device communications in TV white space. IEEE J Sel Areas Commun 34(1):107–121
8. Qin Z, Gao Y, Plumbley M, Parini C (2016) Wideband spectrum sensing on real-time signals at sub-Nyquist sampling rates in single and cooperative multiple nodes. IEEE Trans Signal Process 64(12):3106–3117
9. Ma Y, Gao Y, Cavallaro A, Parini CG, Zhang W, Liang YC (2017) Sparsity independent sub-Nyquist rate wideband spectrum sensing on real-time TV white space. IEEE Trans Veh Technol 66(10):8784–8794
10. Ofcom (2010) Implementing geolocation. Ofcom, London, U.K. https://www.ofcom.org.uk/__data/assets/pdf_file/0035/46889/statement.pdf
11. Dionisio R, Ribeiro J, Marques P, Rodriguez J (2014) Combination of a geolocation database access with infrastructure sensing in TV bands. EURASIP J Wirel Commun Netw 1:1–14
12. Wang N, Gao Y, Evans B (2015) Database-augmented spectrum sensing algorithm for cognitive radio. In: Proceedings of IEEE international conference on communications, London, U.K., pp 7468–7473
13. Tian Z, Giannakis GB (2007) Compressed sensing for wideband cognitive radios. In: Proceedings of IEEE international conference on acoustics, speech, and signal processing, Honolulu, HI, USA, pp 1357–1360
14. Vaswani N, Lu W (2010) Modified-CS: Modifying compressive sensing for problems with partially known support. IEEE Trans Signal Process 58(9):4595–4607
15. Ma Y, Zhang X, Gao Y (2018) Joint sub-Nyquist spectrum sensing scheme with geolocation database over TV white space. IEEE Trans Veh Technol 67(5):3998–4007
16. Mishali M, Eldar YC, Tropp JA (2008) Efficient sampling of sparse wideband analog signals. In: 2008 IEEE 25th convention of electrical and electronics engineers, Israel, Eilat, pp 290–294
17. Donoho DL (2006) Compressed sensing. IEEE Trans Inf Theory 52(4):1289–1306
18. Mishali M, Eldar YC (2009) Blind multiband signal reconstruction: compressed sensing for analog signals. IEEE Trans Signal Process 57(3):993–1009
19. Aswathy GP, Gopakumar K (2019) Sub-Nyquist wideband spectrum sensing techniques for cognitive radio: a review and proposed techniques. AEU Int J Electron Commun 104:44–57
20. Tropp JA, Gilbert AC, Strauss MJ (2006) Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. Signal Process (Special Issue on Sparse Approximations in Signal and Image Processing) 86(3):572–588
21. Tropp JA (2006) Algorithms for simultaneous sparse approximation. Part II: Convex relaxation. Signal Process (Special Issue on Sparse Approximations in Signal and Image Processing) 86(3):589–602
22. Aswathy GP, Gopakumar K (2018) Cognitive radio network with wideband spectrum sensing and reliable data transmission. In: IEEE recent advances in intelligent computational systems, Thiruvananthapuram, India, pp 70–74
23. Aswathy GP, Gopakumar K (2018) Wideband spectrum sensing using modulated wideband converter by revised orthogonal matching pursuit. In: International conference on control, power, communication and computing technologies, Kannur, India, pp 179–184

Anomaly Detection and Safe Transmission of ECG Signals in Point-of-Care Systems

N. S. Akhila and K. Sabeena Beevi

Abstract The increasing global focus on health protection issues draws attention to the importance of point-of-care (POC) technologies. Its ability to provide cost-effective solutions for health maintenance points out their importance. However, against the obvious benefits, POC does not provide any primary diagnosis at the patient side. So, this work aims at a primary diagnosis of cardiac diseases at patient side itself by detecting the electrocardiogram (ECG) abnormality. For perfect diagnosis, confidential data of patients are also required. However, transmission of this confidential data through public network may cause many security concerns. Also, wide privacy problems may occur as private data may be revealed to illicit servers. This paper uses support vector machine (SVM)-based technique for detecting abnormality in ECG signals. And Fast Walsh–Hadamard transform-based steganographic method is used for providing privacy, security and confidentiality of the transferred data.

Keywords Electrocardiogram · Point-of-care systems · Steganography

1 Introduction

Point-of-care systems (POC) are defined as a medical technology which enables patients from villages and rural hospitals to get expert advice from specialist doctors located anywhere in the world [1]. The healthcare industry has been quite slow to adopt new technologies compared to other industries. Due to strict set of rules and the susceptible nature of medical information, the industry has proceeded with concern [1]. However, POC systems have come out as a promising next-generation healthcare system. One of the major advantages of POC is that it provides much faster access to test results, allowing for more rapid medical decisions and more

N. S. Akhila (✉) · K. Sabeena Beevi
Electrical and Electronics Engineering, TKM College of Engineering, Kollam 691005, India
e-mail: akhilans2@gmail.com

K. Sabeena Beevi
e-mail: sabeena3000@tkmce.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_66

appropriate treatments and interventions. Point-of-care systems can make significant stride in cardiovascular diseases (CVD) also. Cardiac point-of-care diagnostic test provides a means to obtain quick and on-site results. Hence, monitoring electrocardiogram (ECG) and performing automatic diagnosis turn out to be particularly important [2]. POC has been integrated into the healthcare systems providing faster results and improved patient outcomes, but primary diagnosis of heart disease is not present in such systems. Also, in POC, the patients' personal and confidential data are required for further analysis of cardiovascular diseases which will cause many security concerns. Moreover, traditional encryption–decryption, watermarking and wavelet-based methods [3, 4] used for safe transmission of ECG signal are prone to attacks. Hence, the objective of this paper is to design a system that ensures primary diagnosis at patient side, end-to-end privacy preservation for the patients and to increase security of data. Initial diagnosis of ECG is done at the patient side itself by using support vector machine (SVM), and in case of any abnormality, the signal will be sent to the medical practitioners. Along with the ECG signals, personal information (name, age, sex, etc.) as well as diagnosis data (temperature, oxygen level, blood pressure, etc.) of the patients are also sent for proper diagnosis of disease. Steganography-based data hiding technique is used to conceal the data.

Organization of the paper is in such a way that Sect. 2 gives a review of the recent literature on anomaly detection and steganography-based methods. The proposed methodology in detail is explained in Sect. 3. Section 4 comprises experimental results and discussions along with comparisons, and Sect. 5 concludes this paper.

2 Literature Survey

The QRS complex is the most prominent waveform in ECG [5]. Its analysis provides the basics for automatic ECG diagnosis. Earlier methods for anomaly detection were based on this QRS detection [5]. Earlier feature extraction-based algorithms were used to classify ECG heartbeat patterns from the ECG signals [6]. Fourier transform analysis was used for finding signal spectrum or range of frequency values within the signal. Nonetheless, Fourier transform [6] only provides the frequency-dependent components, not their time-dependent components. Then, feature extraction was done with the help of wavelet transform. This method first removes noise by use of soft or hard threshold; then, the feature of ECG wave is divided into coefficient vector by optimal wavelet transformation. Also, many approaches were carried out to protect patient data. But, only some were based on steganography techniques that conceal the personal data inside the physiological signal. So, the main challenges are to what limit the method is safe and will there be any distortion on the medical signal after embedding process. A reversible data hiding technique [7] was proposed in order to safeguard the data. But, data are concealed by shifting one bit which causes low capacity. Also, the security of this method depends only on the algorithm. Then, an ECG steganography which was based on wavelet method [8] was introduced. But, the 2D hiding order is very static. Then, a Walsh–Hadamard-based method [9] is used which improves the hiding capacity as well as the performance. A 3D order of the signal was formed by using a key. But, a user-defined key and a rotation factor make this method more strenuous for the user. So, an automatically generated rotation factor method is proposed here which rotates the key automatically and improves the security of the system.

3 Methodology

A suitable balance between the two major concerns has been carefully considered in this paper: (1) to properly detect the anomaly of ECG signal to provide faster medications to growing cardiac diseases and (2) to provide security and continuous protection for the patient secret information as well as the biomedical signal.

Block schematic of our methodology is shown in Fig. 1. ECG data is taken from the PhysioNet, and the patients details are entered manually. To detect abnormalities in ECG signal, support vector machine (SVM) classifiers used. Eight different features like mean, standard deviation, etc., are used in SVM for classification. After classifying whether the signal is normal or abnormal, fast Walsh–Hadamard transform (FWHT) is applied to abnormal ECG signals. The personal data combined with security key with an automatic rotation factor reform the FWHT values into a matrix of 3D order. The encrypted personal information is embedded into the ECG signal and forms a stego signal. This stego signal is then sent to the doctors. At receiver side, only the authorized person who has the key will be able to retrieve the signal and the personal data. Inverse Walsh–Hadamard transform is used for the recomposition of ECG signal at the receiver end.

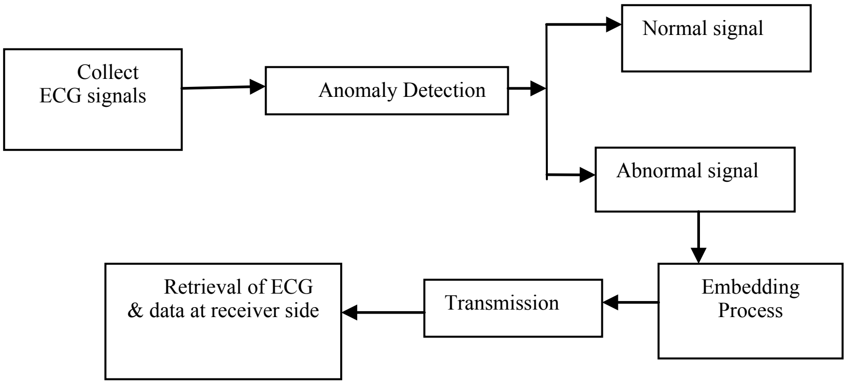


Fig. 1 Block schematic of the method

3.1 Walsh–Hadamard Transform (WHT)

In the proposed scheme, fast Walsh–Hadamard transform (FWHT) [10] is used for encrypting and decrypting the private information into the signal. Coefficients of FWHT are evaluated using Eq. (1) [11].

$$c_n = \frac{1}{M}\sum_{i=0}^{M-1} s_i w(z,k), \quad z = 1, 2 \ldots Z - 1 \tag{1}$$

where c_n is the FWHT matrix coefficients, s_i is the sampled value of ECG and $w(z,k)$ is the transformation matrix. A preconfigured Walsh matrix is used to find the FWHT coefficients. This matrix contains only $+1\backslash-1$.

Patient Data Encryption Advanced Encryption Standard (AES) encrypts the personal data with key. It is a block cipher that uses an encryption key and several rounds of encryption [11]. Varying lengths cipher keys are used in AES. The diagnostic and patient data are combined by using AES-128. Equation (2) [11] shows the encrypted form (*EC*) of the personal information.

$$E_c \Leftarrow D_e(K, S_c) \tag{2}$$

where D_e is algorithm for AES, K is the key, S_c is highly confidential personal information of the patient.

Patient Data Embedding After the encryption process, the key is utilized to convert the physiological signal into a 3D order to embed encrypted patient data. Initially, the key is rotated with a rotation factor that is automatically generated by counting the number of ones in the binary value of key. Random sorting of the key into ascending and descending order is made to reform the signal into a 3D order. To embed the encrypted data one array of $X \times Y \times Z$ is removed and this is transmitted to the authorized persons. Algorithm 1 depicts the overall embedding process.

3.2 Inverse Walsh–Hadamard Transform (IWHT)

For recomposing the original signal and the personal data, the embedded signal should be converted back into its vector format from 3D, and IFWHT is applied for recomposition. It should be noted that stego biomedical signal can also be used for diagnosis purposes. But, the patient's information can be retrieved by legitimate authorities holding the secret key. The IFWHT recomposition is shown by Eq. (3) [11].

$$s_i = \frac{1}{M}\sum_{i=0}^{M-1} cnw\,(z,k), \quad z = 1, 2 \ldots Z-1 \tag{3}$$

where c_n is the FWHT coefficients, s_i is the sampled value of ECG and $w(z, k)$ is the transformation matrix.

Algorithm 1 Algorithm for embedding process

Input: Dataset ECG Samples + personal data

Personal bits array: 128 bits

User-defined key array: 16 byte key

1: Convert ECG into.mat format and apply FWHT transform

2: Binarization of key, personal data and ECG

3: AES-128 operation//combines personal data with key

4: Key rotation using rotation factor//rotation factor − number of ones in the binary key value

5: Sort key and form $X \times Y \times Z$ order

6: Embed data

7: Transmission

4 Results and Discussions

4.1 Dataset

The data is obtained from "PhysioNet" repository [12], which is funded by the National Health Institute (USA) [13]. The data consist of a set of ECG signals sampled at 300 Hz taken from ECG lead recording (with lengths between 30 and 60 s), and each dataset consists of 9000 samples. Figure 2 shows an example of normal and abnormal signal.

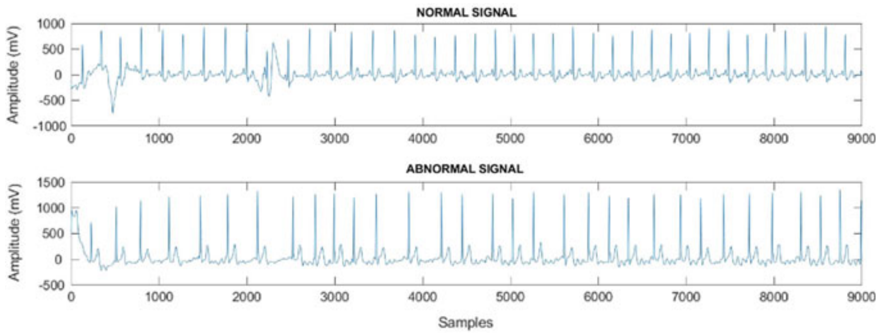


Fig. 2 Normal and abnormal signal

4.2 Performance Measures

The performance of the technique is quantitatively examined by using sensitivity, specificity, accuracy, precision and *F*-score measures. They can be defined as:

Accuracy—Accuracy is simply a fraction of rightly estimated observed values to the total number of observations.
Precision—Precision gives the fraction of rightly estimated positive observed values to the total estimated positive values.
Recall (sensitivity)—Recall gives the ratio of rightly estimated positive observed values to the all observations in actual class.
***F*-score**—*F*-score is the weighted sum of precision and recall.
Specificity (true negative rate)—gives the proportion of actual negatives that are rightly identified as such. Equations (4–8) gives the value of performance measures [14].

$$\text{Sensitivity} = \frac{\text{Number of True positive}}{\text{Number of True positive} + \text{Number of False Negative}} \times 100 \quad (4)$$

$$\text{Precision} = \frac{\text{Number of True positive}}{\text{Number of True positive} + \text{Number of False Positive}} \times 100 \quad (5)$$

$$F\text{-Score} = 2 \times \frac{\text{Sensitivity} \times \text{Precision}}{\text{Sensitivity} + \text{Precision}} \times 100 \quad (6)$$

$$\text{Accuracy} = \frac{\text{Number of True positive} + \text{Number of True Negative}}{\text{Total observations}} \times 100 \quad (7)$$

$$\text{Specificity} = \frac{\text{Number of True Negative}}{\text{Number of False positive} + \text{Number of False Negative}} \times 100 \quad (8)$$

Table 1 presents the performance measures of SVM obtained by the proposed method over the test set, respectively. From the table, it is clear that SVM has high accuracy, precision, sensitivity and *F*-score.

Table 1 Performance evaluation of SVM and BPNN

	Accuracy	Precision	Sensitivity	Specificity	*F*-score
SVM	87.73	87.86	86.6	87.6	87.65
BPNN	80.53	80	81	81.42	80.49

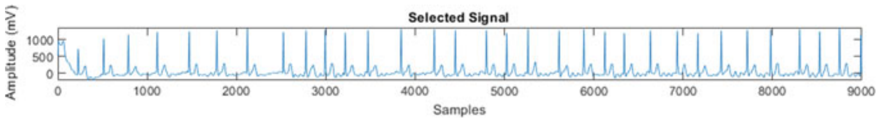


Fig. 3 Abnormal signal for embedding

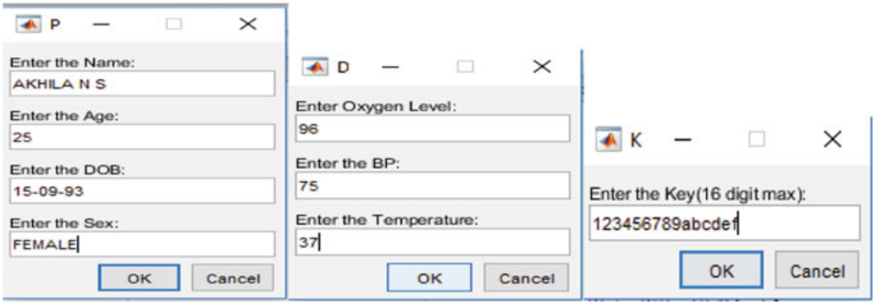


Fig. 4 Personal data and key

4.3 Embedding Process

After classifying the signal as abnormal or normal, the abnormal signal should be sent to medical experts for better treatments in POC systems. So, along with the abnormal signal, patients' details are also sent. To ensure the safety of the patients' information, the details are concealed within the ECG signal sent. Figure 3 depicts a classified abnormal signal which should be sent to the doctors for further analysis.

Along with the ECG signal, patient confidential information is also required like name, age, sex, date of birth and their physiological readings like temperature, pressure, oxygen level so that a better diagnosis of the disease can be done. So, here, in this thesis, these details are entered manually.

Thus, Fig. 4 shows the entered personal data and the user-defined 16 byte key. After this, FWHT is applied to the selected signal for embedding the data. The result of FWHT on the selected signal is shown in Fig. 5. Here, the higher coefficients are removed as it has less significance for reconstruction which also increases the embedding capacity of the signal.

4.4 Retrieval of Confidential Information

The successful retrieval of personal data is shown in Fig. 6. The primary diagnosis of the ECG signal is also shown along with the personal data.

To precisely measure the result of this method on the ECG which is transferred and to verify the stego authenticity, the performance measures of the signal prior to and

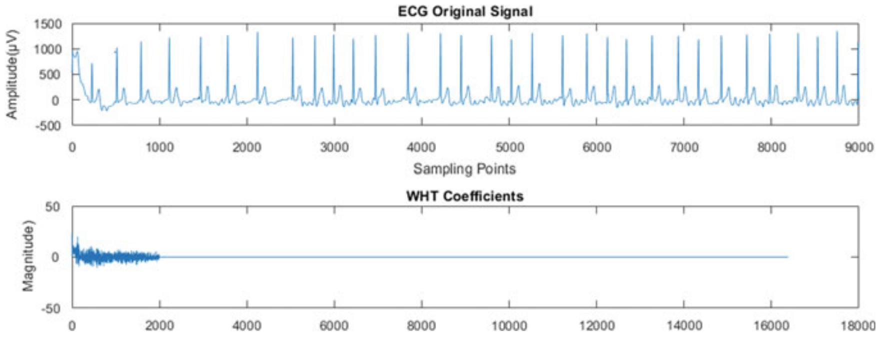


Fig. 5 FWHT applied on selected signal

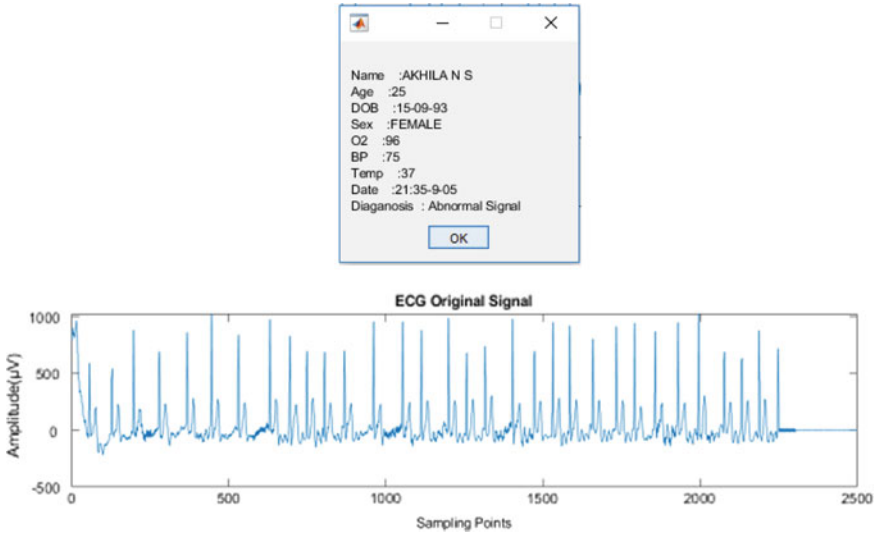


Fig. 6 Reconstructed original signal and personal data

after the stego have been carefully monitored using percentage residual difference (PRD) [8]. PRD gives the measure of variation between the actual ECG host signal and the transformed stego ECG signal as well as reconstructed signals. PRD should be less than 1%. Bit error rate (BER) [11] which gives the reliability of the extracted signal is also calculated. Equation (9) represents the PRD measure where f is the original signal while z the stego or extracted signal.

$$\mathrm{PRD} = \sqrt{\frac{\sum_{k=1}^{m} (f_i - z_i)^2}{\sum_{k=1}^{m} f_i^2}} \tag{9}$$

Table 2 PRD results of ECG samples

Set No	Normal samples		Abnormal samples	
	Stego	Extracted	Stego	Extracted
1	0.532	0.656	0.897	0.876
2	0.456	0.852	0.765	0.762
3	0.225	0.693	0.895	0.467
4	0.894	0.456	0.994	1.332
5	0.687	1.589	0.577	0.678
6	0.745	0.875	0.560	0.589
7	0.925	0.759	0.823	0.756

Table 2 gives the PRD results for both the stego as well as the extracted signals after retrieval of personal information. PRD results for both normal as well as abnormal signals are taken. From Table 2, maximum PRD measured is 0.7%, and the obtained values are >1%. It is clear that, despite different sample values of the ECG signals, all PRDs are less than 1%. This accentuates that this model has a little effect on the actual transferred ECG signals. Also, it provides advanced method for preserving the confidentiality of transferred information as well as the validity of ECG signals.

5 Conclusion

POC system is becoming more and more common nowadays, and the electrocardiogram (ECG) is undoubtedly the most used biological signal in the health care. However, diagnosing any problems in ECG is not an easy task. So, a machine learning-based technique is used to classify the ECG signals at the patient side itself, and a steganography-based FWH algorithm has been used to protect patients' private data in POC systems. An automatic key is used to transform the ECG signal values into a 3D order, and embedding of private data inside biomedical signal is done. Maximum embedding level is ensured by applying FWHT to transform the signals into frequency-based coefficients. To make sure there is lowest distortion, only unimportant coefficients are utilized. Quality of the extracted signal is evaluated from the PRD measures.

References

1. Yager P, Domingo GJ, Gerdes J (2008 Aug) Point-of-care diagnostics for global health. Ann Rev Biomed Eng 10:107–144
2. Wang H et al (2010) Resource-aware secure ecg healthcare monitoring through body sensor network. Wirel Commun IEEE 17(1):12–19

3. Yang S, Yang G (2010) ECG pattern recognition based on wavelet transform and BP neural network, pp 246–249
4. Vijaya V, KishanRao K, Rama V (2011) Arrhythmia detection through ECG feature extraction using wavelet analysis. Eur J Sci Res 66:441–448
5. Castro B, Kogan D, Geva AB (2000)ECG feature extraction using optimal mother wavelet. In: The 21st IEEE convention of the electrical and electronic engineers in Israel, pp 346–350
6. Venkateswarlu D et al (2007) e health networking to cater to rural health care and health care for the aged. In: 2007 9th international conference on e-health networking, application and services, pp 273–276. IEEE (2007)
7. Zheng K, Qian X (2008) Reversible data hiding for electrocardiogram signal based on wavelet transforms. In: International conference on computational intelligence and security, 2008 CIS'08, vol 1, pp 295–299. IEEE
8. Ibaida A, Khalil I, Sufi F (2010) Cardiac abnormalities detection from compressed ECG in wireless telemonitoring using principal components analysis (PCA). In: Proceedings of 5th international conference on intelligent sensors, sensor networks and information processing, Dec 2010, pp 207–212
9. Abuadbba A, Khalil I (2017) Walsh–Hadamard-based 3-D steganography for protecting sensitive information in point-of-care. IEEE Trans Bio-med Eng
10. Abuadbba A et al (2015) Robust privacy preservation and authenticity of the collected data in cognitive radi network Walsh–Hadamard based steganographic approach. Pervasive Mob Comput
11. Ibaida A, Khalil I (2013) Wavelet based ECG steganography for protecting patient confidential information in point-of-care systems. IEEE Trans Biomed Eng 60:3322–3330
12. Goldberger A et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101(23):e215–e220
13. Clifford G, Liu C, Moody B, Lehman H, Silva I, Li Q et al (2017) AF classification from a short single lead ECG recording. The PhysioNet computing in cardiology challenge
14. Beevi KS, Nair MS, Bindu GR (2019) Automatic mitosis detection in breast histopathology images using convolutional neural network based deep transfer learning. Biocybern Biomed Eng 39(1):214–223

Multiband-Loaded Compact Antenna Design for WiMAX/WLAN/UWB Applications

T. K. Sreeja, J. Jayakumari, K. Chandrakala, Remya M. Nair, and Abhilash S. Vasu

Abstract Design and analysis of multiband antenna for WiMAX and WLAN are presented. The designed antenna comprises of a triangular patch and six circular patches that are connected with the triangular patch using rectangular striplines which lies on the top plane of the loaded FR4 substrate. A defected ground structure is developed in this design to get isolation between WiMAX bands. A parametric study on the radius of circular patches and lengths of the rectangular microstriplines has been carried out on the designed antenna to afford all the crucial WiMAX operative bands (2.5/3.5/5.5 GHz). The designed antenna is compact ($23 \times 27 \times 0.8$ mm^3) when compared with the previously proposed multiband antennas. The simulated and measured responses show that the antenna is proficient to operate among the 2.3–2.4, 2.5–2.69, 3.3–3.8, 4–4.7, and 5.425–5.850 GHz frequency bands. The design parameters have been analyzed and measured for validation.

Keywords Defected ground structure (DGS) · MIMO · Worldwide interoperability for microwave access (WiMAX) · Wireless local area network (WLAN) · UWB antenna

T. K. Sreeja (✉) · K. Chandrakala · R. M. Nair · A. S. Vasu
Noorul Islam University, Thuckalay, India
e-mail: sreejaeng07@gmail.com

K. Chandrakala
e-mail: chandrakala.k@hotmail.com

R. M. Nair
e-mail: remya.nair116@gmail.com

J. Jayakumari
Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India
e-mail: jkumaribharat@yahoo.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_67

1 Introduction

Design of multiband antenna for mobile communication systems has become an attractive area of recent research. Accommodating several systems of various frequencies into limited area, including antennas, circuit components are mainly required to be multiple-frequency, efficient and compact. Synchronously operating WiMAX and WLAN technology is the fundamental application of wireless communication systems. The specific spectrum for WiMAX is resonated at 2.5, 3.5, and 5.5 GHz and for WLAN is resonated at 2.4, 5.2, and 5.8 GHz. Plenty of multiband antenna design approaches for compact handheld devices has been accomplished in many literatures. However, almost all the antennas are reasonably outsized and they failed to contribute preferred bandwidths.

Bandwidth enhancement and size compression are achieved by adding different shapes of patches in both of the conducting layers. These characteristics have been obtained by adding multiple resonant paths or many resonant elements in the significant frequency bands in the past literatures [1–5] for WLAN/WiMAX applications, rectangular slot, Y-shaped monopole patch, and two symmetrical inverted L-shaped strips along with split-ring-shaped slot [1], L- and square-shaped slots [2], a modified fork-shaped strip [3], a circular ring, and Y-shaped strip with a loaded ground structure [4], but the antennas are outsized. A multiband monopole antenna is investigated [6] using two inverted L-shaped strips with cross-shaped stripline and DGS, but the antenna efficiency remains to be 60% at 2.4 GHz and 70% at 3.5 GHz. A multiresonator module using a T-shaped stub and a pair of edge resonators including loaded structure was developed in [7] for WiMAX, and the gain measured at lower WiMAX band was low (1.1 dB) and the structure was complicated. Asymmetric coplanar strip-fed antenna using two special open-ended slots was designed [8], but the gain measured at 2.4 GHz is 0.71 dB and 1.95 at 3.5 GHz. A 48 mm $\times$ 18 mm $\times$ 0.8 mm sized multiband structure for GPS/WiMAX/WLAN systems has been investigated in [9], in which the efficiency and gain are lower than the other reported [1–4] monopole WiMAX antennas. A monopole hybrid fractal antenna for multitudinal wireless communication bands is demonstrated in [10] for mobile handheld devices using multiple-input-multiple-output (MIMO) technique. When Koch curves with Minkowski island curve fractal antennas are used, the antenna resonates at one WiMAX band and the antenna efficiency becomes 83%. Introducing a T-shaped patch at the top plane of dielectric layer enhances the isolation between frequency bands and impedance matching of antenna [11]. A novel structure, using star- and triangular-shaped fractal microstrip-fed printed antenna, along with semi-elliptical-shaped DGS is recommended for super-wideband communications in [12], with a gain of 1.4 dBi. It is noticed that enhanced gain and efficiency are obtained by increasing the size of antenna with -22.77 dB reflection coefficient. Hence, boosting the gain and bandwidth of MSA has emerged as a demanding investigation area. It is reported in [13] that, impedance bandwidth is improved up to 50% by incorporating rotated square-shaped slot at the bottom plane of dielectric layer and the gain is enhanced from 3 dBi to 6 dBi. Likewise, several monopole structures have been investigated by using various slots in [14–16]. However, the structures are complicated. Reasonable approaches with loaded structures named DGS have been designed in [17–20]. Circular polarization with dual band in all planes has been achieved using truncated ground structure MPA [17]. It has been proposed in [18] that the cross-polarization level is concealed by MPA with DGS. The equivalent circuit model of MSA as well as DGS has been developed in [20–22]. A log-periodic dipole MSA of 11.5 mm × 78 mm size is demonstrated in [23] for UWB applications. However, the wideband is obtained by enlarging the size. In the present study, a novel efficient patch antenna prototype is demonstrated by constructing a WiMAX antenna. The designed antenna is 23 mm × 27 mm × 0.8 mm sized and resonates between the four WiMAX bands such as 2.3–2.4 GHz/2.5–2.69 GHz/3.3–3.8 GHz/5.425–5.850 GHz, and 4–4.7 GHz WLAN band. The prototype without defected ground structure operates as UWB antenna in the range of 2.6–9.3 GHz. The design parameters have been analyzed and measured for verification. The proposed structure is simple and low cost, providing 100% efficiency, and it is applicable for WiMAX/WLAN applications. The investigation of designed WiMAX/WLAN antenna is interpreted in two sections; in the initial section, the antenna design is illustrated. Second section demonstrates parametric studies and results in comparison with the measured values.

2 Antenna Design

The pictorial representation of proposed coaxial probe-fed multiband-loaded monopole antenna is shown in Fig. 1. The antenna module is composed of a triangular patch and six circular capacitively coupled feeds to boost the impedance bandwidth. A crescent-shaped slot is introduced to minimize the losses. The six circular patches are coupled with six striplines that are linked with a triangular patch

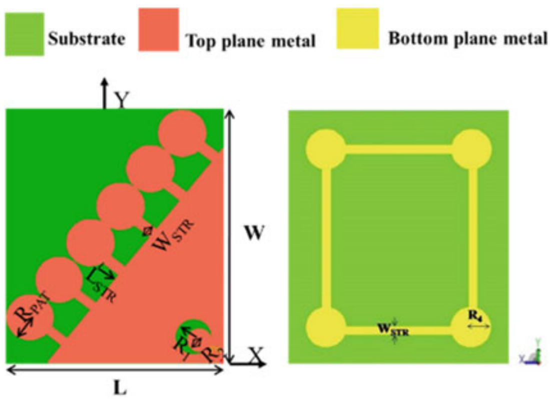


Fig. 1 Geometry of the proposed antenna

(a)Top plane view (b) Bottom plane view

which is on the top plane of the grounded FR4 dielectric sheet. Combination of four-annular patch [24] model DGS is employed in the design to get isolation between WiMAX bands.

The developed antenna design is based on compact and efficient MSA, but with reasonably very high size compared with a guided quarter wavelength at the smallest resonant frequency and does satisfy maximum requirements of WLAN/WiMAX applications. The main objective of this investigation is based on the design of a compact multiband antenna with omnidirectional radiation pattern at all the essential WiMAX frequency bands with maximum efficiency. The design of six microstripline feed circular patches depends on the condition that the six circular patches should be integrated within a quite compact antenna structure whereas they accommodate all the essential WiMAX bands. To achieve this target, the length of the triangular patch is chosen in a method that the entire length (L_{SUB}) of the antenna structure is nearly a quarter wavelength of the guided wave (λ_g) at favorable cutoff frequency using the approximation of effective dielectric constant relation. Therefore, the center cutoff frequency at 3.4 GHz (WiMAX band) is offered by triangular patch length ($L =$ 23 mm) and this corresponds to the quarter of guided wavelength equivalent to the resonant cutoff at 3.4 GHz band. The total width of the antenna structure (W_{SUB}) controls the initial resonant response 2.4 GHz, and this is equivalent to the quarter wavelength of the guided wave at 2.5 GHz. The highest frequency band is varied by the summation length (L_{GND}), width (W_{GND}) of the bottom conducting plane, and radius of the circular patch (R_4). This total length $L_t = L_{\mathrm{GND}} + W_{\mathrm{GND}} + R_{\mathrm{PAT}}$ is nearly the quarter length of the guided wave at 5.7 GHz. A crescent-shaped slot is introduced to minimize the losses at resonant frequencies.

The designed antenna setup is fed by 50 Ω impedance coaxial probe input impedance for the appropriate matching between the input and output, which is moderately loaded by a ground plane structure to establish the connection of the antenna to an external circuitry. Depending on this design conceptualization, the optimization of the proposed antenna dimensions together with the substrate and the ground plane sizes is calculated by a parametric analysis. Tremendous simulation process was carried out for computing reflection coefficient, surface current distributions, far-field radiation patterns, and gain using Agilent Advanced electronic system Design System (ADS) Software. The designed antenna structure imprinted on a 0.8 mm thickness dielectric material (FR4) along with the permittivity constant and loss tangent is 4.5 and 0.024, respectively. Total substrate dimension of the antenna is 23 mm × 27 mm × 0.8 mm. The circular patches of radius 2.5 mm are connected with 2 mm × 1 mm striplines, which are joined with the triangular patch (Fig. 1a). The ground plane is imprinted with an annular patch structure (Fig. 1b). Signal transmission between the antenna and an external device is provided via a 50 Ω SMA connector. The examination of excited current distribution on the body of the proposed antenna is illustrated in Fig. 2.

Figure 2 illustrates the simulated current distributions on the surface, at the three resonant frequencies of 2.4, 3.4, and 5.7 GHz on the edges of the proposed antenna structure. It is observed from the results that in the first and second frequency bands, a huge current density is oriented all along the area of triangular patch near the feed

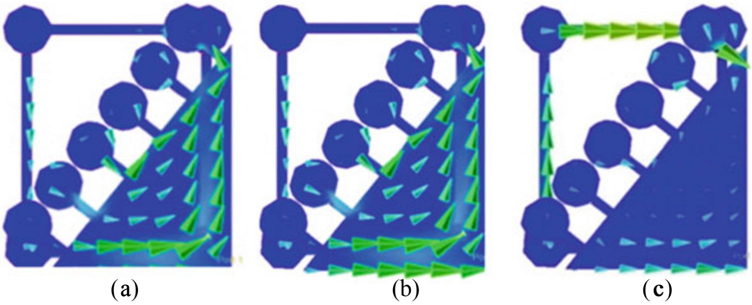


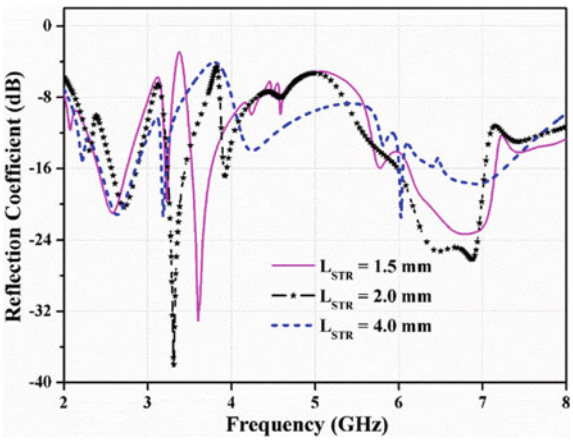
Fig. 2 Current distribution on the surface of proposed antenna at **a** 2.4 GHz **b** 3.4 GHz **c** 5.7 GHz

point (Fig. 2a, b). For upper frequency band, the surface current distribution is more along the edges of the ground plane and in the circular patches (Fig. 2c).

The consequence of stripline length (L_{STR}) and circular patch radius (R_3) on the reflection coefficient is revealed in Fig. 3. An ordinary patch without DGS structure exhibits a wide operating frequency band between 2.6 and 9.3 GHz and it is pictured in Fig. 5. Nevertheless, this wideband uncovers the essential WiMAX bands; precisely, the lower resonant band. To operate the antenna for WLAN/WiMAX wireless communication applications, DGS structure is used and parametric study has been done to get desired frequency bands. Optimizing the sizes of the proposed antenna using parametric analysis and introducing circular patches, along with stripline, accomplish the effective realization of the antenna. The characteristics of the antenna using the optimum size of the circular patches, crescent-shaped slot, and stripline are pictured in Fig. 4.

The optimum parameters of the proposed antenna module are shown in Table 1. The five resonant responses with bandwidths particularly for $S_{11} \leq -10$ dB, of about 100 MHz (2.3–2.4 GHz), 190 MHz (2.5–2.69 GHz), 500 MHz (3.3–3.8 GHz),

Fig. 3 Simulated values of reflection coefficient for various L_{STR} at constant R_{PAT}



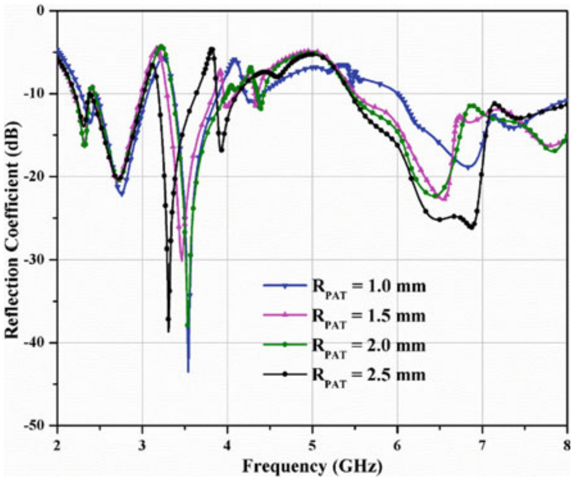


Fig. 4 Simulated values of reflection for various R_{PAT} at constant L_{STR}

Table 1 Design specifications

Parameter	Units (mm)	Parameter	Units (mm)
L	23	W	27
R_1	2	R_2	1.5
L_{STR}	2.5	W_{STR}	1
R_{PAT}	2.5	R_4	2.5
L_1	13	L_2	17
H	0.8	(X_0, Y_0)	(23, 1)

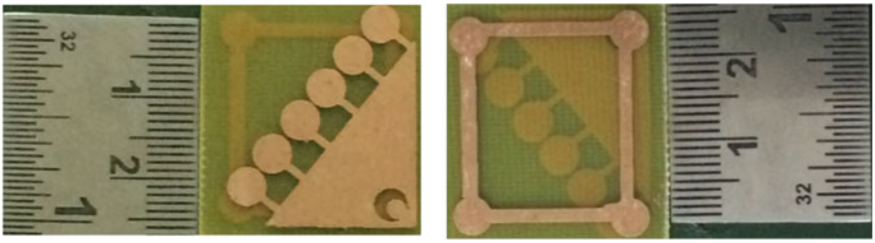
700 MHz (4–4.7 GHz), and (5.425–5.850 GHz) are measured at frequencies 2.3, 2.5, 3.3, 4, and 5.425 GHz, respectively.

3 Results and Discussions

Optimization of the designed antenna dimension has been computed from the parametric analysis. The antenna structure with optimum parameter is constructed, examined, and simulated; investigational results for the magnitude of reflection coefficient along with radiating characteristics are discussed and compared. Length of striplines and radius of patches are the sensitive values in influencing the impedance match. The consequence of L_{STR} on the magnitude of reflection coefficient (S_{11}) is demonstrated in Fig. 3. When the value of L_{STR} increases from 1.5 to 4 mm, the second cutoff frequency diminishes from 3.6 to 3.2 GHz. The WiMAX resonant band from 3.3 to 3.8 GHz is gained by properly tuning L_{STR}. The upper resonant band can be adjusted by suitably controlling the radius of the circular patch. The simulated magnitude of reflection coefficient for various values of patch radius (R_{PAT}) is plotted in Fig. 4.

Reflection coefficient magnitude of upper frequency band of the proposed antenna is enlarged by increasing R_{PAT} from 1.5 to 2.5 mm as plotted in Fig. 4. By conducting a parametric analysis at distinct values of R_{PAT} and $L_{STR,}$ we obtained the ultimate response and impedance matching. The magnitude of the reflection coefficient is measured by gradually increasing R_{PAT} that increases the magnitude of reflection coefficient and decrease in starting frequency of upper resonant band from 6 GHz to 5.48 GHz by maintaining L_{SUB} and W_{SUB} constant. It is observed from Fig. 4 that the required resonant responses are achieved at $R_{PAT} = 2.5$ mm. The proposed antenna was constructed and examined using an Agilent vector network analyzer 8722ES (VNA). The photography of top and bottom view of the designed antenna module is demonstrated in Fig. 5.

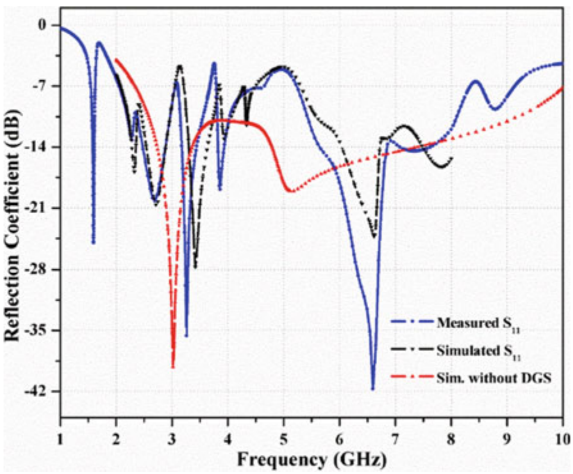
The numerical reflection coefficient values are compared with the measured reflection coefficient values for the constructed antenna as presented in Fig. 6. This can obviously experiment that there is a good concurrence among the examined and simulated reflection coefficient values.



(a) (b)

Fig. 5 **a** Top and **b** bottom view of the fabricated antenna

Fig. 6 Simulated and tested reflection coefficient of the proposed antenna using optimum values



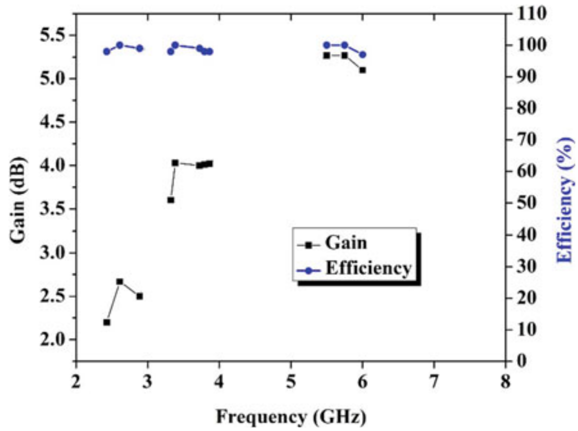


Fig. 7 Gain and efficiency simulated values for the proposed antenna

The negligible deviation among the two results is approximated due to the tolerance levels of dielectric substrate (FR4) specifications. The simulated antenna peak gain and efficiency over the specified bands are plotted in Fig. 7. The examined gain ranges from lower to upper resonant bands are 2.2–2.67, 2.67–4.03, and 4.03–5.27 dBi.

The examined and simulated radiation patterns at all WiMAX bands 2.5, 3.5, and 5.7 GHz for E and H plane are disclosed in Fig. 8. As seen in Fig. 8, the constructed antenna radiation pattern matches reasonably with the simulated radiation pattern results and the fabricated antenna structure offers omnidirectional radiation characteristics for the essential resonant frequency bands. It is noticed from the radiation characteristics that the proposed antenna offers low cross-polarization. It is because of the thin substrate material and the circular patch arrangement. The relative permittivity, substrate thickness, and the antenna arrangement are the critical parameters in deciding the cross-polarization level. However, there is no proper relationship among substrate permittivity and cross-polarization level [24]. Increased relative permittivity can strengthen the electrical length between the ground and microstripline, and hence, high energy is discharged. But the geometrical dimension of the microstripline reduces due to more relative permittivity, and hence, the radiation can be less. Therefore, the cross-polarization value can be diminished by thinning the substrate. However, antenna mechanical behavior will be affected if the substrate thickness is very less.

4 Conclusion

A novel efficient MSA with the loaded ground plane is constructed for multiband utilizations. The designed antenna comprises of six circular patches with stripelines, which are connected with a triangular patch in the top conducting layer, which

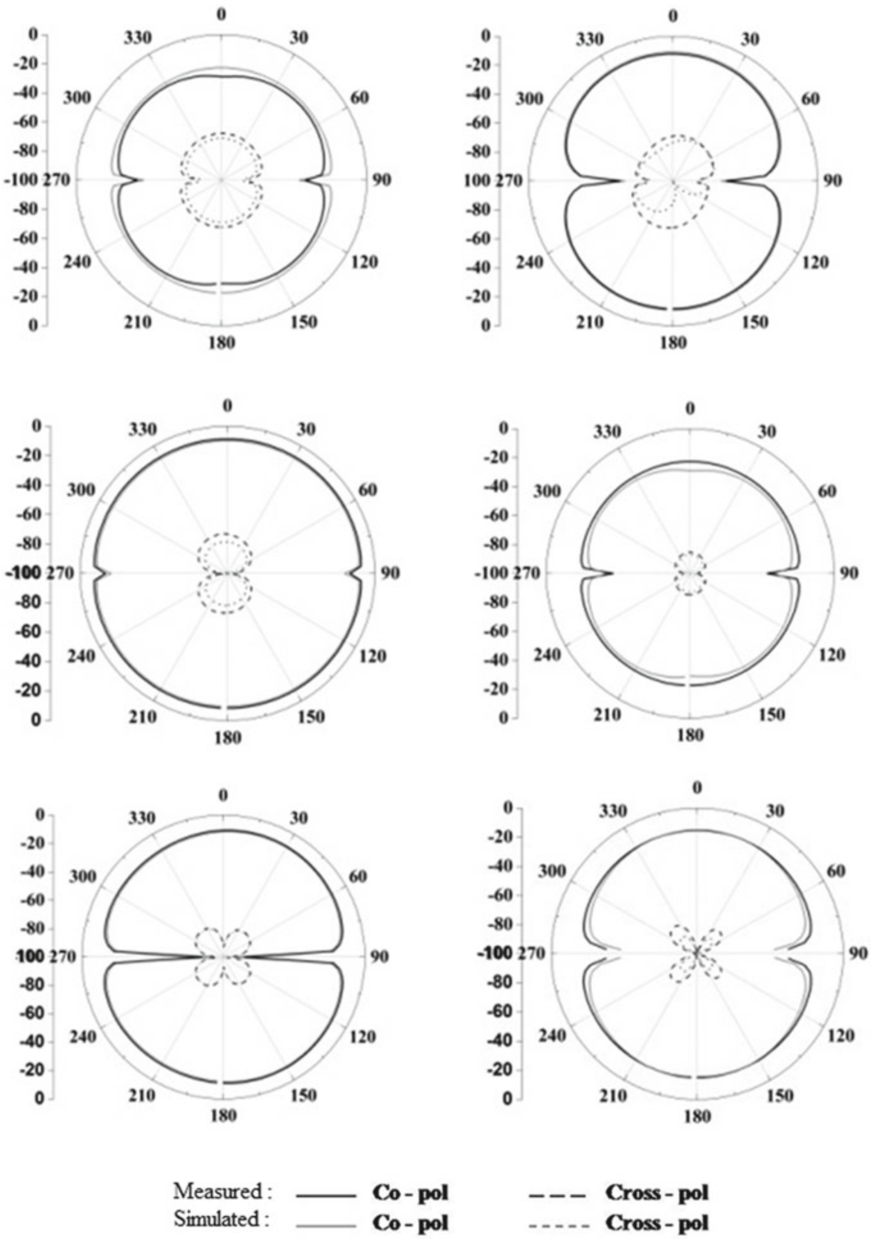


Fig. 8 Tested and simulated E-plane (left) and H-plane (right) radiation characteristics of the proposed antenna module for **a** 2.54, **b** 3.7, and **c** 5.65 GHz

enable proper adjusting of resonant responses. Simulated and tested characteristics are coincided. The preferred bandwidths, and gain for WiMAX (2.5/3.5/5.7 GHz) and WLAN (2.4/5.2 GHz) communication are fulfilled. The designed MSA without loaded ground plane acts as a wideband antenna. The antenna is of reasonably small dimension ($23 \times 27 \times 0.8$ mm^3) and can be a valuable option for UWB/WiMAX and WLAN applications because of its low cost, compact size, omnidirectional radiation characteristics, and multiresonance behavior and is highly efficient over all the preceding operating frequency bands.

References

1. Liu P, Zou Y, Xie B, Liu X, Sun B (2012) Compact CPW-fed tri-band printed antenna with meandering split-ring slot for WLAN/WiMAX applications. IEEE Antennas Wirel Propag. Lett 11:1242–1244
2. Hu W, Yin YZ, Fei P, Xi Y (2011) Compact triband square-slot antenna with symmetrical L-strips for WLAN/WiMAX applications. IEEE Antennas Wirel Propag Lett 10:462–465
3. Xu P, Yan ZH, Wang C (2011) Multi-band modified fork-shaped monopole antenna with dual L-shaped parasitic plane, Electron Lett 6(47)
4. Pei J, Wang AG, Gao S, Leng W (2011) Miniaturized triple-band antenna with a defected ground plane for WLAN/WiMAX applications. IEEE Antennas Wirel Propag Lett 10:298–301
5. Sreeja TK, Sandhya TK, Radha K, Kumari JJ (2012) An efficient multiband antenna for satellite and 4G communication systems. In: International IEEE conference on green technologies (ICGT). Trivandrum, pp 320–324
6. Liu WC, Wu CM, Dai Y (2011) Design of triple-frequency microstrip-fed monopole antenna using defected ground structure. IEEE Trans Antennas Propag. 59:2457–2463
7. Hu W, Yin YZ, Xi Y, Fei P (2013) Compact multi resonator-loaded planar antenna for multiband operation. IEEE Trans Antennas Propag 61:2838–2841
8. Li X, Shi XW, Hu W, Fei P, Yu JF (2013) Compact triband ACS-Fed monopole antenna employing open-ended slots for wireless communication. IEEE Antennas and Wireless Propag Lett. 12:388–391
9. Cao YF, Cheung SW, Yuk TI (2015) A multiband slot antenna for GPS/WiMAX/WLAN systems. IEEE Trans Antennas Prop 63:952–958
10. Choukiker YK, Sharma SK, Behera SK (2014) Hybrid fractal shape planar monopole antenna covering multiband wireless communications with MIMO implementation for handheld mobile devices. IEEE Trans Antennas Propag 62:1483–1488
11. Wang P, Wen GJ, Huang YJ, Sun YH (2012) Compact meander t-shaped monopole antenna for dual-band WLAN applications. Int J RF Microwave Comput-Aided Eng. Wiley Periodicals, Inc., pp 1–7
12. Waladi V, Mohammadi N, Zehforoosh Y, Habashi A, Nourinia J (2013) A novel modified star-triangular fractal (MSTF) monopole antenna for super-wideband applications. IEEE Antennas Wireless Propag. Lett 12
13. Jan JY, Su JW (2005) Bandwidth enhancement of a printed wide-slot antenna with a rotated slot. IEEE Trans Antennas Propag 53:2111–2114
14. Moosazadeh M, Kharkovsky S (2014) Compact and small planar monopole antenna with symmetrical L- and U-shaped slots for WLAN/WiMAX applications. IEEE Antennas Wirel Propag Lett 13:388–391
15. Xu Y, Jiao YC, Luan YC (2012) Compact CPW-fed printed monopole antenna with triple-band characteristics for WLAN/WiMAX applications, Electron Lett 48
16. Liu HJ, Li RL, Pan Y (2014) A multi-broadband planar antenna for GSM/UMTS/LTE and WLAN/WiMAX handsets. IEEE Trans Antennas Propag 62:2856–2860

17. Gautam AK, Kanaujia BK (2013) A novel dual-band asymmetric slit with defected ground structure for circular polarization operation. Microw Opt Technol Lett 55:1198–1201
18. Guha D, Biswas M, Antar YMM (2005) Microstrip patch antenna with defected round structure for cross polarization suppression. IEEE Antennas Wirel Propag Lett 4:455–458
19. Caloz C, Okabe H, Iwai T, Itoh T (2004) A simple and accurate model for microstrip structures with slotted ground plane. IEEE Microw Wirel Comp Lett 14:133–135
20. Hong JS, Karyamapudi BM (2005) A general circuit model for defected ground structure in plannar transmission lines. IEEE Microw Wirel Comp Lett 15:706–708
21. Sreeja TK, Jayakumari J (2018) Design and analysis of compact T shape slotted patch antenna for 4G WiMAX applications. Int J Enterp Netw Manage 9:1–10
22. Sreeja TK, Brightlin BC, Balamurugan S, Kanyakumari J (2017) Multiband layered patch antenna using nanoengineered dielectric substrate BaFe12O19. Adv Sci Eng Med 9:91–97
23. Khandelwal MK, Kanaujia BK, Gautam AK (2013) Low profile UWB log-periodic dipole antenna for wireless communication with notched band. Microw Opt Technol Lett 55:2901–2906
24. Jingjian H, Xiaofa Z, Xie Shaoyi, Weiwei W, Naichang Y (2014) Suppression of cross-polarization of the microstrip integrated balun-fed printed dipole antenna. Int J Antennas Propag. Article ID 765891

A Robust Tamil Text to Speech Synthesizer Using Support Vector Machine (SVM)

A. Femina Jalin and J. Jayakumari

Abstract Speech synthesis systems aim at generating high-quality, natural-sounding speech. Synthesizers become more static due to the inaccessibility of large databases for Indian languages. The prominence of the artificial speech made by these synthesizers is poor. Though statistical-based technique on hidden Markov models (HMMs) and Gaussian mixture model (GMM) is a powerful technique in Text to Speech (TTs) synthesis, recent work in TTS has concentrated on support vector machine (SVM). In this paper, a TTS system is developed for Tamil language using SVM. By using SVM, better sensitivity measures for Tamil Text to Speech are obtained. Output portrays that SVM can produce effectively natural speech compared to HMM and GMM.

Keywords Tamil language · Tamil unicode · SVM · Accuracy · Sensitivity · Precision

A. Femina Jalin (✉)
Department of Electronics and Communication Engineering, Noorul Islam University, kumarakoil, Tamil Nadu, India
e-mail: femina.jalin@gmail.com

J. Jayakumari
Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India
e-mail: jayakumari.j@mbcet.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_68

1 Introduction

Speech embraces an essential role in this developing world not only by means of articulated language, but then it is the most capable way of communication. In recent years, speech is considered as a research topic due to the complete inspiration of humans in their day-to-day lives [1]. Numerous speech synthesis techniques have been considered and studied to see how humans do these tasks effortlessly by using machines. Mainly, the challenge of Text to Speech (TTS) aims at how a text is converted into speech by producing natural-sounding speech.

In Text to Speech, the input text is converted into a sounding speech. Intelligibility, naturalness, and accuracy are the main characteristics of the TTS system. It indicates how the speech is easily understood and sounds natural by means of pronunciation, timing, etc.

There are many TTS systems available in many languages. Compared to other languages, no more research work is done in Tamil language due to the complexity in the creation of database and less knowledge about the Tamil language [2]. Therefore, no proper techniques are available for Tamil language. The challenges faced in Tamil TTS are: The glitches handled by Tamil grammar is very high, familiarity of Tamil language is very less, there is a massive gap in the middle of the Tamil colloquial speech and the pure Tamil which is written, and also Tamil phonetics and the database development is a risky job. Our main task is to develop a Tamil TTS. In our work, support vector machine (SVM) is used for the synthesis process. When compared to hidden Markov model (HMM) and Gaussian mixture model (GMM), SVM produces a better quality of the speech.

1.1 Background

Many research works are completed in the area of Text to Speech synthesis. Paul Taylor [1] describes the concept of how text input is converted into output speech. The basic ideas used in speech processing are explained in this work. Techniques used for speech synthesis are also explained in this paper. Concatenation-based speech synthesis gives better speech quality which is proved in this paper. A number of speech synthesis techniques [3] are described. Vocoder is the first device treated as a speech synthesizer, and it has become the key factor for the present techniques available.

Statistical parametric speech synthesis [4] is a HMM-based speech synthesis, which extracts the parametric exemplification of speech. HMM speech synthesis is open-source tool which provides a wide range of improvement for statistical parametric speech synthesis [5].

Automatic syllable segmentation for Indian languages is mainly concentrated. Initially, the syllable units selected as Indian languages are syllable-centric. For speech segmentation sub-word syllable [6] units are used. In this work, speech is first passed into the filter band. Therefore, it requires fewer phases for segmentation. Then, the analysis of sub-word and delay of segmentation of some known Indian languages are found.

Another work done [5] in this paper tried to add the prosodic features, so that, the synthesized speech will be more close to the natural speech. To get better quality of speech output, linear prediction is done. Thirukural [2] is proposed in this work. They worked to lessen the size and to bring a better quality of the speech. No prosodic features are added. To develop the naturalness of the speech, pitch modification algorithm is introduced.

2 Techniques of Speech Synthesis

For the conversion of TTS, four main synthesis techniques are used. They are (i) articulatory synthesis, (ii) formant synthesis, (iii) concatenative synthesis, (iv) statistical parametric synthesis, and (v) Gaussian mixture model (GMM). They are categorized by what means the speech is parameterized and the storage needed for synthesize.

2.1 Articulatory Speech Synthesis

This method refers to the best refined technique in terms of human speech production mechanisms and computation. It models the human speech production system by fine-tuning the position of tongue, lips, and jaw. Parameters used in such type of synthesizers are vocal cord position and sub-glottal pressure [7]. The drawbacks in this type are attaining precise speech sound is difficult, and the system is demonstrated with an inadequate set of parameters.

2.2 Formant Speech Synthesis

The most popular formant speech synthesis is also known as rule-based synthesis. It is of two types: serial and parallel formant synthesizers [2]. Formant synthesis is flexible and fairly easy to implement. Since the source parameters and the vocal tract model are very difficult to evaluate, unnatural quality of speech is produced by this method. To improve the quality of the synthesized speech, both serial and parallel types of formant synthesizers can be used. Another drawback is, since low memory is required, only small devices can run by this method.

2.3 Concatenative Speech Synthesis

Concatenation of segments of the recorded speech is done in this type of synthesis method. These synthesizers produce high-quality speech. High-quality speech output is produced with a large speech database. Suitable selection of unit is essential for better speech output. Natural-sounding speech and the flexibility get affected by this method. Corpus-based speech synthesis and diphone concatenation synthesis are the two sub-types of concatenative speech synthesis.

2.4 Statistical Parametric Speech Synthesis

The method is also termed parametric and statistic from the time when it defines the speech using parameters and statistics, rather than stored examples. Hidden Markov model is an example. Firstly, the features of speech are extracted from the database. Secondly, to build hidden Markov model, the extracted features and the training data are used. Thirdly, for each word, the speech parameters are produced. Then at last, from the estimated speech parameters, speech waveform is produced. The standard mathematical technique named hidden Markov model (HMM) is the most significant methodologies used for speech synthesis in the latest years [8]. The task performed in HMM is, from the recorded database, the parameters are extracted to generate the synthesizes speech. The welfares of Text to Speech synthesis system have its applications in multimedia, robotics, gesture recognition, and recently in energy desegregating. Observed variables and hidden variables are the two features used in HMM.

2.5 Gaussian Mixture Model

GMM is one of the simple techniques used for TTS with less complexity. It needs efficient feature extraction technique to train the GMM model [3]. When compared to the other techniques such as fuzzy and ANFIS, GMM is one of the accurate predictors. Like HMM, GMM also needs training and testing process. GMM is mainly used to model the feature distribution of speech. Though GMM has flexibility, the converted speech quality is poor than the natural speech. In order to convert the speech parameters, the target and the source speech features are modeled in GMM. The problem in GMM is it produces poor-quality speech output. The reason for the poor quality of speech is over smoothing effect [1, 5] synthesis errors [15] and inaccurate modeling [6].

The GMM function is defined as,

$$\mathrm{PGMM}(\chi) = \sum (\chi, \mu_n N_n = 1), \tag{1}$$

Here, n is the Gaussian component ($n = 1 \ldots N$), and N is the total number of Gaussian components. w_n is the component probabilities $\left(\sum w_n = 1\right)$, also called weights.

$$g(\chi, \mu_n, c_n) = \frac{1}{(2\pi)^{k/2}|c_n|^{1/2}} \mathrm{e}^{-\frac{1}{2}(\chi-\mu_n)^{\mathrm{T}} c_n^{-1}(\chi,\mu_n)} \quad (2)$$

where μ_n is the mean vector, and C_n is the covariance matrix.

3 Support Vector Machine (SVM)

Recent work in machine learning technique mainly focuses on support vector machine (SVM). SVM comes under supervised learning. It is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis [8]. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes comprises the input, making the SVM a non-probabilistic binary linear classifier.

To construct a nonlinear support vector classifier, the linear product (x, y) is replaced by a kernel function $k(x, y)$

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{n} l_i y_i (x_i, x) + m\right] \quad (3)$$

where the l_i are the Lagrange multipliers and m is the bias term.

Still, there is no natural speech produced in many situations by using TTS. To overcome that, machine learning algorithms are used. It generates some outstanding results in Text to Speech synthesis.

4 Proposed SVM Classifier Scheme

The block diagram of our proposed work consists of two phases: training phase and testing phase. Initially, different words are trained in our system. Speech samples from the database are collected with their respective unicode and produce its label output. The label is in numeric value. In the training phase, many words are given. Each and every letter in the word is trained using SVM. A net value or regression model is created. The range of Tamil unicode (Fig. 2) starts from 0BB0 and finishes at 0BFF. The representation of Tamil characters is the Tamil unicode. Each and every letter or symbol or digit is allocated in numeric value [6]. Unicode and their resultant voice signals are stored in the database. After converting to unicode, the corresponding

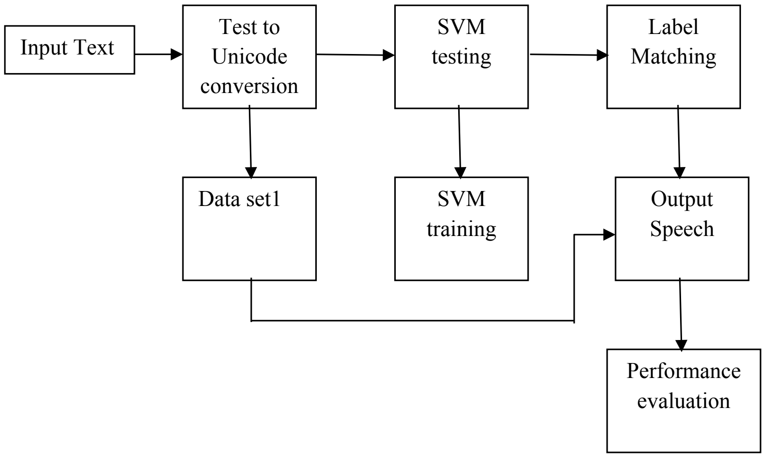


Fig. 1 Block diagram of SVM

voice samples are retrieved from the database and are applied to the trained SVM which will combine the voice samples to form words for producing speech output. In the testing phase, when the unicode is given, label value is generated. By this, particular word is called and the speech output is produced. Speech output is played depending on how exactly the label value occurs, and this gives the accuracy and other parameters (Fig. 1).

5 Results and Discussions

The words are recorded in a closed room with high-quality microphone using the recorded by a female speaker using the tool audacity. The recorded speech files are saved. For recording, the sampling rate used is 16 kHz. Sample words recorded are presented in Table 1. The output speech waveform generated for each word is shown in Fig. 3.

5.1 Senitivity Analysis

Sensitivity analysis is nothing but a tool used to perform quantative risk assessments which evaluate the relationship between the process parameters. Sensitivity analysis is normally executed to check the robustness of the results. Some of the analysis we have evaluated in our work is accuracy, sensitivity, specificity, precision, false measure, and geometric mean. Since SVM classifier is used, the entire system's performance is measured. Mainly, the performance is measured by the number of

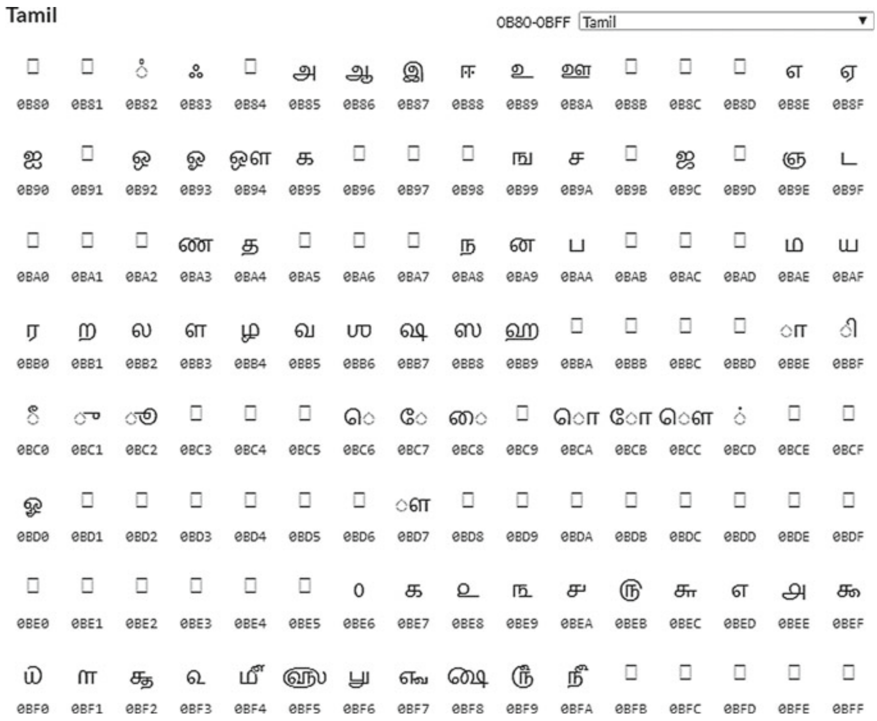


Fig. 2 Tamil unicode

Table 1 Recorded sample words and sentences

Sl No	words
1	அம்மா
2	அப்பா
3	யானை
4	பூனை
5	ஆமை
6	ஊசி
7	அலைவடிவம்
8	கதிர்வீச்சு

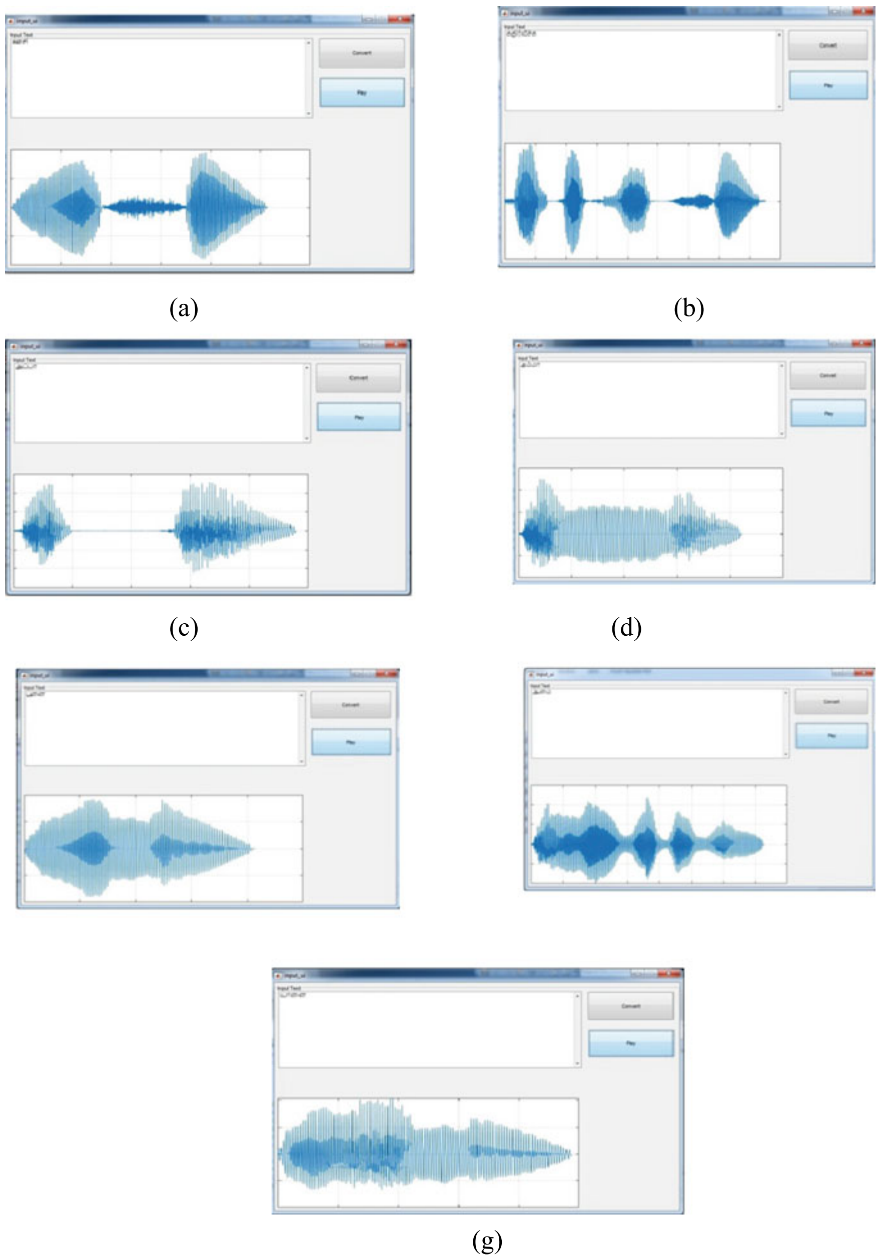


Fig. 3 Output waveforms generated for words using SVM technique

correct pronunciation of the words and the number of incorrect pronunciation of words. The output we are looking for is the positive measure and the other possibility is the negative measure. Performance parameters analyzed in our work are described as follows.

Recall defines what percentage of definite positives was recognized correctly. Mathematical equation for recall is given by

$$\mathrm{R} = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fn}} \tag{4}$$

Mathematical equation for precision is given by

$$\mathrm{Precision} = \frac{\mathrm{tp} + \mathrm{tn}}{\mathrm{tp} + \mathrm{tn} + \mathrm{fp} + \mathrm{fn}} \tag{5}$$

Sensitivity is a data collection and methods used to check how the output data are sensitive to the changes in the input data. Mathematical equation for sensitivity is given by

$$\mathrm{Sensitivity} = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fn}} \tag{6}$$

Specificity is also called true negative rate in which the proportion of the actual negative speech output is identified correctly. Mathematical equation for specificity is given by

$$\mathrm{Specificity} = \frac{\mathrm{tn}}{\mathrm{tn} + \mathrm{fp}} \tag{7}$$

Mathematical equation for accuracy is given by

$$\mathrm{Accuracy} = \frac{\mathrm{tp} + \mathrm{tn}}{\mathrm{tp} + \mathrm{fp} + \mathrm{tn} + \mathrm{fn}} \tag{8}$$

The false measure is a test accuracy measure which means the weighted harmonic mean of recall and precision.

where

tp is the true positive measure which predicts the correct pronunciation of words.
tn is the true negative measure which predicts the incorrect pronunciation of words.
fp is the false positive measure which predicts the correct pronunciation of words.
fn is the false negative measure which predicts the incorrect pronunciation of words (Table 2; Fig. 4).

Table 2 Sensitivity analysis output

Sensitivity measures	Percentage (%)
Accuracy	90
Sensitivity	100
Specificity	88.89
Precision	50
False measure	60.67
Geometric mean	94.28

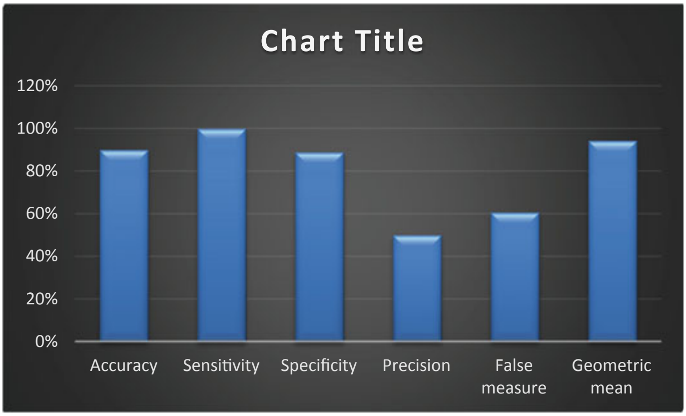


Fig. 4 Sensitivity analysis performance chart

6 Conclusions and Future Work

In this paper, we have developed an efficient TTS synthesis using SVM. Using this method, we could improve the accuracy of the system to 90% compared to HMM- and GMM-based techniques. The quality assessments like accuracy, sensitivity, specificity, precision, false measure and geometric mean we presented in this paper produce good quality speech when compared to HMM and GMM. Better naturalness is also achieved. As future work, we will be incorporating some sentences and will apply neural networks also in the decision process.

References

1. Joachims T (1999) SVMLight: support vector machine. Cornell University. http://www.cs.cornell.edu/People/tj/svm_light/
2. Keller E, Bailly F, Monaghan A (2002) Improvements in speech synthesis, COST 258: the naturalness of synthetic speech. Wiley, Chichester
3. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York

4. Burges CJC (1998) A Tutorial on support vector machines for pattern recognition. Knowl Discov Data Min 2(2):121–167
5. Rajeswari KC, Uma Maheswari P (2014) A novel intonation model to improve the quality of tamil Text to Speech synthesis system. In: Sixth international conference on advanced computing (ICoAC). IEEE, pp 335–340
6. Liu H, Yu L (2005) Toward integrating feature selection algorithms for classification and clustering. IEEE Trans Knowl Data Eng 17:491–502
7. Jelinek F (1976) Continuous speech recognition by statistical methods. Proc IEEE 64(4):532–537
8. Allwein E, Schapire RE, Singer Y (2000) Reducing multiclass to binary: a unifying approach for margin classifiers. J Mach Learn Res 1:113–141
9. Femina JA, Jayakumari J (2017) Text to Speech synthesis system for Tamil using HMM. In: 2017 IEEE international conference on circuits and systems (ICCS). IEEE, pp 447–451
10. Femina JA, Jayakumari J (2019) A novel Text to Speech technique for Tamil language using hidden markov models (HMM). Inter J Innovative Technol Exploring Eng (IJITEE) 8(10):38–47. ISSN: 2278–3075 (Online)
11. Jayavardhana Rama GL, Ramakrishnan AG, Muralishankar R, Prathibha P (2002) A complete Text to Speech synthesis system in Tamil. In: Proceedings of IEEE workshop on speech synthesis, Santa Monica, pp 191–194
12. Jose P, Govindaru V (2013) Malayalam Text to Speech. Inter J Eng Tech Res 1(3)
13. Zhong Yuan X, Yuan Bo-Ling X, Chong Zhi Y (1999) Binary quantization of feature vectors for robust text independent speaker identification. IEEE Trans Speech Audio Process 7(1):70–78

Large Number Multiplication by Repeated Addition

B. Sukrith and A. Sreekumar

Abstract One of the fundamental operations in computer algebra is the multiplication of multi-precision numbers. Some algorithms are available to implement these operations. Naive algorithms like high school mathematics (Neugebauer in the exact sciences in antiquity. Courier Corporation, Chelmsford, 1969) provide simple and good results for smaller numbers, while complex algorithms like FFT provide better results for larger numbers. Almost all algorithms are application-specific. We propose a novel and generic algorithm which is fast and suitable for cryptographic applications.

Keywords Integer multiplication · FFT · Complexity · Computer arithmetic · Finite field · Algorithm

1 Introduction

There are six major algorithms available for integer multiplication; other than school methods, these are versions of Chinese remainder theorem or fast Fourier methods. An amalgamation of many of these algorithms will make a better tool for multiplication of numbers with varied size input. Classical multiplication or graduate school multiplication that we do in our school, it has a complexity of $O(n^2)$. For historical reference, we refer to [1]. Karatsuba [2] algorithms use the divide and conquer approach; it divides the integer into two pieces and uses three evaluation points; hence, it has a complexity of $O(n^{\log 3/\log 2})$.

Toom [3] used more evaluation points and bettered the complexity and with further improvements by Knuth and Bibby [4] and Schönhage and Strassen [5]. Evaluation

B. Sukrith (✉) · A. Sreekumar
Department of Computer Applications, Cochin University of Science and Technology, Cochin, Kerala 682022, India
e-mail: sukrithb@cusat.ac.in

A. Sreekumar
e-mail: sreekumar@cusat.ac.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_69

and interpolation at large sets of points became a bottleneck for further improvements. Rediscovery of fast Fourier transform originally by Gauss [6], then by Cooley and Tukey [7] was efficient for special sets of evaluation interpolation.

Schönhage and Strassen [5] algorithm was the go-to algorithm for thirty years in integer multiplication. It consists of two methods, and the second method, which was primarily called as Schönhage Strassen [5] algorithm wherein it works over $R = \mathbb{Z}/m\mathbb{Z}$, where $m = 2^{2^k} + 1$ is a Fermat number [8]. The first method had its limitation of working but it showed a run time of $O(n \log n \cdots \log^{(l-1)} n(\log^l n)^2)$ for all l. It worked on $R = \mathbb{C}$, while approximating elements of $\mathbb{C}$.

Furer [9] proposed a method that combines the main advantages of Schönhage Strassen. This method reduces the overhead to $2^{O(\log * n)}$ from an overhead factor of $\log \log \log n \ldots 2^{\mathcal{O}(\log * n)}$. Harvey et al. [10] came up with a proposed system of order of complexity $O(n \log n K^{\log * n})$. We are not going in detail about these two algorithms as these are mostly theoretical approaches and do not consider the practical implementation problems and possibilities as well quoted in their own papers.

FFT-based algorithms work well for 2000 or more digits, and it is less efficient for smaller numbers, hence lies the importance of our approach, which is implemented using visual basic C programming.

2 Proposed Algorithm

In this paper, we propose a very simple and efficient algorithm for the multiplication of two integers, which can be implemented in hardware. In this method, we double one of the operands and reduce to half the other operand of multiplication continuously till the latter becomes 0 thus the name multiplication by repetitive addition.

In [4], Knuth and Bibby describe a method to double using Montgomery's algorithm, but we have seen that the proposed method doubles binary coded decimal (BCD) numbers more effectively. We assume that the multi-precision number is in BCD format. So each byte of number consists of two BCD digits. To double a BCD number, it is enough to add 3 to each and every digit of BCD number, which is greater than or equal to 5 and append a zero bit at the right side.

As an example, consider the number and its BCD format

$$\begin{array}{cccccc} 4 & 3 & 6 & 3 & 5 & 8 \\ 0100 & 0011 & \underline{0110} & 0011 & \underline{0101} & \underline{1000} \end{array}$$

Adding 3 with those underlined digits, we get

$$0100\ 0011\ \underline{1001}\ 0011\ \underline{1000}\ \underline{1011}$$

Appending a 0 to it, the number becomes

$$1000\ 0111\ 0010\ 0111\ 0001\ 0110,$$

which is clearly in the BCD equivalent format representation of number

$$2 \times 436358 = 872716,$$

The above addition can be replaced by using a lookup table, which contains entries like $M[01100011_2] = 10010011_2$ and $M[01110110_2] = 10101001_2$

Similarly, reducing the number to half could be done by throwing away the rightmost bit and subtracting 3 from the digits which are greater than or equal to 8. For example, the number 872716 represented in BCD format is

$$1000\ 0111\ 0010\ 0111\ 0001\ 0110$$

Removing last bit (shifting right 1 bit), it becomes

$$0100\ 0011\ \underline{1001}\ 0011\ \underline{1000}\ \underline{1011}$$

Subtracting 3 from the (underlined) digits which are greater than or equal to 8, the number becomes

$$\begin{array}{cccccc} 4 & 3 & 6 & 3 & 5 & 8 \\ 0100 & 0011 & \underline{0110} & 0011 & \underline{0101} & \underline{1000} \end{array}$$

This is BCD equivalent of half of 872716 that is 436358. Similar to the addition of 3 operation, this could be done effectively using another lookup table. Entries in this lookup table will be like $IM[10010011_2] = 01100011_2$.

We also introduce special cases, if the rightmost digit of the number is 0, 1, 5 or 6. In all such cases, we can manipulate 4 bits at once. The details are in Algorithm 1.

The first part of algorithm 1 defines how the multiplication is done by using repeated additions for special cases. Input to this algorithm is two integers a and b, which are predefined to 4 bit binary coded decimal. The advantage of this is that we can pack two digits in a single byte rather than general BCD, where we use one byte to store a digit. P is initially zero, and it is the resultant product of integers in BCD format.

Ending with 0: In this case, $b = 10 \times k$ for some k

$$P = P + a \times b \tag{1}$$

$$= P + a \times (10 \times k) \tag{2}$$

$$= P + (10 \times a) \times k \tag{3}$$

In this case, a is multiplied by 10 by shifting four times to the left, and b becomes k by shifting b four times to the right, similar to that of dividing by ten. In this case, P is unaltered.

Ending with 1: In this case, $b = 10 \times k + 1$ for some k

Algorithm 1 large number multiplication by repeated addition

```
Inputs: Integer a, b where a > b
Assume a and b in BCD format
P ← 0
while b > 0 do
  if b ends with "0000" then
    a ← shiftleft4(a)
    b ← shiftright4(b)
  else if b ends with"0001" then
    P ← P + a
    a ← shiftleft4(a)
    b ← shiftright4(b)
  else if b ends with "0101" then
    a ← shiftleft4(a)
    P ← P+half(a)
    b ← shiftright4(b)
  else if b ends with "0110" then
    P ← P + (a)
    a ← shiftleft4(a)
    P ← P + half(a)
    b ← shiftright4(b)
  else
    if (LSB(b) = 1) then
      P ← P + a
    end if
    a ← ∀digit(a) ≥ 5 add 3
    a ← shiftleft1(a)
    b ← rightshift1 (b)
    b ← ∀ digit(b) ≥ 8 reduce 3
  end if
end while
return P
```

$$\begin{aligned} P &= P + a \times b && (4)\\ &= P + a \times (10 \times k + 1) && (5)\\ &= (P + a) + (10 \times a) \times k && (6) \end{aligned}$$

Thus in this case, a will be added with P in addition to the procedure followed with the case of 0.

Ending with 5: In this case, $b = 10 \times k + 5$ for some k

$$\begin{aligned} P &= P + a \times b && (7)\\ &= P + a \times (10 \times k + 5) && (8)\\ &= \left(P + \frac{10 \times a}{2}\right) + (10 \times a) \times k && (9) \end{aligned}$$

In this case, a and b will be modified same as in the case of "ending with 0," and the new value of a is reduced by half using our proposed algorithm and added with P.
Ending with 6: In this case, $b = 10 \times k + 6$ for some k

$$P = P + a \times b \tag{10}$$
$$= P + a \times (10 \times k + 6) \tag{11}$$
$$= \left(P + a + \frac{10 \times a}{2}\right) + (10 \times a) \times k \tag{12}$$

Table 1 Example of implementation of algorithm with $P = 0$, $a = 436358$, $b = 90165$

Line No.	P	a	b
0	0	436358	90165
12	0	4363580	90165
13	2181790	4363580	90165
14	2181790	4363580	9016
16	6545370	4363580	9016
17	6545370	43635800	9016
18	28363270	43635800	9016
19	28363270	43635800	901
8	71999070	43635800	901
9	71999070	436358000	901
10	71999070	436358000	90
5	71999070	4363580000	90
6	71999070	4363580000	9
22	4435579070	4363580000	9
24	4435579070	43938B0000	9
25	4435579070	8727160000	9
26	4435579070	8727160000	4
27	4435579070	8727160000	4
24	4435579070	BA2A190000	4
25	4435579070	17454320000	4
26	4435579070	17454320000	2
24	4435579070	1A484320000	2
25	4435579070	34908640000	2
26	4435579070	34908640000	1
27	4435579070	34908640000	1
8	39344219070	34908640000	1
9	39344219070	349086400000	1
10	39344219070	349086400000	0

In this case, initially a will be added with P, then a and b will be modified as in previous cases. Finally, half of the value of new a is also added with P.

An illustration of the algorithm for the input $a = 436358$ and $b = 90165$ is given in Table 1. The intermediate values of the variables P, a and b are tabulated as the process progresses along with the corresponding line numbers in algorithm 1.

3 Conclusion

The proposed algorithm can be used for multiplication of two arbitrary large integers effectively as it packs two digits per byte in BCD format. Thus, we can save the total memory used. This algorithm is explicitly proposed for cryptographic application, keeping in mind the range of numbers, which is in 500 digits. Even though FFT-based algorithms outperform in much larger numbers, it takes more time for numbers which are in this range. Our algorithm uses simple addition and shift operation and lookup table so it could be easily implemented in hardware platforms. Major operations are bitwise addition operation, addition by three for doubling and reduction by three for reducing to half, and this operation is made faster by using a lookup table.

References

1. Neugebauer O (1969) The exact sciences in antiquity, vol 9. Courier Corporation, Chelmsford
2. Karatsuba A (1963) Multiplication of multidigit numbers on automata. Sov Phys Dokl 7:595–596
3. Toom AL (1963) The complexity of a scheme of functional elements realizing the multiplication of integers. Sov Math Dokl 3:714–716
4. Knuth DE, Bibby D (1984) The TEXbook, vol 3. Addison-Wesley, Reading
5. Schönhage A, Strassen V (1971) Schnelle multiplikation großer zahlen. Computing 7(3):281–292
6. Heideman MT, Johnson DH, Burrus CS (1985) Gauss and the history of the fast Fourier transform. Arch Hist Exact Sci 34(3):265–277 (cited by 60)
7. Cooley JW Tukey JW (1965) An algorithm for the machine calculation of complex Fourier series. Math Comput 19(90):297–301
8. Harvey D, Van Der Hoeven J (2019) Polynomial multiplication over finite fields in time o (n log n)
9. Fürer M (2009) Faster integer multiplication. SIAM J Comput 39(3):979–1005
10. Harvey D, Van Der Hoeven J, Lecerf G (2016) Even faster integer multiplication. J Complex 36:1–30

A Comparative Performance Analysis of Different Denoising Techniques in Sputum Smear Images

M. Shafeen Nagoor and S. Vinila Jinny

Abstract Tuberculosis (TB) is the deadliest, communicable disease and can nearly affect all the parts of our body. To diagnose this disease, there are various tests available such as chest X-rays and TB skin tests. The most commonly used tool for TB detection is the sputum smear microscopy (SSM) which is less costly. The specimens are stained using Ziehl–Neelsen and are then examined by the technicians for any microbes. Detection of bacilli from stained sputum images manually is a lengthy, far-reaching process which can lead to inaccuracy in the output. Therefore, automatic methods are provided which give an optimal solution in a short time, in the absence of skilled experts in disease diagnosis. This paper gives an overview of available preprocessing methods in various digital images and hence will benefit the researchers working in the smear microscopy field.

Keywords Tuberculosis · Preprocessing · Sputum images · Denoising techniques

1 Introduction

Tuberculosis is a terrible, dangerous disease which can even lead to death if not treated properly. In 2017, there was an estimated rate of 10 m TB cases in which 5.8 m cases were men, 3.2 m were women, and 1.0 m were children [1]. People who are sick with TB expel bacteria into the air, for example, by coughing and sneezing. TB does not spread by shaking someone's hand, sharing food, or sharing toothbrushes [2]. TB disease is classified as latent and active TB. People with latent TB infection have TB germs in their bodies, but they are not sick because the germs are not active. People with TB disease are sick because TB germs remain active, meaning that they are multiplying and destroying tissue in their body [3]. Smear microscopy has remained the cornerstone in the detection of tuberculosis. Sputum smear microscopy [4] is

M. Shafeen Nagoor (✉) · S. Vinila Jinny
Noorul Islam Centre for Higher Education, Kumaracoil, India
e-mail: shafeenb4u@gmail.com

S. Vinila Jinny
e-mail: vinijini@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_70

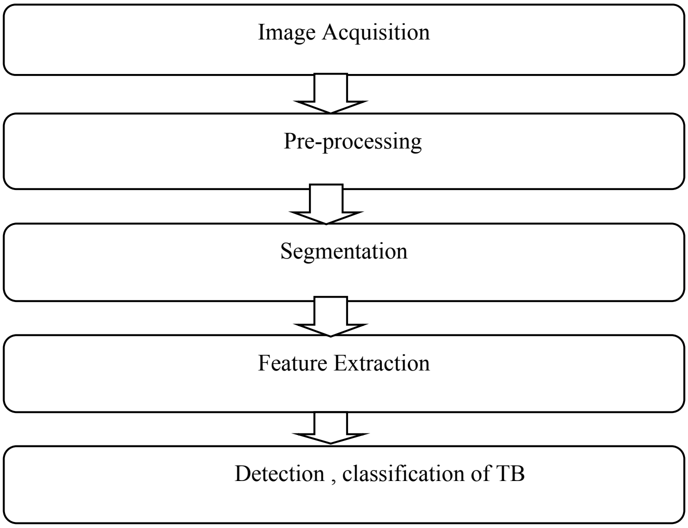


Fig. 1 Basic steps in TB detection

the simple and appropriate technology which is easy to perform. Ziehl–Neelsen [5] method is the recommended procedure for staining tubercle bacilli as it provides good results and leads to a better standards.

There are different steps associated with the early detection of tuberculosis as in Fig. 1. This paper gives a detailed survey of the preprocessing methods.

2 Lıterature Survey

Mithra et al. [6] proposed a model to identify and count the bacilli present in the image. The sputum image is preprocessed using denoising technique and intensity modification. The impulse noise present in the image is removed using adaptive median filtering and then subjected to grayscale for further processing. Location-oriented histogram algorithm (LoH) [7] and speedup robust feature algorithm (SURF) [8] are used in the extraction of intensity-based local bacilli features. Around 275 images from different categories like scanty, less, and touching bacilli of Ziehl–Neelsen image database (ZNSM-iDB) were used. This shows an accuracy of 97.55% which is better. This model removes the unreliable features and gives the infection level of disease more accurately.

Panicker et al. [9] proposed a system for the automatic detection of tuberculosis bacilli by deep learning methods. The input image is initially preprocessed to remove noise using fast non-local means method [10]. The preprocessed image is converted to grayscale, binarized using Otsu's method [11]. Morphological opening and closing are followed to fine-tune, and for the feature extraction, connected component analysis is used. A total of 1817 images such as bacilli and non-bacilli images were used for testing from the INPA laboratory, Brazil. The accuracy of the model is 97.73%. This model is used for predicting the disease accurately in a very limited time. Also, this method works well with single and touching bacilli too.

Chithra et al. [12] proposed a model having an accuracy 94.87%. In this model, the sputum image was subjected to color space transformation to generate the segments. Six features such as the coverage, density, color histogram, texture, bacilli area, and bacilli length were extracted. Ten different categories (empty bacilli, bacilli, and overlapping bacilli) of images from ZNSM-iDB were used for this experiment. This model is accurate, effective and increases the computational speed. The system turns out more complex when dealing with a large number of overlapping bacilli.

Shabut et al. [13] present a real-time expert system for tuberculosis detection. Images were scaled to reduce processing time and then Gaussian blur filtering was applied to the images. Total 18 features were used for training. A total of 252 images were used for processing given by the School of Medical Sciences (SMS), Malaysia. The accuracy of the method is 98.4%. Real-time multiple samples can be easily read and classified as positive and negative. Separation between clusters becomes a difficult task in spite of multistep filtering.

Raju et al. [14] proposed a system to automatically diagnose TB using convolutional neural network. As a preprocessing step, images from the dataset were intensified at the edges, cropped by identifying the foreground from background. The cropped images were then resized, and their pixel values were normalized. The image data set contained 678 images taken from Montgomery, Shenzhen database, India. This model observed a sensitivity of 84.91% and a specificity of 93.02%. There was an improvement in the prediction results due to data augmentation. Noise level was not reduced to a better extent, and the system must be trained on more images to achieve robustness.

Vajda et al. [15] proposed a model with accuracy showing 95.6% for the Shenzhen dataset. Scale-invariant feature transform flow (SIFT) approach was used to compute the corresponding pixels of image pairs. A lung atlas-based segmentation algorithm was used to extract the selected features like shape and curvature descriptor histograms, etc., manually followed by a classifier in the disease diagnosis. The classification error rate is minimized and reduces the inference time. The multistage pipeline used in this model is more complex and requires more development work, not freely usable.

Priya et al. [16] proposed a method to identify and classify sputum images at the image and also object level. The segmentation process is carried out using the active contour method. The preprocessing step is done using the illumination uniform correction method and then segmented to identify the bacilli and outliers. To select the important Fourier descriptors [17], a fuzzy entropy measure was used. The stained slides were prepared at the South African National Health Laboratory Services (NHLS), Cape Town. The accuracy of the model is 91.3%. Improved accuracy and higher sensitivity are achieved. Also, improved classification accuracy at both image and object level was seen but the outliers had no uniform morphology.

3 Image Preprocessing Methods

Image preprocessing improves the image data by applying a set of operations on an image. It is an important step to reduce variations and produce consistent data thereby improving the performance.

3.1 Noise Removal

Noise is simply a disturbance in an image which affects the image during image acquisition. There are several sources through which noise enters an image such as during electronic transmission of image data, presence of dust particles, and environmental conditions. All types of images contain some amount of visual noise. Noisy image gives a snowy, grainy, and textured appearance. Noise identification and algorithm application should be the prior concern. Noise is of various types as described.

Salt-and-pepper noise It is the most frequently seen noise in various images. Random black-and-white dots [18] appear on the images. The image signal when subjected to sharp and sudden changes causes this type of noise and also termed as impulse or spike noise.

Gaussian noise This model is additive [19] and applies the Gaussian distribution. In Gaussian noise, random values are added to the image and give even distribution over the signal.

Speckle noise It is a noise model that multiplies random values with pixel values of the image.

Poisson noise It is also termed as shot noise. The most dominant noise occurs in the lighter parts of an image and is caused by statistical quantum fluctuations.

3.2 Denoising Techniques

Image denoising remains a significant preprocessing step in the analysis of images. Reducing the noise level thereby while preserving the details is the major concern. Denoising methods are of two types: linear and nonlinear. The linear type does not preserve the edges, but are fast with speed, whereas the nonlinear type preserves the image details [20]. To remove noise from images, different filters were used in the system. Image filtering not only improves the image quality but is also used as a preprocessing step before most image processing operations as shown in Fig. 2. Without preprocessing, the other processing would go inappropriate or can even lead to false results.

Fig. 2 Basic steps in preprocessing input image

Median Filter This filter smoothens the data by keeping not only the small but sharp details also. It is simple, powerful, and nonlinear filter. The value of every pixel in the neighborhood is replaced by the median of intensity values [21]. This filter is useful mainly in salt-and-pepper noise.

Mean Filter Mean filter is also an averaging linear filter [22]. It reduces noise and also the amount of intensity variation between various pixels. It acts as a low-pass frequency filter. It calculates the average value and replaces the average value to the central pixel intensity value. The same step is carried out for every pixel values in the image.

Gaussian Filter It is an additive model which follows Gaussian distribution. Image is blurred by removing the noise and detail using this filter. This filter is separable and symmetric and used in image size reduction.

Adaptive Median Filtering It performs spatial processing to detect the image pixels affected by spike noise. A comparison of each pixel in the image to its surrounding neighboring pixels is done during classification level. The size of the neighborhood and threshold is calculated and adjusted. The noisy pixels are then replaced by the median pixel value in the neighborhood pixels.

4 Experimental Evaluation

4.1 Quality Measurements

The main parameter to evaluate the enhancement of images is decided by the peak signal-to-noise ratio (PSNR) and mean squared error (MSE).

Peak signal-to-noise ratio It gives the ratio between the value of the noisy image to that of the original image. It measures the quality between the original and the reconstructed image. The reconstructed image is said to have good quality when the PSNR value is high.

$$\mathrm{PSNR} = 10 * \log 10\,(255^2/\mathrm{MSE}) \tag{1}$$

Mean Squared Error It always remains non-negative and calculates the collective squared error between the reconstructed image and the original image. Lesser value of MSE indicates that the error is very low.

$$\text{MSE} = 1/MN \sum_{y=1}^{M} \sum_{x=1}^{N} \left[I(x, y) - I'(x, y) \right]^2 \quad (2)$$

where,

$I\ (x, y)$—original image and
$I'(x, y)$—reconstructed image
M, N—horizontal/vertical dimensions of the image.

4.2 Results

The filter parameters for different grayscale noisy images during preprocessing are shown in Tables 1, 2, 3 and 4. The quality parameters PSNR and MSE are calculated for estimating the performance analysis on different filters. Input grayscale image with size 256 × 256 is used for analysis as shown in Fig. 3. Performance of four different filters is tested for different noises. The best filtering method will show high PSNR and lower MSE.

In Table 1, the MSE is computed by testing 100 medical images for noise removal process based on Eq. 2. The averaging process decides the final MSE value. Table 1 shows how much MSE is produced after adding noises in the original image. A specified range of noises are added to the input image, and their MSE value is filled in the table. For MSE computation, the original image and noisy images are contributed. The corresponding PSNR values are computed using Eq. 1. Three types of noises are considered for this evaluation, and they are speckle noise, Gaussian noise, and salt-and-pepper noise. The salt-and-pepper noise highly damages the images, because it gains the less PSNR which is 25.3112. The Gaussian noise contaminates the image at less grade because it obtains high PSNR which is 28.2991.

Table 1 Performance parameters of input image on different noise

Noise	PSNR	MSE
Speckle noise	25.6057	178.8594
Gaussian noise	28.2991	100.7331
Salt-and-pepper noise	25.3112	191.4062

Table 2 Peformance parameters of different filters on images with Gaussian noise

Filter	PSNR	MSE
Mean filter	32.4262	209.1787
Median filter	35.5768	143.5727
Gaussian filter	32.4820	67.4698
Adaptive median filter	36.4970	143.5727

Table 3 Performance parameters of different filters on images with salt-and-pepper noise

Filter	PSNR	MSE
Mean filter	33.02128	221.3326
Median filter	44.31463	125.4986
Gaussian filter	40.9339	106.0367
Adaptive median filter	44.4320	125.4986

Table 4 Performance parameters of different filters on images with speckle noise

Filter	PSNR	MSE
Mean filter	32.8493	219.1137
Median filter	29.8782	177.5321
Gaussian filter	33.1563	99.9928
Adaptive median filter	34.7976	177.5321

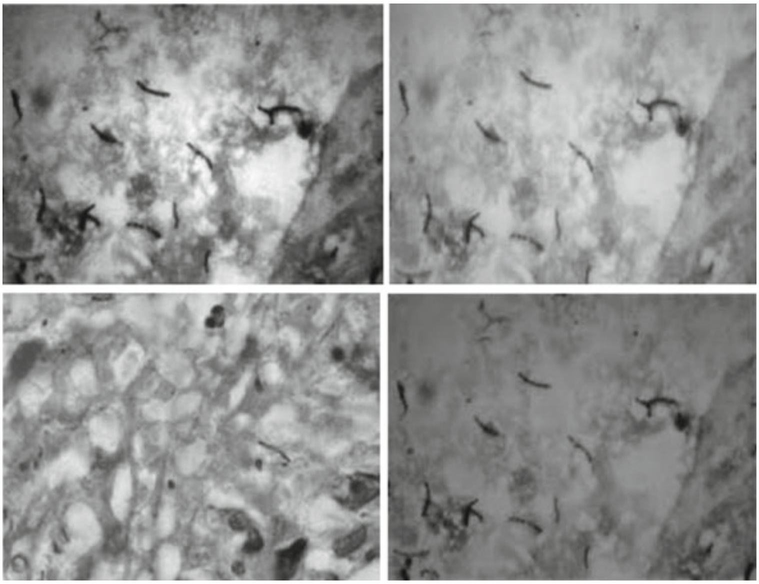


Fig. 3 Sputum smear input images

In Table 2, it is noted that there is a rise in PSNR value and also MSE value is reduced with Gaussian noisy images. Hence, preprocessing the input image with Gaussian filter is said to give better image quality. Also, in the mean filter, the PSNR value is increased with the increase in MSE. Next, with the median filter, PSNR value increases from 28 to 35.5768 and hence the median filter performs well than the average filter. Compared to mean, Gaussian filter, preprocessing with median filter gives better image quality. Lower value of MSE gives less error, and a higher range value of PSNR is preferred. From tables, it is sure that, after noise removal techniques, there remains a variation in the parameter values. Hence, during the preprocessing, the salt-and-pepper noisy image on the adaptive median filter gives the best quality of image when compared to Gaussian, mean, and median filter. However, an adaptive median filter is good for all the filtering process, and the output attained is said to have a better quality of the image.

Table 3 specifies that the adaptive median filter gains high PSNR (44.432 dB) and less MSE (125.4986) in case of salt-and-pepper noise. It proves the best performance of adaptive median filter than the other filters for salt-and-pepper noise. The mean filter reaches less denoising performance because it achieves high MSE (221.33) and less PSNR (33.02 dB).

Table 4 reveals the performance of the four filters when considering the speckle noise. The adaptive median filter shows the best performance by obtaining less MSE and high PSNR than other filters. The values of MSE and PSNR according to the adaptive median filter are 177.5321 and 34.7976 dB, respectively. The least performance providing filter related to speckle noise is mean filter because it gains low PSNR (32.84 dB)

5 Conclusion

The different denoising methods in images are discussed in this paper. The first section describes the literature review done. The next section describes the various types of noise models and different filtering methods used. In the last section, different performance parameters used to compare the effectiveness of filtering techniques are discussed. Mostly, the peak signal-to-noise ratio parameter is used to measure the effectiveness of any filter. Each filter works differently on different types of noises. Performance parameters for different filters are analyzed and conclude that the adaptive median filter is the best for noise removal during the preprocessing stage.

References

1. Organization WH (2018) Global tuberculosis report
2. Centers for disease control and prevention. https://www.cdc.gov/tb/publications/pamphlets/tb_disease_en_rev.pdf
3. Centers for disease control and prevention. https://www.cdc.gov/tb/publications/factsheets/general/tb.htm
4. Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, Urbanczik R, Perkins M, Aziz MA, Pai M (2006) Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis 6:1–3
5. Timalsina B, Kutu B, Pradhan R, Maharjan B (2015) A comparative study on ziehl-neelsen staining (light microscopy), auramine o staining (iLED fluorescent microscopy) and culture on LJ media of sputum samples for the diagnosis of pulmonary tuberculosis. Int J Med Biomed Sci 1:4–14

6. Mithra KS, Sam Emmanuel WR (2019) Automated identification of mycobacterium bacillus from sputum images for tuberculosis diagnosis. Sig. Image Video Process. https://doi.org/10.1007/s11760-019-01509-1
7. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal Mach Intell 27(10):1615–1630
8. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110:346–359. https://doi.org/10.1016/j.cviu.2007.09.014
9. Panicker RO, Kalmady KS, Rajan J, Sabu MK (2018) Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng 38:691–699. https://doi.org/10.1016/j.bbe.2018.05.007
10. Froment J (2014) Parameter-free fast pixelwise non-local means denoising. Image Process Line 4:300–326. https://doi.org/10.5201/ipol.2014.120
11. Smith P, Reid DB, Environment C, Palo L, Alto P, Smith PL (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 20:62–66
12. Chithra RS, Jagatheeswari P (2019) Fractional crow search-based support vector neural network for patient classification and severity analysis of tuberculosis. IET Image Process 13:108–117. https://doi.org/10.1049/iet-ipr.2018.5825
13. Shabut AM, Hoque Tania M, Lwin KT, Evans BA, Yusof NA, Abu-Hassan KJ, Hossain MA (2018) An intelligent mobile-enabled expert system for tuberculosis disease diagnosis in real time. Expert Syst Appl 114:65–77. https://doi.org/10.1016/j.eswa.2018.07.014
14. Raju M, Aswath A, Kadam A, Pagidimarri V (2018) Automatic detection of tuberculosis using deep learning methods. In: Springer proceedings in business and economics, pp 119–129. https://doi.org/10.1007/978-981-13-1208-3_11
15. Vajda S, Karargyris A, Jaeger S, Santosh KC, Candemir S, Xue Z, Antani S, Thoma G (2018) Feature selection for automatic tuberculosis screening in frontal chest radiographs. J Med Syst 42. https://doi.org/10.1007/s10916-018-0991-9
16. Priya E, Srinivasan S (2016) Automated object and image level classification of TB images using support vector neural network classifier. Biocybern Biomed Eng 36:670–678. https://doi.org/10.1016/j.bbe.2016.06.008
17. Luukka P (2011) Feature selection using fuzzy entropy measures with similarity classifier. Expert Syst Appl 38:4600–4607. https://doi.org/10.1016/j.eswa.2010.09.133
18. Rohit V, Ali J (2013) A comparative study of various types of image noise and efficient noise removal techniques. Int J Adv Res Comput Sci Softw Eng 3(10). ISSN 2277–128X
19. Patidar P, Gupta M, Srivastava S, Nagawat AK (2010) Image de-noising by various filters for different noise. Int J Comput Appl 9:45–50. https://doi.org/10.5120/1370-1846
20. Chan RH, Ho CW, Nikolova M (2005) Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. IEEE Trans Image Process 14:1479–1485. https://doi.org/10.1109/TIP.2005.852196
21. Maheswari D (2010) Noise removal in compound image using median filter. Int J Comput Sci Eng 2:1359–1362
22. Al-amri SS, Kalyankar NV, Khamitkar SD (2010) A comparative study of removal noise from remote sensing ımage. Inter J Comput Sci Issues (IJCSI) 7:32–36

Design and Implementation of Compact Economic Kitchen Waste Recycler Bin

D. Sindhanaiselvi and T. Shanmuganantham

Abstract Our country is facing a lot of challenges in clearing the wastes and reducing the pollution. The wastes are created by so many sources and dumped as a landfill in almost all towns and cities and major source of pollution. Out of the total wastes, 20% is from the household kitchen wastes. This is due to the contradiction between high food consumption and low food recycling rate. The food waste or kitchen waste from the house is treated and decomposed properly, to reduce food waste capacity and converting as organic manure. But the customary composting of waste takes six months to a year to decompose. It also requires regular maintenance for the compost pile, and collecting scraps to take outside can be smelly and attack bugs. This project gives an innovative solution to recycle the food or kitchen waste from each house. The solution is to design an eco-friendly compact bin placed in every kitchen of the house that converts food waste to organic fertilizer. Unlike most composting device which dehydrates the scrap is actually a huge technical challenge to provide consistent quality of organic fertilizer every time by this novel method. The organic manure from this recycler bin improves the soil health, reduces the need for additional water and fertilizers. In order to fasten the composting process, the recycler bin is constructed with heating, mixing and grinding sections under controlled conditions.

Keywords Recycler · Waste · Kitchen · Compost · Organic fertilizer

D. Sindhanaiselvi (✉)
Department of EIE, Pondicherry Engineering College, Pondicherry 605014, India
e-mail: sindhanaiselvi@pec.edu

T. Shanmuganantham
Department of Electronics Engineering, Pondicherry University, Pondicherry 605014, India
e-mail: shanmuga.dee@pondiuni.edu.in

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_71

1 Introduction

The day-to-day activity of human gives rise to large variety of different wastes from different sources. Each country is facing lots of challenge in clearing the waste contamination which is more expensive than prevention of this waste generated from its source.

Out of total wastes, 1.3 billion tons food gets wasted or lost in the overall world. To reduce this waste, food recycler is developed that converts food waste into fertilizer. This food or kitchen waste recycler is known as “low cost kitchen waste recycler.” Recent statistical analysis states that food or kitchen waste is growing much more rapidly than other wastes. According to a study, 20% landfill wastes in India are originated from homes. The existing system uses compost pit to degrade the waste to natural manure. It takes about 14 days to degrade due to microbial action. This proposed kitchen waste recycler brings a useful innovation and solution to reduce the landfills from kitchens. The food recycling process is infrastructure free and more reliable for both indoor and outdoor purpose. The decomposing process can be completed in less than two days.

In 2007, Belal Chowdhury, Morshed U. Chowdhury [1] constructed waste management using RFID and sensor module. In the year 2016, Chung-Horng Lung, Ioannis Lambadaris [2] proposed sensor-based automated waste management with cloud computing paradigm. R. Jenifer Prarthana, A. Shankar [3] developed the monitoring system with RFID technology to overcome the problems in the conventional approach. Alharith Almuhaimeed, Hadi Alkatheri [1] designed the compost machine which uses Acidulous microorganism that takes few days to decompose. In 2015, Fachmin Folianto, Yong Sheng Low, Wai Leong Yeow [4] constructed the smart bin system, and the recent research problems on MEMS-based sensors with Arduino-based control of different application are discussed in [5–9].

This kitchen waste recycler overcomes the long time decomposition by combining the three process such as heating, grinding and mixing. It delivers an eco-friendly solution that can be used in each home. It also acts as sustainable organic fertilizer for their outdoor lawns and gardens. The kitchen waste recycling process is compact and more reliable for both indoor and outdoor purpose. The decomposing process can be completed in less than two days.

In this paper, the introduction and proposed kitchen waste recycler to produce organic fertilizer is explained. The block diagram and sections of kitchen recycler are presented in Sect. 2. Section 3 details the hardware interfacing involved in implementation. The different sections of the kitchen waste recycle bin, sample input and organic fertilizer output are shown in Sect. 4.

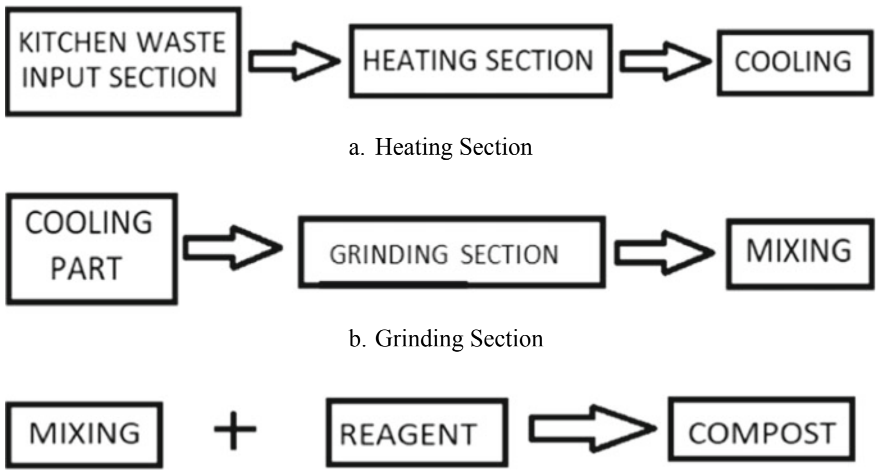
2 Block Diagram of Proposed Recycle Bin

The economic kitchen waste recycler consists of the following sections such as heating section, grinding section and decompost section to produce the organic fertilizer within two days. The low cost kitchen waste recycle consists of three parts is shown in the Fig. 1.

The vegetable waste is collected in a tray. A heating coil is used to heat the metal tray. The main objective of the heating section is to remove the moisture content from the vegetable waste. The waste is heated for about 30 m at a temperature of 80 °C. The waste is grinded in grinding section and then mixed with decomposing additives to make the fertilizer nutrient rich. The vegetable waste is heated and made to cool for 2 h. After the waste is cooled, it is transferred to grinding section. In the grinding section, the waste is grinded for 10 m. After grinding, size of the waste is reduced to one fourth of its initial size, and it is transferred to decompost section for addition of reagent such as baking soda and coconut coir to enhance the decomposition process than the conventional method.

Sodium bicarbonate (known as baking soda) available in all house kitchen is a chemical compound with the formula $NaHCO_3$. It is a salt composed of a sodium caption (Na^+) and a bicarbonate anion (HCO_3^-). This has been flaunted as an active and safe fungicide to reduce the effects of fungal diseases on common ornamental and vegetable plants and causes no harm.

Coir or coconut coir is a natural fiber extracted from the husk of coconut is used as a soil conditioner. It contains several macro and microplant nutrients, including substantial quantities of potassium. After two days, the waste is collected from the tray and used as organic fertilizer in the garden.



a. Heating Section

b. Grinding Section

c. Decompost Section

Fig. 1 Overall block diagram of kitchen waste recycler unit

3 Interfacing of Sensors, LCD Display, Heating Coil and Motor with Relay

The Arduino UNO and the necessary interfacing of the kitchen waste recycler bin are shown in Fig. 2. The Arduino processor is interfaced with temperature and humidity sensor (combined sensor DHT11), LCD display, relay with electric motor and cooling fan, relay with heating coil is further explained in this section.

The interfacing of LCD with Arduino is shown in Fig. 3. The 16 pins JHD162A LCD module is configured in 4-bit mode with 5 V supply. The LCD is used to display the current temperature, humidity and motor condition (ON/OFF) inside the bin and controlled via the Arduino.

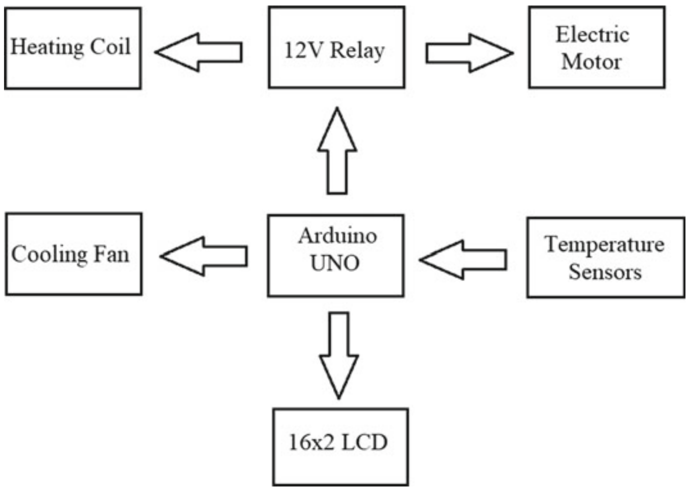


Fig. 2 Controller interfacing block diagram

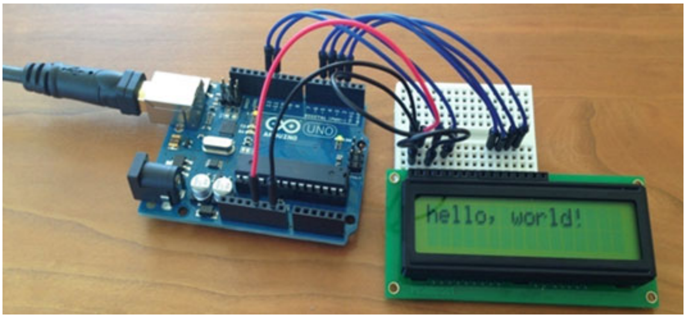


Fig. 3 Interfacing LCD with Arduino

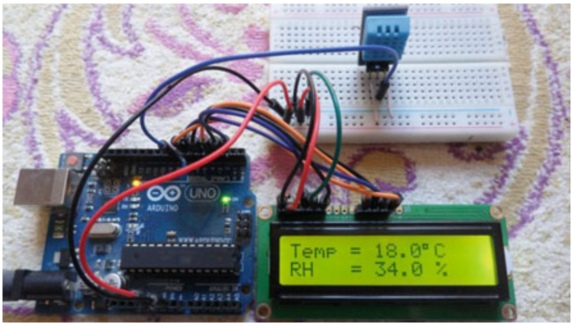


Fig. 4 Interfacing LM35 and DHT11 with Arduino

The LM35 is a precision temperature measuring devices with a linear output voltage is used to sense the temperature. The DHT11 temperature and humidity sensor [5] give digital signal output. The LM35 and DHT11 need no external calibration. The interfacing of LM35 and DHT11 is shown in Fig. 4.

The heating coil is interfaced with Arduino and displays the temperature and humidity which is measured by the LM35 and DHT11 sensor (both temperature and humidity sensor). The ON/OFF of the heating coil is controlled by 12 V relay is also interfaced with Arduino which is shown in Fig. 5. The display shows the heater condition is ON for the time of 2 s and the temperature 25.5 °C and humidity of the bin is around 32.2 °C is displayed in the LCD. The controller is programmed to turn OFF the heater when the temperature exceeds 80 °C.

The motor is used in the grinding section to reduce the size of the kitchen wastes in order to increase the decomposition rate. The interfacing of motor with relay showing the motor status ON for 3 s is shown in the Fig. 6. The time for grinding is fixed based on trial and error of decomposing rate verified with different levels of grounded wastes.

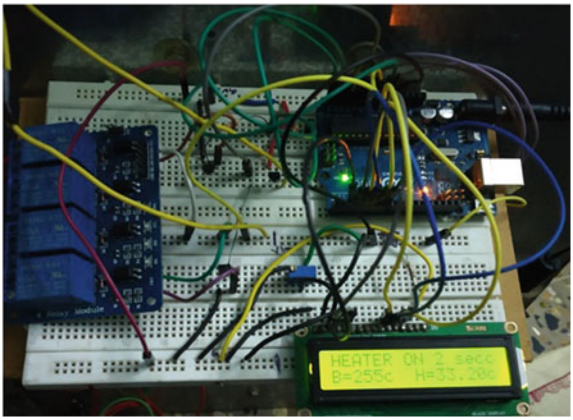


Fig. 5 Interfacing heating coil with Arduino

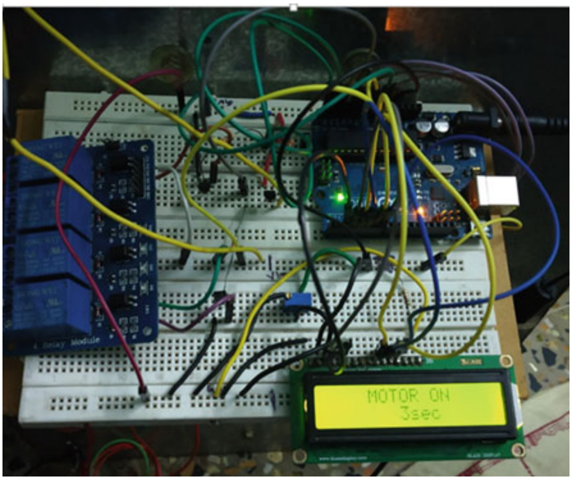


Fig. 6 Interfacing motor with Arduino for grinding section

4 Hardware Implementation

The kitchen waste is dumped inside the compost bin. At first, heating process is carried out to remove moisture from the waste. The heating process is cycled for 20 s over a period of time. The heating section is shown in the Fig. 7. In the figure, the heating coil is placed in one side of the bin. In order to prevent the electrical leakages, the coil is clamped over the hylam sheet. The hylam sheet has good resistance toward electricity. The compost bin is placed in front of the heating coil. The compost bin is placed in the center of the setup closer to the heating coil. The metal clamps are used to hold the bin.

The moisture from the waste is removed, and then, the grinding process is initiated. The waste is grinded for 3 s in a cyclic manner. After a period of time, the waste

Fig. 7 Heating section of the kitchen waste recycler bin

along with additives such as coconut coir and baking soda is shredded to smaller pieces and allowed to decompose. The grinding and decomposing section are shown in Fig. 8. In the figure shown below, the electric motor is placed at the bottom of the setup. The metal clamps are used to hold the electric motor. The electric motor has manually created thread in the shaft. The thread is used to connect the blade with the shaft of the motor. The compost bin is placed above the electric motor. The wooden stand is used to hold the compost bin above the electric motor. And also, metal clampers are used to compost bin moving, while motor is turned on.

The dimension of the setup is (40 × 40 × 70) cm. The top view of the kitchen waste recycler is shown in the Fig. 9. In the figure shown below, the heating coil is placed on the left side of the setup. The coil is placed on hylam sheet. The hylam sheet is used to prevent electrical discharge from coil to the setup. The electric motor

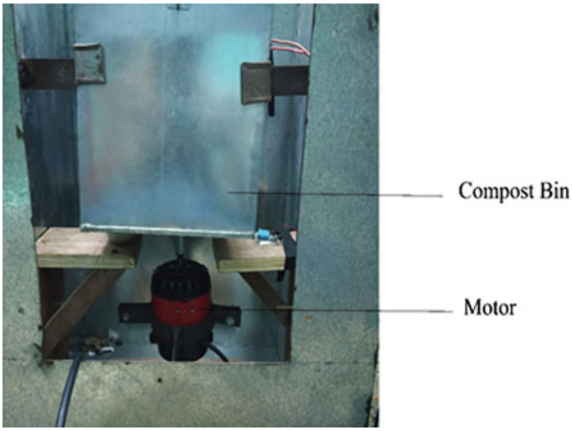


Fig. 8 Grinding and decomposing section of the kitchen waste recycler bin

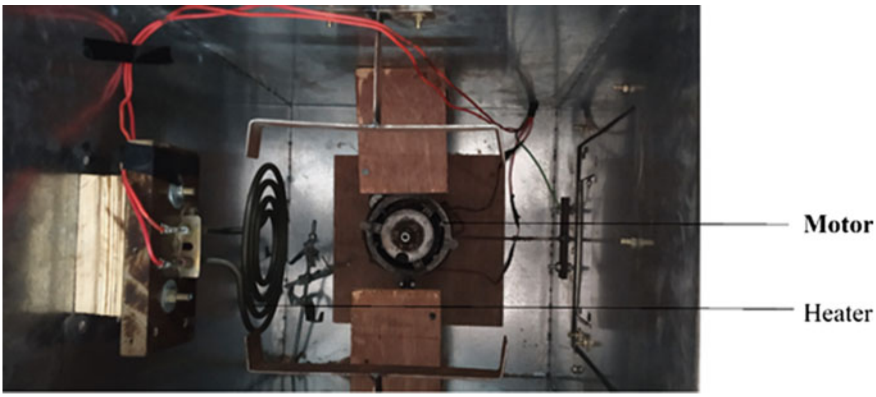


Fig. 9 Top view of the kitchen waste recycler bin

is placed at bottom of the setup. The motor is fixed in the bottom of the setup using the clamps. The clamps are used to hold the electric motor, while it is turned on. The wooden stands and metal clamps in the figure are used to hold the compost bin. The door is available in the right side of the setup, which is used for removing blade from the shaft of the motor and also for maintenance purpose.

The controller unit of the kitchen waste recycler bin is shown in Fig. 10. The controller circuit includes 16 × 2 LCD, relay module and Arduino UNO. The 16 × 2 LCD display is used to show the ongoing process at that instance. The relay module acts as a switch, and it toggles between heating coil and the electric motor periodically. The Arduino UNO is the heart of the controlling part. It connects various components used in the setup. It gets the data from the temperature and humidity sensors and displays it.

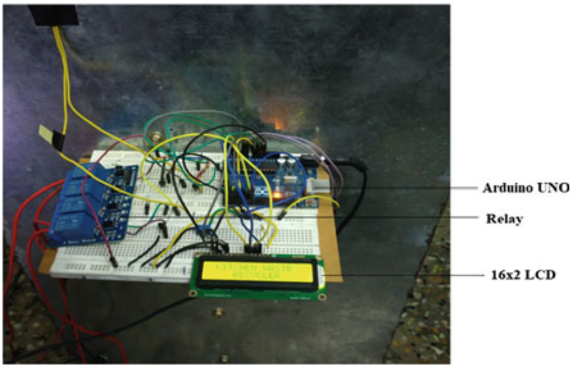


Fig. 10 Controller unit fixed in the side of the kitchen waste recycler bin

Fig. 11 Compact kitchen waste recycler bin

The complete model of compact kitchen waste recycler is shown in Fig. 11 which is very compact and can be placed inside the kitchen itself. This does not produce any bad odor as decomposing takes place within 36 h.

The kitchen waste was collected, and it is shown in the Fig. 12. The waste is dumped into the compost bin. The wasted was heated, grinded and decomposed in the kitchen waste recycler bin. After 36 h, the kitchen waste was turned into nutrient-rich organic manure. The organic output manure is shown in the Fig. 13.

Fig. 12 Sample kitchen waste as input

Fig. 13 Organic manure within 36 h

5 Conclusion

The complete design and implementation of the economic kitchen waste recycler bin are presented which are very compact in size. Sine very fast decomposing rate, there is no bad odor so can be placed inside the kitchen itself. By using the kitchen waste recycler, the biodegradable wastes are separated in the house itself. The dumping of waste in the form of landfill is reduced. In addition to this, we can get low-cost nutrient-rich organic manure is obtained from the households. The organic manure is obtained within 36 h when compared to the microbial decomposition.

References

1. Chowdhury B, Chowdhury MU (2007) RFID-based real-time smart waste management system. IEEE Australasian telecommunication networks and applications conference
2. Aazam KM, St-Hilaire M, Lung C-H, Lambadaris I (2016) Cloud-based smart waste management for smart cities. IEEE international workshop on computer aided modelling and design of communication links and networks
3. Kumar NS, Vijayalakshmi B, Prarthana RJ, Shankar A (2016) IoT based smart garbage alert system using Arduino UNO. IEEE Region 10 Conference (TENCON)
4. Folianto F, Low YS, Yeow WL (2015) Smartbin: smart waste management system. In: IEEE international conference on intelligent sensors, sensor networks and information processing
5. Sindhanaiselvi D, Shanmuganantham T (2018) Performance analysis of embossed diaphragm based MEMS piezo resistive pressure sensor for flood level measurement. J Mater Today 5(10):2214–7853 (Elsevier)
6. Sindhanaiselvi D, Shanmuganantham T (2018) Design and implementation of indoor tracking system using inertial sensor. J Lect Notes Electr Eng 521:465–474 (Springer)
7. Sindhanaiselvi D, Shanmuganantham T (2017) Double boss sculptured diaphragm employed piezoresistive MEMS pressure sensor with silicon-on-insulator (SOI). J Eng Sci Technol 12(7):1740–1754
8. Nallathambi A, Shanmuganantham T, Sindhanaiselvi D (2018) Design and analysis of MEMS based piezoresistive pressure sensor for sensitivity enhancement. J Mater Today Proc 5:1897–1903
9. Sindhanaiselvi D, Shanmuganantham T (2014) Performance analysis of sculptured diaphragm for low pressure MEMS sensors. Appl Mech Mater 592–594:2193–2198
10. Wahab MHA, Kadir AA, Tomari MR, Jabbar MH (2014) Smart recycle bin a conceptual approach of smart waste management with integrated web based system. In: IEEE international conference on IT convergence and security (ICITCS)

Image Denoising Using DnCNN: An Exploration Study

Vineeth Murali and P. V. Sudeep

Abstract Image denoising is a crucial pre-processing step on images to restore the original image by suppressing the associated noise. This paper extends the performance study of the denoising convolutional neural network (DnCNN) architecture on images having the Gaussian noise. The DnCNN is an efficient deep learning model to estimate a residual image from the input image with the Gaussian noise. The underlying noise-free image can be estimated as the difference between the noisy image and the residue image. In this paper, we analyse the performance of DnCNN with data augmentation, batch normalisation and dropout. The experiments are conducted on the Berkeley natural image dataset, and quantitative and qualitative study has been performed. The comparison of the experimental results demonstrates that the DnCNN model converges at a faster rate and works well with a smaller dataset.

Keywords Convolutional neural networks · Deep learning · Denoising · Gaussian noise · Residual learning

1 Introduction

The major sources of noise in images include environmental condition, sensor temperature, dust particles in scanner and interference in the transmission channel [1]. The presence of noise is unavoidable in digital image acquisition, transmission or due to faulty memory locations in hardware and can be measured in terms of the number of pixels corrupted. Hence, the estimation and removal of noise from images are significant for proper interpretation, analysis and accurate parameter estimation. In order to remove noise and enhance image quality, image denoising algorithms have been extensively studied and presented in the literature [2–7]. It is a common pre-processing step in image application such as image restoration, image segmen-

V. Murali (✉) · P. V. Sudeep
Department of Electronics and Communication Engineering, National Institute of Technology Calicut, Calicut, India
e-mail: vineethmurali95@gmail.com

J. Jayakumari et al. (eds.), *Advances in Communication Systems and Networks*, Lecture Notes in Electrical Engineering 656,
https://doi.org/10.1007/978-981-15-3992-3_72

tation, image classification, image registration and capturing images in poor light conditions [8–10].

The conventional image denoising algorithms attempt to recover the clean image by suppressing the noise and preserving actual image discontinuities [5–7]. Nonetheless, most of the conventional methods produce sub-optimal results due to the reasons like setting up of incorrect values for filter parameters, improper noise model and excessive smoothing of the image, etc. [11]

Recently, deep learning architectures such as convolutional neural network (CNN) [12], autoencoders [13] and generative adversarial network (GAN) [14] are proposed for image denoising. These deep learning models have been proposed to denoise images by using the image features learnt automatically during a training phase and to produce better test results compared to the conventional method. However, deep learning models rely on high-performance computing and the high availability of data. In contrast to conventional methods, deep denoising models perform better, and the performance further increases with the amount of data used for training as well as the model complexity increases.

This paper aims to analyse the performance of denoising convolutional neural network (DnCNN) model [15] on the public BSD68 dataset [16]. The output image quality has been measured on the basis of peak signal-to-noise ratio (PSNR) and the mean structural similarity index (SSIM) [15]. In this paper, we present the effect of different regularisation methods such as batch normalisation, data augmentation and dropout with the DnCNN model. Besides, we have studied the impact of epochs on the performance of the DnCNN model [15] in terms of PSNR/SSIM values. In addition, our experiments validate the fast convergence of the DnCNN model.

The remainder of the paper is organised as follows: Sect. 2 describes the noise model and reviews the literature. Section 3 gives an insight into the methodology used. Section 4 analyses the results and discusses the various results obtained, and Sect. 5 includes the conclusion.

2 Background

2.1 Noise Model

In this paper, we have assumed the popular Gaussian distributed noise model. The noisy image I_N is synthesised by adding the Gaussian distributed random variable N having σ with the ground truth (GT) image I and is given by

$$I_N = I + N \tag{1}$$

All the generated images I_N and corresponding noiseless image I have been used to train the neural network model used in this paper. During testing, we provide a new noisy image I_N as input to the model to produce the recovered image $\hat{I}$.

2.2 Image Denoising Methods

In the past, several methods are proposed for image denoising [17–20]. The conventional image denoising methods can be classified as spatial domain methods and transform domain methods [21].

In spatial domain approaches, the filtering operations are performed directly on the pixels of a noisy image. Nonetheless, the spatial filters are not suitable for different kinds of noise. So, the noise removal by applying the spatial filter can reduce the visibility of fine structures in the image and can blur the edges present in it [15].

Compared with spatial domain filtering methods, transform domain filtering methods convert the input noisy image to another domain before applying a denoising procedure. The transform domain filtering methods can be data adaptive or non-data adaptive. In the past, various transform domain filters have been proposed for image denoising [6, 22, 23].

In 2005, Baudes et al. proposed a spatial domain denoising method called non-local means algorithm based on non-local averaging of all pixels present in the image [5]. The NLM attempts to preserve the image features while removing the noise in the image. Pixels in the local neighbourhoods are averaged and hence reduce the redundancy that may be present in the image due to similar features and patterns. As a competent and powerful extension of the NLM approach, block-matching and 3D filtering (BM3D) has been introduced in the literature. The BM3D filter first converts the 2D image data into 3D data arrays and then uses the sparse method to remove noise from the obtained 3D data arrays [6].

In the recent past, deep neural networks have shown great promise in image denoising applications. In 2009, Jain and Seung proposed a convolutional neural network (CNN)-based image denoising approach for enhancing natural images [2]. In 2012, the multilayer perceptron (MLP) model was proposed by Harold et al.. When the method in [3] trained on a large dataset, it was able to outperform different state-of-the-art methods. The image denoising framework by Junyuan et al. used sparse coding and autoencoders to achieve better results than the traditional linear sparse coding algorithms [4]. Trainable nonlinear reaction diffusion (TNRD) model introduced by Chen et al. is a feed-forward network by unfolding a fixed number of gradient descent inference steps [24].

In 2014, Karen Simonyan et al. proposed the VGGNet and investigated the effect of increasing the depth of deep neural networks in large-scale image recognition [25]. They found out that significant improvements can be obtained on the prior art network configurations when the depth of the weight layers was increased to 16–19. However, an increase in the network depth causes the exploding gradient or vanishing gradient problems. In ResNet, the concept of residual network is introduced to address this problem effectively [26]. The deep learning method for denoising using residual learning was able to produce a better denoising output. Recently, generative adversarial networks (GAN) have been proposed by Abeer et al. for image denoising [27]. The GAN model was successful at generating denoising images even when it was trained with the fewer number of images in the dataset.

3 Methodology

3.1 DnCNN Model for Image Denoising

A convolutional neural network (CNN or ConvNet) is the most popular deep learning architecture commonly applied to analyse visual imagery. CNN is a combination of an input layer, multiple hidden layers and an output layer. Hidden layers perform the feature extraction by the process of the convolution operation. The popular CNN architectures are AlexNet [12], VGGNet [25], LeNet [28], GoogleNet [29] ResNet [30], etc. Visual Geometry Group (VGGNet) model was introduced in the ImageNet LSVRC-2014 [31]. There are VGG architectures from eight convolutional layers to 16 convolutional layers along with $three$ fully connected layers resulting in VGG-11, VGG-13, VGG-16, VGG-19 models. Among them, VGG-16 model outperforms all other configurations on the classification task in terms of top-1 error (%) and top-5 error (%) (Fig. 1).

The DnCNN model was proposed by Zhang et al. [15]. The VGG-16 model is modified in order to suit the image denoising application in the DnCNN model [25]. The loss function used is the mean squared error (MSE) between desired residual images and estimated ones from noisy input, i.e. difference between the residue images is taken and loss value is computed and is represented as

$$l(\Theta) = \frac{1}{2N} \sum_{i=1}^{N} \parallel R(y_i; \Theta) - (y_i - x_i) \parallel_F^2 \tag{2}$$

where the parameter $l(\Theta)$ represents the loss function value calculated using the noisy clean training image pairs (x, y) and $R(y_i; \Theta)$ is the residual image. To obtain the denoised image $\hat{I}$, the DnCNN [15] model first finds a residual image $R(y_i; \Theta)$, which is the estimate of the noise map N. Then, the estimate of noise-free image $\hat{I}$ is computed as the difference of $R(y_i; \Theta)$ from the noisy input I_N.

The integration of residual learning and batch normalisation is found to improve the training accuracy of the DnCNN model [15]. They are described below:

(1) *Residual Learning*: ResNet helps in achieving better performance even when we increase the network layer count to hundreds or thousands. As the network

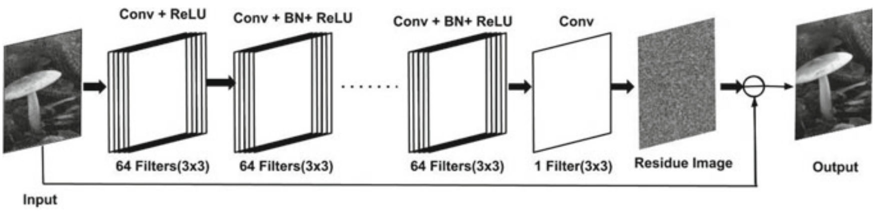
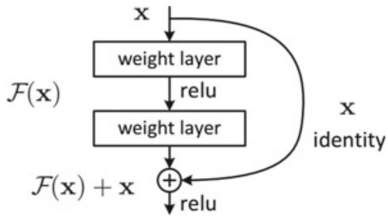


Fig. 1 DnCNN model [15]

Fig. 2 Residual block [15]



depth increases, the network performance drops due to the vanishing gradient problem. The solution to the above problem is the residual networks. The main concept of ResNet is to skip one or more layers using the so-called identity shortcut connection as shown in Fig. 2.

Assuming F as the ReLu activation function, z as the sum of weights and biases, w as the weights, b as the bias, l as the present layer number and a as the output of a particular layer.

$$\begin{aligned} a(l+2) &= F[z(l+2) + a(l)] \\ &= F[w(l+2) \times a(l+1) + b(l+2) + a(l)] \end{aligned}$$

If $w(l+2) = 0$ and $b(l+2) = 0$,

$$\begin{aligned} \text{then, } a(l+2) &= F[0 + a(l)] \\ &= a(l) \end{aligned}$$

Hence, because of this skip connection, it is easy to get

$$a(l+2) = a(l) \tag{3}$$

and this means that adding two layers in the neural network does not hurt the network ability to do things without these layers.

(2) *Batch Normalisation*: The input layers are normalised when we have features in different range so that it must be normalised to speed up learning. Applying the same idea to hidden layers is the basic idea of batch normalisation. Higher learning rates can be used in the network because batch normalisation ensures that activation does not go really high or really low. It decreases the effect of overfitting because it has a slight regularisation effects.

This paper compares the denoising performance of DnCNN model in the presence and absence of the techniques like batch normalisation and dropout. Dropout is the process where random neurons are neglected in the training phase. Theoretically, when the state of the neural network is changed from training to testing, the dropout would shift the variance of the neural unit [32]. Nonetheless, batch normalisation would maintain its statistical variance that is obtained from the entire learning procedure, in the test phase. This variance shift causes the drop in performance of the neural network when dropout is used in combination with batch normalisation [32].

4 Results and Discussion

Google Colab was used to run the program. Google Colab is a cloud service, and it supports GPU and CPU and its free. GPU used in Google Colab: 1xTesla K80, having 2496 CUDA cores, computes 3.7, 12 GB (11.439 GB Usable) GDDR5 VRAM. CPU used in Google Colab: one single core hyper threaded, i.e. (one core, two threads) Xeon Processors @2.3 GHz (No Turbo Boost), 45 MB Cache. For every 12 h or so Disc, RAM, VRAM, CPU cache, etc., data that is on the allocated virtual machine will get erased.

Number of images used for training was 400 of size 180×180. Training and validation datasets have been split in the ratio 80:20. For testing the model, Berkeley segmentation dataset (BSD68) which consisted of 68 natural images was used [16]. Data augmentation was done to increase the number of training images. Data augmentation on the training dataset helped in creating a 238,336 patches of size 40×40. Among them, only 2000 images would be used in 1 epoch, i.e. for training 1600 images were used and for validation 400 images were used. All the neural network models were run for 50 epochs. DnCNN model with $\sigma = 20$ alone was trained for 200 epochs to analyse the model performance with increase in epochs. Epoch is the number of times the whole image database is fed to the training model. Batch size is taken as 128. A total of 128 images will be processed at a time. The loss function used was mean squared error. The optimiser used was $Adam$. The learning rate was 0.001 for epoch 1–30 and 0.0001 for epoch 30–50.

4.1 Qualitative Analysis

In this section, we describe the experimental results on BSD68 dataset. In Fig. 3, we present the denoising results on a natural image of a fish of size 321×481 having Gaussian noise with $\sigma = 20$. For visual comparison, a red rectangular box is used to present a small middle region. It can be observed from the denoised images that most of the noise has been reduced and fine structures have been preserved without blurring the edges. The best image denoising result was obtained using the DnCNN model with kernel size $k = 3$, batch normalisation and without dropout.

Further, we applied the denoising algorithm on an image corrupted by Gaussian noise with $\sigma = 40$ and displayed the obtained results in Fig. 4. Figure 4d shows the residual image, which is the output obtained after processing along the hidden layers. The residual image is the difference between the denoised and noisy image. Residual image is measured in the range [0, 35].

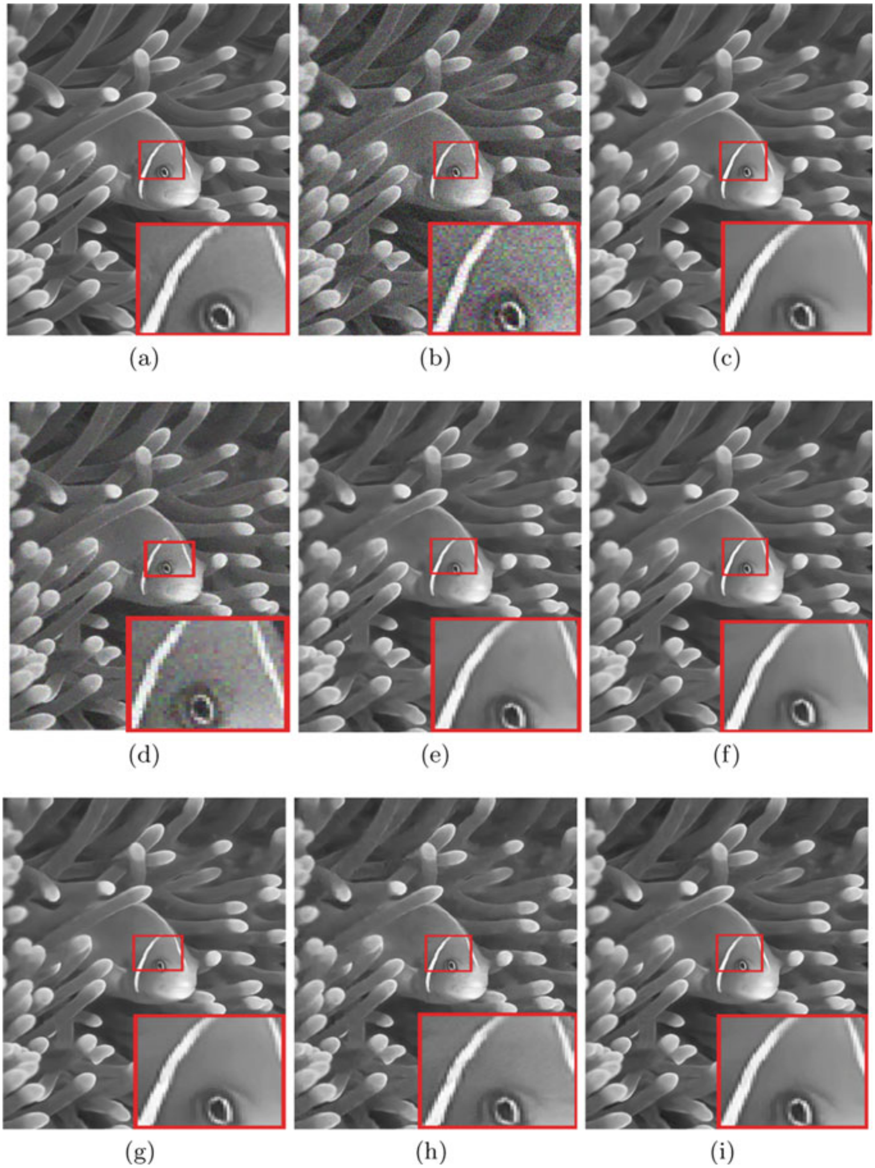


Fig. 3 Denoising results for noise level $\sigma = 20$, **a** GT image; **b** noisy image (PSNR = 23.28 dB); **c** DnCNN denoised image (PSNR = 29.174 dB); **d** NLM denoised image (PSNR = 27.43 dB); **e** BM3D denoised image (PSNR = 28.11 dB); **f** DnCNN without batch normalisation denoised image (PSNR = 28.98 dB); **g** DnCNN with $k = 5$ denoised image (PSNR = 28.61 dB); **h** DnCNN with $k = 7$ denoised image (PSNR = 27.9 dB); **i** DnCNN with kernel size as a combination of 3, 5 and 7 denoised image (PSNR = 28.23 dB)

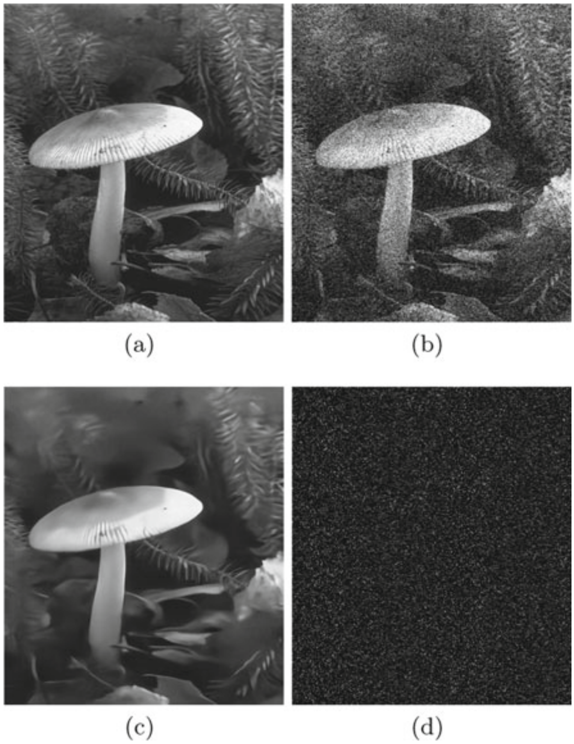


Fig. 4 Denoising results for noise level $\sigma = 40$, **a** GT image; **b** noisy image (PSNR = 16.13 dB); **c** denoised image using DnCNN (PSNR = 26.43 dB); **d** residual image

4.2 Quantitative Analysis

The analysis of quantitative results is based on peak signal-to-noise ratio (PSNR) and mean structural similarity index matrix (SSIM) [33].

$$\text{PSNR} = 10 \times \log_{10} \frac{\text{MAX}_I^2}{\text{MSE}} \tag{4}$$

where MAX_I is the maximum possible pixel value of the image, and MSE is the mean squared error.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{5}$$

where μ_y is the average of y, μ_x is the average of x, σ_y^2 is the variance of y, σ_x^2 is the variance of x, σ_{xy} is the covariance of x and y. c_1 and c_2 are variables to stabilise equation.

Tables 1 and 2 depict the performance of different DnCNN architectures based on the PSNR and mean SSIM measure, respectively. The DnCNN model with $k = 3$,

Table 1 PSNR comparison of different architectures in DnCNN

Model	Epochs				
	10	20	30	40	50
DnCNN [15] ($\sigma = 10$)	33.545	33.538	33.8446	33.892	**33.884**
DnCNN [15] ($\sigma = 30$)	27.614	27.64	27.65	28.339	**28.381**
DnCNN [15] ($\sigma = 40$)	26.685	26.571	**27.193**	26.985	27.073
$\sigma = 20$					
DnCNN [15] ($\sigma = 20$)	29.984	30.206	30.222	30.282	**30.283**
DnCNN with kernel size 3 with data augmented and without BN	30.0057	30.0776	30.1182	30.1585	**30.1683**
DnCNN with kernel size 5 with data augmented	29.2146	29.6592	29.259	29.7351	**29.7372**
DnCNN with kernel size 7 with data augmented	28.1387	29.9942	29.7058	**30.1947**	30.1866
DnCNN with kernel size 3 and dropout 0.4 and data augmentation	22.3462	29.0039	29.1846	29.8679	**29.8871**
DnCNN with kernel size 357 and data augmentation and dropout = 0.4	22.3541	22.353	28.9281	29.0769	**29.0982**

batch normalisation and without dropout has the best performance for the given $\sigma = 20$. From Tables 1 and 2, it can be inferred that for lower noise level (say $\sigma = 10$), the DnCNN network can perform better and give higher PSNR and SSIM values in the range of [32, 34] and [0.94, 0.96], respectively. For $\sigma = 30$, the denoised image has PSNR and SSIM values in the range [27, 29] and [0.86, 0.88] for the denoised images. For $\sigma = 40$, the denoised image has PSNR and SSIM values in the range [26, 27] and [0.83, 0.85] for the denoised images. As a general case, network models with different architectures of DnCNN with $\sigma = 20$ were applied, and performance

Table 2 SSIM comparison of different architectures in DnCNN

Model	Epochs				
	10	20	30	40	50
DnCNN [15] ($\sigma = 10$)	0.9593	0.9591	0.962	0.962	**0.9624**
DnCNN [15] ($\sigma = 30$)	0.8735	0.8738	0.877	0.8829	**0.8838**
DnCNN [15] ($\sigma = 40$)	0.8361	0.8315	**0.836**	0.8319	0.8307
$\sigma = 20$					
DnCNN [15] ($\sigma = 20$)	0.9158	0.9194	0.918	0.9204	**0.9206**
DnCNN with kernel size 3 with data augmented and without BN	0.9159	0.9174	0.9178	0.9184	**0.9189**
DnCNN with kernel size 5 with data augmented	0.9075	0.9149	0.9126	0.9143	**0.9146**
DnCNN with kernel size 7 with data augmented	0.9031	0.9168	0.9127	**0.9208**	0.9205
DnCNN with kernel size 3 and dropout 0.4 and data augmentation	0.6115	0.8905	0.8956	0.8996	**0.8998**
DnCNN with kernel size 357 and data augmentation and dropout = 0.4	0.612	0.6119	0.8913	0.8942	**0.8956**

of the network was noted. All the DnCNN models performed well for the performance measures PSNR and SSIM values when compared to the conventional methods and had the PSNR value range [29, 31] and SSIM value in the range [0.88, 0.93].

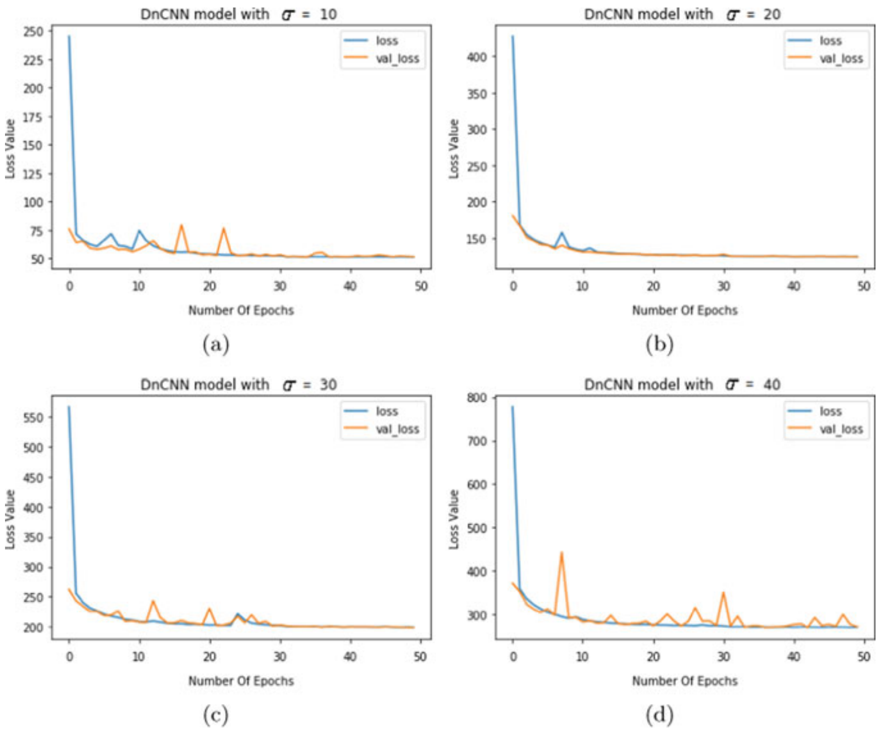


Fig. 5 Training loss and validation loss graph for different σ values

4.3 Discussion

(1) *Loss versus Epoch curve*: From the loss versus epoch graphs of training the DnCNN model, it is seen that both training and validation loss curves follow the same path. During the initial epochs, the training loss is higher, and it converges to a lower loss value within 5 epochs. The faster converging nature is attributed to residual learning and training the model with 1600 patches in a single epoch. As the validation loss curve strictly follows the training loss, underfitting and overfitting are avoided. The loss curves almost remain constant after 30 epochs. The reason why we stopped at 50 epochs is that further training of model by increasing the number of epochs from 50 to 100 resulted only in a 0.03% change in the PSNR value and 0.3% change in SSIM value. So, the optimum number of epochs chosen was 50. The improvement in the quality measures was negligible compared to the computation time required for achieving a higher number of epochs (Fig. 5).

(2) *Impact of kernel size*: DnCNN model has 558,016 parameters for kernel size $k = 3$, 1,541,184 parameters for $k = 5$ and 3,040,736 parameters for $k = 7$. Increasing the kernel size resulted in increase in the number of parameters.

From Tables 1 and 2, it is observed that DnCNN model with $k = 3$ has higher performance compared to other models. DnCNN model with $k = 3$ extracts and learns local features effectively. As the kernel size increases, the model learns generic features resulting in large receptive field and poor weight sharing.

(3) *Dropout*: When dropout and batch normalisation are introduced in the model separately, the model performance increases, but combining both the model performance becomes poor. This is attributed to the shift in variance due to dropout when the state of the neural network is changed from training to testing.

5 Conclusion

In this paper, the performance analysis of a deep convolutional neural network for image denoising on Gaussian noise has been presented. The noisy image is generated by assuming Gaussian noise distribution, and the DnCNN model [15] is used to estimate the denoised image. Residual learning and batch normalisation are the two main features of the DnCNN model. Variants of the DnCNN model [15] were analysed using the BSD68 dataset. The denoising performance is compared using measures such as PSNR and mean SSIM at different noise levels. The experimental results demonstrate that the DnCNN model has the fast converging property and the efficiency to work with small datasets.

References

1. Pandey A, Singh KK (2015) Analysis of noise models in digital image processing. Int J Sci Technol Manag 4(01):140–144
2. Jain V Seung S (2009) Natural image denoising with convolutional networks. In: Advances in neural information processing systems, pp 769–776
3. Burger HC, Schuler CJ, Harmeling S (2012) Image denoising: can plain neural networks compete with BM3D? In: 2012 IEEE conference on computer vision and pattern recognition. IEEE, pp 2392–2399
4. Xie J, Xu L, Chen E (2012) Image denoising and inpainting with deep neural networks. In: Advances in neural information processing systems, pp 341–349
5. Buades A, Coll B, Morel J-M (2005) A non-local algorithm for image denoising. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 2. IEEE, pp 60–65
6. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans Image Process 16(8):2080–2095
7. Sudeep PV, Palanisamy P, Kesavadas C, Rajan J (2015) Nonlocal linear minimum mean square error methods for denoising MRI. Biomed Signal Process Control 20:125–134
8. Guo T, Dong J, Li H, Gao Y (2017). Simple convolutional neural network on image classification. In: 2017 IEEE 2nd international conference on big data analysis (ICBDA). IEEE, pp 721–724
9. Guizar-Sicairos M, Thurman ST, Fienup JR (2008) Efficient subpixel image registration algorithms. Opt Lett 33(2):156–158

10. Haralick RM, Shapiro LG (1985) Image segmentation techniques. Comput Vis Graph Image Process 29(1):100–132
11. Dong W, Zhang L, Shi G, Li X (2012) Nonlocally centralized sparse representation for image restoration. IEEE Trans Image Process 22(4):1620–1630
12. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
13. Baldi P (2012) Autoencoders, unsupervised learning, and deep architectures. In: Proceedings of ICML workshop on unsupervised and transfer learning, pp 37–49
14. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680
15. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process 26(7):3142–3155
16. Roth S, Black MJ (2009) Fields of experts. Int J Comput Vis 82(2):205
17. Norbert W (1950) Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications. MIT Press, Cambridge
18. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: ICCV, vol 98, p 2
19. Yang G-Z, Burger P, Firmin DN, Underwood SR (1996) Structure adaptive anisotropic image filtering. Image Vis Comput 14(2):135–145
20. Yang R, Yin L, Gabbouj M, Astola J, Neuvo Y (1995) Optimal weighted median filtering under structural constraints. IEEE Trans Signal Process 43(3):591–604
21. Aravind BN, Suresh KV (2017) Hybrid image denoising. In: 2017 international conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT). IEEE, pp 46–49
22. Oja E, Hyvarinen A (2000) Independent component analysis: algorithms and applications. Neural Netw 13(4-5):411–430
23. Kaur L, Gupta S, Chauhan RC (2002) Image denoising using wavelet thresholding. In: ICVGIP, vol 2, pp 16–18
24. Chen Y, Pock T (2016) Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. IEEE Trans Pattern Anal Mach Intell 39(6):1256–1272
25. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
26. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence
27. Alsaiari A, Rustagi R, Thomas MM, Forbes AG, et al (2019) Image denoising using a generative adversarial network. In: 2019 IEEE 2nd international conference on information and computer technologies (ICICT). IEEE, pp 126–132
28. LeCun Y, Bottou L, Bengio Y, Haffner P et al (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
29. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
30. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
31. Russakovsky O, Deng J, Hao S, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis (IJCV) 115(3):211–252
32. Li X, Chen S, Hu X, Yang J (2019) Understanding the disharmony between dropout and batch normalization by variance shift. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2682–2690
33. Hore A, Ziou (2010) Image quality metrics: PSNR vs. SSIM. In: 2010 20th international conference on pattern recognition. IEEE, pp 2366–2369