Lecture Notes in Electrical Engineering 248

V. Sridhar
Holalu Seenappa Sheshadri
M. C. Padma
Editors

Emerging Research in Electronics, Computer Science and Technology

Proceedings of International Conference, ICERECT 2012

Springer

Lecture Notes in Electrical Engineering

Volume 248

For further volumes:
http://www.springer.com/series/7818

V. Sridhar · Holalu Seenappa Sheshadri
M. C. Padma
Editors

Emerging Research in Electronics, Computer Science and Technology

Proceedings of International Conference, ICERECT 2012

Springer

Editors
V. Sridhar
Holalu Seenappa Sheshadri
M. C. Padma
Department of Electronics and Communication Engineering
PES College of Engineering
Mandya, Karnataka
India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-81-322-1156-3 ISBN 978-81-322-1157-0 (eBook)
DOI 10.1007/978-81-322-1157-0
Springer New Delhi Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013945802

Printed on acid-free paper

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

Consolidated Report on ICERECT-12

On December 21st, the inauguration function was presided by Dr. H. D. Chowdaiah, Chairman, PET and the keynote address was delivered by Dr. Rangaraj M. Rangayyan of University of Calgary. A book on collection of abstracts of all the selected papers for the conference was also released on this occasion. The chief guest Prof. D. K. Subramanian, TEQIP Mentor, PESCE advised the research scholars to take up innovation experiments and perform meaningful research, which will be useful for the society.

Dr. Rangayyan highlighted the importance of research and gave a thought provoking lecture on Retinopathy. This lecture was followed by an invited talk by Dr. K. N. Bhanuprakash, Scientist ASTAR, Singapore, on Electronic Brain Atlas. Later in the afternoon, the research scholars presented several technical papers.

On December 22nd, Dr. Rajkumar Vedam, Invensys, USA, delivered a talk on Search Directions for Intelligence that exposed the various areas to take up the research. The technical paper presentations were held as per schedule. Nearly 100 papers were presented by the research scholars.

The Valedictory function was held on December 22nd in which Sri. Sharan C. Padashetty, Special Officer, TEQIP, DTE, Karnataka highlighted the financial assistance to various research activities. Principal Dr. V. Sridhar delivered presidential address and requested the delegates to come up with good research and participate actively in research activities. He has also informed that the selected papers will be published in Springer (LNEE) at an early date as a special issue.

On behalf of the executive committee, we render sincere thanks to all the delegates, session chairs, management, participants, reviewers, volunteers, Springer, volunteers, and others who have given us continuous support and encouragement for the success of this conference.

Principal and Organizing Committee
ICERECT-12


P E S College of Engineering Mandya
International Conference On
Emerging Research in Electronics, Computer Science
and Technology (ICERECT – 12) (Under TEQIP)
21st and 22nd December 2012


P E S College of Engineering Mandya
21st and 22nd December 2012

Preface

The first *International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT-12)* is another milestone in the 50 years of existence of PES College of Engineering (PESCE), Mandya. This institution started by Ex. Education Minister, Late K. V. Shankara Gowda, during 1962, is a pioneer institute which is imparting technical education to local people. Presently, the Chairman of PET (R) Dr. H. D. Chowdaiah is standing as a guiding force to bring this college to a higher level of excellence.

The Departments of Electronics and Communication, Computer Science and Engineering, Electrical and Electronics, and Information Science and Engineering, with their rich experience is conducting an international conference on emerging trends to trigger the minds of researchers, faculty members, and also students.

The motivation for this international conference is due to the fact that the present trends in Computers and Communications have received attention recently and noticeable progress towards useful applications to mankind is reported. The conference aims to provide an excellent international forum for sharing knowledge and results in theory, methodology and applications of Electronics, Computer Science Engineering, and Applications. The conference looks for significant contributions to all major fields of the Electronics, Computer Science and Information Technology in theoretical and practical aspects. The aim of the conference is to provide a platform to the researchers and practitioners from both academic as well as industry to meet and share cutting-edge development in the field.

A number of researchers in both academic institutions and industries are contributing to new methods, algorithms, analysis techniques to solve various problems in the above field.

The papers received from various research scholars and students in and around the globe were of good quality and the deliberations are expected to be more useful to the research scholars and also students. Organizers have taken more care to

select the quality papers so as to encourage the researchers to get confidence in their activities. Thanks are due to the editorial board of Springer LNEE for having accepted our conference proceedings to be published in their prestigious series in time.

V. Sridhar
H. S. Sheshadri
M. C. Padma

Acknowledgments

This special issue is a result of collective endeavors from several research scholars and scientists from Electronics and Communication Engineering, Computer Science and Technology from various parts of India and abroad. The authors are indebted to all their efforts and outstanding technical contributions. We would like to express our appreciation to Springer Publishers for helping us to create this special issue on the **Proceedings of the International Conference on Emerging Research in Electronics, Computer Science and Technology** (ICERECT- 12).

We would like to thank Special Officer, TEQIP, Department of Technical Education, and Government of Karnataka for extending their financial support to this conference. We would also like to thank TEQIP Mentor and Coordinator, President and Trustees of PET® for their constant support and encouragement for organizing this conference.

Thanks are also due to the heads of various departments, organizing committee members, the reviewers of the research papers, chairpersons and coordinators of various sessions, staff and student volunteers who made this conference a grand success.

Contents

Abstract

Abstract of invited talk by Dr. Bhanuprakash.

Electronic Brain Atlas: application to medical education and clinical practices.

The research and development in the area of Medical Imaging have been carried out for over many years. The experts and researchers not only view the results as a way to understand the structural and functional aspects of the brain but also try to apply them in the academic and clinical—treatment fields. In my talk, I will focus on the journey of development of various Cerefy Brain Atlases and their applications. Brain atlases can be used for both Academic and Clinical applications. Academic applications include teaching, self—studying and research whereas in clinical application it includes diagnosis, surgical planning, treatment, modeling, and visualization. The talk concludes with demonstration of various atlases like Anatomic Atlas, Vascular Atlas, Comprehensive Brain Atlas, and Atlas—assisted Stroke Image interpretation.

Retinal Vessel Segmentation Using Local Entropy Thresholding

A. P. Manjunath, C. S. Rachana and S. Ranjini

Abstract This paper presents an efficient method for automatic detection and extraction of blood vessels in retinal images. The proposed algorithm is composed of four steps: matched filtering, local entropy thresholding, length filtering, and morphological thinning. The purpose of matched filtering is to enhance the blood vessels. Entropy thresholding can well keep the spatial structure of vascular tree segments. Length filtering is used to remove misclassified pixels. The algorithm has been tested on twenty ocular fundus images.

Keywords Blood vessels · Ocular diseases · MFR image · Local entropy thresholding · Vessel extraction

1 Introduction

The automatic detection of blood vessels in the retinal images can help physicians for the purposes of diagnosing ocular diseases, patient screening, and clinical study, etc. Information about blood vessels in retinal images can be used in grading disease severity or as part of the process of automated diagnosis of diseases. Blood vessel appearance can provide information on pathological changes caused by some diseases including diabetes, hypertension, and arteriosclerosis. The most

A. P. Manjunath (✉) · C. S. Rachana · S. Ranjini
Department of Medical Electronics, Dr Ambedkar Institute of Technology, Bangalore, India
e-mail: manjunath_aps@yahoo.com

C. S. Rachana
e-mail: ranjinisethuram@gmail.com

S. Ranjini
e-mail: rachana.csr@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_1, © Springer India 2014

effective treatment for many eye-related diseases is the early detection through regular screenings. Furthermore, a segmentation of the vascular tree seems to be the most appropriate representation for the image registration applications due to three following reasons: (1) it maps the whole retina; (2) it does not move except in few diseases; and (3) it contains enough information for the localization of some anchor points.

There are many previous works on extracting blood vessels. This paper is based on vessel detection algorithm as in [1] but with slight modifications. In edge detection-based method [2], since local gradient maxima occur at the boundary of the vessels, the significant edges along these boundaries are extracted. The grouping process searches a partner for each edge which satisfies certain criteria like opposite gradient direction and spatial proximity. In tracking-based method [3], each vessel segment is defined by three attributes, direction, width, and center point. The density distribution of cross section of a blood vessel can be estimated using Gaussian-shaped function. Individual segments are identified using a search procedure which keeps track of the center of the vessel and makes some decisions about the future path of the vessel based on certain vessel properties. This method requires that beginning and ending search points are manually selected using cursor. An efficient piecewise threshold probing technique was proposed in [4] where the matched-filter response (MFR) image is used for mapping the vascular tree. A set of criteria is tested to determine the threshold of the probe region and ultimately to decide if the area being probed is a blood vessel. Since the MFR image is probed in a spatially adaptive way, different thresholds can be applied throughout the image for mapping blood vessels.

In this paper, we propose a new algorithm to efficiently locate and extract blood vessels in ocular fundus images. The proposed algorithm is composed of four steps, matched filtering, entropy-based thresholding, length filtering, and vascular intersection detection. Compare with the method in [3], our proposed algorithm does not involve human intervention. Since our algorithm can automatically estimate one optimal threshold value, it requires less computational complexity compared with the method in [4].

2 Proposed Algorithm

The proposed algorithm is composed of four steps. Since blood vessels usually have lower reflectance compared with the background, we apply the matched filter to enhance blood vessels with the generation of a MFR image. Secondly, an entropy-based thresholding scheme can be used to distinguish between vessel segments and the background in the MFR image. A length filtering technique is used to remove misclassified pixels. Finally, morphological thinning is applied to get one-pixel-wide blood vessels.

2.1 Matched-filter Approach

This is our proposed method. Chaudhuri et al. [16] note that the gray-level profiles of the cross sections of retinal vessels have an intensity profile which can be approximated by a Gaussian. A two-dimensional matched-filter approach is proposed to detect the vessels. Vessel segments at various orientations are detected by convolving the image with rotated versions of the matched filter kernel and retaining only the maximum response. At an angular resolution of 15°, a total of 12 convolutions are needed. The resulting image can be thresholded to produce a binary segmentation of the vasculature.

In [5], the gray-level profile of the cross section of a blood vessel can be approximated by a Gaussian-shaped curve. The concept of matched-filter detection is used to detect piecewise linear segments of blood vessels in retinal images. Blood vessels usually have poor local contrast. The two-dimensional matched-filter kernel is designed to convolve with the original image in order to enhance the blood vessels. A prototype matched-filter kernel is expressed as

$$f(x,y) = -\exp\left(-\frac{x^2}{2\sigma^2}\right), \quad \text{for} |y| \leq \frac{L}{2} \tag{1}$$

where L is the length of the segment for which the vessel is assumed to have a fixed orientation. Here, the direction of the vessel is assumed to be aligned along the y-axis. Because a vessel may be oriented at any angles, the kernel needs to be rotated for all possible angles. A set of twelve 16 × 15 pixel kernels is applied by convolving to a fundus image and at each pixel only the maximum of their responses is retained. For example, given a retinal image in Fig. 4a which has low contrast between blood vessels and background, its MFR version is shown in Fig. 4c, where we can see blood vessels are significantly enhanced. The algorithm for obtaining MFR image is as given below.

Step-1: Read the gray-scale retinal image.
Step-2: Initialize the length of the vessel using expression (1).
Step-3: Form the first 15 × 16 kernel.
Step-4: Rotate the above kernel eleven times for step angle of 15° to form 12 kernels.
Step-5: Convolve the gray image with each kernel.
Step-6: Obtain the max response for each pixel.
Step-7: Display the MFR image.

2.2 Local Entropy Thresholding

Secondly, the MFR image is processed by a proper thresholding scheme in order to extract the vessel segments from the background. An efficient entropy-based thresholding algorithm, which takes into account the spatial distribution of gray

levels, is used because image pixel intensities are not independent of each other. Specifically, we implement a local entropy thresholding technique, described in [6] which can well preserve the spatial structures in the binarized/threshold image. Two images with identical histograms but different spatial distribution will result in different entropy (also different threshold values).

The co-occurrence matrix of the image F is an P XQ-dimensional matrix $T = \left[t_{ij}\right]_{\mathrm{PXQ}}$ that gives an idea about the transition of intensities between adjacent pixels, indicating spatial structural information of an image. Depending upon the ways in which the gray level i follows gray level j, different definitions of co-occurrence matrix are possible. Here, we made the co-occurrence matrix asymmetric by considering the horizontally right and vertically lower transitions. Thus, t_{ij} is defined as follows:

$$t_{ij} = \sum_{l=1}^{P} \sum_{k=1}^{Q} \delta \tag{2}$$

where $\delta = 1$ if $\begin{cases} f(l,k) = i \text{ and } f(l,k+1) = j \\ \text{or} \\ f(l,k) = i \text{ and } f(l+1,k) = j \end{cases}$

$$\delta = 0 \text{ otherwise}$$

The probability of co-occurrence p_{ij} of gray levels i and j can therefore be written as

$$p_{ij} = \frac{t_{ij}}{\sum_i \sum_j t_{ij}} \tag{3}$$

If s, $0 \leq s \leq L-1$, is a threshold, then s can partition the co-occurrence matrix into 4 quadrants, namely A, B, C, and D (Fig. 1).

Let us define the following quantities:

$$P_A = \sum_{i=0}^{s} \sum_{j=0}^{s} p_{ij} \tag{4}$$

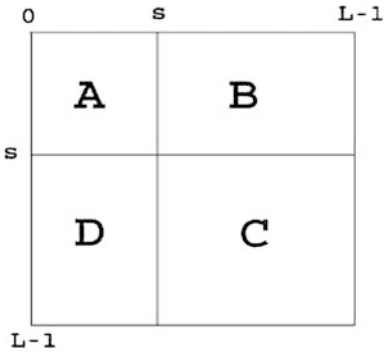


Fig. 1 Quadrants of co-occurrence matrix

$$P_C = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} p_{ij} \tag{5}$$

Normalizing the probabilities within each individual quadrant, such that the sum of the probabilities of each quadrant equals one, we get the following cell probabilities for different quadrants:

$$\begin{aligned} P_{ij}^{A} &= \frac{p_{ij}}{P_A} = \frac{t_{ij} / \sum_{i=0}^{L-1} t_{ij}}{\sum_{i=0}^{s} \sum_{j=0}^{s} t_{ij} / \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \\ &= \frac{t_{ij}}{\sum_{i-0}^{s} \sum_{j=0}^{s} t_{ij}} 0 \le i \le s, \\ &\quad \text{for } 0 \le \mathrm{j} \le \mathrm{s} \end{aligned} \tag{6}$$

Similarly,

$$\begin{aligned} P_{ij}^{C} &= \frac{p_{ij}}{P_C} = \frac{t_{ij}}{\sum_{i-s+1}^{L-1} \sum_{j=s+1}^{L-1} t_{ij}}, \\ &\quad \text{for } \mathrm{s}+1 \le \mathrm{i} \le \mathrm{L}-1,\ \mathrm{s}+1 \le \mathrm{j} \le \mathrm{L}-1 \end{aligned} \tag{7}$$

The second-order entropy of the object can be defined as

$$H_A^{(2)} = -1/2 \sum_{i=0}^{s} \sum_{j=0}^{s} P_{ij}^{A} \log_2 P_{ij}^{A} \tag{8}$$

Similarly,

$$H_C^{(2)} = -1/2 \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} P_{ij}^{C} \log_2 P_{ij}^{C} \tag{9}$$

Hence, the total second-order local entropy of the object and the background can be written as

$$H_T^{(2)}(s) = H_A^{(2)}(s) + H_C^{(2)}(s) \tag{10}$$

The gray level corresponding to the maximum of $H_T^{(2)}(s)$ gives the optimal threshold for object-background classification. For the MFR image shown in Fig. 4c, the entropy-based thresholding result is shown in Fig. 4d where we can see blood vessels are clearly segmented from the background. The algorithm for obtaining entropy threshold image is as given below.

Step-1: Read the MFR image.
Step-2: Compute co-occurrence matrix same as image size using (2).
Step-3: Compute probability of co-occurrence matrix using (3).
Step-4: Divide the matrix into four quadrants using predetermined threshold s.
Step-5: Compute probability of quadrants A and C using (4) and (5).

Step-6: Normalize probability of A and C and compute second-order entropy of A and C using (6) and (7).
Step-7: Obtain optimal threshold value by taking maximum of total entropy (10) for each image pixel.

2.3 Length Filtering

As seen in Fig. 4d, there are still some misclassified pixels in the image. Here, we want to produce a clean and complete vascular tree structure by removing misclassified pixels. Length filtering is used to remove isolated pixels by using the concept of connected pixels labeling. Connected regions correspond to individual objects. We first need to identify separate connected regions. The length filtering tries to isolate the individual objects by using the eight-connected neighborhood and label propagation. Once the algorithm is completed, only the resulting classes exceed a certain number of pixels, for example, 250 are labeled as blood vessels. Figure 4e shows the results after length filtering. An example of connected component labeling is given below. Figure 2 shows a binary image and its corresponding image after component pixel labeling is shown in Fig. 3.

The algorithm for obtaining length-filtered image is as shown below.

Step-1: Read the thresholded image.
Step-2: Label the connected component regions for (1).

Fig. 2 A binary image

1	1	0	1	1	1	0	1
1	1	0	1	0	1	0	1
1	1	1	1	0	0	0	1
0	0	0	0	0	0	0	1
1	1	1	1	0	1	0	1
0	0	0	1	0	1	0	1
1	1	0	1	0	0	0	1
1	1	0	1	0	1	1	1

Fig. 3 Connected component labeling

1	1	0	1	1	1	0	2
1	1	0	1	0	1	0	2
1	1	1	1	0	0	0	2
0	0	0	0	0	0	0	2
3	3	3	3	0	4	0	2
0	0	0	3	0	4	0	2
5	5	0	3	0	0	0	2
5	5	0	3	0	2	2	2

Step-3: Form the length-filtered matrix containing number of pixels in the connected region using (2).
Step-4: If they exceed certain number of pixels, for example 250, then retain it. Discard the others.
Step-5: Obtain the binary image.

2.4 Morphological Thinning

The segmented output Fig. 4e, contains vessels of several pixels wide. Morphological processing is used to reduce it into one-pixel-wide vascular tree. Morphological operations are typically applied to remove imperfections introduced during segmentation. Thinning is a morphological operation that is used to remove selected foreground pixels from binary images, somewhat like erosion or opening. It can be used for several applications, but is particularly useful for skeletonization. In this mode, it is commonly used to tidy up the output of edge detectors by reducing all lines to single pixel thickness. Thinning is only applied to binary images and produces another binary image as output.

The thinning operation is related to the hit-and-miss transform and can be expressed quite simply in terms of it. The thinning of an image A by a structuring element B is

$$\text{thin}(A, B) = A - (\text{hit} - \text{and} - \text{mis}(A, B)) \tag{11}$$

Where the subtraction is a logical subtraction defined by

$$X - Y = X \cap \text{not}\, Y \tag{12}$$

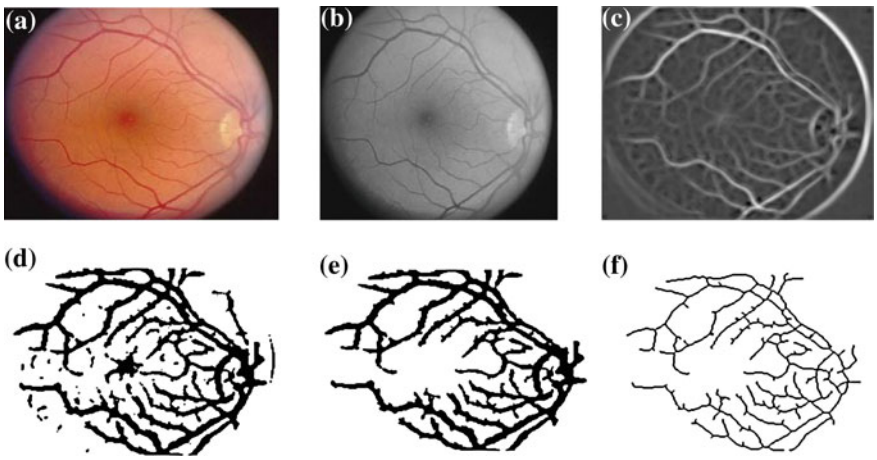


Fig. 4 **a** An original image. **b** Gray-scale image. **c** MFR image. **d** Local entropy thresholding image. **e** Length-filtered image. **f** Morphological thinning image

And the hit-and-miss transform is given by

$$A \otimes B = (A \ominus B_1) \cap (A^c \ominus B_2) \quad (13)$$

The algorithm for morphological thinning is given as follows:

Step-1: Read the length-filtered output.
Step-2: Obtain its hit-and-miss transform using (13).
Step-3: Take the logical subtraction of image using (12).
Step-4: Display the output image.

3 Result and Conclusion

In this paper, we have introduced an efficient algorithm for fully automated blood vessel detection in ocular fundus images using the local entropy thresholding scheme. The proposed method retains the computational simplicity and at the same time can achieve accurate segmentation results in the case of normal retinal images and images with obscure blood vessel appearance. In the case of abnormal retinal images with lesions, some lesions are also mis-detected in addition to blood vessels.

References

1. Chanwimaluang T, Fan G The efficient blood vessel detection algorithm using local entropy thresholding. School of Electrical and Computer Engineering Oklahoma State University
2. Pinz A, Bernogger S, Datlinger P, Kruger A (1998) Mapping the human retina. IEEE Trans Medical Imaging 17(4) Aug 1998
3. Zhou L, Rzeszotarski MS, Singerman L, Chokreff JM (1994) The detection and quantification of retinopathy using digital angiograms. IEEE Trans Med Imaging 13(4) Dec 1994
4. Hoover A, Kouznetsova V, Goldbaum M (2000) Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans Medical Imaging 19(3) March 2000. http://www.ces.clemson.edu/ ahoover
5. Chaudhuri M, Chatterjee S, Katz N, Nelson M, Goldbaum M (1989) Detection of blood vessels in retinal images using two dimensional matched filters. IEEE Trans Medical Imaging 8(3) Sept 1989
6. Pal NR, Pal SK (1989) Entropic thresholding. Signal Process 16:97–108

Implementation of Minutiae Extraction and False Minutiae Elimination for Fingerprint Image on FPGA

S. Gayathri and V. Sridhar

Abstract This paper presents FPGA implementation of minutiae extraction and false minutiae elimination for fingerprint image. A biometric system provides automatic identification of an individual based on a unique feature or characteristic possessed by the individual. Despite the widespread use of fingerprints, there is a little statistical theory on the uniqueness of fingerprint minutiae. A critical step in studying the statistics of fingerprint minutiae is to reliably extract minutiae from the fingerprint images. However, fingerprint images are rarely of perfect quality which requires image preprocessing techniques prior to minutiae extraction and false minutiae elimination. This will provide more reliable estimation of minutiae locations. Minutiae extraction and false minutiae elimination processes are implemented as a single block on FPGA. The FPGA configuration is generally specified using a hardware description language (HDL).

Keywords Biometrics · Fingerprint identification · Minutiae · FPGA · False minutiae · Preprocessing

S. Gayathri (✉)
SJCE, Mysore, Karnataka, India
e-mail: sgmurthy_65@yahoo.com

V. Sridhar
PESCE, Mandya, Karnataka, India
e-mail: venusridhar@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_2,

1 Introduction

Fingerprint-based identity verification is one of the most used biometric systems due to its easiness of acquisition, high distinctiveness, persistence, and acceptance by the public [4]. Usually, the recognition process involves a series of image enhancement and minutiae extraction steps that can be classified as follows:

(a) Fingerprint normalization
(b) Orientation and frequency image estimation
(c) Filtering
(d) Image binarization
(e) Image thinning
(f) Minutiae extraction
(g) False minutiae elimination.

Among these steps, minutiae extraction and false minutiae elimination are important for accurate fingerprint recognition process. The proposed method exhibits improved processing time when compared with MATLAB. The proposed work is carried out at three stages.

Stage 1: To implement two blocks of fingerprint recognition processes minutiae extraction and false minutiae elimination using MATLAB.
Stage 2: To implement these two blocks on FPGA as a single block.
Stage 3: Comparing the results of MATLAB and FPGA implementation.

In high-level fingerprint classification algorithms, extracting the number and precise location of singular points (SP), namely core and delta points are of great importance. The core is defined as the topmost point on the innermost upward recurving ridge, and delta point is defined as the point of bifurcation (in a delta-like region) on a ridge splitting in two branches which extend to encompass the complete pattern area [1]. The identification of these singular points in a typical fingerprint image is depicted in Fig. 1. To detect the core points, the mask must extract rotational regions and to detect the delta points, the mask should extract triangular regions. Most of the approaches proposed in the literature for singularity detection operate on the fingerprint orientation image [2].

Fig. 1 Fingerprint image

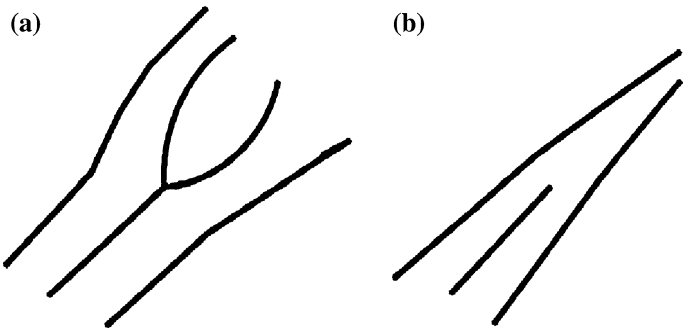


Fig. 2 Typical fingerprint image (**a**) ridgo bifurcation (**b**) ridge ending

Typical fingerprint image depicting the ridge ending and ridge bifurcation is shown in Fig. 2. Ridges (or the valleys) exhibit anomalies of various kinds such as ridge bifurcations, ridges endings, short ridges, and ridge crossovers. For automatic feature extraction and matching, the set of fingerprint features are restricted to two types of minutiae. They are ridge ending and ridge bifurcation. These singular points in a typical fingerprint image as ridge bifurcation and ridge ending are discussed in detail in [3].

In order to make fingerprint image invariant to position, the ridge density between two singular points in a fingerprint is computed [5]. Implementation of minutiae extraction using modified ridge following algorithm can be integrated onto the resource-constrained SoCs [6]. The minutiae extraction algorithm processes the fingerprint image in several stages in order to find the singular points related to bifurcation and termination of ridges. The number of stages and the processing involved in each one differs slightly depending on the algorithm employed [7].

2 Methodology

In image binarization, the pixel values which are above the threshold value are assigned as 1 and those below are assigned as 0. This improves the contrast between the ridges and valleys in a fingerprint image and consequently facilitates the extraction of minutiae.

Image thinning is performed prior to minutiae extraction. The thinning process consists of a set of iterative morphological operations that reduces the width of ridges until they are one pixel wide. Most minutiae extraction algorithms operate on binary images where there are only two levels of interest: the black pixels that represent ridges, and the white pixels that represent valleys. The most commonly employed method of minutiae extraction is the cross number (CN) concept.

Table 1 Properties of a ridge

Sl No	Property
0	Isolated point
1	Ridge ending point
2	Continuing ridge point
3	Bifurcation point

Minutiae extraction is carried out by examining the connectivity of the pixels in the thinned image. The CN for a ridge pixel P is given by:

$$\mathrm{CN} = 0.5 \sum_{i=1}^{8} |P_i - P_{i+1}|, \quad P_9 = P_1 \tag{1}$$

where P_i is the pixel value in the neighborhood of P. For a pixel P, its eight neighboring pixels are scanned in an anticlockwise direction. After the CN for a ridge pixel has been computed, the pixel can then be classified according to the property of its CN value. A pixel with a connectivity of 1 corresponds to an ending minutiae and a connectivity of 3 leads to a bifurcation, whereas otherwise is not a significant pixel. The properties of CN are shown in Table 1.

The minutiae are extracted by scanning the local neighborhood of each ridge pixel in the image using a 3 × 3 window. CN value is then computed, which is defined as half the sum of the differences between pairs of adjacent pixels in the eight-neighborhood.

False minutiae may be introduced into the image due to factors such as noisy images, and image artifacts created by the thinning process. Hence, after the minutiae are extracted, it is necessary to employ a postprocessing stage in order to validate the minutiae. Figure 3 illustrates some examples of false minutiae structures, which include the spur, hole, triangle, and spike structures. It can be seen that the spur structure generates false ridge endings, whereas both the hole and triangle structures generate false bifurcations. The spike structure creates a false bifurcation and a false ridge ending point.

The existence of false minutiae increases both FAR and FRR in fingerprint matching. Therefore, false minutiae elimination forms an important part of the system. False minutiae elimination uses Euclidean distance algorithm. The Euclidean distance between points p and q is the length of the line segment connecting them (pq(bar)). If $p = (p_1, p_2, \ldots, p_n)$ and $q = (q_1, q_2, \ldots, q_n)$ are two points in Euclidean n-space, then the distance from p to q, or from q to p, is given by:

$$\begin{aligned} d(p,q) = d(q,p) &= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} \\ &= \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}. \end{aligned} \tag{2}$$

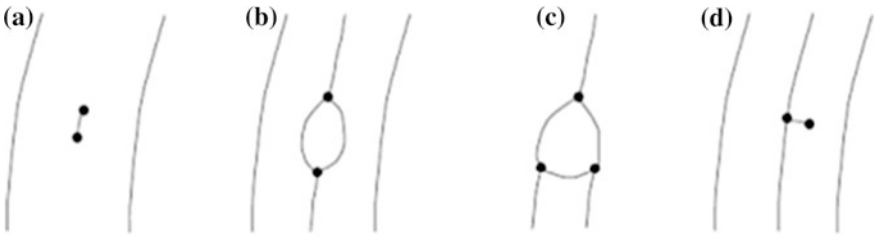


Fig. 3 Typical false minutiae structures. (**a**) Spur, (**b**) Hole, (**c**) Triangle, (**d**) Spike

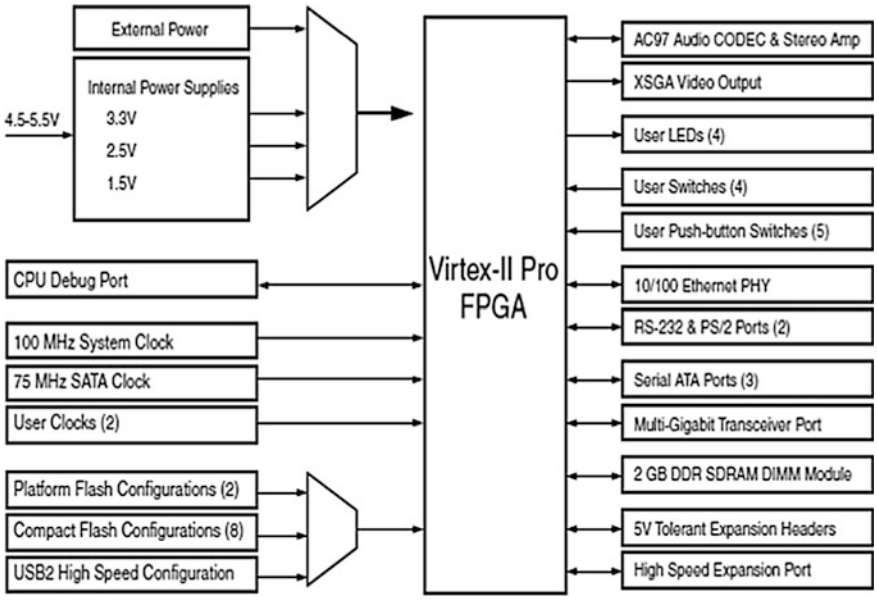


Fig. 4 XUP Virtex-II Pro Development System

3 Implementation of the Proposed Work

(a) Hardware Requirements:

The XUP Virtex-II Pro Development System provides an advanced hardware platform that consists of a high performance Virtex-II Pro Platform FPGA surrounded by a comprehensive collection of peripheral components that can be used to create a complex system and to demonstrate the capability of the Virtex-II Pro Platform FPGA. The block diagram of XUP Virtex-II Pro Development System is shown in Fig. 4.

Advantages of FPGA are as follows:

1. Involves designing hardware to meet specific needs

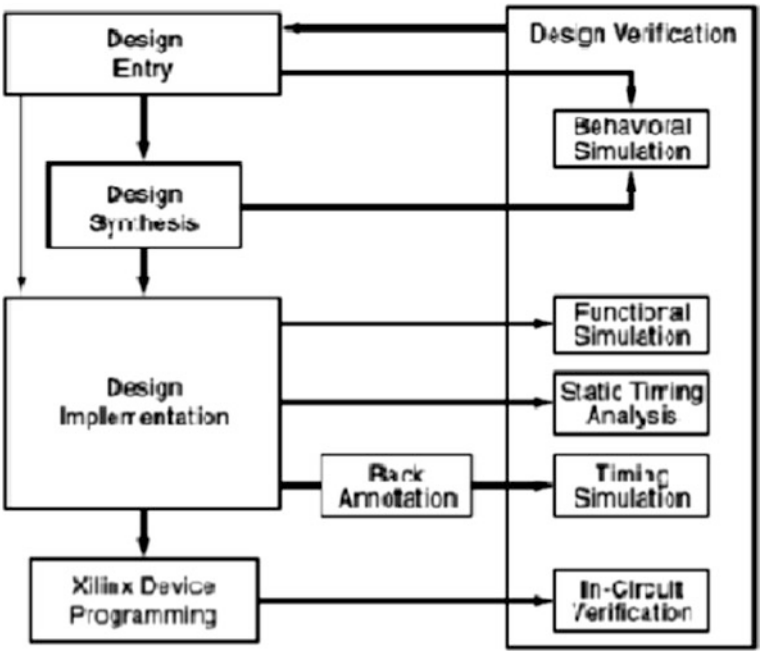


Fig. 5 Xilinx design flow

2. Architecture can be changed whenever required.
3. Portable
4. Lower power consumption
5. Higher processing speed
6. Higher reliability
7. Lower production cost
8. Parallel processing

(b) Software Requirements:

It consists of the following:

1. Xilinx 10.1 simulator
2. MATLAB version R2008b
3. Graphical user interface

Xilinx ISE simulator is a software tool produced by Xilinx for synthesis and analysis of HDL designs, which enables the developer to synthesize (compile) their designs, perform timing analysis, examine RTL diagrams, simulate a design's reaction to different stimuli, and configure the target device with the programmer. The Xilinx design flow is shown in Fig. 5.

(c) Algorithm for Minutiae Extraction:

In general, cross number algorithm is used for minutiae extraction. Consider 3 × 3 matrix at a time.

$$\mathrm{CN} = 0.5 \sum_{i=1}^{8} |P_i - P_{i+1}|, \quad P_9 = P_1 \tag{3}$$

where P_i is the pixel value in the neighborhood of P. For a pixel P, its eight neighboring pixels are scanned in an anticlockwise direction as follows:

P_4	P_3	P_2
P_5	P	P_1
P_6	P_7	P_8

Check for the central element (i.e., [2, 2] element of the 3 × 3 matrix). If the central element is 0, then output = 0. If the central element is 1, then output = (sum of all the elements of the 3 × 3 matrix) 1. As mentioned before, if the output is 1, then the central element of the considered 3 × 3 matrix is the ridge termination point. If the output is 3, then the central element of the considered 3 × 3 matrix is the ridge bifurcation point. This is depicted in Fig. 6.

This method performs the same function as that of the cross number algorithm. This algorithm is written as a function and called with nonlinear filter operations. The nonlinear filter performs sliding and neighboring operations and takes one pixel at a time, and this algorithm is applied to that pixel.

(d) Algorithm for False Minutiae Elimination:

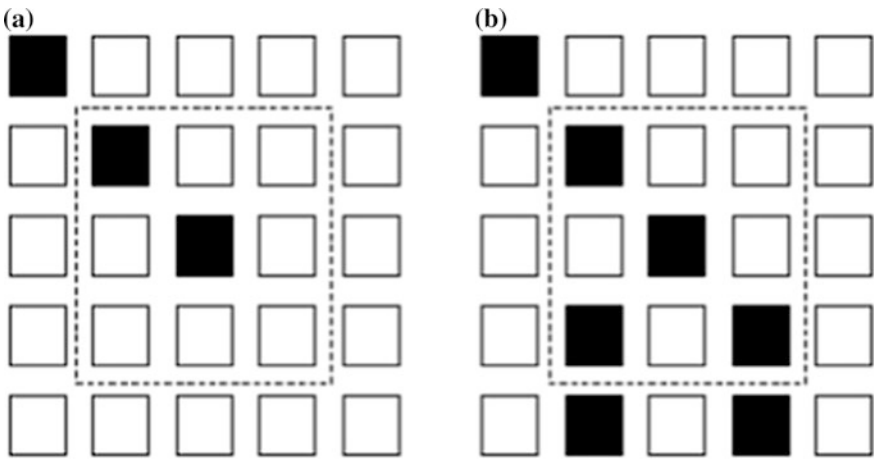


Fig. 6 Ridge ending and bifurcation pixel. (**a**) Crossing number of one corresponds to ridge ending pixel. (**b**) Crossing number of one corresponds to bifurcation pixel

False minutiae elimination uses Euclidean distance algorithm. The Euclidean distance is the ordinary distance between two points that one would measure with a ruler and is given by the Pythagorean formula. The Euclidean distance between points p and q is the length of the line segment connecting them (pq(bar)). If $p = (p_1, p_2, \ldots, p_n)$ and $q = (q_1, q_2, \ldots, q_n)$ are two points in Euclidean n-space, then the distance from p to q, or from q to p, is given by:

$$d(p,q) = d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}. \quad (4)$$

After minutiae extraction, (x, y) coordinates of the minutiae points are saved in two columns matrix form. The number of rows is equal to the number of the minutiae points. Terminations are saved as one dataset, and bifurcations are saved as another dataset. The distance from one point to each and every point is calculated, and the results in 2-D array are saved. The value of D is found, and it is 6 for the images under consideration. The average inter-ridge width D refers to the average distance between two neighboring ridges. Calculation of D value is as follows:

A row of the thinned ridge image is scanned, and all pixels in the row are summed up whose value is one. Then, the row length with the above summation to get an inter-ridge width is divided. For more accuracy, such kind of row scan is performed upon several other rows, and column scans are also conducted. Finally, all the inter-ridge widths are averaged to get the D. The points for which distance between two points is less than 6 and greater than 0 are eliminated. Three types of process are carried out here:

process 1: If the distance between a termination and a bifurcation is smaller than D, remove this minutia.
process 2: If the distance between two bifurcations is smaller than D, remove this minutia.
process 3: If the distance between two terminations is smaller than D, remove this minutia.

(e) Algorithm for creating coefficient file:

The elements in the coefficients file will be very large. Nonlinear operations before loading the image into FPGA are performed. So, the input image is resized to 200 × 200 for FPGA processing. If the image is compressed, there will be loss of data in the image.

1. Image is a 2-D array. Convert this into linear array. For nonlinear filter operations, consider a 3 × 3 matrix of image

1 0 1
0 1 0
0 1 0

2. Now pad the zeros in one line around the matrix. Matrix will become

0 0 0 0 0
0 1 0 1 0
0 0 1 0 0
0 0 1 0 0
0 0 0 0 0

3. Now start the neighboring and sliding operations for the image matrix and save the linear array in a proper format for coefficient file.

4 Experimental Results

Graphical user interface allows users to interact with electronic devices using images rather than text commands. GUI is used to display the results of fingerprint recognition process. The algorithm is implemented both in MATLAB and in FPGA. The original input fingerprint image and the simulation of output of minutiae extraction using MATLAB are shown in Fig. 7 and that of false minutiae elimination in Fig. 8. The processing time of simulation of minutiae extraction and false minutiae elimination using MATLAB are in the order of few microseconds.

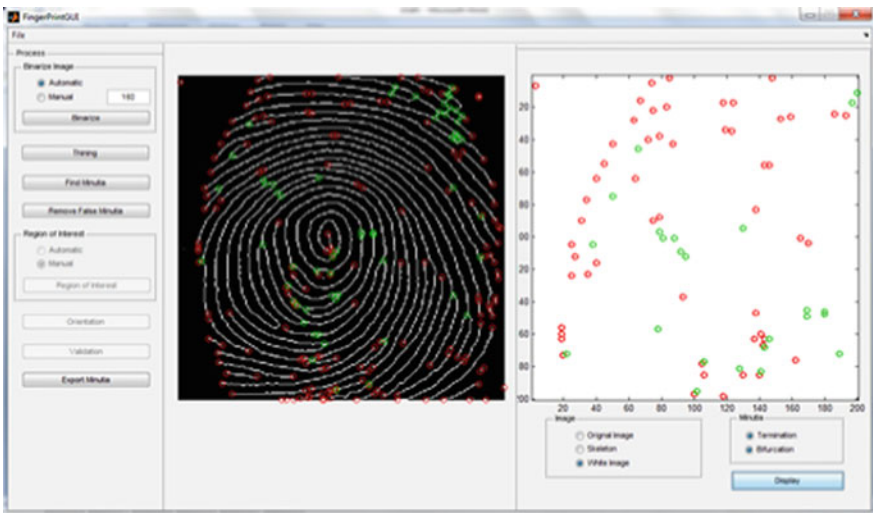


Fig. 7 Output of minutiae extraction

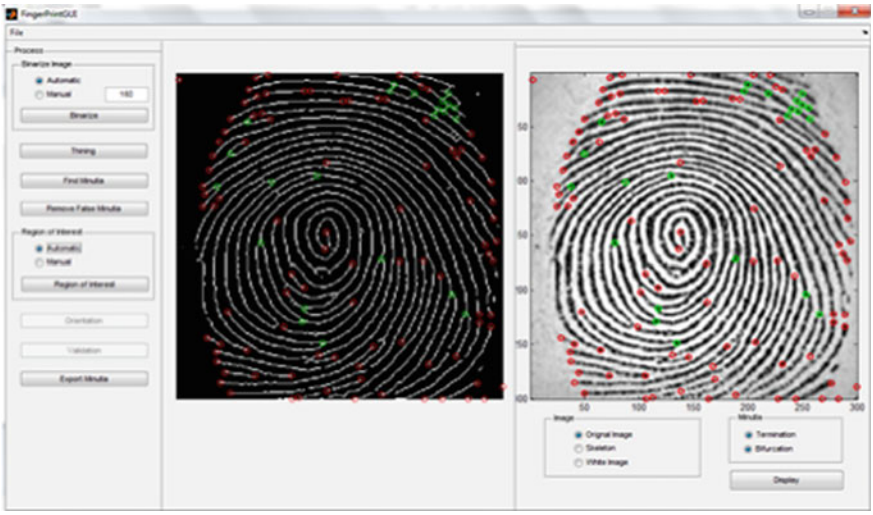


Fig. 8 Output of false minutiae elimination

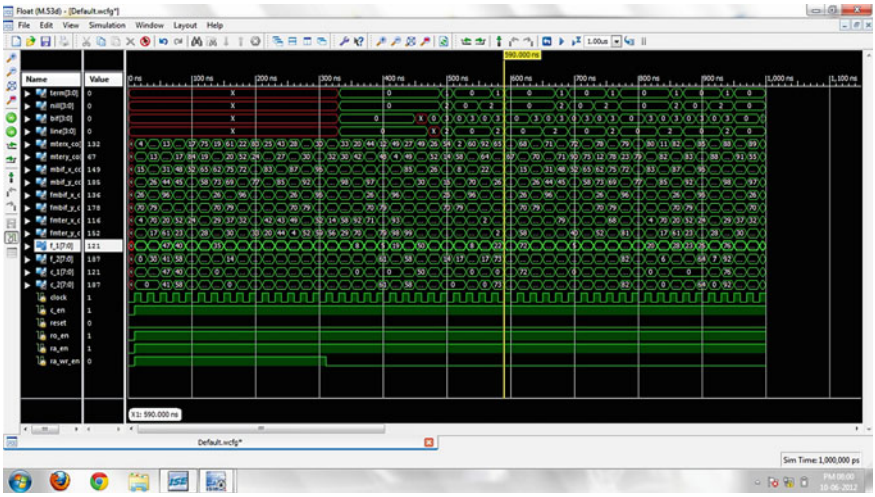


Fig. 9 Simulation timing diagram of minutiae extraction

The simulation results of ModelSim simulator for minutiae extraction and false minutiae elimination are illustrated in Figs. 9 and 10.

The input and the output fingerprint images of FPGA are shown in Fig. 11. The two blocks of fingerprint recognition blocks, namely minutiae extraction and false minutiae elimination, are combined as a single block in FPGA to save the memory

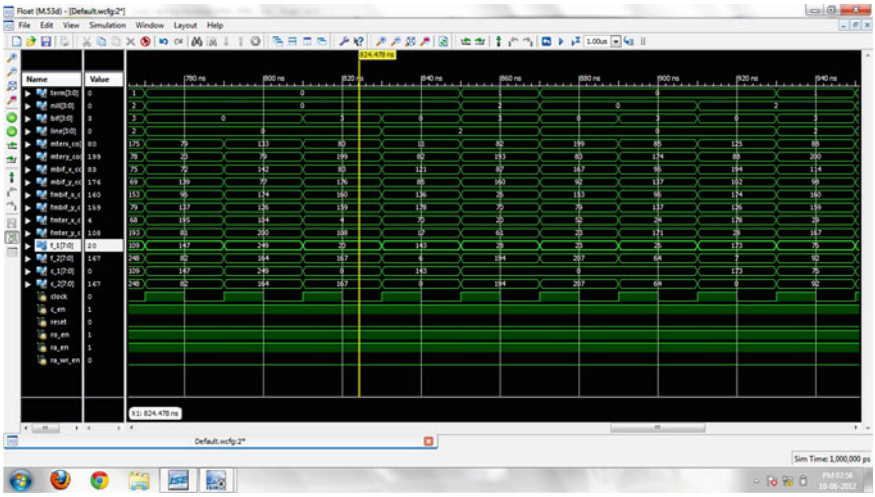


Fig. 10 Simulation timing diagram of false minutiae elimination

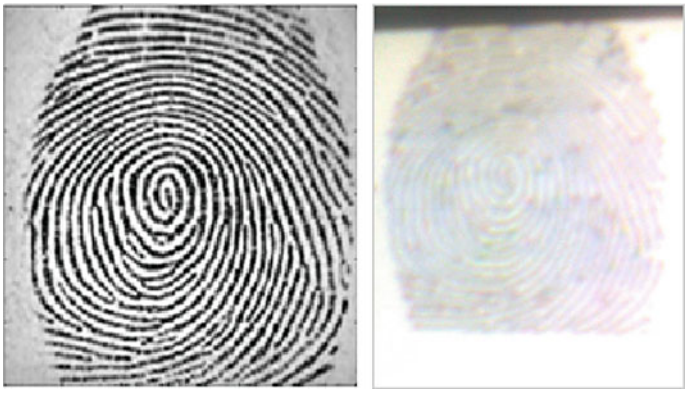


Fig. 11 FPGA input and output images

usage. The output image of FPGA as seen on a VGA monitor is illustrated in Fig. 11. The simulation of minutiae extraction and false minutiae elimination blocks as a single block on FPGA reduces the processing time and is in the order of nanoseconds.

5 Conclusion

In this work, minutiae extraction and false minutiae elimination are implemented using MATLAB. Then, minutiae extraction and false minutiae elimination blocks are implemented as a single block on the Virtex-II Pro FPGA board. The output image is displayed on the monitor using VGA controller for FPGA implementation. Experimental results shows that the processing time will be in the order of nanoseconds for FPGA implementation, whereas processing time will be in the order of microsecond for MATLAB implementation.

6 Future Work

The fingerprint recognition process involves series of image enhancement and minutiae extraction processes. Only a part of the recognition process, namely minutiae extraction and false minutiae elimination, is demonstrated on FPGA. Further, the remaining blocks of the recognition process can be implemented on FPGA, and it has been proposed to integrate all the blocks as a single unit.

References

1. Halici U, Jain LC, Erol A (1999) An introduction to fingerprint recognition. In: Jain LC, Halici U, Hayashi I, Lee SB, Tsutsui S (eds) Intelligent biometric techniques in fingerprint and face recognition. CRC Press, Florida, p 334
2. Mohammadi S, Farajzadeh A (2009) Fingerprint reference point detection using orientation field and curvature measurements. IEEE
3. Maltoni D, Maio D, Jain AK, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
4. Fingerprint Veri_cation Competition 2004 (FVC2004). Available http://bias.csr.unibo.it/fvc
5. Maio D, Maltoni D (1997) Direct gray-scale minutiae detection in fingerprints. IEEE Trans Pattern Anal Machine Intell 19(1):27–39
6. Pan SB, Moon D, Kim K, Chung Y (2006) A VLSI implementation of minutiae extraction for secure fingerprint authentication (vol 2). International conference on computational intelligence and security, pp 1217–1220
7. FBGA Implementation of a minutiae extraction fingerprint algorithm (2008) IEEE 8:1920–1925

Resistive Fractal FSS Based Broadband Radar Absorber

Chandrika Sudhendra, Peter Jose, ACR Pillai and KARK Rao

Abstract A novel broadband circuit analog radar absorber using three-layer ***resistive*** fractal frequency-selective surfaces (FSS) is described. The radar absorber is designed with weight and thickness constraints. Using the innovative concept of embedded passives (EP) resistors, a large number of resistors which are essential for absorption have been designed and integrated into the FSS substrate, totally dispensing with soldering of discrete passives. The ultra wideband radar absorber is simulated using the high-performance 3D electromagnetic (EM) simulation software, HFSS^{TM}. Extensive parametric studies are carried out to study the effect of various design parameters on the performance. The resistive FSS layers are designed and developed as electrically thin printed circuit boards (PCB), and highly accurate photolithographic technology is used for fabrication. Simulation results are verified by performing RCS experiments on the panel RAM. Monostatic RCS experiments are carried out on the panel RAM in microwave anechoic chamber. A minimum radar cross section reduction (RCSR) of −10 dBsm has been realized from 3 to 12 GHz.

Keywords RCS · RCSR · RAM · RAS · FSS · Circuit analog radar absorber

C. Sudhendra (✉) · ACR Pillai
ADE, DRDO, Bangalore, India
e-mail: chansudhtumkur@gmail.com

P. Jose
IDST, Bangalore, India

K. Rao
PES College of Engineering, Mandya, India
e-mail: karkrao@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_3,

1 Introduction

Aircraft/unmanned air vehicle (UAV) design for stealth encompasses various signature reduction methods for reduction in radar, acoustic, IR, thermal and visual signatures. Radar stealth is of primary concern as radar is the primary day and night, all-weather, long-range sensor. Radar stealth design comprises two important methods, namely external geometric shaping and application of radar absorbing materials (RAM). Both methods aim at reducing the monostatic radar cross section (RCS), which is an EM parameter and is the electromagnetic signature of the aircraft. Shape design and application of radar absorbers become *mandatory* for realizing radar stealth.

But, RAM design is fraught with many challenges such as need for ever-increasing absorption bandwidths with weight and thickness constraints. Materials that are used in RAM design should be airworthy, and RAM needs to be designed as a flight worthy structure, capable of sustaining aerodynamic loads. RAM design is highly interdisciplinary and should consider inputs from aerodynamics, structures, and materials. Absorption bandwidths usually are accompanied by added thicknesses. Jaumann radar absorbers are a typical example of multilayer RAM wherein wide bandwidths are plagued by multiple quarter wavelength thicknesses, which limit their use in airborne military applications [1, 2]. Circuit analog dielectric absorber designs address the thickness problem to a large extent wherein the concept of lossy or resistive FSS is used to design RAM for wide absorption bandwidths with thickness constraints. But the design challenge in FSS-based circuit analog absorbers is to realize ***resistive*** FSS since pure FSS does not absorb energy and find application in the design of FSS-based radomes for *out-of-band* RCS reduction in stealth antennae. Broadband circuit analog RAMs are reported in [3, 4]. But these designs employ surface mount chip resistors for realization of resistive FSS. This has limitations such as the unavoidable parasitics associated with discrete passives at microwave frequencies, which cannot be modeled accurately. A large number of discrete resistors in RAM result in reduced reliability due to soldering related defects, cross talk, and interference. In addition, the prohibitive cost of high-frequency resistors and complicated assembly limits their applications. In [5], using lossy FSS comprising of tantalum nitride material, a wideband RAM is described. In our earlier papers [6, 7], we have reported resistive FSS-based wideband RAM, wherein ***embedded passives (EP) resistors*** have been innovatively used for design and fabrication of resistive FSS. Based on this concept, several thousands of discrete resistors are integrated into the FSS PCB, totally eliminating soldering and associated soldering related defects of discrete passives. In this paper, we describe a three-layer circuit analog RAM with RCSR of -10 dBsm from 3 to 12 GHz., with weight constraints. The three resistive FSS layers are realized as electrically thin (thickness $\ll \lambda$, thickness $= 5$ mils) PCBs and are fabricated using conventional PCB fabrication technology.

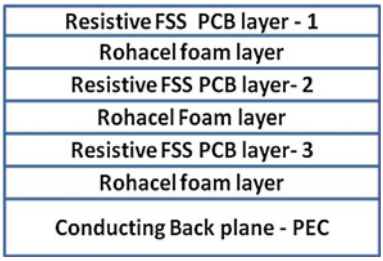


Fig. 1 Dielectric profile of the proposed three-layer circuit analog RAM

2 Electromagnetic Design and EM Simulation

Dielectric profile of the proposed wideband dielectric panel RAM is given in Fig. 1. It comprises three FSS layers, which are designed as lossy/resistive FSS layers, each separated by a low dielectric *Rohacell* foam spacer. The RAM essentially comprises a conducting back plane, modeled as perfect electric conductor (PEC), whose RCS needs to be reduced. The total thickness of RAM is 10.6 mm. The thickness of central foam layer is 4 mm, and the thickness of the top and bottom foam layers is 3 mm. The resistive FSS PCB layers are developed as PCBs using FR4 substrate of thickness 5 mils, with dielectric constant, $\varepsilon_r = 4.4$ and $\tan\delta = 0.002$.

The transmission line equivalent circuit of three-layer RAM is given in Fig. 2. In this figure, $\boldsymbol{Z_0}$ is the input impedance, ***Rs*** is the resistance of each resistive FSS layer, $j\omega L$ is the inductive reactance of the fractal loop layer, and $(1/j\omega C)$ is the capacitive reactance. Each resistive FSS layer is modeled as a series RLC circuit shunted across short-circuited transmission line. The resulting equation is a simple equation for three resonant resistive FSS structures terminated in short-circuited transmission line. The full-wave, 3D high-performance HFSS^{TM} simulation software is used for numerical analysis of the design, and the design is optimized for desired wideband performance. The geometry model of the three-layer RAM in HFSS is given in Fig. 3. The RAM geometry model in HFSS comprises a unit cell of the fractal FSS-based RAM. Applying ***Floquet's theorem*** for periodic FSS,

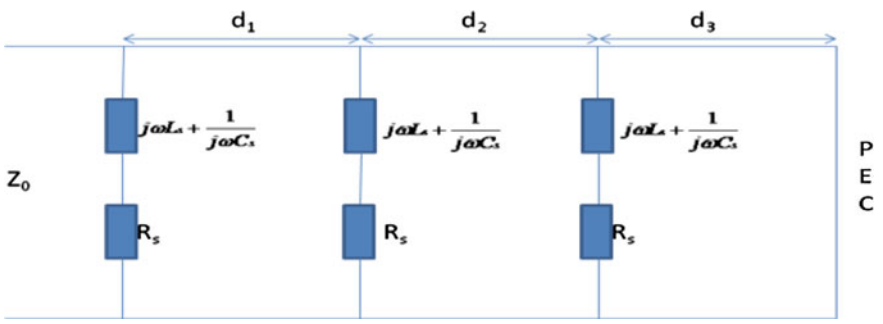


Fig. 2 Equivalent circuit of three-layer circuit analog RAM

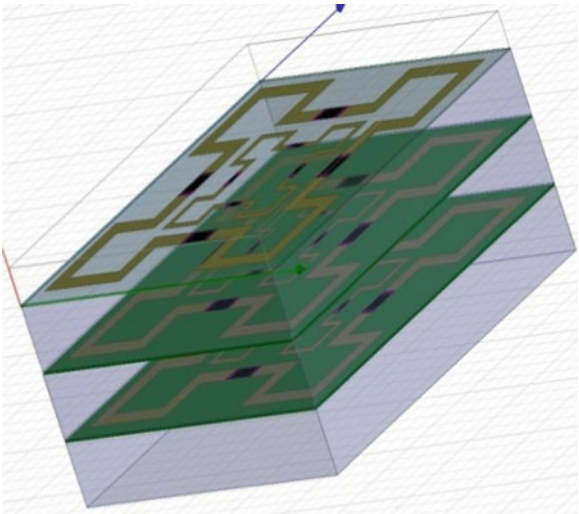


Fig. 3 RAM geometry model in HFSS simulation software. *Black rectangles* at the center of each side are the embedded passives resistors

the complete FSS structure with the conducting backplane can be analyzed by using a unit cell comprising of Minkowski fractal loop with resistors at the center of each side of the fractal FSS loop.

The resistors at the center of each side of the two loops are modeled as EP resistors. Dielectric spacers are modeled as Rohacel^{TM} foam spacers with dielectric constant, $\varepsilon_r = 1.07$ and $\tan\delta = 0.0003$. The interelement spacing is 6.5 mm, which is $<0.5\ \lambda_0$, where λ_0 is the wavelength at highest frequency in the band [8], to avoid the onset of free space grating lobes.

As noted earlier, the **crux of FSS-based circuit analog radar absorber lies in the design and realization of FSS as resistive FSS**. **The novelty of this paper lies in using the concept of embedded passives or integral passives for the design and fabrication of resistors.** The design details of EP resistors with dimensions are given in Fig. 4. In this figure, w_2 is the width of the inner loop, w_1 is the width of the outer loop, R_{in} is the resistor in the inner loop, R_{out} is the resistor in the outer loop, L_{out} is the length of the outer loop, and L_{in} is the length of the inner loop. The resistors used in the two loops for the three layers are listed in Table 1. The total thickness of RAM is 10.6 mm. It is noted from [9] that the least thickness of a 10-dB broadband

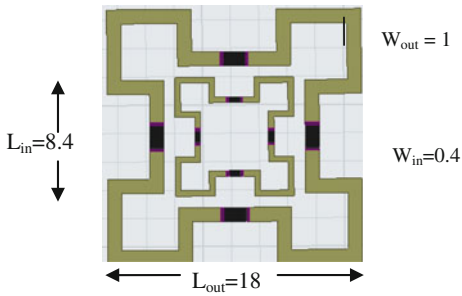


Fig. 4 Design details of unit resistive fractal FSS cell. All dimensions in mm. Drawing not to scale

Table 1 Resistor details in the three fractal FSS loops

Sl. No	Resistive FSS layer	Resistor value in ohms	
		Inner loop (Rin)	Outer loop (Rout)
1	Top layer	240	520
2	Middle layer	200	180
3	Bottom layer	320	500

dielectric radar absorber is (1/17) of the largest operating wavelength. Accordingly, the desired bandwidth may be realized with a thickness not less than 6 mm.

Figure 5 gives the optimized RCSR/***S11*** performance of wideband RAM. Simulation results predict that a minimum RCSR of −10 dBsm may be realized from the three-layer RAM from 3 to 12 GHz.

Extensive parametric studies have been carried out to assess the *sensitivity* of design to all design parameters. A representative result is given in Fig. 6. In this simulation plot, the effect of variation of resistance, $\boldsymbol{R_{out}}$ in the outer loop of top FSS layer on ***S11*** performance is shown. It is observed that with increasing values of resistance, the ***S11*** degrades along with corresponding decrease in the bandwidth at lower edge of the band. An optimal value of 520 Ω is chosen for realizing the desired RCSR bandwidth from 3 GHz. The value of $\boldsymbol{R_{out}}$ in all the three layers affects the low-frequency RCSR. Similarly, resistance in the inner loop affects the RCSR performance at higher band edge frequencies, and an optimal value should be used for realizing the desired RCSR bandwidth.

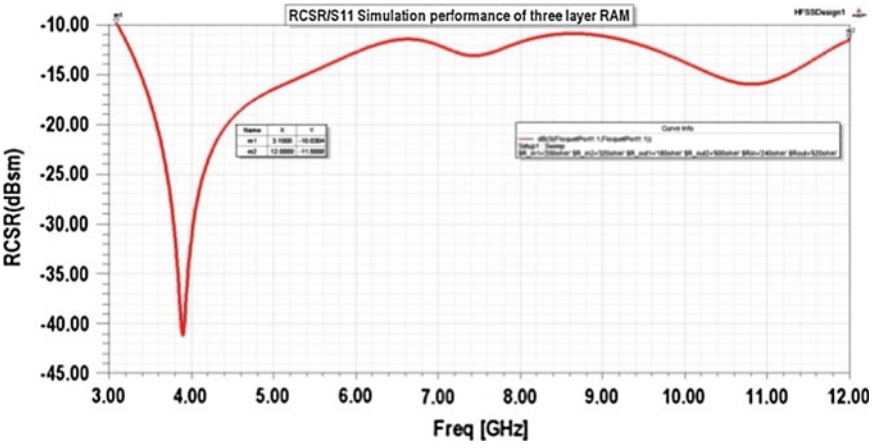


Fig. 5 Optimized simulation performance of three-layer RAM

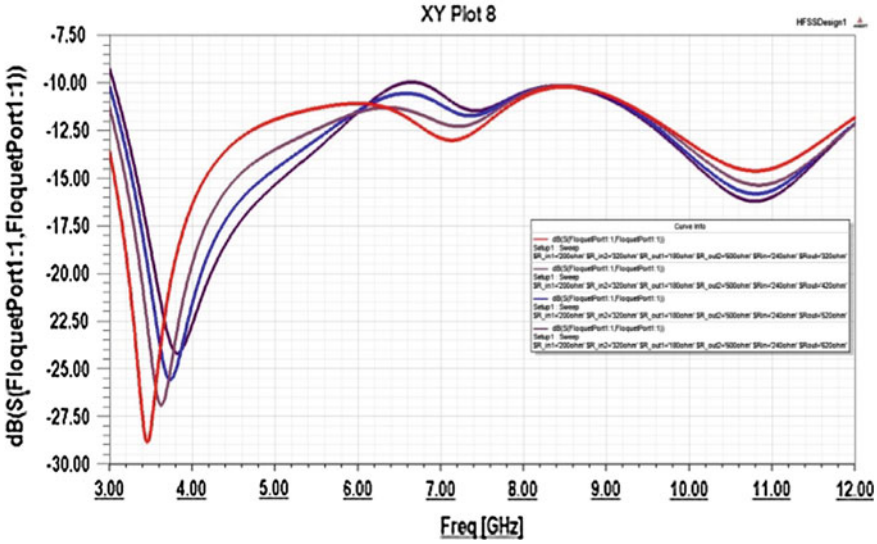


Fig. 6 Parametric simulation studies. R_{out} in the top layer only varying. All other parameters fixed

3 Resistive FSS PCB Layers Design and Fabrication

Each of the three resistive FSS layers is designed as electrically thin, RF/microwave PCB using the PCB design software, **Visula**™ v 2.3. The PCB layout design and fabrication of embedded passives resistors is explained in detail in our earlier paper [6]. It is noted that six different values of resistors are required to be designed for a unit cell for the three FSS layers. *This is enabled with the use of EP resistors*. FR4 substrate of thickness 5 mil is used for fabrication of the EP resistor–based FSS PCBs. A commercially available thin film–based resistive sheet of value 250 Ω/sq. is used for fabrication of resistors. Fabrication of EP resistors requires 'selective etching' and is carried out using highly accurate PCB fabrication technology. A photograph of the assembled three-layer panel RAM is shown in Fig. 7. A blown up single cell is shown in the inset of the same figure. The RAM consists of a total of 2,904 resistors with each resistive FSS PCB layer consisting of 968 resistors, fabricated as EP resistors, without any soldering at all. The three Rohacell™ foam spacers are bonded to the FSS PCB layers using a double-sided FIXON® tape. A carbon fiber reinforced plastic (CFRP) panel of thickness 0.6 mm is used as the conducting backplane of RAM.

Fig. 7 Photograph of the fabricated resistive FSS *top* PCB layer. Inset: A blown up single cell

4 RCS Measurements

RCS experiments are carried out on the panel RAM in microwave anechoic chamber to verify the RCSR performance. Monostatic RCS measurement setup is used for RCS measurements. The panel RAM is placed on an RF transparent thermocol stand and rotated 360° in azimuth using a single-axis positioner. The CFRP conducting backplane serves as a reference with which the reduced RCS return from the absorber side is compared. RCS results are recorded for vertical polarization. High-directivity standard gain horn antennas are used for transmission and reception. Continuously varying phase shifter and attenuator are used in the two sampled ports of directional couplers connected to the transmitting and receiving antenna for performing vectorial cancelation of the background at each measurement frequency. A representative RCSR plot is given in Fig. 8, for 4.5 GHz. Measurements at other frequencies in the band are carried out for vertical polarization, and the results are available. RCS measurements at 3 GHz could not be carried out due to limitations in experimental setup.

5 Discussion of Results

A three-layer RAM based on resistive fractal FSS for RCSR of −10 dBsm from 3 to 12 GHz described in this paper is designed with weight constraints. From RCS experiments, it is observed that the simulated and measured RCS results of RAM agree very well. This is attributed to the use of EP resistors, which have been modeled and fabricated with high accuracy. No shift in the frequency is observed in simulation and experiments due to absence of parasitic effects. The weight of panel RAM is 145 g. This low weight of RAM has been realized by the use of low dielectric foam spacers, and the resistive FSS layers have been designed and

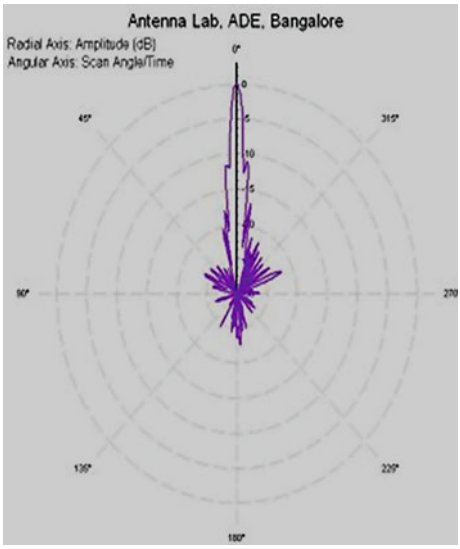


Fig. 8 Measured monostatic RCSR of three-layer RAM. Polarization: VP. RCSR at 4.5 GHz = 22 dB

realized as infinitesimally thin microwave PCBs (thickness = 5 mil). A total of 2,904 numbers of resistors which are mandatory for realizing the RCSR performance from RAM have been realized as EP resistors without any soldering. An RCSR of −12 dBsm is measured at 7, 8, and 9 GHz. At all other frequencies, the measured and simulated RCS values agree to within 0.2–0.5 dB.

6 Conclusion

In order to meet the requirements of broadband RCSR of −10 dBsm (minimum) from 3 to 12 GHz with minimal weight, a resistive FSS-based three-layer RAM is described in this paper. From simulation and RCS experiments on RAM, it is observed that the RAM meets the desired specifications with weight constraints. The thickness of RAM is 10.6 mm and is suitable for airborne stealth application. Due to the two-dimensional symmetry of RAM structure, RAM can be used for circular polarization. Experiments have been carried out to measure the RCSR at ***C***- and ***X***-bands. The simulation and experimental RCS results agree very well. Most important, the innovative use of EP resistors concept for realization of crucial resistive FSS layers results in a quantum leap in reliability. The conducting backplane used in RAM construction is CFRP and is an airworthy structural material. From earlier RCS measurements in a compact antenna/RCS measurement facility, it was found that the RCS returns from aluminum and CFRP panels of same size agreed to within 0.2 dB. The thickness chosen for CFRP meets the skin depth requirements and can be used in airborne applications.

The wideband RAM meets the RCSR bandwidth specifications from 3 to 12 GHz and is suitable for aircraft/UAV stealth applications.

Acknowledgments The authors would like to place on record their grateful thanks to Shri. PS Krishnan, Distinguished Scientist and Director, ADE, for his continued guidance, support, and encouragement. We convey our hearty thanks to Shri. Gurudev SG, Group Director, ADE for his unstinted guidance. We gratefully acknowledge the support of Dr. V. Ramachandra, Scientist G, Head, FTTT division of ADE for RCS measurements in the anechoic chamber, at FTTT division, ADE. Special thanks are due to Ms. R. Nagarathna, STA 'C' for PCB layout design and Mr. Mahalingam, Scientist 'F', OIC, EMI/EMC, PCBD and PCBF groups, Mr. NA Boregowda, TO 'B', Mr. S. Vinod, STA 'B', and Mr. N. Vijaykumar, Tech 'A', for accurate and speedy fabrication of resistive FSS PCBs.

References

1. Knott EF, Shaeffer JF, Tuley M (1993) Radar cross section. Artech House, USA
2. Munk BA (2000) Frequency selective surfaces—theory and design. Wiley, New York
3. Luo Y, Zhuang Y, Zhu S (2009) Thin and broadband Salisbury screen absorber using Minkowski fractal structure. IEEE pp 2573–2574, 978-1-4244-2802-1/2009
4. Yang J, Shen Z (2007) A thin and broadband absorber using double-square loops. IEEE Antennas Wireless Propag Lett 6:388–391
5. Li M, Xiao SQ, Bai YY, Wang BZ (2012) An ultrathin and broadband radar absorber using resistive FSS. IEEE Antennas Wireless Propag Lett 11:748–751
6. Sudhendra C, Mahule V, Nagarathna R, Pillai ACR, Rao K (2010) A broadband circuit analog radar absorber using FSS and embedded passives resistors. 6th international conference on microwaves, antenna propagation and remote sensing, ICMARS 2010, Jodhpur, 14–17th Dec 2010
7. Sudhendra C, Safare SB, Mahule V, Pillai ACR, Rukmini TS, Rao K (2011) A thin wide band circuit analog radar absorber based on novel embedded passives resistors and fractal FSS. In the session on components and devices, in 11th International. Radar symposium India- IRSI - 2011, Bengaluru, India 03–04 Dec 2011
8. Munk B, Munk P, Prior J (2007) On designing Jaumann and circuit analog absorbers (CA Absorbers) for oblique angle of incidence. IEEE Trans Antennas Propag 55(1):186–193
9. Rozanov KN (2000) Ultimate thickness to bandwidth ratio of radar absorbers. IEEE Trans Antennas Propag 48(8):1230–1234

FPGA and ASIC Implementation of 16-Bit Vedic Multiplier Using Urdhva Triyakbhyam Sutra

K. B. Jagannatha, H. S. Lakshmisagar and G. R. Bhaskar

Abstract The ever increasing demand in enhancing the ability of processors to handle the complex and challenging processes has resulted in the integration of a number of processor cores into one chip. Still the load on the processor is not less in generic system. This load is reduced by supplementing the main processor with coprocessors, which are designed to work upon specific type of functions like numeric computation, signal processing, and graphics. The speed of arithmetic logic unit (ALU) depends greatly on the multiplier. In algorithmic and structural levels, numerous multiplication techniques have been developed to enhance the efficiency of the multiplier which concentrates in reducing the partial products and the methods of their addition but the principle behind multiplication remains the same in all cases. Vedic mathematics [1] is the ancient system of mathematics which has a unique technique of calculations based on 16 sutras. Employing these techniques in the computation algorithms of the coprocessor will reduce the complexity, execution time, area, power, etc. Our work has proved the efficiency of Urdhva Triyagbhyam—Vedic method for multiplication which strikes a difference in the actual process of multiplication itself. It enables parallel generation of intermediate products. This sutra is to be used to build a high-speed power-efficient multiplier in the multiply-accumulate (MAC) unit. First, field programmable gate array (FPGA) realization is achieved and next standard cell–based ASIC design of the multiplier is realized. In 180 nm CMOS technology, our speed is 5.2 ns, 257 uW, and its using 1,117 cells.

K. B. Jagannatha (✉) · H. S. Lakshmisagar
Department of Electronics and Communication Engineering, BMS Institute of Technology, Bangalore, Karnataka, India
e-mail: jagan.bmsit@gmail.com

H. S. Lakshmisagar
e-mail: sagar8.hs@gmail.com

G. R. Bhaskar
Department of Electronics and Communication Engineering, Acharya Institute of Technology, Bangalore, Karnataka, India
e-mail: bhaskar.sjbit@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_4,

Keywords Vedic mathematics · Urdhva Triyakbhyam Sutra · Field programmable gate array (FPGA)

1 Introduction

Vedic mathematics [1] was rediscovered in the early twentieth century from ancient Indian sculptures (Vedas). Ancient Indian system of mathematics was derived from Vedic sutras. The conventional mathematical algorithms can be simplified and even optimized by the use of Vedic mathematics. The Vedic algorithms can be applied to arithmetic, trigonometry, plain and spherical geometry, and calculus. In [2], the work presented in this thesis [1] makes use of Vedic mathematics and goes step by step, by first designing a Vedic multiplier, then a multiply-accumulate (MAC) unit. Multiplication basically is the mathematical operation of scaling one number by another. Talking about today's engineering world, multiplication-based operations are some of the frequently used Functions, currently implemented in many digital signal processing (DSP) applications such as convolution, fast Fourier transform, filtering and in arithmetic logic unit (ALU) of microprocessors. Since multiplication is such a frequently used operation, it is necessary for a multiplier to be fast and power efficient and so, development of a fast and low-power multiplier has been a subject of interest over decades. MAC operation is also a commonly used operation in various DSP applications. Now, not only digital signal processors, but also general-purpose microprocessors come with a dedicated MAC unit. When talking about the MAC unit, the role of multiplier is very significant because it lies in the data path of the MAC unit and its operation must be fast and efficient. Two most common multiplication algorithms followed in the digital hardware are array multiplication algorithm and booth multiplication algorithm which has more propagation delay. But "Vedic multiplier" has minimum propagation delay.

The paper is organized as follows. Section 2 presents an overview of Urdhva tiryakbhyam Sutra. In Sect. 3, it is design and analysis of Vedic multiplier. Section 4 shows field programmable gate array (FPGA) results, Sect. 5 shows standard cell–based ASIC design synthesis report, Sect. 6 results and discussion, and Sect. 7 concludes the paper.

2 Urdhva Tiryakbhyam Sutra

The "Urdhva tiryakbhyam Sutra" [3] or "Vertically and Crosswise Algorithm" for multiplication is discussed and then used to develop digital multiplier architecture. This looks quite similar to the popular array multiplier architecture. This sutra shows how to handle multiplication of a larger number (N × N, of N bits each)

Fig. 1 RTL 16 × 16 multiply block

by breaking it into smaller numbers of size (N/2 = n, say) and these smaller numbers can again be broken into smaller numbers (n/2 each) till we reach multiplicand size of (2 × 2), thus simplifying the whole multiplication process. The multiplication algorithm is then illustrated to show its computational efficiency by taking an example of reducing a N × N bit multiplication to a 2 × 2 bit multiplication operation. This work presents a systematic design methodology for fast and area-efficient digit multiplier based on Vedic mathematics, and then, a MAC unit has been made which uses this multiplier (Fig. 1).

3 Design and Analysis of Vedic Multiplier

The hardware realization of 2 × 2 multiplier blocks [7] (Figs. 2, 3, 4).

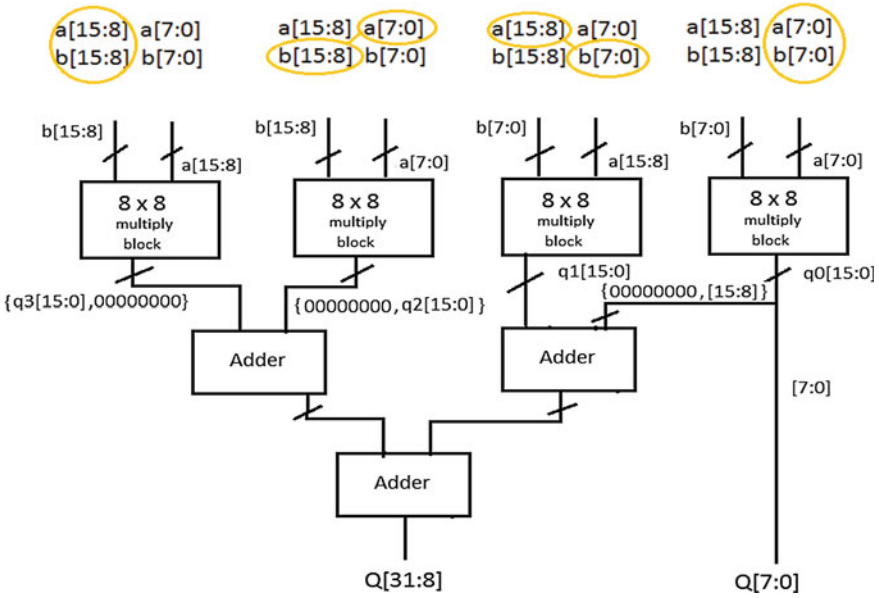


Fig. 2 The hardware realization of 2 × 2 multiplier blocks

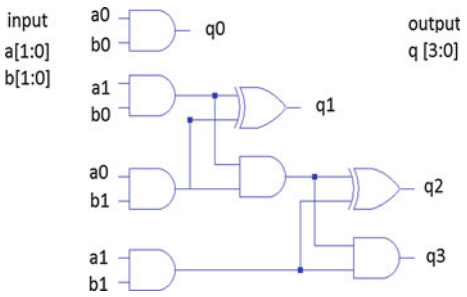


Fig. 3 16 × 16 multiply block

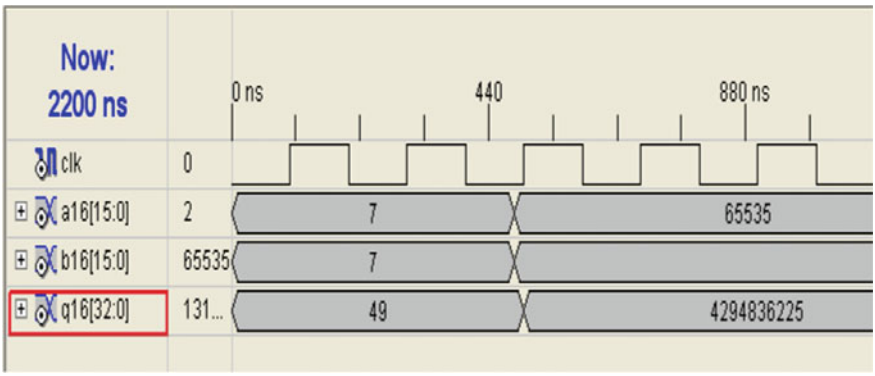


Fig. 4 Simulation waveform for multiplier

4 FPGA Simulation Outputs

Device utilization summary for multiplier		
Number of slices	330 out of 3584	9 %
Number of 4 input LUTs	99 out of 7168	8 %
Number of bonded IOBs	65 out of 97	67 %

5 Standard Cell–Based ASIC Design

In standard cell–based ASIC design, the front-end stages of the design cycle where the designer will write the HDL code and verify the functionality of the design and synthesize the entire HDL code by applying area, power, and delay constraints to get optimized netlist. In back-end design, the netlist, constraints, and library files

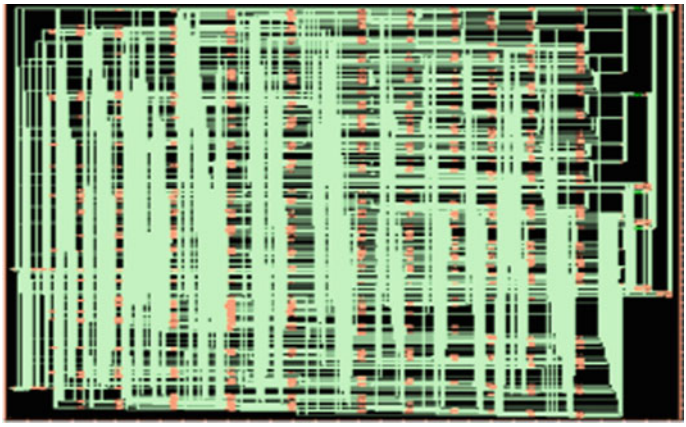
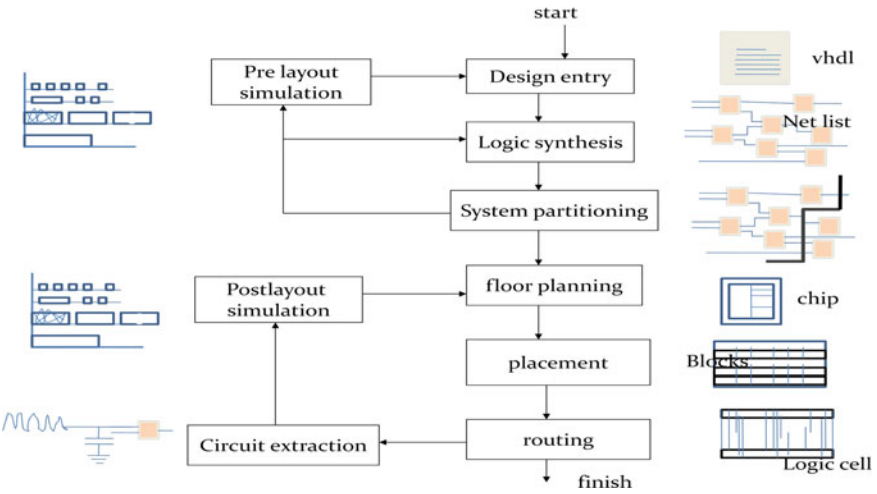


Fig. 5 SYNTHEIS (RTL schematic)

are given to the BACKEND tool which performs Floorplanning, Placement and Routing of design and finally do the RC Extraction and back annotate do the post-layout verification and covert to GDSII format. Figure 5 shows ASIC digital design flow. The entire front-end and back-end design was carried out using Cadence tools, for example, simulation using NCSIM and synthesis using RTL compiler. Floorplanning, Placement, and Routing using Cadence Encounter (Figs. 6, 7, 8).



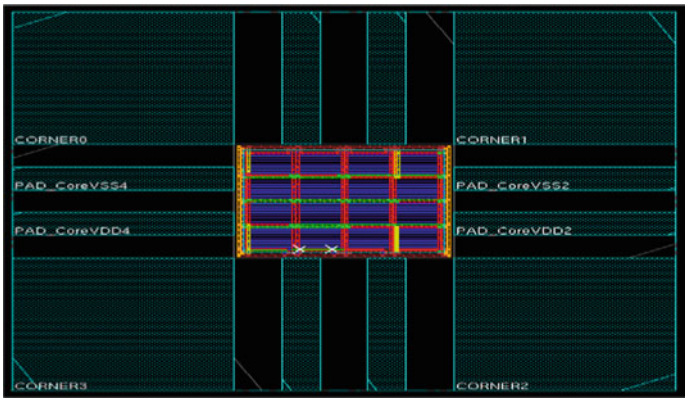


Fig. 6 Floor planning

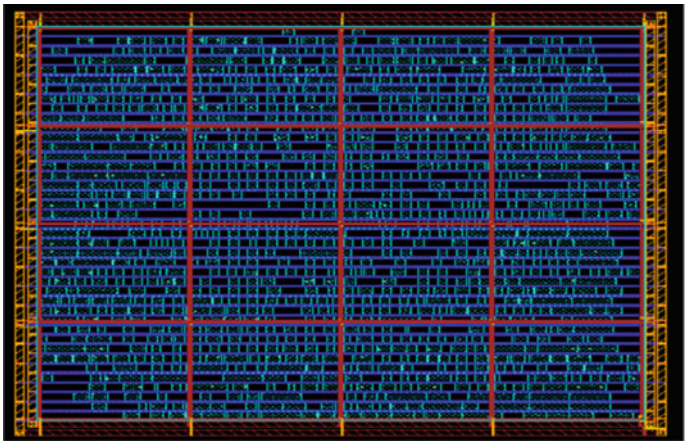


Fig. 7 Placement

6 Result and Discussion

In our design, we are built 16 × 16 bit multiplier block, verified the functionality in XC3STQ144 Xilinx kit, and compared the delay with array and booth multipliers (Table 1), and same multiplier design has been used to MAC unit, using Cadence RTL compiler synthesis of Vedic multiplier has been carried out. Standard cell–based ASIC design of the vedic multiplier is realized, and Table 2 shows the synthesis result for a specific constraint file. The gate-level netlist which generated by RTL compiler, tech files and constraints files is provided to SOC Encounter tool to get place and routed design, and finally design is converted into GDSII format.

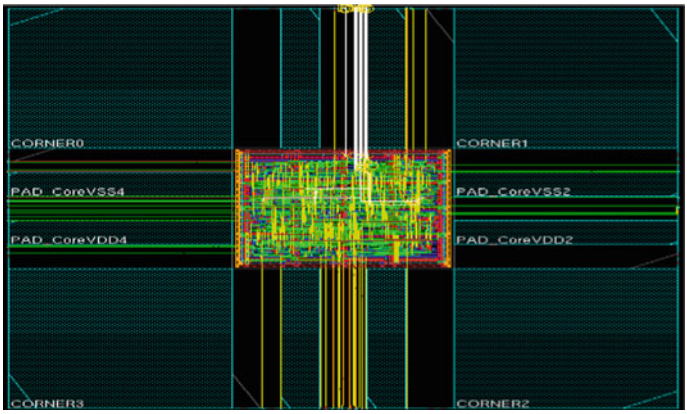


Fig. 8 Routing

Table 1 Comparison of timing delay in multipliers in FPGA

Multiplier (16 × 16)	Time delay (ns)
Array	37
Booth	43
Vedic	*24*

Table 2 ASIC synthesis result table

Parameter	Vedic multiplier (16 × 16)
Area (cells)	1,117
Power (mw)	257
Timing (ns)	5.2

The Vedic multiplier design which has combined advantages of less chip area, improved power dissipation, and timing delay can be used as the building blocks in the design of ALU, DSP such as DFT and FFT, and image processing.

7 Conclusion

In this paper, we presented the implementation of efficient Vedic multiplier which has high speed, less complexity, and consuming less area. The design was simulated, synthesized using Xilinx tool and RTL compiler, and entire back-end design was carried out using Cadence SOC Encounter.

The Vedic multiplier offered 20 and 19 % improvement in terms of propagation delay and power consumption, respectively, in comparison with parallel adder–based implementation. Whereas, the corresponding improvement in terms of delay and power was found to be 33 and 46 %, respectively, with reference to the algebraic transformation–based implementation.

References

1. Thesis by Amandeep Singh (2010) Implementation of 16 bit Vedic multiplier. Thapur University, Patiala. http://dspace.thapar.edu:8080/dspace/bitstream/10266/1109/4/1109.pdf
2. Dhillon HS, Mitra A (2008) A reduced-bit multiplication algorithm for digital arithmetics. Int J Comput Math Sci 2(2):64–69
3. Thapliyal H, Kotiyal S, Srinivas MB (2005) Design and analysis of a novel parallel square and cube architecture based on ancient Indian Vedic mathematics. In: Proceedings of IEEE centre for VLSI and embedded system technologies, International Institute of Information Technology, Hyderabad, India
4. Ramalatha M, Deena Dayalan K, Dharani P, Deborah Priya S (2009) High speed energy efficient ALU design using Vedic multiplication techniques. In Proceedings of IEEE international conference on advances in computational tools for engineering applications, 2009, ACTEA'09
5. Kumar A, Kumar D Siddhi hardware implementation of 16 $\times$ 16 bit multiplier and square using Vedic mathematics. Design Engineer, CDAC, Mohali
6. Saha P, Banerjee A, Bhattacharyya P, Dandapat A (2011) An efficient bit reduction binary multiplication algorithm using Vedic methods. In: Proceedings of IEEE students' technology symposium
7. Chang Y-N, Satyanarayana JH, Parhi KK (1997) Low-power digit-serial multiplier. In: Proceedings of IEEE international symposium on circuits and systems, June 9–12 1997, Hong Kong

Multi-Algorithm Decision-Level Fusion Using Finger-Knuckle-Print Biometric

Harbi AlMahafzah, H. S. Sheshadri and Mohammad Imran

Abstract This paper proposed the use of multi-algorithm feature-level fusion as a means to improve the performance of finger-knuckle-print (FKP) verification. LG, LPQ, PCA, and LPP have been used to extract the FKP features. Experiments are performed using the FKP database, which consists of 7,920 images. Results indicate that the multi-algorithm verification approach outperforms higher performance than using any single algorithm. The biometric performance using feature-level fusions under different normalization techniques as well has been demonstrated in this paper.

Keywords Decision-Level Fusion · Multi-Biometric · Multi-Algorithm · Normalization

1 Introduction

The need for reliable user authentication techniques has increased in the wake of heightened concerns about security and rapid advancements in networking, communication, and mobility. In the absence of robust verification schemes, these systems are vulnerable to the wiles of an impostor. Traditionally, passwords (knowledge-based security) and ID cards (token-based security) have been used to

H. AlMahafzah (✉)
P.E.T. Foundation Research, University of Mysore, Mandya, India
e-mail: hmahafzah@hotmail.com

H. S. Sheshadri
Department E & C Engg, P.E.S College of Engineering, V.T.U, Mandya, India
e-mail: hssheshadri@hotmail.com

M. Imran
Department of Study in Computer Science, University of Mysore, Mysore, India
e-mail: emraangi@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_5,

restrict access to the applications. However, security can be easily breached in these applications when a password is divulged to an unauthorized user or a badge is stolen by an impostor [7]. The emergence of biometrics has addressed the problems that plague traditional verification methods.

Biometric refers to the automatic recognition of individuals based on their physiological and behavioral characteristics. These systems are based on pattern recognition methodology, which follows the acquisition of the biometric data by building a biometric feature set, and comparing versus a pre-stored template pattern. These are unimodal which rely on the evidence of a single source of information for authentication, which have to contend with a variety of problems such as noise in sensed data, intra-class variations, and inter-class similarities [7].

There is a growing interest in resolving some of the inherent difficulties with biometric systems (intra-class variations, etc.)[2]. The techniques surrounding the use of fusing multiple biometric concepts have often been cited as the solution. The fusions of biometric refer as multi-biometrics.

Multi-biometric systems can offer substantial improvement in the matching accuracy of a biometric system depending upon the information being combined and the fusion methodology adopted [7]. *Multi-sensor*: Multiple sensors can be used to collect the same biometric. *Multi-modal*: Multiple biometric modalities can be collected from the same individual, for example fingerprint and face, which requires different sensors. *Multi-instance*: Multiple units of the same biometric are collected, for example fingerprints from two or more fingers. *Multi-sample*: Multiple readings of the same biometric are collected during the enrollment and/or recognition phases; for example, a number of fingerprint readings are taken from the same finger [7]. *Multi-algorithms*: In some cases, invoking multiple feature extraction and/or matching algorithms on the same biometric data can result in improving matching performance [3]. Multi-algorithm systems consolidate the output of multiple feature extraction algorithms, or that of multiple matchers operating on the same feature set.

Thus, this paper evaluates the performance of multi-algorithm approach by fusing the data acquired from different algorithms at decision–score levels. The rest of the paper is organized as follows: Section 2 presents related works, proposed method is given in Sect. 3, detailed experimental results are given in Sect. 4, and conclusion is mentioned in Sect. 5.

2 Related Works

Markus Turtinen et al. [4] have considered of combining local binary pattern (LBP) texture features with a self-organizing map (SOM) classification. Very good performance was obtained in two different classification problems. Julian Fierrez-Aguilar et al. [5] have proposed scheme on the performance of ridge information and minutiae based for fingerprint verification with the effects of image quality. Seyed Mehdi Lajevardi and Zahir M. Hussain [6] have proposed different

approaches to the facial expression recognition task. The logarithmic Gabor filters and the local binary pattern (LBP) operator were used for feature extraction. The results showed that the log-Gabor method outperformed the LBP method producing the largest improvement in the classification accuracy.

Ville Ojansivu and Janne Heikkilä [1] have proposed a new blur insensitive texture classification method, which is based on quantized phase of the discrete Fourier transform (DFT) computed in local image windows, and it is called local phase quantization (LPQ).

Hussian Dawood et al. [8] have proposed a new hybrid approach for texture classification, which uses the joint histogram of the two complementary features, local phase quantization (LPQ), and the variance of the image in spatial domain. Xiaoyang Tan and Bill Triggs [9] have investigated the benefits of combining Gabor wavelets and LBP features. The authors have shown that by combining the two local appearance, it has given considerably better performance than either alone.

3 Proposed Method

In this paper, a hand-based biometric technique, finger-knuckle-print (FKP) has been used, and FKP refers to the image pattern of the outer surface around the phalangeal joint of one's finger, which is formed by bending slightly the finger-knuckle [10]. The experiments are developed for personal authentication using DZhang FKP database. FKP images were collected from 165 volunteers, including 125 males and 40 females.

3.1 Preprocessing

This section describes the region of interest (ROI) extraction; the process involved to extract ROI for each instance is as follows. It is necessary and critical to align FKP images by adaptively constructing a local coordinate system for each image. With such a coordinate system, an ROI can be cropped from the original image using the following steps suggested in [10], as shown in Fig. 1.

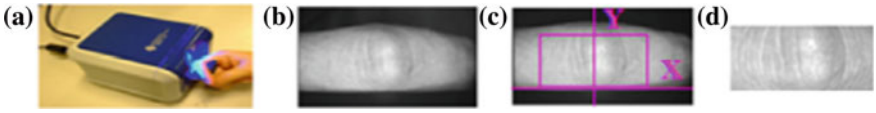


Fig. 1 **a)** Image acquisition device is being used to collect FKP samples. **b)**Sample FKP image. **c)**ROI coordinate system, where the rectangle indicates the ROI area. **d)**Extracted ROI

4 Feature Extraction

As this paper concern about multi-algorithm, the following feature extraction algorithms have been used to extract the features prior to fuse different algorithm combinations.

4.1 Log-Gabor Filters

Log-Gabor filter was proposed by Field [11]. Field suggests that natural images are better coded by filters that have Gaussian transfer functions when viewed on the logarithmic frequency scale. On the linear frequency scale, the log-Gabor filter has a transfer function of the form

$$G(w) = e^{(-log(w/w_0)^2)/(2(log(k/w_0)^2)}$$

where w_0 is the filter's center frequency. To obtain constant shape ratio filters, the term k/w_0 must also be held constant for varying w_0.

4.2 Local Phase Quantization

The local phase quantization (LPQ) method introduced by Ojansivu et al. [1] is based on the blur invariance property of the Fourier phase spectrum. It uses the local phase information extracted using the 2-D DFT computed over a *M*-by-*M* rectangular with neighborhood N_x at each pixel position x of the image *f(x)* defined by

$$F(u,x) = \sum_{y \in N_x} f(x-y) e^{-j2\pi u^T y} = w_u^T f_x$$

where w_u is the basis vector of the 2-D DFT at frequency u, and f_x is another vector containing all M^2 image samples from N_x [1].

4.3 Principal Component Analysis

Principal component analysis (PCA) was invented in 1901 by Karl Pearson. PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

4.4 Locality Preserving Projections

Locality preserving projection (LPP) method introduced by Xiaofei He [12] is a linear projective map that arises by solving a variation problem that optimally preserves the neighborhood structure of the data set.

5 Biometrics Fusion Strategies

In general, a biometric system works in two modes: enrollment and authentication. Verification and identification are the two modes an authentication can be carried out. Fusion of biometric systems, algorithms, and/or traits is a well-known solution to improve authentication performance of biometric systems. Researchers have shown that multi-biometrics, that is, fusion of multiple biometric evidences, enhance the recognition performance.

In biometric systems, fusion can be performs at different levels: *Sensor-Level, Feature-Level, Score-Level, and Decision-Level Fusion* [8].

5.1 Levels of Fusion

- **Sensor-Level Fusion** It entails the consolidation of evidence presented by multiple sources of raw data before they are subjected to feature extraction. Sensor-level fusion can benefit multi-sample systems which capture multiple snapshots of the same biometric.
- **Feature-Level Fusion** In feature-level fusion, the feature sets originating from multiple biometric algorithms are consolidated into a single feature set by the application of appropriate feature normalization, transformation, and reduction schemes [3][8].
- **Score-Level Fusion** The match scores output by multiple biometric matchers are combined to generate a new match score (a scalar).
- **Decision-Level Fusion** Fusion is carried out at the abstract or decision level when only final decisions are available (AND, OR, Majority Voting, etc.).

In all the experiments, the data have been fused at decision level, using "AND" and "OR" rules for two and three algorithm combinations.

6 Results and Discussion

This section deals with the investigation consequences of combining different biometrics feature extraction algorithm at decision-level fusion with "**AND**" **and** "**OR**" rules to measure the performance of multi-algorithms system.

Table 1 Performance analysis of single feature extraction algorithm

Finger-knuckle-print				
FAR (%)	GAR (%)			
	LG	LPQ	PCA	LPP
0.01	70.00	59.66	51.00	38.00
0.10	76.33	71.67	61.00	47.67
1.00	86.33	83.00	79.00	58.66

"**AND**" **and** "**OR**" **Rules** [13][14]: In a multi-biometric verification system, the simplest method of combining decisions output by the different matchers is to use the "AND" and "OR" rules. The output of the "AND" rule is a "match" only when all the biometric matchers agree that the input sample matches with the template. On the contrary, the "OR" rule outputs a "match" decision as long as at least one matcher decides that the input sample matches with the template.

In all the experiments, performance is measured in terms of false acceptance rate (FAR in %) and corresponding genuine acceptance rate (GAR in %). First, the performance of a single algorithm system is measured; later, the results for multi-algorithms biometric system are evaluated. The results obtained from single algorithm biometric system are tabulated in Table 1 and are depicted as receiver operating characteristic (ROC) curve in Fig. 2.

From Table 1, it can be observed that the log-Gabor has higher accuracy over the other algorithms with higher FAR (i.e. FAR = 1), and at the same time, it could be seen that LPQ is outperformed LG for lower values of FAR (i.e., FAR = 0.1 and 0.01) which will be better choice for high security applications. Further, the fusion performance results of two and three algorithms under the score-level fusion are shown. Table 2 shows the decision fusion of two algorithms under "OR" and "AND" rules. From Table 2, one can observe that the fusion of high-performance algorithms has no significant improvement over single algorithm performance; for example, the fusion of LG and LPQ has given a GAR performance under "OR" and "AND" rules is the same as LG itself.

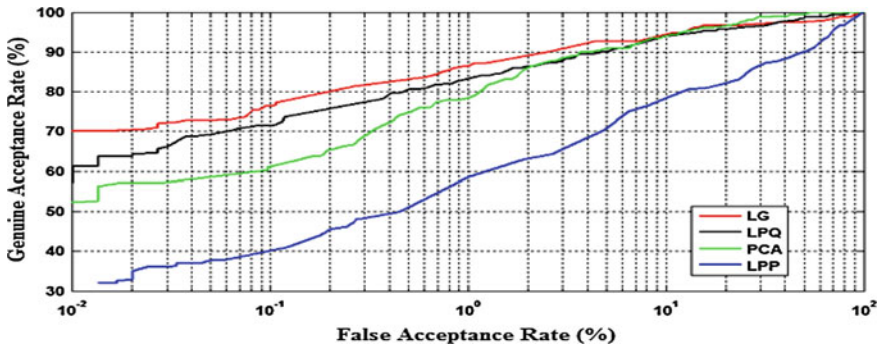


Fig. 2 ROC curve of performance of single algorithm

Table 2 Performance analysis of two feature extraction algorithms at decision-level fusion under "AND" and "OR" rules

FKP decision-level fusion two algorithms						
FAR (%)	GAR (%) With OR rule					
	LG + LPQ	LG + PCA	LG +LPP	LPQ + PCA	LPQ + LPP	PCA + LPP
0.01	66.33	56.00	76.67	56.67	65.33	55.67
0.10	79.67	67.33	83.33	67.67	79.67	67.33
1.00	88.33	79.00	89.67	80.00	83.67	79.33
	GAR (%) with AND rule					
0.01	70.00	70.33	37.00	65.00	37.00	32.00
0.10	80.67	77.33	45.00	73.33	45.00	41.00
1.00	86.33	87.00	59.00	85.67	59.00	58.67

Table 3 Performance analysis of three feature extraction algorithms at decision-level fusion under "AND" and "OR" rules

FKP decision-level fusion three algorithms				
FAR (%)	GAR (%) with OR rule			
	LG + LPQ + PCA	LG + LPQ + LPP	LG + PCA + LPP	LPQ + PCA + LPP
0.01	56.33	66.33	56.00	54.33
0.10	61.67	75.67	61.33	61.67
1.00	80.00	86.33	79.00	80.00
	GAR (%) with AND rule			
0.01	71.00	37.67	37.00	37.00
0.10	80.67	45.67	45.50	41.00
1.00	86.33	58.67	58.67	58.67

For the fusion of high- and low-performance algorithms (LG and LPP), the fusion has a nuance improvement in GAR. For the fusion of low-performance algorithms (PCA and LPP), the fusion result under OR rule is the same as PCA performance and under the "AND" rule is the same as LPP performance. The same thing could be observed for fusion of three algorithms by looking to Table 3. Going beyond three algorithms, that is, fusing of four algorithms has drastically lower performance than single algorithm. The ROC curves fusions of two and three algorithms are portrayed in Figs. 3 and 4, respectively.

The above shown result of fusion is due to the limitation of these two rules as their tendency to result in extreme operating points. When the "AND" rule is applied, the false accept rate (FAR) of the multi-biometric system is extremely low (lower than the FAR of the individual matchers) which produce more error. Similarly, the "OR" rule leads to higher FAR and lower FRR than the individual matchers which also produce more error. Either method of combining biometrics produces many times more errors than if the stronger of the two matchers had been used alone. When one biometric matcher has a substantially higher equal error rate compared to the other matcher, the combination of the two matchers using "AND"

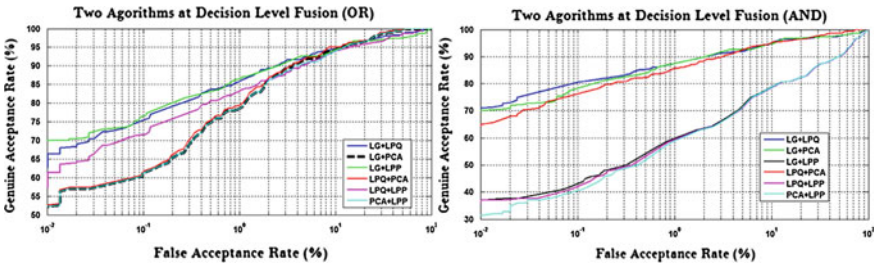


Fig. 3 ROC curve at decision-level fusion combination of two algorithms under "OR" and "AND" rules

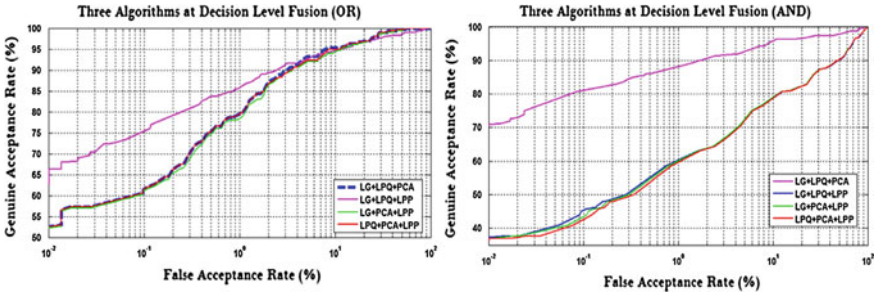


Fig. 4 ROC curve at decision-level fusion combination of three algorithms under "OR" and "AND" rules

and "OR" rules may actually degrade the overall performance [14]. Due to this phenomenon, the "AND" and "OR" rules are rarely used in practical multi-biometric systems.

7 Conclusion

From the analysis of experimental results and observations, it can be concluded that a multi-algorithms biometric fusion at decision-level fusion under "AND" and "OR" rules has no performance improvement over a single algorithm even it could be worst. This is due to the limitation of "AND" and "OR" rules as their tendency to result in extreme operating points. When the "AND" rule is applied, the false accept rate (FAR) of the multi-biometric system is extremely low (lower than the FAR of the individual matchers). Similarly, the "OR" rule leads to higher FAR and lower FRR than the individual matchers. When one biometric matcher has a substantially higher equal error rate compared to the other matcher, the combination of the two matchers using "AND" and "OR" rules may actually degrade the overall performance.

References

1. Ville Ojansivu and Janne Heikkilä. "Blur Insensitive Texture Classification Using Local Phase Quantization". *Proceedings of the 3rd International Conference on Image and Signal Processing (ICISP) 2008, LNCS 5099, pp. 236–243 Springer-Verlag Berlin, Heidelberg 2008*
2. Abhilash Bhargav-Spantzel, Anna C. Squicciarini, Shimon Modi, Matthew Young, Elisa Bertino, and Stephen J. Elliott. "Privacy Preserving Multi-Factor Authentication with Biometrics". *journal of computer security (1875-8924) volume 15, Number 5/2007*
3. J. Stan Z. Li, Anil K. Jain. "Encyclopedia of Biometrics". *Springer*
4. Markus Turtinen, Topi Mäenpää, and Matti Pietikäinen. "Texture Classification by Combining Local Binary Pattern Features and a Self-Organizing Map". *Proceeding SCIA'03 Proceedings of the 13th Scandinavian conference on Image analysis Springer-Verlag, Berlin, Heidelberg 2003*
5. Julian Fierrez-Aguilar, Yi Chen, Javier Ortega-Garcia, and Anil K. Jain. "Incorporating Image Quality in Multi-Algorithm Fingerprint verification". *ICB'06 Proceedings of the 2006 international conference on Advances in Biometrics*
6. Seyed Mehdi Lajevardi, Zahir M. Hussain. "Facial Expression Recognition Using Log-Gabor Filters and Local Binary Pattern Operators". *International Conference on Communication, Computer and Power (ICCCP'09) Muscat, February 15-18, 2009*
7. Harbi AlMahafzah, Mohammad Imran, and H.S. Sheshadri. "Multibiometric: Feature Level Fusion Using FKP Multi-Instance biometric". *IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012*
8. Hussian Dawood, Hassan Dawood and Ping GUO. "Combining the Contrast Information with LPQ for Texture Classification". *Science of Electronics, Technologies of Information and Telecommunications (SETIT) Sousse 21-24 March 2012-Tunisia*
9. Xiaoyang Tan and Bill Triggs. "Fusing Gabor and LBP Feature Sets for Kernel-based Face Recognition". *3rd International Workshop Analysis and Modeling of Faces and Gestures (AMFG '07) 4778 (2007) 235–249*
10. Zhang Lin, Zhang Lei, Zhang David, Zhu Hailong (2011) Ensemble of local and global information for finger–knuckle-print recognition. Elseveir/Pattern Recognition 44:1990–1998
11. D.J. Field. "Relation between the statistics of natural images and the response properties of cortical cells". *J. Opt. Soc. Am. A, 4(12):2379_2394, 1987*
12. Xiaofei He Partha Niyogi "Locality Preserving Projections (LPP)". Advances in Neural Information Processing Systems 16 (NIPS), Vancouver, Canada, 2003
13. A. Ross, K.Nandakumar, and A.K. Jain. "Handbook of multibiometrics". *Springer-Verlag edition, 2006*
14. John Daugman. "Biometric decision landscapes". *Technical Report UCAM-CL-TR-482 ISSN 1476-2986 Number 482 January 2000*

Encryption of Compressed Video Using Efficient RSA-CRT Algorithm

M. Sujatha, Viviet Carol Noronha, Adeline Teena Monteiro, Falon Alice Mascarenhas and Merline Carol Cutinho

Abstract Communication of video streams can be achieved by applying some conventional encryption and compression technique. Encryption helps in secure transmission, and compression reduces the size of a video frames and helps to transfer the information more efficiently. We propose to encrypt compressed video using efficient RSA-CRT algorithm which is more secure and time efficient.

Keywords XOR · MSU · RSA · CRT

1 Introduction

The presence of computer networks has prompted new problems with security and privacy. Having a secure and reliable means for communicating with images and video is becoming a necessity, and its related issues must be carefully considered. Image and video encryption have applications in various fields including Internet communication, multimedia systems, medical imaging, telemedicine and military communication. Encryption is the most effective way to achieve data security. It is

M. Sujatha (✉) · V. C. Noronha · A. T. Monteiro · F. A. Mascarenhas · M. C. Cutinho
Department of Computer Science and Engineering, St. Joseph Engineering College, Vamanjoor, Mangalore, Karnataka 575028, India
e-mail: sujatha_msk@yahoo.co.in

V. C. Noronha
e-mail: viviet14@gmail.com

A. T. Monteiro
e-mail: adi_mon11@yahoo.com

F. A. Mascarenhas
e-mail: falon.firestar@gmail.com

M. C. Cutinho
e-mail: merlinecutinho@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_6, © Springer India 2014

found that the safer encryption methods are not faster enough and if encryption is fast it is not very secure. Hence, we propose efficient RSA-CRT algorithm applied on compressed video frames to guarantee both security and speed.

1.1 Video Encryption

A video stream is a sequence of video frames each of which is a still image. Frames are digitized in a standard RGB format, 24 bits for pixel (8 bits each for Red, Green, and Blue) [1]. Multimedia data are widely used in various kinds of content provide services and information exchange applications. In these services and applications, digital video is transmitted from service provider to end user or exchanged between end users over public communication channels such as satellite, wireless networks, and the Internet. As these public channels are vulnerable to the attack from hackers, video security becomes more important. XOR, DES, AES, RSA are some of the algorithms which can be used for video encryption.

1.2 Video Compression

JPEG is a standardized compression method for full-color and grayscale images. It supports a maximum image size of 65,535 × 65,535. JPEG typically achieves 10:1 compression with little perceptible loss in image quality. JPEG makes image files smaller, and it stores 24-bit per pixel color data instead of 8-bit per pixel data [2].

2 Related Work

The paper [3] is based on RSA encryption algorithm, combining features of digital image, designing and implementing digital image based on RSA encryption algorithm. It is less demanding in terms of time consumption, but also for encryption of color RGB images. After analyzing the [4], it is noted that with CRT it is possible to reconstruct integers in a certain range from their residues modulo a set of pair wise relatively prime modulo. The Singh and Agarwal [4] has used random numbers which are unique for a session of message transmission. At the receiver side, these random numbers are checked and if the number has been received earlier also it means that the packet is replayed and discarded. The paper [5] compresses the encrypted the data. In [6], encryption of video is done using XOR and compression using MSU lossless algorithm. Using motion-compensated lossless video encoder, about 57 % of compression is needed for each unencrypted frame, but for encrypted frames, it is enough if we compress only by 33 %.

3 Proposed Algorithm

The proposed method uses JPEG to compress the video and efficient RSA-CRT algorithm to encrypt the video. The algorithm is as follows:

Step 1: Extract frames from the video
Step 2: Obtain the plaintext.
Step 3: Compress the plain text using JPEG compression algorithm.
Step 4: Use efficient RSA-CRT to encrypt the compressed plaintext and obtain cipher text.
Step 5: If not end of video go to step 2.
Step 6: Combine the cipher frames to obtain cipher video.
Step 7: Send the cipher video through a secure channel.
Step 8: Extract the frames from cipher video.
Step 9: Decrypt the frames using efficient RSA.
Step 10: Decompress the obtained frames using JPEG decompression algorithm.
Step 11: Combine all the frames to obtain original video.

3.1 Efficient RSA-CRT Algorithm

3.1.1 Key Generation

Step 1: Select four prime numbers p, q, r, and s.
Step 2: Compute z = (p−1)*(q−1)*(r−1)*(s−1) where z is Eulers function.
Step 3: Compute g = (h,z) where h is the rank of the matrix and assume p, q, r, s as distinct, sqrt(z)/h < g < sqrt(z)
Step 4: Choose a random integer e where gcd(e,g) = 1.
Step 5: Compute the inverse d where ed = 1 mod g.
Step 6: Determine the receiver's public and private key.
d is the private key and the e is the public key.

3.1.2 Encryption Module

Step 1: Compute ciphertext,

$C_1 = M_1^e \bmod p.$
$C_2 = M_2^e \bmod q.$
$C_3 = M_3^e \bmod r.$
$C_4 = M_4^e \bmod s.$

where M_1, M_2, M_3, and M_4 are four different parts of the frame.

Step 2: Combine C1, C2, C3, and C4 as C to get single cipher frame.
Step 3: Send the encrypted frame C to the receiver.

3.1.3 Decryption Module

Step 1: Obtain the cipher frame from the sender.
Step 2: Compute the plaintext using

$M_1 = C_1^d \bmod p.$
$M_2 = C_2^d \bmod q.$
$M_3 = C_3^d \bmod r.$
$M_4 = C_4^d \bmod s.$

where C_1, C_2, C_3, and C_4 are different parts of the cipher frame.

Step 3: Combine M_1, M_2, M_3, and M_4 as M to get original frame.

4 Experimental Results

We mainly focused on efficient RSA-CRT algorithm applied on compressed video frames. To check the efficiency of RSA-CRT, we first considered the RSA-CRT algorithm applied on images. We have also compared the results with XOR algorithm and RSA algorithm.

4.1 Experimental Study of XOR, RSA, and RSA-CRT Algorithm

4.1.1 Performance Study

If encryption and decryption is performed without compression, more time is consumed for execution and with compression less time is consumed (Tables 1, 2).

The results show that efficient RSA-CRT takes less time when compared to RSA (Table 3).

Experimental results show that efficient RSA-CRT takes less time to encrypt/decrypt each frame than RSA and XOR.

4.1.2 PSNR Ratio Comparison

A graph of different values of n is evaluated, and the PSNR value for the image is computed (Figs. 1, 2 and 3).

The PSNR ratio of XOR algorithm and the PSNR ratio of RSA algorithm are more than the PSNR ratio of efficient RSA-CRT algorithm.

Table 1 Time taken for encryption and decryption using RSA

Image	Size (bytes)	Time to encrypt/decrypt uncompressed data (s)	Time to encrypt/decrypt compressed data (s)
Vegas.jpg	90,160	23	1
Xylophone.jpg (1 frame)	230,454	52	2
Xylophone.jpg (5 frame)	1,152,270	147	10

Table 2 Time taken for encryption and decryption using RSA-CRT

Image	Size (bytes)	Time to encrypt/decrypt uncompressed data (s)	Time to encrypt/decrypt compressed data (s)
Vegas.jpg	90,160	1	1
Xylophone.jpg (1 frame)	230,454	3	1
Xylophone.jpg (5 frame)	1,152,270	20	3

Table 3 Time for XOR, RSA, and efficient RSA-CRT

Frames	XOR	RSA	RSA-CRT
1	5	2	1
5	11	10	3
20	40	39	11

Key value (n)	PSNR
13	8.795
17	8.812
51	9.126
101	11.093
137	13.415

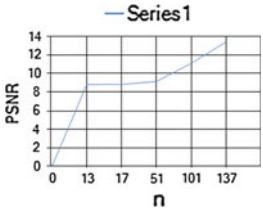


Fig. 1 PSNR ratio for XOR image encryption

Key value (p,q)	PSNR
7,5	7.273
11,13	11.709
13,17	15.352
17,19	19.255

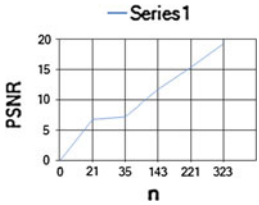


Fig. 2 PSNR ratio for RSA image encryption

Key value (p,q,r,s)	PSNR
31,53,61,73	7.988
71,73,83,101	8.918
101,103,107,97	9.542
101,113,137,179	9.805
163,163,163,163	9.805

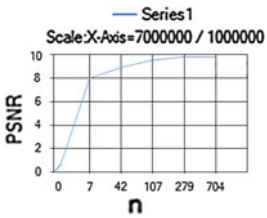


Fig. 3 PSNR ratio for efficient RSA-CRT image encryption

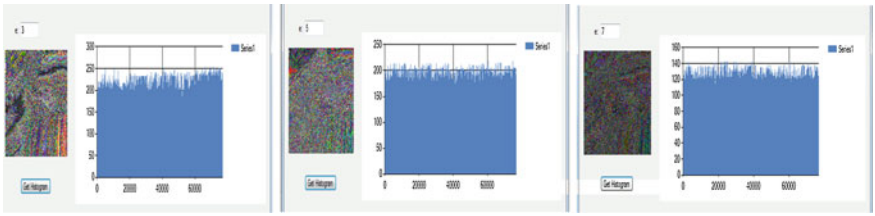


Fig. 4 Sensitivity analysis for encrypted image using RSA when e = 3, 5, and 7

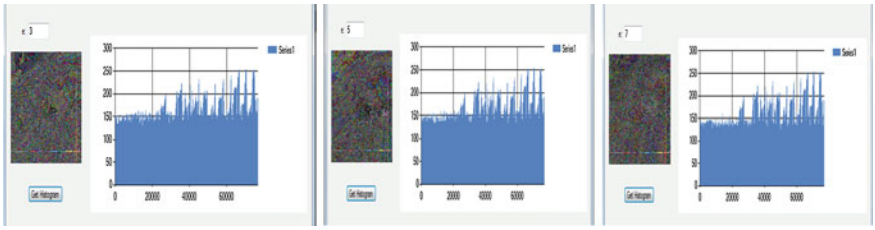


Fig. 5 Sensitivity analysis for encrypted image using RSA-CRT when e = 3, 5, and 7

4.1.3 Sensitivity Analysis

The histogram analysis shows that, RSA-CRT algorithm is more secure than RSA (Figs. 4, 5).

5 Conclusion

It is found that the efficient RSA-CRT encryption algorithm can be effectively used to encrypt video frames. The encryption delay overhead using efficient RSA-CRT is less than the overhead using RSA. If the encryption is done before compressing then the size of cipher text will be more. If the compression is performed prior to encryption, then the size of cipher text will be reduced. Also, the PSNR obtained

using RSA-CRT is less compared to the existing RSA algorithm. Hence, using JPEG for compressing video frames and efficient RSA-CRT for encrypting video frames is a more feasible solution to secure real-time video transmissions in optimum time.

References

1. Zhao G, Yang X, Zhou B (2003) RSA-based digital image encryption algorithm in wireless sensor networks. IEEE lhm Info Theory IT-36:1111–1126, IT-37:995–1005
2. Weinberger MJ, Seroussi G, Sapiro G (2000) The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS. IEEE Trans Image Process 9(8):1309–1324
3. Feder M, Shulman N (2002) Source broadcasting with unknown amount of receiver side information. IEEE lhns Info Theory IT-38:1534–1547
4. Singh S, Agarwal G (2010) Use of Chinese remainder theorem to generate random numbers for cryptography. Int J Appl Eng Res Dindigul 1(1): 115–123
5. Johnson M, Ishwar P, Prabhakaran VM, Schonberg D, Ramchandran K (2004) On com pressing the encrypted data. IEEE Trans Signal Process 52:2992–3006
6. Schonberg D, Draper S, Ramchandran K (1999) On compression of encrypted images. IEEE Trans Inform Theory 21: 268–272

Design of High-Speed Reconfigurable Coprocessor for Next-Generation Communication Platform

S. Mallikarjunaswamy, K. R. Nataraj and K. R. Rekha

Abstract In this paper, we present the high-speed reconfigurable coprocessor for the next-generation communication system, which is highly reliable and more accurate with less delay. The proposed high-speed reconfigurable coprocessor can be easily implemented using various standard operations such as bit shuffle operation, convolutional encoding, fast Fourier transform, interleaving, modulation, scrambling, shift-XOR array, Viterbi decoding, and several other function using the proposed design. The coprocessor has been modeled using VHDL, and synthesis has been performed on model-sim. The gate count is of about 34,000 and critical path of about 0:16 μm technology. The performance comparisons shows that the number of clock cycles can be reduced about 48 % for scrambling and 84 % for convolutional encoding compared with existing DSPs. From the results, the performance of the proposed coprocessor is better compared to conventional DSP (SC140) in terms of number of clocks per cycle.

Keywords Bit shuffle unit structure · Convolutional encoding · Fast Fourier transform · Interleaving · Modulation · Scrambling · Shift-XOR array · Viterbi decoding

S. Mallikarjunaswamy (✉) · K. R. Nataraj · K. R. Rekha
SJB Institute of Technology, Visvesvaraya Technological University, Belgaum, Bangalore, India
e-mail: pruthvi.malli@gmail.com

K. R. Nataraj
e-mail: nataraj.sjbit@gmail.com

K. R. Rekha
e-mail: rekha.sjbit@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_7,

1 Introduction

The recent communication system uses more complicated standards for communication services, which has resulted in number of competing and incompatible standards. There are some standards like frequency-hopping spread spectrum (FHSS) [1, 2], direct-sequence spread spectrum (DSSS) [3] digital audio broadcast (DAB) [4] orthogonal frequency-division multiplexing (OFDM), OSHA's Hazard communication standard (HCS) [5], carrier sense multiple access (CSMA), digital video broadcast (DVB), asymmetric digital subscriber line plus (ADSL++) used in flexible communication system that supports multimedia broadcast and multiband programmable processor. The several other functions performing similar operation under various standards are having different characteristics according to the standards. As a result, some of the flexible processors are developed, producing a solution that is more attractive than the above-said existing methods. This paper proposes a high-speed reconfigurable coprocessor for the next-generation communication system, This system is highly reliable, more accurate, and also has less delay with better performance than existing system. This design has been modeled using VHDL code, model-sim and evaluated performance and comparisons with existing DSP chip. This paper is organized as follows.

2 Reconfigurable Structures for Communication Systems

The operations of communication systems can be classified into three basic operations:

- Multiply, add, and subtract operations.
- Shift and logical operations.
- Bit manipulation operations.

The pervious papers have discussed on design of high-speed reconfigurable coprocessor for multiplier/adder and subtractions operations [6]. Now, this section describes a reconfigurable structure, which performs basic shift operation, logical operations, and bit shuffle unit operations. It also shows comparison of existing shift operation, logical operations, and bit shuffle unit operations. The proposed shift operation, logical operations, and bit shuffle unit operations are as follows.

- Bit shuffle unit structure

Bit shuffle operation includes clearing bit settings, rearranging. The operations are performed by bit AND, bit XOR, bit OR, shift operation, logical operation and so on. Communication system built by scrambler, convolutional encoding, puncturing, interleaving, Viterbi decoding, bit stream multiplexing, and de-scrambling uses bit shuffle unit.

- Puncturing

Figure 1 shows the puncturing operations are performed from ordinary codes by dropping some output bits (i.e., the codes are not sending though a channel). Due to this, the code has higher bit rate (less redundancy) but lower error-correcting capability. The puncturing matrix represents the order of dropping some output bits. The number of output polynomials is equal to the number of rows in a puncturing matrix. The matrix consists 1's and 0's, and the number of 1's in the matrix represents that given bit is transmitted and number of 0's in the matrix is represents the given bit is dropped. The out stream was cyclically applied by puncturing matrix [7]

Let us consider a puncturing matrix for the rate 2/3, which is shown above. It indicates that every second output bit will be transmitted from the first polynomial, but in second polynomial, every bit will be transmitted. Now we describe the reconfiguration coprocessor bit-shuffling unit which can support various operations, it includes bit-shuffling for various standards.

- Existing DSP architectures for bit shuffle unit

Figure 2 details the general structure unit of DSP architecture. It consists of a shifter, logical, and arithmetic units. The shifter and logical units perform repetitive shifting and XOR operation. First, the operation is performed by reading data from the register. Secondly, the shifter shifts the read data, and finally, the XOR operation is performed on shifted data by the logical unit. However, all general DSPs does not perform parallel shift and XOR operation on multiple data. In addition to previous operation, general DSP performs a shift left followed by a shift right; this operation helps in bit extraction operation to extract the file in source. The bit insertion operation can also be performed using AND, shift, or OR operations.

Star-Core SC140 shows in Fig. 3 is a commercial DSP; it supports bit insertion and extraction operation. The extracted instruction is used to extract a bit from a source data register and also to send it to a destination data register, which is zero-extended and right-aligned. The insert instruction used will insert a bit from a source data register into a destination data register. The bits present outside the inserted field in the destination register are unchanged [8].

Fig. 1 The puncturing operations

The standard puncturing matrices for the rate ½ convolutional code are:

RATE	PUNCTURING MATRIX
2/3	$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$
3/4	$P = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$
5/6	$P = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$
7/8	$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

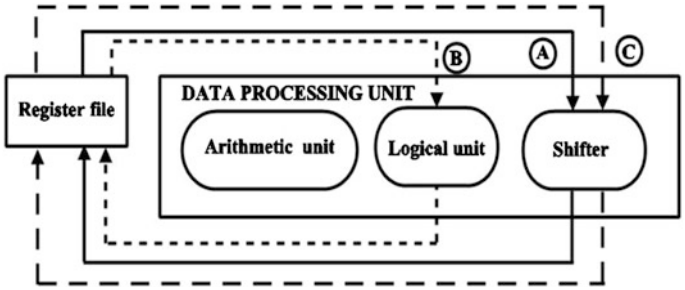


Fig. 2 The general structure unit of DSP architecture

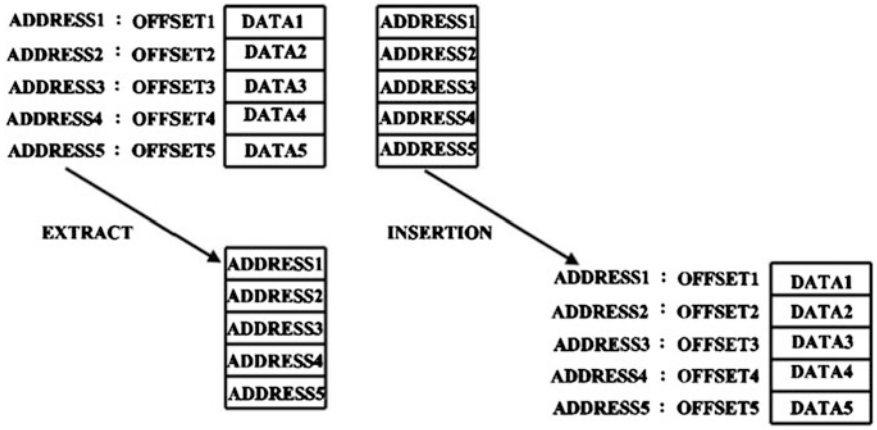


Fig. 3 Star-cores SC140

Figure 4 shows TMS320C6x, Texas Instruments (TI) which supports operations like shuffling and de-shuffling of two input words and the extraction operation takes place depending on the offset [9]. Figure 5 shows TMS320C55x, and it supports expand and extract instructions [10].

- Proposed Bit Shuffle Unit Instructions and Their Hardware Structure

This section details the three instructions required for the bit-shuffling unit operations and their hardware architecture. The proposed instructions include CEN for convolutional encoding, PUNC for puncturing, and SCR for scrambling, multiplexing, and interleaving. Figure 6 shows the proposed bit-shuffling unit which includes a bit extraction/insertion logical operation and shift-XOR array and a bit-storage register. Mask1(Y-bits) and Mask2(X-bits) signals are two masks used to control the bit shuffle unit. The number of bits of Mask1(Y-bits) and that of Mask2(X-bits) can be arbitrarily chosen. Figure 7 shows the proposed shift-XOR operation [11].

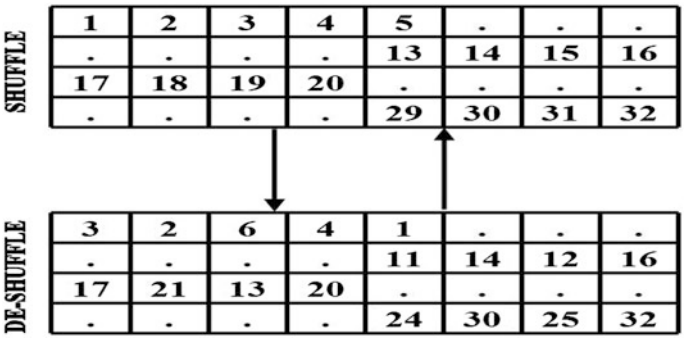


Fig. 4 TMS320C6x

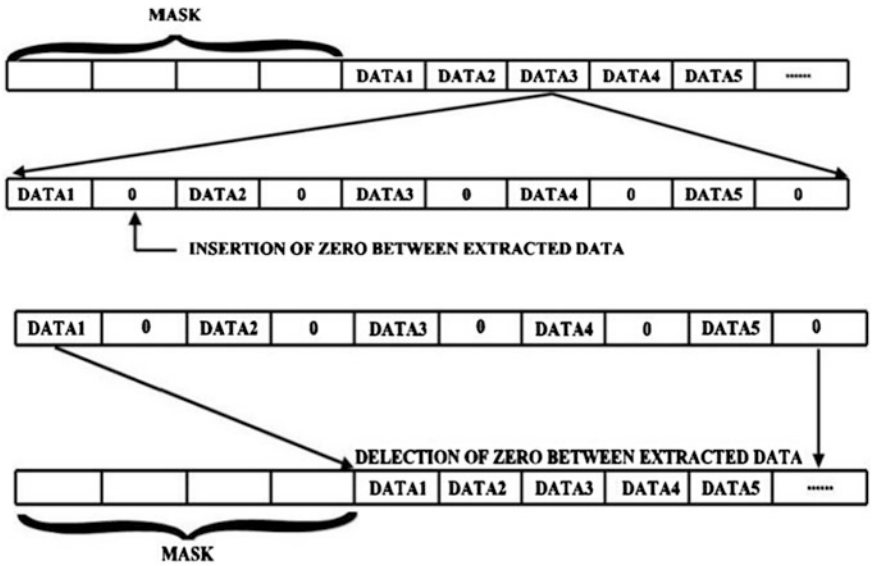


Fig. 5 TMS320C55x

First, it receives X + Y as input data bits and Mask1 and generates the Y-shifted data that are shifted by 1 through Y-bits. Next, it performs parallel XOR operations of the input data and the Y-shifted data selected by Mask1(Y-bits). Consider Mth bit of Mask1 (Y-bits), set to "1", the M-bit-shifted data are XORed with the input data. Hence, the Y output data are generated and transferred to the switching unit. The switching network stores all or some of the Y output data on the registers according to Mask1 (Y-bits). Mask1 (Y-bits) is the selection signal that enables the registers to store the only valid outputs among the Y output data are transferred to the switching network [12].

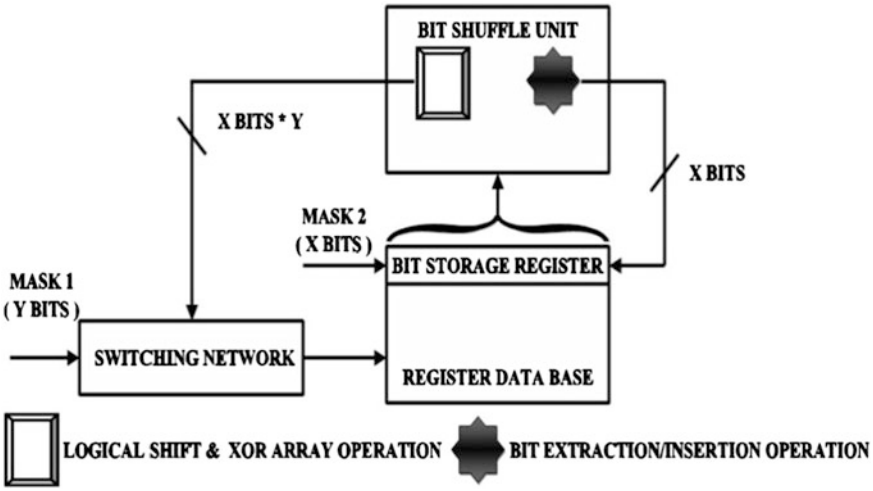


Fig. 6 Proposed bit shuffle unit

- Design of high-speed reconfigurable coprocessor for next-generation communication system

Figure 8 shows the proposed high-speed reconfigurable coprocessor for next-generation communication system that consists of a memory interface for address generation, a data processing unit, data buffering, a configuration manager for instruction decoding, temporary register, configuration registers for storing data path information, generation of configuration information, and configuration signal generator. The operation of proposed reconfigurable coprocessor communication system explained is as follows. First, a microprocessor-encoded configuration instruction sends to the proposed coprocessor configuration manager. The configuration memory contains an encoded instruction by help this proposed coprocessor loads the configuration information according to the encoded instruction. The configuration information contains information about an interconnection between function units and memory access width. The configuration manager saves all information in the configuration registers and controls all function units. The proposed reconfiguration coprocessor for communication system can support various communication system standards.

3 Implementation Results and Comparisons

The proposed design of high-speed reconfigurable coprocessor for next-generation communication system structure has been modeled by VHDL, and synthesis has been performed using the model-sim. The proposed structures gate count is about 17,000 for addition, subtraction, and multiplier units. The proposed bit shuffle unit

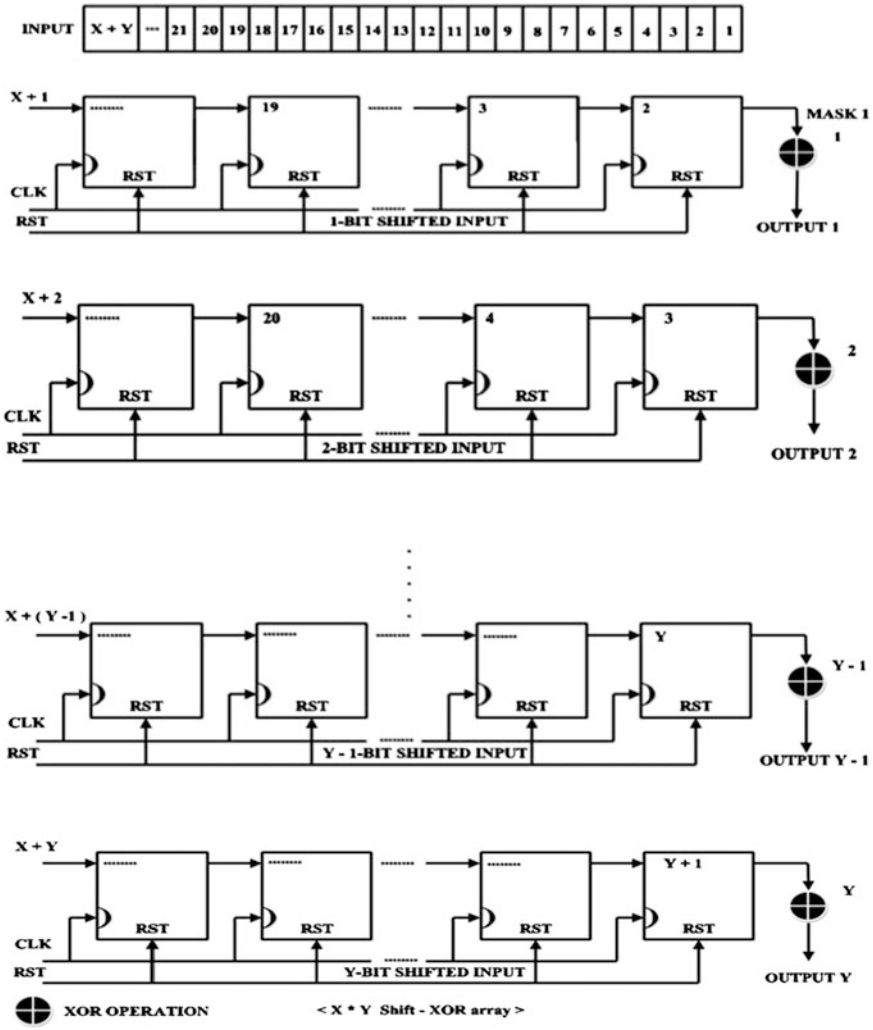


Fig. 7 Shift-XOR array architecture

structure gate count is about 1,500 (N = 6, M = 12). Figure shows the synthesis result of the proposed structures. The critical path of the structures is 3.64 ns with the 0.16 μm technology. Figure 9a shows the realization of top level of proposed design of high-speed reconfigurable coprocessor for next-generation communication system.

The design of high-speed reconfigurable coprocessor for next-generation communication system consists of the data input terminal and the 6-bits-input-setup terminals; the proposed operation is activated by a help of clock input and reset. The reset pin is used to restart the operation when specific time excesses.

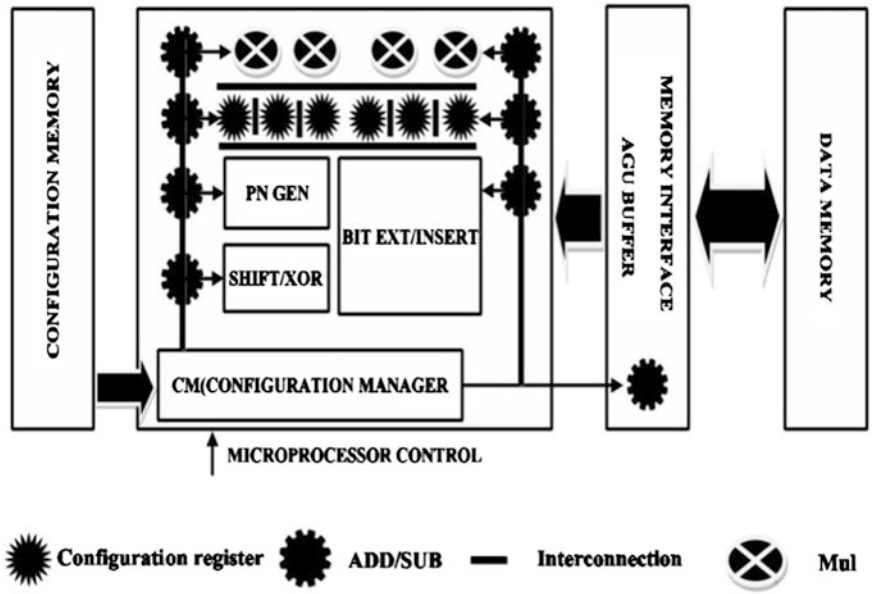


Fig. 8 The proposed high-speed reconfigurable coprocessor for next-generation communication system

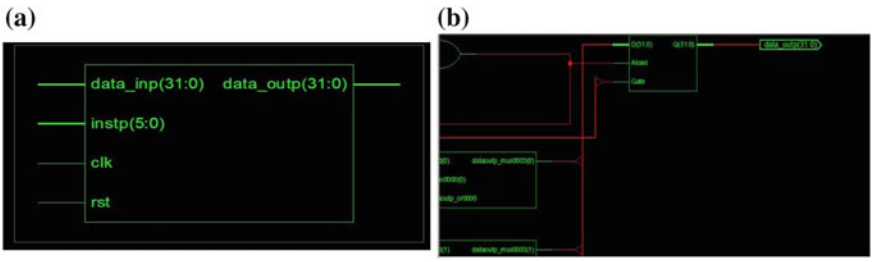


Fig. 9 **a** The realization of top level of proposed design of high-speed reconfigurable coprocessor for next-generation communication system. **b** The internal structure of proposed design of high-speed reconfigurable coprocessor for next-generation communication system

Figure 9b shows the internal structure of proposed design of high-speed reconfigurable coprocessor for next-generation communication system.

The proposed coprocessor has been evaluated for various communication algorithms using the simulator. Table 1 and Fig. 10 show the performance comparisons between the proposed coprocessor and the conventional DSP (SC140) that is VLIW architecture. It shows the proposed coprocessor as better performance than the conventional DSP (SC140) [8].

Table 1 Performance comparisons for various operations

Operation	SC140 (operation/cycle)	Proposed (operation/cycle)
Multiply	4	4
Addition and subtraction	2.5	4
MAC	4	4.5
Complex multiply	1	1.5
Complex MAC	1	1.5
FFT butterfly	1	2

*Numbers represent # of operation per cycle

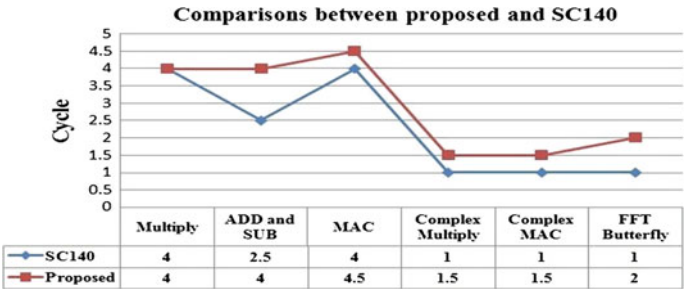


Fig. 10 Comparison between proposed and SC140

Table 2 Performance comparison for bit shuffle unit

Operator	SC140 4 shifter 4ALU	TMS320C6x 4 shifter 4 ALU	Proposed reconfiguration coprocessor Bit shuffle unit
Convolution encoding (cycle) (IS = 95, K = 9, R = 1/2, 192 bits)	463	NA	144
Block interleaving (cycle) 16*6 bits	414	NA	84
Scrambling (MIPS) (802.11a, 12 Mbps)	N A	39 × 106	20 × 106
Convolution encoding (cycle) (802.11a, 12 Mbps)	N A	77 × 106	12 × 106

Table 2 shows the comparison between proposed coprocessor bit shuffle unit and Star-Core SC140. The proposed coprocessor can reduce the clock cycles about 78 % for interleaving and 67 % convolutional encoding. Compare with TMS320C6x, the proposed coprocessor can reduce the clock cycles by about 84 % for convolutional encoding and by about 48 % for scrambling for the IEEE 802.11a standard (12-Mbps data rate).

4 Conclusions

In this work, we have proved the design and the implementation of high-speed reconfigurable coprocessor for next-generation communication system, and the reconfigurable coprocessor can support various communication standards and algorithms. An efficient operation of distribution communication system networks can be achieved by using reconfiguration techniques. The communication system network reconfiguration is carried out by changing the code rate and an instructions status of the sectionalizing switches. The proposed coprocessor can be used for communication operations, such as scrambling, interleaving, convolutional encoding, Viterbi decoding, FFT, and several other functions using the proposed design. The coprocessor has been modeled by VHDL and synthesis has been performed using the model-sim. Performance comparisons shows that the number of clock cycles can be reduced about 48 % for scrambling and 84 % for convolutional encoding compared with existing DSPs. From the results, performance of proposed coprocessor is better than conventional DSP (SC140) in terms of number of clock per cycle. The performance comparisons between the proposed coprocessor and the conventional DSP (SC140) that is VLIW architecture, the proposed coprocessor shows better performance than the conventional DSP (SC140) the number of operations per cycle. For example, SC140 performs 2 trellis butterfly calculations per cycle while the proposed coprocessor performs three butterfly calculations per cycle. The results of the bit shuffle unit performance is compared with Star-Core SC140, the proposed architecture can reduce the clock cycles about 67 % for convolutional encoding and about 78 % for block interleaving. Comparing with TI 62x, the proposed architecture can reduce the clock cycles about 48 % for scrambling and by about 84 % for convolutional encoding.

References

1. Agrwal EN, Nitin EN, Yadav M (2012) Study of fast frequency hopping spread spectrum and jamming systems. Int J Sci Res Pub 2(6):1–3
2. Kim SW (1992) Frequency-hopped spread-spectrum random-access with re-transmission cutoff and code rate adjustment. IEEE J Select Areas Commun 10:344–349
3. Youssef MI, Emam AE, Elghany MA (2009) Direct sequence spread spectrum technique with residue number system. Int J Electr Comput Syst Eng 3(4):223–230
4. Abidi AA (1999) Direct-conversion radio transceivers for digital communications. IEEE J Solid State Circ 30(12):1399–1410
5. Lee JH, Moon JH, Sunwoo MH (2005) Implementation of application-specific DSP for OFDM systems. In: Proceedings of ISAP 2005, Seoul
6. Mallikarjunaswamy S, Nataraj KR (2012) Design of high speed reconfigurable coprocessor for next generation communication platform. IEEE workshop on signal processing systems, 2004, SIPS 2004
7. Fragouli C, Komninakis C, Wesel RD (2001) Minimality for punctured convolutional codes. In: Proceedings of IEEE international conference on communications, 2001, ICC 2001

8. Motorola Semiconductors Inc. (2001) SC140 DSP core-reference manual. Denver, Colo, USA, p 3
9. Texas instruments (2004) TMS320C6x assembly language tools user's guide, pp 1–5
10. Texas instruments (2004) TMS320C55x assembly language tools user's guide, pp 1–10
11. Singh S, Sharma T, Sharma KG, Singh BP (2012) Array multiplier using pMOS based 3T XOR cell, p 2
12. Kim Y (2011) Reconfigurable multi-array architecture for low- power and high-speed embedded systems, p 209

Secret Code Authentication Using Enhanced Visual Cryptography

Rajendra Ajjipura Basavegowda and Sheshadri Holalu Seenappa

Abstract Secret code (password) is widely used in many applications like data transfer, sharing data, and login to emails or internet banking. So a big necessity to have a strong authentication to secure all our applications is as possible. In this paper, we present a new approach to authenticate password using Enhanced visual cryptography (EVC). Visual cryptography is a secret sharing scheme in which an image is converted into shares. No information can be revealed by observing any share. The information about the original image will be obtained on after stacking sufficient number of shares. We have a new approach to authentication secret code using our 2-out-of-2 EVC which provides efficient authentication for e-banking and other internet application.

Keywords Authentication · Enhanced visual cryptography (EVC) · Secret code (password) · Shares · Visual secret sharing scheme

1 Introduction

A password is a form of secret authentication of code that is used to controlled access to a resource. The password is kept secret, from those not allowed to access, and those wishing to gain access are tested on whether or not they know the password and are granted or denied access accordingly. To overcome the vulnerabilities of traditional methods, graphical password schemes have been developed for more authentications. Using EVC, we are providing security to

R. Ajjipura Basavegowda (✉) · S. Holalu Seenappa
PET Research Centre, PES College of Engineering, Mandya, Karnataka, India
e-mail: rajendraab@hotmail.com

S. Holalu Seenappa
e-mail: hssheshadri@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_8,

secret code. Visual cryptography (VC) scheme uses permutation techniques to encode secure code. The idea is to convert the secret code into n shadow images (shares). The decoding only requires selecting some subsets of these n shares, making transparencies of them, and stacking them on top of each other [1, 2]. This paper is organized as follows. Section 2 shows the fundamental principles of 2-out-of-2 VC scheme. Section 3 shows the proposed method to improve the contrast and an approach to secure the secret code. Finally, conclusions are drawn in Sect. 4.

2 Visual Cryptography

VC is a model in which the decryption of the secret image is done by using human visual system without any computational complexity. There are two types of VC schemes: n-out-of-n and k-out-of-n VC schemes [3]. In n-out-of-n VC scheme, image is divided into n shares, and in order to decrypt the image, all n shares are stacked on each other. In k-out-of-n scheme, the shares generated from the image are Xeroxed onto n transparencies and distributed among participants, one for each participant. Any k or more participants can visually reveal the secret image by superimposing any k transparencies together. The secret cannot be decoded by any $k - 1$ or fewer participants, even if infinite computational power is available to them [1], [4]. Following n-out-of-n scheme, we have used 2-out-of-2 scheme. In 2-out-of-2 VC scheme, a secret image is encrypted into two shares such that each share has random binary pattern of pixels. In order to decrypt the image, the two shares need to be overlaid [5].

2.1 Basic Model

Consider a set $Y = \{1, 2 \ldots n\}$ be a set of elements called participants. By applying set theory concept, we have 2^Y as the collection of all subsets of Y. Let $\Gamma_Q \subseteq 2^Y$ and $\Gamma_F \subseteq 2^Y$, $\Gamma_Q \cap \Gamma_F = \theta$ and $\Gamma_Q \, U \, \Gamma_F = 2^Y$, members of Γ_Q are called qualified sets and members of Γ_F are called forbidden sets [4]. The pair (Γ_Q, Γ_F) is called the access structure of the scheme. Γm can be defined as all minimal qualified sets: $\Gamma m = \{A \in \Gamma_Q : A| \notin \Gamma_Q \text{ for all } A| \subset A\}$.

Γ_Q can be considered as the closure of Γ_m, and Γ_m is termed a basis, from which a strong access structure can be derived [1]. Considering the image, it will consist of a collection of black and white pixels. Each pixel appears in n shares, one for each transparency or participant. Each share is a collection of m black and white subpixels. The overall structure of the scheme can be described by an n x m (No. of shares $\times$ No. of subpixels) Boolean matrix $S = [S_{ij}]$, where

$S_{ij} = 1$, if and only if the jth subpixel in the ith share is black.
$S_{ij} = 0$, if and only if the jth subpixel in the ith share is white.

Following the above terminology, let (Γ_Q, Γ_F) be an access structures on a set of n participants. A $(\Gamma_Q, \Gamma_F, \alpha)$-VCS with the relative difference α and set of thresholds $1 \leq k \leq m$ is realized using the two $n \times m$ basis matrices S_w and S_b, if the following condition holds:

1. If $Y = \{i_1, i_2, \ldots i_p\}$ Γ_Q, then the "or" V of rows $i_1, i_2, \ldots i_p$ of S_w satisfies $H(V) \leq k - \alpha \cdot m$, whereas, for S_b, it results that $H(V) \geq k$.
2. If $Y = \{i_1, i_2, \ldots i_p\}$ Γ_F, then the two $p \times m$ matrices obtained by restricting S_w and S_b to rows $i_1, i_2, \ldots i_p$ are identical up to a column permutation [6, 7].

The first condition is called contrast and the second condition is called security. The collections C_w and C_b are obtained by permuting the columns of the basis matrices S_w and S_b in all possible ways [8, 9]. The important parameters of the scheme are as follows:

- m, the number of subpixels in a share. This represents the loss in resolution from the original image to the shared one. The m is computed using the equation:

$$m = 2^{n-1} \tag{1}$$

- α, the relative difference. It determines how well the original image is recognizable. This represents the loss in contrast. The α is to be as large as possible and is calculated using the equation:

$$\alpha = |n_b - n_w|/m \tag{2}$$

where n_b and n_w represent the number of black subpixels generated from the black and white pixels in the original image.

- β, the contrast. The value β is to be as large as possible. The contrast β is computed using the equation:

$$\beta = \alpha \cdot \mathrm{m} \tag{3}$$

The minimum contrast that is required to ensure that the black and white areas will be distinguishable if $\beta \geq 1$ [3].

2.2 Generation of Shares

In order to generate the shares in the 2-out-of-2 scheme, we have the following mechanism (Table 1).

An original black pixel is converted into two subpixels for two shares, shown in 1st row. After stacking the two shares, we will get a perfect black. Similarly, we have other combinations for two subpixels generated shown in 2nd row. For

Table 1 Pixel pattern for 2-out-of-2 VC scheme

Pixel color	Original Pixel	Share1	Share2	Share1+ Share2
Black	■	■□	□■	■■
Black	■	□■	■□	■■
White	□	■□	■□	■□
White	□	□■	□■	□■

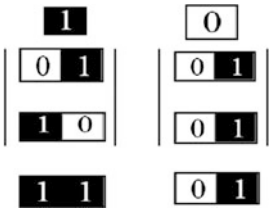


Fig. 1 Basis matrices construction

original white pixel, also we have two subpixels for each of the two shares, but after stacking the shares, we will not get exact white. We have a combination of black and white subpixels. This results in the loss of the contrast. Considering the following Fig. 1, we can generate the basis matrix:

The basis matrices for white and black pixels are given as:

$$s_w = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$s_b = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

In general, if we have $Y = \{1, 2\}$ as set of number of participants, then for creating the basis matrices S_w and S_b, we have to apply the odd and even cardinality concept of set theory. For S_w, we will consider the even cardinality and we will get $S_w = \{\Theta, \{1, 2\}\}$ and for S_b, we have the odd cardinality $OS_b = \{\{1\}, \{2\}\}$. In order to encode the black and white pixels, we have collection matrices which are given as

$C_w = \{$Matrices obtained by performing permutation on the columns of $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\}$

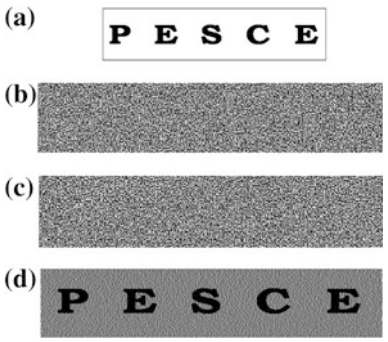


Fig. 2 Visual cryptography scheme. **a** Original image. **b** Share 1. **c** Share 2. **d** Decrypted image

$C_b = \{\text{Matrices obtained by performing permutation on the columns of } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\}$

So finally, we have,

$$c_w = \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

$$c_b = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Now, to share a white pixel, randomly select one of the matrices in C_w, and to share a black pixel, randomly select one of the matrices in C_b. The first row of the chosen matrix is used for share 1 and the second for share 2.

2.3 Stacking of Shares

The Fig. 2 shows the stacking of the shares. Figure 2a shows the original image, and Fig. 2b, c are the shares generated from the original image. Figure 2d shows the decrypted image after stacking the two shares. From the Fig. 2d, it can be observed that contrast in the decrypted image is less. In order to improve the contrast, an analysis on the relative contrast value is required.

3 Experimental Results and Discussion

3.1 Proposed Method

Based on the analysis on the relative contrast, we have the following observation:

Table 2 Relative contrast value v/s number of subpixels

Shares n	Subpixels $m = 2^{n-1}$	Relative contrast (α)	Contrast $\beta = \alpha \cdot m$
2	2	0.50	1
3	4	0.25	1
4	8	0.125	1
5	16	0.0625	1
6	32	0.03125	1

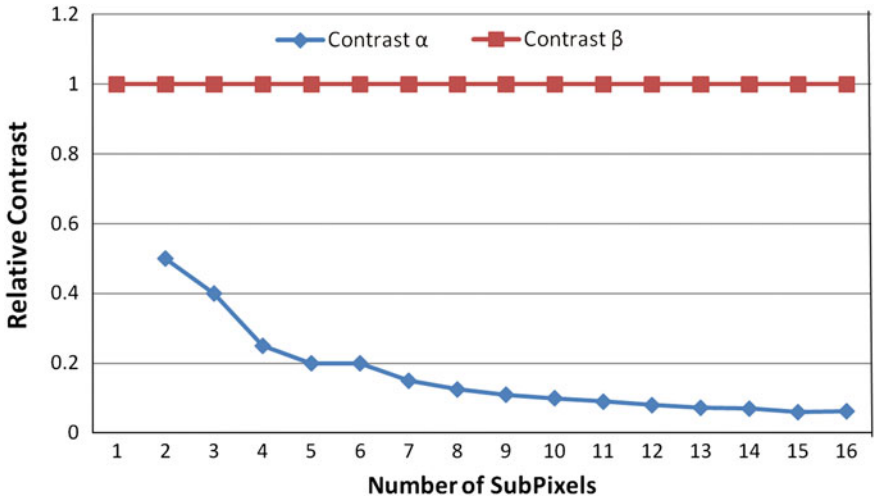


Fig. 3 Relative contrast versus number of subpixels

From the Table 2, we can see that the relative contrast value decreases as the number of subpixels increases. The following Fig. 3 depicts the same.

So considering the same 2-out-of-2 EVC in order to increase the relative contrast value, we have used an additional matrix along with the basis matrices. The additional matrix is used to share the white pixels in the reconstructed secret image. The additional matrix can be formed in the following manner:

Let Y be the set which is given by

$Y = \{i_1, i_2 \ldots i_n\}$ of n elements.
We define an additional matrix AS_w with order n × m such that
$AS_w = [AS_{ij}]$ where
$AS_w = 0$ if and only if $1 \leq i \leq n$ and $j = 1, 2$.

The collection matrices will be obtained in the following manner:

$$C_w = \{\text{Basis Matrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + \text{Additional Matrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\}$$

$$C_b = \{\text{Matrices obtained by performing permutation on the columns of} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\}$$

Now the value of α will be equal to 3/4. This result shows that the relative difference of proposed method is better compare to the existing one.

3.2 Stacking of the Shares

The Fig. 4 shows the stacking of the shares. Figure 4a shows the original image, and Fig. 4b, c are the shares generated from the original image. Figure 4d shows the decrypted image with better contrast. With this better contrast, we can apply our approach in various fields for achieving the security objectives.

In the Fig. 5, the visual secret sharing application will take this secure code and splits the secure code into shares. Our proposed method uses 2-out-2 EVC which improves the contrast. So that after decryption, the image is clearly visible. The shares will be generated which will be in encrypted form. The shares transmitted in different channels in secured way. Authorized persons will receive the shares from different channels, and then, shares are overlaid to get the secret code.

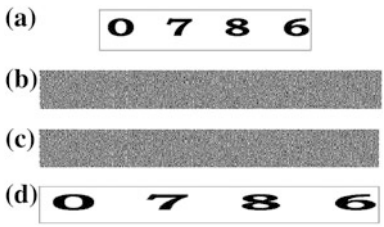


Fig. 4 Enhanced visual cryptography scheme

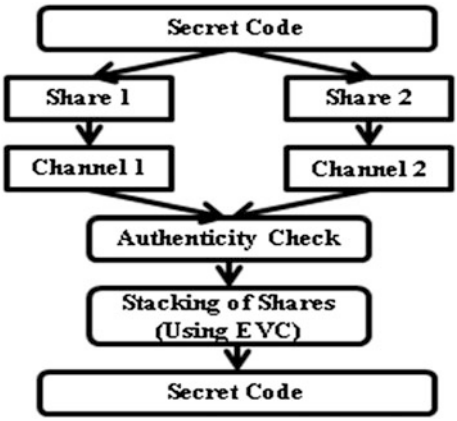


Fig. 5 Software architecture of the proposed method

4 Conclusion

In this paper, a new approach is proposed to transmit secret code. The secret code is divided into two shares using 2-out-of-2 EVC, and it is sent in different channels. Where the contrast of the stacked share (decrypted secret code) is comparatively better than earlier methods, it provides security to devices, to share secret data among the group and split shares in different channels to protect from hackers. This application can also be used for sending secret codes to a customer of the e-banking and in similar highly secured applications.

This is a part of our approach to certain modifications in the methods of encrypting medical data in an image for secured telemedicine. Further the research work is going on graphical password with visual cryptography.

Acknowledgments I thank my guide Dr. H S Sheshadri, professor and dean (research) for supporting to carry out this work.

References

1. Shamir A (1979) How to share a secret. Comm ACM 22(11):612–613
2. Naor M, Shamir A (1995) Visual cryptography. In: Proceedings of advances in cryptology EUROCRYPT' 94, LNCS, Springer, pp 1–12
3. Monoth T, Babu Anto P (2009) Achieving optimal contrast in visual cryptography schemes without pixel expansion. Int J Recent Trends Eng 1(1)
4. Ogiela MR, Ogiela U (2009) Linguistic cryptographic threshold schemes. Int J Future Gener Comm Network 2(1):33–40
5. Rajendra AB, Sheshadri HS (2012) A study on visual secret sharing schemes using biometric authentication techniques. AJCST 1:157–160
6. Manimurugan S, Porkumaran K (2011) A new fast and efficient visual cryptography scheme for medical images with forgery detection. In: Proceedings of IEEE international conference on emerging trends in electrical and computer technology (ICETECT) 2011, pp 594–599
7. Blundo C, University of Salerno, De Santis A, Stinson DR, University of Nebraska-Lincon (1996) On the contrast in visual cryptography scheme
8. Radha N, Karthikeyan S (2010) A study on biometric template security. ICTACT J Soft Comput 01
9. Stinson D (1999) Visual cryptography and threshold schemes. IEEE Potentials 18(1):13–16

A Data Acquisition (DAQ) Architecture for FPGA Implementation

Prashant Kramadhari, C. R. Anand Kumar, Vinayak Hungund and H. S. Sheshadri

Abstract As the need of high-precision application and availability of high-bandwidth ADC increases, the conventional method of interfacing the high-bandwidth ADC to processor becomes uneconomical (processor+ memory), inefficient (performance of CPU unused), rigid (lacking repeated customization). Hence, a dedicated piece of hardware which serves the purpose of the interface is needed and FPGAs prove to be one such reliable option due to their features like customization, power-consumption, redundancy, flexible IOs, and reconfiguration. In this paper, we present one such implementation for acquiring data through the FPGA coprocessor.

Keywords Data acquisition · FPGA · ADC · Bandwidth · Reconfiguration

1 Background

Data acquisition (DAQ) is the process of sampling signals that measure real-world physical conditions and converting samples into digital numeric values that can be processed by a processor/computer. The quality of the signals output at the DAQ

P. Kramadhari (✉)
IAV Systems Pvt. Limited, Bangalore, India
e-mail: kramadhari.prashant@gmail.com

C. R. Anand Kumar
HCL Technologies Ltd, Noida, India
e-mail: kumar.anand743@gmail.com

V. Hungund
Continental Automotive Systems Pvt. Ltd, Bangalore, India
e-mail: v.k.hungund@gmail.com

H. S. Sheshadri
PESCE, Mandya, Karnataka, India
e-mail: hssheshadri@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_9,

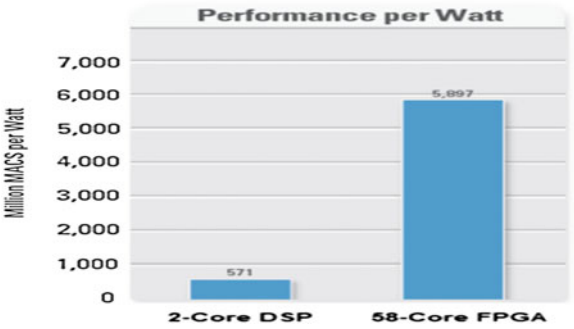


Fig. 1 Performance comparison (FPGA vs. conventional DSP processor). [+] *MACS*—Multiply–accumulate operations per second (measure of DSP performance)

subsystem directly affects further processing. A DAQ application developed on a DSP processor serves the purpose of controlling the DAQ process. Due to sequential processing of application software, the throughput achieved may not be suited for applications requiring high-speed DAQ from multiple channels. To overcome some of the drawbacks of usage of software as application layer, we have introduced a dedicated digital circuitry as an accelerator to overcome low-throughput bottlenecks of software applications.

Model-based design is the new concept used in many upcoming areas like automotive and aeronautics which helps to cut the design, development, and test effort considerably. Matlab offers a wide variety of support for the development of various environments like C.VHDL. The system generator toolbox provided by Xilinx is used for generating the VHDL code for our design [1] (Fig. 1).

The figure above shows the performance characterization of Xilinx[R] Spartan 3E FPGA with a typical DSP processor signifying the enhanced performance per watt providing enough leverage for acquisition and preprocessing operations inside the FPGA instead of being restricted to glue logic.

FPGA has got IOs that can support a wide range of signal standards for interfacing conducive for DAQ applications [2] such as multi-voltage, multi-standard IO interface pins signal standards (3.3, 2.5, 1.8, 1.5, and 1.2 V signaling).

Providing flexibility in acquiring data of various standards and at the same time being configurable in nature.

In this paper, we have described the DAQ logic implemented on a Spartan-3 FPGA for a delta-sigma ADC, ADS 1252 which is a 24-bit serial ADC with an effective resolution of 19 bit up to 40 KHZ taking advantage of the aforementioned potential performance enhancement of FPGA.

1.1 ADS 1252 ADC Chip

The block diagram of ADS 1252 is as shown in the figure below.

The ADS 1252 is a precision, high-dynamic range, 24-bit, delta-sigma A/D convertor capable of achieving very high-resolution digital results at high data

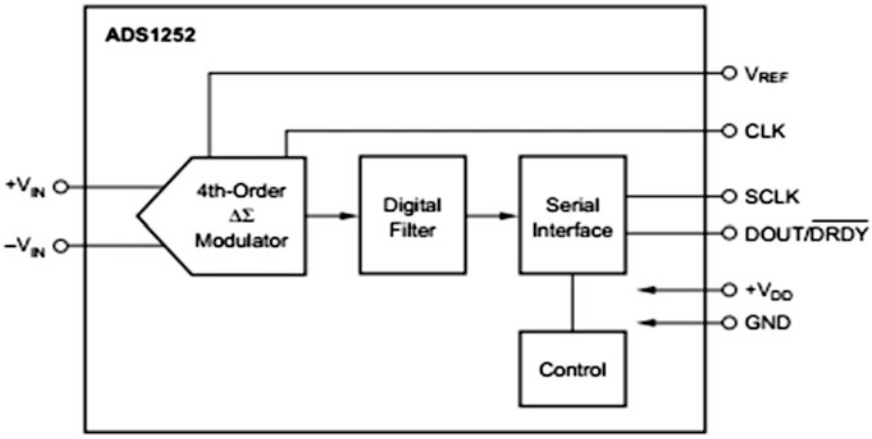


Fig. 2 Block diagram of ADS 1252 chip

rates [1]. The analog data is sample at the rate determined by frequency of the system clock (CLK). The sampled analog input is modulated by the delta-sigma A/D modulator, which is followed by a digital filter. A sinc digital low-pass filter the output of delta-sigma modulator and the writes the result into a data output register. The DOUT/DRDY is pulled low, indicating that the new data are available to be read by external FPGA. As shown in the Fig. 2, the main functional blocks of ADS 1252 are including fourth-order delta-sigma modulator, a digital filter, control logic, and a serial interface.

The ADS1252 has 384-cycle sub-period of operation for every period of A/D conversion. A 384-cycle sub-period comprises of 24 high, 6 low, and 6 high; this indicates that the analog data are converted and ready for transmitting out of ADS 1252. We need to provide at SCLK for 24 cycles with the frequency of range between 4 and 16 MHZ. At every rising edge of SCLK, we get the serial data latched out of ADC.

The timing diagram for the above-mentioned operation is as shown below.

When we have multiple ADS 1252 in a system, we need to synchronize them by holding SCLK high in the beginning of operation for 4–20 cycles. At the end of 20 cycles, all the A/D in the system are synchronized and we can use common SCLK for all ADCs. This would be very useful so that we have a single-clock domain in terms of SCLK else we will have multiple SCLK, and it would be very difficult to manage data.

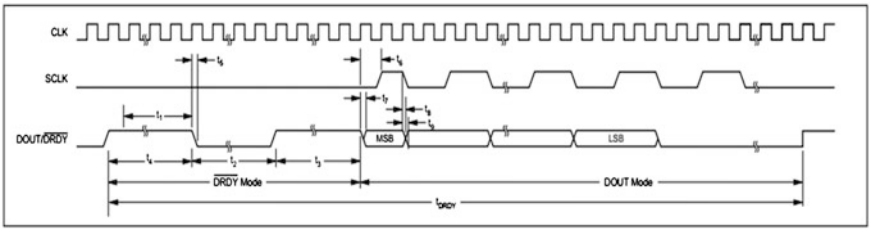


Fig. 3 DOUT/DRDY timing of ADS 1252

2 FPGA Logic Design

The system consists of a DSP-FPGA-based controller which needs to fetch serial ADC DATA from ADS 1252 ADC chip and convert it to a parallel format before we transmit it to DSP via EMIF bus. DSP would read the data from EMIF bus via DMA routine and transfer the DATA to its internal RAM memory.

The objective is to detect whether A/D has performed the conversion and this can be performed by reading 24-high, 6-low, and 6-high DRDY/DOUT signals. This is always an indication of DATA conversion.

FPGA can provide SCLK (refer Figs. 3 and 4) after 24 + 6 + 6 and before 384 cycles and can read 24-bit serial data into its registers. FPGA would than convert the 24-bit serial data in DRDY/DOUT line shown in Fig. 4 to parallel data of 16-bit precision (EMIF bus interface is configured to be 16 bits) for subsequent DSP operations.

2.1 Design Considerations

The design consideration would be based on the bandwidth of ADS 1252 which is around 40-KHZ data rates. The CLK and SCLK frequencies [3] are based on the above bandwidth and requirement of the ADS 1252 and the control loop frequency of our system which is around 3 KHZ. This scenario would lead us with 12 MHZ

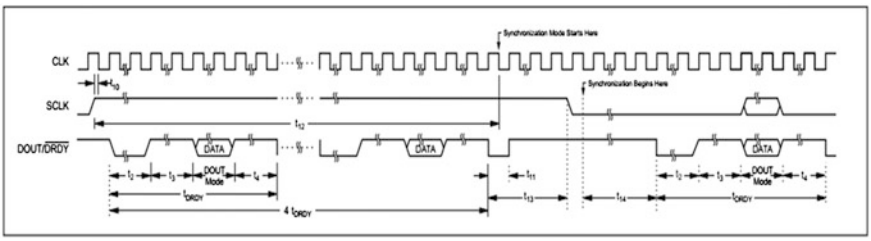


Fig. 4 Synchronization mode timing

CLK with an output data rate of 30 KHZ which was very easily perceptible from FPGA. There is also a possibility of increasing the number of channels, that is, we can connect more number of ADC to FPGA which could all work in parallel with synchronization of ADS 1252. The limit of connected of ADCs is only dependent on number of free IO pins of FPGA.

2.2 FPGA Logic Design Functional Description

The VHDL code is generated using the system generator toolbox [4] available from Xilinx which reduces the development time and also helps in testing/comparing the functional requirement with more ease. The generated VHDL code could be tested by Model-sim or Xilinx ISE Simulator (Figs. 5 and 6).

The block diagram in Fig. 7 representing the FPGA chip top-level module shows the implementation of a four-channel input single output interface for the ADS 1252 IC. For the ADS 1252 as many as 32 channels could be simultaneously used (384/12 = 32) [3, 5]. A typical implementation of 4 channels and its usage in terms of the FPGA resource and capability of a typical FPGA (Spartan 3E) is discussed in this paper.

The blocks marked in yellow represent the FPGA boundary and interface with external devices. The 2 blocks on the left side in Fig. 7 (labeled 'Counter 1' and 'ROM 1') are for simulating the data or bit sequence of the ADS 1252 IC which is connected to all the 4 input channels of the FPGA. The Channel_1 block in the middle is the main interface which converts the serial data from the input channels into single parallel output (16 * 2 * 2) which could be spit out in any data length using a FIFO.

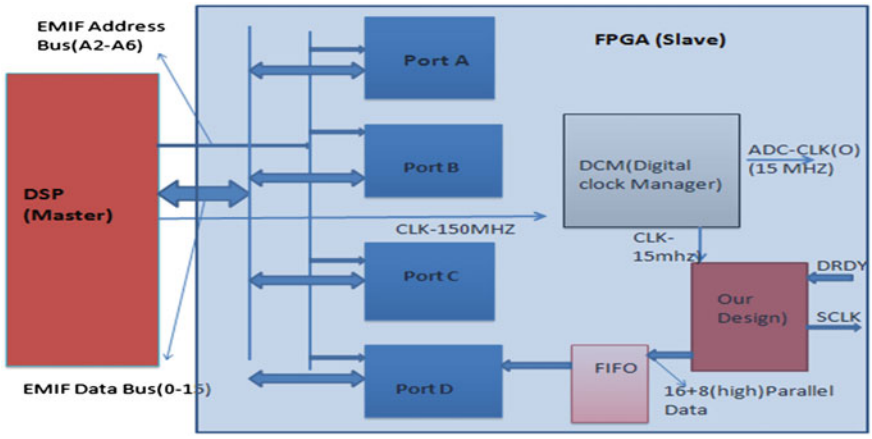


Fig. 5 DAQ System architecture

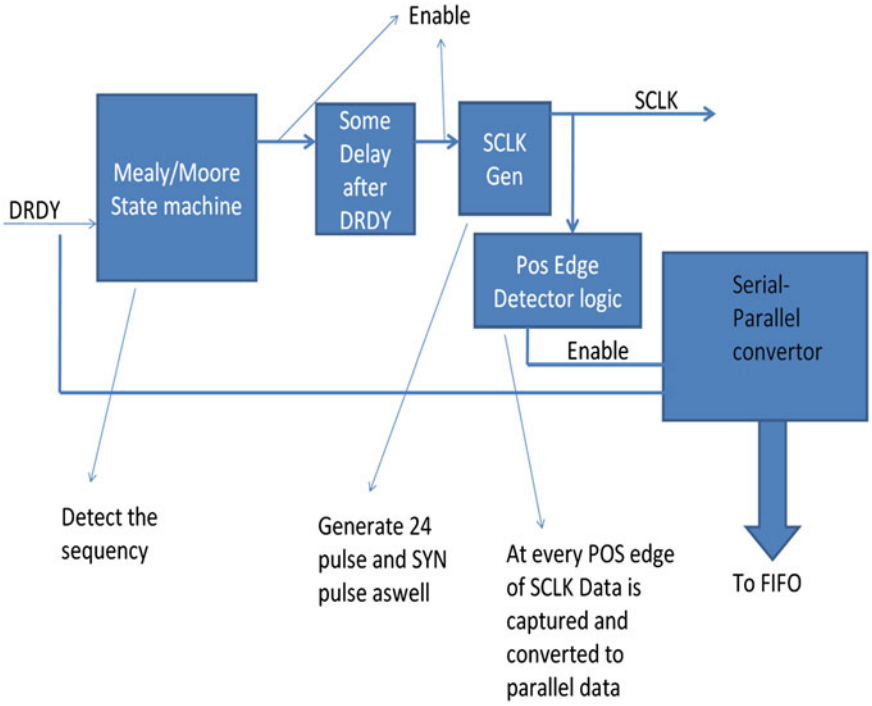


Fig. 6 Functional representation of customized FPGA logic

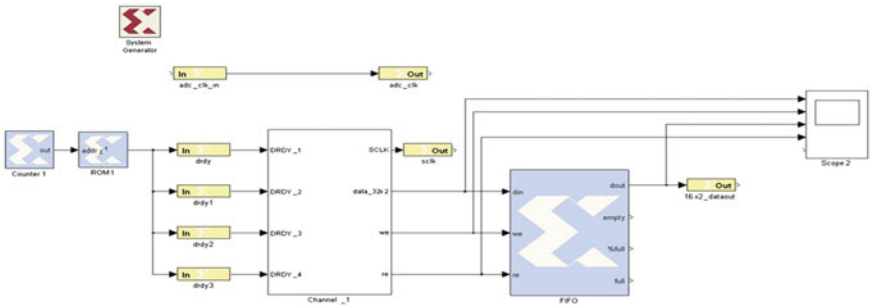


Fig. 7 FPGA chip top level module

2.3 FIFO Depth Calculation

The design might run the risk of data overflow/underflow if a FIFO of optimum depth is not incorporated. The depth of the FIFO was calculated taking into consideration the following parameters [6] (Figs. 8, 9, 10 and 11).

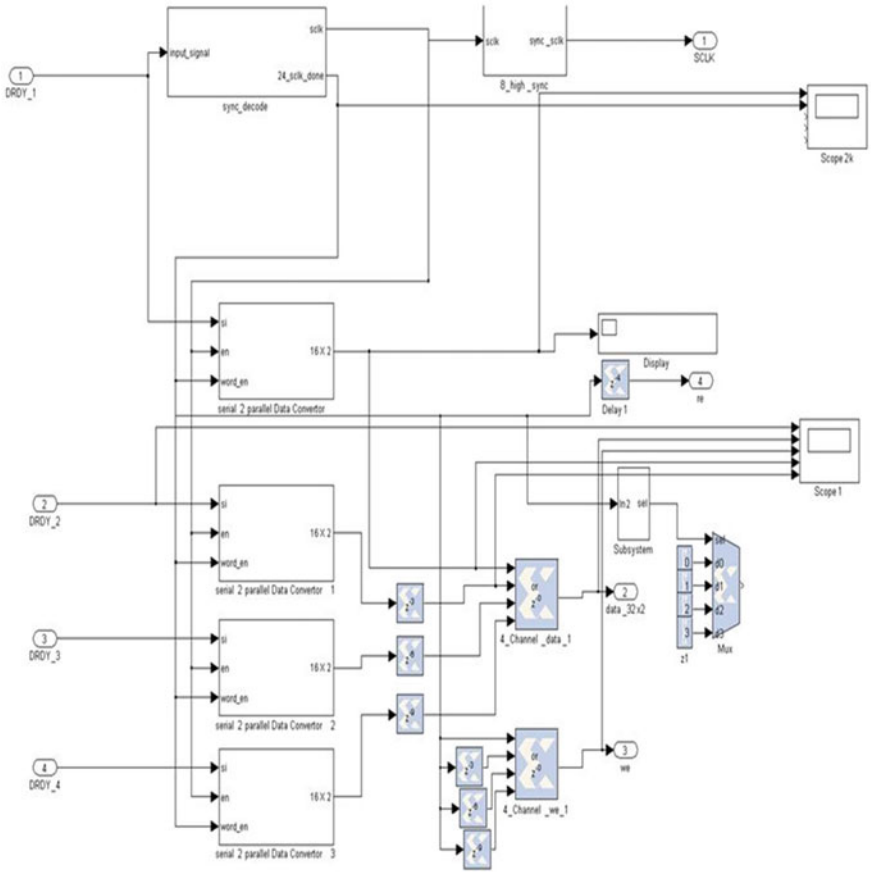


Fig. 8 Functional representation of a single channel

- A synchronized FIFO
- Writing clock 25 MHz—F1
- Reading clock 75 MHz—F2, Idle cycles = 6
- Writing burst size—B.

$$\text{FIFO depth} = B - B * F2/(F1) = B - B * (75/(25 * 6)) = B - B/2 = B/2.$$

3 FPGA Resource Consumption

Figure 12.

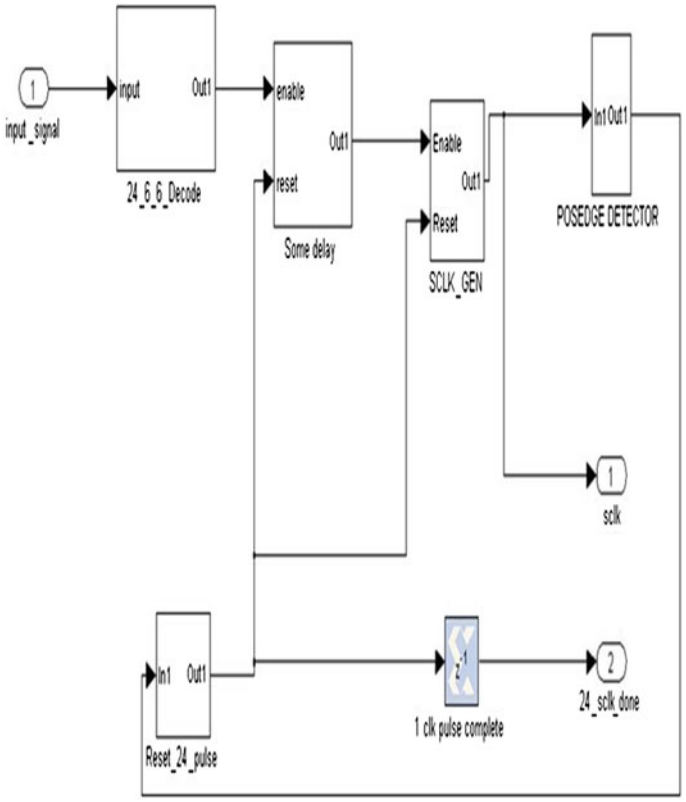


Fig. 9 Decoding and SCLK generation logic

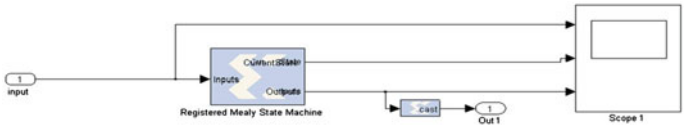


Fig. 10 Control FSM

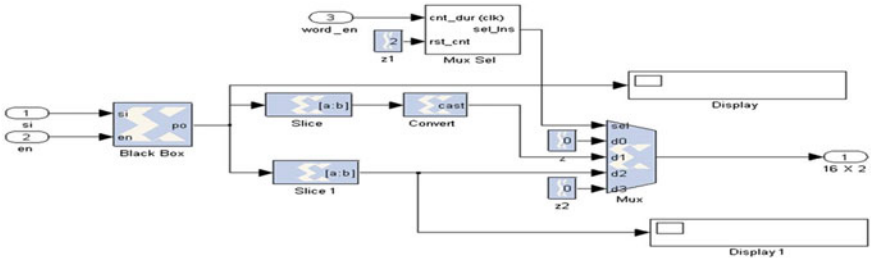


Fig. 11 Data output control

Device Utilization Summary				
Logic Utilization	Used	Available	Utilization	Note(s)
Total Number Slice Registers	1,742	7,168	24%	
Number used as flip flops	1,726			
Number used as Latches	16			
Number of 4 input LUTs	1,440	7,168	20%	
Logic Distribution				
Number of occupied Slices	1,459	3,584	40%	
Number of Slices containing only related logic	1,459	1,459	100%	
Number of Slices containing unrelated logic	0	1,459	0%	
Total Number of 4 input LUTs	1,607	7,168	22%	
Number used as logic	1,440			
Number used as a route thru	167			
Number of bonded IOBs				
Number of bonded	87	97	89%	
Number of BUFGMUXs	3	8	37%	
Number of DCMs	1	4	25%	
Number of RPM macros	19			

Fig. 12 FPGA resource consumption (Xilinx ISE tool)

4 Conclusion

This paper depicts the DAQ mechanism implemented through on an FPGA which provides flexibility and at the same time enhances the functionality of the system as it acts as a coprocessor which accelerates the preprocessing thereby off-loading some computation intensive tasks from DSP to FPGA DAQ logic. Exploitation of parallelism of FPGA has also been illustrated by parallel implementation of 4 such DAQ functionalities on a single FPGA chip which equally demonstrates the scalability of the design limited only by the availability of IOs on a given FPGA chip.

Model-based design is an efficient way to design complex algorithms. Although not yet widely used, it has many advantages. Designs can be prototyped extremely fast, without the need for a low-level description and waveform simulations. Standard HDL languages are automatically generated from an abstract model. A designer can concentrate on the behavior, not on the detailed logic implementation. Model-based design offers a possibility to drastically reduce the design time. The full capability of very complex system-on-chip hardware can be explored. When using standard HDL languages, debugging and simulation usually take more time than the creation of a description of the actual hardware that has to be built.

Acknowledgments The authors express their sincere gratitude to the director of BiSS Research Dr R Sunder and Dr. Bruno Paillard, senior designer convergence instruments for their logistical and technical support during our work.

References

1. Xilinx-DSP model based design solution—Avnet XFest (2007)
2. Spartan -3 FPGA user guide, http://www.xilinx.com/support/documentation/data_sheets/ds099.pdf
3. ADS1252 datasheet. http://www.ti.com/product/ads1252
4. Van Beeck K, Heylen F, Meel J, Goedemé T (2010) Comparative study of model-based hardware design tools. Campus De Nayer, Association K. U. Leuven, Jan De Nayerlaan 5, 2860 Sint-Katelijne-Waver, Belgium
5. Texas Instruments TMS320C5502 data sheet
6. Calculating FIFO depth. http://www.asic-world.com/tidbits/fifo_depth.html
7. Sudeep KC, Anand Kumar CR, Vinayaka S, Solanki P, Dr. Majumdar J (2011) A novel architecture for FPGA implementation of real time automatic thresholding. In: Proceedings of the 1st CCSO, Dr.AIT, 13 May 2011
8. Signal Ranger SR-MK2 user guide. http://www.latticesemi.com/documents/doc26686x11.pdf
9. Interfacing analog to digital convertor with FPGA. http://www.softdb.com/dsp-products-SR-MK2.php

Finite-Dimensional Realization of Lavrentiev Regularization for Nonlinear III-posed Equations

Suresan Pareth

Abstract A finite-dimensional realization of the two-step Newton method is considered for obtaining an approximate solution (reconstructed signals) for the nonlinear ill-posed equation $F(x) = f$ when the available data (noisy signal) is f^δ with $f - f^\delta \leq \delta$ and the operator F is monotone. We derived an optimal-order error estimate under a general source condition on $x_0 - \bar{x}$, where x_0 is the initial approximation to the actual solution (signal) $\bar{x}$. The choice of the regularization parameter is made according to the adaptive method considered by Pereverzev and Schock (2005). 2D visualization shows the effectiveness of the proposed method.

Keywords Newton method · Lavrentiev regularization · Monotone operator · Ill-posed problems · Nonlinear analysis

1 Inverse Problems and Applications

In this paper, we consider the nonlinear equation

$$F(x) = f \tag{1}$$

that arise from the study of nonlinear inverse problems, where $F : D(F) \subseteq X \to X$ is a nonlinear monotone operator, and X is real Hilbert space.

Inverse problems are problems where causes for a desired or an observed effect are to be determined. They have, nearly always driven by applications, been studied for nearly a century now. An important feature, both theoretically and numerically, of inverse problems is their ill-posedness. We call the problem (1) ill-posed in the sense that its solution does not depend continuously on the

S. Pareth (✉)
Department of Mathematical and Computational Sciences, National Institute of Technology, Surathkal, Karnataka 575025, India
e-mail: sureshpareth@rediffmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_10,

right-hand side, that is, they do not fulfill Hadamard's classical requirements of existence, uniqueness, and stability, under data perturbations, of a solution. Solutions of an inverse problem might not exist for all data (e.g., a consistent temperature history exists only for a very smooth final temperature in the model of the classical heat equation), it might not be unique (which raises the practically relevant question of identifiability, that is, the question if the data contain enough information to determine the desired quantity), and it might be unstable with respect to data perturbations. The last aspect is of course especially important, since in real-world problems, measurements always contain noise (an other source of noise being errors in numerical procedures), and approximation methods for solving inverse problems which are as insensitive to noise as possible have to be constructed, so-called regularization methods. The proposed method will concern "Emerging Applications of Inverse Problems Techniques to Signal Reconstruction and Imaging Science". These include the problem of restoration or approximation of the signal from the noisy data [8], tomography in medical imaging, microscopy, atmospheric imaging, biometrics, integrating optics and imaging, and computer vision techniques. Other applications are in applied optics and remote sensing sciences. Atmospheric turbulence blur arising in applied optics and remote sensing is due to long-term exposure through the atmosphere where turbulence in the atmosphere gives rise to random variations in the refractive index.

The organization of this paper is as follows. A brief introduction is given in Sect. 2. The method and its convergence analysis are given in Sect. 3; the error bounds and the regularization parameter choice strategy are given in Sects. 4 and 5 deals with implementation of the adaptive choice rule. Numerical results are given in Sect. 6. Finally, the paper ends with a conclusion in Sect. 7.

2 Introduction

In [1], the authors proposed a two-step Newton Lavrentiev method for approximately solving the ill-posed operator Eq. (1). In this paper, we consider the finite-dimensional realization of the method considered in [1].

Throughout this paper, the inner product and the corresponding norm on the Hilbert space X are denoted by $\langle,\rangle$ and $\|.\|$. Recall that [10, 11] F is monotone, if $F(x) - F(y), x - y \geq 0$, $\forall x, y \in D(F)$, the domain F. We assume that $f^\delta \subseteq X$ are the available noisy data with $\left\|f - f^\delta\right\| \leq \delta$where δ denotes the noise level and (1) has a solution $\bar{x}$. Since (1) is ill-posed, the regularization methods are used to obtain a stable approximate solution for (1).

For monotone operators, one usually uses the Lavrentiev regularization method (see [5, 9–11]) for solving (1). In this method, the regularized approximation x_α^δ is obtained by solving the operator equation

$$F(x) + \alpha(x - x_{0)} = f^{\delta}. \tag{2}$$

From the general regularization theory, it is known that the Eq. (2) has a unique solution x_{α}^{δ} for any $\alpha > 0$ and $x_{\alpha}^{\delta} \to \bar{x}$ as $\alpha \to 0, \delta \to 0$, provided α is chosen appropriately (see, [9, 11]).

3 The Method and its Convergence

Let $\{P_h\}_{h>0}$ be family of orthogonal projections on X. Let $\in_h := F'(P_h x_0)(I - P_h)$ and $\{b_h : h > 0\}$ be such that $\lim_{h\to 0} \frac{\|(I-P_h)x_0\|}{b_h} = 0$ with $\lim_{h\to 0} b_{h=0}$. We assume that $\in_h \to 0$ as $h \to 0$. The above assumption is satisfied if $F(P_h x_0)$ is a compact operator. Further, we assume that $\in_h \leq \in_0$, $b_h \leq b_0$ and $\delta \in (0, \delta_0]$.

Analogous to the iterative scheme (1.5) and (1.6) in [1], we define an iterative scheme to obtain an approximate solution for the Eq. (2) in the finite-dimensional subspace of X by

$$y_{n,\alpha}^{h,\delta} = x_{n,\alpha}^{h,\delta} - \mathrm{R}_{\alpha}^{-1}\left(x_{0,\alpha}^{h,\delta}\right) P_h \left[F\left(x_{n,\alpha}^{h,\delta}\right) - f^{\delta} + (x_{n,\alpha}^{h,\delta} - x_0)\right] \tag{3}$$

$$x_{n+1,\alpha}^{h,\delta} = y_{n,\alpha}^{h,\delta} - \mathrm{R}_{\alpha}^{-1}\left(x_{0,\alpha}^{h,\delta}\right) P_h \left[F\left(y_{n,\alpha}^{h,\delta}\right) - f^{\delta} + (y_{n,\alpha}^{h,\delta} - x_0)\right] \tag{4}$$

where $x_{0,\alpha}^{h,\delta} := P_h x_0$ is the projection of the initial iterate x_0 on to $R(P_h)$, the range of P_h and $R_{\alpha}(x) = P_h F'(x) P_h + \alpha P_h$ with $\alpha > \alpha_0 > 0$. Here, the regularization parameter α is chosen from the finite set $D_N = \{\alpha_i : 0 < \alpha_0 < \alpha_1 < \ldots < \alpha_N\}$. Note that even though the proposed method has local linear convergence, it requires, for its merit, the computation of the Fréchet derivative $F'(.)$ only at Note, that even though the proposed method has local linear convergence, it require for its merit, the computation of the Frechet derivative only at$P_h x_0$.

3.1 Convergence Analysis

We need the following assumptions for the convergence analysis.

Assumption 1 (*cf.* [10], *Assumption* 3) There exists a constant $\mathrm{k}_0 > 0$, $\mathrm{r} > 0$, such that for every $x, u \in B_r(x_0) \cap B_r(\bar{x}) \subset D(F)$ and $v \in X$, there is an element $\phi(x, u, v) \in X$, such that

$$[F'(x) - F'(u)]v = F'(u)\ \phi(x, u, v), \|\phi(x, u, v)\| \leq k_0\ \|v\|\ \|x - u\|.$$

Assumption 2 (*cf.* [7]) There exists a continuous, strictly monotonically increasing function $\varphi : (0.a) \to (0, \infty)$with $a > \|F'(\bar{x})\|$ satisfying $\lim_{\lambda\ >\ 0} \varphi(\lambda) = 0$ and $v \in X$ with $\|v\| \leq 1$ such that $x_0 - \bar{x} = \varphi(F'(\bar{x}))v$ and

$$\sup_{\lambda \geq 0} \frac{\alpha\varphi(\lambda)}{\lambda + \alpha} \leq \varphi(\alpha),\ \forall \lambda \in (0, a].$$

$$\text{Let } e_{n,\alpha}^{h,\delta} = \left\| y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta} \right\|,\ \forall n \geq 0. \tag{5}$$

Hereafter, assume that $\|F'(P_h x_0)\| \leq C_F$, $\delta_0 < \frac{\alpha_0}{4k_0(1+\frac{\in_0}{\alpha_0})}$, for some $\alpha_0 > 0$, $b_0 < \frac{\sqrt{1+\left(\frac{1}{2(1+\frac{\in_0}{\alpha_0})^2} - \frac{2\delta_0 k_0}{(1+\frac{\in_0}{\alpha_0})\alpha_0}\right)} - 1}{k_0}$ and $\|x_0 - \bar{x}\| \leq \rho$ where

$$\rho \leq \frac{\sqrt{1+\left(\frac{1}{2(1+\frac{\in_0}{\alpha_0})^2} - \frac{2\delta_0 k_0}{(1+\frac{\in_0}{\alpha_0})\alpha_0}\right)} - 1}{k_0}. \tag{6}$$

$$\text{Let } \gamma_\rho := \left(1 + \frac{\in_0}{\alpha_0}\right)\left[\frac{k_0}{2}(\rho + b_0)^2 + (\rho + b_0) + \frac{\delta_0}{\alpha_0}\right] \tag{7}$$

$$\text{and } q = \left(1 + \frac{\in_0}{\alpha_0}\right) k_0 r \tag{8}$$

$$\text{If, } r \in \left(\frac{1 - \sqrt{1 - 4k_0\left(1 + \frac{\varepsilon_0}{\alpha_0}\right)\gamma_\rho}}{2k_0\left(1 + \frac{\varepsilon_0}{\alpha_0}\right)}, \frac{1 + \sqrt{1 - 4k_0\left(1 + \frac{\varepsilon_0}{\alpha_0}\right)\gamma_\rho}}{2k_0\left(1 + \frac{\varepsilon_0}{\alpha_0}\right)}\right) \quad \text{then } \frac{\gamma_\rho}{1-q} < r. \tag{9}$$

In due course, we shall make use of the following lemma extensively.

Lemma 1 *Let $P_h x_0 \in D(F)$. Then $\left\|R_\alpha^{-1}(P_h x_0) P_h F'(P_h x_0)\right\| \leq \left(1 + \frac{\in_0}{\alpha_0}\right)$.*

Proof Note that

$$\begin{aligned}
\left\|R_\alpha^{-1}(P_h x_0) P_h F'(P_h x_0)\right\| &\leq sup_{\|v\| \leq 1} \left\|(P_h F'(P_h x_0) P_h + \alpha P_h)^{-1} P_h F'(P_h x_0) v\right\| \\
&\leq sup_{\|v\| \leq 1} \left\|(P_h F'(P_h x_0) P_h + \alpha P_h)^{-1} P_h F'(P_h x_0) P_h v\right\| \\
&\quad + sup_{\|v\| \leq 1} \left\|(P_h F'(P_h x_0) P_h + \alpha P_h)^{-1} P_h F'(P_h x_0)(I - P_h) v\right\| \\
&\leq \left(1 + \frac{\in_h}{\alpha}\right) \leq \left(1 + \frac{\in_0}{\alpha_0}\right).
\end{aligned}$$

□

Lemma 2 (cf. [2] Lemma 2) *Let $e_0 = e_{0,\alpha}^{h,\delta}$ and γ_ρ be as in (7). Then $e_0 \leq \gamma_\rho \leq 1$.*

Theorem 1 *Let $y_{n,\alpha}^{h,\delta}$, $x_{n,\alpha}^{h,\delta}$ and $e_{n,\alpha}^{h,\delta}$ be as in (3), (4), and (5), respectively with $\delta \in (0, \delta_0]$ and $\alpha \in D_N$. Then*

(a) $\left\|x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}\right\| \le q\left\|y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}\right\|$;

(b) $\left\|y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\right\|\| \le q^2\left\|y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}\right\|$;

(c) $e_{n,\alpha}^{h,\delta} \le \mathrm{q}^{2\mathrm{n}}\gamma_\rho$;

(d) $x_{n,\alpha}^{h,\delta}, y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$.

Proof Observe that if $x_{n,\alpha}^{h,\delta}, y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$, then by Assumption 1 and Lemma 1, we have,

$$\begin{aligned} x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta} &= y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta} - \mathrm{R}_\alpha^{-1}(P_h x_0)P_h\left[F\left(y_{n-1,\alpha}^{h,\delta}\right) - F\left(x_{n-1,\alpha}^{h,\delta}\right) + \alpha(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta})\right] \\ &= \mathrm{R}_\alpha^{-1}(P_h x_0)P_h \int_0^1 [F'(P_h x_0) - F'(x_{n-1,\alpha}^{h,\delta} + t(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}))](y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta})\mathrm{d}t \\ &= \mathrm{R}_\alpha^{-1}(P_h x_0)P_h F'(P_h x_0)\int_0^1 \phi((P_h x_0, x_{n-1,\alpha}^{h,\delta} + t(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}), y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta})\mathrm{d}t \end{aligned}$$

and hence $\left\|x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}\right\| \le \left(1 + \frac{\epsilon_0}{\alpha_0}\right)k_0 r\left\|y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}\right\|.$ (10)

□

Again observe that if $x_{n,\alpha}^{h,\delta}, y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$ then by Assumption 1, Lemma 1, and (10) we have

$$\begin{aligned} y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta} &= x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta} - \mathrm{R}_\alpha^{-1}(P_h x_0)P_h\left[F\left(x_{n,\alpha}^{h,\delta}\right) - F\left(y_{n-1,\alpha}^{h,\delta}\right) + \alpha(x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta})\right] \\ &= \mathrm{R}_\alpha^{-1}(P_h x_0)P_h \int_0^1 [F'(P_h x_0) - F'(y_{n-1,\alpha}^{h,\delta} + t(x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}))](x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta})\mathrm{d}t \\ &= \mathrm{R}_\alpha^{-1}(P_h x_0)P_h F'(P_h x_0)\int_0^1 \phi((P_h x_0, y_{n-1,\alpha}^{h,\delta} + t(x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}), x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta})\mathrm{d}t \end{aligned}$$

$$\begin{aligned} \text{and hence} \left\|y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\right\| &\le \left(1 + \frac{\epsilon_0}{\alpha_0}\right)k_0 r\left\|x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}\right\| \\ &\le q^2\left\|y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}\right\|. \end{aligned}$$ (11)

Thus, if $x_{n,\alpha}^{h,\delta}, y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$ then (a) and (b) follow from (10) and (11), respectively. And (c) follows by Lemma 2 and (11). Now using induction, we shall prove that $x_{n,\alpha}^{h,\delta}, y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$. Note that $x_{0,\alpha}^{h,\delta}, y_{0,\alpha}^{h,\delta} \in B_r(P_h x_0)$ and hence by (10)

$$\begin{aligned} \left\|x_{1,\alpha}^{h,\delta} - P_h x_0\right\| &\le \left\|x_{1,\alpha}^{h,\delta} - y_{0,\alpha}^{h,\delta}\right\| + \left\|y_{0,\alpha}^{h,\delta} - P_h x_0\right\| \\ &\le (1+q)e_0 \\ &\le \frac{\gamma_\rho}{(1-q)} < r. \end{aligned}$$

that is, $x_{1,\alpha}^{h,\delta} \in B_r(P_h x_0)$, again by (11)

$$\begin{aligned}\left\|y_{1,\alpha}^{h,\delta} - P_h x_0\right\| &\le \left\|y_{1,\alpha}^{h,\delta} - x_{1,\alpha}^{h,\delta}\right\| + \left\|x_{1,\alpha}^{h,\delta} - P_h x_0\right\| \\ &\le q^2 e_0 + (1+q)e_0 \\ &\le \frac{\gamma_\rho}{(1-q)} < r.\end{aligned}$$

that is, $y_{1,\alpha}^{h,\delta} \in B_r(P_h x_0)$. Suppose $x_{k,\alpha}^{h,\delta}$, $y_{k,\alpha}^{h,\delta} \in B_r(P_h x_0)$ for $k > 1$. Then since

$$\left\|x_{k+1,\alpha}^{h,\delta} - P_h x_0\right\| \le \left\|x_{k+1,\alpha}^{h,\delta} - x_{k,\alpha}^{h,\delta}\right\| + \left\|x_{k,\alpha}^{h,\delta} - x_{k-1,\alpha}^{h,\delta}\right\| + \cdots + \left\|x_{1,\alpha}^{h,\delta} - P_h x_0\right\| \quad (12)$$

We shall first find an estimate for $\left\|x_{k+1,\alpha}^{h,\delta} - x_{k,\alpha}^{h,\delta}\right\|$. Note that by (a) and (b), we have

$$\begin{aligned}\left\|x_{k+1,\alpha}^{h,\delta} - x_{k,\alpha}^{h,\delta}\right\| &\le \left\|x_{k+1,\alpha}^{h,\delta} - y_{k,\alpha}^{h,\delta}\right\| + \left\|y_{k,\alpha}^{h,\delta} - x_{k,\alpha}^{h,\delta}\right\| \\ &\le (1+q)\left\|y_{k,\alpha}^{h,\delta} - x_{k,\alpha}^{h,\delta}\right\| \\ &\le (1+\mathrm{q})\mathrm{q}^{2\mathrm{n}}\mathrm{e}_0\end{aligned}$$

Therefore by (12), we have

$$\begin{aligned}\left\|x_{k+1,\alpha}^{h,\delta} - P_h x_0\right\| &\le (1+q)[q^{2k} + q^{2(k-1)} + \cdots + 1]e_0 \\ &\le (1+q)\frac{1-q^{2k+1}}{1-q^2}e_0 \\ &\le \frac{\gamma_\rho}{(1-q)} < r.\end{aligned} \quad (13)$$

So by induction $x_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$ for all $n \ge 0$. Again by (a) and (b), we have

$$\begin{aligned}\left\|y_{k+1,\alpha}^{h,\delta} - P_h x_0\right\| &\le \left\|y_{k+1,\alpha}^{h,\delta} - x_{k+1,\alpha}^{h,\delta}\right\| + \left\|x_{k+1,\alpha}^{h,\delta} - P_h x_0\right\| \\ &\le q^{2(k+1)}e_0 + (1+q)[q^{2k} + q^{2(k-1)} + \cdots + 1]e_0 \\ &\le (1+q)\frac{1-q^{2k+3}}{1-q^2}e_0 \\ &\le \frac{\gamma_\rho}{(1-q)} < r.\end{aligned}$$

Thus, by induction $y_{n,\alpha}^{h,\delta} \in B_r(P_h x_0)$ for all $n \ge 0$.

The main result of this section is the following theorem.

Theorem 2 *Let* $y_{n,\alpha}^{h,\delta}, x_{n,\alpha}^{h,\delta}$ *and* $e_{n,\alpha}^{h,\delta}$ *be as in* (3), (4), *and* (5), *respectively, and assumptions of Theorem 1 hold. Then* $\left(x_{n,\alpha}^{h,\delta}\right)$ *is a Cauchy sequnce* in $B_r(P_h x_0)$. *and*

converges to $x_\alpha^{h,\delta} \in B_r(\bar{P}_h x_0)$. *Further,* $P_h[F(x_\alpha^{h,\delta}) + \alpha(x_\alpha^{h,\delta} - x_0)] = P_h f^\delta$ *and* $\left\|x_{n,\alpha}^{h,\delta} - x_\alpha^{h,\delta}\right\| \le C\, q^{2n}$, *where* $C = \frac{\gamma_\rho}{1-q}$.

Proof Using the relation (b) and (c) of Theorem 1, we obtain

$$\begin{aligned}\left\|x_{n+m,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\right\| &\le \sum_{i=0}^{m-1}\left\|x_{n+i+1,\alpha}^{h,\delta} - x_{n+i,\alpha}^{h,\delta}\right\| \\ &\le \sum_{i=0}^{m-1}(1+q)q^{2(n+1)}e_0 \\ &\le \mathrm{C}\ \mathrm{q}^{2\mathrm{n}}\end{aligned}$$

Thus $\left(x_{n,\alpha}^{h,\delta}\right)$ is a Cauchy sequnce in $B_r(P_h x_0)$ and hence, it converges, say to $x_\alpha^{h,\delta} \in B_r\left(\overline{P_h x_0}\right)$. Observe that

$$\begin{aligned}\left\|P_h\left[F\left(x_{n,\alpha}^{h,\delta}\right) - f^\delta + \alpha\left(x_{n,\alpha}^{h,\delta} - x_0\right)\right]\right\| &= R_\alpha\left\|\left(x_{0,\alpha}^{h,\delta}\right)\left(y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\right)\right\| \\ &= \left\|P_h F'(x_{0,\alpha}^{h,\delta})P_h + \alpha P_h\right\| e_{n,\alpha}^{h,\delta} \qquad (14) \\ &\le (C_F + \alpha)q^{2n}\gamma_\rho\end{aligned}$$

Now by letting $n \to \infty$ in (14), we obtain

$$P_h[F(x_\alpha^{h,\delta}) + \alpha(x_\alpha^{h,\delta} - x_0)] = P_h f^\delta \qquad (15)$$

4 Error Bounds Under Source Conditions

The objective of this section is to obtain an error estimate for $\left\|x_{n,\alpha}^{h,\delta} - \overline{x}\right\|$ using the Assumption 2.

Proposition 1 (cf. [2], Proposition 1) $F : D(F) \subseteq X \to X$ *is a nonlinear monotone operator in X. Let* $x_\alpha^{h,\delta}$ *the solution of* (15) *and* $x_\alpha^\delta := x_\alpha^{h,0}$. *Then* $\left\|x_{n,\alpha}^{h,\delta} - \overline{x}\right\| \le \frac{\delta}{\alpha}$.

Theorem 3 (cf. [2], Theorem 4) *Let* $\rho < \frac{2}{k_0(1+\frac{\epsilon_0}{\alpha_0})}$. *Assumptions 1 and assumptions in Proposition 1 be satisfied. Then* $\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| \le \bar{\mathrm{C}}\left(\varphi(\alpha) + \frac{\epsilon_h}{\alpha}\right)$ *where* $\bar{\mathrm{C}} = \frac{\max(1,\rho+\|\bar{x}\|)}{1-\frac{k_0}{2}(1+\frac{\epsilon_0}{\alpha_0})\rho}$.

Theorem 4 (cf. [2], Theorem 4) *Let $x_{n,\alpha}^{h,\delta}$ be as in (4), the assumptions in Theorem 2 and Theorem 3 hold. Then* $\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| \leq Cq^{2n} + \max\{1, \bar{C}\}\left(\varphi(\alpha) + \frac{\delta+\in_h}{\alpha}\right)$

Proof Observe that $\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| \leq \left\|x_{n,\alpha}^{h,\delta} - x_{\alpha}^{h,\delta}\right\| + \left\|x_{\alpha}^{h,\delta} - x_{\alpha}^{\delta}\right\| + \left\|x_{\alpha}^{\delta} - \bar{x}\right\|$

So by Proposition 1, Theorem 2, and Theorem 3, we obtain

$$\begin{aligned}\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| &\leq \mathrm{C}\ \mathrm{q}^{2\mathrm{n}} + \frac{\delta}{\alpha} + \bar{\mathrm{C}}\left(\varphi(\alpha) + \frac{\in_h}{\alpha}\right) \\ &\leq \mathrm{C}\ \mathrm{q}^{2\mathrm{n}} + \max\{1, \overline{\mathrm{C}}\}\left(\varphi(\alpha) + \frac{\delta + \in_{\mathrm{h}}}{\alpha}\right)\end{aligned}$$

$$\text{Let} \quad \mathrm{n}_\delta = \min\left\{n : q^{2\mathrm{n}} \leq \frac{\delta + \in_h}{\alpha}\right\} \tag{16}$$

$$\text{and} \quad C_o = C + \max\{1, \overline{\mathrm{C}}\}. \tag{17}$$

□

Theorem 5 *Let $x_{n,\alpha}^{h,\delta}$ be as in (4), the assumptions in Theorem 4 be satisfied. Let n_δ and C_o be as in (16) and (17), respectively. Then*

$$\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| \leq C_o\left(\varphi(\alpha) + \frac{\delta + \in_h}{\alpha}\right). \tag{18}$$

4.1 A Priori Choice of the Parameter

Note that error estimate $\left(\varphi(\alpha) + \frac{\delta+\in_h}{\alpha}\right)$ in (18) is of optimal order if $\alpha_\delta := \alpha(\delta, h)$ satisfies, $\varphi(\alpha_\delta)\alpha_\delta = \delta + \in_h$.

Now as in [3], using the relation $\psi(\lambda) := \lambda\varphi^{-1}(\lambda), 0 < \lambda < a$, we have $\delta + \in_{\mathrm{h}} = \varphi(\alpha_\delta)\alpha_\delta = \psi(\varphi(\alpha_\delta))$, so that $\alpha_\delta = \varphi^{-1}(\psi^{-1}(\delta + \in_{\mathrm{h}}))$.

Theorem 6 *Let $\psi(\lambda) := \lambda\varphi^{-1}(\lambda)$, for, $0 < \lambda < a$ and assumptions in Theorem 5 hold. For $\delta > 0$, let $\alpha_\delta = \varphi^{-1}\left(\psi^{-1}(\delta + \in_h)\right)$ and let n_δ be as in (16). Then* $\left\|x_{n,\alpha}^{h,\delta} - \bar{x}\right\| = \mathcal{O}(\psi^{-1}(\delta + \in_h))$

4.2 An Adaptive Choice of the Parameter

In this subsection, we present a parameter choice rule based on the balancing principle studied in [6, 9]. In this method, the regularization parameter α is selected from some finite $D_N(\alpha) = \alpha_i = \{\mu^i\alpha_0, i = 0, 1, 2, \ldots, N\}$ where $\mu > 1$, $\alpha_0 > 0$ and

let $n_i = \min\{n : q^{2n} \leq \frac{\delta+\epsilon_h}{\alpha}\}$. Then, for $i = 0, 1, 2, \ldots, N$, we have $\left\| x^{h,\delta}_{n_i,\alpha_i} - x^{h,\delta}_{n_i,\alpha_i} \right\| \leq C\left(\frac{\delta+\epsilon_h}{\alpha_i}\right)$.

Let $x_i = x^{h,\delta}_{n_i\alpha_i}$, $i = 0, 1, 2, \ldots, N$. Proof of the following is analogous to the proof of Theorem 3.1 in [10].

Theorem 7 (c.f. [10], Theorem 3.1) *Assume that* $\exists\, i \in \{0, 1, 2, \ldots N\}$ *such that* $\varphi(\alpha_i) \leq \frac{\delta+\epsilon_h}{\alpha_i}$. *Let assumptions of Theorem 5 and Theorem 6 hold and let* $l := \max\left\{i : \varphi(\alpha_i) \leq \frac{\delta+\epsilon_h}{\alpha_i}\right\} < N, k := \max\left\{i : \left\|x_i - x_j\right\| \leq 4C_o\left(\frac{\delta+\epsilon_h}{\alpha_j}\right), j = 0, 1, 2, \ldots, i\right\}$. *Then* $l \leq k$ *and* $\|\overline{\mathrm{x}} - \mathrm{x_k}\| \leq \mathrm{c}\ \psi^{-1}(\delta + \epsilon_h)$ *where* $c = 6C_o\mu$.

5 Implementation of Adaptive Choice Rule

The following steps are involved in implementing the adaptive choice rule:

- Choose $\alpha_0 > 0$such that $\delta_0 < \frac{\alpha_0}{4k_0(1+\frac{\epsilon_0}{\alpha_0})}$ and $\mu > 1$.
- Choose $\alpha_i = \mu^i\alpha_0, i = 0, 1, 2, \ldots, N$.

Finally, the adaptive algorithm associated with the choice of the parameter specified in Theorem 7 involves the following steps.

5.1 Algorithm

1. Set i = 0.
2. Choose $n_i = \min\{n : q^{2n} \leq \frac{\delta+\epsilon_h}{\alpha}\}$.
3. Solve $x_i = x^{h,\delta}_{n_i\alpha_i,}$ by using the iteration (3) and (4).
4. If $\left\|x_i - x_j\right\| \geq 4C_o\left(\frac{\delta+\epsilon_h}{\alpha_j}\right), j < i$, then take $k = i - 1$ and return$\mathrm{x_k}$.
5. Else set $i = i + 1$ and go to step 2.

6 Numerical Example

We apply the algorithm by choosing a sequence of finite-dimensional subspace (V_n) of X with $\dim(V_n) = n + 1$. Precisely we choose V_n as the linear span of $\{v_1, v_2, \ldots v_{n+1}\}$ where $v_i, i = 1, 2, ..n + 1$ are the linear splines in a uniform grid of $n + 1$ points in $[0, 1]$. Note that $x^{h,\delta}_{n,\alpha}, y^{h,\delta}_{n,\alpha} \in V_n$, so $y^{h,\delta}_{n,\alpha} = \sum_{i=1}^{n+1} \xi^n_i v_i$ $x^{h,\delta}_{n,\alpha} = \sum_{i=1}^{n+1} \eta^n_i v_i$, where ξ^n_i and η^n_i are some scalars. Then, from (3) we have

$$(P_hF'(P_hx_0)P_h + \alpha P_h)(y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}) = P_h\left[f^\delta - F\left(x_{n,\alpha}^{h,\delta}\right) + \alpha\left(x_{0,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\right)\right] \quad (19)$$

Observe that $(y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta})$ is a solution of (19) if and only if $\overline{(\xi^n - \eta^n)} = (\xi_1^n - \eta_1^n, \xi_2^n - \eta_2^n, \ldots, \xi_{n+1}^n - \eta_{n+1}^n)^T$ is the unique solution of

$$(Q_n + \alpha B_n)\overline{(\xi^n - \eta^n)} = B_n[\overline{\mu^n} - F_{h1} + \alpha(X_0 - \overline{\eta^n}], \text{where} \quad (20)$$

$$\begin{aligned}
Q_n &= \left[\left\langle F'\left(x_{0,\alpha}^{h,\delta}\right)v_i, v_j\right\rangle\right], \quad i,\ j = 1, 2, \ldots n+1\\
B_n &= \left[\langle v_i, v_j\rangle\right], \quad i,j = 1, 2, \ldots n+1\\
\overline{\mu^n} &= \left[f^\delta(t_1), f^\delta(t_2), \ldots f^\delta(t_{n+1})\right]^T\\
F_{h1} &= \left[F\left(x_{n,\alpha}^{h,\delta}\right)(t_1), F\left(x_{n,\alpha}^{h,\delta}\right)(t_2), \ldots .F\left(x_{n,\alpha}^{h,\delta}\right)(t_{n+1})\right]^T\\
X_0 &= \left[x_0(t_1), x_0(t_2), \ldots x_0(t_{n+1})\right]^T
\end{aligned}$$

and $t_1, t_2, \ldots t_{n+1}$ are grid points. Further from (4), it follows that

$$(P_hF'(P_hx_0)P_h + \alpha P_h)(x_{n+1,\alpha}^{h,\delta} - y_{n,\alpha}^{h,\delta}) = P_h\left[f^\delta - F\left(y_{n,\alpha}^{h,\delta}\right) + \alpha\left(x_{0,\alpha}^{h,\delta} - y_{n,\alpha}^{h,\delta}\right)\right] \quad (21)$$

and hence $(y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta})$ is a solution of (21) if and only if $\overline{(\eta^{n+1} - \xi^n)} = \left(\eta_1^{n+1} - \xi_1^n, \eta_2^{n+1} - \xi_2^n, \ldots \eta_{n+1}^{n+1} - \xi_{n+1}^n\right)^T$ is the unique solution of

$$(Q_n + \alpha B_n)\overline{(\eta^{n+1} - \xi^n)} = B_n\left[\overline{\mu^n} - F_{h2} + \alpha\left(X_0 - \overline{\xi^n}\right)\right], \quad (22)$$

where $F_{h2} = \left[F\left(y_{n,\alpha}^{h,\delta}\right)(t_1), F\left(y_{n,\alpha}^{h,\delta}\right)(t_2), \ldots F\left(y_{n,\alpha}^{h,\delta}\right)(t_{n+1})\right]^T$. Note that (20) and (22) are uniquely solvable as Q_n is a positive definite matrix (i.e., $xQ_nx^T > 0$ for all nonzero vector x) and B_n is an invertible matrix.

Example 1 (see, [10], Sect. 4.3) Let $\mathrm{F} : \mathrm{D}(F) \subseteq \mathrm{L}^2(0,1) \to \mathrm{L}^2(0,1)$ defined by $\mathrm{F(u)} := \int_0^1 \mathrm{k(s,t)u^3(s)ds}$, where $\mathrm{k(s,t)} = \left\{\begin{array}{ll}(1-\mathrm{t})\mathrm{s}, & 0 \le \mathrm{s} \le \mathrm{t} \le 1\\ (1-\mathrm{s})\mathrm{t}, & 0 \le \mathrm{t} \le \mathrm{s} \le 1\end{array}\right\}$. The Fréchet derivative of F is given by

$$F'(u) = 3\int_0^1 k(s,t)u^2(s)\mathrm{d}s \quad (23)$$

□

In our computation, we take $f(t) = \frac{6\sin\pi t + \sin^3\pi t}{9\pi^2}$ and $f^\delta = f + \delta$. Then, the exact solution $\bar{x}(t) = \sin\pi t$. We use $x_0 = \sin\pi t + \frac{3(t\pi^2 - t^2\pi + \sin^2\pi t}{4\pi^2}$ as our initial guess, so that $x_0 - \bar{x}$ satisfies the source condition $x_0 - \bar{x} = \varphi(F'(\bar{x}))\frac{1}{4}$ where $\varphi(\lambda) = \lambda$.

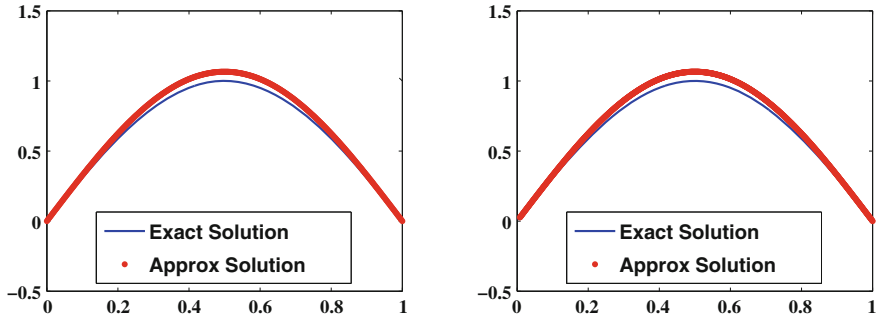


Fig. 1 2D visualization of the plot

For the operator $F'(.)$ defined in (23), $\in_n = \mathcal{O}(n^{-2})$ (c.f. [4]). Thus, we expect to obtain the rate of convergence $\mathcal{O}(\delta + \varepsilon_n)^{1/2}$. We choose $\alpha_0 = 1.5\delta, \mu = 1.5$ and $q = 0.51$. The exact solution and the approximate solution $x_{n,\alpha}^{h,\delta}$ obtained for $n = 2^i, i = 9, 10$ are given in Fig. 1.

7 Conclusion

In this paper, we considered the finite-dimensional approximation of the iterative method considered in [1] for obtaining an approximate solution (signal) for a nonlinear ill-posed operator equation $F(X) = f$ where $F : D(F) \subseteq X \to X$ is a monotone operator defined on a real Hilbert space X when the available data (signal) f^δ with $f - f^\delta \leq \delta$. The proposed method has a local linear convergence to the actual solution $\bar{x}$ for properly chosen parameter α. The choice of the regularization parameter is made according to the adaptive method considered by Pereverzev and Schock (2005) and obtained an optimal-order error estimate. A numerical example and the corresponding computational results are exhibited to confirm the reliability and effectiveness of our method.

Acknowledgments S.Pareth thanks National Institute of Technology Karnataka, India, for the financial support.

References

1. George S, Pareth S (2012) Two step modified Newton method for nonlinear Lavrentiev regularization, ISRN applied mathematics, 2012, Article ID 728627, doi:10.5402/2012/728627
2. George S, Pareth S (2012) An application of Newton type iterative method for Lavrentiev regularization for Ill-posed equations: finite dimensional realization. IAENG Int J Appl Math 42(3):164–170

3. George S, Nair MT (2008) A modified Newton-Lavrentiev regularization for nonlinear ill-posed hammerstein-type operator equation. J Complex 24:228–240
4. Groetsch CW, King JT, Murio D (1982) Asymptotic analysis of a finite element method for Fredholm equations of the first kind. In: Baker CTH, Miller GF (eds) Treatment of integral equations by numerical methods. Academic Press, London, pp 1–11
5. Jaan J, Tautenhahn U (2003) On Lavrentiev regularization for ill-posed problems in Hilbert scales. Numer Funct Anal Optim 24(5–6):531–555
6. Mathe P, Perverzev SV (2003) Geometry of linear ill-posed problems in variable Hilbert scales. Inverse Prob 19(3):789–803
7. Nair MT, Ravishankar P (2008) Regularized versions of continuous Newton's method and continuous modified Newton's method under general source conditions. Numer Funct Anal Optim 29(9–10):1140–1165
8. Lu S, Pereverzyev SV (2008) Sparsity reconstruction by the standard Tikhonov method, RICAM-Report No. 2008-17
9. Perverzev SV, Schock E (2005) On the adaptive selection of the parameter in regularization of ill-posed problems. SIAM J Numer Anal 43:2060–2076
10. Semenova EV (2010) Lavrentiev regularization and balancing principle for solving ill-posed problems with monotone operators. Comput Methods Appl Math 4:444–454
11. Tautanhahn U (2002) On the method of Lavrentiev regularization for nonlinear ill-posed problems. Inverse Prob 18:191–207

Video Signals Demodulator for Satellite Communication

P. Srividya, K. R. Nataraj and K. R. Rekha

Abstract This paper proposes a VLSI architecture to design a QPSK demodulator of video signals for satellite communication using baseband processing. The demodulator mainly consists of a multiplier, a root-raised cosine (RRC) filter designed using distributed arithmetic (DA) technique, and an adder to add I-channel and Q-channel signals. The main advantage of this architecture is that it does not employ any MAC unit, whose operational speed is, generally, a bottleneck for high filter throughput. Instead, it makes extensive use of LUTs and hence is ideally suited for FPGA implementation. The VHDL hardware description language is used to code the entire model. The modulator and demodulator have been coded in MATLAB in order to validate the simulation results. The simulated and MATLAB results compare favorably.

Keywords QPSK modulator and demodulator · A/D conversion · Video signals · MATLAB · Xilinx · VHDL and IP core generator

P. Srividya (✉)
Research Scholar, Visvesvaraya Technological University, Belgaum, Karnataka, India
e-mail: psrividya@sjbit.edu.in

K. R. Nataraj · K. R. Rekha
Department of Electronics and Communication Engineering, SJBIT, Bangalore, Karnataka, India
e-mail: krnataraj@sjbit.edu.in

K. R. Rekha
e-mail: krrekha@sjbit.edu.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_11,

1 Introduction

A communication satellite is basically an electronic communication package placed in orbit whose prime objective is to initiate or assist communication transmission of information or message from one point to another through space, covering a wide area over the surface of the earth. The information transferred most often corresponds to voice (telephone), video (television), and digital data and can use both analog and digital modulation techniques.

A repeater satellite only changes frequency and power parameters on a digitally modulated signal. It makes no attempt to detect the digital data. That means that if there is distortion on the uplink signal, this distortion will be amplified and translated in frequency to the downlink, thus lowering the quality of the signal even more. This has a negative impact on the received downlink signal power and total C/N ratio and results in inferior bit error rate.

To overcome these limiting factors on basic bent-pipe satellites, new types of transponders were conceived, which allow a number of different approaches to avoid these problems. These types of transponders are used on "smart" satellites called onboard processing (OBP) satellites.

Onboard processing is divided roughly into two parts: carrier processing and baseband processing. The basic concept of the carrier processing technique is where the uplink-modulated carriers from different sources are distributed directly to output port (downlink) antennas. By this technique, no attempt is made to recover the baseband data during the processing, and no changes in modulation formats are allowed.

In baseband processing, the uplink carriers are first down-converted and then demodulated to reconstruct the binary information stream to correct any errors that may have happened during the uplink. A new waveform is generated with the clean binary sequence, and after that, the signal is again modulated and up-converted to be transmitted on the downlink, free from any uplink degradation, thus improving the Eb/No ratio and, therefore, the BER at the receiving earth station.

2 Digital Transmission of Video Signals

Digital modulation is the obvious choice for satellite transmission of signals that originate in digital form and that are used by digital devices. Virtually all signals sent via satellites are now digital.

Baseband digital signals are usually represented as serially transmitted logical ones and zeros. In computer circuitry, a logical zero may be represented by a low voltage and a logical one by a higher voltage. This arrangement is inconvenient for transmission over any significant distance and is not used. All circuits that carry the signal must have a frequency response that extends to DC, and this is difficult to achieve since many communication circuits contain transformers. To avoid this

problem, digital modulators usually accept their input in a polar non-return-to-zero (NRZ) format: Logical ones and zeros are transmitted as plus or minus a stated value.

Satellite links always carry RF signals, which require the data to be modulated onto a radio frequency carrier for transmission. The most popular choice of modulation for a digital satellite communication is phase-shift keying (PSK).

Band-pass (or radiofrequency) transmission of digital data differs from base-band transmission only because modulation from an RF wave is required. The receiver demodulates the modulated RF wave to recover the baseband data stream. Thus, intersymbol interference will occur at the receiver due to band-limiting of the modulated waveforms, unless filters that satisfy the Nyquist criterion are used.

3 QPSK Technique

In QPSK, the phase ϕ of the carrier can take on one of the four values: 45°, 135°, 225°, or 315°. The angular separation between any two adjacent phasors of a QPSK is 90°. QPSK modulator block diagram is shown in Fig. 1. The incoming serial bit stream enters the bit splitter (demultiplexer), where it is converted to a parallel, two-channel output: the I (in-phase)-channel and the Q (in-quadrature)-channel. Consequently, the data (bit) rate, Rb, in each of the channels is Rb/2. The data in each channel are fed to the channel's balanced modulator with input carrier's relative phase of 0° and 90°, respectively. The outputs of the product modulators are fed to the bit-combining circuit (linear summer), where they are converted from the I- and Q-data channels to a single binary output data stream.

The demodulation technique reverses the modulation process by converting the analog baseband signals back into a series of digital pulses that the terminal device can use. In general, a coherent demodulator with a decoder is employed (Fig. 2) in receiving systems for digital satellite communications.

The decoding operation involves reconstructing the data bit sequence encoded onto the carrier for which phase-coherent decoding is used. It requires the decoder

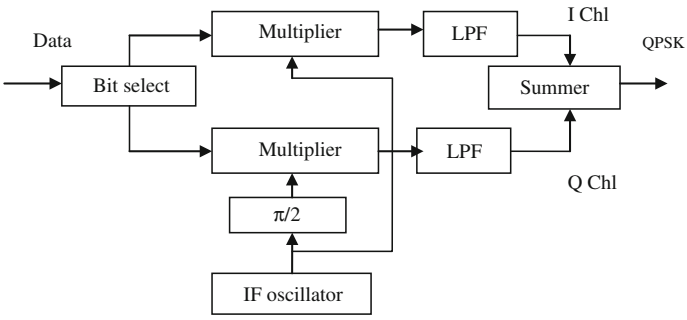


Fig. 1 QPSK modulator

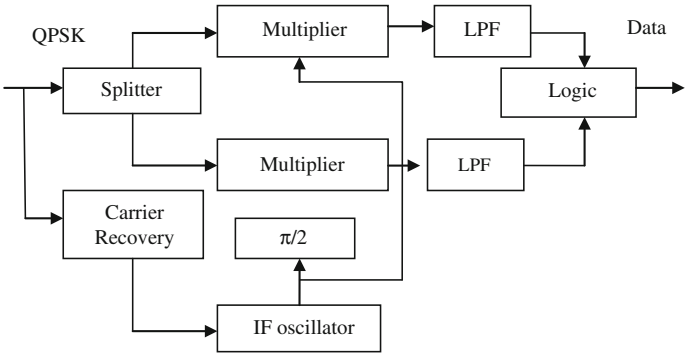


Fig. 2 QPSK demodulator

to use a referenced carrier at the same frequency and phase as the received modulated carrier during each bit time.

4 Distributed Arithmetic–Based Filters

FIR filters are most widely used in FPGA implementations owing to its linear phase property. When compared with IIR filters, FIR filters have simple and regular structures which are easy to implement on hardware. However, FIR filters require higher number of taps compared to IIR filters for achieving the same frequency specification.

Distributed arithmetic is a popular architecture for implementing FIR filters without the use of multipliers. DA realizes the sum-of-products computation required for FIR filters efficiently using LUTs, shifters, and adders. Since these operations map efficiently onto an FPGA, DA is a favoured architecture on these devices. DA relies on the fact that the filter coefficients, c[n], are known, so multiplying c[n]x[n] becomes a multiplication with a constant. This is a pre-requisite for a DA design.

5 Implementation Steps

1. Selecting the baseband signal frequency for video signal (4 MHz).
2. Selecting the sampling frequency (8 MHz).
3. Sampling the input signal.
4. Quantizing the input signal.
5. Encoding the quantized signal. For each sample, 4 bits are used for encoding.
6. Obtain QPSK-modulated signal in MATLAB.

7. The QPSK-modulated wave is stored as ROM input and is used in VHDL coding. This is accomplished using IP core generator of Xilinx.
8. The multiplier, DA-based FIR filter, and the adder are designed in VHDL.

6 Simulation Results

Figure 3 shows the original 4 MHz signal and the quantized signal. Figure 4 shows the input binary data after analog-to-digital conversion.

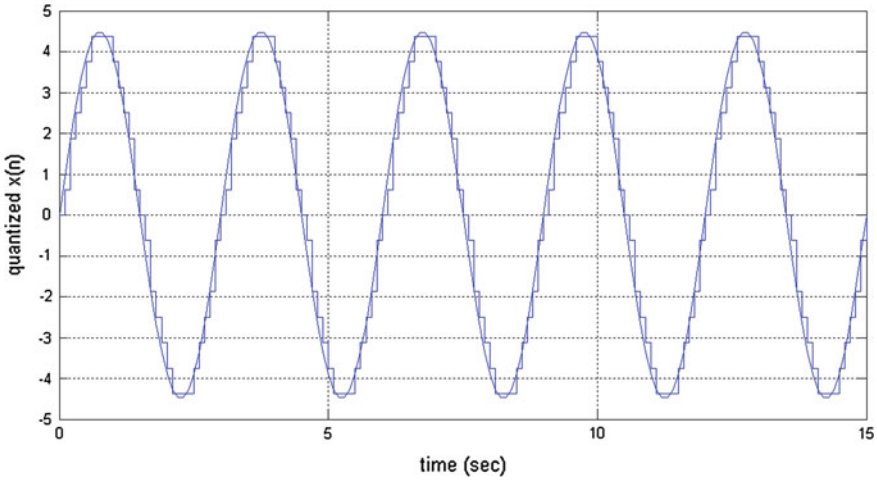


Fig. 3 Input signal and the quantized signal

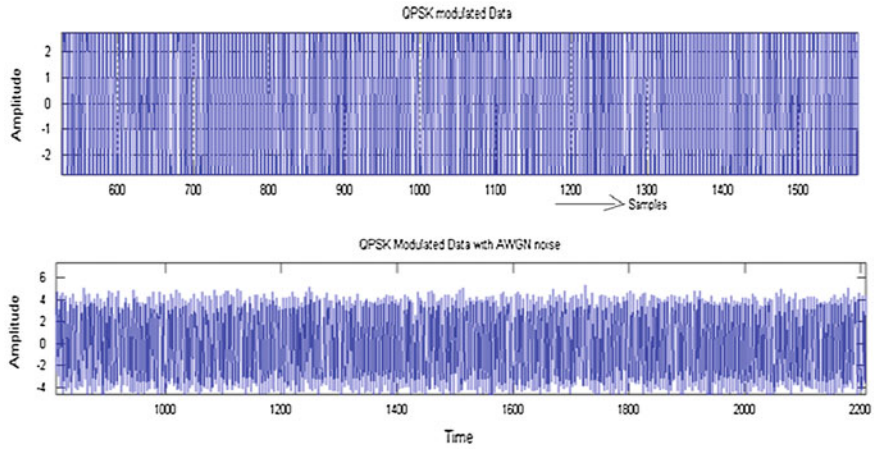


Fig. 4 Modulated wave and the uplink signal to the satellite obtained in MATLAB

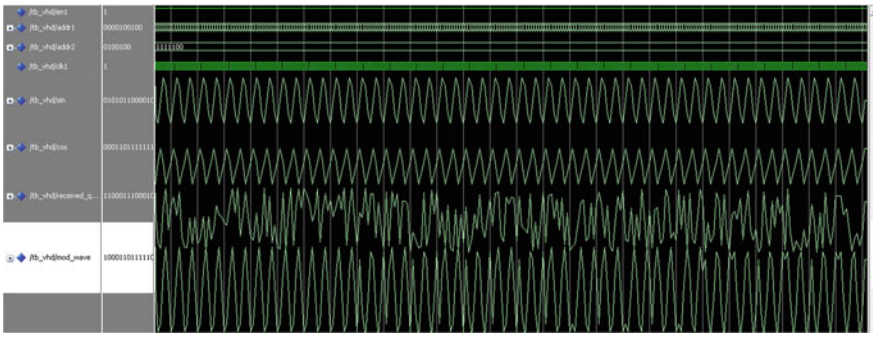


Fig. 5 Modelsim output of cosine and sine carrier waves, uplink signal, and modulated wave

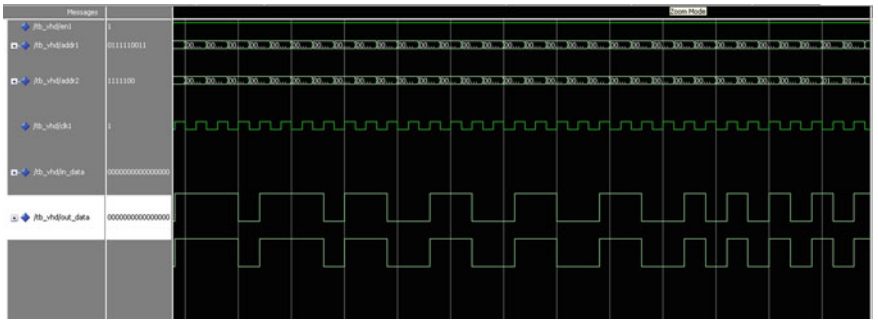


Fig. 6 The input signal and the demodulator output signals obtained using Modelsim

7 Conclusion

A QPSK modulator is implemented in MATLAB. The modulator generates QPSK signal for video signals of different frequencies. The input frequency of 4 MHz and carrier frequency of 12 MHz are considered to generate the QPSK signal samples which are then converted to hexadecimal and used as ROM input to the demodulator. The demodulator is coded using VHDL, and the output is verified using Modelsim simulator (Fig. 5, 6).

References

1. Proakis J (1995) Digital Communications, 3rd edn. McGraw-Hill, International Edition
2. Amir M, Sodagar S (2000) A novel architecture for ROM less sine – output direct digital frequency synthesizers by using 2 m order parabolic approximation, IEEE IA International frequency control symposium, pp 284–289

3. Allred DJ, Huang W, Krishnan V, Yoo H, Anderson DV (2004) An FPGA implementation for a high throughput adaptive filter using distributed arithmetic, IEEE symposium on field programmable custom computing machines IEEE computer society, pp 324–325
4. Longa P, Miri A (2006) Area efficient FIR filter design on FPGAs using distributed arithmetic. In: IEEE international symposium on signal processing and information technology. pp 248–252
5. Sen W, Bin T (2007) Distributed arithmetic for FIR filter design on FPGA, ICCCAS, pp 620–623
6. Eshtawie MAM, Othman MB (2007) An algorithm proposed for FIR filter coefficients representation, World Academy of Science, Engineering and Technology. pp 57–63
7. Shantala SY (2009) Kulkarani: high speed and low power FPGA implementation of FIR filter for DSP applications. Eur J Sci Res 31:19–28
8. Nataraj KR, Ramachandran S, Nagabushan BS (2009) Development of algorithm for demodulator for processing satellite data communication, Int J Comput Sci Netw Secur 9(6):233–243
9. Nataraj KR, Ramachandran S, Nagabushan BS (2009) Development of algorithm, architecture and FPGA implementation of demodulator for processing satellite data communication, Int J Comput Sci Netw Secur 9:137–147
10. Kolawole MO (2002) Satellite communication engineering, Marcel Dekker, Inc
11. Sakla T, Jain D, Gautham S (2010) Implementation of digital QPSK modulator by using VHDL/Matlab. Int J Eng Sci Technol 2(9):4827–4831

Importance of On-Demand-Modified Power-Aware DSR in MANETs

Shivashankar, G. Jayanthi, G. Varaprasad, Khurram J. Mohammed and G. Devaraju

Abstract A recent trend in ad hoc network routing is the reactive on-demand philosophy where routes are established only when required without power-aware routing. This paper proposes a new scheme called modified power routing DSR (MPRDSR) to improve existing on-demand routing protocols by introducing the Power efficient scheme in whole mobile ad hoc network (MANET). The scheme establishes quick adaptation to distributed processing, dynamic linking, and low processing at all times. This scheme uses the concept of power awareness among route selection nodes by checking power status of each node in the topology which insures fast selection of routes with minimal efforts and faster recovery. In route discovery phase, EPRDSR selects the bandwidth and energy constraints are built into the DSR route discovery mechanism. The results indicate that EPRDSR is able to discover the required path with lesser overheads, the network lifetime increased by around 60–65 %, the packet delivery ratio improved and the packet experienced a low average delay. Moreover, the local repair mechanism was able to find an alternate path in most of the cases enhanced the network lifetime and delayed the repair and reconstruction of the route. The scheme is incorporated with the DSR protocol and the performance has been studied through simulation and

Shivashankar (✉) · G. Jayanthi · G. Devaraju
Ambedkar Institute of Technology, Bangalore, India
e-mail: nshivu_ss@rediffmail.com

G. Jayanthi
e-mail: jayanthi_g_15@yahoo.co.in

G. Devaraju
e-mail: g_devaraju_05@yahoo.co.in

G. Varaprasad
BMS College of Engineering, Bangalore, India
e-mail: drvaraprasad@gmail.com

K. J. Mohammed
Ghousia College of Engineering, Bangalore, India
e-mail: khurramashrafi@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_12, © Springer India 2014

scheme performs very well. The main goal of EPRDSR is not only to extend the lifetime of each node, but also to prolong the lifetime of each connection. Using the ns-2 simulator, we compared EPRDSR against the minimum total transmission power routing (MTPR).

Keywords MANET · EPRDSR · Network lifetime · Packet delivery ratio

1 Problem Definition and Introduction

The rapid evolution in the field of mobile computing is driving a new alternative way for mobile communication, in which mobile devices form a self-creating, self-organizing, and self-administering wireless network, called a *mobile ad hoc network*. Its intrinsic flexibility, lack of infrastructure, ease of deployment, auto-configuration, low cost, and potential applications make it an essential part of future pervasive computing environments. As a consequence, the seamless integration of mobile ad hoc networks with other wireless networks and fixed infrastructures will be an essential part of the evolution toward future fourth generation communication networks. From a technological point of view, the realization of this vision still requires a large number of challenges to be solved related to devices, protocols, applications, and services. The concise discussion in this paper shows that despite the large efforts of the MANET research community and the rapid progress.

The nodes in an ad hoc network are constrained by battery power for their operation. To route a packet from a source to a destination involves a sufficient number of intermediate nodes. Hence, battery power of a node is a precious resource that must be used efficiently in order to avoid early termination of a node or a network. Thus, energy management is an important issue in such networks. Efficient battery management, transmission power management [1], and system power management are the major means of increasing the life of a node.

A few reasons for energy management in MANETs are limited energy of the nodes, difficulties in replacing the batteries, lack of central coordination, constraints on the battery source, selection of optimum transmission power, and channel utilization. Finally at the network layer, issues which are open areas, designing of an efficient routing algorithm increases the network lifetime by selecting an optimal relay node. The prime concern of this paper is to develop an efficient routing protocol for the ad hoc networks which may take care of energy needs and as well as proper handling of real and non-real-time data as per their need.

The following section discusses some of the power efficient routing algorithms. In general, a routing protocol which does not require large tables to be downloaded or greater number of calculations is preferable, also, reducing the amount of data compression that is done before transmission may decrease the communication power by ultimately increases the number of computation tasks. Hence, a balance

must be reached between the number of computation and communication tasks performed by the node, which are contradictory to each other. Many research efforts have been devoted for developing power-aware routing protocols. Transmission power control and load distribution are two approaches to minimize the active communication energy, and sleep/power-down mode is used to minimize energy during inactivity. The primary focus of the above two approaches is to minimize energy consumption of individual node. The load distribution method balances the energy usage among the nodes and maximizes the network lifetime by avoiding overutilized nodes at the time of selecting a routing path. In transmission power control approach, stronger transmission power increases the transmission range and reduces the hop count to the destination, while weaker transmission power makes the topology sparse, which may result in network partitioning and high end-to-end delay due to a larger hop count. The following are the applications of MANET with efficient power-aware routing protocol.

- Search and rescue operations
- Disaster recovery
- Replacement of fixed infrastructure in case of environmental disasters
- Policing and fire fighting
- Supporting doctors and nurses in hospitals
- Military communication and operations
- Vehicular services: road or accident guidance, transmission of road and weather conditions, taxi cab network, and intervehicle networks.

2 Some of the Related Research Work

The MANET topology changes as mobile hosts move about (mobility) or exhaust their battery source (energy depletion.) Furthermore, due to medium access control (MAC layer protocols), multi-path fading, noise, and signal interference, the capacity of a wireless link may be degraded over time. Ad hoc networks require the support of multi-hop communication schemes in order to relay massages through other hosts when there is no direct access between the sending and receiving hosts. This is not only because a "dead" host cannot provide service to the local service requests, but also because a "dead" host cannot relay messages from another host to its destinations, hence, adversely affecting the network connectivity and service completions for other hosts in the network. The power consumption of a mobile host consists of the communication power and computation power. Communication power refers to the power used by a host to communicate with the other mobile hosts. A mobile host also expends computation power to provide service to various requests and execute network algorithms. Many energy-aware routing protocols have been developed to reduce or balance communication power among mobile hosts [2, 3].

Results proved that the number of updates required is smaller than the highest-degree and lowest-id heuristics. Since node weights were varied in each simulation

cycle, computing the cluster-heads becomes very expensive and there are no optimizations on the system parameters such as throughput and power control [4, 5]. The *highest-degree algorithm* is also known as connectivity-based algorithm [6]. This algorithm is based on the degree of nodes assumed to be the number of neighbors of a given node. Whenever the election procedure is needed, nodes broadcast their identifier (ID) which is assumed to be unique in the same network. According to the number of received IDs every node computes its degree and the one having the maximum degree becomes cluster-head. Major drawbacks of this algorithm are the degree of a node changes very frequently, the CHs are not likely to play their role as cluster-heads for very long. In addition, as the number of ordinary nodes in a cluster is increased, the throughput drops and system performance degrades. All these drawbacks occur because this approach does not have any restriction on the upper bound on the number of nodes in a cluster.

Nagpal et al. [7] presents a strong reason for the use of variable transmission range in future mobile ad hoc network, and we believe variable transmission range is more suitable to the needs of mobile ad hoc network, their devices, and applications. In short, variable transmission range is keystone of the design of future mobile ad hoc network. Energy efficiency continues to be a key performance metric as efficient utilization of energy increases the network lifetime and capacity. So efforts are made to reduce the energy consumption in different ways. To utilize the residual battery power of nodes in an efficient manner, researchers have taken variable transmission range [4, 8, 9] of nodes into accounts as follows:

It finds the minimal radius for connectivity and to increase it up to the target one. On this basis, the transmission range is computed; thus, energy consumption is minimized. (ii) The target radius and dominating sets–based (TRDS) protocol gives a unique solution to two different problem minimum energy broadcasting and activity scheduling. In paper [10, 11], a graph search algorithm is proposed. They used variable transmission range to prolong the lifetime of a node and save it from becoming dead. Their simulation results have confirmed the same. Chauhan and Chopra [12] proposed three distributed algorithms for self-adjusting the transmission range of nodes in a wireless sensor network. The main objective is to vary the transmission range in such way that it gives lower energy consumption for sending a packet from source to destination.

3 Design and Implementation of Proposed EPRDSR

3.1 Existing Dynamic Source Routing (DSR) Protocol Mechanism

DSR is a routing protocol designed for MANETs. It is a source routing protocol and composed of two main mechanisms: route discovery and route maintenance. In the process of route discovery, a route request is send from a node *S* to a node

D by broadcast, only when *S* attempts to send a packet to *D* and has no available route to *D* in its route cache. So the node *S* does not always know a route to *D* and the route request proceeds completely on demand to reduce the routing overhead. Intermediate nodes piggyback their ID into the source route included in the route request message and relay that route request by broadcast, if they do not know an available path to *D*. When the route request reaches *D* or some intermediate node which knows the route to *D* (by checking its route cache), a route reply is unicasted to *S* with the complete path from *S* to *D*. Nodes could receive the same route request more than one time, but only the first one will be handled. Node *S* could also receive multiple routing replies. The first arrival route is used immediately. If in the following routing replies, a shorter path from *S* to *D* is included, the new route will be used instead of the old one. If *S* cannot receive a route reply after a period of time, the route request will be resent until a path to *D* is finally discovered. Route maintenance is the mechanism by which node *S* is able to detect, while using a source route to *D*, if the network topology has changed such that it can no longer use its route to *D* because a link along the route no longer works. When route maintenance indicates a source route is broken, *S* can attempt to use any other route it happens to know to *D* or can invoke route discovery again to find a new route for subsequent packets to *D*. Route maintenance is used only when *S* is actually sending packets to *D*.

3.2 Minimum Total Transmission Power Routing Mechanism

In a non-partitioned ad hoc network, there exists at least one path for a node to communicate with any other node. So theoretically, any node can reach any other node through a random forwarding path. However, the power consumption along different paths varies, due to its dependence on variations of distance between directly communicating nodes and noise interference levels. The greater the values these parameters hold, the larger amount of power is demanded to transmit. Minimum total transmission energy, such as MTPR, focuses on end-to-end energy efficiency. Generally, the route selected by conserving energy is the shortest distance path or minimum-hop path. Even though some nodes may be dissipating more energy due to dynamics of link characteristics such as distance or error rate, the end-to-end shortest path naturally leads to conservation of energy in transmission.

The majority of energy-efficient routing protocols for MANET try to reduce energy consumption by means of an energy-efficient routing metric, used in routing table computation instead of the minimum-hop metric. This way, a routing protocol can easily introduce energy efficiency in its packet forwarding. These protocols try either to route data through the path with maximum energy bottleneck or to minimize the end-to-end transmission energy for packets, or a weighted combination of both. A first approach for energy-efficient routing is known as

MTPR. That mechanism uses a simple energy metric, represented by the total energy consumed to forward the information along the route. This way, MTPR reduces the overall transmission power consumed per packet, but it does not directly affect the lifetime of each node (because it does not take into account the available energy of network nodes). However, minimizing transmission energy only differs from shortest hop routing if nodes can adjust transmission power levels, so that multiple short hops are more advantageous, from an energy point of view, than a single long hop. In 802.11, we do not have access to this capability, so that, in a fixed transmission power context, this metric corresponds to a shortest path routing.

3.3 Implementation of Proposed EPRDSR Mechanism

In order to compare DSR to EPRDSR, each node tracks its current energy level using battery power. When a node receives a route reply or an ACK, it updates its energy field in the packet path. Once the source node receives either a route reply or an ACK, it updates its energy table with the energy values in the path. When choosing a path, the DSR implementation chooses the path with the minimum number of hops. For EPRDSR, however, the path is chosen based on energy. First, we calculate the bottleneck energy for each path, that is, the lowest hop energy of the path. The path is then selected by choosing the path with the maximum lowest hop energy. For example, consider the following scenario. There are two paths to choose from. The first path contains three hops with energy values 22, 18, and 100, and the second path contains four hops with energy values 20, 35, 25, and 80. The score for the first path is 18, while the score for the second path is 20. Because 20 is greater than 18, the second path would be chosen.

3.4 Power-Aware Model in EPRDSR

The proposed algorithm maximizes the network lifetime and minimizes the power consumption during the source to destination route establishment. This algorithm takes special care to transfer both real-time and non-real-time traffic by providing energy-efficient and less congested path between a source and destination pair. This model is discussed below.

The energy is calculated by using this formula.

$$\text{Energy} = \text{Power} * \text{Time} \tag{1}$$

The energy consumption is measured by the transmitting power or receiving power multiply the transmitted time.

$$P_t = \frac{(8 * \text{PacketSize})}{\text{Bandwidth}} \tag{2}$$

The transmitting energy E_{tx} is defined as

$$E_{\text{tx}} = P_{\text{tx}} * P_t \tag{3}$$

The receiving energy E_{rx} is defined as

$$E_{\text{rx}} = P_{\text{rx}} * P_t \tag{4}$$

Energy consumption of a node after time t is calculated using the following equation:

$$E_{\text{con}}(t) = N_{\text{t}} * C_1 + N_{\text{r}} * C_2 \tag{5}$$

where

$E_{\text{con}}(t)$	Energy consumed by a node after time t.
N_t	No. of packets transmitted by the node after time t.
N_r	No. of packets received by the node after time t.
C_1 and C_2	Are constant factors having a value between 0 and 1.

Let E be the initial energy of a node, the residual energy E_{Res} of a node at time t can be calculated by using the formula

$$E_{\text{Res}} = E - E_{\text{con}}(t) \tag{6}$$

where E_{Res}: residual energy and E_{con}: consumed energy.

Total energy consumption of all nodes is measured as the summation of all node's residual energy plus the product of initial energy and number of nodes.

$$T_{\text{Econ}} = N * \text{Initial Energy} - E_{\text{Res}} \tag{7}$$

where T_{Econ}: total consumed energy and N: total number of mobile nodes in MANET.

The following algorithm illustrates the route discovery process.

3.5 EPRDSR Algorithm

3.5.1 Route Discovery

Step 1. Source node *S*.

(a) Create the RREQ packet with field values set as SA = *S*, DA = *D*, Seq.No = I, TTL = *T*, Hops = *H*, Min_energy = Initial energy.

(b) Broadcast the RREQ packet to next neighbor node whose Max_.Network_ lifetime = Min_.transmission_ power.

Step 2. If the intermediate node will receive the RREQ packet.
(a) The Min_power field in RREQ is updated by initial energy.
(b) Forward the RREQ packet to node 2.
(c) Calculate node's residual energy(remaining energy).
(d) This residual energy value is compared with Min_power value in routing table. The route is selected on the basis of Min_power >=Renergy. Otherwise the link between intermediate nodes is unavailable.
Step 3. If the node receiving the RREQ packet in D, then the node D.
(a) Generates the RREP packet for uni-casting to source. The bandwidth field of the RREP packet is updated with the cumulative bandwidth of the path and Energy field should be updated by cumulative Energy.
(b) D uni-casts all the node disjoint paths back to the source node *S*.

In route discovery procedure, the EPRDSR builds a route between source to destination using a route request and route reply query cycle. When a source node wants to send a packet to destination for which it does not already have a route, it forward a route request (RREQ) packet to all the neighbors across the network. The performance of EPRDSR is improved by adding energy model parameters in RREQ packet, two additional fields are added in the RREQ header information such as bandwidth and energy constraints.

3.5.2 Route Selection

When RREQ receives at the neighbor node, it forwards a RREP packet back to the source. Otherwise, it rebroadcasts the RREQ. If they may receive a processed RREQ, they discard the RREQ and do not forward it. If RREQ of multiple paths is received at source node, it stored by the hop-count value. In EPRDSR, the route is selected on the basis of minimum number of hops. But the EPRDSR protocol select the best path by sorting multi-route in descending order of nodal residual energy and bandwidth, and the data packets are forwarded by using the maximal nodal residual energy. The extended route request packet of EPRDSR is shown in Table 1.

In EPRDSR-routing discovery process, the source node in the network sends the extended route request (RREQ) message to the destination node through number of intermediate nodes. The data transmission in wireless network can be directly within one hop or through number of intermediate nodes. The extended RREQ message contains the source and destination node IP address, advertised hop-count value, timeout value, bandwidth of the link (bandwidth), and minimum energy value. The computed bandwidth and minimal nodal energy are greater than the threshold value of bandwidth and energy then only the RREQ message forwards to the next neighbor node otherwise it discarded. When the RREQ message

Table 1 Extended route request message

SA	DA	Seq. No	Hop count	Timeout	Bandwidth	Min_Energy

arrives at next node, the bandwidth and minimal nodal energy are updated into the route list entries.

At the initial stage, the source node's initial energy is entered into minimum energy field, and the residual energy is computed at every node in the network. This residual energy is compared with minimum energy field of RREQ packet. If this value is less than the minimum energy field, then it replaced by residual energy. While selecting the best path, the minimum energy should be kept as the lowest among all the nodes in this route. Once the RREQ packet is received by the destination node, the node will produce route reply (RREP) packet and send back to the source node. RREP packet is also included two additional fields bandwidth and minimum energy, and the RREP packet records the routing information from the source to destination. The duplicate packet ID is received by the destination node; then it responds with a maximum of RREP packets to the source node.

4 Analysis of Experimental Results

For evaluating EPRDSR, simulation results are obtained by experiments with different energy levels, different traffic loads, and different movement patterns of nodes. In the first set of experiments, every node is given a battery with full capacity initially. In this scenario, no nodes turn off due to running out of energy. This is to test the routing performance of EPRDSR in a regular network scenario. A total of 300 CBR streams are generated within the 1,500 s simulation time. Extensive simulations were conducted using NS-2.33. The simulated network consisted of 100 nodes randomly scattered in a 1,000 × 1,000 m area at the beginning of the simulation. The tool *setdest* was used to produce mobility scenarios, where nodes are moving at six different uniform speeds ranging between 0 and 25 m/s. Table 2 shows the simulation parameter setting for the protocol evaluation. These were generated using the tool *EPRDSR..tcl*, with the following parameters:

4.1 Power-Aware Metrics

4.1.1 Packet Delivery Ratio

The ratio of the data packets delivered to the destination to those generated by CBR sources. This metric illustrates the effectiveness of best effort routing protocols. This performance measure also determines the completeness and correctness of the routing protocol. If F is fraction of successfully delivered packets, C is total number of flows, f is id, R is packets received from f, and T is transmitted from f, then F can be determined by

Table 2 Simulation parameters

Topographical area	1,000 × 1,000
Channel type	Wireless channel
Radio propagation model	TwoRayGround
Network interface type	Phy/Wirelessphy
MAC type	802_11 g
txPower	0.2 W
rxPower	0.1 W
idlePower	0.001 W
Initial energy of a node	2,000 J
Routing protocols	DSR/MTPR/EPRDSR
Number of mobile nodes	100
Mobility	0–25 m/s

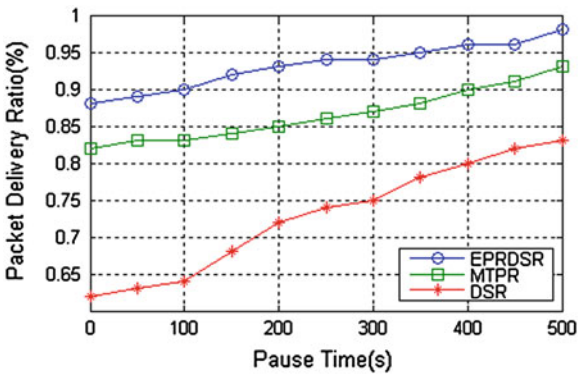
$$F = \frac{1}{C}\sum_{f=1}^{c}\frac{R_f}{T_f} \tag{8}$$

It has been found that in all cases, EPRDSR perform better than MTPR and DSR in packet delivery ratio. Figure 1 shows the throughput in packet delivery ratio for 100 nodes. As is clear this scheme improves the throughput performance of EPRDSR as well.

4.1.2 End-to-End Delay

Average end-to-end delay is the delay experienced by the successfully delivered packets in reaching their destinations. This is a good metric for comparing protocols. This denotes how efficient the underlying routing algorithm is, because delay primarily depends on optimality of path chosen.

Fig. 1 Packet delivery ratio



$$\text{Average End to End delay} = \frac{1}{S}\sum_{i=1}^{s}(r_i - s_i) \tag{9}$$

where S is number of packets received successfully, r_i is time at which packet is received, s_i is time at which it is sent, and i is unique packet identifier. Simulations have been conducted for 100 nodes. As the mobility increases, the performance gain by alternate routes becomes more significant. Due to the presence of neighboring nodes which are in active state, nodes attempt to use multiple alternate paths for data delivery in the presence of route breaks, the protocol is able to deliver more packets to the destination than EPRDSR. The proposed scheme has some packet losses. Alternate paths may be broken as well as the primary route because of mobility or power status of nodes or be unavailable and not discovered during the route reply phase. Moreover, packets can be lost because of collisions and contention problems.

End-to-end delay is presented in Fig. 2. As expected, new scheme has longer delays than EPRDSR. One can only measure delays for data packets that survived to reach their destination. Proposed scheme delivers more data packets, and those packets that are delivered in power-aware routing but not in EPRDSR, take alternate and possibly longer hop routes. So power-aware routing scheme having longer delays than EPRDSR does not represent its ineffectiveness since these protocols use the same primary route.

Figure 3 highlights the energy consumed by routing protocols. EPRDSR offers the best performance while DSR shows the worst results. Typically, on-demand protocol DSR presents an energy descendent trend as the motion rate drops, and the MTPR presents an energy consumption that remains practically constant as motion rate varies.

Figure 4 shows the results when varying the maximum node speed among values 0, 1, 5, 15, and 25 m/s. These results confirm the constant behavior of EPRDSR even for high motion patterns. The energy consumption of the three on-demand protocols increases as the maximum motion speed grows. As the

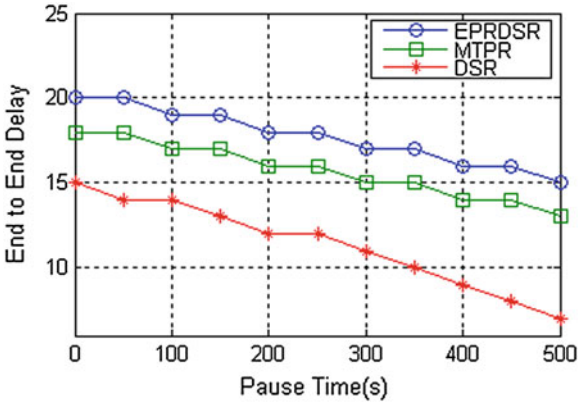


Fig. 2 End-to-end delay

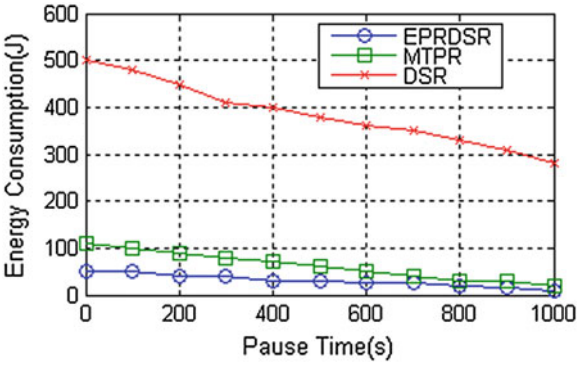


Fig. 3 Routing energy consumption comparison for pause time

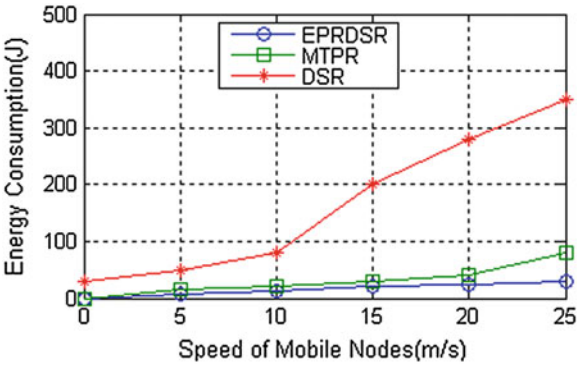


Fig. 4 Routing energy consumption comparison for maximum node speed

motion speed moves from a humans walking MANET scenario to a road cars MANET, the difference between DSR and MTPR grows of a factor of 3 to a factor of 5. Finally, when speed grows, EPRDSR performs better than MTPR and DSR.

Figure 5 shows the simulation results when varying the number of nodes while maintaining the traffic load. When passing from 20 to 60 nodes, the protocol

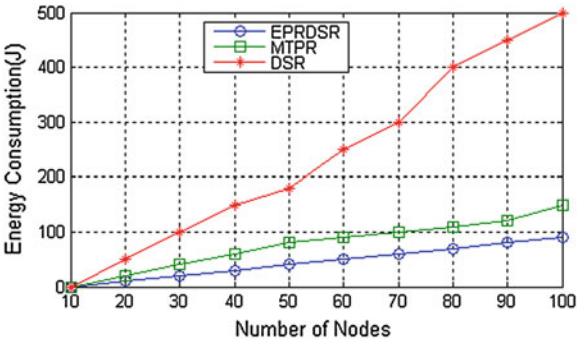


Fig. 5 Routing energy consumption comparison for varying node number

suffers an increment of 51.8 %. This characteristic makes this protocol not scalable. With respect to the DSR and MTPR, the energy increment for a MANET of 100 nodes with respect to a MANET of 60 nodes is quite similar and is about 65 %. This increment is mainly due to route maintenance process; with EPRDSR, the increment is mainly due to the propagation of power aware with maximum lifetime between nodes.

5 Conclusion

A scheme has been proposed that utilizes power status of each mobile node and alternate paths. This scheme can be incorporated into any ad hoc on-demand unicast routing protocol to improve reliable packet delivery, end-to-end delay, and throughput in the face of node movements and route breaks. As a case study, it has been applied to EPRDSR and performance has been studied via simulations. Simulation results have indicated that new technique provides robustness to mobility and enhances protocol performance. Its performance has been found much better than other existing protocols (DSR and MTPR) in finding the active routes increases. The performance of EPRDSR protocol increases the network lifetime by 60–65 % as compared DSR and MTPR and its counterparts and causes slight overhead in route selection initially. Overall, we conclude that our mechanism demonstrates significant benefits at high-traffic and high-mobility scenarios. We expect that these scenarios will be common in ad hoc networking applications.

References

1. Shankar S, Sivakumar B, Varaprasd G Implementing energy aware issues in MANET by designing efficient routing protocol. Int J Comput Appl (IJCA) ISSN: 0975–8887 6(9) ProQuest CSA Technol Res J Cite-Seer. NASA ADS USA
2. Maleki M, Dantu K, Pedram M (2003) Lifetime prediction routing in mobile ad hoc networks. IEEE Wireless Commun Networking Conf 2:1185–1190
3. Royer E, Toh CK (1999) A review of current routing protocols for ad hoc mobile wireless networks. IEEE Pers Commun 6(2):46–55
4. Muthuramalingam S, RajaRam R, Pethaperumal K, Devi VK (2010) A dynamic clustering algorithm for manets by modifying weighted clustering algorithm with mobility prediction. Int J Comput Electr Eng 2(4):1793–8163, 709–714
5. Hussein AH, Abu Salem AO, Yousef S (2008) A flexible weighted clustering algorithm based on battery power for mobile ad hoc networks. IEEE Int Symp Ind Electron
6. Yu JP, Chong PHJ (2005) A survey of clustering schemes for mobile ad hoc networks. IEEE Commun Surv Tutorials 7(1):32–48
7. Nagpal CK, Kaur M, Gupta S, Bhushan B (2011) Impact of variable transmission range on MANET performance. Int J Ad hoc Sens Ubiquit Comput (IJASUC) 2(4):59–66
8. Zabian A, Ibrahim A (2008) Power saving mechanism in clustered ad-hoc networks. J Comput Sci 4(5):366–371

9. Bein D, Datta AK, Sajja P, Zheng SQ (2009) Impact of variable transmission range in all-wireless networks. In: proceedings of the 42nd Hawaii international conference on system sciences
10. Shew JP, Hsieh KY, Cheng YK (2009) Distributed transmission power control algorithm for wireless sensor networks. J Inf Sci Eng 25:1447–1463
11. Tanejaa S, Kushb A, Makkarc A, Bhushand B (2011) Power management in mobile adhoc network. Int Trans J Eng Manage Appl Sci Technol 2(2)
12. Chauhan RK, Chopra A (2010) Energy efficient routing in mobile ad hoc network with capacity maximization, IJCA special issue on "mobile ad-hoc networks" MANETs

Studies on Quasi-Z-Source-Based Isolated DC/DC Converter for Distributed Power Generation

H. C. Vijaya Kumari, H. N. Suresh and S. Dhamodharan

Abstract Distributed generation is an emerging concept in the electricity sector, which represents good alternatives for electricity supply instead of the traditional centralized power generation concept. In particular, fuel cell power generation due to its inherent advantages is expected to play a big part as a distributed generation scheme. The present chapter describes novel quasi-Z-source-based isolated DC/DC converter for distributed power generation. It is used to interconnect low-DC-voltage-producing fuel cell to residential loads. The system contains voltage-fed quasi-Z-source inverter with continuous input current on the primary side, a single-phase isolation transformer, and a voltage doubler rectifier. The model of the quasi-Z-source-based isolated DC/DC converter is considered for simulation studies using *MATLAB*. Simulated data related to gate signals of quasi-Z-source inverter, voltage from fuel cell, DC-link voltage, primary and secondary voltage from the isolation transformer, and output voltage from the voltage doubler rectifier are used to test the performance of the quasi-Z-source-based DC/DC converter.

Keywords DC/DC converter · Fuel cell · Quasi-Z-source inverter · Rectifier

H. C. Vijaya Kumari (✉) · H. N. Suresh
Malnad College of Engineering, Hassan, India
e-mail: viji.kumari10@gmail.com

H. N. Suresh
e-mail: suresh_mce@rediffmail.com

S. Dhamodharan
Adroit Solutions, Bangalore, India
e-mail: info@adroitsolutions.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_13,

1 Introduction

Distributed generation represents a small-scale electric source which is a good alternative for electric supply instead of traditional centralized power generation concept. Its services include standby generation, peak saving, base-load generation, or cogeneration. For businesses, distributed generation can reduce peak demand charges and overall energy use and emissions and at the same time ensure good power quality and reliability. For large utilities and power producers, distributed generation can augment overall system reliability, avoid large investments in transmission line upgrades and the associated transmission losses, closely match capacity increases to demand growth, and supply demand more efficiently in remote areas [1].Various technologies that are available for distributed generation include microturbines, solar panels, wind turbines, and fuel cells. In particular, fuel cell power generation is recognized as the modern approach to distributed power generation due to its several advantages.

The utilization of fuel cell for distributed power generation requires the development of a special voltage-matching converter that converts a fuel cell low DC voltage to residential load voltage. Figure 1 represents the typical structure of two-stage interface converter suitable for such purposes.

The interface converter is realized within the DC/DC/AC concept for safety and dynamic performance. The most challenging issue in this is design of the front-end isolated DC/DC converter, because this stage is the main contributor of interface converter efficiency, weight, and overall dimensions. The low voltage provided by the fuel cell is always associated with high currents in the primary part of the DC/DC converter. These high currents lead to high conduction and switching losses in the semiconductors and therefore reduce the efficiency. Moreover, the large voltage boost factor requirement presents unique challenge to the DC/DC converter design.

The conventional voltage-matching converters are voltage-source pulse-width-modulated inverters with the step-up transformer and rectifier. The drawback of this scheme is the forbidden shoot-through period.

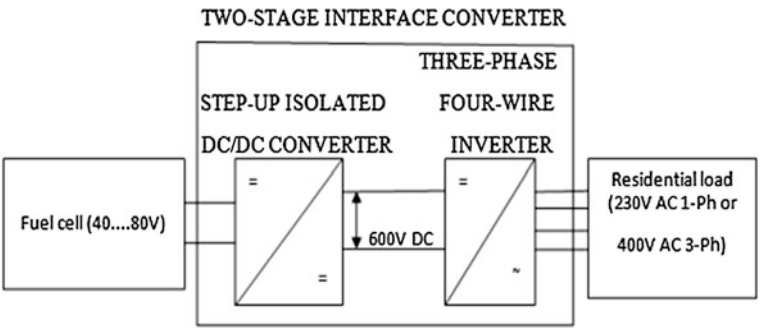


Fig. 1 Typical structure of the interface converter for residential fuel cell–powered systems

Alternately, combination of the front-end boost converter with the voltage-source pulse-width-modulated inverter can be resorted for a similar purpose [2]. The drawback associated with this alternate scheme is the complicated control and protection algorithm and reduced reliability due to increased number of switching devices. In this context, the impedance-source (Z-source) inverter-based topologies provide a modern approach to the step-up voltage conversion techniques. The main feature of the Z-source inverter [3] is that it can boost the input voltage by introducing a shoot-through operation mode, which is forbidden in traditional voltage-source inverter.

The work reported in this chapter describes development of a novel quasi-Z-source-based single-phase isolated DC/DC converter for distributed generation which contains voltage-fed quasi-Z-source inverter with continuous input current on the primary side, a single-phase isolation transformer, and voltage doubler rectifier.

2 Overview of Quasi-Z-Source-Based Single-Phase Isolated DC/DC Converter

The quasi-Z-source-based single-phase isolated DC/DC converter circuit topology is shown in Fig. 2. The quasi-Z-source inverter is derived from a traditional Z-source inverter [4] and has two distinctive advantages as compared to Z-source inverter, such as continuous DC current drawn from the source and lower operating voltage of the capacitor. The quasi-Z-source network consists of two capacitors C1 and C2, two inductors L1 and L2, and a diode D1.

The high-frequency step-up isolation transformer provides the required voltage gain and the galvanic isolation of the input and output sides of the converter. Transformer's primary winding is connected to the output terminals of quasi-Z-source inverter, while the secondary side is connected to the voltage doubler rectifier. The voltage doubler rectifier is derived from the full-bridge rectifier by the replacement of diodes in one leg by the capacitors (C3 and C4). The resulting

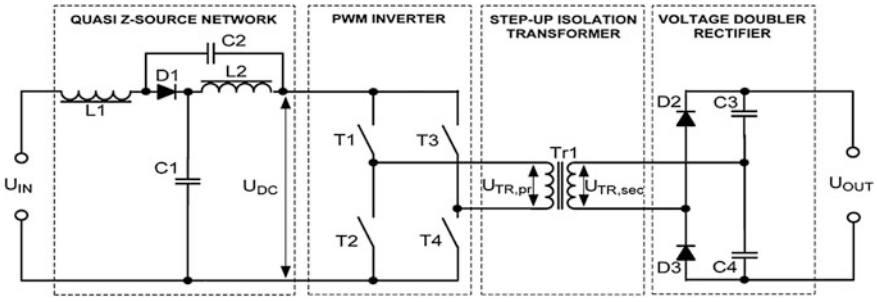


Fig. 2 Single-phase quasi-Z-source-based isolated DC/DC converter scheme

advantages of the voltage doubler rectifier over the traditional rectifier scheme are the doubling effect of the secondary winding voltage of the isolation transformer and reduced power dissipation due to smaller number of rectifying diodes and full elimination of the smoothing inductor.

2.1 Operating Modes of the Quasi-Z-Source-Based Single-Phase DC/DC Converter

The central idea implemented in the quasi-Z-source-based converter is to keep the DC-link voltage (U_{DC}) constant despite the variation in voltage of the fuel cell. By keeping the DC-link voltage constant, the PWM inverter can be operated with a fixed duty cycle. In accordance with the input voltage (fuel cell voltage), the operating modes of the DC/DC converter of the present work can be broadly categorized as non-shoot-through and shoot-through operation modes as shown in Fig. 3. During non-shoot-through operation mode, the discussed quasi-Z-source DC/DC converter acts as a traditional voltage-source inverter.

The control principle of the single-phase quasi-Z-source inverter in the shoot-through (voltage boost) operation mode is shown in Fig. 4. The switching patterns of the traditional single-phase voltage-source inverter are shown in the Fig. 4a.

To control the shoot-through states, two reference signals (Up and Un) are introduced as shown in Fig. 4b. If the triangular waveform is greater than Up or lower than Un, the inverter switches turn into the shoot-through state. Thus, in the proposed shoot-through PWM control method, to reduce the current stress of the switches, the shoot-through states are created by simultaneously turning ON all inverter switches. The shoot-through time is evenly split into two intervals of half the duration. During the shoot-through states, the voltage across inverter bridge drops to zero. The resultant DC-link voltage is shown in Fig. 4c.

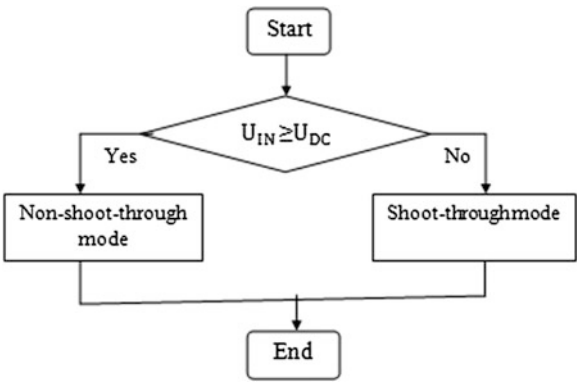


Fig. 3 Flowchart of operating modes of quasi-Z-source-based isolated DC/DC converter

3 Simulink Model of Quasi-Z-Source-Based DC/DC Converter

The operating principle discussed in Fig. 4 is reflected in a block diagram of gate signal generator developed for *MATLAB* implementation as shown in Fig. 5. It consists of control of active and zero states and control of shoot-through states. The shoot-through states are generated using triangular waveform generator and two relational operators in the Simulink model. The active and zero states are generated using 50 % pulse width pulse generators and NOT logic gates. In the control of shoot-through states, triangular waveform is generated by repeating sequence block. The reference signals, Up and Un, are set using constant block.

The simulation model of quasi-Z-source single-phase isolated DC/DC converter with DC machine load developed in the present work is shown in Fig. 6. The following are the quasi-Z-source single-phase DC/DC converter parameters assigned for simulation purposes: $P = 500$ W, C1 = C2 = 240 μF, L1 = L2 = 50 μH, and C3 = C4 = 10 μF. The isolation transformer has the turns' ratio of 1:3.75. The parametric values assigned are as mentioned in the literature [5].

The *Continuous* block indicated in the simulation model of Fig. 6 is a power graphical user interface. The power graphical user interface block is necessary for simulation of any Simulink model containing SimPower system blocks. It is used to store the equivalent Simulink circuit that represents the state-space equations of the model.

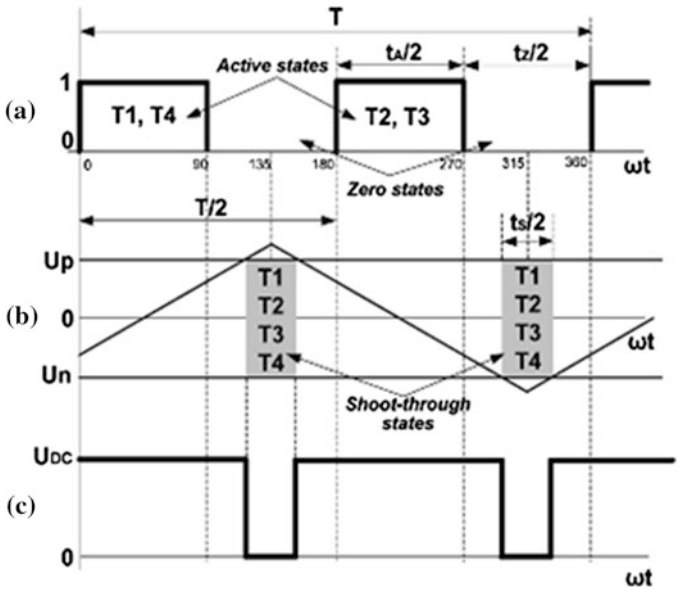


Fig. 4 Operating principle of the single-phase quasi-Z-source inverter in the shoot-through mode

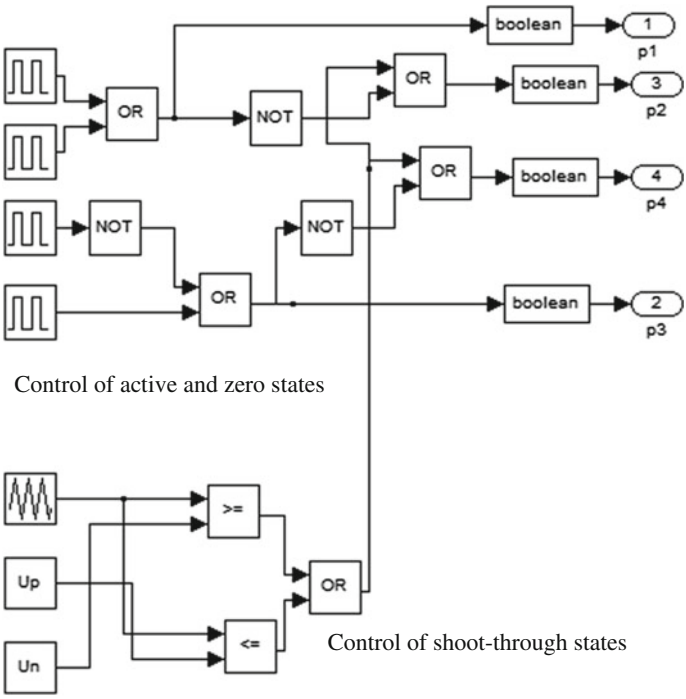


Fig. 5 Simulation model of gate signal generator developed in the present work

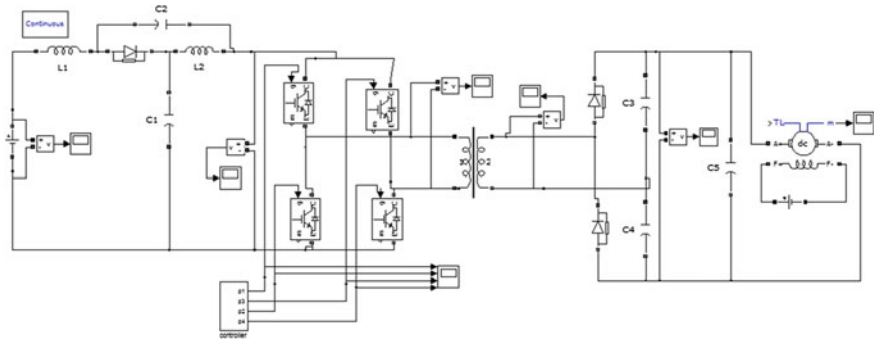


Fig. 6 Simulation model of quasi-Z-source-based single-phase isolated DC/DC converter

4 Results and Discussion

In single gate signal generator circuit, shoot-through period and non-shoot-through period gate signals are obtained. Depending on the gate signal type, the converter operates as quasi-Z-source-based isolated DC/DC converter or traditional voltage-source inverter.

4.1 Gate Signals of Quasi-Z-Source Inverter

The gate signals to IGBT/diode of quasi-Z-source inverter are generated by a special gate signal generator circuit (Fig. 5). Figure 7a shows the gate signals with shoot-through period given in the literature [5], and Fig. 7b shows the shoot-through period signals obtained in the present work. The gate signals have active states, zero states, and shoot-through states. The shoot-through periods are present in the zero states of transistor T2 and transistor T4 gate signals. The shoot-through period is evenly split into two intervals of half the time duration. The gate signals have 0.25 duty cycle shoot-through period. The gate signals without shoot-through period obtained in *MATLAB* simulation are shown in Fig. 7c. The gate signals have active and zero states. The operating frequency of the transistors T1, T2, T3, and T4 is equal to that of isolation transformer (5 kHz).

In Fig. 7a, the time frame window amounts to 0.2 ms or 2×10^{-4} s (i.e., the start instant is 39.50 ms and end instant of signal capture is 39.70 ms). Similarly, in Fig. 7b, c, the time frame window of signal capture is over 0.2 ms or 2×10^{-4} s. In both figures, the time duration of signal capture is maintained to be 0.2 ms (amounting to a frequency of 5 kHz), such that relative comparisons are valid.

Similarly, in the next section, other basic operating waveforms of the DC/DC converter that are simulated using *MATLAB* are presented. The wave shapes obtained are as described in the literature [5].

4.2 Fuel Cell Input Voltages

To demonstrate the performance of quasi-Z-source-based isolated DC/DC converter, consider 40 V DC input voltage supply from a fuel cell. For the performance studies of the present work, input voltage of 40 V is applied as shown in Fig. 8a. The 40 V input voltage is needed when the quasi-Z-source inverter is

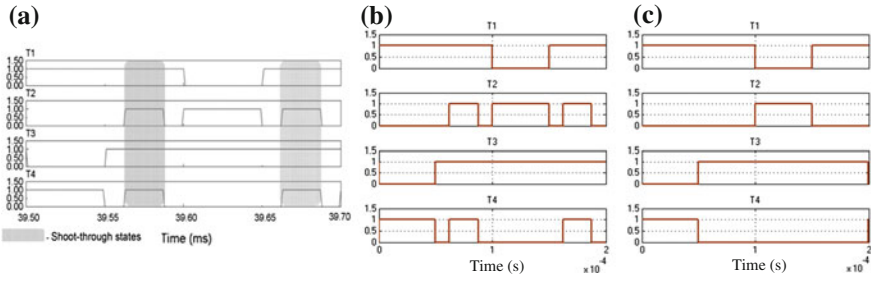


Fig. 7 Gate signals **a** with shoot-through period presented in the literature, **b** with shoot-through period, and **c** without shoot-through period obtained in the present work

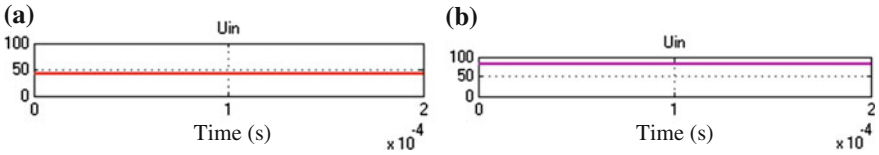


Fig. 8 Input voltage waveform **a** 40 V input during shoot-through operation mode and **b** 80 V input during non-shoot-through operation mode

operated with shoot-through period gate signals. The output voltage from fuel cell is then applied to the quasi-Z-source network. This 40 V output is boosted to 80 V DC-link voltage in quasi-Z-source network due to the effect of shoot-through period gate signals.

A fuel cell of 80 V is needed when the converter operates with non-shoot-through period gate signals. Figure 8b shows the 80 V DC input supply used in the *MATLAB* simulation model. The 80 V input leading to 80 V DC-link voltage shows that the converter is operated as a traditional voltage-source inverter.

4.3 DC-Link Voltages

The effect of voltage boost depends on the shoot-through period in the gate signals. 80 V DC-link voltage obtained in the present work is shown in Fig. 9a.

The quasi-Z-source network operated with gate signals without shoot-through period ensures 80 V DC-link voltage only (when voltage input to the converter is maximum, i.e., 80 V). Figure 9b shows the 80 V DC-link voltage obtained through *MATLAB* simulation of the present work.

4.4 Isolation Transformer Primary and Secondary Winding Voltages

The isolation transformer is located after quasi-Z-source network and before voltage doubler rectifier in the quasi-Z-source-based single-phase isolated DC/DC

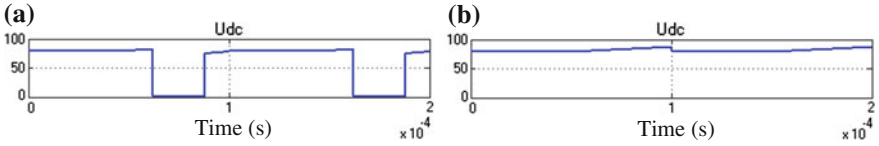


Fig. 9 DC-link voltage waveform obtained **a** during shoot-through operation mode and **b** during non-shoot-through operation mode in the present work

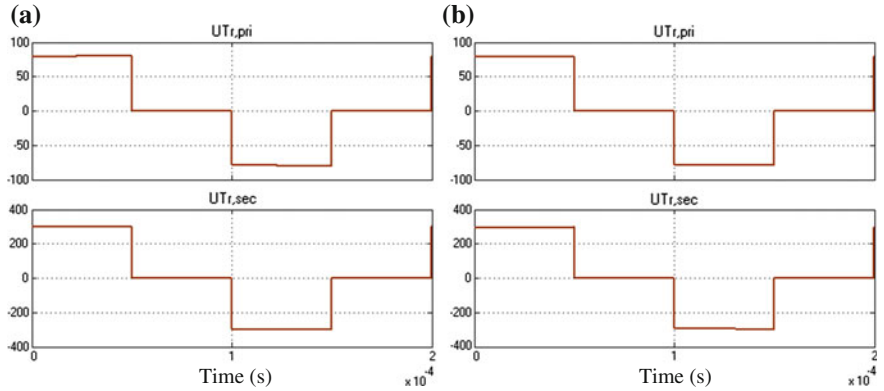


Fig. 10 Primary voltage and secondary voltage of the isolation transformer obtained **a** during shoot-through operation mode and **b** during non-shoot-through operation mode in the present work

converter model (Fig. 6). It has a turns' ratio of 1:3.75 and is used to boost the output voltage from the quasi-Z-source network. Figure 10 shows the primary voltage of 80 V and secondary voltage of 300 V obtained in the present work.

4.5 Output Voltage from Voltage Doubler Rectifier

The voltage doubler rectifier doubles the output voltage from the secondary winding isolation transformer and is situated after isolation transformer in the quasi-Z-source-based single-phase isolated DC/DC converter model (Fig. 6). The voltage doubler rectifier provides the demanded voltage doubling effect of the peak voltage of the secondary winding of the isolation transformer, thus ensuring the DC output voltage 600 V. Figure 11a shows the 600 V DC output voltage obtained in the present work during shoot-through operation mode.

600 V DC output voltage obtained in the present work is shown in Fig. 11b. The waveform shown corresponds to the proposed converter acting as voltage-source inverter. The results obtained upon simulation studies show that the

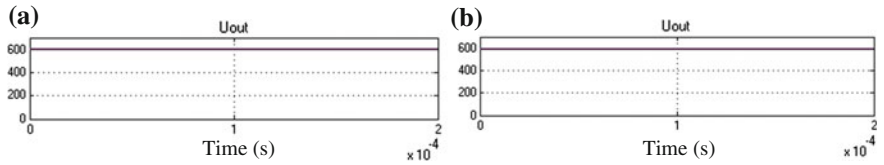


Fig. 11 600 V DC output voltage waveform obtained **a** during shoot-through operation mode and **b** during non-shoot-through operation mode in the present work

quasi-Z-source-based DC/DC converter developed provides a required output voltage (600 V) utilizing minimum 40 V input voltage.

5 Conclusions

In the present work, the quasi-Z-source-based isolated DC/DC converter has been successfully applied for distributed power generation. Important conclusions that are drawn out of the investigations made in the present work are as follows:

- The implemented quasi-Z-source inverter can successfully boost the input voltage by introducing shoot-through operation mode, which is forbidden in traditional voltage-source inverter.
- The quasi-Z-source inverter implemented on the primary side of the quasi-Z-source-based isolated DC/DC converter is observed to provide both voltage boost and buck functions with no additional switches, only by the use of a special control algorithm.
- The high-frequency step-up isolation transformer provides the required voltage gain and input–output galvanic isolation.
- The number of turns of the secondary winding of the isolation transformer is reduced by 62 % (turns ratio of 1:3.75 in the case of voltage doubler rectifier instead of 1:10 of traditional full-bridge rectifier) due to the voltage doubling effect available with the voltage doubler rectifier.

References

1. Xu H, Kong L, Wen X (2004) Fuel cell power system and high Power DC–DC converter. IEEE Trans Power Electron 19(5)
2. Choi W, Enjeti P, Howze JW (2003) Fuel cell powered UPS systems: design considerations. Available at:http://enjeti.tamu.edu/journal-papers/ fuelcellupsdesign.pdf
3. Peng FZ (2003) Z-source inverter. IEEE Trans Ind Appl 39(2):504–510
4. Li Y, Anderson J, Peng FZ, Liu D (2009) Quasi-Z-source inverter for photovoltaic power generation systems. In: proceedings 24th annual IEEE APEC, 15–19 Feb 2009, pp 918–924
5. Vinnikov D, Roasto I (2011) Quasi-Z-source-based isolated DC/DC converters for distributed power generation. IEEE Trans Ind Electron 58(1):192–201

Security in Pervasive Computing: Risk Probability Assessment Model

A. C. Yogeesh and G. Manoj Someshwar

Abstract Pervasive computing (also known as ubiquitous computing) brings to light a new genre of computing where the computer completely pervades the life of the user. Pervasive computing applications promise seamless integration of digital infrastructure with their interconnected devices and services into our everyday lives. Pervasive computing applications often demand physical location and other context information about users and resources to enhance the user experience. The necessity of sharing resources and collaborate brings out new types of interactions among users as well as between the virtual and physical worlds. In this context, it becomes difficult to separate physical security from digital security. The focuses of current paper demonstrate a risk assessment model for security in pervasive computing applications.

Keywords Ubiquitous computing · Mahalanobis · Clusters · Kohonen self-organizing map

1 Introduction

With pervasive computing, we refer to a scenario in which computing is omnipresent, and particularly in which devices that do not look like computers are endowed with computing capabilities [1]. "A computer on every desk" does not

A. C. Yogeesh (✉)
Government Engineering College, Kushalnagar, Karnatka, India
e-mail: Yogeesh13@gmail.com

G. M. Someshwar
Anwarul-uloom College of Engineering & Technology, Yennepally, Vikarabad, RR district, Andhra Pradesh, India
e-mail: Manojgelli@yahoo.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_14,

qualify as ubiquitous computing, having data processing power inside light switches, door locks, fridges and shoes, instead, does.

A global ubiquitous computing infrastructure is envisioned in which billions of autonomous entities must interact in a decentralized and ad hoc manner. In this type of environment, traditional security mechanisms based on a centralized authorization model will not scale.

The interactions between these autonomous entities are similar to those interactions in human networks [1, 2]. Humans must often make ad hoc decisions regarding interaction with partially known or unknown persons in situations where complete information is unavailable and where no trusted third party exists. Similarly, entities in the ubiquitous computing environment are both autonomous and mobile and must be capable of dealing with unforeseen circumstances ranging from unexpected interactions with other unknown entities to disconnected operation. Human society has developed the concept of trust to overcome initial suspicion and gradually evolve privileges in these scenarios. Recently, there has been an increased interest in the development of security mechanisms for this type of environment based on the human notion of trust.

2 Risk Assessment Model

In Risk assessment model, we must characterize the extent to which risk is associated with a privilege that may be assigned to an unknown entity such that the entity's trustworthiness may then be used to decide whether the risk is acceptable in granting said privilege [3, 4]. Risk is the possibility of something adverse happening, and risk management is the process of assessing risk, taking steps to reduce risk to an acceptable level and maintaining that level of risk.

The SECURE research project [5] focuses on the integration of trust and risk in making security decisions in the pervasive computing environment. As illustrated in the SECURE framework diagram presented in Fig. 1, risk evaluation is fundamental component in performing trust-based access control.

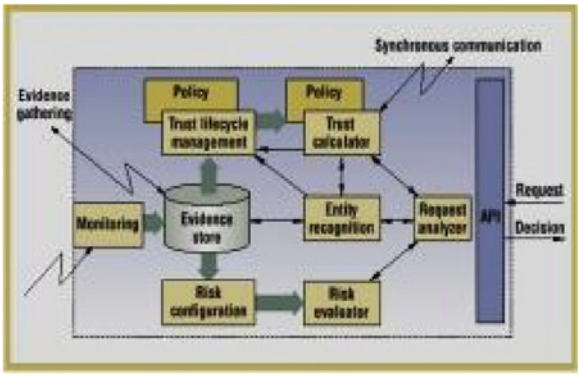


Fig. 1 SECURE framework

Risk is commonly defined as the hazard level combined with the likelihood of the hazard leading to an accident and the hazard exposure or duration (latency) [5]. Mahalanobis is used to measure the similarity between any two interactions precisely and take correlation of the different features into account.

2.1 Risk Probability Estimation Model

We assume that any interaction in ubiquitous computing environments can be expressed as a feature vector and that the vector elements are comparable. This is reasonable as different discrete numbers can denote even different non-numeric features. The features must be able to describe an interaction precisely and completely. Therefore, the features describing the context of and the principals engaged in an interaction must be specified as precisely as possible to measure the risk associated with an interaction, the following general risk assessment formula [5] is presented.

$$R = F(x1,\ x2,\ x3,\ xm) + Z \tag{1}$$

or

$$R = F(X) + Z \tag{2}$$

where R is the probability of risk of a certain interaction, or indeed the risk value itself. X is the feature vector, and xi (i 1, 2, 3, m) are its elements which consist of known parameters for this interaction. The feature elements specify the context of the interaction, participants and relevant historical memory. Their values are derived from observation or collected data. Z is the random disturbance factor, and normally, we assume it to be zero. F maps the current context and participant features to the risk value or risk probability. Its specific format may be known or unknown depending on different contexts.

When the map F is known, which might be a linear or non-linear function, estimating R is not difficult. For example, consider the probable risk of vehicular death by driver's age. If it is a normal distribution, F has a $\frac{1}{\sigma\sqrt{2\pi e}}^{-\frac{1}{2}\left(\frac{x-u}{\sigma}\right)}$ according to simple data, the parameters, u, σ are easy to confirm. These are the traditional parameters used in an estimating procedure.

If the dimension of vector X is high, up to 3, estimating R is also easy as we can plot out the data and see the pattern underlying the huge data set.

However, in fact, the map F is often unknown and the dimension of the features vector is very high. For this kind of situation, F is just like a black box passing the current data and historical data as inputs and outputting a risk value. It is worthwhile to note that the historical data are necessary for estimating the risk of the current interaction. For example, it is difficult to imagine that a child could give risk estimation for driving, as he would have not had enough experience in this context, that is, historical data.

However, even for the historical interactions, we know only if the result of each interaction is unexpected or not. The risk probability associated with the interactions remains unknown. If the historical features vector with confirmed risk probability were known to us, there would be many means to find the map F and predict the probability of risk for a new interaction. The non-linear regression, regression splines and neural network provide the mathematical solutions to it.

In Sect. 2.1.1, we will present a general risk probability definition aimed to solve this problem.

2.1.1 Risk Probability Definition

We define a general risk probability as follows:

$$R(X) = \frac{U(N_r(X))}{\|N_r(X)\|} \tag{3}$$

where $N_r(X)$ is an r-neighbourhood of point X, $N_r(x) = \{y : \rho\ (x, y) < r\}$ The r-neighbourhood is not necessary a super sphere.

$\|N_r(X)\|$ is the number of all the points in this neighbourhood.

$U(N_r(X))$ is the number of unexpected points in corrects for correlation between the different features.

2.2 Risk Probability Estimation

According to our risk probability definition, when a cluster [5] is confirmed, the number of unexpected points in this cluster is easy to collect for historical data, from which the risk value is easy to calculate. So clustering historical interactions according to patterns is necessary as a kind of pre-processing in which distinct subclasses of patterns are discovered whose members are more similar to each other than they are to other patterns.

For each cluster, the rate between the number of unexpected results of some interactions and the numbers of elements in the cluster is defined as the average loss rate (ALR) [5]. We aim to make up risk probability for each vector in this cluster by ALR. Intuitive idea is to see how close each vector X is to the average vector in the cluster. The most common measure of similarity for two vectors is

$$\mathrm{Sim}(X_i, \bar{X}) = \frac{\sum_{n=1}^{m} x_{\mathrm{in}} \times \overline{x_n}}{\sqrt{\left(\sum_{n=1}^{m} {x_{\mathrm{in}}}^2\right)\left(\sum_{n=1}^{m} \overline{x_n}^2\right)}}. \tag{4}$$

Essentially, this is Euclidean distance between two vectors. However, every element in a vector expresses a feature of the interaction. We do not know how important each feature playing a role in the interaction may be. At the same time, as all the elements have been discrete, they scatter in different ranges. The above formula cannot account for these issues. In our experiment, we give another measure of the similarity based on Mahalanobis distance.

$$\mathrm{MD}_t^2 = (X - m_t)C_t^{-1}(X - m_t)' \tag{5}$$

where X is one of the feature vectors in cluster t; m_t is the mean vector in cluster t; and C_t is the covariance matrix for X. It can be shown that the surfaces on which r is constant are ellipsoids that are centred about the mean m_t. In the special case where the features are uncorrelated and the variances in all directions are the same, these surfaces are spheres, and the Mahalanobis distance becomes equivalent to the Euclidean distance.

Since we can make up risk probability for each vector of historical data, the trained architecture is also able to apply to new input. We briefly summarize the training procedure as follows:

- Abstract the features from the historical data to generate the features vectors.
- Using the feature vectors as input vectors, start the clustering procedure.
- After clustering, select every cluster set from the clustering results.

3 Experimental Results and Discussion

We perform an experiment to verify risk probability estimation using a subset of the data which contains 34,065 records. The 34,065 pieces of data are divided into 24 clusters using a Kohonen self-organizing map [5]. The probability distribution for the data is given in Fig. 2. Each record is just a raw TCP/IP dump and therefore every feature vector has 41 elements, denoted F1 through F41. Only 8 clusters contain unexpected connections (attacks), and the other 16 clusters are all normal connections sets with ALRs of 0. The cluster 14 is the largest one which has almost 25 % connections while it contains no attacking connection. The risk of connections in this cluster is 0 that means a big part of connections are normal.

Let us examine how this risk estimation component might be integrated with the trust-based security architecture presented in Fig. 1. We cannot give risk estimation for some initial interactions for which we have no historical data. However, the monitoring component works to observe the result of the interactions and to collect relevant data to the distributed evidence store. When the data stored are enough to cover as many observed outcomes as possible, the risk evaluator is invoked to perform risk estimation using our algorithm. The historical data clustering procedure could work offline, while the estimation of risk for current interactions would work in real time as security decisions need to be made.

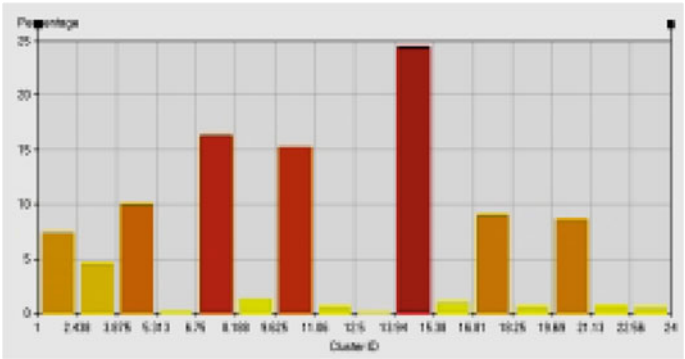


Fig. 2 Distribution of clustering data

This procedure is also similar to the human trust establishing and risk assessment. After people accumulate enough experience and identify the context and participants, they can recollect as similar as possible interactions which involves searching for data in similar context and participants to produce a precise estimation on risk.

4 Conclusion

Pervasive computing is all about facilitating our lives through digital environments which are sensitive, adaptive and responsive to human needs. For more than mobile computing, this technology will fundamentally change the nature of computing, allowing most objects we encounter in daily life to be "aware", interacting with users in both the physical and virtual worlds. While research challenges remain in all areas of pervasive computing, all the basic component technologies exist today. In hardware, we have mobile devices, sensors and even smart appliances. Supporting software technologies include digital signal processing and object-oriented programming. Hence, in the near future, we can expect our homes to network intelligent devices that transparently support our information and communication needs [6].

Risk assessment is a very important component in trust-based security strategy especially when the trustworthiness of some entity is unknown and no recommendation information is available. Currently, most traditional risk assessment procedures follow a fairly static process and cannot satisfy the requirements of the ubiquitous computing environment. A more flexible, dynamic risk assessment is needed. In this paper, we described a risk assessment model and proposed an estimator of risk probability that forms the core part of risk assessment in the ubiquitous computing environment. This estimator is based on a general definition inspired by traditional PDF approximation and implemented by a clustering

procedure [7]. To take the distribution of points into account, we adopt Mahalanobis distance to calculate similarity of interactions. We are currently developing the secure framework into which this risk probability estimator is embedded.

Acknowledgments This is a piece of research work carried out for PhD degree at CMJ University Shillong, Meghalaya. I would like to thank second author and my guide Dr. G Manoj Someshwar for his motivation, encouragement and support.

References

1. Ranganathan A, Al-Muhtadi J, Biehl J, Ziebart B, Campbell RH, Bailey B (2005) Towards a pervasive computing benchmark. In: Proceedings of the 3rd international conference on pervasive computing and communications workshops (PerCom 2005 Workshops)
2. Mattern F, Sturm P (2003) From distributed systems to ubiquitous computing the state of the art, trends, and prospects of future networked systems. In: Irmscher K, Fähnrich K-P (eds) Proc. KIVS 2003
3. Kindberg T, Zhang K, Shankar N (2002) Context authentication using constrained channels. In: Proc. WMCSA: 4th IEEE workshop on mobile computing systems and applications, IEEE Computer Society, pp 14–21
4. Fletcher S, Jansma R, Lim J, Halbgewaches R, Murphy M, Wyss G (1995) Software system risk management and assurance. In: Proceedings of the 1995 new security paradigms workshop, 22–25 Aug 1995, San Diego
5. Bacon J, Dimmock N, Ingram D, Moody K, Shand B, Andrew T (2003) Definition of risk model. SECURE Deliverable 3.1, 2003
6. Weiser M (1993) Ubiquitous computing. IEEE Comput 26(10):71–72
7. Satyanarayanan M (2001) Pervasive computing: vision and challenges. IEEE Pers Commun 8(4):10–17

CPU Load–Based Countermeasure Technique for Intelligent DoS Attack Targeting Firewalls

B. R. Vatsala and C. Vidya Raj

Abstract DoS attacks pose one of the most serious threats to firewalls which are considered as the first line of defence in any network. An attacker may use potential probing technique to discover last matching rules of a firewall rule set and hence trigger low-rate DoS attack that floods the firewall with traffic targeting rules at the bottom of policy. We propose a countermeasure technique in which DoS attack is detected by observing CPU load and ask all clients to solve a computational difficult problem and servicing the clients that provide the result.

Keywords Firewall · Firewall policy · Denial of service · CPU load

1 Introduction

A firewall [1] is a combination of hardware and software that deals with network packets according to its policy. The firewall policy consists of a list of rules with each rule representing a set of conditions. If an incoming packet matches all conditions of a particular rule, then a certain action is taken. If a rule is matched, the firewall matching engine skips remaining rules. Typically default rules and last matching rules are placed at the bottom of the rule set. Discovering how to trigger these rules by an outside attacker can be disastrous. The attacker can then launch a DoS [2] attack consisting of single flow of packets to trigger these last matching rules which require more CPU cycles of the server containing firewall.

B. R. Vatsala (✉) · C. Vidya Raj
Department of Computer Science and Engineering, The National Institute of Engineering, Mysore 570008, India
e-mail: vat_mys@yahoo.com

C. Vidya Raj
e-mail: vidya_rajc@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_15,

A feature-based detection scheme can be used to detect this type of intelligent attack by inspecting change in distributional aspects of packet size. Considering the fact that different application traffic has different packet size distributions and these distributions change during DoS attacks, an IP packet size entropy (IPSE)–based scheme in which the attacks are detected by observing time series of packet size entropy is used to distinguish between legitimate and attack traffic. The legitimate traffic will then be placed in a queue with high priority than the queue holding attack traffic.

We propose a model in which DoS attack is detected when CPU load is beyond the threshold load and in turn throw a question for each request to solve a computationally hard problem which steals the CPU cycles of the attacker.

This paper is organized as follows. Section 2 elaborates on the probing technique used to discover the last matching rules of a firewall policy. Section 3 discusses the entropy-based countermeasure scheme. Section 4 discusses our proposed model to endure DoS attack, based on CPU load.

2 Probing Technique to Discover Last Matching Rules

The key idea [3] behind probing technique for remotely discovering last matching rules is to send a number of back to back probing packets or a probing packet train (PPT) and then measure train stretch. A PPT length would stretch more if significant processing is required. To measure the train stretch, a non-cacheable HTTP request destined to Web server in DMZ is sent just after sending PPT as in Fig. 1, the time interval between sending HTTP request and receiving reply will reflect the PPT length.

Once the attacker finds the PPT that has long stretch, He sends these packets in a single flow and hence triggers low-rate DoS attack.

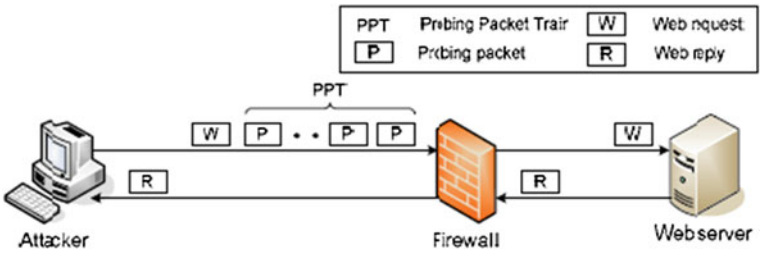


Fig. 1 Detection of last matching rules

3 Entropy-Based Counter Measure Technique

Many Applications have typical packet sizes with respect to request and response or data and acknowledgements (e.g., FTP traffic mostly consists of data packets with 1,500 bytes and acknowledgement packets with 40 bytes); on the other hand, attacks produce packets independent of response from the victim hence packet size distribution is changed [4]. Analysis of packet size distribution is useful to detect low-rate DoS attack; Fig. 2 depicts normal data-flow distribution during FTP transfer and ICMP attack.

Packet size distribution can be described by entropy which measures the degree of dispersal or concentration of a distribution. By observing time series of the entropy of packet size, we can expose the changes in packet size distribution and detect attack points.

When a packet arrives, the entropy is updated according to the following equation:

$$E(t) = \sum (nl/s)\log(nl/s) \tag{1}$$

where

- s is the observation window size (number of packets in the window),
- nl is the number of times packets of size l appear in the observation window.

The entropy takes small values when size distributions of observed packets are concentrated (i.e., all packets are of same size) and takes large value when the size distribution is dispersed. The threshold of the entropy for reporting an alarm can be deduced by self-learning from legitimate traffic data for a period of time. Once threshold is established, an entropy smaller than the threshold is considered as an

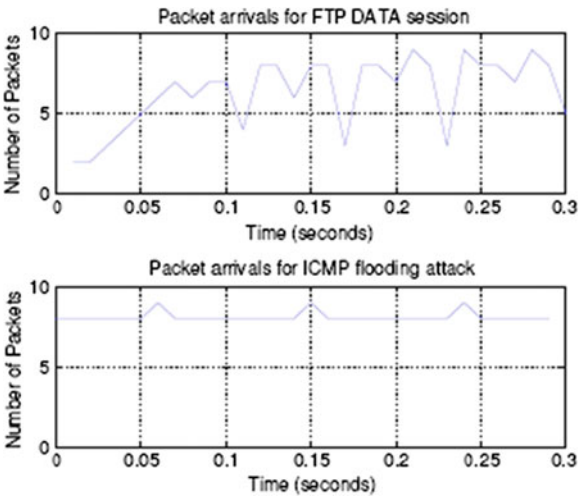


Fig. 2 Packet size distribution of legitimate and attack traffic

indication of possible DoS attack. This approach thus distinguishes between legitimate and attack traffic.

The countermeasure [5] is to endure attack attempts without denying service to legitimate clients. The legitimate packets are placed in a queue with high priority, and attack packets are placed in another queue with low priority. The firewall keeps processing the packets from the first queue whenever there are available packets. The packets in the second queue are processed when there are no packets in the first queue. This approach gives the legitimate packets high importance and they are processed faster. However, the suspected packets are never blocked. Instead, such packets are delayed to give way for legitimate packets to be processed first.

4 CPU Load–Based Countermeasure Technique

CPU load [6] is the measure of the amount of work that a CPU performs. The CPU load increases as there is increase in number of processes. CPU load of the system containing firewall will be observed, and a threshold of the CPU load for reporting an alarm can be deduced by self-learning from legitimate traffic; the increase in CPU load beyond the threshold indicates the possibility of DOS attack.

We can observe the CPU load in UNIX system by inspecting the load average value which represents the average system load over a period of time. It conventionally appears in the form of three numbers which represent the system load during the last one-, five-, and fifteen-minute periods.

We can query the current CPU load from a Unix shell by running the `uptime` command:

```
$ uptime
09:53:15 up 119 days, 19:08, 10 users, load average: 1.73
0.50 7.98
```

Interpretation of load average of "1.73 0.50 7.98" on a single-CPU system is as follows:

- during the last minute, the CPU was overloaded by 73 % (1 CPU with 1.73 runnable processes, so that 0.73 processes had to wait for a turn),
- during the last 5 min, the CPU was under loaded 50 % (no processes had to wait for a turn),
- during the last 15 min, the CPU was overloaded 698 % (1 CPU with 7.98 runnable processes, so that 6.98 processes had to wait for a turn).

For example, if the threshold of CPU load (load average) is 3, then Dos is detected during time interval between 5 and 7 time units (Fig. 3).

The countermeasure is to ask all clients to solve a computation difficult problem that takes more processing during possibility of DOS attack for each request. The clients that provide the result are serviced. Obviously, the client generating attack

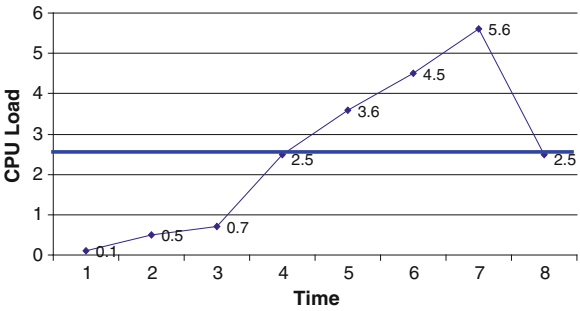


Fig. 3 CPU load–based DOS detection

traffic gets more number of NP-hard problems to be solved as they send more requests and its CPU cycles are used to solve this problem and may run out of resources thus stops generating attack traffic.

The algorithm for finding threshold of the CPU load is

1. Run the uptime command and find load average.
2. Observe the load average value for a period of time during legitimate traffic and identify a the threshold value for CPU load.
3. Whenever the CPU Load is beyond the threshold value DoS is detected.
4. Whenever DoS is detected, issue a computational difficult problem to all client requests.
5. Service the one who provides the answer.

5 Conclusion

The potential probing technique can be used to detect last matching rules of a firewall; an attacker can trigger these rules to endure low-rate DoS attack on firewalls. Entropy-based countermeasure technique distinguishes between legitimate and attack traffic, and high priority is given to legitimate traffic. This approach is not free from false-positive results and consumes resources of the server since entropy calculation has to be done in regular time intervals also different queues, one for assumed attack traffic and one for legitimate traffic has to be maintained along with scheduling the processes. CPU load–based countermeasure detects DoS attack based on CPU load and asks all clients to solve a computational difficult problem. This approach steals CPU cycles of the attacker, but resources of legitimate clients are wasted and do not solve IP spoofing attack.

Acknowledgments We acknowledge the support of the National Institute of Engineering in the completion of this work and also acknowledge authors whose paper has been referenced.

References

1. Wikipedia a free encyclopedia http://en.wikipedia.org/wiki/Firewall_(networking)
2. Wikipedia a free encyclopedia http://en.wikipedia.org/wiki/Denial of service_(networking)
3. Salah K, Sattar K, Sqalli M, Al-Shaer E (2008) A probing technique for discovering last matching rules of a network firewall. In: The 5th IEEE international conference on innovations in information technology, Innovation'08
4. Ping D, Shunji A (2007) Detecting DoS attack using packet size distribution. In: Bio-inspired models of network, Information and computing systems
5. Al-Haidari F, Sqalli M, Salah K, Hamodi J (2009) An entropy-based countermeasure against intelligent DoS attacks targeting firewalls In: IEEE international symposium on policies for distributed systems and networks
6. Wikipedia a free encyclopedia http://en.wikipedia.org/wiki/Load_(computing)

A Practical Current Measurement Using Rogowski Coil and Haar Wavelet

T. V. Sushma, H. K. Muttanna Kadal, D. Jatin, T. S. Sheshadri and N. Ramesh

Abstract Digital signal processing using 'Fourier transforms' and Haar wavelets is well known. However, literature does not seem to contain references with detailed practical approach to digital signal processing. In this article, detailed description of a current pulse transient captured using a digital oscilloscope and improvised Rogowski coil is used to show that other tools such as 'PSpice simulation' and some heuristic reasoning coupled with domain knowledge are necessary for a successful application of the Fourier and Haar analysis. The simplicity of the circuit used for the exploding wire setup for plasma creation enables a practical demonstration of the need for a careful modeling of the system. System identification is confirmed with alternate practical measurements of one of the parameters.

Keywords Haar wavelet · Rogowski coil · Plasma current pulse · Transient measurement

1 Introduction

The wire explosion technique is one of the well-known methods of plasma generation. This involves exploding the wire by applying a high voltage to the wire in a short duration. This is a transient process. Determination of the power delivered to the load and hence the efficiency of system require accurate estimate of the resistance offered by the system. Manual measurement resulted in a value of

T. V. Sushma · N. Ramesh
S. J. C. Institute of Technology, Chickballapur, India

H. K. Muttanna Kadal (✉) · D. Jatin · T. S. Sheshadri
A. P. D. L., Department of A. E., Indian Institute of Science, Bangalore, India
e-mail: muttanna.kadal@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_16,

0.2 Ω. A more accurate estimate required experimental analysis. An emf waveform captured was used for the estimation. However, it was found that the signals captured in real time were corrupted by various external parameters, leading to erroneous measurements. Computation of the resistance using the noisy signal resulted in a resistance value that was away from the measured values by about 64 %. This prompted the use of signal processing tools on the captured waveforms like Fourier analysis and wavelet transforms for better values. It was seen, however, that the use of signal processing tools could be exploited better by gaining knowledge of the underlying system. This required modeling the system first to identify the events inherent to the process followed by the application of signal processing tools.

2 Wire Explosion Process

The wire explosion process is one of the well-known techniques of plasma generation for electrothermal chemical (ETC) guns. This was carried out in a plasma parametric chamber and involved application of a high voltage of the order of 2 kV in a short duration using a surge generator. The process is shown in the Fig. 1.

The process described is a transient process. In order to determine the efficiency of the system, it is required to estimate the resistance offered by the system. The resistance offered by the circuit determined by manual measurement at various points of the circuit was approximately 0.2 Ω. An accurate value of resistance had to be determined experimentally. One approach was to determine the current flowing through the system and then compute the resistance by using Ohm's law. To determine the resistance offered, a reduced system was used in which the load was shorted, that is, the wire was removed and the experiment was repeated by the application of a voltage of 500 V. Since the current through the circuit cannot be measured directly, a Rogowski coil was designed for the measurement of current.

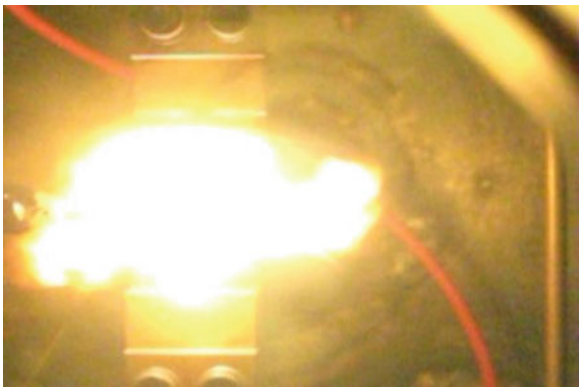


Fig. 1 Plasma explosion process

Rogowski coil is a well-known method of transient current measurement and measures the emf induced in the coil due to the current flowing through the conductor. The induced emf is proportional to the rate of change of current, d*i*/d*t* through the circuit. The current can be obtained by integrating the current derivative waveform. The emf induced in the Rogowski coil is given by Eq. 1.

$$e = -\frac{\mu NA}{l}\frac{\mathrm{d}i}{\mathrm{d}t} \tag{1}$$

where,

μ	Magnetic permeability of the material
N	The number of turns in the coil
A	Area of the coil
l	Length of the coil
d*i*/d*t*	Rate of change of current through the coil.

The waveform was captured by using a Tektronix Digital Storage Oscilloscope TDS100, a 60 MHz scope that provides a sampling rate of 1 G samples/s. The waveform captured is as shown in the Fig. 2.

It is seen that the captured waveform is distorted due to which the analysis of the signal becomes difficult. However, analysis was performed on the waveform recorded. Integrating the d*i*/d*t* values obtained from the emf waveform recorded, the peak current was computed to be 6.96 kA, which resulted in a resistance value of 0.07 Ω. This is away from the measured value by about 64 %. To ensure correctness of the value computed, it is essential to identify the components which are not inherent to the process and eliminate them. Proper modeling of the system is necessary to explain the recorded waveform. Signal processing tools can then be employed to clean up the recorded waveform to generate waveforms closer to ideal conditions for accurate estimation of the electrical parameters.

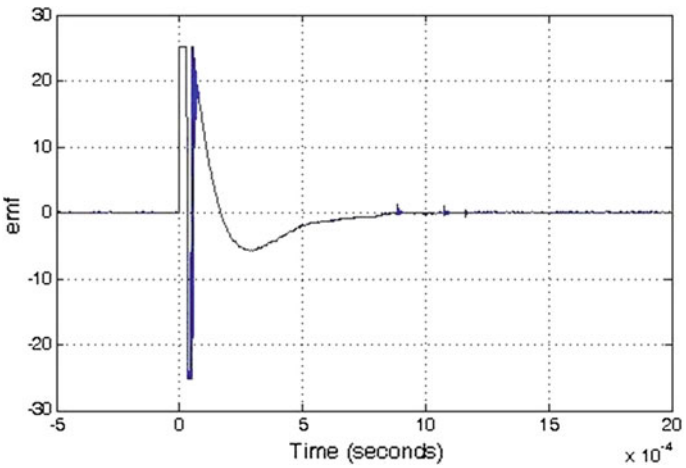


Fig. 2 Emf waveform captured using Rogowski coil

3 Modeling of the System

The system used for plasma generation through wire explosion involves the application of a high voltage in a short duration using a surge generator as stated above. This involves charging the capacitors to the required voltage and then discharging it through the load. This indicates that the system could be a simple RC circuit. The circuit simulations were done using the PSpice software.

4 RLC Circuit Simulation

The RC circuit results in a voltage that decreases monotonically. The presence of negative portion in the emf waveform shown in Fig. 2 suggests that the current waveform must be oscillatory in nature. This indicates the presence of an inductance in the circuit and that the system must be an RLC circuit instead of the previously assumed simple RC circuit. The system was modeled as an RLC circuit with a small value of L of 10 μH using PSpice. The di/dt waveform is as shown in Fig. 5.

The di/dt waveform can be obtained directly from the voltage across the inductor waveform by dividing it by the inductance value or by integrating the current waveform. The di/dt waveform in Fig. 3 accounted for the major portion of the plot obtained from the Rogowski coil shown in Fig. 2 except for the high-frequency portion occurring at the vertical trigger point.

5 RLC Circuit with Switch Bounce

One of the possible sources for the occurrence of the high-frequency component was considered to be the reflections in a transmission line which was picked up by the Rogowski coil. In order to verify this line of thinking, the transmission line model was simulated using PSpice. However, it was not possible to obtain

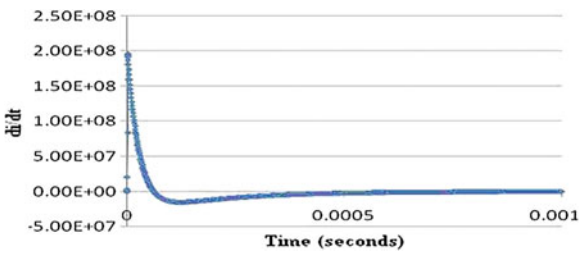


Fig. 3 Derivative of the current waveform, di/dt

satisfactory results using this assumption. Hence, the other possible source of occurrence of the high-frequency component was considered to be the result of switch bouncing which results in the variation in resistance. The effect of switch bouncing was simulated using PSpice. The circuit and the simulation results are shown in Figs. 4 and 5, respectively.

It is seen that the RLC circuit with the switch bounce effect in Fig. 6 is able to simulate the waveform captured using Rogowski coil to a reasonable extent, thus indicating that the high-frequency portion in the emf plot in Fig. 2 is indeed due to the switch bounce and not a part of the process. In other words, it indicates that the high-frequency component in the captured waveform is noise resulting from external factors and not inherent to the process. This component must be removed in order to simulate the emf waveform obtained using Rogowski coil without switch bounce more precisely.

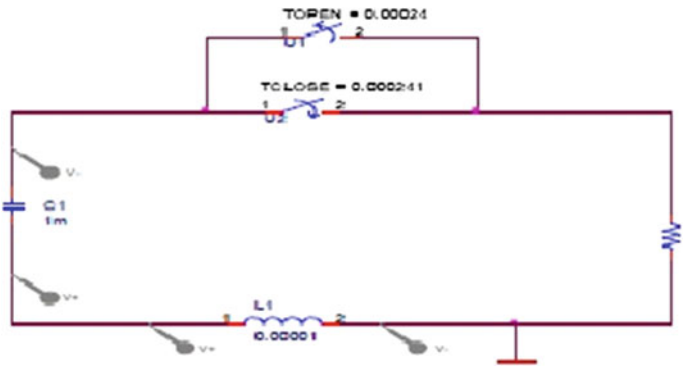


Fig. 4 RLC circuits with switch bounce effect

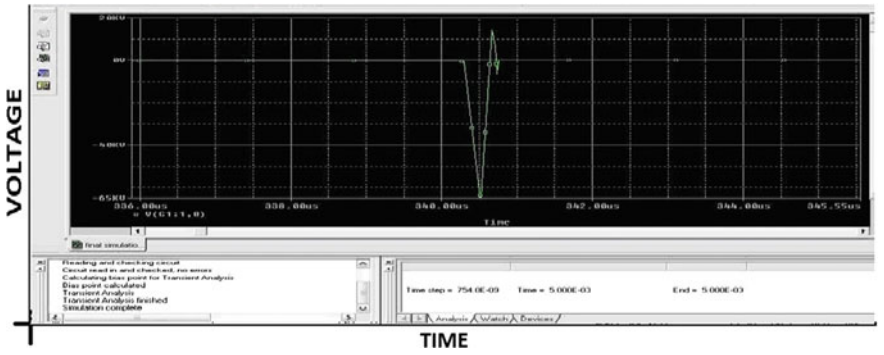


Fig. 5 Voltage waveform

6 Need for Signal Processing

In order to find out the sources and causes of harmonic distortion, one can detect and localize those disturbances [1]. The commonly employed tools are the Fourier transform and the wavelet transform. These tools are used for identifying the spectral components present in the signal. The unwanted frequencies can then be removed. The Fourier transform is one of the oldest signal processing tools used which provides information about the spectral components but does not provide the time information at which they occur. The wavelet transform, however, which is based on the Fourier transform provides an advantage of time localization of frequency component. The application of the two signal processing tools on the electrical signal aiding in an accurate estimation of the resistance has been presented.

7 Application of Fourier Transform

The Fourier transform of a signal shows the spectral components present in the signal. The unwanted spectral components can be removed by setting it to zero and then reconstructing the signal. However, the Fourier transform removes the frequency components throughout the signal and does not localize it to the location at which it is occurring in the signal. The discrete Fourier transform of the emf plot shown in Fig. 2 is shown in the Fig. 6.

From the Fourier transform of the recorded waveform, it is seen that most of the energy is concentrated in the low-frequency region with a large magnitude of around 2,000 units. Considerable signal energy with a magnitude of 1,000 units is found to occur at a frequency of 10 kHz. Reconstruction of the signal after setting the magnitude of the high-frequency spectral components to zero using the inverse Fourier transform results in the signal as shown in Fig. 7.

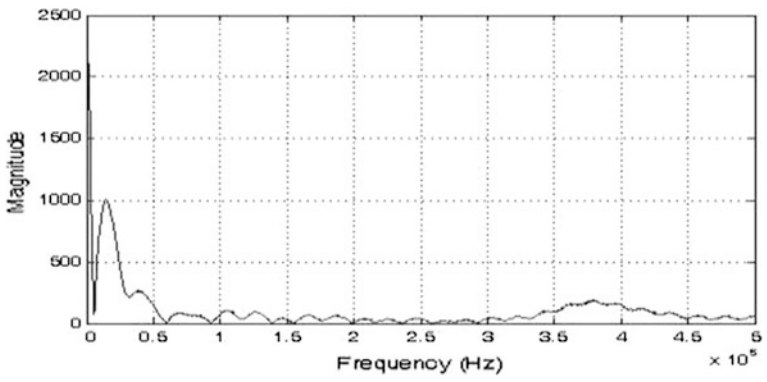


Fig. 6 Fourier transform of the emf waveform recorded using Rogowski coil

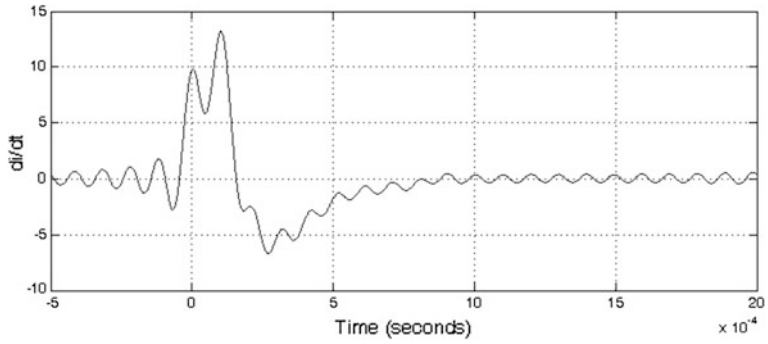


Fig. 7 Reconstructed signal after the removal of high-frequency components

It is seen that the removal of the high-frequency components has resulted in a drastic change in the shape of the signal as compared to the original signal. This is because the Fourier transform does not provide time localization of the frequency and removes the spectral components throughout the signal. This shows that it is not possible to localize the spectral component that causes the distortion which needs to be eliminated. This prompted the use of wavelet transform which provides time localization of the spectral components enabling the identification and elimination of the particular spectral component that causes the distortion.

8 Application of Wavelet Transform

The wavelet transform provides the advantage of time localization of the spectral components of the signal, thus aiding in the removal of the spectral component at the required time instant and hence found to have wide spread use in analysis of transient signal. The analysis is capable of revealing aspects of data that other signal analysis techniques miss, including aspects such as trends and breakdown points [2]. The basic Haar wavelet was applied to the signal. The magnitude of the highest frequency component of the signal obtained using Haar wavelet is shown in Fig. 8.

It is seen that most of the energy is distributed in the low-frequency components with a spike of large magnitude occurring in the high frequency at around 0.3 μs. Setting the high-frequency components represented by the detailed coefficients to zero and performing the wavelet analysis for 10 levels result in the waveform as shown in Fig. 9.

It is seen that the processed signal shown in the Fig. 9 is reasonably closer to the signal obtained under ideal conditions. The rate of change of current is obtained by dividing the computed emf by the mutual inductance of the coil. The positive portion of the resulting di/dt plot is integrated to obtain the current flowing through the circuit over time. A peak current of 2.89 kA is obtained. This resulted

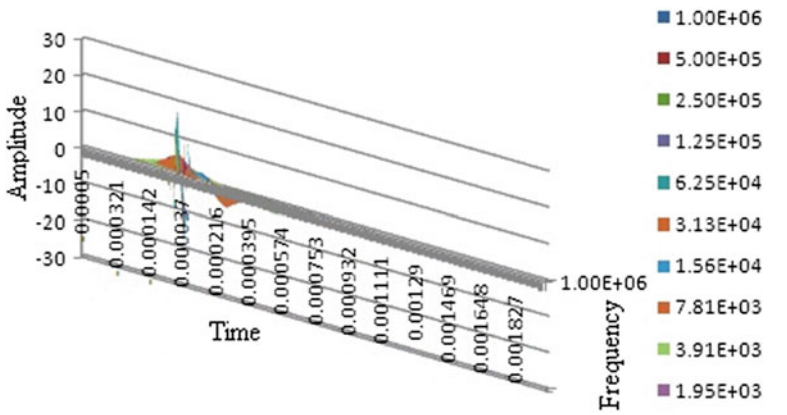


Fig. 8 Haar wavelet transform of the emf

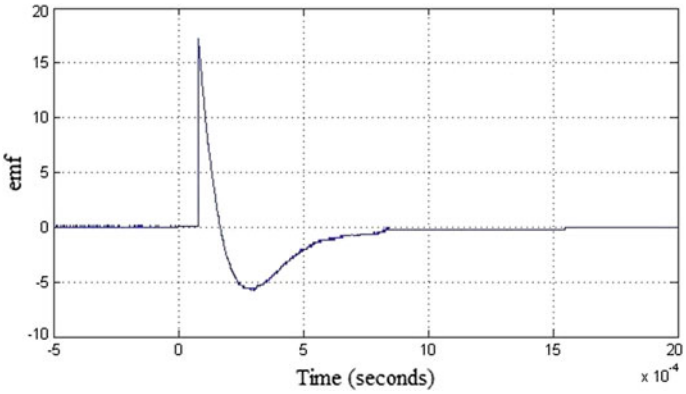


Fig. 9 Emf waveform processed by using Haar wavelet

in a resistance value of 0.17 Ω which is much closer to the manual estimation. System modeling thus proves to be useful to identify the components that are inherent to the process. The signal processing tools can then be exploited to obtain waveforms closer to that explained by the process.

9 Results

The waveform captured using the Rogowski coil has been simulated using an RLC circuit with switch bounce. The occurrence of the high-frequency component in the emf waveform is deduced to be a result of switch bounce. The removal of the high-frequency component using Fourier transform causes a drastic change in the

signal since it does not provide time localization of the frequency components and removes the spectral component throughout the signal. Hence, Haar wavelet which provides time localization of the frequency component has been applied.

The signal values of the emf waveform at different time instants and the waveform itself have been recorded in the Tektronix oscilloscope and imported to Microsoft Excel. Code to determine the Fourier transform and application of Haar wavelet on the signal captured have been done using MATLAB version 7. The resulting values on application of Haar wavelet have been imported to Microsoft Excel to plot the frequency components occurring at various instants as shown in Fig. 8.

The high-frequency component is found to occur at 0.3 μs and is removed by setting the detailed coefficients to zero followed by the application of wavelet transform. Resistance value is estimated to be 0.17 Ω which is closer to that estimated using manual measurement.

10 Conclusion

The paper shows that the use of the domain knowledge followed by the application of signal processing tools aids in detecting and classifying the components inherent to the signal. This aids in better analysis of the signal. Modeling the system by using the PSpice allowed the testing of the various parameters that could have resulted in the observed process.

The nature of the process helps in employing the appropriate signal processing tool. The wavelet transform is found to be a better tool for the analysis of the transient signal than the Fourier transform since it provides a time localization of the frequency component. Computation using the recorded waveform without processing resulted in a value of resistance that was away from the manually measured values by 64 %, indicating that the estimation must be erroneous. The application of wavelet transform followed by computation resulted in much closer values. Thus, domain knowledge is necessary for exploiting the signal processing tools.

References

1. Montaño JC, Borrás D et al (2007) Wavelet-Fourier analysis of electric signal disturbances, ICREPQ
2. Ward DA, Exon JLT (1993) Using Rogowski coils for transient current measurement. Eng Sci Educ J. http://www.onepetro.org/mslib/servlet/onepetropreview?id=SPE-88559-PA&soc=SPE

Improved Accuracy of Higher-Order Legendre Polynomial Roots in MATLAB

D. Jatin, H. K. Muttanna, T. S. Sheshadri and N. Ramesh

Abstract Legendre polynomial (LP) has found extensive use in solutions of various physical phenomena. The roots of LP up to 44th order can be obtained using the popular and widely available MATLAB (7.5.0 R2007b) library function 'roots' which yields real roots only up to order 44. The solution is also found to have large errors due to limited precision in MATLAB. To obtain accurate roots of LP in MATLAB, it is very important to obtain accurate LP coefficients. It is possible that other mathematical software like Maple do not have this limitation. This article explores the roots of 44th-order LP used in Gaussian quadrature in MATLAB. The accuracy of higher-order LP roots has been improved using the variable precision integer (VPI) format in MATLAB. MATLAB's 'roots' function and VPI method are also compared.

Keywords Higher-order Legendre polynomial · Gaussian quadrature · Variable precision integer (VPI)

1 Introduction

Dealing with huge numbers and higher-order polynomial has always been limited by the accuracy available in the computer since it can deal with only 0–2n in the positive range. Because of this limit, for higher-order Legendre polynomial (LP), the coefficients are rounded off. A LP solution of 44th has been published in Ref. [1]. However, on verification, we found errors as the roots on substitution in the

D. Jatin (✉) · H. K. Muttanna · T. S. Sheshadri
A. P. D. L., Dept, A. E.,I.I.Sc, Bangalore, India
e-mail: jatinspace2000@gmail.com

N. Ramesh
S.J.C. Institute of Technology, Chickballapur, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_17,

polynomial did not yield zero. In this paper, we expand on this further. A discussion on the difficulties in calculating the higher-order polynomial coefficients and its solution has been provided. A method of improving the root accuracy has also been presented.

1.1 Application

LP has an important role in image processing, signal processing, and control like a fingerprint ridge orientation model based on LP [1] and [2]. LP is applied to problems in mathematics and physics, for example the solution of Laplace's equation in spherical co-ordinate system. An example for such a case is the quantum mechanics for solving the potential at a point due to an electrical charge 'q'. It is also applied to determine the guided waves in homogeneous infinitely long cylinders utilizing elastic materials of cylindrical anisotropy [3]. Higher-order polynomial is also used in thin piezoelectric structure, which is an important part of active noise cancellation devices and smart structures [4].

2 The Legendre Polynomial ($P_n(x)$)

The LP is named after Adrien-Marie Legendre. The ordinary differential equation associated with LP is frequently encountered in physics and other technical fields [5]. The Gaussian integration method that is widely used as an integration solution can be applied using various orthogonal polynomials, among which the Gaussian–Legendre integration method is the most widely used. LP can be defined in a very elegant and compact way by the Rodrigue's formula [Eq. (1)]. [5].

$$P_n(x) = \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(x^2 - 1\right)^n \tag{1}$$

The determination of the LP and the coefficient using the above formula is not accurate. Better results can be obtained using the following iteration formula in a much simpler way [Eq. (2)].

$$P_n(x) = \frac{2n-1}{n} \cdot x \cdot P_{n-1}(x) - \frac{n-1}{n} \cdot P_{n-2}(x) \tag{2}$$

where $Po = 1 \quad \& \quad P1 = x$. This formula is very popular, but from the point of view of numeric calculus, it has one disadvantage. Its coefficients are decimal, and this causes round-off errors, leading to inaccuracy for higher polynomial degree. It is convenient to rearrange the iterative formula to avoid fractional coefficients. But as the order of polynomial increases, the accuracy of the Legendre coefficients will decrease because of the limited accuracy supported by computer. A solution to find

the coefficients of any order LP with very high accuracy has been provided later in this paper. More information about the integer calculation of coefficient can be found in [6].

3 Explanation of VPI

MATLAB is able to deal only up to 16 decimal places, after which MATLAB rounds off the number. MATLAB variable precision integer (VPI) is a MATLAB tool, which was developed by John D'Errico which deals with huge integers of arbitrary size [7]. VPI allows us to deal with any arbitrary size number.

4 Theoretical Methods of Generating Coefficients

The LP of order n is denoted as $P_n(x)$. The polynomials are either odd or even functions of x for even or odd orders of n. The general form of an LP of order n is given by the summation [8]

$$P_n(x) = \sum_{m=0}^{M} (-1)^m \frac{(2n-2m)!}{2^n m!(n-m)!(n-2m)!} \times x^{n-2m} \tag{4}$$

$M = n/2$ or $n - 1/2$, whichever is an integer.

One of the properties of LP is that the summation of all the coefficients is equal to 1.

$$\sum_{i=0}^{n} a_n = 1 \tag{5}$$

n order of LP
a_n the coefficient of LP

After the 44th order, coefficients calculated using Eq. (4) in MATLAB do not follow the property mentioned in Eq. (5). The error in the coefficients at different orders is mentioned in Table 1. The calculated error increases exponentially for higher-order LP.

The inserting the randomly selected roots in respective order of polynomials in MATLAB is shown in Table 2. All the roots are calculated using MATLAB library function call 'roots'. After 44th order, roots calculated by MATLAB become imaginary. The plots of evaluation of these roots of orders 5 and 44 are shown in Fig. 1. It can be concluded that the calculated roots using MATLAB library function in the range of [−1,−0.7]–[0.7,1] have more error compared to other.

Table 1 Error in coefficient calculation

Order	Root	P_n(root)
25	−0.99555697	3.40E−08
43	0.072152991	4.8E−13
43	0.072152991	4.8E−13
44	0.035289237	1.9E−13
44	−0.997979603	0.0078125
45	−0.996354097477772 + 0.000919574626438i	0.272949218750000 + 0.007003784179688i
45	0.06898698	3.677E−12

Table 2 Some roots of $P_n(x)$ and P_n(root)

Order	Summation of coefficients
1	1
25	0.999999993
30	1.000002488
44	1.01159668
50	14.109375
60	298560
70	755482624
80	4.03E + 12
90	1.20E + 16
100	−1.11E + 20

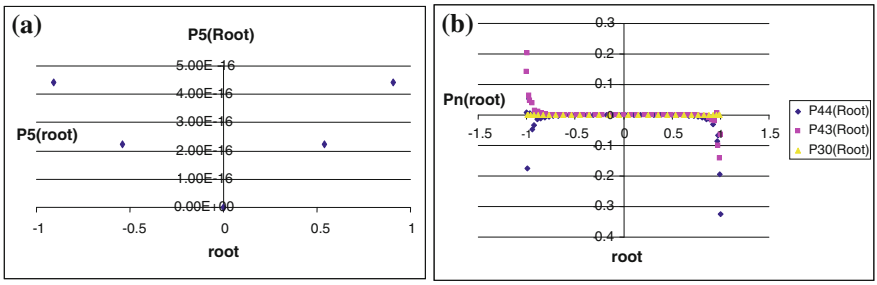


Fig. 1 Evaluation of roots. **a** 5th-order LP. **b** 30th-, 43rd-, and 44th-order P_n(root)

5 Integer Coefficient Calculation Method [6]

It is convenient to rearrange the iterative formula to avoid fractional coefficients. Let us assume that the LP can be written in fractional form as

$$P_n(x) = \frac{1}{k_n} L_n(x) \tag{6}$$

where k_n is an integer number and $L_n(x)$ is a polynomial having integer coefficients. The Legendre' polynomial $P_n(x)$ is then fully defined by the couple $[k_n, L_n(x)]$. Now, 'K0 = 1, L0 = 1' and 'K1 = 1, L1 = x' and substituting in the recurrent Eq. (2),

$$\left(\frac{L_n(x)}{k_n}\right) = \frac{2n-1}{n} x \left(\frac{L_{n-1}(x)}{k_{n-1}}\right) - \frac{n-1}{n}\left(\frac{L_{n-2}(x)}{k_{n-2}}\right) \tag{7}$$

$$\left(\frac{L_n(x)}{k_n}\right) = \frac{k_{n-2}\cdot(2n-1)\cdot x\cdot L_{n-1}(x) - k_{n-1}\cdot(n-1)\cdot L_{n-2}(x)}{n\cdot k_{n-1}\cdot k_{n-2}} \tag{8}$$

Separating the numerator and denominator, we obtain the following two recurrence equations

$$\begin{aligned} L_n(x) &= k_{n-2}\cdot(2n-1)\cdot x\cdot L_{n-1}(x) - k_{n-1}\cdot(n-1)\cdot L_{n-2}(x) \\ k_n &= n\cdot k_{n-1}\cdot k_{n-2} \end{aligned} \tag{9}$$

$$\begin{aligned} V_n(x) &= k_{n-2}\cdot(2n-1)\cdot x\cdot L_{n-1}(x) - k_{n-1}\cdot(n-1)\cdot L_{n-2}(x) \\ b_n &= n\cdot k_{n-1}\cdot k_{n-2} \\ C_n &= \gcd\left(b_n, \mathrm{coef}(V_n)\right) \end{aligned} \tag{10}$$

where the coef() operator returns the vector of all coefficients of the polynomial $V_n(x)$ and the GCD is the greatest common divisor. The final **LP** is

$$k_n = \frac{b_n}{C_n}, \quad L_n(x) = \frac{1}{C_n} V_n(x) \tag{11}$$

It can be seen from Table 1 that the accurate estimation of the coefficients is difficult. But using the VPI MATLAB format and the method for calculating integer LP coefficients, it is possible to determine the LP coefficients accurately.

6 Results

Here, we will see the accuracy between the roots calculated using two methods.

- **In the 1st method**, floating number coefficients are calculated as per Eq. (4) and the roots of LP are calculated using MATLAB's library function.
- **In the 2nd method**, integer number coefficients are calculated using VPI format as per Eq. (8) and the roots are searched based on sign change detection. Because of the VPI format, the MATLAB library function cannot be used to do the arithmetic operation. So, separate MATLAB code is developed to deal with VPI numbers. Synthetic division is generally used to speed up the algorithm, but it reduces the accuracy of coefficients. To avoid this error, the numerator and denominator are stored separately.

Table 3 Comparison between MATLAB roots and VPI format roots for 20th-order LP

a_n	Coefficient (Kb = 262144)	Roots calculated by 2nd method	1st method root	Evaluation of the roots calculated by 1st method [P_{20}(root) x 1e-9]
0	46189	−0.9931285991850949247861	−0.993128599192866	0.931322574615478
1	0	−0.9639719272779137912672	−0.963971927258452	0.603904481977224
2	−9699690	−0.91223442825132590 5868	−0.912234428271958	0.243744580075145
3	0	−0.8391169718222188233949	−0.839116971809323	0.129602995002642
4	334639305	−0.746331906460150792614	−0.746331906464916	0.027682744985214
5	0	−0.636053680726515025453	−0.636053680725781	0.004368061468085
6	−4461857400	0.9931285991850949247861	0.993128599193708	0.858562998473644
7	0	0.9639719272779137912672	0.963971927256709	0.589352566748858
8	30117537450	0.91223442825132590 5868	0.912234428273316	0.298314262181520
9	0	0.8391169718222188233949	0.839116971808666	0.122781784739345
10	−116454478140	0.746331906460150792614	0.746331906465230	0.035583980206866
11	0	0.636053680726515025453	0.636053680725632	0.004789058039023
12	273491577450	−0.51086700 1950827098004	−0.510867001950672	−0.000746763761938 −0.000478311834584
13	0	0.51086700 1950827098004	0.510867001950714	
14	−396713057400	−0.373706088715419560672	−0.373706088715502	−0.000325135209554
15	0	0.373706088715419560672	0.373706088715495	−0.000305929110170
16	347123925225	−0.22778585114164507808	−0.227785851141637	−0.000031077518123
17	0	0.22778585114164507808	0.227785851141637	−0.000028893978877
18	−167890003050	−0.076526521133497333754	−0.076526521133498	−0.000000643824125

The Table 3 presents roots calculated by two different methods for 20th order. The LP coefficients for 44th-order LP are given in Appendix. In both the cases, a zero band is defined as < 1e-15.

The 3rd column of Table 3 presents the roots of 20th order calculated by integer coefficient (2nd method). Here, Kb = 262144 is the GCD of 20th-order LP integer coefficients. If we sum all the integer coefficients given in column 2, then it is equal to Kb. The last two columns presented the roots calculated by the 1st method and evaluation of the root in P20 polynomial function, respectively. In the 2nd method, zero band is defined as < 1e-15.

The root '$x = 0.963971927256709$' is taken from the 1st method and substituted in 20th-order LP. As same way, the root '$x = 0.9639719272779137912672$' calculated by the 2nd method is substituted in 20th-order LP, which has integer coefficients. It has been observed that there is a huge error in root calculation in the 1st method.

$P_{20}(0.963971927256709)$ [1st method]
=

1.4666622175551324331718658907936598841198025063278469469018617310928152392751320282622741257763132579641653759926741916606809856796775630876190324393455201351257354678953080900159269713685946023128

$P_{20}(0.\ 963971927279137912672)$ [2nd method]
=
3.224068593922932134010646622869335519879460983677625784294537912961405016641140514210329174185577968947784529056564942681561904721779750596573642053369400737516054964176452685665378328427278 **E-15**

The roots calculated using two different methods of order 44th are compared in Appendix. The integer coefficients of 44th-order LP are shown in Appendix Table A.1.

7 Conclusion

Now, we have a solution to calculate the accurate coefficients for any order of LP. We also presented the more precise and accurate roots up to 44th-order LP. As we have a real root for any order, it is possible to improve root to any desired accuracy. The presented VPI format is not only used in LP, but it can be used wherever higher digital accuracy is required. The roots of 44th order are highly accurate as compared to MATLAB 'roots' function. The presented roots of 44th order can be directly used in 'Gaussian–Legendre' integration method. The roots obtained using VPI format are expected to improve the result in the field of physics

problem, signal and image processing, control system, quantum mechanics, etc. It is also possible to calculate the root to any accuracy so that the remainder of $P_n(x)$ approaches to zero.

Table A.1 Integer coefficient of 44th-order LP

a^n (Coefficient position)	Coefficients (Kb = 2199023255552)
0	263012370465
1	0
2	−260382246760350
3	0
4	42832879592077575
5	0
6	−2798414800015734900
7	0
8	96845140757687397075
9	0
10	−2053116984062972817990
11	0
12	29085823940892114921525
13	0
14	−291497488286742953938800
15	0
16	21497939761147292852986 50
17	0
18	−11999503631385482415980700
19	0
20	51724176179498474203622070
21	0
22	−17465306242428056224599 6600
23	0
24	46637430074889411005543 2950
25	0
26	−99014851543611364904076 5340
27	0
28	16738224903800968829022 46170
29	0
30	−22471547916827047807239 35088
31	0
32	23785408581318951812098 10325
33	0
34	−19587983537556783845257 26150
35	0
36	12281354757674491458534 31475
37	0
38	−56602545398101496907042 9300

(continued)

A.1 Appendix

Table A.1.

Table A.1 (continued)

a^n (Coefficient position)	Coefficients (Kb = 2199023255552)
39	0
40	18069274107855477858786781 5
41	0
42	−35676847832002685667755550
43	0
44	32810631727106063986202 25

References

1. Ram S, Bischof H, Birchbauer J (2008) Curvature preserving fingerprint ridge orientation smoothing using Legendre polynomials. IEEE computer society conference on computer vision and pattern recognition, pp 1–8, 23–28 June 2008
2. Aburdene M (1996) Recursive computation of discrete Legendre polynomial coefficients. Department of Electrical Engineering, Bucknell University, Lewisburg, PA 17837, pp 221–224
3. Elmaimouni L, Lefebvre JE, Zhang V, Gryba T (2005) A polynomial approach to the analysis of guided waves in anisotropic cylinders of infinite length. Université de Valenciennes, 18 Jan 2005
4. Hauck A, Kaltenbacher M, Lerch R (2006) 5E − 3 simulation of thin piezoelectric structures using anisotropic hierarchic finite elements. Dept. of Sensor Technol., IEEE Ultrasonics Symposium, pp 476–479, 2–6 Oct 2006
5. http://en.wikipedia.org/wiki/Legendre_polynomials, 07/03/2009
6. http://digilander.libero.it/foxes/poly/Poly_Legendre_Tables.htm, 6/04/2009
7. http://www.mathworks.com/matlabcentral/fileexchange/22725, 2/04/2009
8. Mughal AM, Ye X, Iqbal K (2006) Computational algorithm for higher order Legendre polynomial and Gaussian quadrature method. In: Proceeding of the international conference on scientific computing, Las Vegas, NV, USA, 26–29 June 2006

Stability Analysis of Speech Synthesis Filter of CELP-Based AMR-WB Codec

D. Jatin, T. S. Sheshadri, N. Ramesh and H. K. Muttanna

Abstract The speech codec analyzes the speech using A(z) (analysis filter) and synthesizes back at decoder side using linear prediction coefficients (LPC). These LP coefficients are sensitive and cannot be sent directly in a transmission channel. A small corruption in LPC values during transmission destroys the synthesized speech at the decoder side. We have presented new results on the stability and sensitivity of LPC based on changes in speech input pitch length, sign bit, and LPC values during transmission (or for any other reason) consecutively and simultaneously. Present analysis will help to add varying dynamic range to LSF coding. For this each individual LPC need to be related to each LSF. All the speech inputs considered in this study are voiced speech, which has been separated manually. For a specific order, we analyzed the numbers of LPC which are more responsible for increase in prediction error at decoder side when they are corrupted by noise. Present analysis provides the reference for number of bits required for quantization of LPC or line spectral pairs (LSF).

Keywords Codec · Formant frequency · All-pole synthesis filter · Linear prediction coefficient · Stability and sensitivity of LPC · Line spectral frequency · Symmetric and anti-symmetric polynomial

1 Introduction

The adaptive multi-rate wideband (AMR-WB) codec is the first codec adopted as a standard for both wireless and wire-line services [1]. The code-excited linear prediction (CELP)-based AMR-WB interoperates with the latest 3GPP2 (Third

D. Jatin (✉) · T. S. Sheshadri · H. K. Muttanna
A. P. D. L. Department A. E., Indian Institute of Science, Bangalore, India
e-mail: jatinspace2000@gmail.com

N. Ramesh
S. J. C. Institute of Technology, Chickballapur, Bangalore, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_18,

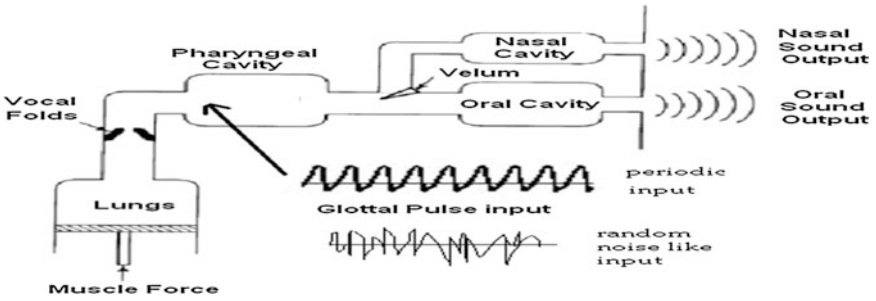


Fig. 1 Simplified human speech production system

Generation Partnership Project) wideband standard [2, 3]. The simplified human speech mechanism is shown in Fig. 1. The frequencies which pass through vocal tract and peak in spectrum are called formant frequencies. These formant frequencies are used in 'prediction filter A(z)' to compress the speech as in Eq. 1 [4].

$$A_p(z) = 1 + \sum_{k=1}^{P} a_k(k) \cdot z^{-k} \tag{1}$$

$A_p(z)$ Predication function or speech prediction filter (analysis filter).
a_k LP coefficient, k = 1, 2,…, p. where p is the order of synthesis filter.

A(z) is an analysis filter and its coefficients are called the LPC, which predict the next sample using previous 'P' samples of the input stream [5]. A small change in LPC during transmission can destroy the synthesis of speech at the decoder side. Further, the roots of A(z) are complex, and such sensitivity parameters cannot be quantized with a small error [6]. Therefore, for improved performance of the AMR-WB, its stability has to be considered. This chapter is organized as follows:

1. Introduction of the LPC calculation using 'Levinson–Durbin algorithm' described in Sect. 2.
2. In Sect. 3, a new study on sign bit and magnitude corruption by the noise and its effect on sensitivity of A(z) is presented.
3. In Sect. 4, the sensitivity of A(z) with variation in pitch is analyzed. Stability improvement in A(z) is also shown in this section.

2 LPC Calculation [7, 8, 9]

An all-pole synthesis filter '1/A(z)' is a digital representation of the vocal tract filter.

$$\hat{S}(n) = a_1 \cdot S(n-1) + a_2 \cdot S(n-2) + \cdots + a_p \cdot S(n-p) \tag{2}$$

The error between the predicted and the actual can be written in Eq. (3) as

$$e(n) = S(n) - \hat{S}(n) = S(n) - \sum_{i=1}^{p} a_i \cdot S(n-i) \tag{3}$$

where $S(n)$ is the actual nth sample and $e(n)$ is the prediction error in nth sample $\hat{S}(n)$ predicted nth sample. The LPCs are determined using the mean square error criterion as in Eq. (4).

$$E = \sum \left[e_n^2\right] = \sum \left(s(n) - \sum_{i=1}^{p} a_i S(n-i)\right)^2 \tag{4}$$

Here, 'i' is varying from 0 to p. The optimum value of the $\hat{S}(n)$ is found, when $E\left[e_n^2\right]$ is minimized. It can be done by partial differential of mean square error with respect to each 'a_i', which can solved using faster 'Levinson–Durbin algorithm' [6]. After evaluating the LPC, it needs to be converted into LSF for quantization purpose and stability. The LPC to LSF can be calculated using [10–13].

3 Effect on Residual Error due to Corrupt Sign Bit and Value of LPC

3.1 Sign Bit Corruption

The LP coefficients require 16/32 bits, in which MSF (most significant bit) is a sign bit. The sign bit (+/−) of LP coefficients is corrupted by flipping consecutively on voice speech inputs for order 16. One of them has been shown in Fig. 2. The sign bit corruption in a_1 to a_5/a_6 causes 1,000 times more prediction error compared to remaining LPCs. The sensitive window region is shown in Fig. 3. This means that the sample of voice segment is more correlated with its past 0.375 ms of data as compared to data further back in time.

3.2 Corrupted LPC Value

- To corrupt the LPC, they are changed simultaneously as per Eq. (5) by 10 % (*n1*) to 160 % from their original values. Figure 4 shows effect on an average prediction error per sample.

$$a_i = a_i + a_i \cdot (n \times 0.1) \tag{5}$$

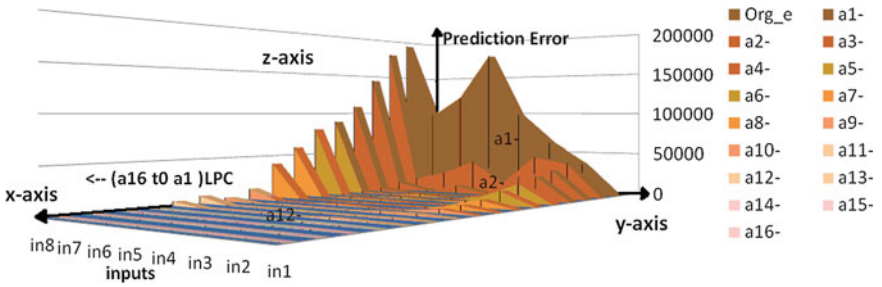


Fig. 2 Average prediction errors due to corrupt sign bit of each coefficient consecutively (*x-axis* LPC, *y-axis* 8 inputs of voice speech at 16 kHz, *z-axis* prediction error in increment in percentage with respect to original prediction error)

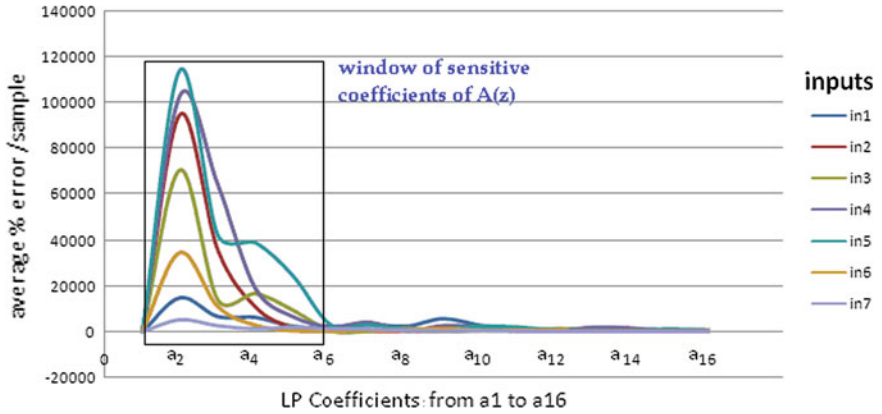


Fig. 3 Sensitive window analysis with respect to an average prediction error/sample due to consecutive flip of a sign bit of LPC for individual inputs

where i = 1, 2, 3,…,16 and n = 1,…,16.

In Fig. 4, x-axis presents the value of 'ni' and y-axis shows the average prediction error per sample caused by simultaneously changing all LPC of various inputs.

- In this case, only one LPC is corrupted at a time and the remaining LPCs are kept unchanged. LPCs are consecutively changed from 10 to 80 % of their original values to find out the average prediction error per sample as shown in Fig. 5. The prediction error increases if LPCs are corrupted. Based on the above study, we have calculated the average sensitivity of each LPC toward the increasing prediction error as shown in Fig. 6. The LPCs a_1, a_2, and a_3 are more sensitive toward corruption compared to remaining LPCs. The corrupted 'a_1' increases 52 % of average prediction error per sample compared to remaining LPCs. The a_2 and a_3 are the next most sensitive LPC toward corruption.

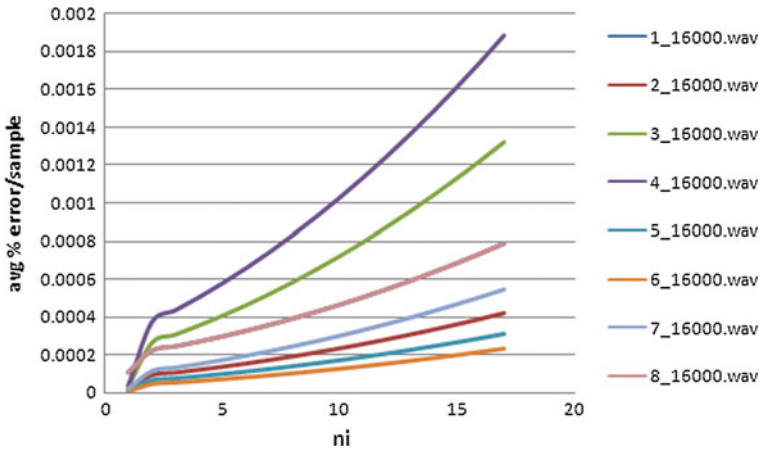


Fig. 4 Residual caused by simultaneously changing all LPC of inputs as per Eq. (5)

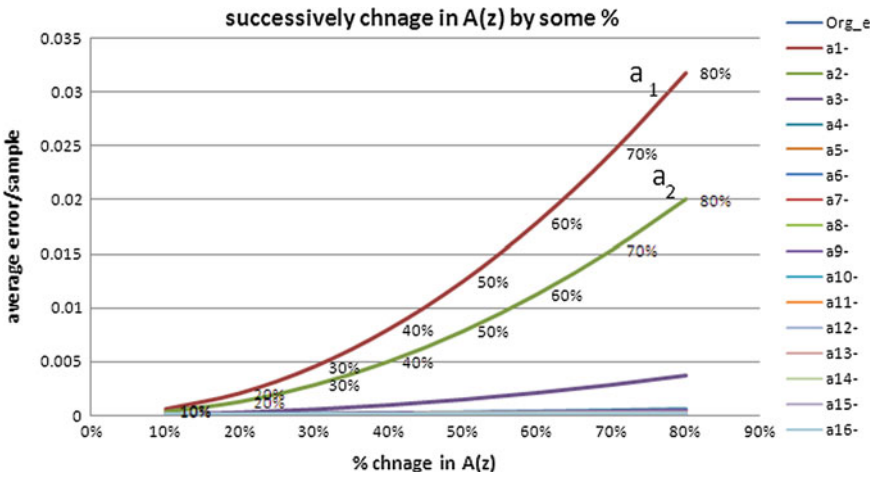


Fig. 5 Percentage change in A(z) successively versus average prediction error/sample

4 Pitch Sensitivity of LPC and Stability Improvement

4.1 Pitch of Input Frame and A(z)

The effect of pitch variation on each of the A(z) coefficients has been studied here. A numbers of voice speech segments have been experimented for effect on LP coefficients due to pitch variation. One of input has been shown in Fig. 7.

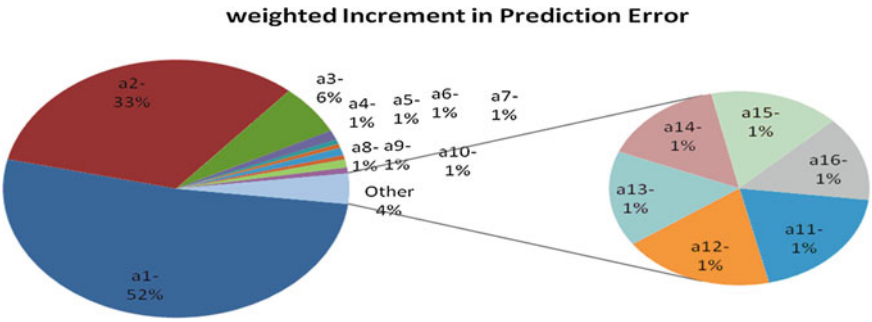


Fig. 6 Overall percentage prediction error weighted of a_1 to a_{16} as these increase from 10 to 80 % consecutively

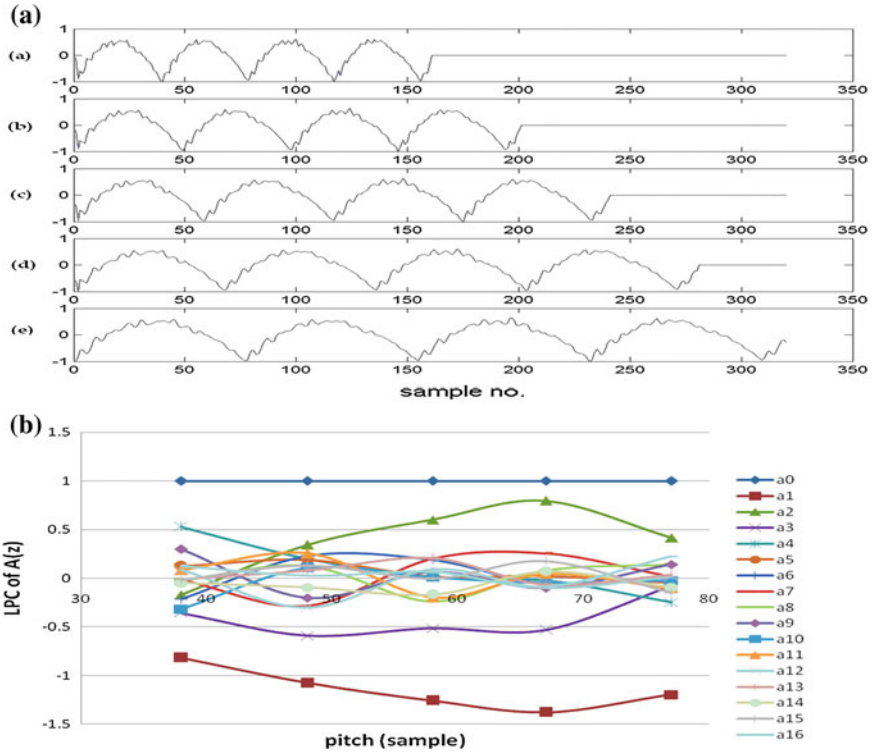


Fig. 7 Input frame with various **a** pitch length and **b** LPC distribution

It can be shown from Fig. 7b that coefficients a_1 and a_2 deviate less from their original value as compared to a_3 to a_{16}, when the pitch of the input frame changes as shown in Fig. 7a. The deviation is more in coefficient $a3$ to $a16$.

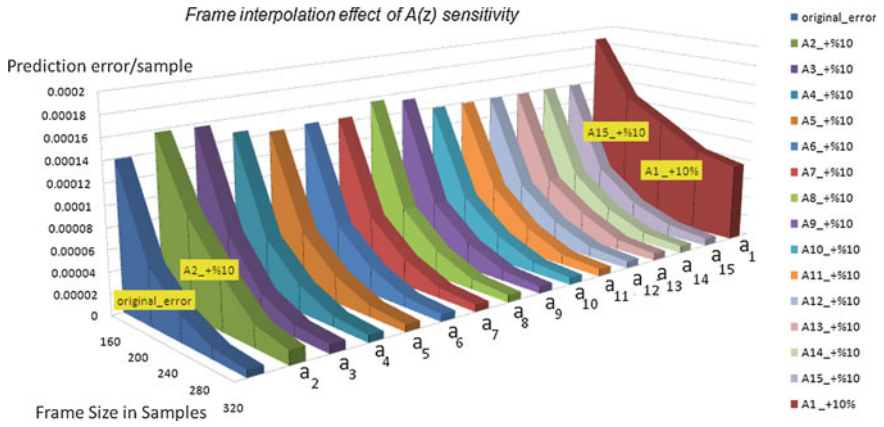


Fig. 8 Interpolation versus the sensitivity of A(z) for prediction of sample

4.2 Stability Improvement in LPC

- The 1st blue bar from left in Fig. 7 presents the prediction error versus interpolation or a pitch increment without any corruption in LPC. An input frame has been interpolated in such a way so that the frame lengths in sample become 240, 280, and 320 as shown in Fig 7a.
- Now, if the a_2 of LPC is increased by 10 % for each interpolation frames, then the resultant average prediction error per sample is shown by 2nd green bar in Fig. 8.
- The LPC from interpolated frame provides the reduced sensitivity to corruption and enhanced prediction capability as compared to original frame size.
- If such an interpolated LPC is used for LSF conversation, then it provides the stable synthesize filter.

5 Conclusion

The sign bit corruption in a_1 to a_5/a_6 causes 1,000 times more prediction error compared to remaining LPCs. So, while quantization, a_1, a_2, and a_3 LPCs require more number of bits as compared to remaining ones. If we find out how the each LPC is related to LSP, then we can improve the accuracy of LPF while transmitting to channel. From the sensitivity window and weight analysis, we can allocate the different quantization levels for each LSF. From Sect. 4, it can be concluded that 'as the interpolation/pitch increases the frame size, then corrupted LP coefficients produce less prediction error/sample' as compared to without interpolation. This means that if the same percentage corruptions happen during transmission with both the LPCs, then LPC calculated with interpolated data

provides more stable A(z) (less sensitivity or less prediction error) as compared to original LPC (no interpolation) calculated from original data frame length. Basically, analysis presented in this chapter provides stable A(z) filter LP coefficients with reduced sensitivity to corruption during transmission.

References

1. Wideband Speech Coding Standards and Applications by VoiceAge, 2004–2008
2. Laflamme C, Adoul J-P, Salami R, Morissette S, Mabilleau P (1991) 16 KBPS wideband speech coding technique based on ALGEBRAIC CELP. IEEE Acoust Speech Signal Process 1:13–16
3. Adaptive Multi-Rate Wideband speech codec (AMR-WB/G.722.2) by VoiceAge Corporation
4. Atal BS (2006) The history of linear prediction. IEEE Signal Process Mag 23(2):154–161
5. Flanagan JL, Coker CH, Rabiner LR, Schafer RW, Umeda N (1970) Synthetic voices for computers. IEEE Spectr 7(10):22–45
6. Viswanathan V (1982) Design of a robust baseband LPC coder for speech transmission over 9.6 Kbit/s noisy channels. IEEE Trans Commun 30:663–673
7. Chu WC (2003) Mobile Media Laboratory. DoCoMo USA Labs. San Jose, California
8. Yu R, Lin X, Ko CC (2002) A multi-stage Levinson-Durbin algorithm. IEEE 1:218–221
9. Atal BS, Hanauer S (1971) Speech analysis and synthesis by linear prediction of the speech wave. J Acoust Soc Am 50:637–655
10. Grassi S, Dufaux A, Ansorge M, Pellandini F (1997) Efficient algorithm to compute LSP parameters from 10th –order LPC coefficients. IEEE Int Conf 3:1707–1710
11. Kabal P, Ramachandran RP (1998) The computation of line spectral frequencies using Chebyshev polynomials. IEEE Trans Acoust Speech Signal Process 34(6):1419–1426
12. Semenov V (2007) A novel approach to calculation of line spectral frequencies based on inter-frame ordering property. IEEE J Comput 2(7):1072–1075
13. Kibey Sameer A, Kulkarni Jaydeep P, Sarode Piyush D (2004) A fast LSF search algorithm based on interframe correlation in G.723.1. EURASIP J Hindawi Publishing Corporation Appl Signal Process 8:1107–1112

Performance Analysis of AODV and AOMDV Routing Protocols on Scalability for MANETs

B. Rekha and D. V. Ashoka

Abstract In the world of computer networks, routing protocols of mobile ad hoc networks (MANETs) draw an increasing attention. One of the challenging components of routing in MANETs is accomplishing scalability which defines the performance ability of a routing protocol when one or more network parameters grow in value. In this chapter, two reactive routing protocols—Ad hoc On-demand Distance Vector (AODV) and Ad hoc On-demand Multipath Distance Vector (AOMDV) are chosen to show their performance on scalability by changing number of nodes. Also their comparative analysis is performed through simulation using ns-2 on different metrics like end-to-end delay, throughput, packet delivery fraction (PDF). Results show that AOMDV performs better to AODV when the number of nodes is increased.

Keywords MANET · AODV · AOMDV · Scalability · Performance evaluation · Delay · Throughput · PDF

1 Introduction

MANET is characterized by self-configuring, decentralized, high dynamic topology and easily broken without any infrastructural components. Each node acts both as a host and as a router. MANET is represented by distributed system with wireless mobile nodes which move freely and self-organize forming ad hoc

B. Rekha (✉)
Department of Information Science and Engineering, SJB Institute of Technology, Karnataka, Bangalore, India
e-mail: rekha.sjbit@gmail.com

D. V. Ashoka
Department of Information Science and Engineering, JSS Academy of Technical Education, Karnataka, Bangalore, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_19,

network topologies without pre-existing infrastructure. So each node acts both as a host and as a router. Traditionally, we find MANET applications in tactical networks, but now technologies like IEEE 802.11, Hyperlan, and Bluetooth enable commercial deployments apart from military domain. Today, we see a revolutionary change in our information society by the use of wireless communication devices such as laptops, cell phones, personal digital assistants (PDAs). Due to mobility of nodes, routing is one of the key challenges faced by researchers. In this chapter, an attempt has been made to analyze protocols performance on scalability by varying number of nodes as large-scale ad-hoc networks find applications in consumer owned networks, tactical military networks, natural disaster recovery services and vehicular networks [1].

2 Overview of MANET Routing Protocols

Based on how routing information is acquired and maintained by mobile nodes, existing MANET routing protocols are classified into three categories: proactive (table driven), reactive (on demand), and hybrid.

2.1 Proactive Routing Protocols

Proactive protocols create the routes (shortest path) periodically upon changes in topology and maintain it in their routing tables. Hence, route to a particular node is available at any moment. They maintain fresh lists of destinations and their routes by periodically distributing routing tables throughout the network. But drawbacks increased control overhead upon frequent changes in topology due to mobility of nodes, respective amount of data for maintenance, and slow reaction on restructuring and failures. Few examples in this category are Wireless Routing Protocol (WRP), Destination-Sequenced Distance Vector (DSDV) Routing Protocol, Fisheye State Routing (FSR) Protocol, and Optimized Link-State Routing (OLSR) Protocol.

2.2 Reactive Routing Protocols

Reactive protocols create the routes by discovery mechanism on requirement basis. These protocols find a route on demand by flooding the network with route request packets. After finding the route to any destination node, it is maintained until it is no longer required or destination node is unavailable. The distinct advantage is less control overhead and thus better scalability compared to proactive ones. However, source nodes experience delay in route discovery process before they send data packets, that is, high latency time in route finding, and

excessive flooding can lead to network clogging. Examples include Dynamic Source Routing (DSR), AODV, AOMDV, Temporally Ordered Routing Algorithm (TORA), Dynamic MANET On-demand Routing Protocol (DYMO) are few examples of on-demand routing protocols.

2.3 Hybrid Routing Protocols

Hybrid protocols combine the advantages of proactive and of reactive routing. The routing is initially established with some proactively prospected routes and then serves the demand from additionally activated nodes through reactive flooding. Examples of hybrid algorithms include Zone Routing Protocol (ZRP), Hybrid Wireless Mesh Protocol (HWMP), Hybrid Routing Protocol for Large Scale (HRPLS) Mobile Ad Hoc Networks with Mobile backbones, etc.

3 Related Work

AODV and AOMDV in [2] and [3] are compared and found that AOMDV incurs more routing overhead and delay than AODV, but it has a better efficiency in packet drop and delivery. However, maximum connections considered are 50. Comparative analysis [4] is also given for AODV and AOMDV. Authors have chosen mean link failure rate as a function of mean node speed. They have proved that AOMDV can cope with mobility-induced route failures, reducing packet loss by 40 percent with improved delay and reduced overhead. Authors in [3, 5–8] have chosen various pause times to analyze the performance of routing protocols. However, the number of nodes is varied in the range 50–100. Comparison of these two protocols in heterogeneous hybrid cluster routing is done in [7] to show the increased performance in terms of throughput. They have concluded that AOMDV gives a better performance. However, they have chosen 50 nodes for simulation.

Since nodes keep moving with different speed in MANETs, main challenge is to route packets with low overhead and achieve high throughput and low packet loss. In this chapter, an attempt has been made to check the performance of selected two protocols in terms of scalability of the MANET along with performance metrics such as packet delivery fraction, throughput, and average end-to-end delay.

4 Overview of AODV

AODV is a single-path, reactive, loop-free distance vector routing protocol. It is capable of both unicast and multicast routing which uses route discovery approach of DSR and DSDV. A source node to find the destination initiates route discovery by flooding route request (RREQ) packets and waits for route reply (RREP) packet

from destination node. Upon receiving the first copy of RREQ packet, intermediate nodes set up a reverse path to the source node by using previous hop of the RREQ as the next hop of the reverse path. If an intermediate node knows the path to the destination node, it unicasts RREP to the source node along the reverse path. Thus, AODV uses hop-by-hop routing approach. Sequence numbers are used to ensure loop freedom in AODV. Each node will have a monotonically increasing number for itself in its routing table. Whenever a node sends out any message, it increases its own sequence number [9]. The highest sequence number is maintained for each destination by each node which signifies a fresher route [9]. It determines the freshness of routing information. The node with higher sequence number has more recent information. Route error (RERR) packets are used by MANET nodes to maintain the routes. If any intermediate node detects a failure over any of its links, it invalidates all destinations which are unreachable. Then, it generates RERR packets which are propagated toward the traffic sources having a route through failed link. The source after receiving RERR initiates a new route discovery process. Figure 1 outlines the routing of AODV.

4.1 Advantages and Disadvantages

The main advantage of AODV protocol is the routes that are established on demand, and destination sequence numbers are applied to find the latest route to the destination. The connection setup delay is lower. Disadvantage of AODV protocol is that intermediate nodes can lead to inconsistent routes if the source sequence number is very old and the intermediate nodes have a higher but not the latest destination sequence number, thereby having stale entries. Also, multiple route reply packets in response to a single route request packet can lead to heavy control overhead. Another disadvantage of AODV is unnecessary bandwidth consumption due to periodic beaconing.

5 Overview of AOMDV

AOMDV [4] a reactive routing protocol, is an extension to AODV protocol. It computes multiple loop-free and disjoint paths. The routing entries for each

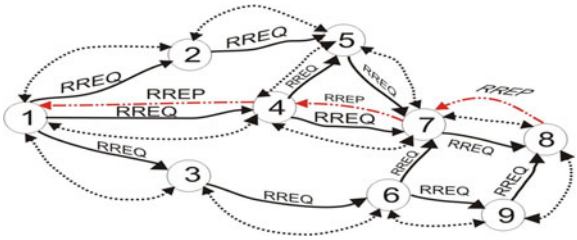


Fig. 1 Route request packets flooding from node 1 to node 8 and path establishment through route reply packets

destination contain a list of the next hops along with the corresponding hop counts, where all the next hops have the same sequence number. This helps to keep track of a route [4]. For each destination, a node maintains the advertised hop count which is the maximum hop count for all the paths. It is used to send route advertisements of the destination. The duplicate advertisement defines an alternate path to the destination. Loop freedom is ensured for a node by accepting alternate paths to destination if it has a less hop count than the advertised hop count for that destination. The advertised hop count is not changed for the same sequence number as the maximum hop count is used. When a route advertisement is received for a destination with a greater sequence number, the next hop list and the advertised hop count are reinitialized [4]. AOMDV can be used to find node-disjoint or link-disjoint routes. To find node-disjoint routes, each node does not immediately reject duplicate RREQs. Each RREQ arriving via a different neighbor of the source defines a node-disjoint path. This is because nodes cannot broadcast duplicate RREQs. So any two RREQs arriving at an intermediate node via a different neighbor of the source could not have traversed the same node. In an attempt to get multiple link-disjoint routes, the destination only replies to RREQs arriving via unique neighbors. After the first hop, the RREPs follow the reverse paths, which are node disjoint and thus link disjoint. The advantage of using AOMDV is that it ensures loop-free multiple and disjoint paths [4].

6 Simulation Environment

The simulations have been performed using network simulator ns-2 [10]. The network simulator ns-2 is discrete-event simulation software. A network of nodes placed within a 11,000 × 11,000 m area is considered. The performance is evaluated by keeping the network speed and pause time constant and varying the network size (number of mobile nodes). Figure 2 shows the network with 1,000 mobile nodes in the NAM console which is a built-in program in ns-2-allinone package.

Table 1 shows the simulation parameters used in this evaluation.

Performance Metrics

While analyzing, packet delivery fraction (PDF), average end-to-end delay, and throughput are the performance metrics which are considered in simulation.

Packet Delivery Fraction

It is a ratio of the number of packets received by the destination to the number of packets send by the source, and this illustrates the level of delivered data to the destination. The greater value of packet delivery fraction means better performance of the protocol.

PDF = Σ No. of packet received/Σ No. of packet sent

Average end-to-end delay

It is defined as average time taken by data packets to propagate from source to destination across the network. This includes all possible delays caused by

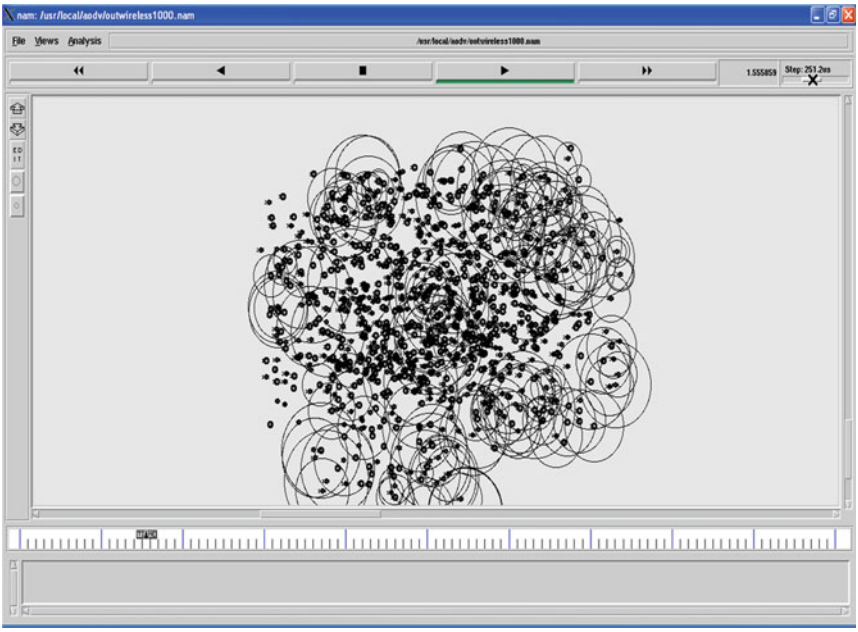


Fig. 2 MANET with 1,000 nodes

Table 1 Parameter values for simulation

Simulator	Ns-2.35
Protocols	AODV and AOMDV
Simulation duration	20 s
Simulation area	11,000 × 11,000 m
Number of nodes	500, 600, 700, 800, 900, 1,000
Movement model	Random waypoint
MAC layer protocol	IEEE 802.11
Link type	Duplex link
Queue size	50
Pause time	5 s
Packet size	1,500 bytes/packet
Application type	FTP
Agent type	TCP

buffering during routing discovery latency, queuing at the interface queue, and retransmission delays at the MAC, propagation, and transfer times; the lower value of end-to-end delay means the better performance of the protocol.

End-to-end delay = Σ (arrive time—send time)

Throughput

Throughput is the average number of messages successfully delivered per unit time, that is, average number of bits delivered per second.

Throughput = Σ Total number of received packets at destination/time taken

7 Simulation Results and Analysis

From Fig. 3, it is noted that AOMDV incurs more delay than AODV due to the fact that on link failure in AOMDV, it tries to find the alternate path from backup route which results in additional delay.

Throughput is better in AOMDV than in AODV as it is shown in Fig. 4, except at one point which is because AODV is single-path protocol. On link failure in , packets are not delivered to the destination. Since AOMDV is multipath routing protocol, it finds an alternate path and delivers the packets.

AOMDV has better PDF than *AODV* on increasing number of nodes. This is because AOMDV finds different path on link break which is seen in Fig. 5.

8 Conclusion

AODV and AOMDV routing protocols are evaluated by increasing number of nodes in the range 500–1,000 using ns-2. Comparison is based on end-to-end delay, throughput, and packet delivery fraction. By comparative study of simulation, it is found that AOMDV is preferred to AODV when throughput and PDF are concerned as it has got better throughput and PDF. But AOMDV incurs more delay when compared to AODV. Hence, when delay is concerned, AODV is

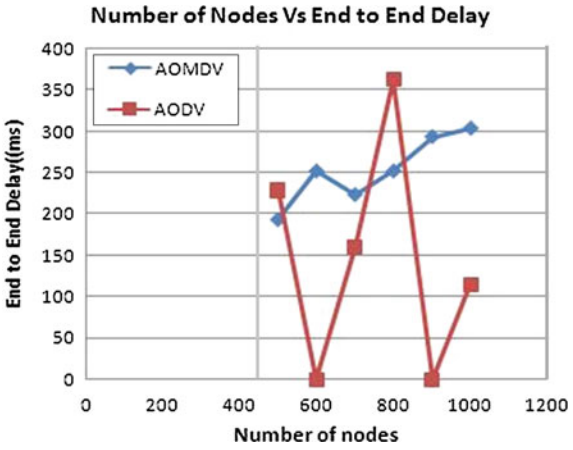


Fig. 3 End-to-end delay

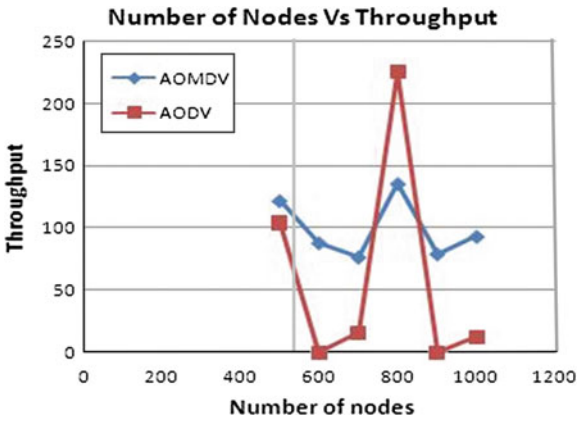


Fig. 4 Throughput

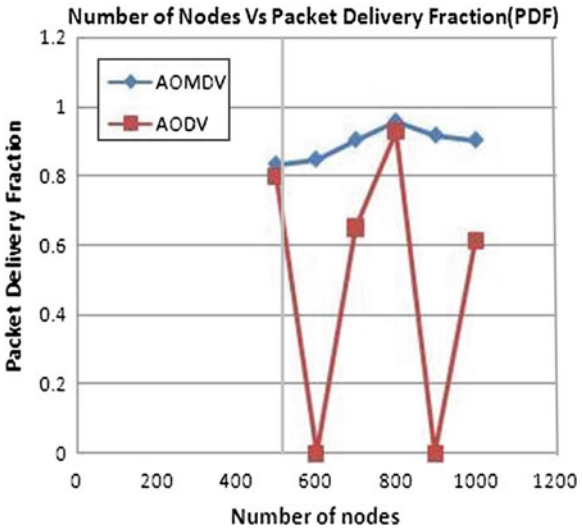


Fig. 5 Packet delivery fraction

preferred, but still in the above all cases, it can also be observed that values of AODV are inconsistent.

References

1. Eriksson J (2005) Routing scalability in MANETs. University of California, Riverside
2. Biradar SR et al (2010) Performance evaluation and comparison of AODV and AOMDV. Int J Comput Sci Eng 02(02):373–377
3. Jacob J, Seetalakshmi V (2012) Performance analysis and enhancement of routing protocol in MANET. Int J Mod Eng Res (IJMER) 2(2):323–328

4. Marina MK, Das SR (2006) Ad-hoc on-demand multipath distance vector routing. Wireless Commun Mob Comput 6:969–988
5. Singh EP (2012) Brief description of routing protocols in MANETs and performance analysis (AODV, AOMDV, TORA). Int J Adv Res Comput Sci Softw Eng 2(1)
6. Kumawat R (2011) Comparative study of on-demand routing protocols for MANETs. Int J Comput Appl 27(10)
7. Singh M (2012) Simulation and comparison of AODV, DSR and AOMDV routing protocols in MANETs. Int J Soft Comput Eng (IJSCE) 2(3)
8. Sridevi B (2011) Performance comparison using AODV and AOMDV protocols in heterogeneous hybrid cluster routing using partial authority nodes in MANETs. Eur J Sci Res 58(4):542–549
9. Klein-Berndt L A quick guide to AODV routing. Wireless Commun Technol Group National Inst Stand Technol
10. Network simulator NS-2. http://www.isi.edu/nsnam/ns/

Feature-Level Image Fusion Using DWT, SWT, and DT-CWT

G. Siddalingesh, A. Mallikarjun, Harihar Sanjeevkumar and S. Kotresh

Abstract Image fusion is the process of combining information from two or more sensed or acquired images into a single composite image that is more informative and becomes more suitable for visual processing or computer processing. Image fusion fully utilizes much complementary and redundant information of the original images. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite image that contains a better description of the scene than any of the individual source images. Feature-level image fusion (FLIF) algorithms (both in spatial and in frequency domain) were developed and evaluated using fusion quality evaluation metrics. The images to be fused are passed through joint segmentation algorithm to get the common segmentation map. Salient feature, viz. standard deviation, is computed for corresponding segments (both the images), and the segment was chosen based on best salient feature. It was done for all the segments. Five different image sets were used to evaluate the proposed fusion algorithm. To compare the performance of this algorithm, three different pixel-level image fusion algorithms, viz. DWT, SWT, and DT-CWT, were also implemented and evaluated. From this study, it is concluded that FLIF provides a good fused image at the cost of execution time and also it requires a good segmentation map. Most of the time DT-CWT provides good fusion results since it considers the edge information in six directions. In all

G. Siddalingesh (✉) · A. Mallikarjun
BITM, Bellary, India
e-mail: siddug123@gmail.com

A. Mallikarjun
e-mail: mtechmalli@gmail.com

H. Sanjeevkumar
CIT, Gubbi, Tumkur, India
e-mail: sanjeevkumar.harihar@gmail.com

S. Kotresh
RYMEC, Bellary, India
e-mail: kotresh2000@rediffmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_20,

cases, the DWT-based pixel-level image fusion algorithm does not provide good results since it does not consider the edge information and lack of shift invariant. SWT-based image fusion algorithm provides good results in some cases where there are no much edges in the images to be fused, it is shift invariant, and it does not consider the directional edge information.

Keywords Image segmentation · Image fusion · Entropy · DT-CWT · FLIF · Pixel

1 Introduction

Image fusion is the process of combining information from two or more sensed or acquired images into a single composite image that is more informative and becomes more suitable for visual processing or computer processing. Image fusion fully utilizes much complementary and redundant information of the original images. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite image that contains a better description of the scene than any of the individual source images. The objective is to reduce uncertainty, minimize redundancy in the output, and maximize relevant information pertaining to an application or a task. In image fusion, the image data appear in the form of array of number, which represents brightness (intensity), color, temperature, distance, and other scene properties [1].

One level higher than pixel-level image fusion is Feature-level image fusion. One technique of achieving this is with a region-based fusion scheme. Initially, an image is segmented to produce a set of regions. Various region properties can be calculated [2]. The properties can be used to determine which features from which images are used in the fused image. Feature-level image fusion has some advantages over pixel-level image fusion as more intelligent semantic fusion rules can be considered based on actual feature in the image rather than on single pixel. Feature is very important than the actual pixel. Hence, it is better to incorporate the feature information in the process of fusion [3].

Segmentation algorithm plays a vital role in region-based image fusion process. Features should be segmented as single regions. Feature may split into more than one region, and each region has to be treated separately. If possible, less number of regions should be generated to reduce the computational burden. Segmentation can be done either separately (uni-model segmentation) or jointly (joint segmentation). Uni-model segmentation method may create many regions than joint segmentation since different images have different features [4]. In this paper, joint segmentation using dual-tree complex wavelet transform (DT-CWT) and several image fusion methods are studied. The images used in image fusion should already be registered. Misregistration is a major source of error in image fusion.

2 Literature Survey

With the recent rapid developments in the field of sensing technologies, multisensor systems have become a reality in a growing number of fields such as remote sensing, medical imaging, machine vision, and the military applications for which they were first developed. The result of the use of these techniques is a great increase in the amount of data available. Image fusion provides an effective way of reducing this increasing volume of information while at the same time extracting all the useful information from the source images. Multisensor data often presents complementary information about the region surveyed, so image fusion provides an effective method to enable comparison and analysis of such data. The aim of image fusion, apart from reducing the amount of data, is to create new images that are more suitable for the purposes of human/machine perception and for further image-processing tasks such as segmentation, object detection, or target recognition in applications such as remote sensing and medical imaging [1]. For example, visible-band and infrared images may be fused to aid pilots landing aircraft in poor visibility.

The image fusion can be categorized as follows:

- Low-level fusion (pixel-level fusion).
- Mid-level fusion (Feature-level fusion).
- High-level fusion (symbol-level fusion).

The past decade gave rise to a considerable number of different approaches to fuse visual image information. Multisensor data fusion with MATLAB [1] in this book give complete information about image fusion methods explains about pixel- and Feature-level image fusion methods, it also explains about various performance metrics by using these metrics we are comparing various fusion methods. Image fusion can be performed using different methods. One such publication is region-based image fusion using complex wavelets [3]. This paper focuses on region-based image fusion using complex wavelet in this characteristic of each region is calculated, and region-based approach is used to fuse the images. In Feature-level image fusion, first we need to segment the images; one such paper deals with the segmentation is [4, 5], and these paper deals with the segmentation of images into disjoint regions, in a manner consistent with human perception of the content, segmentation can be performed by using dual-treecomplex wavelet transform and watershed transform.

Image fusion can be performed in wavelet domain, and the advantage of wavelet transform is that it can be analyze signal in time domain and frequency domain, respectively. If we fuse a high-resolution panchromatic image and a multispectral image by DWT, the fused image conserves more spectral characteristic of the multispectral image. DWT has two main disadvantages:

- Lack of shift invariance—this means that small shift in the input signal can cause major variations in the distribution of energy between DWT coefficients at different scale.

- Poor directional selectivity.

After the fusion, we need to calculate the performance metrics, one paper deals with it [6], by using these performance metrics, we are calculating various metrics and comparing the fusion methods. Image fusion using complex wavelet [7] in this paper deals with the fusion of images by using dual-tree complex wavelet transform and discrete wavelet transform and compares the DWT versus DT-CWT; from this paper, we can notice that dual-tree complex wavelet transform provides better result than the discrete wavelet transform.

3 Segmentation of Image

One level higher than the pixel-level image fusion is Feature-level image fusion in which image is segmented to regions and the corresponding regions are fused. Feature is more important than the single pixel. Hence, it is better to incorporate the feature information in the fusion process. Segmentation algorithm plays a vital role in region-based image fusion process. Features should be segmented as single regions. Feature may split into more than one region, and each region has to be treated separately. If possible, less number of regions should be generated to reduce the computational burden [4].

Segmentation can be performed by applying DT-CWT and watershed transforms [8]. The segmented images are used in the Feature-level image fusion process. The DT-CWT provides the six subbands oriented at ± 15, ± 45, ± 75. Complex wavelets are shift invariant and retain the properties of scale and orientation sensitivity. Image segmentation is the process of separating out mutually exclusive homogeneous region of interest [9]. DT-CWT is used to detect the texture boundaries.

3.1 Texture Representation

Generally, Gabor filter had been used to for texture representation because of facts from psychophysical experiments. Human visual system decomposes the visual field into perceptual channels. These channels are evenly spaced in angle. Gabor filter representation is computationally burdensome. Complex wavelets are alternative to Gabor functions for texture analysis. Complex wavelets are shift invariant and retain the properties of scale and orientation sensitivity.

The detail coefficients of dual-tree complex wavelet transform is used for texture process. Denoting the detail coefficients at level i, orientation θ by $D_{i,\theta}(x, y)$ and retain the complex magnitude $\left|D_{i,\theta}(x, y)\right|$ for further analysis.

3.2 Texture Post processing

Simple gradient calculation of complex magnitude gives rise to a double edge in the gradient magnitude. Application of watershed algorithm produces a spurious narrow region along the boundaries. It can be avoided using median filter before gradient operator. Median filter is edge preserving smoothing filter that can suppress isolated noise without blurring sharp edges. Specifically, the median filter replaces a pixel by the median of all pixels in the neighborhood and it is computationally burdensome. The solution is separable median filter and it has to be chosen with care: the first filter removes the double edge effect of the subbands, and the second filter parallel to subband orientation removes noise of the sub-bands [10]. Considering both scale and orientation, the subband resulting from the filtering is

$$S_{i,\theta}(x,y) = MedFilt_i\left(MedFilt_{(\theta+0.5\pi)}\left(\left|D_{i,\theta}(x,y)\right|\right)\right) \tag{1}$$

The order of the median filter is chosen as $(7 + 2n)$, where n is the current level of the wavelet transform, and the constant term is equal to the size of wavelet filters.

3.3 Texture Gradient

Gaussian derivative function is used as gradient operator. The texture gradient magnitude of each sub-band is

$$\mathrm{TG}_{i,\theta}(x,y) = \sqrt{\left(S_{i,\theta}(x,y) * G'_x\right)^2 + \left(S_{i,\theta}(x,y) * G'_y\right)^2} \tag{2}$$

where G'_x and G'_y are the Gaussian partial derivative filters in x and y directions, and * denotes convolution.

Weighted sum of the magnitudes is

$$\mathrm{TG}(x,y) = \sum_{i,\theta} \mathrm{int}\ erp\left(w_{i,\theta} \times \mathrm{TGH}_{i,\theta}(x,y)\right) \tag{3}$$

where $\mathrm{TGH}_{i,\theta}(x,y) = \dfrac{\mathrm{TG}_{i,\theta}(x,y)}{\max\left(\mathrm{TG}_{i,\theta}(x,y)\right)}$

$$w_{i,\theta} = \frac{N_i}{\sum_{x,y} \mathrm{TGH}_{i,\theta}(x,y)^2}$$

N_i is number of pixels in sub-band at level i,
int erp () is interpolation function.

Up sampling is done using interpolation, and it is performed separable.

3.4 Gradient Combination

Texture and intensity gradients are combined to get final gradient capturing all perceptual edges in the image [11]. The combined gradient will be dominated by intensity gradient in smooth regions and texture gradient in textured regions. The activity measure is

$$\text{activity}(x,y) = \exp\left(R_{\text{half}}\left(\frac{E_{\text{tex}}(x,y)}{\alpha} - \beta\right)\right) \tag{4}$$

where $\alpha = 2$ and $\beta = 7$ are chosen based on intuition, $R_{\text{half}}(\varsigma)$ is the half wave rectification to suppress the negative exponents as:

$$R_{\text{half}}(\varsigma) = \begin{cases} 0: & \varsigma < 0 \\ \varsigma: & \varsigma \geq 0 \end{cases} \tag{5}$$

Texture energy E_{tex} is calculated on up sampled subbands. Texture features respond slightly larger area than the desired because of the involved spatial integration. Morphological erosion E is used to overcome the problem and strel used in this function is a square neighborhood of nine pixels. The texture energy is computed as

$$E_{\text{tex}} = \sum_{i,\theta} \text{int}\, erp\left(E\left(\frac{S_{i,\theta}(x,y)}{2^i}\right)\right) \tag{6}$$

The denominator 2^i is used to correct the DC gain of the wavelet filters. Finally, the weighted sum of texture and modulated intensity gradient is computed as

$$\text{GS}(x,y) = \frac{|\text{IG}(x,y)|}{\text{activity}(x,y) \times w_I} + \frac{\text{TG}(x,y)}{w_T} \tag{7}$$

where

w_I is to be four times the median intensity gradient,
w_T is the median value of the texture gradient.

4 Feature-level Image Fusion

One level higher than pixel-level image fusion is Feature-level image fusion. One technique of achieving this is with a region-based fusion scheme. Initially, an image is segmented to produce a set of regions. Various region properties can be calculated. The properties can be used to determine which features from which images are used in the fused image. Feature-level image fusion has some advantages over pixel-level image fusion as more intelligent semantic fusion rules can

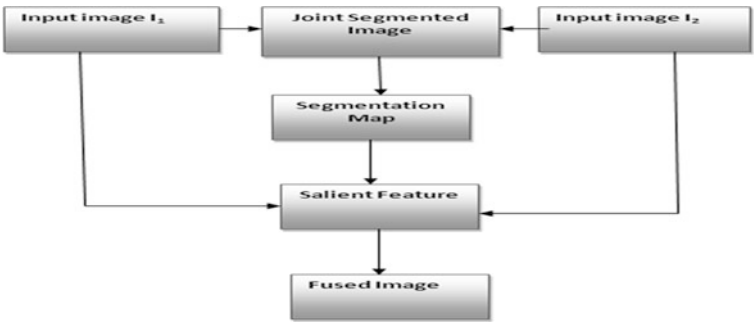


Fig. 1 Block diagram of Feature-level image fusion

be considered based on actual feature in the image rather than on single pixel. Feature is very important than the actual pixel. Hence, it is better to incorporate the feature information in the process of fusion. Segmentation algorithm plays a vital role in region-based image fusion process. Features should be segmented as single regions. Feature may split into more than one region, and each region has to be treated separately. If possible, less number of regions should be generated to reduce the computational burden [4, 5].

The block diagram of Feature-level image fusion is as shown in Fig. 1. The input images are joint segmented by using DT-CWT and watershed transform as discussed in Sect. 3 (Image segmentation). The joint segmented image is shown in Fig. 3.13 which is used as the segmentation map [12]. By using the segmentation map, we are calculating the salient feature like standard deviation; if the standard deviation of the segmented part of the input image I_1 is greater than the standard deviation of the segmented part of the input image I_2, then the fused image part comes from input image I_1, otherwise it is from input image I_2. Figure 2 shows the segmented part of visible and infrared images. The fused image without filter as

Fig. 2 Example of segmented part of visible and infrared images

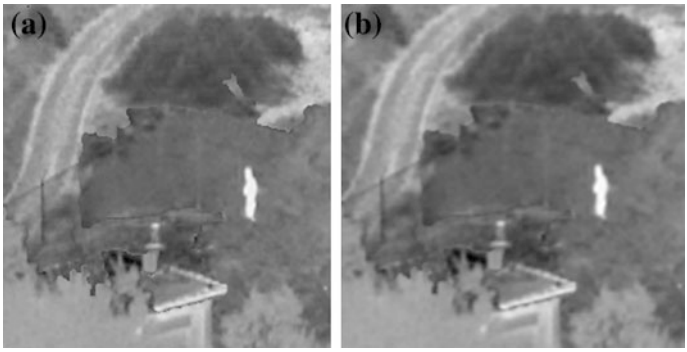


Fig. 3 Fused images: **a** without filter, **b** with filter

Fig. 4 Feature-level fused image (wavelet domain)

some ridge lines is shown in Fig. 3a which can be reduced by using the MATLAB function imfilter. The fused image with filter is shown in Fig. 4b.

The above fusion is based on the spatial domain, for the wavelet domain consider Fig. 2 shows the segmented part of visible and infrared image for that we are calculating the wavelet coefficients by using DT-CWT; if the standard deviation of segmented part of input image I_1 is greater than the standard deviation of segmented part of input image I_2, then the wavelet coefficients comes from input image I_1, otherwise wavelet coefficient from input image I_2. After that, the inverse dual-tree complex wavelet transform was performed to get the fused image shown in Fig. 4.

5 Result

5.1 Medical Images

Figure 5a shows the CT image (left side image) and MR image (right side image, Fig. 5b). CT image provides clear bones information, but it does not provide soft tissues information. The MR image provides soft tissue information, but it does not

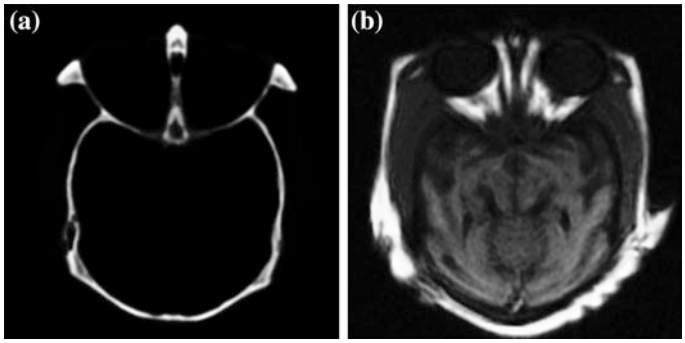


Fig. 5 Medical images: **a** medical A and, **b** medical B

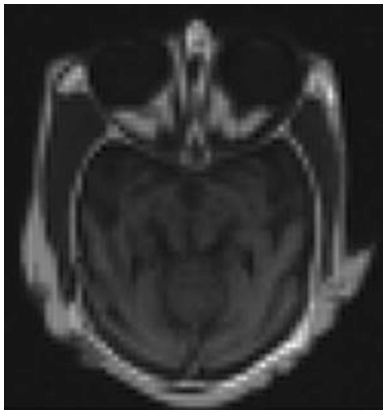


Fig. 6 DWT-based fused image

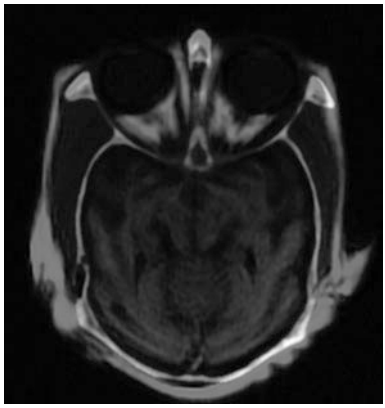


Fig. 7 SWT-based fused image

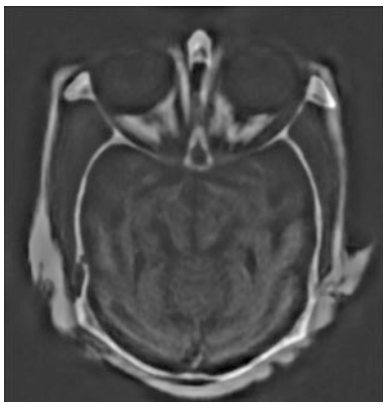


Fig. 8 DT-CWT-based fused image

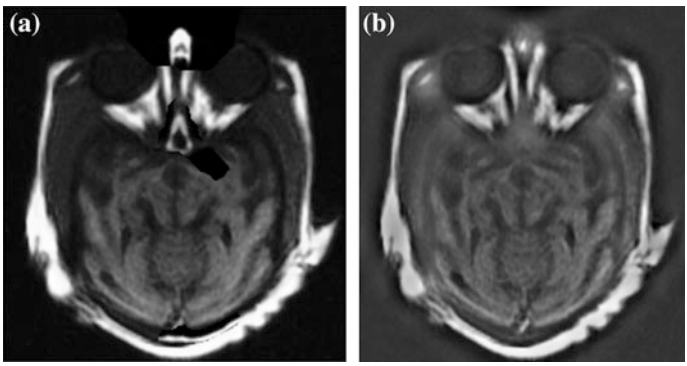


Fig. 9 Feature-level fused images: **a** spatial domain and, **b** wavelet domain

Table 1 Performance metrics for medical image

Methods	DWT		SWT		DT-CWT		Feature-level			
							Spatial domain		Wavelet domain	
Metrics	Without filter	With filter	Without filter	With filter	Without filter	With filter	Without filter	With filter	Without filter	With filter
E	6.006	6.094	6.022	5.994	6.01	6.022	5.805	**6.67**	6.5069	6.547
S	33.727	1488	35.247	1542	36.06	1562	59.09	**2606**	42.223	1845
SF	9.4348	8.368	11.83	6.623	13.95	11.01	16.48	18.14	13.788	**97.99**
CE	2.2835	2.857	1.720	2.816	1.96	2.830	0.981	**1.587**	2.2653	2.814
MI	2.162	2.0	2.486	2.07	2.359	2.058	**2.896**	2.03	2.1827	2.058
ET (sec)	0.43		0.68		0.39		29.3		26.0	

provide bones information. It shows that both CT and MR provide complementary information [13]. The fusion of these two images occurs, and the resultant fused image now contains both the bones information and tissues information, which cannot be seen in the individual CT or MR image as shown in Figs. 6, 7, 8, 9 DWT,SWT, DT-CWT and FLIF in spatial and wavelet respectively, gives better result compared to other fusion methods (Table 1).

6 Conclusion

Feature-level image fusion (FLIF) algorithms (both in spatial and in frequency domain) were developed and evaluated using fusion quality evaluation metrics. The images to be fused are passed through joint segmentation algorithm to get the common segmentation map. Salient feature, viz. standard deviation, is computed for corresponding segments (both the images), and the segment was chosen based on best salient feature. It was done for all the segments. Five different image sets were used to evaluate the proposed fusion algorithm. To compare the performance of this algorithm, three different pixel-level image fusion algorithms, viz. DWT, SWT, and DT-CWT, were also implemented and evaluated.

From this study, it is concluded that FLIF provides a good fused image at the cost of execution time and also it requires a good segmentation map. Most of the time DT-CWT provides good fusion results since it considers the edge information in six directions. In all cases, the DWT-based pixel-level image fusion algorithm does not provide good results since it does not consider the edge information and lack of shift invariant. SWT-based image fusion algorithm provides good results in some cases where there are no much edges in the images to be fused, it is shift invariant, and it does not consider the directional edge information.

Only standard deviation is used as a salient feature to select the best segment. More salient features can be used along with some fuzzy logic or neural networks to choose the best segment. It can also be extended for color image fusion. There is a scope to develop robust joint segmentation algorithm and hence better Feature-level image fusion.

References

1. Raol JR, Naidu VPS (2010) Multi-sensor data fusion with matlab. CRC press, ISBN 978-1-4398-0003-4
2. Naidu VPS (2010) Discrete cosine transform—based image fusion. Defense Sci J 60(1):48–54
3. Lewis JJ (2004) Region-based image fusion using complex wavelets. In: proceeding of the 7th international conference on information fusion, pp 555–562
4. Callaghan RJO', Bull DR (2005) Combined morphological-spectral unsupervised image segmentation. IEEE Trans Image Proc 14(1):49–62

5. Luo F (2010) Wavelet-based image registration and segmentation framework for the quantitative evaluation of hydrocephalus. Int J Biomed Imaging
6. Naidu VPS, Raol JR (2008) Pixel-level image fusion using wavelets and principal component analysis a comparative analysis. Defence Sci J 58(3):338–352
7. Hill P (2002) Image fusion using complex wavelets. Int Conf Inf Fusion 504–510
8. Naidu VPS (2010) Image fusion using the measure of focus, MSDF report No:1011/ATR07, 06th May 2010
9. Anwaar-uli-Haq M (2010) A novel color image fusion QoS measure for multi sensor night vision applications, IEEE paper 978-1-4244-7755-5
10. Zheng Y (2007) Effective image fusion rules of multi-scale image decomposition. In: proceedings of the 5th international symposium on image and signal processing and analysis, pp 362–366, Proc. ISPA 2007
11. Gonzalez RC, Woods RE (2007) Digital image processing, 3rd edn, ISBN 978-81-317-2695-2
12. Kingsbury NG (1999) Image processing with complex wavelets. Philos Trans R Soc London A Math Phys Sci 357(1760):2543–2560
13. Vekkot S, Shukla P (2009) A novel architecture for wavelet based image fusion. World Acad Sci Eng Technol 57

Aircraft Recognition System Using Image Analysis

K. Roopa and T. V. Ramamurthy

Abstract This paper describes the aircraft recognition system implemented using the moment invariant technique and the discrete cosine transform (DCT) feature method using MATLAB. The main objective of this paper is to provide a performance comparison between the two methods by using the static images from a database of images of aircraft and the captured image of the test aircraft. The DCT technique is found to yield a better overall recognition result for the set of database and test images considered. The recognition performance can be improved by including enough images in different viewpoints in the database. The features used here can also be used conveniently in real-time image recognition.

Keywords Aircraft recognition system · Moment invariant technique · DCT · MATLAB

1 Introduction

Moment invariants were derived by Hu [1] and were used by others [2–5]. Dudani et al. [2] described an automatic recognition of aircraft types from optical images by extracting the invariant moments from binary television images which are then used for automatic classification. The test involves 132 images of six aircraft types.

K. Roopa (✉)
Department of Telecommunication, Sir M. Visvesvaraya Institute of Technology, Bangalore 562157, India
e-mail: roopakodnad@yahoo.co.in

T. V. Ramamurthy
National Aerospace Laboratories, Department of Electronics and Communication Engineering, Reva Institute of Technology & Management, Bangalore 560064, India
e-mail: drtvramamurthy@revainstitution.org

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_21,

In their paper, two distinct decision rules were used in classification, viz., a Bayes decision rule and a distance-weighted k-nearest-neighbour rule.
Chen [3] described improved moment invariants for shape discrimination computed using only the shape boundary to reduce the computations. Mercimek et al. [4] described the recognition performance of classifiers in conjunction with invariant moment-based feature sets. Ramteke [6] described an invariant moment-based feature extraction for handwritten Devanagari vowels recognition.

Some researchers have used an affine invariant feature extraction for object recognition [5, 7]. Affine is related to a transformation of coordinates, resulting in translation, rotation and scaling. Flusser et al. [5] described a feature-based recognition of affine-deformed objects and patterns by using the affine moment invariants as features. Yu et al. [7] described a fully affine invariant image comparison method to identify features that have undergone very large affine distortions.

Several techniques were used by different researchers for aircraft identification in particular [8–10]. Hsieh et al. [11] described a symmetry-based algorithm to estimate an aircraft's optimal orientation for rotation correction. Then, distinguishable features are derived from each aircraft for recognition. Then, a hierarchical recognition scheme is used to recognize the types of aircraft. Molina et al. [8] described an identification of aircraft in the airport movement area. Video cameras were used for image capture of the tail. Camera should be deployed near the runways. When aircraft passes in front of the camera, an image of its tail is captured. Then, the aircraft tail number is recognized through an optical character recognition (OCR) algorithm. Then, the aircraft identification algorithm uses the OCR results and a database containing the tail numbers of all aircraft in the airport to identify the aircraft among the potential candidates. Rihaczek et al. [12] described an identification system of large aircraft by using features related to length and wing span.

Some researchers have used neural networks for classification [13–15]. An aircraft identification system using back propagation neural network is presented by Somaie et al. [9]. To avoid the problem of rotation of the aircraft in 2D plane, they have used the Hotelling transform to align aircraft pattern with its principal axis. Saghafi et al. [10] described aircraft type recognition from its optical or infrared images by using an area-based feature extraction method combined with a multilayer perceptron neural network for classification. When a large number of training data are available, and when proper classification approach is employed, recognition of images can be done with simple feature vectors. This reduces the computation and also maintains the performance of aircraft recognition within an acceptable limit. Kim et al. [16] used multilayer neural networks for identification and orientation estimation of aircraft. The translation and rotation invariant feature space is formed by determining the centroid and the perimeter of the aircraft.

The proposed problem in general is similar to the content-based image retrieval (CBIR). It refers to the problem of searching for a specific digital image in large image databases. Examples of features used in such techniques are colour, texture, shape and spatial location. The colour is a widely used visual feature in colour

image retrieval. A colour histogram can be used for representing colour feature. Texture is an inherent property of surfaces which contains information about the structural arrangement of surfaces. The ability to match using texture similarity can often be useful in distinguishing between areas of images with similar colour. The shape representations can be boundary based or region based. Examples for shape features are aspect ratio and consecutive boundary segments. Spatial location feature involves retrieving pictorial data by its position within an image.

Jayaprabha et al. [13] in their paper analysed challenges of CBIR techniques and described colour-based feature extraction. Datta et al. [14] described CBIR approaches and their trends. Vadivel et al. [15] described the characteristics of weighted feature vector in CBIR applications. Rui et al. [17] provided a comprehensive survey in the area of CBIR. Eakins et al. [18] described various CBIR methods. Jain et al. [19] provided a survey on CBIR systems using clustering techniques for large datasets. Srinivasa Rao et al. [20] described a CBIR system using Exact Legendre Moments and Support Vector Machine classifier.

2 Invariant Moments [21]

The invariant moments are not affected by translation, scale changes, mirroring and rotation. The 2D moment, m_{pq}, of order $(p + q)$ of a digital image $f(x, y)$ of size $M \times N$ is defined as

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$

The corresponding central moment of order $(p + q)$, that is, μ_{pq} is defined as

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

where

$p = 0, 1, 2,\ldots$, $q = 0,1,2,\ldots$, are integers and $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$

The normalized central moments, η_{pq}, are given by

$\eta_{pq} = \frac{\mu_{pq}}{\mu'_{00}}$ where $\gamma = \frac{p+q}{2} + 1$ for $p + q = 2, 3,\ldots$.

A set of seven invariant moments can be derived from the second and third moments. The first invariant moment is given by $\phi_1 = \eta_{20} + \eta_{02}$.

In this work, the first invariant moment is used as a feature for each database image and stored. The first moment is also calculated for the test image, and then, feature matching is done to recognize the test image.

3 The 2D Discrete Cosine Transform [22, 23]

The two-dimensional DCT for an M-by-N input image A is:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\Pi(2m + 1)p}{2M} \cos \frac{\Pi(2n + 1)q}{2N}, \quad \begin{matrix} 0 \le p \le M - 1 \\ 0 \le q \le N - 1 \end{matrix}$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2}/\sqrt{M}, & 1 \le p \le M - 1 \end{cases} \qquad \alpha_q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2}/\sqrt{N}, & 1 \le q \le N - 1 \end{cases}$$

The values B_{pq} are called the DCT coefficients of A. The DCT tends to concentrate information, making it also useful for image compression applications.

In this work, the first value in the coefficient matrix is used as the feature of an image. This value is extracted for all the database images and stored. The same feature is extracted for the test image as well, and then, feature matching is done to recognize the test image.

4 Block Diagram of the Aircraft Recognition System Using Image Analysis

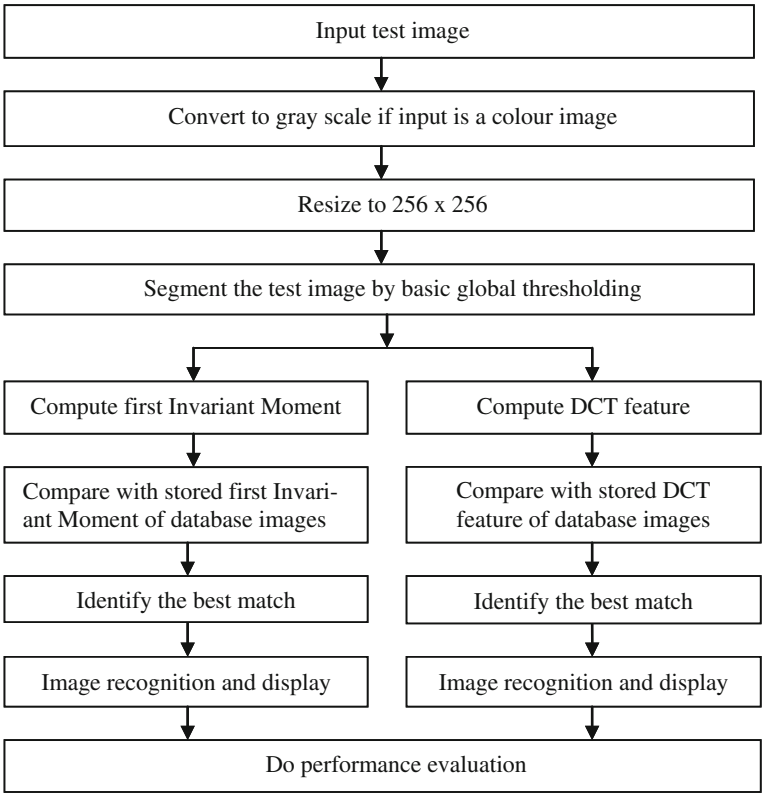


Fig. 1 Chengdu_J10

Fig. 2 Chengdu_J20

Fig. 3 Ef_Typhoon

5 Implementation

In the proposed system, nine classes of fighter aircraft images are maintained in the database: viz., Chengdu_J10 [24], Chengdu_J20 [25], EF_Typhoon, JF-17, LCA_Tejas, Mig35 [24], FA18, French_Rafale, Sukhoi_Su30 [AeroIndia 2011, Bangalore].

The images of aircraft which include translation and scaling were captured during the manoeuvres with different orientation of pitch, yaw and roll axes, in the air show, 'AeroIndia 2011', Bangalore. Thirteen such images for simulation purpose form a part of the database. A total of 23 such test images are used. Noiseless images were used here for analysis. It is assumed that the test images belong to the same category as those used in the database. The database images are given in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19. The experimental results are tabulated in Table 1.

Fig. 4 FA18(1)

Fig. 5 FA18(2)

Fig. 6 FA18(3)

Fig. 7 FA18(4)

Fig. 8 FA18(5)

Fig. 9 French_Rafale

Fig. 10 JF-17

Fig. 11 LCA_Tejas

Fig. 12 Mig35

Fig. 13 Sukhoi_SU30(1)

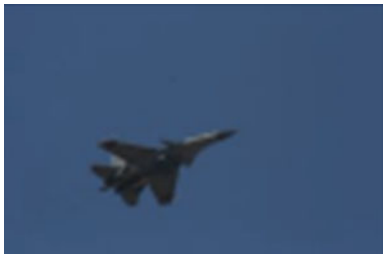


Fig. 14 Sukhoi_SU30(2)

Fig. 15 Sukhoi_SU30(3)

Fig. 16 Sukhoi_SU30(4)

Fig. 17 Sukhoi_SU30(5)

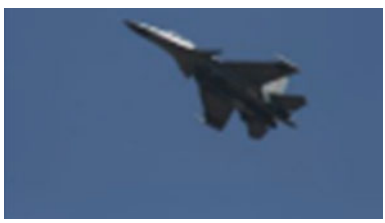


Fig. 18 Sukhoi_SU30(6)

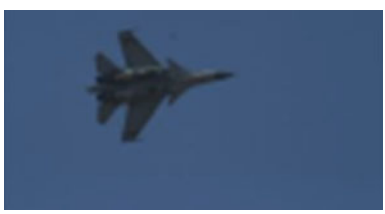


Fig. 19 Sukhoi_SU30(7)

Table 1 Results of moment invariance and DCT methods

No. of images in the database	No. of test images used	No. of test images recognized		Overall % recognition = Total no. of images recognized/Total no. of test images	
		MI	DCT	MI	DCT
19	23	14	20	(14/23) 60.8	(20/23) 86.9

6 Conclusions

In this study, aircraft recognition system using image analysis is described. The presented technique gives a computationally simpler algorithm for aircraft recognition using invariant moment and DCT features compared to other methods. The system is found to give an overall recognition of 60.8 and 86.9 %, respectively, for the two methods. The percentage recognition can be increased by including multiple view images for a single class of aircraft in the database. If the images have noise, then pre-processing steps for noise removal must be done.

References

1. Hu MK (1962) Visual pattern recognition by moment invariants. IRE Trans Inf Theor, pp 179–187
2. Dudani SA, Breeding KJ, McGhee RB (1977) Aircraft identification by moment invariants. IEEE Trans Comput, C-26(1):39–46
3. Chen CC (1993) Improved moment invariants for shape discrimination. Pattern Recogn, 26(5):683–686
4. Mercimek M, Gulez K, Mumcu TV (2005) Real object recognition using moment invariants. Sadhana, 30(6):765–775
5. Flusser J, Suk T (1993) Pattern recognition by affine moment invariants. Pattern Recogn 26(1):167–174

6. Ramteke RJ (2010) Invariant moments based feature extraction for handwritten devanagari vowels recognition. Int J Comput Appl 1(18):1–5
7. Yu G, Morel JM (2009) A fully affine invariant image comparison method. IEEE Int Conf Acoust, Speech and Signal Process, ICASSP, pp 1597–1600
8. Molina JM, García J, Berlanga A, Besada J, Portillo J (2002) Automatic video system for aircraft identification, ISIF, pp 1387–1394
9. Somaie AA, Badr A, Salah T (2001) Aircraft image recognition using back propagation, Proc CIE Int Conf Radar, pp 498–501
10. Saghafi F, Khansari Zadeh SM, Etminan Bakhsh V (2008) Aircraft visual identification by neural networks. JAST 5(3):123–128
11. Hsieh JW, Chen JM, Chuang CH, Fan KC (2005) Aircraft type recognition in satellite images. Vision, image and signal process, IEEE Proc 152(3):307–315
12. Rihaczek AW, Hershkowitz SJ (2001) Identification of large aircraft, IEEE Transactions on Aerospace and Electronic Systems, 37(2):706–710
13. Jayaprabha P, Somasundaram RM (2012) Content based image retrieval methods using graphical image retrieval algorithm (GIRA). Int J Inf Commun Technol Res, 2(1):22–28
14. Datta R, Li J, Wang JZ (2005) Content-based image retrieval—approaches and trends of the new age, Copyright ACM 1-58113-940-3/04/0010
15. Vadivel A, Majumdar AK, Sural S (2004) Characteristics of weighted feature vector in content-based image retrieval applications. Proc Int Conf Intell Sensing Information Process, 127–132
16. Kim DY, Chien SII, Son H (1991) Multiclass 3-D aircraft identification and orientation estimation using multilayer feed forward neural network, IEEE Int Joint Conf on Neural Networks, 1(18–21):758–764
17. Rui Y, Huang TS, Chang SF (1999) Image retrieval: current techniques, promising directions, and open issues. J Vis Commun Image Represent 10:39–62
18. Eakins J, Graham M (1999) Content-based image retrieval. Joint Inf Syst Committee Appl Programme, Report 39
19. Jain M, Singh SK (2011) A survey on: content based image retrieval systems using clustering techniques for large data sets. Int J Manag Inf Technol (IJMIT), 3(4):23–39
20. Srinivasa Rao C, Kumar SS, Mohan BC (2010) Content based image retrieval using exact legendre moments and support vector machine. Int J Multimedia Appl (IJMA) 2(2):69–79
21. Gonzalez RC, Richard E (2009) Woods, digital image processing, Pearson (3rd edn)
22. http://www.mathworks.in/help/toolbox/images/f21-16366.html
23. Jain AK (2002) Fundamentals of digital image processing, Prentice-Hall, pp 150–153
24. http://www.fighter-planes.com
25. http://www.combataircraft.com/en/Military-Aircraft/Chengdu/J-20/

Design of Physical Layer Transmitter and Receiver for Proximity-1 Space Link Protocol Using VHDL

Nikhil Kulkarni and S. Rohith

Abstract Proximity-1 is a bidirectional space link protocol to be used by space applications. It consists of a physical layer and a data link layer. This protocol has been designed to meet the requirements of space applications for efficient transfer of space data over various types and characteristics of proximity space links. In this paper, transmitter and receiver of physical layer have been designed. Here, VHDL language is used for designing the protocol, and Modelsim 6.1 simulator is used for coding and verifying the simulation results, respectively.

Keywords Physical layer · Data link layer · Mode · Transceiver · Proximity · Protocol · ASM · PCM · MIB

1 Introduction

The physical layer is responsible for the ultimate transmission of data over any network communication media. It operates with data in the form of bits that are sent from the physical layer of the sending (source) device and received at the physical layer of the destination device. The physical layer consists of the basic networking hardware transmission technologies of a network. It is a fundamental layer underlying the logical data structures of the higher-level functions in a network [1, 2]. In this paper, physical layer design includes two parts: first, the sender part and second, the receiver part.

N. Kulkarni (✉) · S. Rohith
Department of Electronics and Communication, Nagarjuna College of Engineering and Technology, Bangalore, Karnataka, India
e-mail: nikhil.bvb85@gmail.com

S. Rohith
e-mail: rohithvjp2006@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_22, © Springer India 2014

The rest of the paper is organized as follows. Section 2 provides the overview of various layers of the protocol. In Sect. 3, overview of the physical layer is discussed. Design of physical layer is discussed in Sect. 4. The simulation results of the design are given in Sect. 5. In Sect. 6, the conclusion and scope of the future work are discussed. The details of references needed to design and develop this protocol are mentioned in Sect. 6.

2 Overview of Proximity-1 Space Link Protocol Layers

The Proximity-1 is a bidirectional space link protocol used for space application. There are two main layers like physical layer and data link layer [1]. Each layer's activities are divided into a sender side and a receiver side.

2.1 Physical Layer

Physical layer of Proximity-1 space link protocol. The specification for the channel connection process, provision for frequency bands and assignments, modulation, data rates, and performance requirements are defined in this layer. Currently, the physical layer only defines operations at UHF frequencies for the Mars environment [1].

2.2 Data Link Layer

Data link layer is responsible for providing coded symbols to the physical layer at the transmitter side. On the receiver side, data link layer accepts the serial coded data from the physical layer. This layer is subdivided into the five sublayers and as shown in Fig. 1 [1, 3].

2.2.1 Frame Sublayer

The frame sublayer includes frame validation procedures, such as transfer frame header checks, and supervisory data processing for supervisory frames.

2.2.2 Data Services Sublayer

The data services' sublayer defines the frame acceptance and reporting mechanism for proximity links (FARM-P) (receiver side) and the frame operation procedures

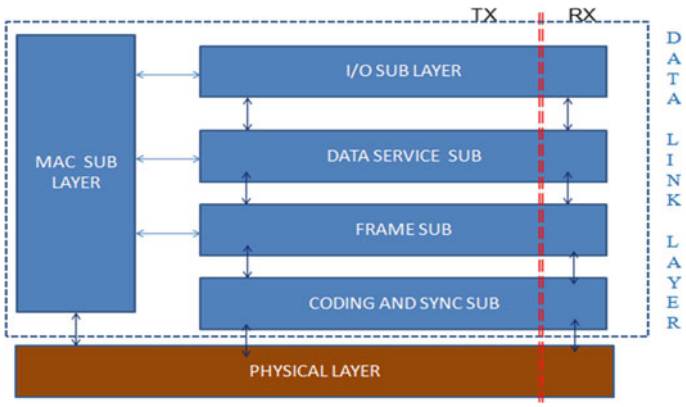


Fig. 1 Overview of proximity-1 layers

for proximity links (FOP-P) (sender side) associated with the expedited and sequence-controlled data services including how the FOP-P and FARM-P (COP-P) operate in the sequence-controlled service.

2.2.3 Input/ Output Sublayer

The input/output (I/O) interface sublayer provides the interface between the transceiver and the onboard data system and their applications.

2.2.4 Medium Access Control Sublayer

The medium access control (MAC) sublayer defines how a session is established, maintained (and how characteristics are modified, e.g., data rate changes), and terminated for point-to-point communications between proximity entities. This sublayer builds upon the physical and data link layers' functionality. The MAC controls the operational state of the data link and physical layers. It accepts and processes supervisory protocol data units (SPDUs) and provides the various control signals that dictate the operational state.

2.2.5 C&S Sublayer

The coding and synchronization sublayer includes PLTU delimiting and verification procedures. In addition, this sublayer performs the following operations.

3 Physical Layer Overview

Proximity-1 activities are divided between a sender side (TX) and a receiver side (RX). The sender side is concerned with the transmitted physical channel and also with the acquisition of the received physical channel in order to establish a Proximity-1 link. The operation of the transmitter is state driven. The receiver side is concerned with the reception of data on the received physical channel. The following Fig. 1 shows the overview of Proximity-1 layers.

The functions of physical layer are as follows:

1. Physical layer to data link layer interfaces.
2. Configuration of physical layer.

1. *Physical Layer to Data Link Layer Interfaces*

The prime function of the physical layer is to establish and maintain a communication channel upon which the data can flow. To enable a physical channel connection, the physical layer goes through a series of actions to establish a communication channel. The physical layer accepts control variables (MODE, DUPLEX, TRANSMIT, MODULATION) from the MAC sublayer of the data link layer for control of the transceiver. Figure 2 shows the data and control flows between the physical layer and elements of the data link layer [1].

2. *Configuration of Physical Layer*

The establishment of the communication channel depends on the configuration of the following physical layer parameters, frequency, polarization, modulation, acquisition and idle sequence and symbol rates, such that common operating characteristics exist in both communicating entities. The MAC sublayer sets the local transceiver to the desired physical configuration, under the control of the directives [1] SET TRANSMITTER PARAMETERS and SET RECEIVER PARAMETERS. A SET PL EXTENSION's directive is the mechanism by which additional physical layer parameters defined outside of the Proximity-1 physical layer can be enabled or disabled. The format and content of these and other Proximity-1 directives are specified in an annex of reference.

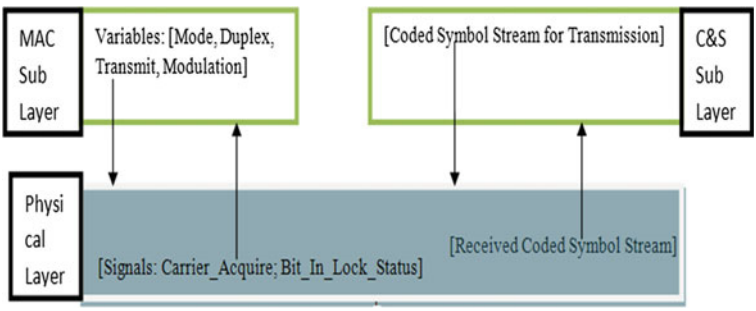


Fig. 2 Control variables, signals, and data transfers

4 Design of Physical Layer

Physical layer design includes two parts: first, the sender part and second, the receiver part; these are briefed as follows, and design of transmitter part is shown in Fig. 3 [1].

4.1 On the Sender Side of the Physical Layer

It accepts control variables from the MAC sublayer of the data link layer for control of the transceiver and accepts a coded symbol stream from the coding and synchronization sublayer of the data link layer for modulation onto the radiated carrier.

It provides an output bit clock to the coding and synchronization sublayer in order to receive the output bit stream. It provides status, that is, Carrier_Acquired and Bit_In_Lock_Status signal to the media access control sublayer.

The sender side of the physical layer consists of the following blocks:

- PCM Signal Generator Block (RF Out): RF_OUT represents all of the possible signal outputs to the communication partner from the physical layer. These consist of off (no signal), carrier only, idle data, etc., as shown in Fig. 3.
- Clock Generation Block: It is used for generating a few KHz clock signal from a range of MHz which is suitable for the design as shown in Fig. 3 [4].

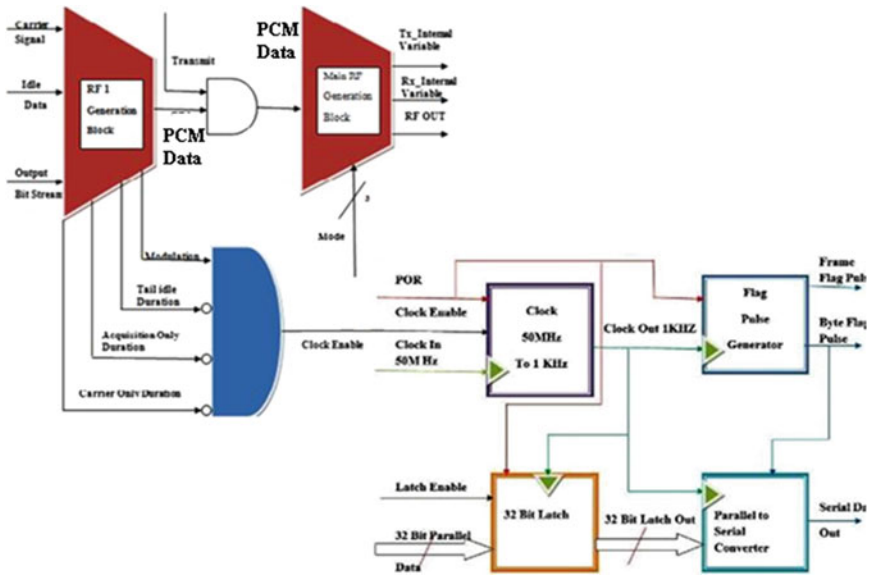


Fig. 3 Design of sender side of the physical layer

- Latch and Flag Pulse Generator: Latch is used to latch out the parallel data from C&S sublayer, and it generates two pulses one at every 32 clock cycles for indicating transfer of 32 bit data and another at ever 1,024 clock cycles to indicate the transfer of frame as shown in Fig. 3.
- Parallel-to-Serial Converter: Here, the parallel data from coding and synchronization sublayer is converted into serial data for further transmission as shown in Fig. 3.

4.2 On the Receiver Side of the Physical Layer

It provides status signals to the MAC sublayer of the data Link layer, that is, it provides the received bit Clock/Data to the coding and synchronization sublayer. It is shown in Fig. 4.

The receiving side of the physical layer consists of the following blocks:

- Serial-to-parallel converter: It is used to convert serial data to parallel data at the receiver.
- 32 bit latch: It is used to latch out the parallel data only when ASM pulse appears.
- ASM Comparator: It compares the 24 bit ASM value with the 32 bit parallel data.
- D flip-flop: It is used to indicate the ASM pulse with respect to clock.

5 Simulation Results of Physical Layer

The Modelsim 6.1 simulator tool is used for coding and simulating the design of the protocol using VHDL style of coding, since Modelsim is one of the industry standard tools, and hence, it can be used for the above-mentioned purpose.

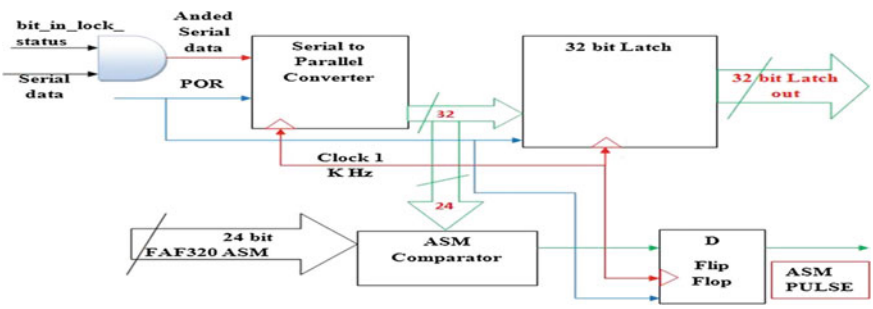


Fig. 4 Physical layer receiver module

The following Figs. 5, 6, 7, 8, 9, 10 show the simulation results of transmitter blocks discussed in Sect. 3 designs of physical layer [1]. The Figs. 5 and 6 show the simulation results of transmitter-state RF generator block, where the output will be in carrier state, ideal data, and serial output bit stream. Figures 7 and 8 show the clock generator output, and here, the input clock is operating at 50 MHz and output clock is operating at 1 kHz, and Fig. 9 shows the simulation results of pulse generator where we get two flag pulses to indicate 32 bit flag pulse (shown as combination of pulses in Fig. 9) and frame flag pulse at every 1,024th clock cycle to indicate one frame. Figure 10 shows the simulation results of parallel data converted into serial data coming from coding and synchronization layer.

The Figs. 11, 12, 13, 14 show the entire receiver simulation results Fig. 11 shows serial-to-parallel converter in which serial data is $(3E0F8000)_H$. Figure 12 shows the receiver latch where the parallel data are latched to the physical layer of receiver only when ASM pulse is present, ASM comparator is shown in Fig. 13 where it compares the 24 bit ASM value $FAF320_H$ with parallel data to latch out the parallel data at that instant and entire physical layer receiver block results are shown in Fig. 14.

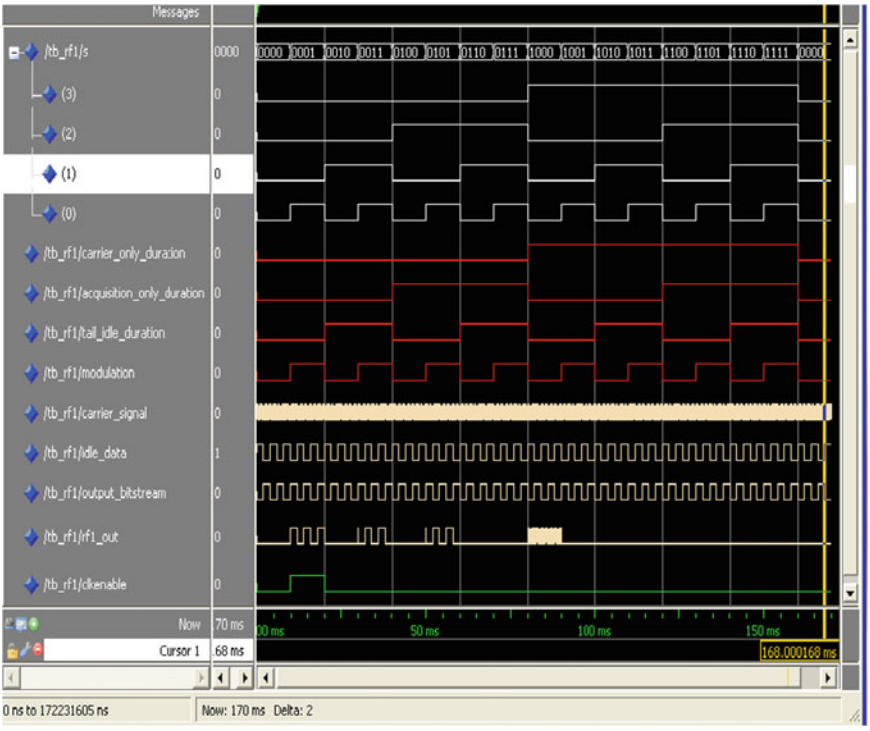


Fig. 5 PCM signal generator

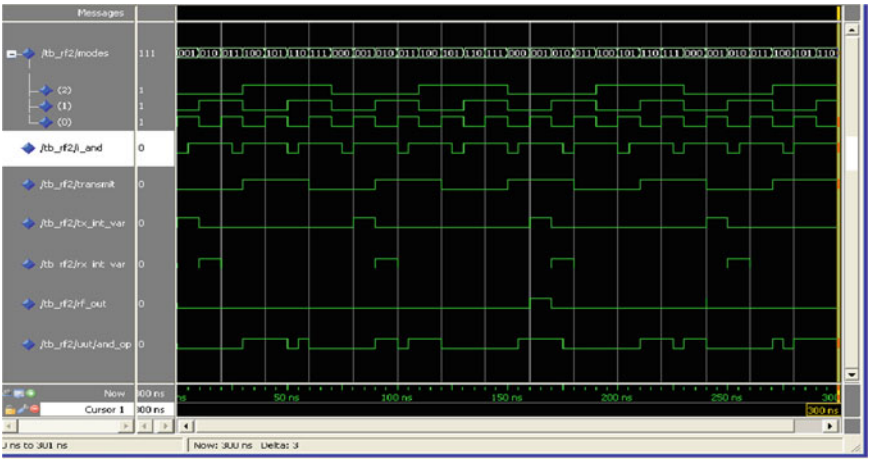


Fig. 6 RF signal generator

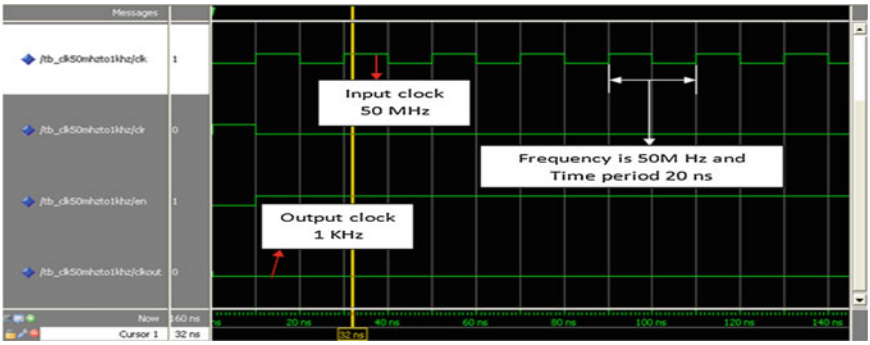


Fig. 7 Clock generator input 50 MHz

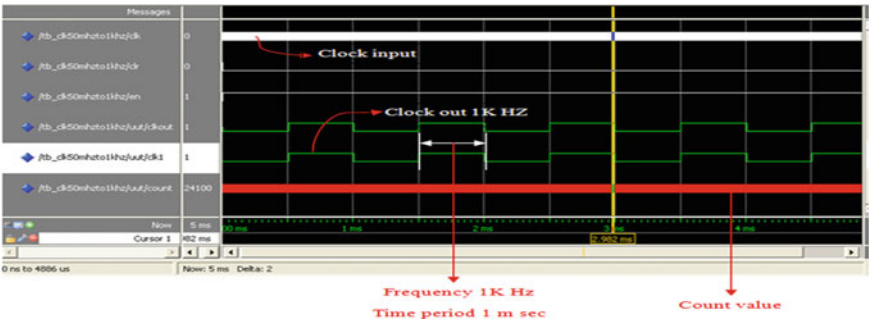


Fig. 8 Clock generator output 1 kHz

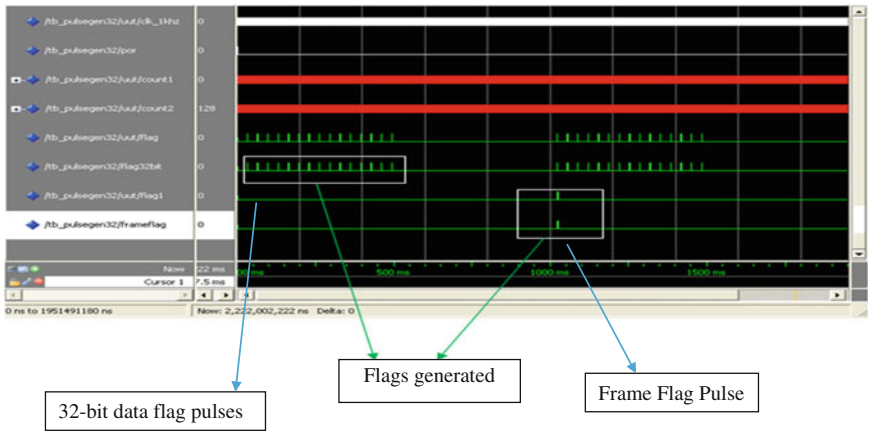


Fig. 9 Pulse generator signals

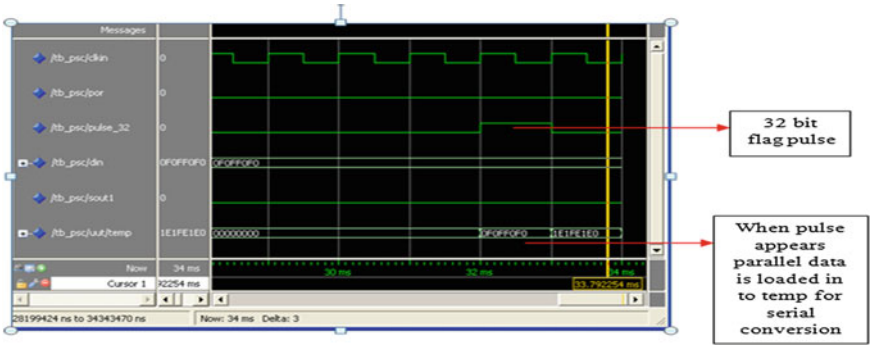


Fig. 10 Parallel-to-serial converter

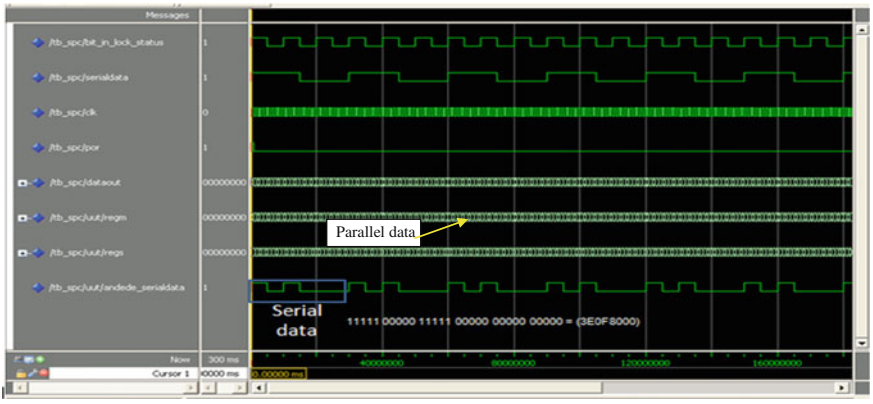


Fig. 11 Serial-to-parallel converter

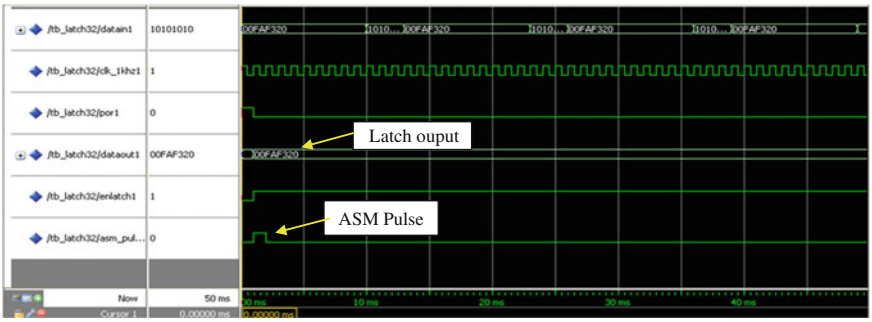


Fig. 12 Receiver latch

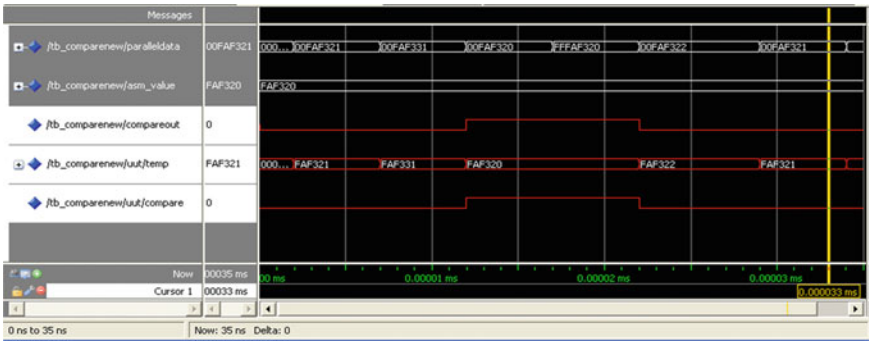


Fig. 13 ASM comparator

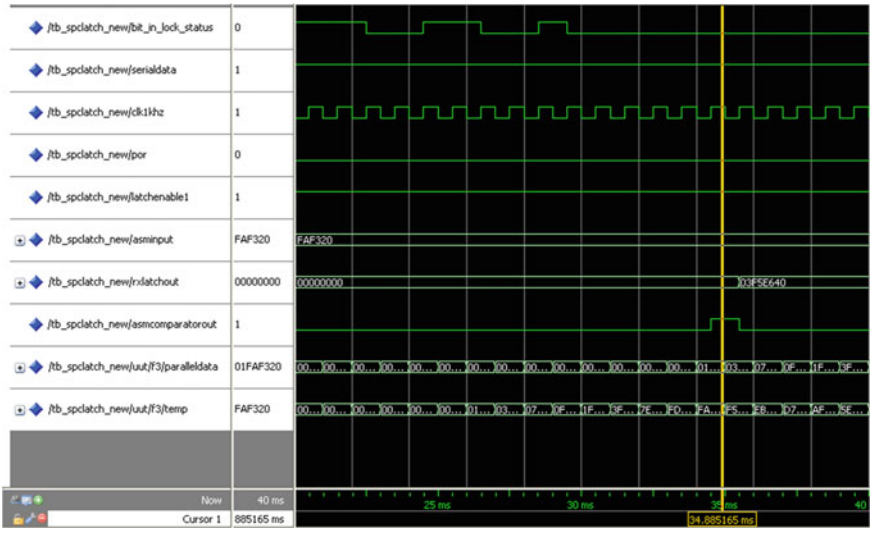


Fig. 14 Entire receiver results

6 Conclusion and Further Enhancements

The design and simulation of transmitter and receiver parts of the physical layer for Proximity-1 space link protocol are analyzed, designed, and simulated. In this, RF signal generator, clock generator, pulse generator and parallel-to-serial converter, serial-to-parallel converter, ASM comparator, and D flip-flop are designed using VHDL and simulated by Modelsim simulation tool and results are given in Sect. 5 . Since the protocol is an open source protocol, it can be designed as per the individual requirements. In future, we planned to implement the design using FPGA.

References

1. Proximity-1 Space Link Protocol—Physical layer (2011) Recommendation for space data system standards, CCSDS 211.1-B-3. Blue Book. Issue 3. Washington, D.C.: CCSDS
2. Tanenbaum AS (2003) Computer networks, 4th edn. Prentice Hall, New Jersey
3. Proximity-1 Space Link Protocol—Coding and Synchronization Sub layer. (2003) Recommendation for space data system standards, CCSDS 211.2-B-1.Blue Book. Issue 1. Washington, D.C.: CCSDS
4. Rohith S, Kulkari N (2012) Design of physical layer of proximity-1 space link protocol using VHDL. In: Proceedings of ICVSP-2012 KSIT-Bangalore, pp 4–5

Author Biographies

Nikhil Kulkarni received his BE degree in electronics and communication engineering in 2009 from B.V.B.C.E.T Hubli and M. Tech degree in VLSI design and embedded systems in 2012 from N.C.E.T Bangalore, Karnataka. He is currently working as a lecturer at V.D.R.I.T, Haliyal. His main area of interest includes analog design, embedded coding and VLSI design.

Rohith.S received his BE degree in electronics and communication engineering in 2006 and M. Tech degree in VLSI design and embedded systems in 2008 from Visvesvaraya Technological University, Karnataka. He is currently working as an assistant professor at Nagarjuna College of Engineering and Technology, Bengaluru. His main area of interest includes digital watermarking, steganography, error control coding, cryptography, and VLSI design.

Semantic Analysis of Precise Detection Rate in Multi-Object Mobility on Natural Scene Using Kalman Filter

D. Pushpa and H. S. Sheshadri

Abstract Detection counting as well as gathering features to perform analysis of behavior of natural scene is one of the complex processes to be design. The current work focuses not only to detect and count the multiple moving objects but also to understand the crowd behavior as well as exponentially reduce the issues of inter-object occlusion. The image frame sequence is considered as input for the proposed framework. Unscented Kalman filter is used for understanding the behavior of the scene as well as for increasing the detection accuracy and reducing the false positives. Designed on Matlab environment, the result shows highly accurate detection rate.

Keywords Multiple Object Detection · Object tracking · Object counting · Foreground · Background · Kalman Filter

1 Introduction

In the last few years, object detection and localization have become very popular areas of research in computer vision. Although most of the categorization algorithms tend to use local features, there is much more variety on the classification methods. In some cases, the generative models show significant robustness with respect to partial occlusion and viewpoint changes and can tolerate considerable intra-class variation of object appearance [1]. However, if object classes share a high visual similarity, then the generative models tend to produce a significant

D. Pushpa (✉) · H. S. Sheshadri
Maharaja Institute of Technology, Mysore, PES College of Engineering, Mandya, India
e-mail: pushpad15@gmail.com

H. S. Sheshadri
e-mail: hssheshadri@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_23, © Springer India 2014

number of false positives. On the other hand, the discriminative models permit us to construct flexible decision boundaries, resulting in classification performance often superior to those obtained by only generative models [2, 3]. However, they contain no localization component and require accurate localization in positions and scale. In the literature, the standard solution to this problem is to perform an exhaustive search over all positions and scales. However, this exhaustive search imposes two main constraints. One of them is the detector's computational complexity. It requires large computational time for relatively large number of objects. The second is the detector's discriminance, since a large number of potential false positives need to be excluded. The current research work has been experimented with monocular-adjusted image capturing device, which facilitates to design a distinctive connection between the subjects walking in real-time and image plane. The phenomenon is considered for isometric view of the subjects walking. In Section 2, we review recent work on multi-object detection. In Sect. 3, problem description is discussed in brief followed by proposed system in Sect. 4. Section 5 discusses about research methodology followed by Sect. 6 that discusses about implementation. Result is discussed in Sect. 7 and finally concludes by summarizing the entire research work in Sect. 8.

2 Related Work

The prior works done in the line of the proposed techniques are tabulated below (Table 1):

3 Problem Description

Object recognition is a computationally expensive process as well as challenging because a combination of factors must be considered to identify objects. These factors may include limitations on allowable shapes, the semantics of the scene context, and the information present in the image itself. Given a video clip, the initial problem is segregating it into number of frames. Each frame is then considered as an independent image, which is in RGB format and is converted into gray-scale image. Next, the difference between the frames at certain intervals is computed. This interval can be decided based on the motion of moving object in a video sequence. If the object is moving quite fast, then the difference between every successive frame has to be considered, posing a difficulty in designing smart object identification system.

Table 1 Existing approaches used in moving object detection

Years	Authors	Problem focused	Techniques used	Remark
2010	Nicolas [4]	Counting people	Skelton graph, background subtraction	Objects other than human are not considered
	Conte et al. [5]	Counting people	Motion vector, support vector regressor	Objects other than human are not considered
	Burkert et al. [6]	Monitoring people based on aerial images	Microscopic & mesoscopic parameter, trajectory	Inter-object occlusion not focused on
	Conte et al. [7]	Counting moving people	Detection using interesting points	Objects other than human are not considered
	Widhalm [8]	Flow of pedestrian	Growing neural gas, dynamic time warping	Objects other than human are not considered, inter-object occlusion not focused on
2011	Dehghan et al. [9]	Automatic detection & tracking of pedestrians	Static floor field, boundary floor field, dynamic floor field	Addressed occlusion, good results obtained, not focused on multiple-objects moving
	Spampinato [10]	Event detection in crowd	Lagrangian particle dynamics	Results found with issue in noises in frames affecting the performance of flow map estimation
	Suhr [11]	Moving object detection	Background compensation method using 1-D feature matching & outlier rejection	Not focused on inter-object occlusion, multiple objects are not considered
	Lei et al. [12]	Maritime surveillance	Trajectory pattern mining	Better result but cannot be used in crowded scenes
	Sirmacek [13]	Tracking people from airborne images	Kalman filter	Not focused on inter-object occlusion, multiple objects are not considered
2012	Cui et al. [14]	Moving object detection	Graph cuts	Not focused on inter-object occlusion, multiple objects are not considered
	Chen et al. [15]	Counting people in crowd	2-stage segmentation	Better result, but impact of occlusion & multiple objects are not considered
	Gowsikhaa [16]	Human activity detection	Artificial neural network	Topic is narrowed down to face recognition system, occlusion issue is not addressed
	Rosswog [17]	Object detection in crowd	SVM	Occlusion issue is not addressed
	Arif [18]	Counting moving object	Neural network	Illumination state & occlusion issues are not addressed

4 Proposed System

The prime aim of the proposed work is to design a robust framework that is capable of identifying mobile objects using sequences of frames captured from video. The research work is designed not only for its utility in crowd detection that was presented in [19], but now, the proposed framework will be used for extensive analysis of the video dataset. Various knowledge and data about the coordinates and characteristics of the mobile object at various instance of time as the fundamental of identifying abnormal mobile object detection pertaining to its movement.

5 Research Methodology

The proposed system uses the frames captured from the still camera, where the visibility is restricted to a defined area of focus of the aperture of the camera. This method is adopted to analyze the preciseness of the mobility of the object, which could not be done on mobile video sensors. The system uses distinctive technique to filter out the foreground objects from the background for better accuracy. Basically, the background model is considered as prime module as the experimental framework designs for every pixels of the captured frame as a distribution of the typical values considered by background module. The main concept behind this is the most frequent values of a pixel that are corresponding to the background metaphors. The system is designed in such a way that the iterative mobility in the background is also captured. That will mean that even a slightest displacement on any object on the considered visibility screen will also be captured.

The categorization of the pixel into background or foreground is designed depending on the collected module of background using outlier detection. According to this design, the pixels with the intensity values sufficiently distinctive from the background distribution are indexed as foreground. The framework will also use gradual alteration in the illumination factor by spontaneously upgrading the background model of every pixel in the frames. The same phenomenon will also be applicable to the static objects in the considered scene. An example input image and the corresponding pixel labeling computed from a background model can be seen on the left where foreground pixels were labeled as white. The person crossing the street can be clearly seen in Figs. 1 and 2.

From Figs. 1 and 2, by lumping large connected foreground regions into blobs, foreground objects, such as the person seen crossing the street is identified. However, the left images also show that an additional foreground object was erroneously detected in the top right corner due to violent motion of the trees in the background. Such occasional misdetections are taken into account at later stages of the object detection algorithm. The consecutive stage will include a sequence of the object blobs together across various frames of the scene. In order to accomplish

Fig. 1 Showing original scene, Mask

Fig. 2 Showing original scene, Mask

this task, various attributes like position, velocity, and size are used to represent every object block. Unscented Kalman filter is specifically used in this work for this purpose for determining the actual coordinates of the mobile object identified on the scene frames. The main purpose of adopting this step is to truncate the possibility of noise from blob estimations which could possibly lead to generation of the false-positives values.

6 Implementation

The proposed work is performed on 32-bit Windows OS. In order to smoothly run the heavy computation from large datasets of frames, it is highly essential to use high-configuration system with minimum 2.20-GHz processor with Intel core-i3 along with 4-GB RAM size. The programming is done on Matlab environment due to its faster computation of complex problems. The image sequence datasets has been taken from [19].

By deploying the computed indexing, foreground objects are identified; however, direct use of these blobs for tracking is sensitive to noise and fails to determine object identities at different times in a sequence. In order to couple a sequence of state measurements, that is, relevant blob properties such as the centroid, centroid velocity between consecutive frames, or size into a denoised-state trajectory, the filter operates as a two-stage process switching between absorbing evidence from the most recent foreground blob and forming a predictive distribution for the next. In the context of tracking multiple objects simultaneously, the key difficulty lies in determining which unscented Kalman filter should absorb which new state measurement.

Algorithm: Detection of object mobility
Input: Image sequence datasets
Output: Appropriate detection of object movement in visualization
Start
1 Load set of image sequence
2 Returns data matrix
3 Reading Indexed individual frames
4 Reading RGB components of the scene
5 Reading frame number
6 Estimate the log of the Gaussian distribution for each point in x using Mu and Sigma.
7 x = matrix of column-wise data points.
8 Mu = column vector
9 Sigma = Covariance matrix of respective size
10 Estimate the log probability for each column in x-centered.
11 Deploy other parameters
12 weight = matrix of mixing proportion
13 Sigma = Covariance
14 Active_Gaussian = components that were either matched or replaced.
15 Background_Threshold = Threshold standard deviation for outlier detection on each Gaussian.
16 K = Quantity of components used in the mixtures.
17 Alpha = How fast for adapting the component weights
18 Rho = How fast to adapt the component means and covariances
19 Threshold_Deviation = Threshold used for finding matching versus unmatched components
20 INIT_Var = Initial variance for newly placed components
21 INIT_MixProp = Initial prior for newly placed components
22 COMPONENT_THRESH: Filter out connected components of smaller size
23 BACKGROUND_THRESH: Percentage of weight that must be accounted for by background models
24 Design a helper function for structure of Kalman filtering
25 Define Kalman State type
26 Create best background image

27 Computes a beta probability density on image coordinates pos
28 set the state transition matrix
29 Update using position only
30 Update using position and velocity
31 Compute a cost matrix
32 Computes the squared distance between the centroid predictions of a list of objects and a list of observed blob centroids.
33 Compute a cost matrix
34 Compute the square distance between Kalman filter predictions and observations
35 Create foreground_map
36 Binary matrix
37 Define chromaticity coordinates
38 Show visualization of object detection
End

7 Result Discussion

In order to accomplish the project work, the whole detection protocol was assessed on a number of different sequences in order to analyze its accurateness and weakness. The previous work [19] done is able to detect objects from crowded scene, but false positives are quite moderate in number. Hence, this paper will use the technique of foreground detection that can be used for extracting useful information under various distinctive conditions, particularly when the data stream is recoded in different color space. The interesting thing is that it also shows every minute movement even for shaking leaves of tree. Manual inspection of various detection results suggested that multiple-object tracking using the devised association technique performs adequately in many situations.

The assessment is done considering three case scenarios, for example, (1) successful detection rate of object: Fig. 3 shows very effectively both truck and

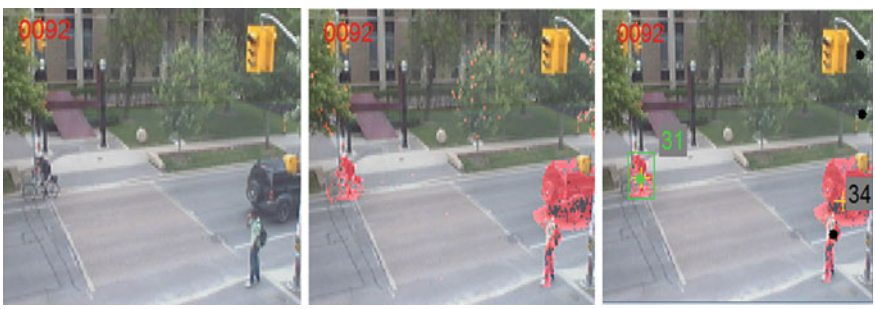


Fig. 3 Result showing successful detection rate

Fig. 4 Result showing successful detection of multiple object and counting

Fig. 5 Result showing preciseness in multiple object detection counting

human along with the person riding bicycle. (2) successful detection of multiple-object and counting: Fig. 4 shows very distinctively the human as well as truck as a moving objects. (3) Preciseness in multiple object detection counting: Fig. 5 shows an accurate detection of a person riding bicycle, a person walking, and a person walking along with pram. All the three objects are accurately shown and counted too. The above-implemented result shows better detection rate for multiple object with precise bounding box over the object. It successfully counts the objects too. The previous model has some detection issue due to inter-occlusion. However, 95 % of the inter-occlusion issues are overcome in this model with highly precise detection rate.

8 Conclusion

This paper has presented a novel and highly precise model of detecting multiple object in real-time street scene with highly uncertainty of the types of objects, their movements, and their count rates. The previous model presented has been addressed for object detection from crowded scene with better detection rate, but

even some moderate false positives were also raised due to inter-occlusion and variances in illumination. However, using unscented Kalman filter, such issues were overcome in this model. The scope of the application for this model is now much higher. If the previous model could be used to detect multiple object, then this model can be used to analyze the behavior of object in the real-time scene. The simulation result shows highly precise detection rate. However, these will eventually limit the specification of the model in terms of segregating the targeted object with not so important object. For example, due to proposed algorithm that has used both foreground and background model, the framework detects even the minor movement of leaves of trees. This property might be little unwanted when we want to design an application specific to analyze pedestrian or car or some major moving object visually that keeps highest importance in application. Hence, the next further work will be enhanced the model using homographic transformation.

References

1. Josef S, Russell BC, Alexei A, Zisserman A, William T (2005) Discovering objects and their location in images. Proceedings ICCV'05, Beijing, China, 370–377 2005
2. Ng AY, Jordan MI (2001) On discriminative versus generative classifiers: a comparison of logistic regression and naive bayes. In: Proceedings NIPS'01, Vancouver, British Columbia, Canada, 841–848 2001
3. Ferrari V, Fevrier L, Jurie F, Schmid C (2008) Group of adjacent contour segment for object detection. PAMI 30(1):30–51
4. Merad D, Aziz KE, Thome N (2010) Fast people counting using head detection from skeleton graph. In: 2010 Seventh IEEE international conference on advanced video and signal based surveillance
5. Conte D, Foggia P, Percannella G, Tufano F (2010) A method for counting people in crowded scenes, EURASIP Journal on Advanced Signal Processing, 2010
6. Burkert F, Schmidt F, Butenuth M, Hinz S (2010) People tracking and trajectory interpretation in aerial image sequences. IAPRS 38:209–214
7. Conte D, Foggia P, Percannella G, Tufano F, Vento M (2010) Counting moving people in videos by salient points detection. In: International conference on pattern recognition, 2010
8. Widhalm P, Brandle N (2010) Learning Major Pedestrian Flows in Crowded Scenes. In: International conference on pattern recognition, 2010
9. Dehghan A, Idrees H, Zamir AR, Shah M (2011) Keynote: automatic detection and tracking of pedestrians in videos with various crowd densities. Springer, Berlin
10. Spampinato C, Faro A, Palazzo S (2011) Event detection in crowds of people by integrating chaos and lagrangian particle dynamics. In: Proceedings 3rd international conference on information and multimedia technology (ICIMT 2011), Dubai, UAE, 28–30 Dec 2011
11. Suhr JK (2011) Moving object detection for static and Pan-Tilt-Zoom cameras in intelligent visual surveillance, A doctorial thesis, Graduate School of Yonsei University, 2011
12. Lei PR, Su IJ, Peng WC, Han WY, Chang CP (2011) A framework of moving behavior modeling in the maritime surveillance, Journal of Chung Cheng Institute of Technology, Vol. 40, No.2, pp 33–42
13. Sirmacek B, Reinartz P (2011) Kalman filter based feature analysis for tracking people from airborne images. In: ISPRS workshop high-resolution earth imaging for geospatial information, Hannover, Germany, Jun 2011

14. Cui Y, Zeng Z, Cui W, Fu B (2012) Moving object detection based frame difference and graph cuts. J Comput Inf Syst 8(1):21–29
15. Chen CH, Chen TY, Wang DJ, Chen TJ (2012) A cost-effective people-counter for a crowd of moving people based on two-stage segmentation. J Inform Hiding Multimedia Signal Process 3(1):12–25. ISSN 2073-4212
16. Gowsikhaa D, Manjunath AS (2012) Suspicious human activity detection from surveillance videos, (IJIDCS). Int J Internet Distrib Comput Syst 2(2):141–149
17. Rosswog J, Ghose K (2012) Detecting and tracking coordinated groups in dense, systematically moving, crowds, SIAM-2012
18. Arif M, Saqib M, Basalamah S, Naeem A (2012) Counting of moving people in the video using neural network system. Life Sci J 9(3):1384–1392
19. Pushpa D, Sheshadri HS (2012) Precise multiple object identification and tracking using efficient visual attributes in dense crowded scene with regions of rational movement. IJCSI Int J Comput Sci 9(2): 2, pp 322–331, March 2012

EEG Seizure Analysis Using Fractal Dimensions During Electroconvulsive Therapy

S. Kotresh, G. B. Mukartihal, B. N. Gangadhar and G. Siddalingesh

Abstract Seizure is the first sign of neurological disease or dysfunction in the newborn. An electroencephalogram (EEG) is a commonly used noninvasive technique for studying brain activities. Waveforms are planar curves–ordered collections of (x, y) point pairs where the x-values increase monotonically. The fractal characterization may be especially useful for analyzing and comparing complex waveforms such as EEGs. Running fractal dimensions were measured on four channels of an EEG recorded from a volunteer. The changes in the background activity due to eye closure were clearly differentiated by the fractal method. The fractal method was also successful in detecting low amplitude spikes and the changes in the patterns in the EEG. The effects of different window lengths and shifts on the running fractal dimensions have also been studied.

Keywords Electroencephalogram · Electromotive force · Fractal dimension · Electroconvulsive therapy · Seizure · Electrical activity · Charge · Sampling

S. Kotresh (✉) · G. Siddalingesh
Information Technology Department, RaoBahadur Y. Mahabaleshwarappa Engineering College, Bellary, Karnataka, India
e-mail: kotresh2000@rediffmail.com

G. Siddalingesh
e-mail: siddug123@gmail.com

G. B. Mukartihal
Basaveshwara Engineering College, Bagalkot, Karnataka, India
e-mail: mukartihal2002@yahoo.co.in

B. N. Gangadhar
Department of Psychiatry, NIMHANS, Bangalore, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_24,

1 Introduction

Human brain generates electric potentials, which can be recorded by placing electrodes on the intact skull or scalp surface. These records are called electroencephalogram and the method itself the electroencephalography (EEG). The electrical activity of the brain is complicated when compared to that of single nerve fiber or neuron. This is due to large number of neurons and synapses.

EEG is interpreted in terms of the background activity, paroxysmal activity, transients, and symmetry across the hemispheres in terms of location of channels. EEG is used to study the relationship between the changes in electric potentials in different area of cortex and subcortical structures. The frequency contents of EEG are crucial importance in its assessment. In an EEG, the potential difference between an electrode position and its reference point is amplified and then recorded by pen writers on continuously moving paper. A typical recording session lasts approximately for about 20 min and generates 55 m of paper tracing. The frequency varies greatly with different behavioral states, and the nature of the wave varies over the different parts of the scalp [1].

The common neurological disorders in which the EEG pattern is altered are epilepsy, which occurs due to excessive discharge of impulses from cerebral cortex. The disorders of mid-brain affect the ascending reticular activating system and subdural hematoma during which there is collection of blood in the subdural space over the cerebral cortex [2].

2 Brief Overview of Fractal Dimensions and Electroconvulsive Therapy

Fractal dimensions are the mathematical concepts which measures geometrical complexity of an object, and the algorithms for estimating the fractal dimension calculate a numerical value using as data a time series for the specific problem. Fractal dimensions are powerful techniques for modeling complex irregular figures that cannot be described with classical geometry. Fractals are used in image synthesis application and used to measure length, area or volume of curves and surfaces.

The procedure of electroconvulsive therapy (ECT) essentially involves application of a low-frequency electrical stimulus across two electrodes applied over the scalp to induce a generalized seizure for therapeutic purposes. The stimulus is an electromotive force (volts-V) generated from the ECT device, which drives a flow of electrons. The current across the two electrodes applied on the patient's scalp. The intensity of this current is inversely dependent on the electrical impedance (ohms) offered in the circuit between the two electrodes. The stimulus is applied for a finite period (seconds-T) which can be controlled by an electronic timer. This is expressed in joules (J) or Watt second and is derived from the formula $J = VIT$.

ECT devices employed to generate a range of user selectable electromotive forces from 90 to 160 V. This voltage is derived from stepping down the mains AC voltage using a transformer. The current conducted in the circuit through the brain structures between the electrodes is governed by the inter-electrode impedance that comes from the electrode-skin contacts as well as the rest of the anatomical structures. Inter-electrode impedance is variable both across individuals and within individuals across ECT sessions. The current yields the charge traversing the brain tissue in units of coulombs (C) from the formula $C = IT$. The method of arriving at the stimulus dose in units of coulombs is referred as coulometry.

The conventional sine-wave stimulus is crude, and the slow rising and trailing edges of the sine wave do not produce efficient cerebral stimulation. This unnecessarily continues stimulation even while the neuronal tissue is in a post-depolarization refractory period. The sine wave was hence modified to eliminate the rising and trailing edges, retaining only the peaks in the form of 'square wave' pulses. The pulses are brief (1–2 m s) with no stimulus in between the pulses. Intermittent stimulation using pulses is more efficient as it allows the tissue to recover from post-depolarization refractory period. Hence, the preference is given for brief pulse stimulus.

The stimulus sources available today evaluate the inter-electrode impedance and dynamically adjust the voltage to ensure a constant current throughout the stimulus application. In the sine-wave stimulus, the amplitude of current or voltage gradually rises and trails in each cycle.

Some of the specifications of ECT device

Signal parameter	Voltage (V)	Current (I)	Frequency (F)	Width	Duration	Impedance (Z)	Energy	Charge
Pulse	160 V	0.8 A	100 PPS	1 ms	1 s	200 Ω	12.8 J	80 mC
Sine	120 V	0.6 A	50 Hz	–	0.5 s	200 Ω	36 J	300 mC

Following are the terms used in characterizing brief pulse stimuli:

Frequency refers to number of cycles per second, in the transformer-based sine-wave stimuli, and the line frequency is about 50 Hz. In the pulse stimulus, on the other hand, the operator varies the frequency but some devices keep this factor constant. It is a better convention to use the term pulses per second (PPS) in this form of stimulus.

1. *Pulse width*: It is a measure relevant to only pulse-wave stimuli. The recommended width is 1–2 m s.
2. *Stimulation duration*: It is the total period for which the stimulus train is applied. In the sine-wave stimulus, it is usually less than 1 s, and in pulse wave, it is between 0.5 and 0.8 s. It should be remembered in the latter waveform, the actual duration of stimulus application is given by this equation (pulse width x pulses per second x total duration of train), and this value is usually less than 10 % of the duration.

3. *Charge*: The total amount of electrons traversing the inter-electrode tissue including the brain is the product of current and actual duration of stimulus, and it is measured in milli coulombs. Stimulus estimation by charge is used in contemporary ECT practice [3].

3 Methodology

3.1 Data Collection

EEG signals from volunteer were recorded with EEG amplifiers. Four channels of unipolar are placed on right frontal, right temporal and left frontal, left temporal regions.

3.2 Filtering and Sampling

Raw EEG signals were filtered by using FIR filter and then sampled at 256 samples/s. By initialization of order and cutoff frequency, filter order $(M) = 40$, $Wc = 32$ rad/s, evaluating the transfer function of the filter. FIR filter have linear phase characteristics, good delay characteristics, depend on previous input samples, has only zeros and easier to implement [4].

$$h(n) = \text{Sin}\{Wc(n-\tau)\}/\pi(n-\tau);\ \tau = (M-1)/2 \quad \text{where } n = 0 \text{ to } M-1;$$

Convolving $h(n)$ with the input signal $x(n)$ to get the filtered out put as $y(n)$

$$y(n) = x(n) * h(n);$$
$$y(n) = \sum x(k)h(n-k) \text{ between the limits } k = 0 \text{ to } m-1$$

The filtered EEG signals were then processed for quantification of fractal dimensions.

3.3 Concept of fractal dimensions

The fractal dimension (D) is represented by an individual number (often a fraction) that can be used as a fundamental quantification of even the most complex of shapes.

Fractal dimension 'D' of a planar curve is $D = \text{Log}\ (L)/\ \text{Log}\ (d)$. where L is total length of the curve, and d is diameter of the curve.

$L = \text{Sum}\ [\text{dist}\ (i, i+1)]$, that is, distance between two point pairs $d = \text{Max}\ [\text{dist}\ (1, i)]$, that is, average distance between successive points using average step 'S'

$$D = \text{Log}\ (L/S)/\text{Log}(d/S)$$

$$D = \text{Log}(n)/\text{Log}(n) + \text{Log}(d/L)$$

where $n = L/S$ is the number of steps in the curve [5].

4 Results and Discussion

The seizure phases are classified into early, mid-, and post-seizures in the EEG data. In each of the phases, 4 s of data will be used for analysis. The early phase of seizure is the first 8 s following the stimulus offset, and the post-seizure is the first 8 s after seizure termination. Seizure termination is a clinical judgment of absence of all seizure discharges. Mid-seizure is defined as part of seizure EEG with maximum amplitude (Figs. 1, 2, 3, 4).

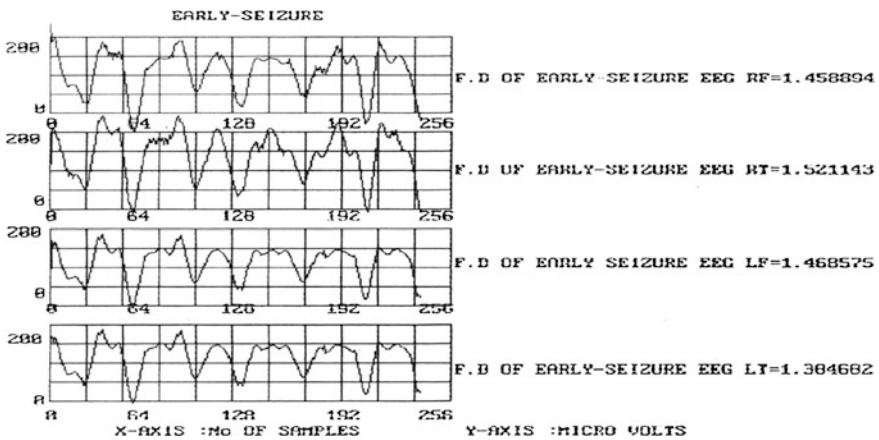


Fig. 1 Early seizure detection

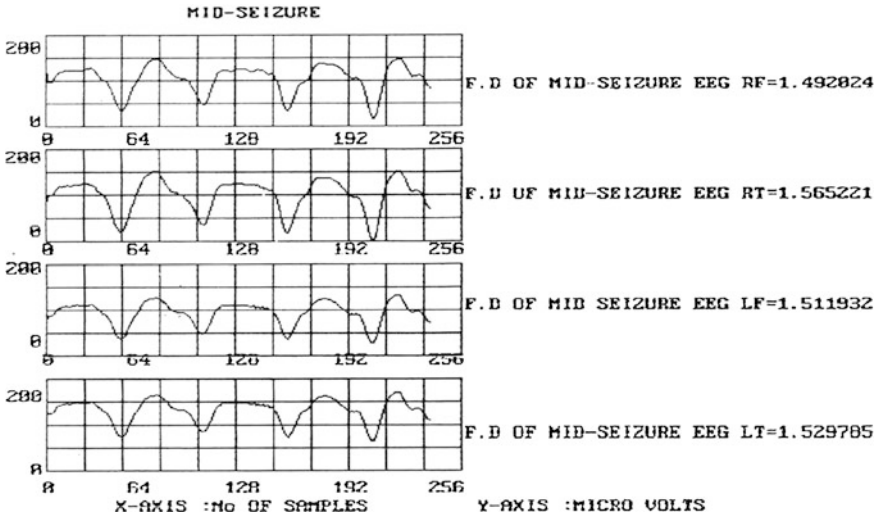


Fig. 2 Mid-seizure detection

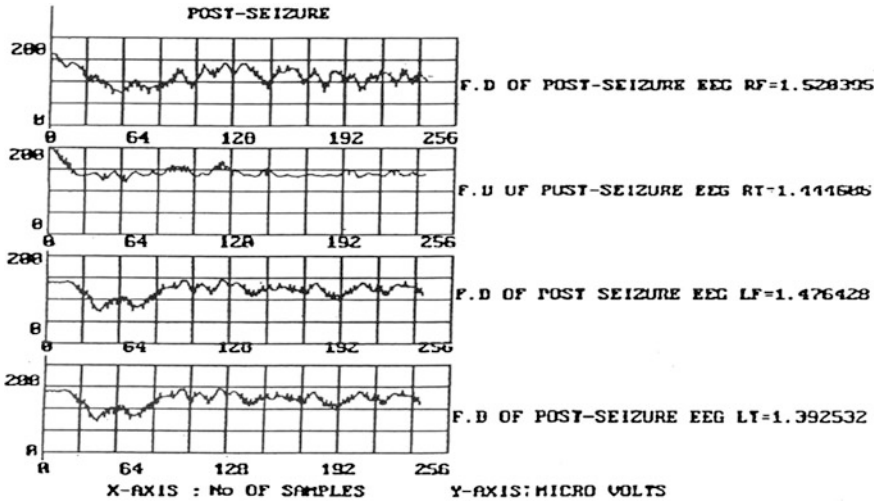


Fig. 3 Post-seizure detection

Fig. 4 Table of fractal dimensions

S.No	Type of seizure	Location of channel	Fractal dimension value
1	Early seizure	RF	1.4589
		LF	1.5211
		RT	1.4685
		LT	1.3847
2	Mid seizure	RF	1.4820
		LF	1.5652
		RT	1.5119
		LT	1.5297
3	Post seizure	RF	1.5203
		LF	1.4446
		RT	1.4764
		LT	1.3925

5 Conclusion

The present study suggests that running fractal dimension may be an effective method of presenting large-scale EEG data in a comprehensive single-line plot. The fractal values are higher with more complex and space filling properties than with those of simple curves. The nature of fractal values of biosignals may be of great help in the future in understanding its dynamic properties. The study of the analysis of the EEG has been made here, keeping in mind the nonlinear deterministic nature of the EEG. The obtained results suggest that fractal dimension can reflect changing pattern in EEG.

The fractal dimensions have a better potential for the real-time implementation. The fractal dimension value rises from early seizure to mid-seizure and had pronounced drop in the post-seizure in all four channels.

References

1. Pradhan N, Narayana Dutt D (1993) Use of running fractal dimensions for the analysis of changing patterns in electroencephalograms. Comput Biol Med 23(5):381–388
2. Sembulingam K, Sembulingam P Essentials of medical physiology
3. Gangadhar BN (1990) Proceedings of the national workshop on ECT; priorities for research and practice in India
4. Salivahan S, Vallavaraj A, Jnanapriya C A text book of digital signal processing
5. Katz JM (1988) Fractals and the analysis of wave forms. Comput Biol Med 18(3):145–156
6. Khandpur RS A textbook of bio-medical instrumentation
7. Mandelbrot BB (1983) The fractal geometry of nature. Freeman, New York
8. Peitgen HO, Jurgens H, Saupe D Chaos and fractals new frontiers of science
9. Jasper HH (1958) Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr Clin Neurophysiol 10:370–375

10. Hyrman V, Palmer LH, Cernic J, Jetlina J (1988) The search for the perfect stimulus. Biol Psychiatry 20:634–645
11. Endler NS (1988) The origins of electro convulsive therapy
12. Ottosson JO (1962) Electro convulsive therapy of endogenous depression; an analysis of the influence of the therapy. J Mental Sci 108:694–703
13. Grassberger P, Procacia I (1983) Measuring the strangeness of strange attractors. Physica 9D:189–208

Medical Image Denoising Using Wavelet-Based Ridgelet Transform

P. Vetrivelan and A. Kandaswamy

Abstract Noise is undesired or contaminated information present in images. During transmission, data may be affected due to noise and further processing of the same data does not produce good result. This affects the quality of the image which results in image blurring. Therefore, noise should be removed to get the clear information. In order to remove the noise, various denoising algorithms may be used. To preserve the details of the image, ridgelet transform uses hard thresholding algorithm. Ridgelet transform is used as it is concentrated near the edges of the image and it represents one-dimensional singularity in two-dimensional spaces. Wavelet is good in representing point singularities. When wavelet is linked with ridgelet, denoised image quality will be improved. Parameter like PSNR is calculated in order to measure the performance.

Keywords Ridgelet transform · Radon transform · Peak signal-to-noise ratio (PSNR) · Wavelet-based ridgelet transform (WBRT)

1 Introduction

Digital images play an important role in image processing fields such as image denoising [1], computer graphics, signal processing, pattern recognition, etc. Noise is mainly caused by sensors, imperfect instruments, problems occurring with data acquisition process; interference of natural phenomenon results in degradation

P. Vetrivelan (✉)
Department of Electronics and Communication Engineering, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India
e-mail: pvvslm@yahoo.co.in

A. Kandaswamy
Department of Biomedical Engineering, PSG College of Technology, Coimbatore, Tamilnadu, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_25,

of quality of images. Noise present in images may be either additive noise or multiplicative noise. Wavelet is the powerful transform which is used for representing zero-dimensional singularities, whereas ridgelet is a powerful transform for representing one-dimensional singularities in two-dimensional spaces.

2 Noise

Noise commonly present in an image is undesired information that contaminates the image. In the image denoising process, information about the type of noise present in the original image plays a significant role. Noise in imaging systems is usually either additive or multiplicative. The Gaussian additive white noise has a frequency spectrum that is continuous and uniform over a specified frequency band. It is spatially uncorrelated, and the noise for each pixel is independent and identically distributed.

2.1 Additive and Multiplicative Noises

Typical images are corrupted with noise modeled with either a Gaussian, uniform, or salt and pepper distribution. Another typical noise is a speckle noise, which is multiplicative in nature. Noise is present in an image either in an additive or in a multiplicative form. An additive noise follows the rule which is given in Eq. (1)

$$w(x, y) = s(x, y) + n(x, y) \tag{1}$$

The multiplicative noise follows the rule as it is given in Eq. (2).

$$w(x, y) = s(x, y) * n(x, y) \tag{2}$$

where

$s(x, y)$	original image
$n(x, y)$	noise introduced into the signal
$w(x, y)$	corrupted image
(x, y)	pixel location.

Image addition also finds application in image morphing. By image multiplication, the mean of brightness of the image is varied.

2.2 Gaussian Noise

Gaussian noise is evenly distributed over the signal. This means that each pixel in the noisy image is the sum of the true pixel value and a random Gaussian distribution, which has a bell-shaped probability function.

2.3 Salt and Pepper Noise

Salt and pepper noise is an impulsive type of noise, which is also referred to as intensity spikes. This is caused generally due to errors in data transmission. It has only two possible values: a and b. The probability of each is typically less than 0.1. The corrupted pixels are set alternatively to the minimum or to the maximum value, giving the image a "salt and pepper" like appearance. Unaffected pixels remain unchanged. For an 8-bit image, the typical value for pepper noise is 0, and for salt noise, it is 255. The salt and pepper noise is generally caused by malfunctioning of pixel elements in the camera sensors, faulty memory locations, or timing errors in the digitization process.

2.4 Speckle Noise

Speckle noise is a multiplicative noise. This type of noise occurs almost in all coherent imaging systems such as laser, acoustics, and synthetic aperture radar (SAR) imagery.

The source of this noise is attributed to random interference between the coherent returns. Fully developed speckle noise has the characteristics of multiplicative noise. Speckle noise follows a gamma distribution.

2.5 Poisson Noise

Poisson noise is frequently present in an image. It is not an additive noise rather it is pixel intensity dependent. Due to this noise, bright pixels in an image are statistically corrupted with noise than dark pixels. This causes a fluctuation in an pixel value of an image which leads to an complication in denoising process.

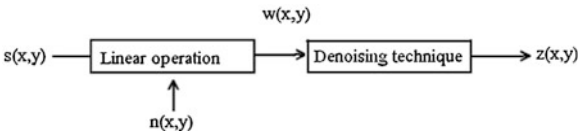


Fig. 1 Denoising concept

3 Image Denoising

Removing noise from the original image or signal is still a challenging problem for researchers [2]. The purpose of image denoising is to maintain the main features of the original image and remove noise from the image. Image denoising also has wide application domain in medicine diagnosis. Medical images have different species such as CT, MRI. These different images have their respective application. Medical denoised image is to get the clear information without noise. The image $s(x, y)$ is blurred by a linear operation, and noise $n(x, y)$ is added to form the degraded image $w(x, y)$. This is convolved with the restoration procedure $g(x, y)$ to procedure the restored image $z(x, y)$. The denoising concept is shown in Fig. 1.

4 Ridgelet Transform

Ridgelet transform has ridge functions which avoid wrap-around effect and exact reconstruction and ridge properties. If wavelet is applied to two-dimensional basis, liability is lost which is overcome by the ridgelet transform. It is developed to break the limitation of wavelet transform. During the process of denoising, wavelet transform produces more wavelet coefficients at each scale of decomposition. These coefficients are difficult to be faced and are overcome by ridgelet transform. To characterize texture in efficient way, the transform that capture singularities along lines and edges, namely ridgelet transform, should give better representation [3].

4.1 Finite Ridgelet Transform

In order to overcome the weakness of wavelet in higher dimensions, Candes pioneered ridgelet transform for representation of linear singularities in images [4]. After the finite radon transform (FRAT) mapping line singularities in image domain to point singularities in FRAT domain, a collection of orthogonal transforms are applied to FRAT coefficients column-wise to accomplish the FRIT.

4.2 Radon Transform

The ridgelet transform can be represented in terms of radon transform [5]. The radon transform of an image $f(x1, x2)$ is defined in the given Eq. (3)

$$R_f(\theta, t) = \int f(x_1, x_2)\, \delta(x_1 \cos\theta + x_2 \sin\theta\, t)\, \mathrm{d}x_1\, \mathrm{d}x_2 \tag{3}$$

where

f	Collection of line integrals
(θ, t)	$\in [0, 2\pi]$
δ	Dirac function
$x_1 \cos\theta + x_2 \sin\theta$	Constant

The continuous ridgelet transform (CRT) is simply the application of a mono-dimensional wavelet to the slices of the radon transform. The ridgelet coefficients CRT f(a.b.θ) of an object f are given by analysis of the radon transform as given in Eq. (4)

$$\mathrm{CRT}_f(a, b, \theta) = \psi(c, \delta)\, R_f(\theta, t)\ \mathrm{d}t \tag{4}$$

where $\psi_{(a,\ b)}$—mono-dimensional wavelet

As a consequence, wavelets are very effective in representing objects with isolated point singularities, while ridgelets are very effective in representing objects with singularities along straight lines. In radon domain, when 1D Fourier transform is applied instead of a wavelet, 2D Fourier transform is obtained. Compute the one-dimensional inverse fast Fourier transform (IFFT) on each line, that is, for each value of the angular parameter. Numerical results show that the FRIT is more effective than the wavelet transform in approximating and denoising images with straight edges.

4.3 Discrete Ridgelet Transform

To complete the ridgelet transform, apply a one-dimensional wavelet transform along the radial variable in random space in the form of flowchart. The DRT of an image of size $n \times n$ is an image of size $2n \times 2n$, introducing a redundancy factor equal to 4. To overcome the weakness of wavelets in higher dimensions, a new system of representations is named ridgelet, which deal effectively with line singularities in 2D [6]. This idea is to map a line singularity using the radon transform. Then, the wavelet transform can be used to effectively handle the point singularity in the radon domain. Their initial proposal was intended for functions defined in the continuous space. For practical application, the development of discrete versions of the ridgelet transform that leads to algorithmic implementation is a challenging problem. Due to the radial nature of ridgelets, straight forward

implementations based on discretization in polar coordinates results in transforms that would be either redundant or cannot be perfectly reconstructed.

5 Denoising by Hard Thresholding Theorem

Suppose that one is given noisy data of the form described in the Eq. (5)

$$I(x,y) = I(x,y) + \sigma Z(x,y) \tag{5}$$

where $Z(x, y)$—unit variance and zero-mean Gaussian noise

Denoising a way to recover $I(x, y)$ from the noisy image $I(x, y)$ as proper as possible.

5.1 Wavelet-Based Ridgelet Transform (WBRT)

The ridgelet transform has been shown to be very effective for representing global lines in an image. Ridgelet analysis may be constructed as a wavelet analysis in the radon domain. The 2D radon transform of an object is the collection of line integral indexed by $(\theta, t) \in [0, 2\pi]\ x\ R$ which is given by

$$R_f(\theta,t) = \int R^2 f(x1,x2)d(x_1\cos\theta + x_2\sin\theta - t)dx_1dx_2 \tag{6}$$

where δ is the Dirac distribution.

Then, the ridgelet transform is precisely the application of a 1D wavelet transform to the slices of the radon transform where the angular variable θ is constant and it is varying. For each scale $s > 0$, position $t \in R$, and angle $\theta \in [0, 2\pi]$, the 2D ridgelet function ψ, s, t, θ is defined from a 1D wavelet function ψ as

$$\Psi, s, t, \theta(x1,x2) = s^{-1/2}\psi((x_1\cos\theta + x_2\sin\theta - t)/s) \tag{7}$$

A ridgelet constant along the lines $x_1 \cos\theta + x_2 \sin\theta = \text{const}$.

Transverse to a ridge is wavelet. Thus, the basic strategy for calculating the continuous ridgelet transform is first to compute the radon transform $R_f(t, \theta)$ and second to apply a 1D wavelet transform to the $R_f(t, \theta)$ slices. Different digital ridgelet transforms can be derived depending on the choice of both the radon algorithm and the wavelet decomposition.

5.2 Algorithm

Thresholding is a simple nonlinear technique [7] which operates on one wavelet coefficient at a time. The two types of thresholding are hard thresholding and soft thresholding. The hard thresholding technique is defined in Eq. (8).

$$\mathrm{THR}_H\left(d_j(k), \delta\right) = \begin{cases} 0, & \text{if } d_j(k) \leq \delta \\ d_j(k), & \text{if } d_j(k) > \delta \end{cases} \tag{8}$$

Here, the hard thresholding algorithm is used for an image with the different noise. The soft thresholding is defined in Eq. (9).

$$\mathrm{THS}_S\left(d_j(k), \delta\right) = \begin{cases} 0, & \text{if } d_j(k) \leq \delta \\ \mathrm{sgn}\left(d_j(k)\right)\left|(d_j(k) - \delta\right|, & \text{if } d_j(k) > \delta \end{cases} \tag{9}$$

5.3 Procedure

1. Apply multiscale wavelet transform to the noisy image and get the scaling coefficient and multiscale wavelet coefficients.
2. Choose the threshold and apply thresholding to the multiscale wavelet coefficients (leave the scaling coefficient alone).
3. Reconstruct the scaling coefficient and the multiscaling ridgelet coefficients are thresholded and get the denoised image.

Table 1 Performance of ridgelet transform for medical images

Images type of noise	PSNR (dB)			
	CT-skull	CT-vertebrae	MRI-brain	MRI-knee
Poisson	25.0973	26.1986	21.2944	21.6660
Gaussian	19.3426	22.4657	19.5749	20.1945
Speckle	20.2624	25.5629	20.4856	20.1331
Salt and pepper	18.2443	19.5534	18.3394	18.7073

Table 2 Performance of wavelet-based ridgelet transform for medical images

Images type of noise	PSNR (dB)			
	CT-skull	MRI-brain	MRI-heart	MRI-knee
Poisson	26.381	24.3553	21.0097	23.7466
Gaussian	25.0322	23.2751	19.8274	22.8904
Speckle	25.0615	23.8023	20.2304	22.7624
Salt and pepper	24.3473	22.3692	18.4556	21.9208

With the help of hard threshold algorithm in MATLAB tool, the Peak signal-tonoise ratio (PSNR) values are tabulated in Table 1 for ridgelet transform and Table 2 for wavelet-based ridgelet transform. Some of the sample simulation results for wavelet-based ridgelet transform are shown in Figs. 2, 3.

6 Simulation Results

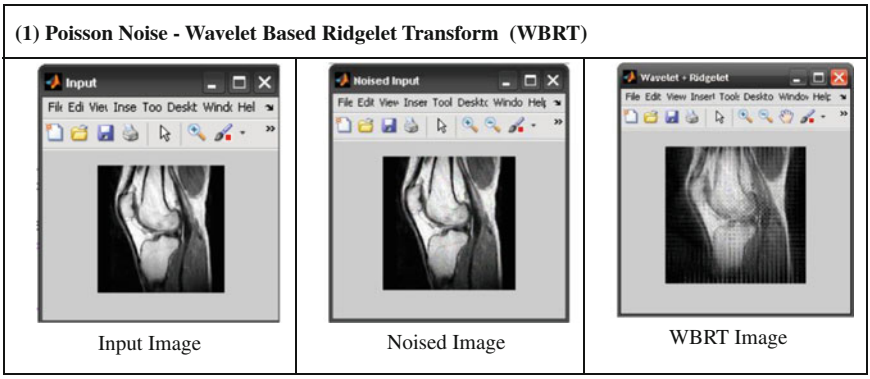


Fig. 2 MRI knee image

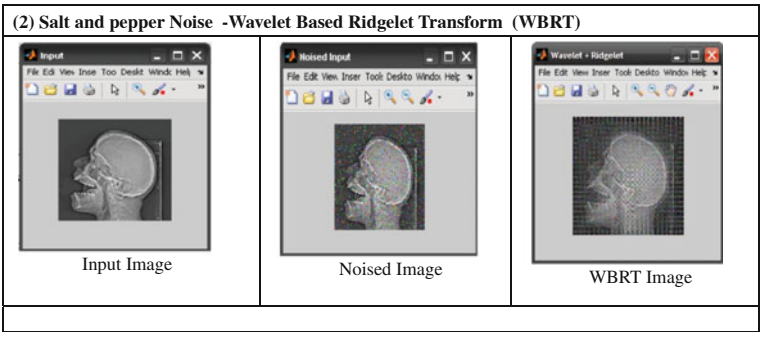


Fig. 3 CT –skull-image

7 Conclusion

Thus, the proposed algorithm has very good denoising effect along the straight edges of the image. The noise should be predicted in order to get useful information for images but also in the case of medical images. Experimental results indicate that the finite ridgelet transform offers an efficient representation for images that are smooth away from line discontinuities or straight edges. In the case of ridgelet transform, the proposed algorithm uses simple and sophisticated technique which has effective denoising effects for medical images than to general images, especially along the straight edges of the image. In addition to that, it restores sparsity by reducing redundancy across scale. When wavelet coefficient are taken and are applied with ridgelet transform using hard thresholding algorithm, Images have better denoising effects and have high Peak signal-tonoise ratio (PSNR) value than ordinary ridgelet transform and performance shown in terms of PSNR.

References

1. Ruikas SD, Doye DD (2011) Wavelet based image denoising technique (IJACSA). Int J Adv Comput Sci Appl 2(3)
2. Freiman M, Werman M, Joskowicz L (2011) A curve let based passion-specific prior for accurate multi model brain image rigid registration. Med Image Anal 15
3. Candes EJ (1999) Ridgelet and representation of mutilated sobolev functions. SIAM Math Anal
4. Candes EJ, Schumaker LL (1999) Ridgelet and their derivatives: representation of images with edges curves and surfaces. Vanderbilt University Press, Nashville
5. Goal P, Sinha HP, Sing H (2011) Ultra sound image denoising using multiscale ridgelet transform with hard and neigh coefficient thresholding. Int J Comput Commun Technol, Vol 2.
6. Do MN, Vetterli M (2003) The finite ridgelet transform for image representation. IEEE Trans Image Process 12(1):16–28
7. Xian-hua J, De-heng H (2011) An improved algorithm noise elimination to image threshold. Comput Teach Dept Anyang Norm Univ 11(9)

Real-time Retrieval System for Heritage Images

Sumit Mishra, Jit Mukherjee, Prasenjit Mondal, Shashaank M. Aswatha and Jayanta Mukherjee

Abstract This paper presents a real-time retrieval system of similar images in a large database. The similarity in images is determined by feature matching technique. The Speeded-Up Robust Features (SURF) are computed for all the images in database (pre-computed) and the query image. Along with SURF, the color information of the images is also used for obtaining an efficient similarity among the images. Principal component analysis (PCA) has been carried out to enhance the efficiency of the system, in terms of time and space, which is followed by SR-tree-based multidimensional indexing of the pre-computed image features. The proposed system is experimented on the distributed and centralized computing environments. Experimental results show the performance of the proposed system in the distributed environment in real-time image retrieval process to be a satisfactory format.

Keywords Image retrieval · SURF · Principal component analysis · SR-tree · Distributed computing

S. Mishra (✉) · J. Mukherjee · P. Mondal · S. M. Aswatha · J. Mukherjee
Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India
e-mail: sumitmishra97@gmail.com

J. Mukherjee
e-mail: jitm@sit.iitkgp.ernet.in

P. Mondal
e-mail: prasenjitm@cse.iitkgp.ernet.in

S. M. Aswatha
e-mail: shashaankama@cse.iitkgp.ernet.in

J. Mukherjee
e-mail: jay@cse.iitkgp.ernet.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_26,

1 Introduction

Real-time image retrieval in a large database is a challenging area of research. Such retrieval systems use image features, to determine similar images in the system, for the retrieval process. Many existing efficient retrieval systems use the local image features, like Scale Invariant Feature Transform (SIFT) [1], Affine SIFT (ASIFT) [2], Oriented FAST and Rotated BRIEF (ORB) [3], Binary Robust Invariant Scalable Keypoints (BRISK) [4], to compute the key points in the images. One of the well-established key-point detection (and description) techniques for image matching is Speeded-Up Robust Features (SURF) [5]. Although SIFT and ASIFT perform better than the SURF in capturing local features, unlike SURF, these techniques are too much time-consuming to be used in real-time systems.

The main difficulty found in matching two images, in real time, with varying spatial resolutions is that it takes a lot of time for computing the similarity among them. This motivated us to design an image retrieval system that works almost in real time, while retaining the robustness and reliability of an efficient similarity matching using SURF technique. The real-time performance of our system is attributed to the existing techniques of dimensionality reduction and efficient multidimensional indexing. Also, it is better suited for matching images of fixed structures such as archaeological structures. Similarity matching using SURF is achieved by obtaining a set of key points extracted for each of the images being compared. From each key point, a fixed multidimensional feature comprising SURF descriptor and color descriptor is computed. It is a computation-intensive and time-consuming task to index millions of such high-dimensional descriptors for retrieval task. Therefore, the multidimensional features are compressed by applying principle component analysis (PCA) [6] technique on them. The resulting feature set is then indexed using the multidimensional indexing structure, SR-tree (spherical–rectangle tree) [7], which enables us to perform efficient range query over the tree.

The performance of the proposed system has been tested on supervised and unsupervised clusters of images. In supervised technique, indexers are computed for each of the clusters by applying the PCA technique independently. For unsupervised scheme, indexers are generated on all the clusters together. It is observed that the performance of the proposed technique on supervised clusters is better than the unsupervised ones. Supervised scheme is also helpful in determining the probable image clusters for a given image. The performance of the proposed scheme has also been implemented in the distributed environment.

2 Methodology

The overall methodology of the proposed system is discussed as follows. At first, the image clusters are formed by supervised clustering algorithm. Each of the clusters represents either the images of a common concept or images of a common

object. For each cluster, an indexer is built using multidimensional structure SR-tree that contains the multidimensional feature vectors for each of the key points detected. These feature vectors comprise SURF and color features, whose dimensions are reduced by PCA. The indexers are generated in three steps as discussed below. The above-mentioned steps are depicted in Fig. 1.

2.1 Image Feature Computation

The feature key points should be invariant to scale and rotation, to achieve a robust matching of images. In this paper, SURF is used to detect and describe the feature key points. The success of SURF features to effectively provide a high degree of repeatability and distinctiveness has been proved [8]. With the advantage of integral image representation, the SURF feature extraction procedure is very fast and robust. Figure 2 shows an example of heritage image with its SURF feature representation. SURF key point is represented as a 64-dimensional vector. Along with SURF, the color information of the images is also used for obtaining an efficient and robust similarity among the images. Considering 64-dimension SURF and a 4-dimension color feature, each image feature vector is represented as a 68-dimensional vector.

SURF Key-Point Detection and Description: The local maxima of the Hessian matrix determinant are selected as an interest point. Their localization is achieved by fast approximation of Hessian matrix. A local response maxima process is carried over several octaves of Gaussian scale space to obtain scale invariance of the features. The non-maxima suppression is applied in $3 \times 3 \times 3$ neighborhood of the interest point in scale space [9]. The maxima of the Hessian matrix determinant, which signify the scales of the interest points, are interpolated in scale and image space. Around each interest point, a circular neighborhood of radius six times the scale of the detected interest point is considered. By using a sliding window (of size 60°), the Haar wavelet responses within this circular neighborhood are obtained. The Haar wavelet responses are efficiently obtained by integral image computations. A direction, which maximizes the sum of the Haar wavelet responses, is assigned as the dominant orientation of that key point.

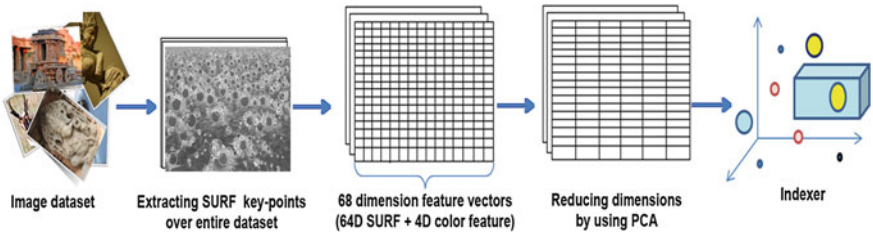


Fig. 1 Computation of the indexer for a set of images

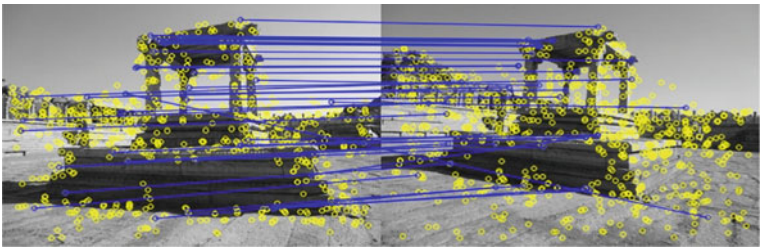


Fig. 2 Matching between two images using the key points detected by SURF. The key points (*yellow circles*) and their correspondences are shown

For each detected key point, a square neighborhood, which is aligned along the dominant orientation of that key point, of size 20 times the scale of those key points, is considered. To make the descriptor also scale invariant, this square neighborhood region is split into 4×4 sub-regions. Four values from the first-order Haar wavelet responses are extracted in a 5×5 grid in the block for each of these sub-regions [9].

2.2 Computation of Color Feature

The key points detected using SURF are invariant to scale and rotation. To make the key points invariant to different illuminations, color information of each key point is considered as an additional feature. For that, a neighborhood region around each key point is considered. In LUV color space, the mean and standard deviation of U and V color channels in the neighborhood are computed. The use of UV color channels ensures invariance to different illuminations and also minimizes the effect of shadows. So, after concatenating the color feature of four dimensions (the mean and standard deviations) with the 64 dimensions of the SURF feature, we get 68-dimension feature representation for each key point.

2.3 Principal Component Analysis

For each key point, the feature descriptor has 68 dimensions. Indexing feature vectors of 68 dimensions for a large number of images is computationally intensive and time-consuming task. For example, for an image with spatial resolution 400–500, approximately 400–1,000 key points are detected, in the described technique. To reduce the computational complexity, before indexing, PCA technique has been used to significantly reduce the number of dimensions of the feature vectors from the feature set. The PCA is an orthogonal linear transformation technique that captures most of the information of feature samples and

represents them with fewer dimensions (with some loss of information). The real-world data are considered to be nonlinear. Therefore, the loss of information is inevitable if feature representation for all the images in the database is compressed using PCA technique.

In each image cluster, the dimensionality of the feature vectors is reduced from 68 to 9 using PCA. From the scree plot in Fig. 3, it can be observed that 9 principal components are sufficient enough to represent the information content with affordable data loss and considerable data reduction. The use of PCA is justified by a huge reduction in data size. The redundant information (not much significant information), which caused an intensive computation complexity and consumed system memory, is reduced by the application of PCA.

2.4 Indexing Technique

For context-based image retrieval on large database, a multidimensional index structure is needed. One-dimensional index structure uses a single search key and retrieves results that match the search key value which is inefficient for spatial data, high-dimensional point data. Supporting partial match queries, range queries, and nearest neighbor queries, multidimensional index structure outperforms one-dimensional index structure in this scenario. SR-tree indexing structure is used here, which uses a region by intersection of bounding sphere and bounding rectangle (which outperforms SS-tree, KDB tree, etc.) for indexing key points [7]. In our approach, the feature sets of the training images are indexed by SR-tree. The index file is stored in the secondary storage, which is used in the server program to load the feature set of the training images. It is also observed that the retrieved image set contained more of similar images when the indexing of dimension-reduced features

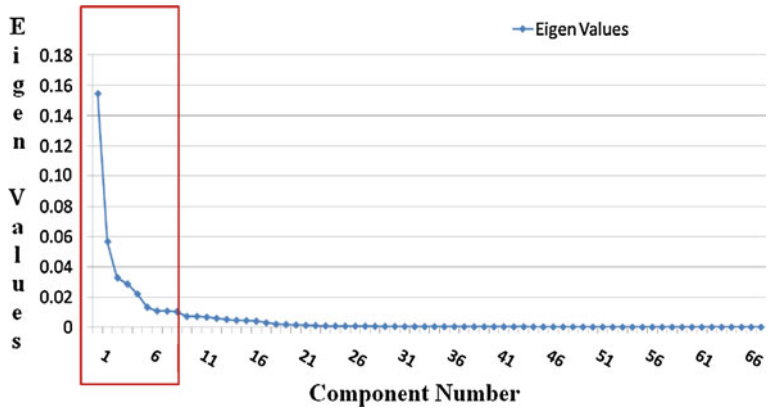


Fig. 3 A plot for dimensionality reduction using PCA, showing the first nine principal components capturing most of the information

is carried on clusters of images, rather than on the entire dataset. This accuracy has been achieved with multiple clusters and partitioning of the index files. Also, the advantage of using multiple index files is being able to use the system in distributed environment, as discussed in the following section.

2.5 Distributed Image Retrieval System

It has been observed that the searching a query image using the conventional image retrieval system (centralized system) takes a lot of time. In this environment, the image database and the retrieval system reside in a single computer. Hence, when some malfunction arises in the centralized system, the retrieval system suffers. To make the proposed system scalable and fault tolerant, the proposed system is implemented in a distributed computing environment. A distributed system consists of multiple autonomous computer systems that communicate with each other through a communication network. Distributed computing refers to using the distributed systems to solve a computational problem.

In our work, we have used four individual computers that are connected to each other, thereby forming a distributed system. For proper functioning of this system, at first, the whole image dataset is randomly divided into four groups. The indexers are computed for each of these four groups using the techniques mentioned in Sect. 2.4. The index contains the key-point locations along with their respective descriptors. Each indexer is logically partitioned into different clusters based on the similarity of the images (supervised). For the distributed computing, each indexer is assigned to one of the computer system in the distributed system. Whenever a new image query request arrives, the task is distributed to four different computers. The feature key points are determined for the input query image. The indexer in each computer is searched for the key points of the query image for similarity matching. The similar procedure is followed for all the other systems. Finally, the individual results from each of the indexers residing in the different systems are gathered, and all the corresponding retrieved images for the given query image are shown. The process of searching for a query image in the distributed computing environment with four computers is shown in Fig. 4.

3 Results and Discussion

The proposed system has been implemented in C language with Open CV application programming interface on a standard PC (Intel Core 2 Quad CPU, 3.0 GHz, 4 GB RAM) in Linux environment. The system has been tested on a database of 40,000 images. From all the images, we have collectively obtained over 30 million key points. We experiment our system on a number of sample query images. It has been observed that most of the retrieved images closely match

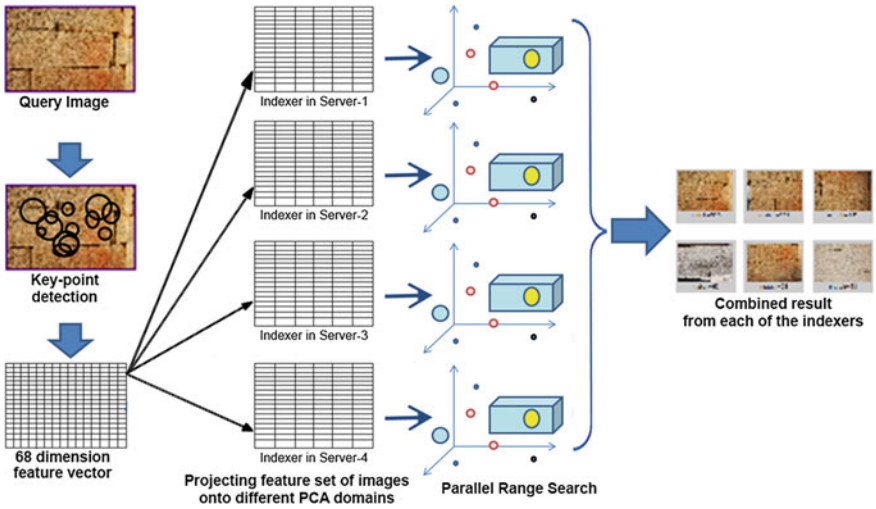


Fig. 4 The process of searching a query image in a distributed environment

with the query image. Also, the proposed system has been tested on a variety of images using the distributed architecture. Figure 5 depicts two query images and their corresponding images as retrieved from the database of images. The image retrieval time and the number of key points matched with the query image are also shown in the image. It has been noticed that the average image retrieval time for a query image with spatial resolution 500 × 500 is reduced to approximately 1 s. Thus, we have achieved a speed gain of nearly 12 times with distributed computation. We have compared the image retrieval systems using distributed system with the centralized system. The systems are tested with a number of query images

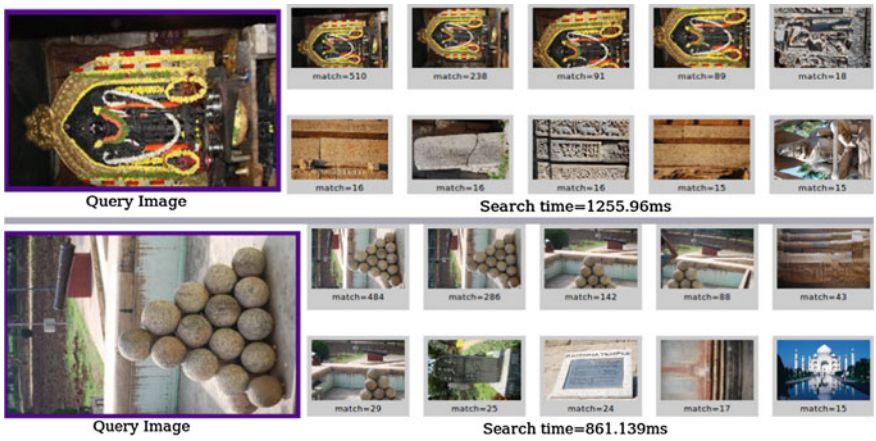


Fig. 5 Sample query images and the retrieved images as obtained using the proposed technique

Table 1 Distributed versus centralized image retrieval system

Spatial resolution	Average image retrieval time (s)	
	Centralized system	Distributed system
250 × 250	0.9	0.5
500 × 500	3	1.5
1000 × 1000	8	4
3000 × 3000	46	24

For images with different spatial resolutions, the average time (in seconds) taken to retrieve images is shown

of varying spatial resolutions as tabulated in Table 1, which shows the average time taken to retrieve similar images with varying spatial resolutions. It has been observed that the time consumed for retrieving an image in a distributed system is much lesser, as compared to the centralized system. The average precision and recall for the distributed image retrieval system with multiple image clusters are computed as 99 and 70 %, respectively (found to be 82 and 31 % for centralized system, respectively).

4 Conclusion

In this paper, we have introduced a fast, real-time image retrieval system. The features used for image matching consist of the key-point descriptors obtained from the SURF and color information in the neighborhood of the key points. The combined feature provides invariance to rotation, scale, and illumination. A reduction in redundant memory and computation complexity has been achieved by principal component analysis, before feature indexing. A better accuracy of retrieved results has been achieved with multiple clusters of images by partitioning the index files, which enabled the use of the system in a distributed environment. The system is experimented for its real-time performance, along with satisfactory retrieval of similar image sets.

Acknowledgment This research is partially supported by the Department of Science and Technology, Government of India, through sanction no. NRDMS/11/1586/2009. This retrieval system is deployed online and can be accessed for public use at the Web site http://imedix.iitkgp.ernet.in/SMARAK/.

References

1. Lowe D (2004) Distinctive image features from scale-invariant key points. Int J Comput Vision 60(2):91–110
2. Yu G, Morel JM (2009) ASIFT: an algorithm for fully affine invariant comparison. SIAM J Imaging Sci 2(2):438–469

3. Rublee E, Rabaud V, Konolige K, Bradski G (2011) ORB: an efficient alternative to SIFT or SURF. Comput Vision (ICCV), IEEE Int Conf, pp 2564–2571
4. Leutenegger S, Chli M, Siegwart RY (2011) BRISK: binary robust invariant scalable keypoints. In: Proceedings of the IEEE international conference on computer vision (ICCV)
5. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (SURF). Comput Vis Image underst 110(3):346–359
6. Ke Y, Sukthankar R (2004) PCA-SIFT: a more distinctive representation for local image descriptors. Comput Vision Pattern Recogn, IEEE Comput Soc Conf, 2:506–513
7. Katayama N, Satoh S (1997) The SR-tree: an index structure for high-dimensional nearest neighbor queries. ACM SIGMOD Int Conf Manage Data 26:369–380
8. Kalantidis Y, Tolias G, Spyrou E, Mylonas P, Avrithis Y, Kollias S (2011) VIRaL: visual image retrieval and localization. J Multimedia Tools Appl 51(2):555–592
9. Herbert B, Andreas E, Tinne T, Gool LV (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110:346–359

PDMWSN: A Probable Approach to Detect Malicious Node in Wireless Sensor Network

N. Ambika and G. T. Raju

Abstract Wireless sensor networks are prone to attacks as they are deployed in unattended and insecure environments that are prone to different types of attacks. Hence, security algorithms have to be implemented. This paper generates group-based encryption keys and provides a validated report to the base station which guarantees the integrity and authentication of the data that are being transmitted. This paper secures the nodes from Sybil attack by 21 %, sinkhole attack by 25 %, and wormhole attack by 30 %.

Keywords Security · Malicious nodes · Wireless sensor network · Group key authentication

1 Introduction

Wireless sensor networks are a special type of ad hoc networks, which do not have fixed infrastructure and whose topology is liable to change due to node failures. Wireless sensors are devised primarily to sense the environment, record readings, process it, and forward the processed data to the base station. These devices are application specific, allowing application-dependent node platforms, communication protocols, data aggregation, and in-network processing and decision making. Hence, the number of nodes varies depending on the application in use. Wireless sensor networks are used in a variety of applications like home automation [1, 2],

N. Ambika (✉)
Dayananda Sagar College of Engineering, Bharathiar University, Coimbatore, India
e-mail: ambika.nagaraj76@gmail.com

G. T. Raju
Department of Computer Science, RNSIT, Bangalore, India
e-mail: drgtraju_rnsit@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_27,

environmental monitoring, industrial monitoring [3] and control, military security management [4], traffic control, health care [5].

Security [6, 7] becomes one of the crucial issues in any sort of application, where there is necessary to shield private data. As wireless sensor network is application specific, security level to any application varies. References [8–10] propose a security scheme for leach. In this paper, count of packets is analyzed at every level of transmission. Section 2 narrates the previous work similar to the present work. Section 3 describes the work in detail. Section 4 portrays the implemented work. Section 5 concludes by briefing the work.

2 Related Work

The framework presented in [11] works in a distributed environment to detect intrusions by collaborating with the neighboring nodes. It works in two modes: online prevention allows safeguarding from those abnormal nodes that are already declared as malicious, while offline detection finds those nodes that are being compromised by an adversary during the next epoch of time. Whenever a node senses any message, online prevention validates the packet whether it is coming from legitimate neighboring node or not. If it is received from the normal node, the sensor node performs normal task; otherwise, it discards it immediately.

Li et al. [12] proposed a group-based intrusion detection scheme that meets the efficient, lightweight, and flexible requirements by partitioning the sensor networks into many groups in which the sensors in each group are physically close to each other and are equipped with the same sensing capability. The algorithm takes into consideration multiple attributes of the sensor nodes simultaneously to detect malicious attackers precisely. The algorithm can decrease the false alarm rate and increase the detection accuracy rate compared with existing intrusion detection schemes while lowering the power consumption.

The algorithm presented in [13] uses a packet count–based sliding window approach. At every node, only the last N packets received from each neighbor are used to calculate the statistics for that neighbor and each arriving packet is compared against these values. We call N the main packet buffer length. If the packet conforms to the statistics of the neighbor, it is accepted as normal and is used for new calculations. The oldest packet's values are removed from the list. We record the arrival time and receive power of each incoming packet.

To monitor receive powers for anomalous packets, minimum and maximum values of packet's receive power are updated with each regular packet reception. An anomalous packet is a packet whose receive power is below the minimum or above the maximum of receive power currently kept in the main packet buffer of length N.

3 PDMWSN Model

3.1 System Architecture

A heterogeneous network is being considered, possessing different kinds of nodes necessary to perform different functions. The detector and the mobile agent deployed are Waspmote. Waspmote with a microcontroller Atmel ATmega1281, transceiver ZigBee/802.15.4/DigiMesh/RF, 2.4 GHz/868/900 MHz, program–data memory of 8 K SRAM, external memory of 128 K flash ROM, 4 K EEPROM, and 2 GB SD card is utilized. Waspmote integrates with a GPS that can deliver accurate position and time of information. Waspmote is either in sleeping mode or in hibernating mode. Following the predefined interval programmed by the user, Waspmote wakes up, reads from the sensor, implements the wireless communication, and retracts to the sleep mode. Each device is powered with rechargeable batteries and solar panel, making the system completely autonomous. The main characteristic of Waspmote is its low power consumption.

Other nodes in the network are Tinynode 584. This is a compact sensor node platform based on MSP430 microcontroller and Xemics XE1205 multichannel radio transceiver. It possesses 10 k of SRAM, 48 k of flash ROM, and 512 k of serial storage. One lithium battery and 2 alkaline batteries together change the mode to the required working state.

3.2 Embedding the Keys Inside the Sensors and its Deployment in the Network

The base station is responsible to generate the keys and deploy the keys into the ROM of the nodes. The keys stored in these nodes are mutually exclusive from each other. The nodes contain one or more wasp nodes, and the other nodes making up the cluster are Tinynode. The nodes are deployed in the form of group consisting of waspmote acting as a detector of the cluster. It is responsible for spying the activities of the cluster members and updating the activities of the cluster to the base station. Other node in the cluster does the sensing, partial processing, encrypting the data being sensed, and transmitting it to the cluster head. All the nodes inside the cluster have a buffer size set. The uncompromised nodes keep the level of buffer size the same and transmit the data of same length every time.

3.3 Formation of Cluster

The nodes deployed in the network broadcast HELLO message. The nodes within the transmission range respond to the message. The nodes within the transmission range form the cluster. These nodes generate a group session key for each session and utilize the key to encrypt the sensed data. After the formation of the group, the cluster head is chosen depending on the energy it possesses and its trust level portrayed by the detector of the respective cluster.

3.4 Task of the Detector

The detector of the node has to keep a track of activity of every node inside the cluster. The detector sends encrypted periodical reading to the base station along with the location it is residing. The detector contains the table which stores the node id, number of packets it has dispatched during the time interval, and the total number of packets it received/forwarded to the next cluster.

3.5 Task of the Mobile Agent

Mobile agent is stored with a predefined path by the base station. The number of mobile agents depends on the application and the size of the network that it has to do survey. The mobile agent does no interaction with any of the nodes in the network. The task of the mobile agent is to observe the activity of the nodes in the network. The time interval is noted along with the time the node transmits the data. These data are encrypted by the mobile agent by using its own encryption key, dispatched to the base station along with its location information.

3.6 Evaluation by the Base Station

The base station compares the result obtained by the mobile agent and the detector, and cross-verification is done against the data obtained by the cluster. The base station provides the path mobile agent has to follow from the location data obtained by the detectors of the cluster.

4 Simulated Results

In [13], the false alarm rate gets reduced as the time intensifies. The false alarm rate eliminates the uncompromised nodes from the network, minimizing density of nodes. While the proposed work takes care of this issue and keeps the false alarm rate to a minimum of 0.01, this paper also takes care of Sybil attack, sinkhole attack, and wormhole attack which gives an added advantage.

4.1 Energy Consumption

The simulation is done using ns-2. Energy consumption becomes a vital issue in sensor networks as the batteries in them cannot be recharged. Hence, the energy has to be consumed judiciously. But the integrity of the data also has to be taken care. Hence, additional security measures have to be applied. The proposed paper consumes 27.3 % extra energy compared to a network without any security and 18.9 % more energy compared to [13]. But this energy is utilized by the mobile agent (if independent mobile agent can be deployed, the energy of the static nodes will not decline). Compared to [11], the work consumes 5.5 % less energy (energy consumed by mobile agent not considered Fig. 1).

4.2 Sybil Attack

Sybil attack [14, 15] is where the identities of the nodes are duplicated. The node will be able to have multiple identities. The nodes that are compromised and are

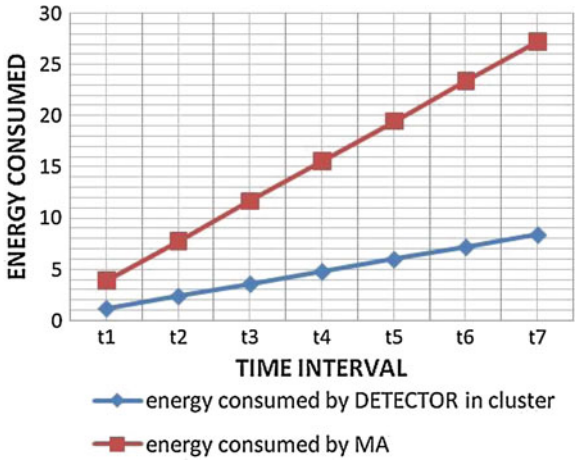


Fig. 1 Total energy consumed by PDMWSN

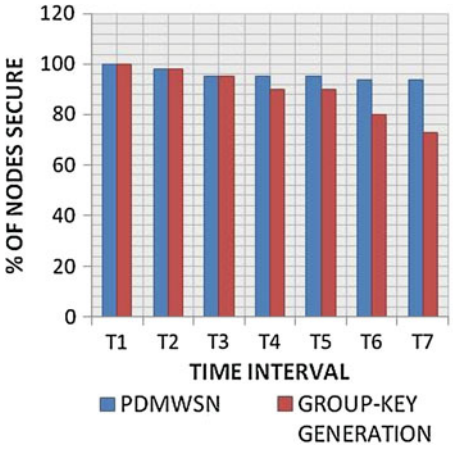


Fig. 2 Effect of Sybil attack

under the control of adversaries try to duplicate identification. These duplicated nodes gain trust of the neighboring nodes. In this paper, the node keys are mutually exclusive and the detector deployed along with other nodes in the group provides an exact location where they are located. Hence, nodes are protected from Sybil attack by 21 % compared to network that generates group key without any security (Fig. 2).

4.3 Sinkhole Attack

Sinkhole attacks [16, 17] usually attract the surrounding nodes either by spoofing or by replaying the advertisements to an extremely high quality route to the base station. These attacks usually attract the surrounding nodes either by spoofing or by replaying the advertisements for an extremely high quality route to a base station. The continuous surveillance of the detector and random observation of the mobile agent detect the compromised nodes. Hence, nodes are secured by 25 % compared to group key generation without any security against sinkhole attack (Fig. 3).

4.4 Wormhole Attack

A wormhole attack [18, 19] is where the attacker records the packet at one location in the network, tunnels them to another location, and retransmits it into the network. Surveillance of detector and mobile agent keeps track of the number of incoming and outgoing packets. Hence, the compromised node will not be able to hide from these nodes for a longer time. This paper secures the nodes by 30 % compared to network without security from wormhole attack (Fig. 4).

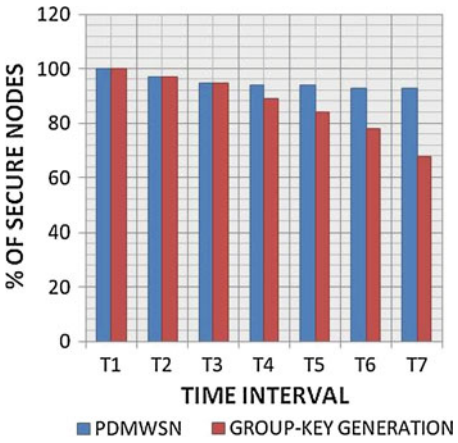


Fig. 3 Effect of Sinkhole attack

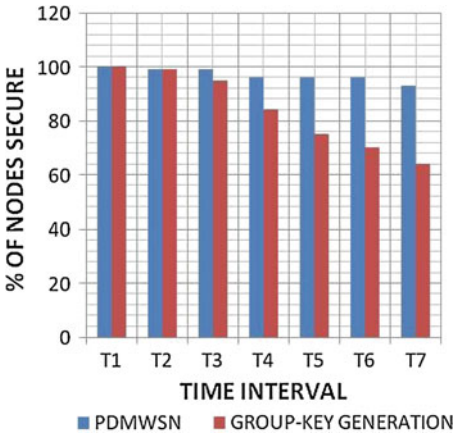


Fig. 4 Effect of Wormhole attack

5 Conclusions

Integrity and authentication are the primary necessities of sensor network. This paper secures the nodes from sinkhole attack, Sybil attack, and wormhole attack by using group key encryption and continuous monitoring by detector of the cluster and random monitoring by a mobile agent.

References

1. Hussain S, Schaffner S, Moseychuck D (2009) Applications of wireless sensor networks and RFID in a smart home environment, Seventh annual communication networks and services research conference, 153–157, doi:10.1109/CNSR.2009.32
2. Zatout Y, Campo E, Llibre J-F (2009) WSN-HM: energy-efficient wireless sensor network for home monitoring "5th International conference on intelligent sensors, sensor networks and information processing (ISSNIP):367–372, doi:10.1109/ISSNIP.2009.5416757
3. Gungor VC, Hancke GP (2009) Industrial wireless sensor networks: challenges, design principles, and technical approaches. IEEE Trans Ind Electron, 56(10):4258–4265, doi:10.1109/TIE.2009.2015754
4. Lee SH, Lee S, Song H, Lee HS (2009) Wireless sensor network design for tactical military applications: remote large-scale environments, IEEE military communications conference, 1–7, doi:10.1109/MILCOM.2009.5379900
5. Akyildiz E, Su W, Sanbrrsubrarnaoiam Y, Cuyirci E (2002) A survey OD SCOSUT network. IEEE Commmicariom Mngeirre, 40(8):102–114
6. Karlof C, Wagner D (2002) Secure routing in wireless sensor networks: attacks and countermeasures, Elsevier's ad hoc network journal, special issue on sensor network applications and protocols
7. Kambourakis G, Klaoudatou E, Gritzalis S (2007) Securing medical sensor environments: the code blue framework case. In: Wagner R, Tjoa AM et al. (eds) 2nd international conference on availability, reliability, and security—1st international symposium on frontiers in availability, reliability and security, pp 637–643, Vienna, Austria, IEEE Computer Society Press
8. Su CC, Chang KM, Kuo YH, Horng MF (2005) The new intrusion prevention and detection approaches for clustering-based sensor networks. In: IEEE wireless communications and networking conference, pp 1927–1932
9. Krontiris I, Dimitriou T (2007) Towards intrusion detection in wireless sensor networks. In: 13th European wireless conference, Paris
10. Roman R, Zhou J, Lopez J (2006) Applying intrusion detection systems to wireless sensor networks. In: 3rd IEEE consumer communications and networking conference, pp 640–644
11. Farooqi AH, Khan FA, Wang J, Lee S (2012) A novel intrusion detection framework for wireless sensor networks, pers Ubiquit Comput, doi:10.1007/s00779-012-0529-y
12. Li G, He J, Fu Y (2008) A group based intrusion detection scheme in wireless sensor networks. In: The 3rd international conference on grid and pervasive computing workshop, pp 286–291
13. Wang Q, Wang S, Meng Z (2009) Applying an intrusion detection algorithm to wireless sensor networks, Second international workshop on knowledge discovery and data mining
14. Newsome J, Shi E, Song D, Perrig A (2004) Detecting sybil attacks in wireless sensor networks using neighboring information. In: Proceedings of the 3rd international symposium on Information processing in sensor networks, ACM, New York, doi:10.1145/984622.984660
15. Misra S, Myneni S (2010) On identifying power control performing sybil nodes in wireless sensor networks using RSSI, IEEE global telecommunications conference, doi:10.1109/GLOCOM.2010.5683728
16. Teng L, Zhang Y (2010) SeRA: a secure routing algorithm against sinkhole attacks for mobile wireless sensor networks, Second international conference on computer modeling and simulation, pp 79–82, doi:10.1109/ICCMS.2010.95
17. Chen C, Song M, Hsieh G (2010) Intrusion detection of sinkhole attacks in large scale wireless sensor networks, IEEE international conference on wireless communications, Networking and information security (WCNIS), pp 711–716, doi:10.1109/WCINS.2010.5541872
18. Hu YC, Perrig A, Johnson DB (2006) Wormhole attacks in wireless networks. IEEE J Selected Areas Commun, 24(2):370–380, doi:10.1109/JSAC.2005.861394
19. Zhao Z, Wei B, Dong X, Yao L, Gao F (2010) Detecting wormhole attacks in wireless sensor networks with statistical analysis. WASE Int Conf Inf Eng (ICIE), pp 251–254, doi:10.1109/ICIE.2010.66

Study of Partial Discharge in Pressboards

S. Soumya, Eranna, S. Vasudevmurthy and B. V. Sumangala

Abstract Study of partial discharge plays an important role in the aging and the rupturing process of solid and mixed insulation systems. The PD analysis is a predictive test which indicates insulation degradation in advance, which may lead to the failure of the system, and it is performed under normal operating conditions. Pressboards are solid insulating materials made by compressing layers of paper together and drying them. They are used in bushings and in transformers as insulating barriers or supporting materials. The main advantage of this material is that it provides good mechanical support even at higher temperatures up to 120 °C. It has a high degree of flexibility, simultaneously possessing high tensile and compression strength. This paper presents the experimental study conducted on the prepared sample using pressboards as the solid insulating sheet and the effects of different parameters such as pressure, temperature, curing period, and number of conducting layers on partial discharge in the insulating sheet.

Keywords Partial discharge · Dielectric · Pressboard · Insulation

1 Introduction

Partial discharge (PD) is localized ionization within electrical insulation that is caused by high electrical field. PD occurs in voids in solid insulation (paper, polymer, etc.), in gas bubbles in liquid insulation or around an electrode in a gas.

S. Soumya (✉) · Eranna · S. Vasudevmurthy · B. V. Sumangala
New Horizon College of Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India
e-mail: soumyas1210@gmail.com

Eranna
e-mail: eranna_1964@yahoo.com

B. V. Sumangala
e-mail: sumangala_bv@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_28,

PD is the intermediate step before breakdown occurs [1]. After initiation, the PD propagates and develops into electrical trees. The review of literature suggests that increasing temperature results in an increase in geometry of voids and the pressure within them [2, 3]. The appearance of space charge would distort the electric field distribution in the void [4]. Once begun, PD causes progressive deterioration of insulating materials, ultimately leading to electrical breakdown.

Pressboards are made by compressing layers of paper together and drying them. They are used in bushings and in transformers as insulating barriers or supporting materials. The application depends upon the thickness and density of paper used. The electrical strength of pressboard is higher than that of resins or porcelain. The main advantage of this material is that it provides good mechanical support at higher temperatures up to 120 °C.

2 Partial Discharge Detection and Measurement

The different causes of breakdown in insulation lead to PD activity which includes design and manufacturing defects, incorrect installation of components, faulty materials, and mechanical damage. PD can occur in a gaseous, liquid, or solid insulating medium. PD activity can initiate under normal working conditions in high-voltage equipment where the insulation condition has deteriorated with age and/or has been aged prematurely by thermal overstressing.

PD dissipates energy, generally in the form of heat, but sometimes as sound and light as well, like the hissing and dim glowing from the overhead line insulators. Heat energy dissipation may cause thermal degradation of the insulation, although the level is generally low. The effects of PD within high-voltage cables and equipment can be very serious, ultimately leading to complete failure. The cumulative effect of PD is called treeing. Damage is caused by the energy dissipated by high-energy electrons or ions, ultraviolet light from the discharges, ozone attacking the void walls, and cracking as the chemical breakdown processes liberate gases at high pressure.

PD can be prevented through careful design and material selection. In critical high-voltage equipment, the integrity of the insulation is confirmed using PD detection equipment during the manufacturing stage as well as periodically through the equipment's useful life. PD prevention and detection are essential to ensure reliability and operation of high-voltage equipment.

PD testing is particularly important where HV plant has a high criticality. Quite often, PD activities initiate in high-voltage-condition equipment, where the condition of the insulation has deteriorated in time, and this is why PD testing is highly recommended. The PD analysis is a predictive test that indicates insulation degradation of the failure of the system in advance, and it is performed under normal operating conditions. Furthermore, PD testing does not affect the equipment under test, and most of the times the occurrence of PD can be identified within the test object, thus allowing the solving of the problem.

It is essential to detect and monitor partial as it represents a diagnostic approach which uses measurements of PD in order to evaluate the integrity of the equipment. The PD measurements can be taken continuously or intermittently, and they can be detected online or offline. The results of this analysis are used to predict the electrical equipment which needs to be replaced or repaired.

3 Insulation

An insulator, also called a dielectric, is a material that resists the flow of electric current. An insulating material has atoms with tightly bonded valence electrons. These materials are used in parts of electrical equipment, intended to support or separate electrical conductors without passing current through themselves. Some materials such as glass, paper, or Teflon are very good electrical insulators. A much larger class of materials, for example rubber-like polymers and most plastics, are still "good enough" to insulate electrical wiring and cables.

The strength of insulation is mainly determined by the impregnating material used. By impregnating with liquid or molten (dissolved) materials, an attempt is made to exclude the ionizing gases and to protect the carefully dried fiber materials against the effect of the atmospheric moisture. Complete removal of the air trapped between the pores (capillaries) is therefore a precondition for the impregnating agent to permeate effectively. This produces the ideal mixed dielectric of oil–cellulose, so that the prolific absorption of oil by the pressboard will also continue to ensure the prominence of this material in the construction of oil-filled transformers and capacitors.

3.1 Breakdown in Solid Dielectrics

The dielectric strength of solid materials is affected by many factors, viz., ambient temperature, humidity, duration of test, impurities, or structural defects whether AC, DC, or impulse voltages are being used, pressure applied to these electrodes, etc. The mechanism of breakdown in solids is again less understood. However, it is convenient to divide the timescale of voltage application into regions in which different mechanisms operate [5]. The various mechanisms are intrinsic breakdown, electromechanical breakdown, breakdown due to treeing and tracking, thermal breakdown, electrochemical breakdown.

3.1.1 Breakdown Due to Treeing and Tracking

The treeing phenomenon can be observed in all dielectric wherever non-uniform fields prevail. If we have two electrodes separated by an insulating material and the

assembly is placed in an outdoor environment, some contaminants in the form of moisture or dust particles get deposited on the surface of the insulation and leakage current starts between the electrodes through the contaminants, say moisture. The current heats the moisture and causes breaks in the moisture films. These small films then act as electrodes, and sparks are drawn between the films. The sparks cause carbonization of the insulation and lead to the formation of permanent carbon tracks on the surface of insulations known as tracking [6]. For tracking to occur, the insulating material must contain organic substances. For this reason, for outdoor equipment, tracking severely limits the use of insulation having organic substances.

3.2 Sample Pressboards

Pressboards are made by compressing layers of paper together and drying them. They are used in bushings and in transformers as insulating barriers or supporting materials. The application of these pressboards depends upon the thickness and density of paper used. For high-frequency capacitors and cables, usually low density paper (0.8 gm/cm^3) is used, whereas medium density paper is used for power capacitors and high density papers are used in and energy storage capacitors. The electrical strength of pressboard is higher than that of resins or porcelain. The main advantage of this material is that it provides good mechanical support even at higher temperatures up to 120 °C. Since the need for high-voltage electrical transformers, there has been a need for an insulating material that could withstand the high electrical and physical stresses experienced around a core and windings. Because of its exceptional properties as a pure fiber material, containing no bonding agents, pressboard is used in a wide range of applications in transformers, capacitors as well as electrical machinery and equipment. It has a high degree of flexibility, simultaneously possessing high tensile and compression strength. Since pressboard is a natural product, it poses no problems for the recycling or disposal of stamping residues, edge cuts, and much more.

4 Experimental Model

As tests on practical systems are not always feasible due to cost and time factors, model system tests are gaining importance. In the present work, a model insulation system consisting of pressboard with uniform field electrode (parallel plate) arrangement has been tested for its behavior under PDs.

Details of samples: Four samples using pressboard as the insulating system are prepared, and the details are as follows:

- Pressboard thickness—0.1 mm
- Aluminum foil thickness—11 micron (food grade)
- Dimension of square sample—13 × 13 cm
- Diameter of the aluminum foil—10 cm
- Number of layers of pressboard—13
- Adhesive used—fevicol SR

SAMPLE 1 13 layers of pressboard—with no aluminum foil in the interspaces.

SAMPLE 2 13 layers of pressboard—with aluminum foil in the 6th interspace.

SAMPLE 3 13 layers of pressboard—with aluminum foils in 3 interspaces: aluminum foil at 1st interspace, at 12th interspace, and one at the center.

SAMPLE 4 13 layers of pressboard—with aluminum foils in 6 interspaces: aluminum foils in the alternate interspaces.

SAMPLE 5 13 layers of pressboard—with 12 interspaces and 12 aluminum foils: aluminum foils in all interspaces.

The pressboard layers are bound together with the help of a thin layer of the adhesive mentioned above, with the respective sample type as given. Then, the samples containing 13 layers each are compressed with the help of the universal testing machine (UTM) as shown in Fig. 3, with pressure of 300 and 500 kg force (kgf).

5 Experimental Setup

The experimental set up consists of the following:

5.1 Shielded Room

The shielded room is of the dimension 10-by-10 ft. The room is plywood laminated with aluminum sheets, and the total area of the aluminum is 600 sq. Ft. The thickness of the aluminum sheets determines the ability to suppress the airborne interference.

5.2 Partial Discharge-Free Transformer

Rating of the transformer: 230/100 kV, 5kVA. The transformer is filled with 45 l of good grade oil.

5.3 Partial Discharge-Free Capacitor

The tubular structure which is made up of aluminum is provided to minimize the PD interference that occurs due to the electric field at the sharp objects and corners. The edges are also rounded for the same reason.

5.4 Partial Discharge Meter

The PD meter serves to verify very-high-frequency PD pulses appearing in the high-voltage dielectrics. The PD meter is a valuable instrument for the timely detection of these disturbing phenomena and to discard faulty apparatus under test. The display of PD intensity is in picocoulomb (pC).This equipment has a built-in oscilloscope, which is used to display power frequency test voltage with the PD pulses in an elliptical shape.

5.5 Electrode Setup

The electrodes are made up of brass. The diameter of the electrodes is 10 cm. They are placed 2 ft above the ground. The position of the electrodes is adjustable as required. It is ensured that the electrodes are perfectly in level. Overall in the setup, a niche is carved for every bolt, to minimize the effect of the electric field around these sharp objects which interferes in the PD measurement and gives errors in readings. All the lights and fans are switched off before the measurement procedure because the small sparks that they produce also interfere in the PD measurement (Fig. 1).

Fig. 1 PD measurement room

6 Experimental Details

The samples are prepared as explained in Sect. 4 and are placed between the parallel plate electrodes as shown in Fig. 2. The PD occurrence in various samples is measured by varying:

- pressure
- temperature
- curing time
- Number of conducting layers.

The pressure is varied up to 500 kgf as shown in Fig. 3 using UTM.

Fig. 2 Electrode

Fig. 3 Sample compression

7 Results and Discussion

7.1 Effect of Pressure (Sample 3)

From the Fig. 4, it can be inferred that as the pressure applied on the sample is increased from 300 to 500 kgf, the PD is found to decrease. This can be explained from the fact that the more the compression pressure on the sample is, the lesser will be the voids created, and hence, this reduces the PD effect.

7.2 Effect of Temperature

From the Fig. 5a and b, it can be inferred that as the temperature increases from 40° to 50°, the value of PD is found to be diminished by a considerable value. On the other hand, when the temperature was increased beyond 50, the PD value increased to a great extent. This can be explained from the fact that at very high temperature, the sample loses its ability to stay bonded with the layers, which therefore increases the conductivity between the layers.

7.3 Effect of Curing Time (Sample 3)

From the Fig. 6, it can be inferred that as the curing time increases from ¼ to 1 h, the value of PD is found to increase. This can be explained from the fact that when the sample is subjected to high voltages for prolonged durations, there will be

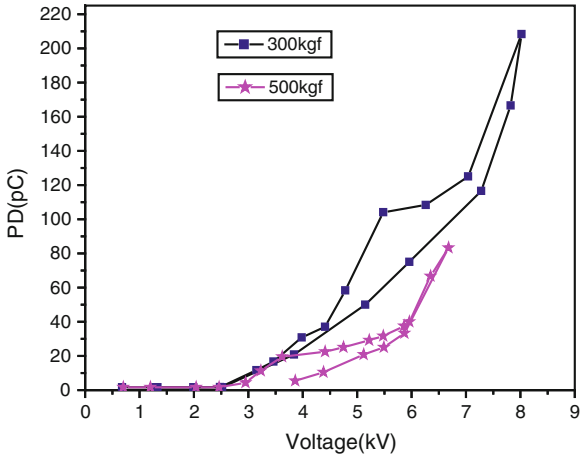


Fig. 4 PD versus V (pressure)

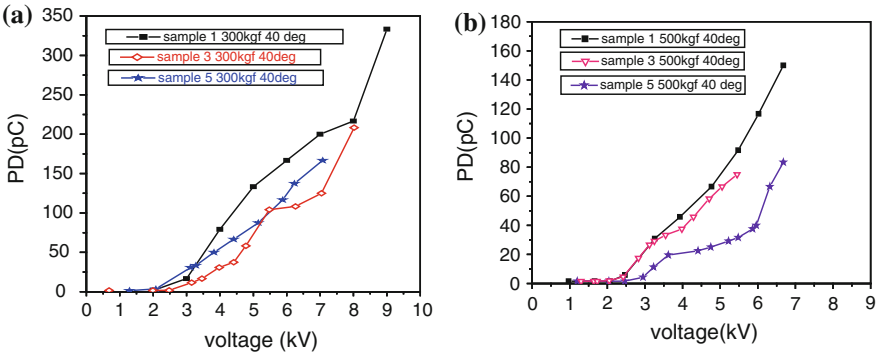


Fig. 5 **a** PD versus V (500 kgf, 40°). **b** PD versus V (300 kgf, 40°)

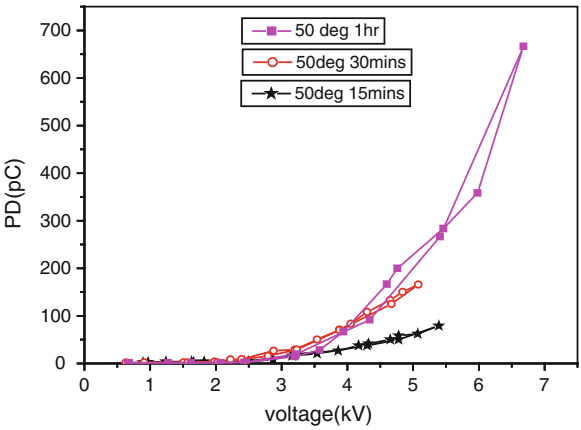


Fig. 6 PD versus V (different curing time)

increase in the trees formed, which increases the conductivity of the sample and further leads to charring. Thus, the PD is found to increase with the increase in curing time.

7.4 Effect of Number of Conducting Layers (300 kgf)

From the Fig. 7, it can be inferred that as the number of conducting layers increases from no aluminum foil (conducting) to 12 foils, the value of PD is found to decrease. This can be explained from the fact that as the number of foils increases, the electric field between the electrode plates becomes more uniform, due to which the PD is found to decrease.

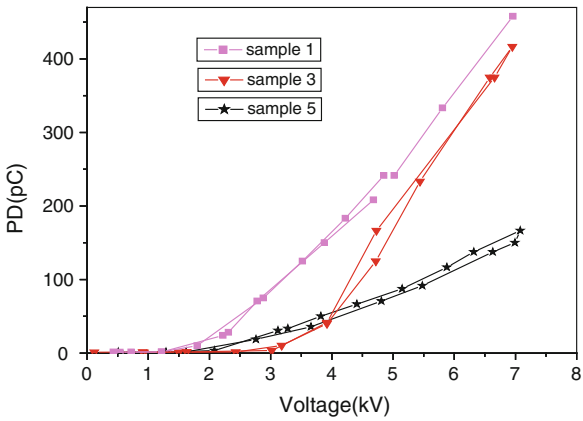


Fig. 7 PD versus V (different number of conducting layers)

8 Conclusion

Experiments were conducted on different samples for different parametric variations, and the following are the conclusions with reference to the Figs. 4, 5, 6, 7:

- **PRESSURE:** It is observed that the compression of samples plays a major role in reducing the PD as the air gap between the layers is reduced.
- **TEMPERATURE:** Increase in temperature leads to expansion of the sample layers, thus increasing the PD.
- **NUMBER OF LAYERS:** By increasing the number of conducting layers, a more uniform field can be achieved, resulting in less PD.

On the whole, the characteristic behavior of the selected pressboards for various arrangements of pressure, temperature, conducting layers, etc. is highly nonlinear. For higher voltages, the PD variation appears to be linear which calls for further detailed investigations.

References

1. Fruth B Partial discharge mechanisms in solid insulation systems
2. Karady GG, Ro RR (1995) Effect of temperature on the partial discharge initiation voltage of capacitors. IEEE Trans Dielectr Electri Insul 2(3):499–502
3. Devins JC (1984) The physics of partial discharges in solid dielectrics. IEEE Trans Electri Insul 18–21:613–617, 19:475–495
4. Blackbum TR, Liu Z, Phung BT, Morrow R (1999) Investigation of partial discharge mechanisms in a void under ac electric stress. High voltage engineering symposium, 22–27 Aug 1999, Conference Publication No. 467, O IEE, 1999
5. Zeller HR (1987) Breakdown and pre-breakdown phenomena in solid dielectrics. IEEE Trans Electri Insul El-22(2):115–122 April
6. Cavallini A, Conti M, Contin A, Montanari GC (2003) Advanced PD inference in on-field measurements. Part 2: identification of defects in solid insulation systems. IEEE Trans Dielectri Electri Insul 10(3):528:538 June

3DS^3: A Framework for 3D Static Scene Streaming

V. Vani, R. Pradeep Kumar and S. Mohan

Abstract The development of networked virtual environment (NVE) over the Internet has become more predominant with a very rich *3D interactive computer graphics*. Therefore, the effective rendering of 3D scenes with necessary interactions in the client machine has become the need of the hour. The well-established techniques for video and audio streaming which are already in place have become the primary inspiration to focus our research toward 3D static scene streaming (3DS^3). In this approach, the part of a demanded 3D data alone is brought to the client based on the interactions subject to the bandwidth and memory capacity. The main objective is to reduce the *streaming and rendering time*, that is, overall turnaround time (TT). The well-proven concept of *multi-threading* and *serialization* processes along with the *predictive model* is adapted in this work to achieve the stated objective. The experimental results and analysis carried out affirm that there is a remarkable reduction in rendering latency and streaming time with the serialization and multi-threading predictive framework.

Keywords Predictive modeling · 3D scene streaming · 3D rendering · Serialization · Multi-threading · Networked virtual environment

V. Vani (✉)
Department of Information Technology, Dr.N.G.P.IT, Affiliated to Anna University, Coimbatore, India
e-mail: vasudevan.vani@gmail.com

R. Pradeep Kumar
Department of Computer Science Engineering, Adithya IT, Affiliated to Anna University, Coimbatore, India
e-mail: rpk.ind@gmail.com

S. Mohan
Department of Computer Science Engineering, Dr.N.G.P. IT, Affiliated to Anna University, Coimbatore, India
e-mail: s.mohan77@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_29,

1 Introduction

Over the last decade, interactive 3D computer graphics (in other words 3D real-time graphics) [1] has gained attention over the network because of its emotional connection with the mankind, particularly on multi-user platform. Interactive 3D computer graphics is needed to simulate and study some of the systems in real time and understand the working principles of it. It is a replica of real world, and whatever is not possible to do in real world, like walking through the hazardous place virtually, real-time simulation of the flight crash, etc., is made possible in interactive 3D computer graphics. Interactive 3D computer graphics over the Web has given rise to the development of 3D virtual environments (VE) [2, 3], which is available on the Web and accessed by multiple users at the same time. While transmitting the 3D scenes across the Web, the streaming [4] approach is promising to deliver the content as per the client request. Though many developments are evident in 3D Web, a few attempts are made toward 3D streaming over the network as it gives initial response quickly to the client without complete download of the 3D data.

In this paper, an attempt is initially made on the server side to transfer a complete 3D static scene (which consists of multiple objects) that is requested by the clients. Here, several dynamic threads are created based on the number of objects in the scene. These dynamic threads are used to handle the objects in the scene, and as soon as the data pertaining to an object are read, it is sent to the client through TCP/IP protocol suite. On the client side, the transmitted data are received and rendered with necessary interactions. Secondly, instead of transferring entire 3D static scene to the client and then rendering it on the client machine, only visible portion of the scene is determined and packed to the client using dynamic threading. On the client side, transmitted packet is unpacked and then rendered with necessary interactions. Based on the interaction (user move), next packet with only the faces and vertices that are not referred already in the scene is streamed to the client, buffered, and rendered. In this approach, 3D streamed data are buffered instead of storing it into the file and render it. Finally, a predictive agent module is added to the server side in order to predict in advance the next client move. Based on the offline analysis carried out on the collected user profiling, the required number of vertices and faces along with the initial static scene is streamed to the client. In all the three approaches, the turnaround time from streaming to rendering is calculated for every static scene selected from the server database and the results are analyzed to exhibit the efficiency of our methods.

2 3D Streaming: An Overview

3D streaming [5–8] is the process of delivering 3D content in real time for the users over a network. 3D streaming is carried out in such a way that the interactivity and visual qualities of content may match as closely as if they were stored

locally. The main resource bottleneck here is usually assumed to be the bandwidth and not rendering or processing power of the clients [9]. To achieve this goal, simplification of the model and transmission of the content based on the user's view are two dominant strategies.

2.1 Importance of 3D Streaming in NVE

Due to the demand for more realistic worlds, networked virtual environments (NVEs) have become larger and more dynamic. Currently, almost all massively multi-user online games (MMOGs) [10, 11] require users to install the entire 3D content used for game operation on their local machines priory, due to the large data volume that easily extends to over hundreds and even 1,000 of megabytes. However, prior installations are undesirable and even unpractical under two likely future scenarios:

Large Volume of Dynamic Data: MMOGs have a few GBs of relatively static content (e.g., World of Warcraft (WoW) [12] is over 5 GB with 11 million monthly subscribers and holds the Guinness World Record for the most popular massively multi-user online role player game of all time). However, as content becomes larger and more dynamic, streaming will save both the installation time and update time. In fact, a prior installation is already unsuitable for the social MMOG Second Life [13, 14], which depends on streaming to deliver over 34 TB of user-created models, textures, and behavior scripts. Also, the virtual globes such as Google Earth [15] and NASA World Wind currently have terabytes of data. Extensions into 3D may only be a matter of time, as shown by initiatives such as X3D Earth [16]. Preinstallation thus is unpractical, given the size of the data and the user base.

In realizing such planet-scale virtual environment [17], large-scale 3D streaming could be the basis for next-generation virtual globe applications.

Large Count of Virtual Environments: As imagined by proponents of Web 3D, the future Internet could very well be three-dimensional where diverse environments exist to various socializing, shopping, and learning experiences. If millions of 3D sites were to exist, prior installations for each one of them would be frustratingly inconvenient and unpractical.

As content becomes easier to create and cheaper to host, more developers, even individuals, are now building virtual worlds (e.g., Second Life hosts 34 terabytes of user-generated contents). The trends toward larger worlds and larger numbers of worlds thus are beginning to show the inadequacy of current VE installation methods.

Even the audio and video data are also large in size, but thanks to streaming delivery, users need not wait for a complete download before using the content. The possibility to stream audio and video has been crucial in their widespread popularity. Similarly, when broadband networks and 3D acceleration cards were adopted by average users, streaming delivery of 3D data [11, 17] began to appear

as viability, raising hopes for the widespread adoption of 3D content streaming for various 3D virtual environments.

3 Proposed Work

The proposed work attempts to build a framework for 3D static scene streaming by adopting the improvised approach which includes predictive-view-dependent streaming by exploiting threading, buffering, packing, and unpacking techniques. The framework is used to conduct the experiment with multiple 3D static scenes, and the experimental results are studied to verify the enhancements achieved.

3.1 3DS^3 Framework

Figure 1 depicts the overview of 3DS^3 framework. The 3D static scene streaming indicates that the scene comprises of multiple 3D objects in predetermined positions and the entire scene remains fixed. In Fig. 1, server package has streaming module and the 3D static scene database. The streaming module in turn comprises of dynamic threads, packing, predicting, and transmitting modules. Based on the client inputs

1. request to the server to stream a specific 3D static scene and
2. the initial eye position (viewpoint).

The server-side program would dynamically create number of threads which are directly proportional to the number of objects in the scene. The threads are created to read the objects from the 3D static scene database which contains

1. geometry information (vertices) [18, 19] and
2. connectivity information (faces). Based on the initial eye position that is being sent by the client, identify only the visible portion of the 3D static scene, pack it, and then transmit it using stream socket [20] to the respective client.

The client package that depends on sever package has rendering module apart from client input module. As specified earlier, client input module sends the updated eye (camera) position based on the user interactions on the rendered scene to the server. The rendering module comprises of receiving and unpacking modules. The receiving and unpack modules receive the transmitted data from the buffer, unpacks it into vertices and faces, and finally renders it to the client display unit.

As soon as the initial view of the 3D scene is streamed and rendered, the client can interact with the 3D scene by using the keys UP, DOWN, LEFT, RIGHT, PgUP, PgDown, PLUS, MINUS to perform rotation operation on specific axis

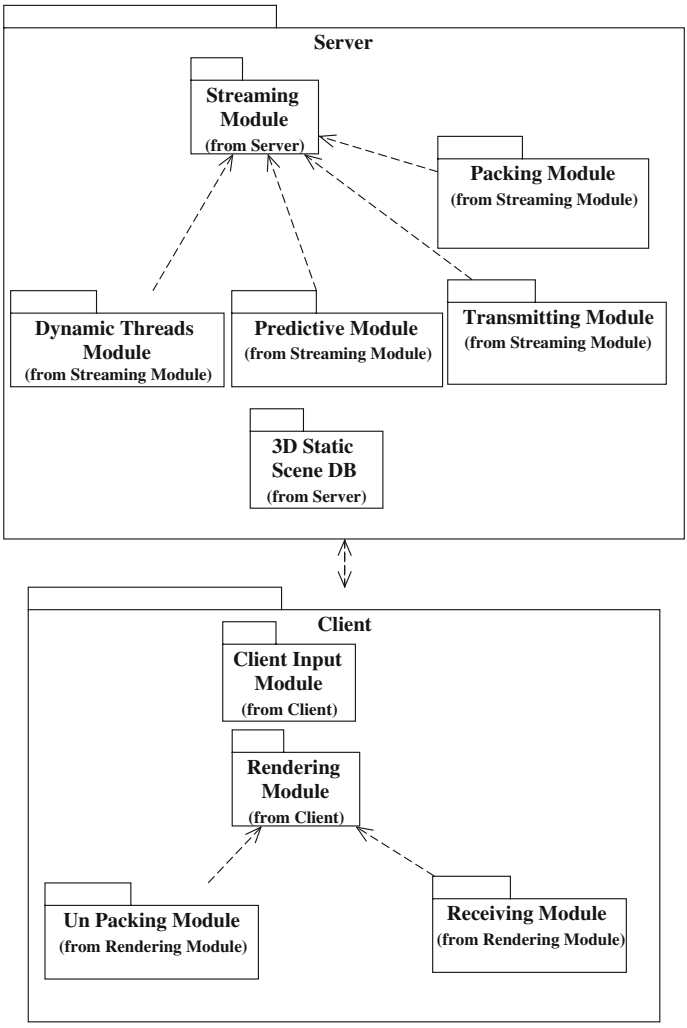


Fig. 1 3DS[3] framework overview

(+Y, −Y, +X, −X, +Z, −Z) and Zoom in /Zoom out, respectively. Based on the interaction, the camera position is altered and it is sent to the server along with the client's keystroke. The server in turn would determine the vertices and faces that need to be transmitted. While transmitting, server would transmit only the vertices and faces which are not transmitted already. This would save remarkable amount of transmission time which ultimately contributes to the reduction in overall turnaround time. Also, with the help of predictive module, it is possible to determine the next user move well in advance. This is used to transmit the vertices and faces that are not demanded along with the demanded vertices and faces. The predictive module is incorporated to reduce the rendering latency.

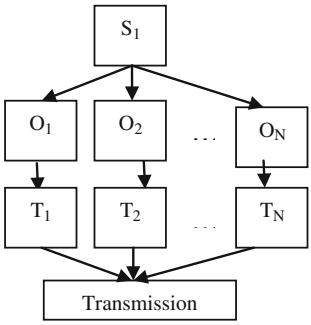


Fig. 2 3DS[3] framework: multi-threading model

3.1.1 Multi-Threading (T)

Assume a 3D static scene (S_1) with 3 objects, namely O_1, O_2, and O_3. In order to stream these objects dynamically, 3 threads are created, namely T_1, T_2, and T_3. Though based on the size of the object and the thread priority, streaming and rendering sequence would vary, and it is ensured to maintain synchronization among threads. This would enable the client to render the objects correctly. The following Fig. 2 depicts the multi-threading model implemented in 3DS[3] framework. The advantage of threading is that the user's waiting time is further reduced by streaming the object whose initial data are computed and packed in the instantiated threads without waiting for all other threads to complete.

3.1.2 Serialization

Serialization (S) refers to packing and unpacking of vertices and faces. Packing happens at the server side where the needed objects in the chosen 3D static scene are packed into the buffer and transmitted to the client with the help of dynamically created threads. Here, only the unreferred vertices and faces are packed and transmitted. The vertices are floating-point values, whereas the faces and indices are integer values represented by the vertex positions. These values are packed into the character buffer which is then transmitted to the client using stream socket.

As soon as the transmitted packets are received by the client, unpacking process begins. This process unpacks the character buffer and stores back the data into the vertices and face vectors. Since only the vertices and faces that are not previously transmitted are brought, reordering of vertices is also carried out before the data are sent to rendering.

3.1.3 Predictive Model

The streaming module sends the client input to predictive model for further process. Initially, the user interactions performed by various users across different 3D

static scenes are collected in order to derive a probable user patterns. With the help of this offline analysis, the predictive model is built along with other modules in streaming. This is attempted to send the predicted vertices and faces in advance before the user does any key press. The prediction leads to reduced rendering latency whenever the demanded faces/vertices are already rendered. This prediction leads to rendering latency reduction.

4 Experimental Results

Experiments are conducted to study the performance of the various enhancements made to come out with 3DS[3] framework. The results are overwhelming, and it affirms the stated objective. The experimental setup was run on an Intel Core 2 Duo CPU P8600 @ 2.4 GHz with 4 GB RAM and ATI Radeon 1-GB graphics card system for both server and client. For experimental purpose, we have constructed the 3D static scenes comprising of multiple objects which are from simple to complex in terms of the vertices and faces. Table 1 and Fig. 3 show the TT required to transmit and render the complete 3D static scene. Initially, the TT is calculated to transmit and send the 3D static scene from the server to client without threading and serialization approaches. The results shown in Fig. 3 reveal that the initial approach takes more time for streaming and rendering. Therefore, the dynamic multi-threading along with packing and unpacking (serialization) approach is attempted for streaming and rendering 3D static scenes in which each scene comprises of multiple objects. The result of second approach proves to be better than the initial one. Table 2 and Fig. 4 highlight the merit of 3D streaming by transmitting only the required data from the client which reduces the rendering latency and ultimately the turnaround time. Table 3 and Fig. 4 show the time required to render the first version of the scenes based on the camera position which are consistently sent by the client to the server during user navigation. Table 3 shows the initial saving achieved while transmitting partial data based on the camera viewpoint. This highlights the importance of streaming instead of

Table 1 TT with and without threading and serialization for complete 3D static scene transmission (V—vertices)

Scene	V	Faces	TT without T and S	TT with T and S	Time saved
Scene 1	10,430	19,866	0.782	0.098	0.684
Scene 2	10,190	19,074	0.341	0.087	0.254
Scene 3	31,933	63,844	2.413	0.232	2.181
Scene 4	17,418	34,824	0.533	0.146	0.387
Scene 5	8,709	17,412	0.292	0.078	0.214
Scene 6	8,709	17,412	0.675	0.077	0.598
Scene 7	5,755	11,128	0.436	0.051	0.385
Scene 8	73,500	147,008	3.644	0.564	3.08

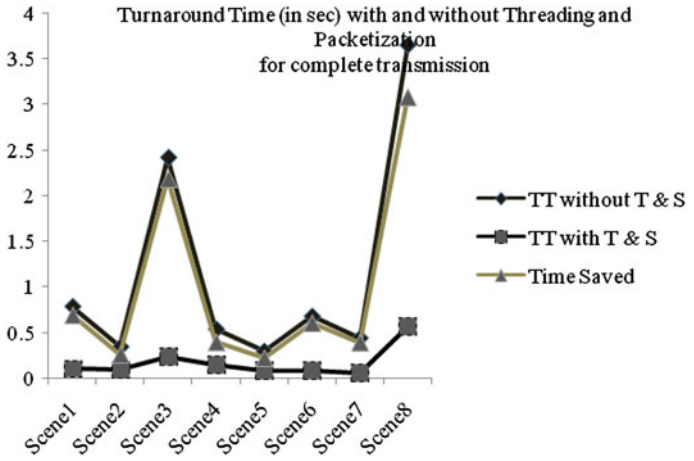


Fig. 3 TT with and without threading and serialization for complete 3D static scene transmission

Table 2 TT with and without threading and serialization for partial 3D static scene transmission (V—vertices)

Scene	V	Faces	TT without T and S	TT with T and S	Time saved
Scene 1	7,533	9,780	0.154	0.037	0.117
Scene 2	7,001	9,391	0.141	0.034	0.107
Scene 3	20,033	30,262	0.936	0.068	0.868
Scene 4	12,214	19,601	0.264	0.046	0.218
Scene 5	6,016	9,600	0.146	0.038	0.108
Scene 6	5,982	9,712	0.258	0.039	0.219
Scene 7	3,358	5,299	0.079	0.037	0.042
Scene 8	43,173	72,978	1.023	0.137	0.886

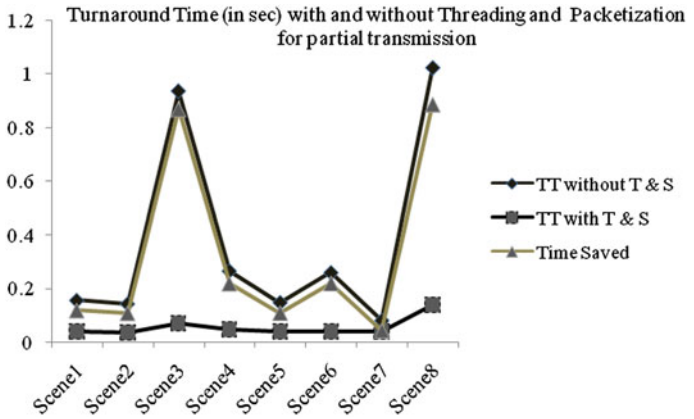


Fig. 4 TT with and without threading and serialization for partial 3D static scene transmission

Table 3 Vertices and faces saved in percentage between complete and partial transmission

Scene	Vertices saved (%)	Faces saved (%)
Scene 1	28	51
Scene 2	31	51
Scene 3	37	53
Scene 4	30	44
Scene 5	31	45
Scene 6	31	44
Scene 7	42	52
Scene 8	41	50

Table 4 Time saved percentage between complete and partial transmission

Scene	Turnaround time with threading and serialization		Time saved (%)
	Complete transmission	Partial transmission	
Scene 1	0.098	0.037	38
Scene 2	0.087	0.034	39
Scene 3	0.232	0.068	29
Scene 4	0.146	0.046	32
Scene 5	0.078	0.038	49
Scene 6	0.077	0.039	51
Scene 7	0.051	0.037	73
Scene 8	0.564	0.137	24

Table 5 Average turnaround time after multiple interactions with threading, serialization, and prediction for streaming

Scene	Accumulated TT	Average turnaround time
Scene 1	0.194	0.0388
Scene 2	0.226	0.0452
Scene 3	0.411	0.0822
Scene 4	0.325	0.065
Scene 5	0.293	0.0586
Scene 6	0.234	0.0468
Scene 7	0.2	0.04
Scene 8	0.556	0.1112

complete file download to the client machine. In partial transmission, the transmitted data from the server to the client are buffered and then rendered. Table 4 shows the time saving achieved between complete and partial transmission. In partial transmission, only the visible parts of the geometry and connectivity information are sent to the client. The experimental results clearly convey the fact that the client could see the initial scene more quickly than before. Finally, Table 5 and Fig. 5 show the average TT achieved after multiple client interactions. Here, the TT after each move is randomly selected, and its average is taken. The

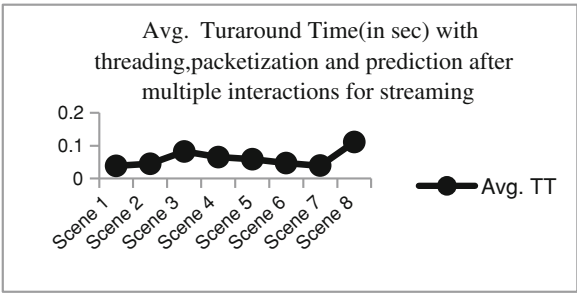


Fig. 5 Average TT with threading, serialization, and prediction after multiple interactions for streaming

experimental results shown are quite encouraging when we compare the TT required for complete 3D static scene transmission with that of average turnaround time. The accumulated TT over multiple interactions is still lesser than the TT required for complete 3D static scene transmission.

5 Conclusion and Future Directions

In this paper, an attempt has been made to build the framework for streaming 3D static scenes. The framework is built by exploiting the multi-threading, packing, and unpacking concepts along with the Internet sockets. Also, the framework has an additional capability of predicting the possible next move of the client. Based on this prediction, possible data are streamed along with the data that are needed immediately. It is demonstrated with experimental results that the rendering latency and streaming time are reduced to a greater extent. The results achieved are quite encouraging, and this framework can further be extended to support 3D dynamic scene streaming and rendering with state management.

References

1. Brutzman D, Macedonia M, Rhyne TM, McCanne S (1997) Internetworked 3D computer graphics: beyond the bottlenecks and roadblocks. ACM SIGCOMM '97
2. Blanc AL., Bunt, J, Petch J, Kwok Y (2005) The virtual learning space: an interactive 3D environment. In: Proceedings of the tenth international conference on 3D web technology, Bangor. Web3D '05. ACM, New York, NY, pp 93–102
3. Cabral M, Zuffo M, Ghirotti S, Belloc O, Nomura L, Nagamura M, Andrade F, Faria R, Ferraz L (2007) An experience using X3D for virtual cultural heritage. In: Proceedings of the twelfth international conference on 3D web technology (Perugia, Italy). Web3D '07. ACM, New York, pp 161–164
4. Deb S, Narayanan PJ (2004) Design of a geometry streaming system. In: Proceedings of ICVGIP, pp 296–301
5. De Paor DG, Whitmeyer SJ (2011) Geological and geophysical modeling on virtual globes using KML, COLLADA, and Javascript. Comput Geosci 37(1):100–110. doi:10.1016/j.cageo.2010.05.003

6. Cheng W, Ooi WT, Mondet S, Grigoras R, Morin G (2007) An analytical model for progressive mesh streaming. In: Proceedings of the 15th international conference on Multimedia MULTIMEDIA '07). ACM, New York, 737–746. doi:10.1145/1291233.1291399
7. Cheng W, Liu D, Ooi WT (2009) Peer-assisted view-dependent progressive mesh streaming. In: Proceedings of the 17th ACM international conference on Multimedia (MM '09). ACM, New York, pp 441–450. doi:10.1145/1631272.1631333
8. Tang Z, Guo X, Prabhakaran B (2011) Receiver-based loss tolerance method for 3D progressive streaming. Multimedia Tools Appl 51(2):779–799. doi:10.1007/s11042-010-0634-8
9. Teler E, Lischinski D (2001) Streaming of complex 3d scenes for remote walkthroughs. EUROGRAPHICS 20(3)
10. Knutsson B, Lu H, Xu W, Hopkins B (2004) Peer-to-peer support for massively multiplayer games. In INFOCOM, p 96–107
11. Hu SY, Jiang JR, Chen BY (2010) Peer-to-peer 3D streaming. IEEE Internet Comput 14(2):54–61
12. Wyman MT (2011) World of warcraft, making great games. Focal Press, Boston, pp 23–38, ISBN 9780240812854. doi: 10.1016/B978-0-240-81285-4.10002-9
13. Rosedale P, Ondrejka C (2003) Enabling player-created online worlds with grid computing and streaming. Gamasutra Resource Guide
14. Jin SAA (2011) It feels right. therefore, i feel present and enjoy: the effects of regulatory fit and the mediating roles of social presence and self presence in avatar-based 3D virtual environments. Presence 20(2):105–116 by the Massachusetts Institute of Technology
15. Wu H, He Z, Gong J (2010) A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. Comput, Environ Urban Syst 34(4):291–298, ISSN 0198-9715. doi: 10.1016/j.compenvurbsys.2009.12.001
16. Anslow C, Marshall S, Noble J (2006) X3D Earth in the software visualization pipeline. X3D Earth requirements' 06 naval postgraduate school, Monterey Nov, pp 14–15
17. Hu SY (2006) A case for 3d streaming on peer-to-peer networks. In: Proceedings of the eleventh international conference on 3D web technology (Web3D), pp 57–63
18. Botsch M, Pauly M, Rossl C, Bischoff S, Kobbelt L (2006) Geometric modeling based on triangle meshes. In ACM SIGGRAPH Courses (SIGGRAPH '06). ACM, New York Article 1. doi:10.1145/1185657.1185839
19. Botsch M, Pauly M, Kobbelt L, Alliez P, Lévy B, Bischoff S, Rössl C (2007) (2007) Geometric modeling based on polygonal meshes. In ACM SIGGRAPH courses (SIGGRAPH '07). ACM, New York Article 1. doi:10.1145/1281500.1281640
20. Brian Beej Jorgensen Hall,"Beej's guide to Network Programming using internet sockets", Version 3.0.14, chapters 1, 2, 3,5 and 7)

An Efficient Approach to Detect Clone Attacks in Wireless Sensor Networks Using Random Walk Protocol

S. Sheela and R. C. Manjunath

Abstract Wireless sensor networks (WSNs) deployed in hostile environments are exposed to clone attacks. In such attack, an introducer compromises a few nodes, replicates them, and inserts arbitrary number of replicas into the network. Consequently, an introducer can carry out many internal attacks. Previous solutions on detecting clone attacks have several drawbacks. Existing protocol, Randomized Multicast, is non-deterministic and fully distributed (NDFD) and fulfills the requirements, but it has very high communication overhead. Then, based on random walk, two new NDFD-based protocols are proposed, RAndom WaLk (RAWL) and Table-Assisted RAndom WaLk (TRAWL), which fulfill the requirements while having only moderate communication and memory overheads. The random walk strategy outperforms previous strategies because it distributes a core step, the witness selection, to every passed node of random walks and then the adversary cannot easily find out the critical witness nodes. Trace the walk and calculate the probability of detection. Implement TRAWL which is based on RAWL. A table is used to record random walks. Plot the graph to analyze the algorithm.

Keywords Wireless sensor networks · Computer network security · Clone attacks · Node replication · Random walk

S. Sheela (✉)
Department of TCE, DBIT, Mumbai, India
e-mail: sheelatcdbit@gmail.com

R. C. Manjunath
Department of ECE, AIT, Gurgaon, India
e-mail: manjunathrc@acharya.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_30,

1 Introduction

Wireless sensor network (WSN) is a wireless network consisting of spatially distributed autonomous devices using sensors to cooperatively monitor physical or environmental conditions. These sensor devices are placed in a designated area. The function of these sensors is to collect and transfer the data to the base station (BS) directly or through other sensors. The sensors coordinate with each other to establish a network so that they can monitor the designated geographical area. The collected data will be reported continuously to BS through the network. WSN is a large network of resource-constrained sensor nodes with multiple preset functions, such as sensing and processing, to fulfill different application objectives. However, WSNs are susceptible to many types of attacks because they are deployed in an open and unprotected environment. The sensors send the collected data within their communication range to the BS without failure. But the BS may not receive all data sent by these nodes due to transmission problems or node malfunctions. Therefore, it is our duty to identify and report the malfunctioning nodes.

Clone attack [3] (also called node replication attack) is a severe attack in WSNs. In this attack, an adversary captures only a few of nodes, replicates them, and then deploys arbitrary number of replicas throughout the network. The capture of nodes is plausible [4, 5] because sensor nodes are usually unprotected by physical shielding due to cost considerations [3] and are often left unattended after deployment. If replicas are not detected, then the network will be vulnerable to a large class of internal attacks [6]. Various approaches have been proposed to detect clone attacks [3]. In this paper, firstly, it is shown that in order to avoid the drawbacks of existing approaches, replica-detection protocols must be non-deterministic and fully distributed (NDFD) and fulfill requirements on witness selection. The first protocol, RAndom WaLk (RAWL), starts several random walks randomly in the network for each node a and then selects the past nodes as the witness nodes of node a. Analysis shows that $O(\sqrt{n} \log n)$ walk steps are sufficient to detect clone attacks with high probability. The second protocol, Table-Assisted RAndom WaLk (TRAWL), is based on RAWL and adds a trace table at each node to reduce memory cost.

2 Methods Used in this Paper

2.1 Problem Definition

Wireless sensor networks (WSNs) deployed in hostile environments are vulnerable to clone attacks. In clone attack, an adversary compromises a few nodes, replicates them, and inserts arbitrary number of replicas into the network. Consequently, the adversary can carry out many internal attacks. Adversary can inject false data or suppress legitimate data, spread blame for abnormal behavior and revoke

legitimate nodes using aggregated voting, and monitor communication. It can capture only few nodes at a time and then offline attack to extract node's secrets and transfers secrets to generic nodes to deploy clones. Clones will know everything the compromised node knew. The capture of attack is plausible because sensor nodes are usually unprotected by physical shielding due to cost considerations and are often left unattended after deployment. If replicas are not detected, the network will be vulnerable to a large class of internal attacks.

2.2 Existing System

It roughly classifies existing approaches into two categories: central control involved and fully distributed. Approaches in the first category usually need a central control, for example, the BS, in some critical steps. Approaches in the second category do not need any central control. Central Control Involved Approaches: Finding conflicting location claims is a common method to detect replicas. In this approach, each node sends a list of its neighbors and their location claims to the BS [3]. Since there is a common assumption that the replicas cannot have different IDs, the BS can easily find and revoke nodes with the same ID but with different location claims. Fully Distributed Approaches: All existing approaches in this category detect replicas based on finding conflicting location claims. An example approach is node-to-network broadcasting. When executing the protocol, each node broadcasts its location claim to the whole network and all nodes store the location claims of their neighbors only. Then, if a node receives two conflicting claims of same node ID, it can revoke that node by flooding the network with the two claims. The disadvantages of existing protocols such as LSM, RED, and Randomized Multicast are prone to central control and deterministic and vulnerable to simple witness compromising attacks, and also adversary could learn the critical witness nodes to start smart attacks, which are solved by the proposed protocols RAWL and TRAWL.

2.3 Proposed System

Two algorithms have been stated: RAWL-based detection and TRAWL-based detection.

RAWL-based Detection: At a high level, RAWL works with following steps in each execution: Each node broadcasts a signed location claim. Each of the node's neighbors probabilistically forwards the claim to some randomly selected nodes. Each randomly selected node sends a message containing the claim to start a random walk in the network, and the passed nodes are selected as witness nodes and will store the claim. If any witness receives different location claims for a same node ID, it can use these claims to revoke the replicated node.

TRAWL-based Detection: A method to reduce the memory cost of RAWL protocol is TRAWL, because sensor nodes usually have limited size of memory, for example, on the order of a few kilobytes, which is also a precious resource. Basic idea is to employ a trace table at each node to record the traces (represented by "digests") of random walks.

2.4 Implementation

Randomly nodes are deployed to form WSN. For this, after taking the number of nodes from the user, random x and y values are generated, which becomes the position of the node. A sink node is placed. Data aggregation is used to report the detection of event to the sink node. A copy of my node function is used for all the nodes. Standard distance vector routing protocol is used to find the shortest path between two nodes, which will be further used for information exchange in minimum possible cost. Two classes, namely my node representing the nodes and sink representing the sink node, are written. The my node class consists of basic parameters of a node in ad hoc wireless sensor network like node ID, sensing range, neighbor list, energy, data, delay, deployment details.

As per the requirement, objects of this class are made for the nodes of ad hoc WSNs. As we run the code depending on the number of type 1 and type 2 nodes, requested objects of my node are made. One sink node is created. Then, these nodes are deployed in random order. Then, these nodes will exchange beacon messages which uses send_pckt.m to exchange the neighboring node information. Then, the list of neighbors is updated. Thus, an ad hoc wireless sensor network is simulated. Then, as per the number of type 2 nodes entered, clone nodes or intruders are created. Then, each node broadcasts a signed location claim. Each of the node's neighbors probabilistically forwards the claim to some randomly selected nodes.

Each randomly selected node sends a message containing the claim to start a random walk in the network, and the passed nodes are selected as witness nodes and will store the claim. If any witness receives different location claims for a same node ID, it can use these claims to revoke the replicated node. When a randomly chosen node starts a random walk, all the passed nodes will still become witness nodes. However, now they do not definitely store the location claim; instead, they store the location claim independently with probability. Then, the required data are stored and are used to generate the graph. The clone node is simulated. The clone attack is simulated. The attack is made completely random by implementing RAWL. Track the walk and calculate the probability of detection. Implement TRAWL which is based on RAWL. A table is used to record random walks. Plot the graph to analyze the algorithm.

2.4.1 Steps for RAWL Protocol

The initial concept of RAWL is as follows:

- Each node broadcasts a signed location claim. (Its location in the form of x and y coordinates is in the claim).
- Each node has private and public key pairs stored in non-volatile storage of respective nodes.
- Each of the node's neighbors probabilistically forwards the claim to some randomly selected nodes.
- Each randomly selected node sends a message containing the claim to start a random walk witness selection in the network, and the passed nodes are selected as witness nodes and will store the claim.

2.4.2 Steps for TRAWL Protocol

The TRAWL protocol is implemented as follows:

- They do not definitely store the location claim; instead, they store the location claim independently with probability $c2/\sqrt{n} \log n$.
- Each witness node will create a new entry in its trace table for recording the pass of a location claim.
- The trace table maintained by each node has every entry of the table corresponds to the pass of a random walk.
- The table has the two columns: node ID and claim digest. The node ID is the ID field of a claim. The claim digest is a truncated message authentication code (MAC) of the whole location claim.
- When receiving a location claim, a node will first find the entries which have the same node ID as the claim in its trace table.
- When the two digests are different, the node detects a clone attack. If the node stored the location claim of the entry, it will flood the network with the two location claims to revoke replicas.

3 Experimental Analysis

In Fig. 1, each node broadcasts a signed location claim to its neighbors. Each neighbor verifies the signature and checks the probability. The distance between two neighbors cannot be bigger than the transmission range. Each chosen node that receives the claim of a first verifies the signature and then stores the claim and becomes a witness node of a. Also, it will start a t-step random walk in the network. The neighbor will also become a witness node of a. When node finds a collision, that is, two different location claims with a same ID, then the node will

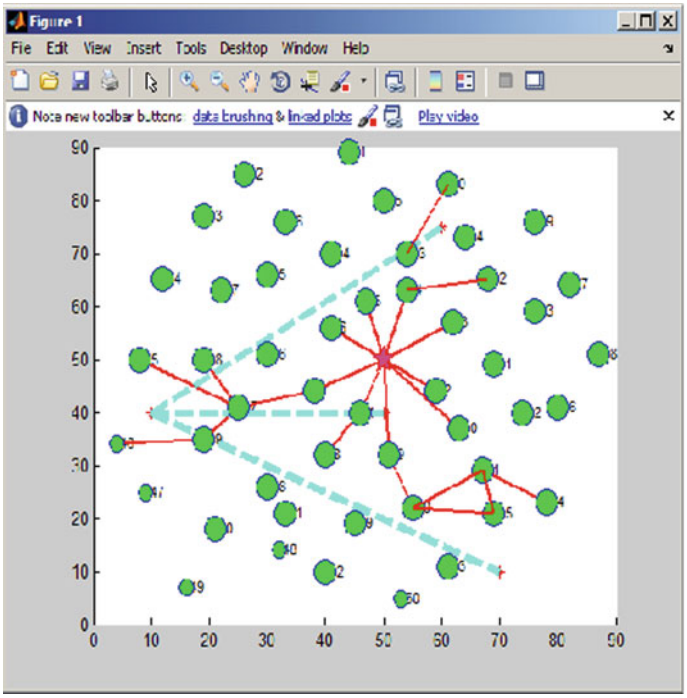


Fig. 1 The RAWL protocol

broadcast the two conflicting claims as evidence to revoke the replicas. Each node receiving the two claims independently verifies the signatures. If the two signatures are valid, it terminates the links with replicas.

Figures 2 and 3 show the probability detection of type 1 nodes and type 2 nodes, respectively, on broadcast reachability and connectivity in heterogeneous WSN. Here, reachability refers to coverage area, and connectivity is getting connection for data exchange. Heterogeneous refers to different abilities of nodes to take decisions; they are dependent. A type 2 node refers to introducer. The probabilities of reachability of type 1 nodes are higher than type 2 nodes.

Figure 4 shows the memory overheads of RAWL and TRAWL protocols. Here, the size of a location claims and size of a trace table are in bytes. And the number of bytes, each node stores in its memory. The probability of detection of TRAWL is higher than RAWL. The mechanism of TRAWL is used to reduce the memory overhead of RAWL using a table to cache the digests of location claims. The number of bytes used in TRAWL is less than RAWL as shown in the Fig. 4.

Figure 5 shows the probability of detection of RAWL and TRAWL protocols with different number of walk steps. Here, walk steps refer to number of nodes required to find the attacker. TRAWL has nearly same probability of detection with RAWL with less than 0.1 differences.

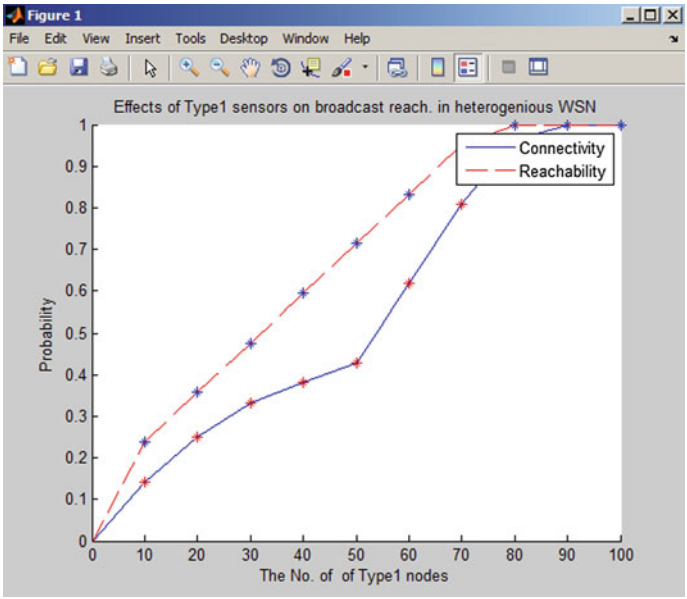


Fig. 2 Effects of type 1 sensors on broadcast reachability in heterogeneous WSN

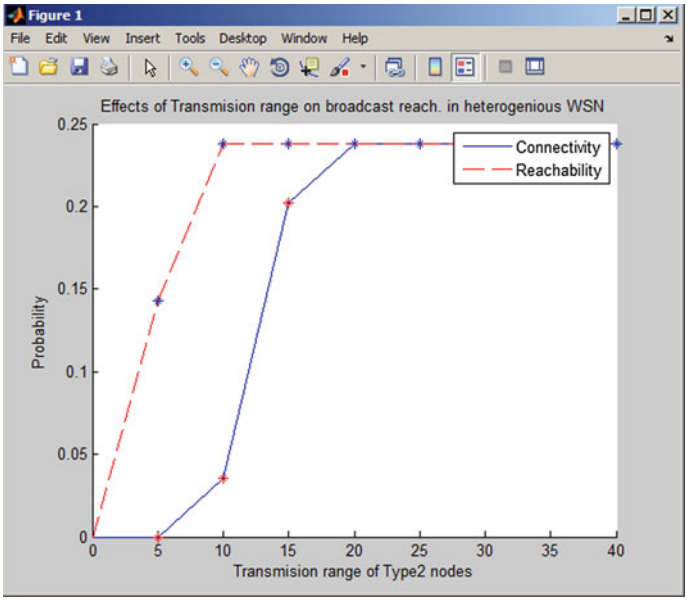


Fig. 3 Effects of type 2 sensors on broadcast reachability in heterogeneous WSN

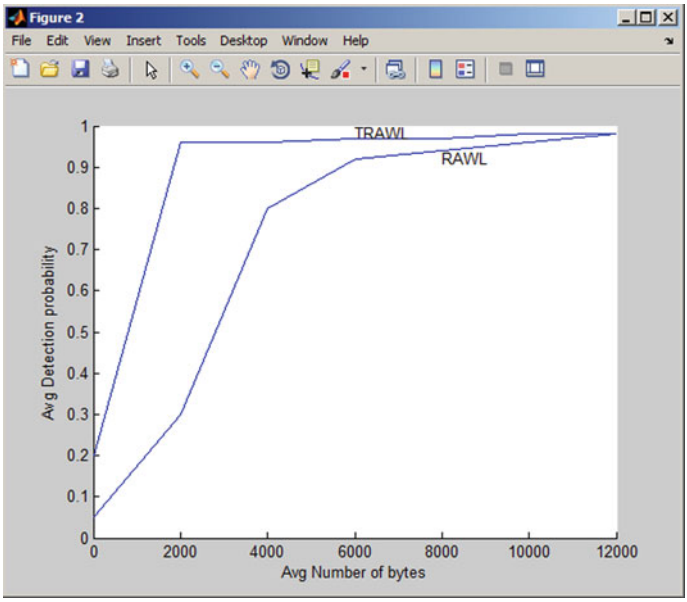


Fig. 4 Average probability detection of RAWL and TRAWL versus number of bytes

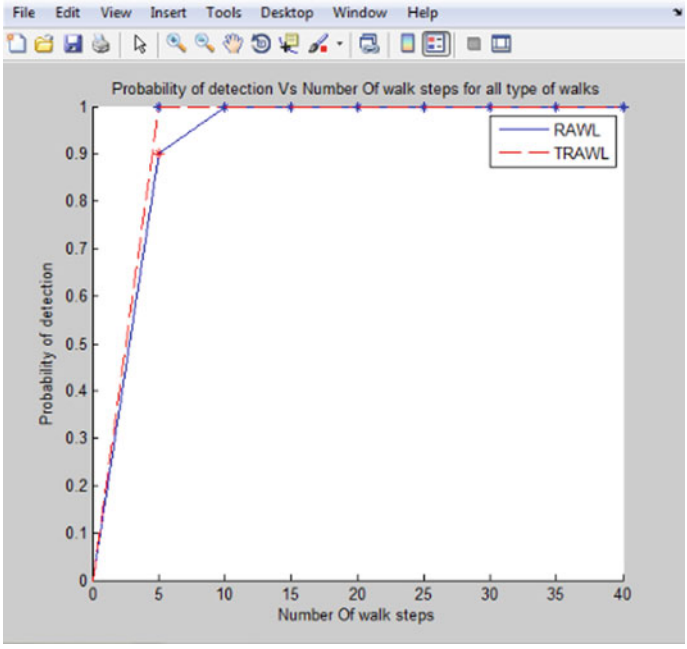


Fig. 5 Probability of detection of RAWL and TRAWL versus number of walk steps

4 Conclusion and Future Scope

It has been discussed that various approaches are used to detect node replication, which shows how centralized approaches place excessive trust in the BS and excessive load on those nodes near it. Local voting schemes are well equipped to detect distributed node replication. It is found that existing solutions have several drawbacks which greatly limit the usages. Firstly, it is shown that in order to avoid the drawbacks of existing approaches, replica-detection protocols must be NDFD and fulfill security requirements on witness selection. Based on random walk, two new NDFD-based protocols, RAWL and TRAWL, are proposed to fulfill the requirements and have comparably high communication overhead. But they provide a better trade-off between communication overhead and security properties than the other protocols. Finally, TRAWL is used to reduce the memory overhead of RAWL. Due to limitation of time and constraint of resource, the current project work is restricted to specific functionality only. But in case, if such obstruction is overcome, the current project work could be extended with following futures: The time complexity of detecting the intruder could be improved; the memory overhead could be reduced; and there could be a usage of a secret key with which the detection of intruder could be informed to other nodes, but this set of keys must be used only for this purpose and not for data communication.

References

1. Zeng Y, Cao J, Zhang S, Guo S, Li X (2010) Random-walk based approach to detect clone attacks in wireless sensor networks. Proc IEEE Commun 28(5)
2. Akyildiz I, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: a survey. Comput Netw 38(4):393–422
3. Parno B, Perrig A, Gligor V (2005) Distributed detection of node replication attacks in sensor networks. In: Proceedings of the IEEE symposium on security and privacy (S&P'05), pp 49–63
4. Liu F, Cheng X, Chen D (2007) Insider attacker detection in wireless sensor networks. Proc IEEE INFOCOM, pp 1937–1945
5. Chan H, Perrig A, Song D (2003) Random key predistribution schemes for sensor networks. In: Proceedings of the IEEE symposium on security and privacy (S&P'03), May 2003, pp 197–213
6. Yang Y, Shao M, Zhu S, Urgaonkar B, Cao G (2008) Minimum network traffic in sensor networks. In: Proceedings of the first ACM conference wireless network security (WiSec'08), pp 77–88
7. Melchor CA, Ait-Salem B, Gaborit P, Tamine K (2009) Active detection of node replication attacks. Int J Comput Sci Netw Secur 9(2):13–21
8. Wang H, Sheng B, Tan CC, Li Q (2007) Key LED-transmitting sensitive data over out-of-band channels in wireless sensor networks. College of William and Mary, Computer Science, Williamsburg, VA, Technical report WM-CS-2007-11
9. Liu A, Ning P (2008) An efficient key management scheme based on elliptic curve signcryption for heterogeneous wireless sensor networks. In: Proceedings of the seventh international conference on information processing in sensor networks (IPSN'08), pp 245–256

Size and Shape-Based Ovarian Abnormality Detection of Ultrasound Images

B. S. Usha, S. Sandya and G. Shruthi

Abstract Incidence of ovarian abnormalities like infertility, ovarian cysts, and risk of malignancy is rapidly increasing. Monitoring the ovaries is necessary for general and reproductive health of women. Transvaginal ultrasound (TVUS) image is preferred in detection of ovarian abnormalities as it has a wider field of vision and has a better diagnosis rate. It allows better visualization in shorter examination times and can detect masses as small as 1 cm. Analysis of ovary is done by analyzing major axis, minor axis, and shape of the ovary which is done manually by the expert. Manual analysis is time-consuming and prone to subjective error. Hence, there is a need for computer-assisted diagnostic support system for detection of ovarian features to aid the experts in fast and accurate diagnosis. In this paper, a novel, automated method for extracting ovarian parameters major and minor axis and shape is proposed. The proposed technique is tested on 30 TVUS images of ovaries. The experimental results are validated against the manual measurements done by the expert, and the results obtained are in good agreement with the experts. The proposed approach could achieve sensitivity of 60 %, specificity of 92 %, and an error rate of 14.9 %.

Keywords Transvaginal ultrasound · Ovary · Ovary parameters · Speckle filters · Image segmentation · Image enhancement · Feature extraction

B. S. Usha (✉) · S. Sandya · G. Shruthi
R N S Institute of Technology, Bangalore 560098, India
e-mail: bsusha@gmail.com

S. Sandya
e-mail: sandya9prasad@gmail.com

G. Shruthi
e-mail: shruthig29111988@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_31,

1 Introduction

Ovaries are the female pelvic reproductive organs that house the ova and are also responsible for the production of estrogen and progesterone. They are oval in shape with uneven surface, located on the left and right sides of the uterus. Ultrasound is the default imaging method used by the gynecologist as a diagnostic tool to 'image' or see structures in the female pelvis. TVUS image is preferred in detection of ovarian abnormalities as it has a wider field of vision and has a better diagnosis rate, can detect masses as small as 1 cm, and distinguish solid lesions from cysts. Ovary is most frequently scanned organ by ultrasound in many abnormalities like infertility, irregular menstruation cycles, polycystic ovary syndrome detection (PCOS). Determination of ovarian shape and size is one of the first steps in evaluation of the health of the ovary. The ovary is imaged for its morphology (major axis, minor axis in the lateral view) and shape analysis. The normal range of ovary measurement is listed in Table 1.

The first signature of ovarian abnormality is the increase in the major-axis and minor-axis lengths of the ovary and change in the shape of the ovary which can be readily analyzed using a TVUS image.

Measurement of these parameters is taken manually by the doctor who requires 20–25 min per patient for complete diagnosis.

Hence, the proposed method for an automated measurement of these parameters saves doctor's time and helps in making a fast diagnosis.

2 Background Work

The ultrasound images are of poor quality and low contrast due to the presence of speckle noise which is visible as a dense granular noise that is spread throughout the image. In computer-assisted processing of ultrasound images, speckle reduction is one of the preprocessing techniques used to improve the image quality and possibly the diagnostic potential of the image. Shruthi et al. [1] discuss various speckle reduction filters and concludes that bilateral filter and cascaded linear filters best suit for denoising ultrasound images.

For automated measurement of ovarian parameters, it is first required to segment the ovary from the ovarian ultrasound images. Segmentation of ovary from its background is complex because of the high homogeneity between the region of

Table 1 Normal ovary measurement

Normal ovary	Normal size
Major-axis length	3–5 cm
Minor-axis length	1.5–3 cm
Depth (cross-section)	1.5–2 cm
Shape	Oval (almond)

interest (ROI) ovary and background. Literature is available on computer-assisted approaches for ovarian follicle segmentation, but literature on computer-assisted approaches to extract the ovary is very scarce. In [2], edge detection method for extraction of follicles from ovarian ultrasound image is proposed. PotoEnik and Zazula [3] proposed optimal thresholding and Krish's operators for segmentation of follicles. But the algorithm required more execution time, and recognition rate was approximately 70 %. The recognition rate was increased to 88 % by using region growing methods. Krivanek and Sonka [4] proposed a technique for detection of follicle wall inner and outer borders which involved watershed segmentation, interactive selection of the follicle of interest, inner border detection, and outer border detection. PotoEnik and Zazula[5] proposed prediction correction algorithm and cellular neural networks (CNN) for detection of follicular region. They could achieve sensitivity of 78 % and specificity of 15 %. Hiremath et al. [6, 7] proposed horizontal scaling thresholding (HST) and vertical scaling thresholding (VST) techniques for extraction of follicles. Then, the results obtained from HST and VST were fused by image fusion. Area, ratio of major-axis length to minor-axis length, centroid, compactness, and extent were the features used for classification. The classification rate for the proposed method was 83.33 %, false acceptance rate (FAR) was 2.08 %, and false rejection rate (FRR) was 16.66 %. In [8], edge-based methods like Canny, Sobel, and Prewitt were applied on histogram-equalized image to get the follicle region. The results showed the classification rate of 75.2 % with the FAR 22.5 % and FRR 24.1 %. The results were further improved by using active contour methods [9, 10] with detection rate increased to 94.2 %, FAR 12.6 %, and FRR 5.7 %.

The above literature [2–10] presents methods for segmentation of ovarian follicle from ovarian ultrasound image. This paper discusses a new algorithm for segmentation of ovary from ovarian transvaginal ultrasound (TVUS) images and extraction of its geometric features, namely major-axis and minor-axis lengths and eccentricity (shape parameter). The experimental results of proposed method are validated against the manual measurements taken by the expert.

This paper is organized as follows: Section 3 provides the description of the algorithm where speckle reduction filters, contrast enhancement methods, segmentation techniques, and feature extractions are explained in detail. In Sect. 4, the experimental results are discussed along with manual measurements taken by the expert, and in Sect. 5, conclusion and future scope of work is discussed.

3 Proposed Method

The proposed approach using image processing techniques is shown in Fig. 1. The input is the TVUS image of ovary as shown in Fig. 2a. The ultrasound images are taken from GE-LOGIQ Book XP ultrasound machine, which is a high-performance multipurpose color hand-carried imaging system designed for abdominal, cardiac, obstetrics, gynecology, vascular, musculoskeletal, small parts, pediatric

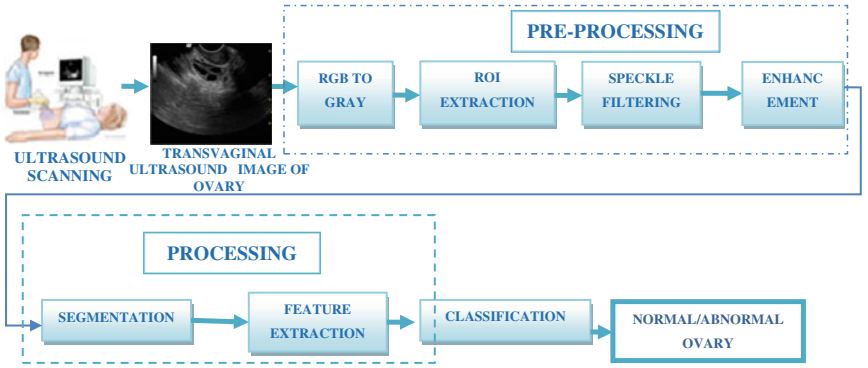


Fig. 1 Block diagram of proposed ovarian abnormality detection system

neonatal, and intraoperative applications. The images are obtained in JPEG-LS 24-bit format (as the machine can work in color Doppler mode).

3.1 Preprocessing

3.1.1 RGB to Gray Conversion

The input is TVUS image of ovary which is in 24-bit RGB format which is converted into 8-bit intensity image using Eq. (1)

$$\mathrm{I}(x, y) = 0.299 * R(x, y) + 0.587 * G(x, y) + 0.114 * B(x, y) \tag{1}$$

3.1.2 ROI (Ovary) Selection

TVUS image (Fig. 2a) contains a fan-shaped area which shows the image of interest obtained by the ultrasound scan and a marginal area containing other information such as the patient's name and date (Fig. 2b). Since the margin area is of little significance, it can be removed. Hence, a subimage which contains only the bounding box of the fan area is extracted from the original image.

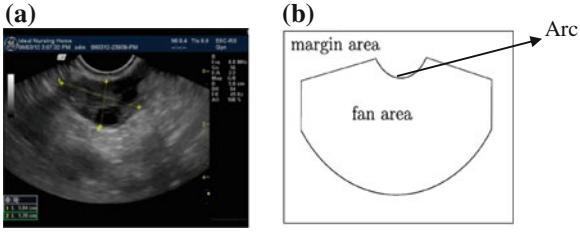


Fig. 2 **a** Ultrasound image of ovary and **b** Image showing fan area and margin area

3.1.3 Speckle Reduction Filter

Speckle filtering algorithm is applied on ROI to remove the speckle noise. Bilateral filter [11–13] is a convolution-based linear filter which gives better smoothing and detail preservation.

The window size chosen is 5 × 5 and sigma value as 10. The algorithm to implement bilateral filter is as follows:

- Read the input image.
- Choose the values of window size W and sigma (window size = 5 and sigma = 10).
- For chosen window, find the dimensions, that is, [M,N] = size(window).
- Compute the Gaussian distance weights using the formula

$$G = \exp(-(M.\hat{}2 + N.\hat{}2)/(2 * \text{sigma}\hat{}2)) \tag{2}$$

- Create a null matrix B with dimension equal to size of input image.
- Now convolve the window on input image and by extracting the local regions.
- Compute the Gaussian intensity weights using the equation

$$H = \exp\left(\frac{-(W - I(x,y))\hat{}2}{2 * \text{sigma}^2}\right) \tag{3}$$

where W is the convolution mask and I is the input image.

- Calculate the bilateral filter response using the formula

$$F = H. * G \tag{4}$$

- The bilateral filtered image is obtained by using the equation

$$I_f(x,y) = \frac{\text{sum}(F(x,y). * I(x,y))}{\text{sum}(F(x,y))} \tag{5}$$

In bilateral filter, the choice of sigma is very important. If sigma is high, the filter will act as a smoothing filter and will blur the edges. If sigma is too low, the noise cannot be removed.

3.1.4 Contrast Enhancement

The image appears smoothened after filtering, but in order to segment ovary from its background, a good contrast between the ovary and the background is required. Hence, contrast stretching (Eq. 6) and power law transform [14] (Eq. 7) are applied.

$$I_c(x,y) = \left(I_f(x,y) - c\right)\left(\frac{a-b}{c-d}\right) + a \tag{6}$$

where 'a' and 'b' are the allowable maximum and minimum pixel value limits, and c and d are the maximum and minimum pixel value limits of the input image.

Power law transform is given by

$$I_u(x,y) = C * \left(I_c(x,y)^{\gamma}\right) \tag{7}$$

where C and γ are positive constants. In general, a value of $\gamma > 1$ enhances the contrast of high-value portions of the image at the expense of low-value regions, while we see the reverse for $\gamma < 1$.

3.2 Processing

The processing stage involves segmentation and feature extraction. Segmentation process extracts ovary from its complex, homogeneous background. Morphological operators are applied on the segmented image to remove spurious regions. This is followed by extracting the features of the ovary.

3.2.1 Segmentation

Adaptive thresholding (local region–based thresholding) and global thresholding methods are used as shown in Fig. 3. The main idea of adaptive thresholding is to divide an image into subimages and threshold these individually. Since the threshold for each pixel depends on its location within an image, this technique is said to be adaptive. In other words, local adaptive thresholding selects an individual threshold for each pixel based on the range of intensity values in its local neighborhood. Given a window size $W \times W$ and a user-defined threshold value 'T', the adaptive threshold is calculated as given in Eq. 8.

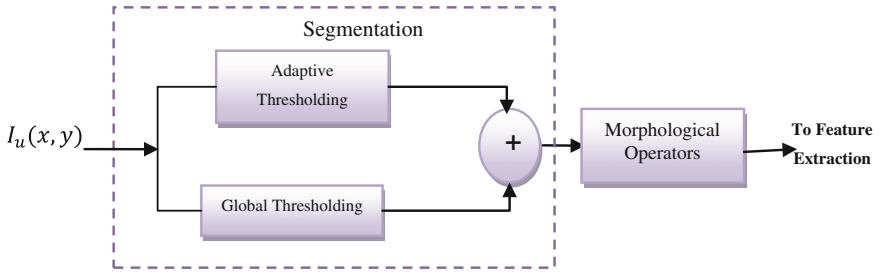


Fig. 3 Block diagram of segmentation module

$$\text{threshold} = \text{mean}_{\text{window}(W)} + T \tag{8}$$

There are some regions which still show homogeneity between extracted region and background. In order to extract ovary accurately from background, global thresholding is applied. This method is used to create a mask which helps in getting correct boundary of ovary ROI. The bi-level thresholding is given by Eq. 9.

$$\begin{aligned} I_g(x,y) = & 1 \quad \text{if } I_u(x,y) > T \\ & 0 \quad \text{if } I_u(x,y) \leq T \end{aligned} \tag{9}$$

The segmented output is the sum of adaptive thresholding and global thresholding. But this still contains the spurious regions which are removed using morphological operators.

3.2.2 Morphological Operators

The process of segmentation is followed by morphological operators [15] such as region filling, closing, erosion, and dilation to get the complete and exact boundary of ovary and to remove spurious objects. This is followed by labeling to exactly extract ovary. A connected component labeling (CCL) algorithm finds all connected components in an image and assigns a unique label to all points in the same component [15, 16]. The area of each labeled components is calculated, and the component with the largest area is the ROI (ovary).

3.3 Feature Extraction and Classification

The goal of feature extraction is to obtain representative features that can be used to classify the ovary. The main aim of geometric feature extraction is to recognize the geometric properties of ovary in ovarian ultrasound images. The parameters considered are major-axis length, minor-axis length (Table 1), and eccentricity. The major axis and minor axis of an ellipse and ovary are shown in Fig. 4a and b

The shape is measured using shape parameter, eccentricity, which is the ratio of the distance between the foci of the ellipse and its major-axis length. The eccentricity equation is given by Eq. 10

$$\text{eccentricity} = \sqrt{\left(1 - \frac{a^2}{b^2}\right)} \tag{10}$$

where 'a' is the semi-major-axis length and 'b' is semi-minor-axis length.

Classification of ovary as normal/abnormal is based on the features extracted such as major-axis length, minor-axis length, and eccentricity. For the normal

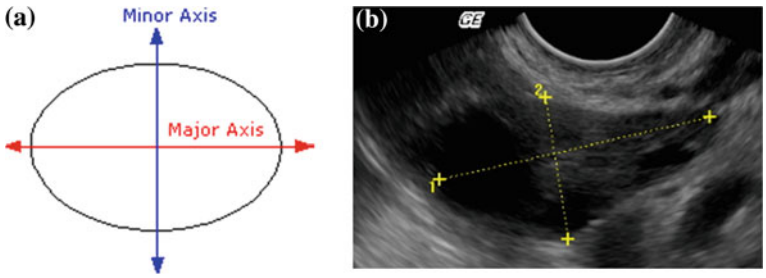


Fig. 4 **a** An ellipse, **b** ovary region (*1* major axis, *2* minor axis)

ovary, the major-axis length should be within the range of 3–5 cm, minor-axis length should be within the range of 1.5–3 cm, and eccentricity should be between 0.78 and 0.9. Any variations in these measurements are considered as abnormal.

4 Results and Discussion

The input is an RGB (24 bits, 532 × 434 × 3) JPEG-LS image which is converted into intensity image (8 bits, 532 × 434). The ROI is just below the arc (Fig. 2b). In our experiments, the window size of 381 × 251 below arc holds the ROI. Hence, all images are cropped to 381 × 251 for further processing. The cropped ROI is filtered using bilateral filter which provides smoothing and detail preservation. The window size chosen is 5 × 5 and sigma value as 10. From filtered image, it can be seen that the ovary region is not distinguishable from background (Fig. 5).

Hence, contrast enhancement techniques contrast stretching and power law transform are applied to obtain a good contrast between the ovary region and background.

Segmentation is done on enhanced image to extract ROI from its homogeneous background. The segmentation uses adaptive thresholding and global thresholding. In adaptive thresholding, image is divided into subimages and threshold value is selected as in Eq. 9. The global thresholding is performed to get the image mask by setting the threshold value equal to 0.128. The outputs of adaptive thresholding and global thresholding (∼ Image Mask- Negation of Image mask) are summed to get the boundary of ovary. From the result (Fig. 5), it can be seen that there are some spurious objects that must be removed. The spurious regions are removed using morphological operators region filling, closing, erosion, and dilation. But there still exists some noise, and hence, to remove these regions, labeling of components is done. The area of each labeled component is calculated, and the largest area corresponds to ovary region, and hence, this component is chosen from the labeled components, thus removing the spurious objects other than ROI (ovary) in the image. The next step is to extract the features of the ovary such as

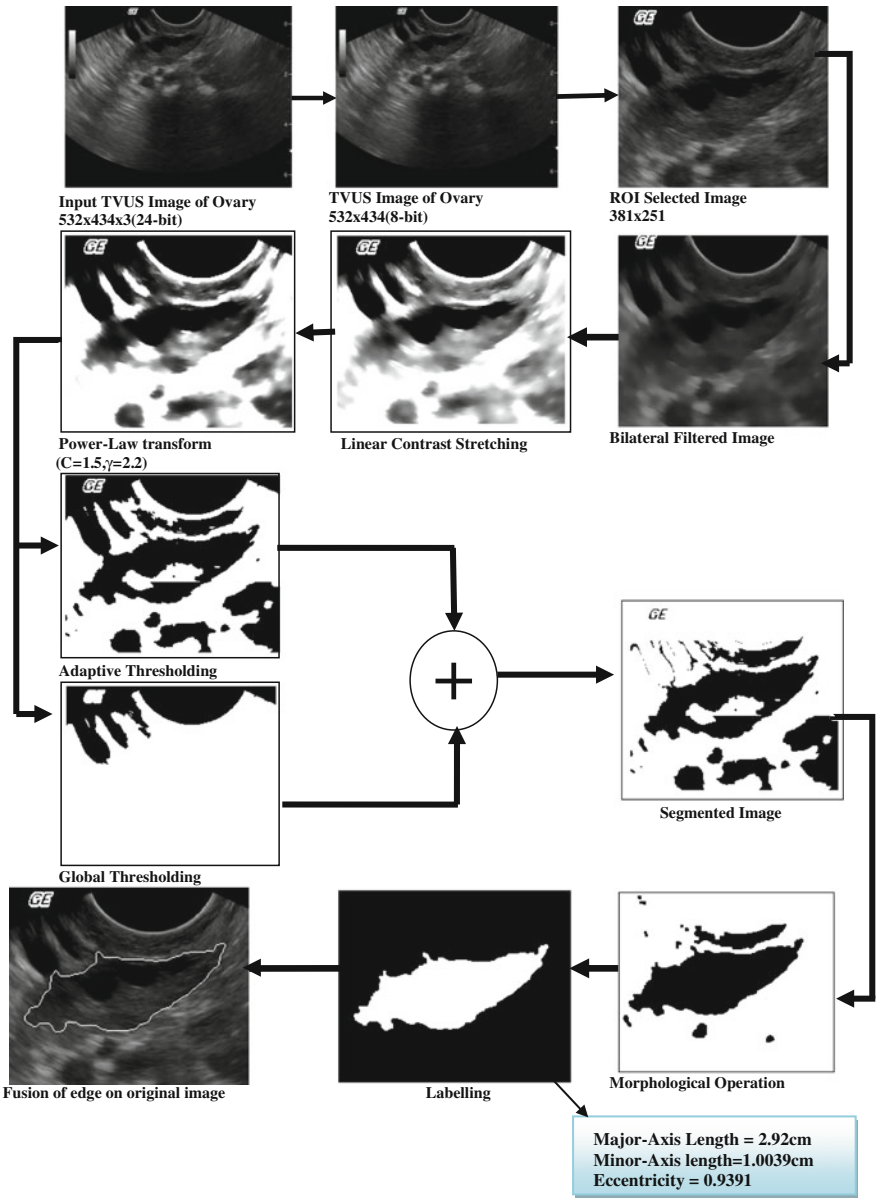


Fig. 5 Extraction of ovarian features from TVUS of Ovary

major-axis length, minor-axis length, and eccentricity which is used for classification of ovary as normal/abnormal. For the image shown in Fig. 5, the features obtained are as follows:

Major-axis length = 184, minor-axis length = 63.2427, and eccentricity = 0.9391. The values obtained are in terms of pixel distance which has to be converted into centimeters for which the resolution of the machine should be known.

For the TVUS images used in this project, 63 pixels correspond to 1 cm. Therefore, major-axis length = 184/63 = 2.92 cm, minor-axis length = 63.2427/63 = 1.0039 cm, and eccentricity = 0.9391. Since the values are within the normal range of the ovary and eccentricity is >0.78, ovary is classified as normal.

5 Calculation of Sensitivity, Specificity, and Error Percentage

A dataset of 30 images is taken for the proposed ovarian abnormality detection system. Among the database of 30 images, 25 images are normal and 5 are abnormal images. Sensitivity and specificity are the important parameters used to measure the accuracy of the algorithms. The accuracy can be calculated based on four values, namely the true-positive (TP) rate, the false-positive (FP) rate, the false-negative (FN) rate, and the true-negative (TN) rate.

TP is when an 'abnormal' image is correctly identified as 'abnormal.'

FN is when an 'abnormal' image is incorrectly identified as 'normal.' TN is when a 'normal' image is correctly identified as 'normal,' and FP is when a 'normal' image is incorrectly identified as 'abnormal.' The equations of sensitivity and specificity are given by Eq. 11 and Eq. 12

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{12}$$

Out of 25 normal images, 23 images are classified as normal and 2 images are classified as abnormal. Among 5 abnormal images, 3 images are classified as abnormal and 2 images are classified as normal. Hence, true positive = 3, true negative = 23, false positive = 2, and false negative = 2

$$\text{Sensitivity} = \frac{3}{3+2} \times 100 = 60\ \% \qquad \text{Specificity} = \frac{23}{23+2} \times 100 = 92\ \%$$

The error percentage is calculated based on the results obtained and expert values.

The error percentage for length and width is calculated by using the Eq. 13 and Eq. 14.

$$\text{Error Percentage} = \frac{\left|\left((\text{length})_{\text{obtained value}} - (\text{length})_{\text{Expert value}}\right)\right|}{(\text{length})_{\text{Expert value}}} \times 100 \tag{13}$$

Table 2 Comparison of obtained results and expert values

Images	Obtained results		Expert value			Inference		
	Major-axis length (in cm)	Minor-axis length (in cm)	Major-axis length (in cm)	Minor-axis length (width) (in cm)	Eccentricity		Error % major axis	Error % minor axis
1	2.60	1.96	2.11	1.98	0.6599	Abnormal	23.2	1
2	4.36	2.31	4.02	2.47	0.8473	Normal	8.04	6.4
3	2.44	1.05	3.02	1.86	0.9023	Normal	19.2	43.5
4	4.15	1.29	3.96	1.58	0.9501	Normal	4.5	18.3
5	5.60	1.99	4.08	1.14	0.9349	Abnormal	27.1	42.7
6	4.11	2.15	3.87	2.41	0.8525	Normal	5.8	10.27
7	2.80	0.94	2.98	1.02	0.9419	Normal	6.04	7.8
8	2.34	1.14	2.51	1.24	0.8725	Normal	6.7	8.06
9	4.20	2.65	4.38	2.75	0.5625	Abnormal	4.1	3.6
10	3.12	1.60	3.04	1.70	0.8123	Normal	2.5	5.8
11	6.56	4.06	6.24	4.27	0.7843	Abnormal	4.8	4.9
12	3.59	2.11	3.38	2.08	0.8096	Normal	5.8	1.4
13	3.91	2.01	3.70	1.90	0.8133	Normal	5.3	5.4
14	2.77	1.77	2.96	2.07	0.7699	Normal	6.4	14.4
15	1.96	0.98	2.14	1.72	0.8249	Normal	8.4	43.02
16	2.34	1.44	2.08	2.20	0.7902	Normal	12.1	34.54
17	3.53	1.85	2.70	2.85	0.8500	Normal	23.5	35.08
18	3.74	2.25	2.38	1.88	0.8003	Normal	36.3	16.44
19	2.79	1.75	3.21	2.42	0.7788	Normal	37.8	27.68
20	3.44	1.56	3.12	1.49	0.9801	Normal	9.3	4.4
21	3.58	1.55	3.64	1.83	0.9005	Normal	1.6	15.3
22	1.86	1.09	2.41	1.63	0.8370	Normal	22.8	33.12
23	3.96	1.66	4.60	2.34	0.9080	Normal	13.9	29.05

(continued)

Table 2 (continued)

Images	Obtained results		Expert value			Inference		
24	2.26	1.05	2.97	1.48	0.8837	Normal	23.9	29
25	4.08	3.18	4.10	4.32	0.8190	Normal	0.4	26.3
26	2.87	1.68	3.00	2.19	0.7926	Normal	4.3	23.2
27	5.20	3.33	5.52	3.23	0.7250	Abnormal	5.8	3.0
28	5.62	4.69	7.02	4.98	0.5503	Abnormal	19.9	21.9
29	6.74	5.70	8.17	6.53	0.5642	Abnormal	17.5	12.7
30	7.45	5.50	7.20	5.38	0.6744	Abnormal	3.3	2.18
						Average %	11.9 %	17.4 %

Similarly for error percentage of width,

$$\text{Error Percentage} = \frac{\left|\left((\text{width})_{\text{obtained value}} - (\text{width})_{\text{Expert value}}\right)\right|}{(\text{width})_{\text{Expert value}}} \times 100 \qquad (14)$$

For example, consider the readings of image 2,

$$\text{Error Percentage (for length)} = \frac{4.36 - 4.02}{4.02} = 8.04\ \%$$

$$\text{Error Percentage(for width)} = \frac{2.47 - 2.31}{2.31} = 6.4\ \%$$

The average error percentage for major-axis length (length) obtained is 11.9 %. The average error percentage for minor-axis length (width) obtained is 17.4 %. The average error percentage for major-axis length (length) and minor-axis length (width) obtained is 14.9 %.

6 Conclusion

The experiments were conducted on 30 images, and the results obtained are in good agreement with the manual ovary detection by experts. The proposed approach could achieve sensitivity of 60 %, specificity of 92 %, and an error rate of 14.9 %. The average error percentage for major-axis length (length) obtained is 11.9 %. The average error percentage for minor-axis length (width) obtained is 17.4 %. The average error percentage for major-axis length (length) and minor-axis length (width) obtained is 14.9 %.

7 Future Enhancements

The database is limited to 30 images which is significantly small. The error percentage obtained is 14.9 %, which could be decreased if large database is considered.

Spatial linear speckle reduction filters are used for denoising. The algorithm can be tested after using frequency domain filters for denoising. The segmentation can be improved using region-based methods, contour tracing active contour segmentation methods.

Another parameter which experts usually prefer for classification is through evaluation of the volume of ovary. Hence, this volume parameter can be included along with major-axis length, minor-axis length, and eccentricity. To find the volume of ovary, the cross-sectional and lateral sectional image is required.

The other parameters such as area, perimeter, and solidity can be used to increase the specificity and sensitivity of the system. Algorithms like elliptic curve fitting for shape analysis can be used to obtain precise shape of ovary.

Acknowledgment Our sincere thanks to Dr. Sriprada Vinekar, Gynecologist and Sonographist, Ideal Nursing Home, Bangalore, for supporting with the TVUS image database of ovaries. We thank her for the continued discussions and for rendering manual measurement of ovaries and classification of ultrasound images of ovaries. We thank Dr. Vipula singh, HOD, ECE, RNSIT, for her valuable technical inputs. The authors are also thankful to RNS Institute of Technology for the excellent infrastructure support provided in executing the work.

References

1. Shruthi G, Usha BS, Sandya S (2012) A novel approach for speckle reduction and enhancement of ultrasound images. Int J Comput Appl 45(20):0975–8887
2. Mahmood NH, Ahmmad SNZ, Hashim H, Rani SNNA (2012) Ovary ultrasound image edge detection analysis: a tutorial using MATLAB. Int J Eng Res Appl (IJERA) 2(3):1635–1642, ISSN: 2248-9622
3. PotoEnik B, Zazula D Automated computer-assisted detection of follicles in ultrasound images of ovary. In: Tenth IEEE symposium on computer-based medical systems 1063-7125/97
4. Krivanek A, Sonka M (1998) Ovarian ultrasound image analysis: follicle segmentation. IEEE Trans Med Imaging 17(6):935–944
5. Potočnik B, Zazula D (2000) Automated ovarian follicle segmentation using region growing. V: LONCARIC, Sven (ur). IWISPA 2000, Pula, Croatia, June 14–15, 2000. Zagreb: University Computing Center, University of Zagreb, 2000, str. 157-162. [COBISS.SI-ID 5424918]
6. Hiremath PS, Tegnoor JR Recognition of follicles in ultrasound images of ovaries using geometric features. IEEE Trans Med Imaging 17(6), ISBN: 978-1-4244-4764
7. Hiremath PS, Tegnoor JR Follicle detection in ultrasound images of ovaries using active contours method. IEEE Trans Med Imaging 18(9), ISBN: 978-1-4244-8594-9
8. Hiremath PS, Tegnoor JR (2010) Automatic detection of follicles in ultrasound images of ovaries using edge based method. IJCA Special issue on recent trends in image processing and pattern recognition RTIPPR
9. Hiremath PS, Tegnoor JR (2010) Follicle detection in ultrasound images of ovaries using active contours method. International Conference on Signal and Image Processing (ICSIP 2010) pp 286–291. doi:10.1109/ICSIP.2010.5697484
10. Mehrotra P, Chakraborty C, Ghoshdastidar B, Ghoshdastidar S (2011) Automated ovarian follicle recognition for polycystic ovary syndrome. 2011 International conference on image information processing (ICIIP 2011)
11. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Proceedings of 6th international conference on computer vision, pp 839–846. doi:10.1109/ICCV.1998.710815
12. Phelippeau H, Talbot H, Akil M, Bara S (2008) Shot noise adaptive bilateral filter. In: Proceedings of the 9th international conference on signal processing, pp 864–867
13. Barash D (2002) Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. IEEE Trans Pattern Anal Mach Intell 24:844–847
14. Gonzalez RC, Woods RE (2001) Digital image processing. 2nd edn. Addison- Wesley, Boston. ISBN: 0201180758

15. Harrington N (2008) Thesis report on segmentation of human ovarian follicles from ultrasound images acquired in vivo using geometric active contour models and a naïve's Bayes classifier
16. Sonka M, Halvic V, Boyale R (2007) Image processing, analysis and machine vision. Chapman and Hall, London. ISBN: 049508252X

Rate Quality Factor Optimization for VP8 Video Codec for Mobile Communication Applications

S. Basavaraju and B. Sivakumar

Abstract In this paper, mainly, we are proposing a better rate control algorithm optimization technique; it is applying for VP8 video coding for improving the performance and also quality of video codec data for mainly mobile communication applications. Actually, rate control plays an better role in video coding and transmission to provide the best video quality at the receiver end, and our proposed algorithm technique mainly exploits the existing constant quality control, which is governed by a parameter called quality factor (QF) to give a constant bit rates. For this purpose, a new mathematical model called the rate–quality factor(R–Q′) is derived to generate optimum QF for the current coding frame using the bitrates resulting from the encoding of the previous frame in order to meet the target bitrates. The process of calculating the QF is simple, and further calculation is not required for each coded frame. It also provides the rate control solution for both intra-frame-only and inter-frame coding modes. Our all experimental results show that the proposed scheme generates coding bits very close to target bits and provides improved coding efficiency at low bit rates also introducing simple complexity to measure for the video content. So that, in this proposed method considering previous results and comparing with our results then confirm by comprehensive experiments results.. The quality control parameter can be derived from Lagrangian multiplier and hence can be used in any type of encoder that uses RDO. Experimental results have shown that the proposed algorithm can control the bitrates within 1 % of the target bitrates on average, and it has better bitrates regulation over each GOP than the rate control algorithm of H.264. This is an advantage that is crucial in real-time multimedia data streaming in preventing buffer overflow or underflow. Performance of this video data is analyzed using

S. Basavaraju (✉)
ECE Department, Sapthagiri College of Engineering, Bangalore, Karnataka, India
e-mail: rajhunsur@yahoo.co.in

B. Sivakumar
Telecommunication Department, AIT, Bangalore, Karnataka, India
e-mail: sivabs2000@yahoo.co.uk

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_32,

PSNR and bitrates which needs to calculate once this new technique is implemented on specified plat form.

Keywords VP8 · Rate distortion · RCA · QS · RD · DCT · Video codec · Quality · SAD · MSE

1 Introduction

In Video coding and transmission of this data in multimedia, communication is very much important for today's multimedia's applications, and also digital video applications are increasing day by day and most of the applications have very different constraints. Major challenges in designing a good digital video system [1] are to provide encoded video with good visual quality and compression performance. For high compression, video coding is mainly needed operational control of the video encoder. Essentially, the rate control part of the encoder tries to regulate varying bitrates characteristics of coded bit streams in order to produce high-quality decoded frame at the receiver for a given target bitrates so that compressed bit streams can be delivered through the available channel bandwidth without causing buffer overflow and underflow.

So that in this experiment, a rate control algorithm, the performance of the proposed rate control algorithm in VP8 Video codec is compared with the other rate control scheme in H.264 and also in order to justify the proposed work since there is no previous work in VP8 Video codec as far as the rate control mechanism is concerned. Rate control algorithm (JVT-G012) of H.264 reference software (JM11) [2] consists of three tightly consecutive components: group of pictures (GOP) level rate control, picture level rate control and the optional basic unit level rate control.

The present VP8 Video codec architecture controls constant quality rather than bitrates by using a single QF as quality indicator to maintain the desired quality. The algorithm presented in this paper mainly exploits this idea by considering QF as a varying parameter in order to achieve constant bitrates. It has the advantage of giving stable quality while delivering the desired constant bitrates.

So the choosing of the predictor and also the weights of the neighboring pixels has a direct bearing on the efficiency of the algorithm. Usually, the design of video data compression schemes involve trade-offs among various factors, including the degree of compression, the amount of distortion introduced (e.g., when using lossy data compression) and the computational resources required to compress and uncompress the data. Considering any kind of video coding, the all prediction errors in each of the DCT blocks (8×8) are DCT transformed and then quantized (according to specified QP), zigzag scanned and lastly run-length coded. In this way, rate distortion optimization techniques have been mainly applied to video encoders [3], and it will provide good potential for substantial improvements in compression efficiency.

2 Rate Distortion Optimization for VP8 Video Codec

(a) Present VP Video codec

The complete block diagram of VP8 Video codec [2] is shown in Fig. 10, at this moment, libvpx is the only software library capable of encoding VP8 video streams and also the same libvpx is capable of decoding VP8 video streams. The VP8 Video codec's good strength is it offers the "Better quality real-time video delivery", This Video codec also started to give the attention (attract) for wide interest in the video coding research community from both industry and as well as academia.

One more important factor is VP8 uses three different types of reference frames for inter prediction: the "last frame", a "golden frame" (one frame worth of decompressed data from the arbitrarily distant past) and also last frame e called as "alternate reference frame".

And major features of VP8 Video codec is it will support for low bandwidth applications, and major features are Web video format, hybrid transform with adaptive quantization, flexible reference frames, efficient intra prediction and interpretation, sub-pixel interpolation, adaptive in-loop deblocking filtering, even frame-level adaptive entropy coding and friendly data partitioning.

Basically, VP8 video compression codec WebM video format is available. It has three different versions (0–3) are defined as four different profiles with different decoding complexity.

There are two decoders (1) VPX decoder and (2) simple decoder, and the only difference between the two is that simple decoder just simply decodes a stream and nothing else, while the VPX decoder can do plenty of other things.

(b) Proposed Rate control algorithm optimization

So that the rate control plays very much an important role in all video standards. Without this, rate control technique [4] may be underflow, and overflow of the end buffer may occur due to the mismatching between the source bit rate and the available channel bandwidth for delivering a compressed bit stream. In other words, without rate control, any video coding encoder is difficult to use. Besides these standard-recommended rate control schemes [3], many improved algorithms have also been developed in the previous years. Basically, for the transmission style, rate control method can be considered as two major categories.

Constant-bit-rate (CBR) control for the constant-channel-bandwidth video transmission [5] and variable-bit-rate (VBR) control for the variable-channel-bandwidth video transmission [6]. In terms of the unit of rate control operation, these rate control schemes can be classified into macro block- [7], [8] or frame-layer [2] rate control. These rate control [7] schemes usually resolve two main problems. The first is how to allocate proper bits to each coding unit according to the buffer status, that is, rate allocation, and the second is how to adjust the encoder parameters to properly encode each unit with the allocated bits, that is, quantization parameter adjustment. Many parameters affect the rate distortion streaming

performance including the prediction dependency structure, which determines coding efficiency and random access to images; the accuracy of the geometry information which is used in coding. In this paper, mainly, we are proposing as we mentioned in above, the present google's VP8 Video codec mainly controls firm and constant good quality instead of bitrates by using a user-defined parameter, QF, as this quality indicator to maintain the targeted good quality. Our proposed algorithm exploits this idea by choosing QF as a varying parameter in order to achieve average bitrates which is constant over each GOP. But Since the QF, it plays an important role in controlling the quality of the encoded video sequence or the number of bits generated in the encoding process of VP8 Video codec, finding the optimum QF for this set of target bitrates and also test sequences could lead to an algorithm that controls the output bitrates of the encoder. The complexity of each frame in the sequence could be changing all the time. So, it is practically impossible to use a constant QF to encode the entire video sequence to achieve constant bitrates over a GOP because optimum QF for a previous frame would no longer be optimum for the current and the following frames. However, bitrates controlling govern a GOP could be possible by adaptively changing the QF of each frame according to a certain type of algorithm before we encode. Based on this idea, a relationship between the bitrates, R, and the QF, which can be used to estimate the QF for a given target bitrates, is derived. This model is known as the R–QF model.

3 RD Proposed Algorithm for VP8 Video Codec

Figure 10 shows the complete block diagram of VP8 Video codec in this using the generated number of bits required to encode a frame as the feedback.

The rate control procedure in TM5 works in three steps. First of all, target bit allocation estimates the number of bits available to code the next picture. It is performed before coding the picture. In the next step, by means of a virtual buffer, the reference value of the QP for each MB is set. In the final step, adaptive quantization is carried out by modulating the reference value of the QP from the previous step according to the spatial activity in the MB to derive the final QP, which is used to quantize the MB. The detailed procedure of TM5 can be found in. In the proposed rate control method, only the modified version of TM5's step one, which is target bit allocation, is used to estimate the optimum number of bits required from the different types of frame in order to meet the target bit rate. Finally, we employed our proposed rate control algorithm in order to achieve bitrates close to the target bitrates for both types of frame coding available in VP8, which are intra-frame-only coding, where there is only intra-frame type and inter-frame coding which is IL2L2L1 or IBBP coding.

3.1 The Rate–QF (R–F) Model

This section presents the derivation of the relation between rate and QF in the R–QF model. Since R and D are inversely proportional to each other as below:

$$R \propto 1/D \tag{1}$$

$R = K/D$ where K is constant

$$D = KR^{-1} \tag{2}$$

For high data rate, this distortion [5] considered by the quantization usually approximately is uniformly distributed and also the power of the quantization noise P is $Q2/12$. Considering the distortion of the quantization, noise at low bit rates is not exact uniform. So that defining a distortion D for the quantization parameter Q by (Fig. 1)

$$D = P \times \mathrm{Q} \tag{3}$$

And here P states that distortion value. And our scheme for this Video coding can expressed as

$$R(D) = \log_2(\sigma/P \times D) \tag{4}$$

4 Frame-Level and Macro Block-Level Rate Control

To find out the quantization parameter, we should allocate target bits for the current image frame. For this purpose, need to find number of bits allocated to I and P frames like in same manner of MPEG 2 TMS Model. And also In order to decide the quantization parameter for the current macro block, considering the rate-quantization model statistical characteristics of the source data for the current macro block should be known to calculate the quantization parameter.

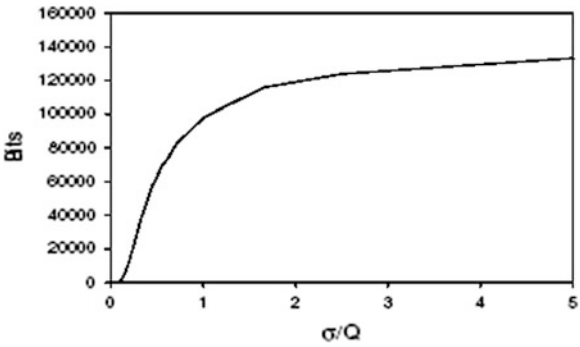


Fig. 1 For high data rate, this distortion

In order to decide the quantization parameter for the current macroblock, use the rate-quantization model also need some information to calculate the quantization parameter from the proposed rate-quantization model. Similarly, to determine the quantization parameter, considering distortion parameter, which can be obtained from the quantization parameter, And then Predicting the distortion parameter by distortion parameters of the left and the top macro blocks of the current macro block, respectively.

4.1 Adjustment of Quantization Parameter

The quantization parameter is adjusted by considering the type of the current macro block and the difference between the previous quantization and the current quantization parameters. And final quantization parameter is used for encoding. Also to minimize fluctuation of picture quality and blocking artifacts, apply different adjustment methods.

Usually, another method for rate distortion optimization is carried out to determine the encoding mode for each packet and channel capacity, and also in order to minimize the overall expected end–end distortion. Also adaptive RDO model will improve the performance.

So this model-based kind for improving the approaching speed of a blind method such as the bisection method also in general adapted model fits well, and a reduced number of trail and errors can be greatly improved. Model based for approaching the critical rates is based on the monotonically non-decreasing property. Even despite being more efficient and more accurate other methods, it does not explicitly make full appreciation of the distortion rate model.

From a rate distortion theoretic point of view [9], a good coding design is very much important to find a set of encoding and decoding algorithmic to minimize the actual RD cost. And even mode selection for motion estimation will conduct based on actual RD cost in a macro block by macro block manner.

PSNR can be calculated using below Eq:

$$p(X, Y) = 10 \log_{10} \frac{255^2 \text{ nm}}{\sum_{i=1,j=1}^{n,m} \left(X_{ij} - Y_{ij}\right)^2} \tag{5}$$

Fig. 2 VP8 video output without RD

Fig. 3 VP8 video output with RD

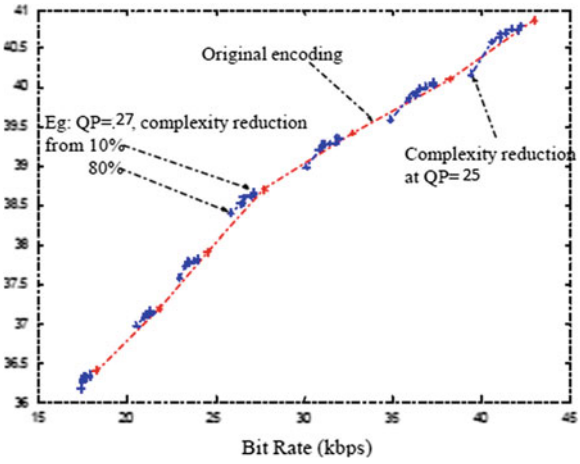


Fig. 4 Performance analysis of VP8 Codec (below)

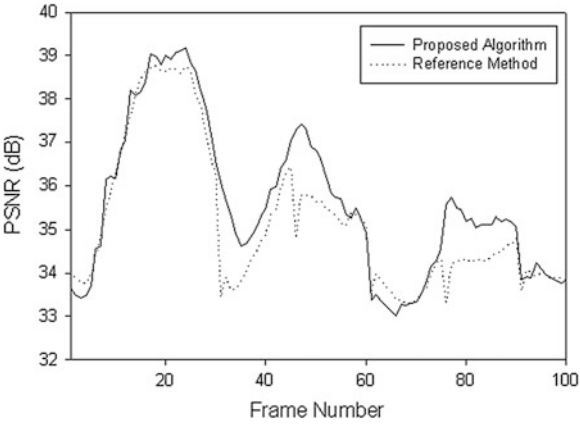


Fig. 5 PSNR for "vp80-00-comprehensive-002" (frame rate = 15 fps, target bit rate = 64 kbps)

Above Fig. 10 shows VP8 Video encoder and decoder blocks. And tested results for the are shown in Figs. 2, 3, 4, 5, 6, 7, 8, 9, whereas Figs. 3, 7 and 9 are with rate distortion algorithm. And Figs. 4, 5, 11,12 and 13 show performance results of VP8 Video codec with considering rate distortion algorithm. Table 1

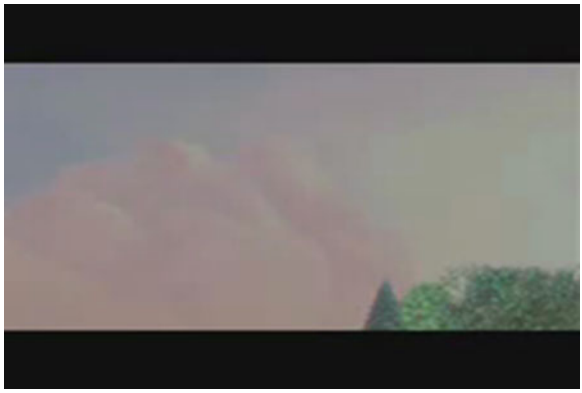


Fig. 6 VP8 video output without RD

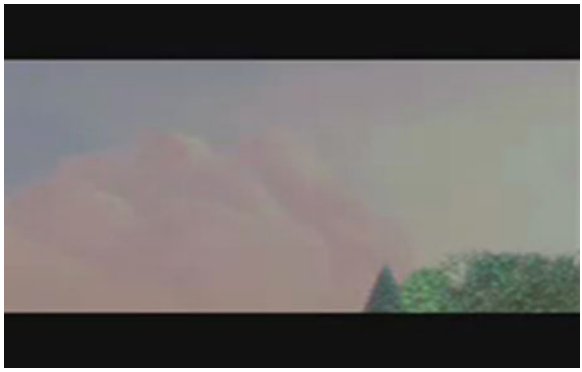


Fig. 7 With rate distortion algorithm for VP8 codec

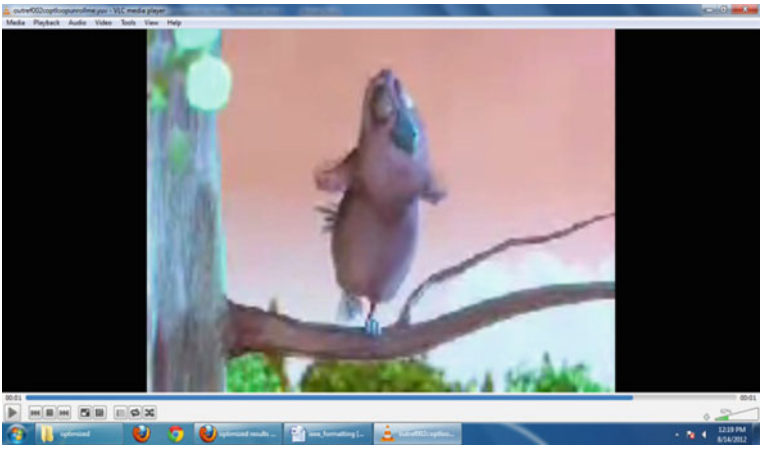


Fig. 8 VP8 video output without RD

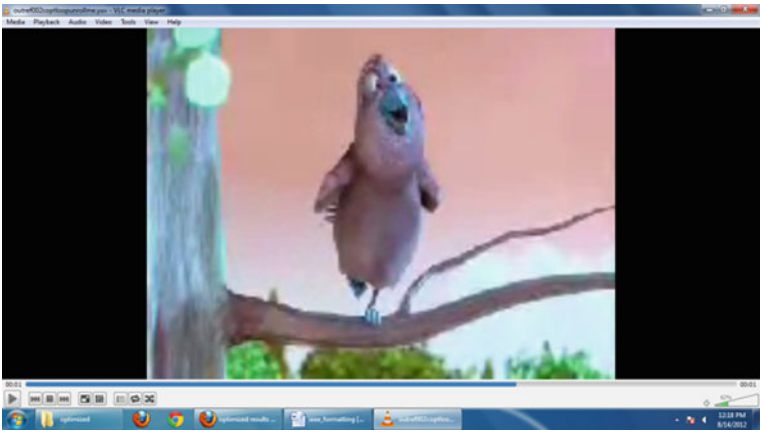


Fig. 9 With rate distortion algorithm for VP8 codec

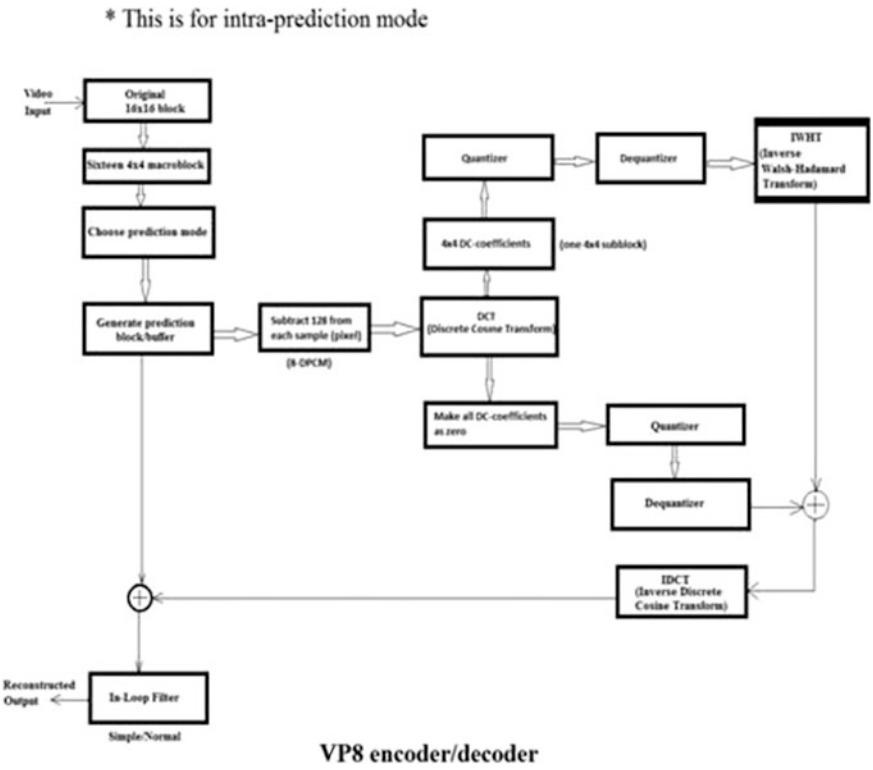


Fig. 10 VP8 Encoder/Decoder(below)

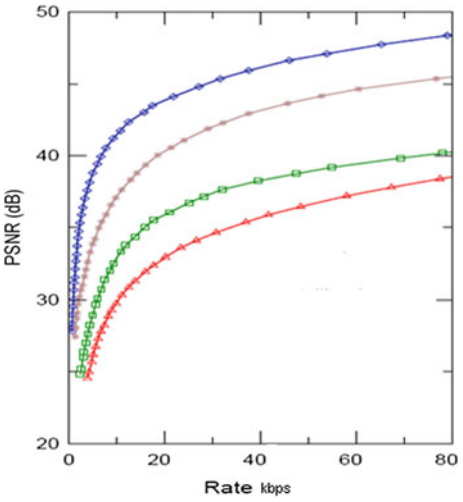


Fig. 11 Bitrates versus PSNR for different VP8 Video test streams

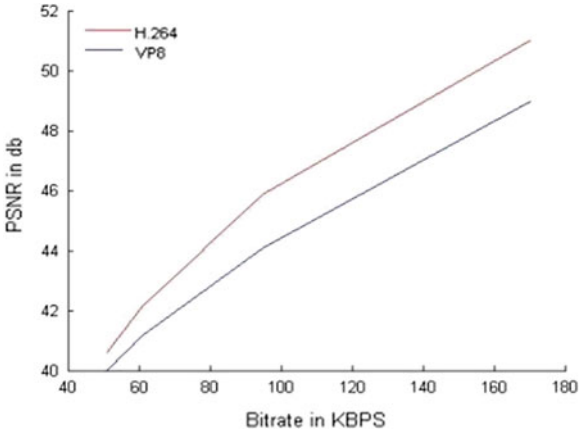


Fig. 12 Bitrates and PSNR comparison for H264 and VP8 test case 1

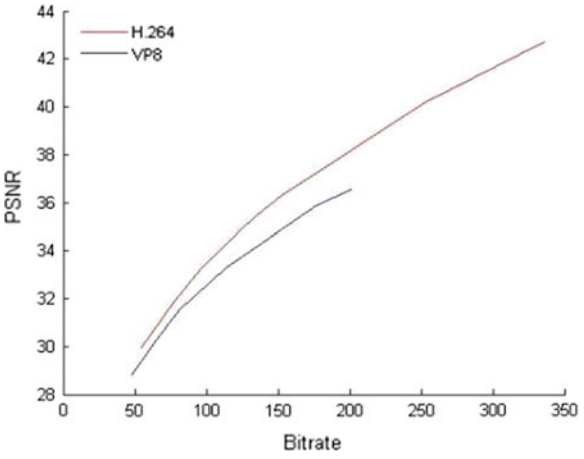


Fig. 13 Bitrates and PSNR comparison for H264 and VP8 test case 2

Table 1 Rate and distortion of encoded sequences

Sequence	PSNR	Bit rate (Kbps)
vp80-00-comprehensive-002	29.67	33.95
vp80-00-comprehensive-007	37.57	51.85
vp80-00-comprehensive-015	39.32	67.34

shows rate and distortion of encoded sequences. Here, after considering rate distortion algorithm, VP8 codec output looks like more promising compared to without using rate distortion algorithm which is shown in fig and also speed is more when considered for rate distortion algorithm for this VP8 Codec since rate distortion algorithm will do the more optimization.

5 Conclusion and Future Work

We observed that after implementing rate distortion algorithm like rate-quantization model based on the rate distortion function of the Laplacian distribution and successfully show and prove that the bit rate curve can be modeled by the logarithm function. And also our simulation results show that the introduced rate control method generates coding bits very close to target bits and provides improved coding efficiency at low bit rates.

In this along with optimization quality of the data also looks good, this shows that performance of VP8 Video codec along with rate distortion algorithm yields high-quality and optimized data. In future, there is lot of scope for improving this work like considering this rate distortion algorithm for more resolution data, in this way still we are working.

References

1. Strachan D (1996) Video compression. SMPTE J 105(2):68–73
2. Bankoski J, Wilkins P, Xu Y (2011) VP8 data format and decoding guide. http://www.ietf.org/internet-draft/draft-bankoski-vp8-bitstream-01.txt. Jan 2011
3. Berger T (1971) Rate distortion theory-a mathematical basis for data compression. Prentice-Hall, Englewood Cliffs
4. Wiegand T, Lightstone M, Mukherjee D, Campbell TG, Mitra SK (1996) Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard. IEEE Trans Circuits Syst Video Technol 6(2):182–190
5. Sullivan GJ, Weigand T (1998) Rate distortion optimization for video compression. IEEE Signal Process Mag 15(6):74–90
6. Yang E, Yu X Rate distortion optimization in H.264
7. Wang S, Rehman A SSIM-motivated rate distortion optimization for video coding
8. MPEG-2, MPEG-2 Test Model5 (TM5) Doc.ISO/IEC/JTC1/SC29/WG11/N0400, Test Model Editing Committee, April 1993
9. Kim S, Ho YS Rate control algorithm for H.264/AVC video coding standard based on rate-quantization model

Author Biographies

Basavaraju S. holds the Master degree from VTU, Karnataka, and currently pursuing PhD at PRIST University Thanjavur, India, and earlier worked at several software industries in signal and multimedia processing domain and currently working as Assistant Professor in Sapthagiri College of engineering, Bangalore.

Dr B Sivakumar holds PhD degree from Anna University Tamil Nadu and currently working as Professor and head of the Telecom dept at Dr AIT.

Congestion Control by Heuristics in High-Speed Networks Using ANN

S. V. Uma and K. S. Gurumurthy

Abstract A primary reason for Congestion in communication networks is the presence of preprocessing and postprocessing queues (buffers) in routers and switches. Congestion in the networks reduces the quality of service (QoS), hence causing delays and loss of data. Most of the existing congestion control techniques are incapable of controlling the congestion that might occur in the future, as they do not use any prediction algorithms. To overcome this inability, in this paper, a new heuristic congestion control technique is proposed to determine the present status of congestion in the network and also predict the congestion in future. The proposed heuristic congestion control technique maintains a historical database, containing the congestion control parameters such as round-trip time (RTT), average queue size (AQS), available bandwidth, and network speed. Based on this historical database, the current status of congestion is determined and also its future value is predicted in networks with the help of an AI technique using a feed-forward backpropagation neural network (FFBNN). Finally, the sending rate of the packets to the network is adjusted using these congestion control parameter values, so as to reduce congestion in the network. An optimal congestion window is designed using an optimization algorithm called evolutionary programming (EP) algorithm. Simulation results have shown that the proposed heuristic congestion control technique efficiently controls the congestion in the networks, increases the throughput by 20 %, reduces the propagation delay by 22 % when compared to the existing technique [10], and also improves the performance of the network by 20 %.

S. V. Uma (✉) · K. S. Gurumurthy
Department of ECE, Bangalore University, Bangalore, Karnataka, India
e-mail: umakeshav2000@gmail.com

K. S. Gurumurthy
e-mail: drksgurumurthy@gmail.com

S. V. Uma
ECE department, RNSIT, Bangalore, Karnataka, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_33,

Keywords Congestion control · Quality of service (QoS) · Artificial neural networks (ANN) · Round-trip time (RTT) · Feed-forward backpropagation neural network (FFBNN) · Evolutionary programming (EP)

1 Introduction

In the past decade, there has been an exponentially growing research effort in the field of communication networks. The goal is to broadcast multimedia traffic over a "uniform" communication medium [1]. Today, the Internet is growing rapidly. The networks, such as Abilene and ESNet, provisioned with a large amount of bandwidth ranging from 1 to 10 Gbps, are now extended to connect various research organizations throughout the world. However, the round-trip delay of those networks is greater than 200 ms when they run over a long distance [2]. Also, the stability of the Internet till date is maintained due to the presence of congestion control and avoidance algorithms [3]. Without appropriate congestion control mechanisms installed and running in networks, there would be ineffectual utilization of resources, eventually resulting in network collapse [4].

Congestion refers to a sudden increase in traffic in the paths of a communication network, when one or more nodes reach beyond their capacity limit, and consequently, they either drop the incoming packet or buffer them for a later transmission, which produces delay in the arrival of packets at the receiver [5]. Congestion in networks also happens due to surpass in aggregate demand when compared to the available capacity of the network. Network congestion will often increase when network speed increases. Thus, efficacious congestion control techniques are essential, mainly to deal with "bursty" traffic of today's high-speed networks [6].

In congestion control, the load produced by the network traffic is controlled in order to enhance the overall performance of the network. Based on the current load condition of the network, congestion control is done by controlling the sending rate of the data streams of each source, which not only avoids congestion but also leads to high consumption of the available bandwidth [7]. The standard TCP has been extremely successful in performing congestion avoidance and control, hence preventing severe congestion of traffic in the current low-speed networks [8]. Also, innumerable high-speed TCP variants have been proposed, including the end-to-end approaches such as High Speed TCP(HSTCP), scalable TCP(STCP), CUBIC TCP, FAST TCP, compound TCP (CTCP), TCP Illinois and the router-based approaches such as XCP and VCP. Also, some researches focus on the application-level schemes over UDP to achieve congestion control for high-speed networks, such as UDT [9].

Congestion in the networks reduces the quality of service (QoS), hence causing delays and loss of data. In this research work, a heuristic congestion control technique is proposed to reduce the congestion from high-speed networks. Here, a historical database is used, which includes the values of congestion control parameters and the network congestion levels corresponding to them. These values

are given to the feed-forward backpropagation neural network (FFBNN) to perform the training process. The same parameter values are then collected from the current network, and such values are given to the trained FFBNN, which then predicts the new congestion value. The sending packet rate will be reduced based on this congestion value from the trained FFBNN. Before performing the FFBNN process, the parameter values are checked with the user-defined threshold value. If the parameter values are greater than the threshold value, then a congestion window is designed or else, the sending packet bit rate is reduced based on FFBNN output value. An optimal congestion window is dynamically designed using the EP algorithm, and the window size values are given to the FFBNN. By the aforesaid process, the congestion in the high-speed networks is reduced.

The rest of the paper is organized as follows: Sect. 2 describes the proposed heuristic, Sect. 3 discusses the experimental setup, Sect. 4 gives a summary of the implementation results, and Sect. 5 concludes the paper.

2 Proposed Heuristic Congestion Control Technique

The role of congestion control is to control the congestion as well as to maintain the offered load below the capacity of network. Here, a method for identifying the onset of congestion, reducing it, and prediction of the congestion levels in the future in high-speed networks is being proposed. The proposed approach is shown in Fig 1. It is composed of three stages:

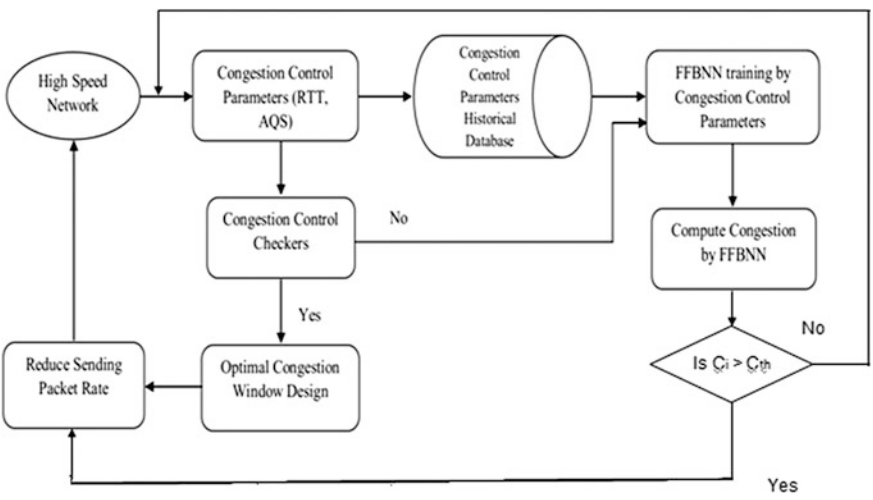


Fig. 1 Structure of the proposed heuristic congestion control technique in high-speed networks

- Congestion control by checkers.
- Congestion control by FFBNN.
- Optimal congestion window design.

The proposed heuristic technique reduces the congestion levels in the high-speed networks by using the historical parameter values of the same network kept in a database. This database contains parameter values, namely round-trip time (RTT), average queue size (AQS), transmission speed, and bandwidth. In addition to this, congestion values related to such parameters are also present in the database. Thus, the historical database is represented as $D = \{R_i, Q_i, S_i, B_i, C_i\}$, $i = 1, 2, \ldots, N$, where N denotes the number of row values in the database, R_i represents the RTT value, Q_i is the AQS, S_i and B_i represent the transmission speed and bandwidth of the networks, respectively. Also, the congestion value related to such parameters is represented as C_i, which represents the percentage of congestion in the networks. These historical database values are given to the FFBNN for performing the training process, which is briefly described in the Sect. 2.2.

Here, a wide area network with M number of nodes has been designed, and each node in the network represents any one of the LAN areas. The designed wide area network is represented as W, and the nodes in W are denoted as s_j; where $j = 1, 2, \ldots, M$. If a source node sends a packet to the destination node without any interruption, then the network performance is automatically increased. However, such high-performance values cannot be obtained by all the networks during the data transmission because the data rate of packets sent by the user may be higher than the network capacity, leading to large delay or loss of data before it reaches its target. So, QoS determining parameters like RTT, AQS, transmission speed, and bandwidth should be collected from the current network W. Based on these parameter values, the network congestion needs to be determined, and necessary steps should be taken to avoid such congestion.

2.1 Congestion Control by Checkers

When a packet is sent from the source node to destination node, the parameter values such as RTT and AQS are calculated. Then, these computed values are checked with two checkers, namely RTT checker and AQS checker. The RTT and AQS values are computed by the following formulas:

$$R_n = (w^* R_o) + ((1 - w)^* R_n) \tag{1}$$

$$Q = \left(Q_o^* \left(1 - \left(\frac{1}{2}\right)^k\right)\right) + \left(Q_c + \left(\frac{1}{2}\right)^k\right) \tag{2}$$

Equation (1) calculates the RTT values from the network W before sending the packet to the destination node. In Eq. (1), R_o and R_n represent the old and new RTT measurement values, respectively, and w is the smoothing factor, which has values between 0 and 1. The AQS is calculated by using the formula given in Eq. (2), where Q_o and Q_c are the old and current queue size values, and k represents the exponential weight factor, which is defined by the user. Both these RTT and AQS values are calculated from the network W before sending the packet p through the node s_j. After that, the checkers check these computed values and take appropriate action as described below.

2.1.1 RTT Checker

RTT checker will check whether the computed RTT value R_n is greater or less than the predefined threshold value T_R. The RTT checker has considered two rules during the comparison process,

- If R_n is greater than T_R, then the node s_j reduces the packet sending rate and the congestion window designing process is also carried out, which is described in the below sections.
- If R_n is less than T_R, the congestion control by FFBNN is executed to compute the current and future congestion value.

2.1.2 Average Queue Size Checker

The computed AQS value Q is verified by the AQS checker, which checks whether this computed Q value is greater or less than the predefined threshold value of AQS T_Q. Similar to RTT checker, this AQS checker also considers the same two rules as mentioned above during the comparison process.

2.2 Congestion Control by FFBNN

The congestion control via FFBNN is carried out by means of performing FFBNN training and testing process by utilizing the network parameter values. The parameter values from the historical database D are utilized as an input to the FFBNN for accomplishing the training process. In training phase, the four inputs that are given to the FFBNN are the network RTT value R_i, AQS value Q_i, transmission speed S_i, and the network bandwidth B_i, The FFBNN is well trained by these network parameter values and predicts the future value of congestion C_i in the network for the given input parameter values.

The FFBNN comprises of four input units, one output unit, and N_d hidden units. Initially, the input records are transmitted to the hidden layer and then to the output layer. This is termed as the forward pass of the backpropagation algorithm. Each node in the hidden layer obtains its input from the input layer, which are multiplexed with suitable weights and then summed. The input value computation function in the hidden layer is called as bias function, which is described below:

$$Z_i = \beta + \sum_{h=0}^{N_d-1} \left(w_{ih}R_{ih} + w_{ih}Q_{ih} + w_{ih}S_{ih} + w_{ih}B_{ih}\right) \tag{3}$$

In Eq. (3), R_i, Q_i, S_i, and B_i are the input values taken from the i′th row of the database. The output of the hidden layer is the nonlinear transformation of the resulting sum. The same process is followed in the output layer. The subsequent Eq. (4) represents the activation function of the output layer. The output values from the output layer are compared with the target values, and then, the learning error rate for the neural network is computed by using the Eq. (5):

$$\chi = \frac{1}{1 + e^{-Z_i}} \tag{4}$$

$$\ell = \frac{1}{N_d} \sum_{h=0}^{N_d-1} D_h^i - O_h^i \tag{5}$$

where ℓ is the learning error rate of the FFBNN, D_h^i is the desired output, and O_h^i is the actual output. The error between the nodes is transmitted back to the hidden layer. This is termed as the backward pass of the backpropagation algorithm. The reduction in error through backpropagation algorithm is depicted in the following steps:

- Initially, the weights are allocated to hidden layer neurons. The input layer has a constant weight, while the weights for output layer neurons are selected randomly. Then, the bias function and output layer activation function are calculated by using the Eqs. (3) and (4).
- Subsequently, the backpropagation error is computed for each node, and the weights are adjusted using the following Eq. (6).

$$w_{ih} = w_{ih} + \Delta w_{ih} \tag{6}$$

where the weight Δw_{ih} is changed using the relation,

$$\Delta w_{ih} = \eta . X_{ih} . E^{(\phi)} \tag{7}$$

where η is the learning rate that normally ranges from 0.2 to 0.5, and $E^{(\phi)}$ represents the BP error. The bias function, activation function, and BP error calculation process are continued until the BP error gets reduced to $E^{(\phi)} < 0.1$. If the BP error reaches the lowest value, then the FFBNN is well trained by the input

network parameter values for performing the congestion control. The trained FFBNN can now provide appropriate future congestion values for the corresponding input values, and the congestion control is done effectively by this FFBNN.

After the FFBNN training process, similar input network parameter values are gathered from the network W. Based on the input parameter values, the output congestion value C_i is obtained. Then, based on this C_i value, the sending packet (p) rate is reduced by the sender s_j. Subsequently, the reduced packet p is sent to the corresponding receiving node travelling through many intermediate nodes. Thus, a congestion checking process is carried out in each intermediate node before the packet p arrives at the receiving node.

2.3 Optimal Congestion Window Design

The optimal congestion window designing process is carried out when R_n is greater than T_R and Q is greater than T_Q. We have used an optimization technique called evolutionary programming (EP) for selecting an appropriate congestion window size. The congestion window has an initial value 1 and a maximum value c_s. The designed optimal congestion window is updated constantly. This process is repeated until the congestion window reaches the maximum window size c_s. Once the congestion window reaches the maximum value c_s, then again the EP algorithm is executed to select an optimal congestion window size.

2.3.1 Evolutionary Programming

EP is a popular optimization approach that depends on the natural generation. It involves random number generation at the initialization process. EP begins with the definition of the problem to optimize and creating an objective function to compute the possible candidate solutions. More precisely, objective function is the way of identifying the individual that gives the best result. To obtain the optimal solution, that is, best congestion window size, the EP performs four crucial operations: (1) Chromosome generation, (2) Computing fitness function, (3) Mutation, and (4) Termination.

3 Experimental Setup

The proposed heuristic congestion control technique is implemented with MATLAB version 7.12. The sample test network for the proposed technique is given in Fig. 2.

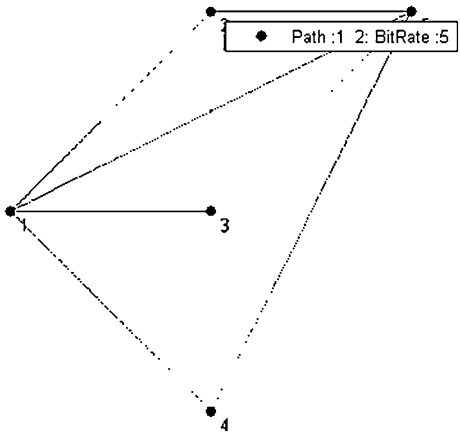


Fig. 2 Proposed test network with five nodes

The proposed congestion control technique utilizes a test network for packet transmission, as shown in Fig. 2. This network is developed with five nodes (from 1 to 5) and a number of edges (i.e. channel) to connect those nodes.

4 Results and Discussion

Now, the proposed technique is analyzed with different size of packets that are transmitted at different time intervals adhering to multimedia traffic requirements. While transmitting such variable size packets in high-speed networks, congestion may happen, hence increasing the transmission delay. This new technique could control the congestion in high-speed networks, simultaneously reducing the transmission delay. The average queue length and RTT values for the proposed and existing techniques [10] for different time intervals are shown in Fig. 3. The normalized threshold levels for the same were taken as 1 and 0.1.

A high value of propagation delay indicates that the network has low performance in congestion control because it takes more time to reach the destination node, whereas a lower value indicates that the network is efficient in transferring the data within less time and the congestion control is more effective. The performance of the proposed technique is validated by comparing the congestion control process with a recently developed earlier congestion control technique [10]. The comparison result of the proposed and existing congestion control techniques [10] in terms of their propagation delay and throughput are shown in Fig. 4.

In Fig. 4, it can be seen that the packets are sent with varying size from sender to receiver at different time intervals. The sending and receiving times of each packet are noted, to calculate the propagation delay of the current network in both congestion control methods. When the packet size is increased, the propagation delay values are also increased in the existing method. But this increase is

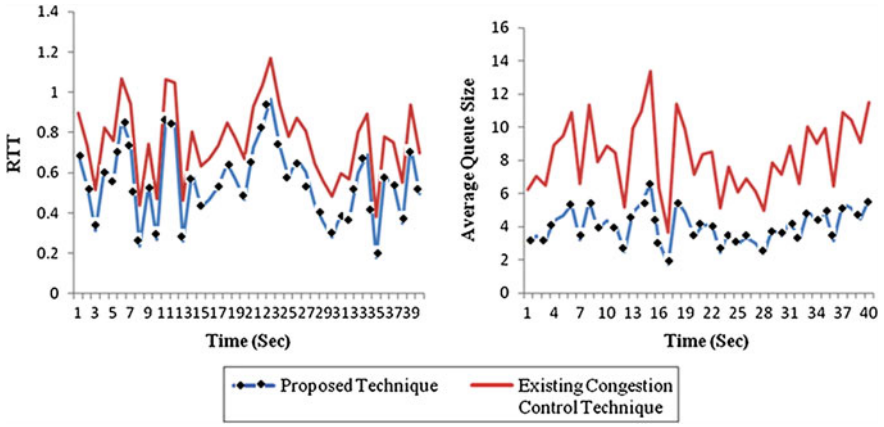


Fig. 3 Proposed and existing congestion control techniques. (**a**) RTT and, (**b**) AQS values at different time intervals

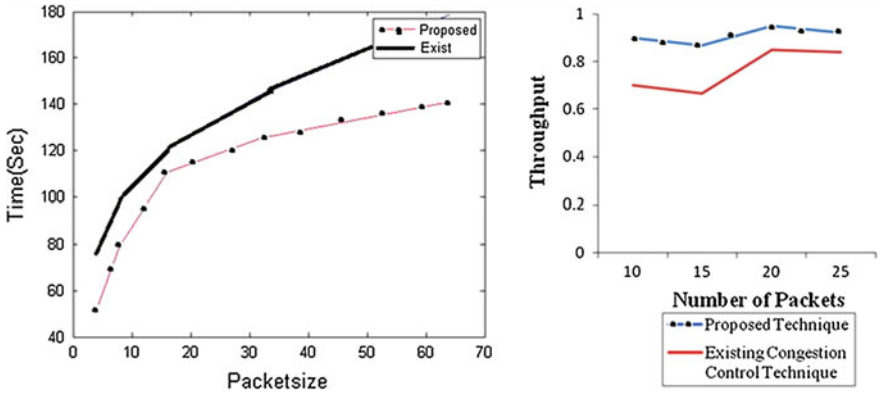


Fig. 4 Comparison of proposed and existing congestion control technique performance in terms of propagation delay and throughput

proportionately lower in the newly proposed method. The significant network parameters such as RTT, average queue length, transmission speed, and network bandwidth from the historical database have made our proposed heuristic congestion control technique more precise in controlling the congestion in the high-speed networks with varying packet rate.

The normalized threshold level of throughput was 0.5, and delay was 40.

Figure 4 also shows the performance of the proposed and existing congestion control techniques in terms of the throughput values as simulated for the test network. The throughput values are calculated from the network by calculating the successful delivery of packets at the receiver. It can be seen that the proposed technique transfers the packets from sender to receiver with more number of successful packet deliveries at the receiver when compared to the existing

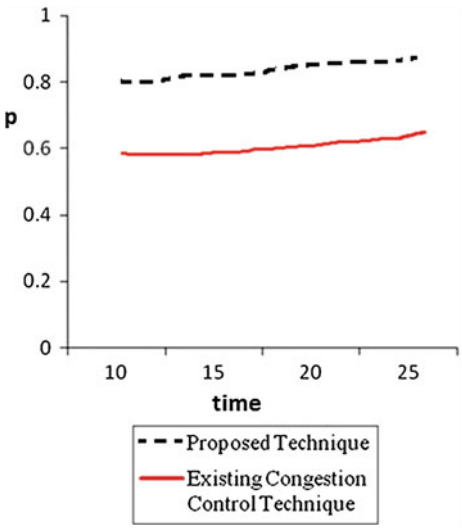


Fig. 5 Comparison of the proposed technique with the existing technique in terms of overall performance *p*

technique, without loss in sending packet rate. Also, the proposed congestion control technique has given high throughput value than the existing method [10].

With this technique, we were able to achieve better system performance (*p*) in terms of increased throughput, speed and reduced packet losses, and the results are as shown (Fig. 5).

5 Conclusion

In this paper, the proposed heuristic congestion control technique was depicted thoroughly with implementation results. In the proposed methodology, the congestion control was done using checkers and via the trained FFBNN with network parameter values collected in the historical database. The parameters estimated from the current network were given to the FFBNN to find the congestion level in the current network and predict its value in the future. An optimal congestion window was designed based on the RTT parameter value using the EP algorithm, which makes the proposed congestion control technique more expedient to perform the congestion control process in the high-speed networks. All these steps have enhanced the efficacy of the proposed congestion control technique by at least 20 % when compared to the existing technique [10]. The simulation results have revealed that the performance of proposed heuristic technique was superior compared to prior techniques. When compared to the prior technique [10], the congestion control performance of proposed technique has offered better performance in transmitting the packets with higher speed, therefore being more potent in high-speed networks and also more pertinent for transferring data in less time.

References

1. Xu L, Harfoush K, Rhee I (2004) Binary increase congestion control for fast, long distance networks. In: IEEE conference on INFOCOM, pp 1–13
2. Mascolo S (1997) Smith's principle for congestion control in high speed ATM networks. In: 36th IEEE conference on decision and control, San Diego, CA, USA, vol 5, pp 4595–4600
3. Bansal D, Balakrishnan H (2001) Binomial congestion control algorithms. In: IEEE INFOCOM, Anchorage, AK, pp 1–10
4. Haider A (2004) Improved congestion control for packet switched data networks and the internet. Technical Report
5. Jasem HN, Zukarnain ZA, Othman M, Subramaniam S (2009) The new AIMD congestion control algorithm. World Acad Sci Eng Technol 50(130):769–774
6. Ahmad S, Mustafa A, Ahmad B, Bano A, Hosam AS (2009) Comparative study of congestion control techniques in high speed networks. Int J Comput Sci Inf Secur 6(2):222–231
7. Reddy S, Reddy LC (2008) A survey on congestion control mechanisms in high speed networks. Int J Comput Sci Netw Secur 8(1):187–195
8. Liu S, Basar T, Srikant R (2006) TCP-Illinois: a loss and delay-based congestion control algorithm for high-speed networks. In: 1st international conference on Performance evaluation methodologies and tools, Pisa, Italy
9. Dangi R, Shukla N (2012) A new congestion control algorithm for high speed networks. Int J Comput Technol Electron Eng (IJCTEE) 2(1):218–221
10. Kaura S, Vatsa AK (2011) A novel architecture and mechanism for congestion control in high speed network. Int J Next-Gener Netw 3(1):21–35

Low-Power Sum of Absolute Difference Architecture for Video Coding

D. V. Manjunatha and G. Sainarayanan

Abstract Motion estimation is the power-hungry block in the video compression system. The sum of absolute difference (SAD) is the most repeated operation in block matching algorithms of the motion estimation subsystem. This paper identifies low-power full adder that is used in the proposed SAD architecture. The proposed SAD method consumes low power as compared to the existing SAD architectures. The designs are implemented using ASIC flow (synthesis) using the Cadence tools and the simulations are done using the Modelsim simulation tools. The synthesis is done using TSMC 180 nm and 45 nm technology libraries; the synthesis is mainly focused on the power consumed, area acquired and latency. This paper implements 4 × 4 (16 input samples) SAD algorithm, and the proposed architecture shows 26.3 % improvement in leakage power, 67.58 % improvement in dynamic power, and 65.6 % improvement in the total power with the 180 nm technology and 13.9 % improvement in the leakage power, 62 % improvement in the dynamic power, and overall 60 % improvement of total power with the 45 nm technology, as compared to the existing architecture.

Keywords SAD · MAD · WED · NCC · Leakage power · Dynamic power etc.

D. V. Manjunatha (✉)
Don Bosco Institute of Technology, Bangalore, India
e-mail: dvmanjunatha@gmail.com

G. Sainarayanan
New Horizon College of Engineering, Bangalore, India
e-mail: sai.jgk@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_34,

1 Introduction

Video compression is one of the techniques in video processing system, which is used to reduce resource usage. The two primary challenges addressed during video compression are:

- Limited network bandwidth.
- Limited storage capacity.

Hence, the two important metrics of a video encoder are low computational complexity and low-power hardware implementation. In the present-day world, compression ratio plays a major role in the field of video processing. The motion in the video scene will reduce the efficiency of the compression ratio. The efficiency of the compression ratio can be increased by exploiting the similarities between the video frames. The simple metric system is the SAD algorithm, where the absolute differences between the corresponding elements are added up. In this work, we first identify the low-power architecture at the level of 1-bit addition (full adder); here, five full-adder architectures are synthesized based on which architecture is giving the low-power solution. Such an adder is used in the ripple-carry parallel adder, and this ripple parallel adder is used in the SAD architecture for achieving the low power in video coding (compression).

This paper is organized as follows: Sect. 2 describes the various existing gate level realizations of 1-bit binary full adder. The higher adder architectures are mentioned in the adder exploration. Section 3 gives details about the existing and proposed SAD architectures. Section 4 describes the results and discussion; finally, we conclude in Sect. 5.

2 Adder Architectures

The full adder is the most basic element of many critical data paths of the digital arithmetic circuits and digital signal processors. A 1-bit full adder is basically a combinational logic circuit that performs binary addition operation on 3 single bit binary numbers and produces two outputs called sum and carry; there are many efficient full-adder architectures in the literature, namely (1) 2 EXNOR gate and 1 MUX architecture, (2) 1 EXNOR- and 2 MUX-based full adder, (3) EXOR, AND, and OR gate architecture, (4) EXOR and Nand gates architecture (Figs. 1, 2, 3, 4).

2.1 Adder Architecture Exploration

The addition of two binary numbers is the fundamental and most often used arithmetic operation on digital signal processors. Hence, binary adders are critical

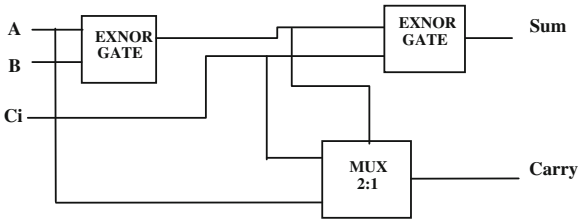


Fig. 1 EXNOR and MUX architecture of full adder

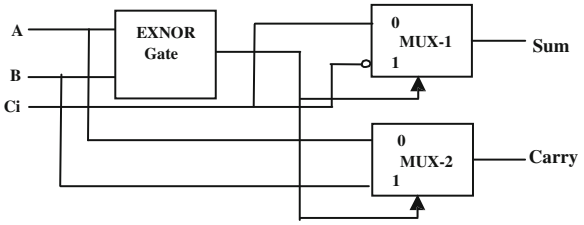


Fig. 2 1 XNM EXNOR- and MUX-based full adder

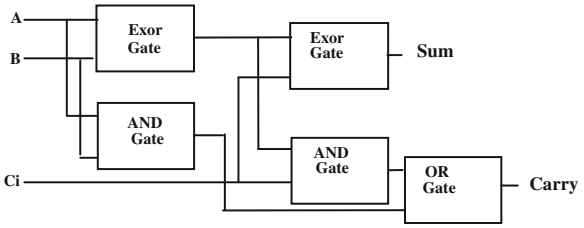


Fig. 3 EXOR, AND and OR gate full adder

building blocks in most of the systems. Key challenges addressed in the adder design are the carry propagation operation involving all operand bits. Various different architectures for binary addition have been proposed, covering a wide range of performance goals. But, the suitability of adder architectures for cell-based design and hardware synthesis has been very important for the low-power addition. Based upon the delay, there are two fundamental adder architectures:

Ripple Carry It is the simplest architecture for an n-bit adder; intermediate carries are generated sequentially. It has the smallest area and the longest delay, and consumes lowest power.
Carry Look Ahead Here, the carry is generated in parallel to reduce the processing delay. Hence, it is faster than ripple-carry architecture. It consumes more area and power.

The ripple-carry adder architecture is shown in Fig. 5.

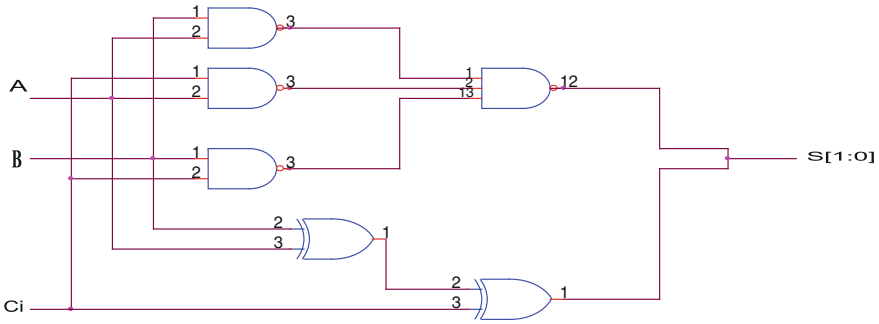


Fig. 4 EXOR and Nand gates full adder

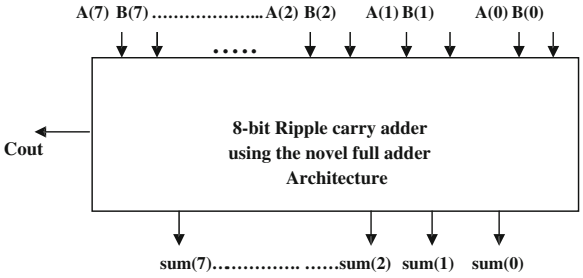


Fig. 5 8-bit ripple-carry adder

3 Sum of Absolute Difference Architecture

Sum of absolute difference (SAD) is primarily used to detect motion in the motion estimation subsystem of video compression system [1]. It requires only two basic mathematical operations addition and shifting. Wong et al. [2] describe this parallel hardware implementation of the SAD operation in field-programmable gate arrays (FPGAs). A novel SAD 16 unit that performs a 16 × 1 SAD operation is proposed by Hisham et al. [3]. The SAD algorithm is the simplest metric that considers all the pixels in the block for computation and also separately, which makes its implementation easier and parallel. Due to its simplicity, this algorithm is one of the fastest and can be used widely in motion estimation and object recognition. Typical steps involved in fully parallel SAD architecture are:

Perform absolute difference of all the pixels (of a block of video).
Perform Sum of all the absolute differences.
Select block with minimum difference value.

The SADs are implemented using the following architecture (Fig. 6):

Motion estimation uses SAD [4, 5]. The SAD measures the absolute difference between each pixel of the template image (previous) and the search (current) image,

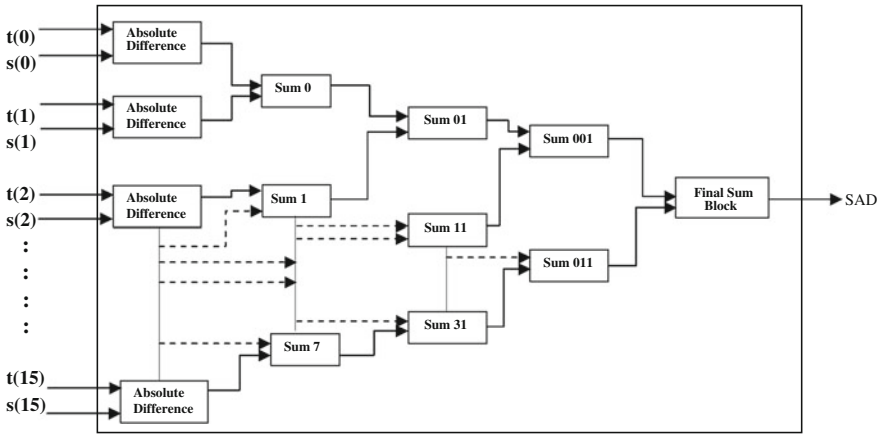


Fig. 6 4 × 4 SAD architecture

and it finds lot of applications in the hardware implementation of image/video processing blocks in the video codec such as object recognition, motion estimation.

Like SAD, there are various metrics available in the literature, and some of them are mean absolute difference (MAD), weighted Euclidian distance (WED), and normalized cross-correlation (NCC). Usually, SAD is used because of its simplicity in the hardware implementation, another alternative to SAD is MAD, WED is used in the medical image motion estimation as used in comparing [6] the two patterns of IRIS, and NCC [7] is used to compare the acquired image with a database images for good match.

In sum of absolute differences first absolute difference between the pixels of template image block and the search image block and these are added up to find the best match. Here, the SAD implementation is done in two ways: One is the existing sum of absolute algorithm, which uses the normal EXOR gates and the Nand gates in its adder architecture; and the other is sum of absolute algorithm, which uses the EXOR, And and OR gates; the latter gives the best results in terms of power as shown in Sect. 4.

4 Results and Discussion

The existing and proposed parallel SAD architectures were implemented in ASIC methodology. The architectures are modeled using Verilog, functionally verified using Modelsim, and synthesized using RTL Compiler. The results are tested using 180 and 45 nm technology libraries of Cadence EDA tools; the results are presented at the levels of 1-bit adder, 8-bit ripple-carry adders, and 4 × 4 SAD levels of the existing and the proposed architectures as below (Fig. 7 and Tables 1, 2):

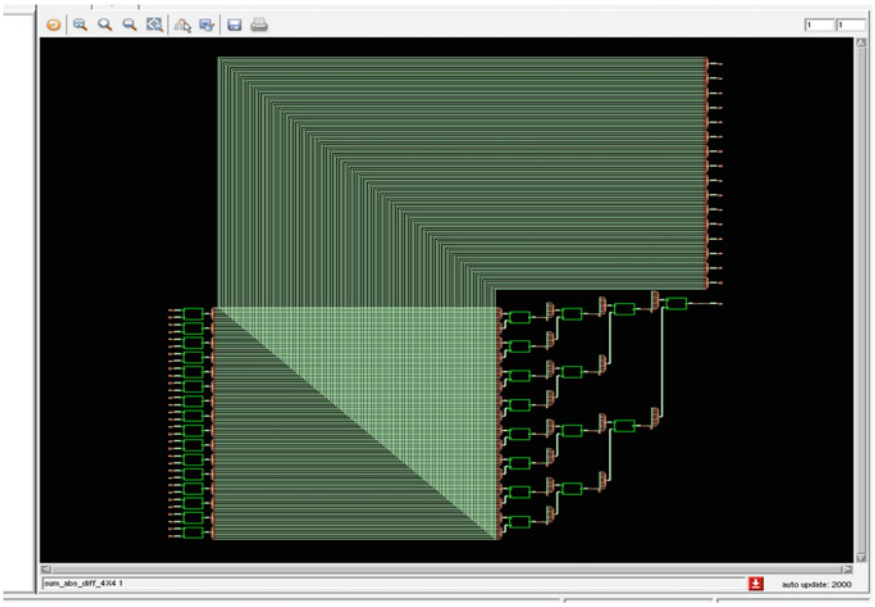


Fig. 7 Synthesis snapshot of proposed 4 × 4 sum of absolute difference

Table 1 Full-adder synthesis results

Particulars	Tech (nm)	LP nW	DP nW	Total power nW	T in Ps	A in Sq microns
EXOR and Nand gates existing	180	96.144	492.446	588.590	157	24
	45	20.643	279.643	300.551	528	18
Proposed EXOR, AND, and OR gates	180	51.390	314.024	365.414	276	19
	45	17.528	84.534	102.062	637	14

Table 2 Synthesis results of SAD 4×4

Particulars	Tech (NM)	LP nW	DP nW	Total power nW	T in Ps	A in Sq microns
SAD 4 × 4 existing	180	25149.932	500144.707	525294.390	04490	6397
	45	10387.000	323237.000	333624.846	10416	1742
SAD 4 × 4 proposed	180	18551.848	162128.067	180679.915	02666	5492
	45	08934.580	122501.911	131436.491	06143	1560

Note LP leakage power, *DP* dynamic power, *T* delay, *A* area

5 Conclusion

In this paper, we have proposed a low-power SAD architecture that performs 4 × 4 SAD operation. The existing and proposed SAD architectures have been simulated using Modelsim version 6.5 synthesized using TSMC 180 nm technology and 45 nm technology using the Cadence RTL compiler, and results are compared with existing architecture in terms of leakage power, dynamic power, total power, latency, and the area. The results show 26.3 % improvement in leakage power dissipation, 67.58 % improvement in dynamic power dissipation, and 65.6 % improvement in the total power with the 180 nm technology, and 13.9 % improvement in the leakage power, 62 % improvement in the dynamic power, and overall 60 % improvement of total power with the 45 nm technology. This shows that the proposed SAD architecture reduces the power consumed. As future work, the methodology can be extended for further improvement in the power consumption and for more input samples.

Acknowledgments The authors would like to thank Dr. Venkatesvarlu and Dr. Siva Yellampalli of UTL Technologies for their support in the laboratory, and the authors are thankful especially to the management, Don Bosco Institute of Technology, Bangalore, for their constant encouragement and support.

References

1. Vanne J, Aho E, Hamalainen TD, Kuusilinna K (2006) A high-performance sum of absolute difference implementation for motion estimation. IEEE Trans Circuits Syst Video Technol 16(7):876–883
2. Wong S, Vassiliadis S, Cotofana S (2009) A sum of absolute differences implementation in FPGA hardware. Int J Electr Comput Eng 4:9
3. Hisham C, Komal K, Mishra AK (2009) Low power and less area architecture for integer motion estimation. Int J Electr Comput Eng 4:9
4. Wong S, Vassiliadis S, Cotofana S (2002) A sum of absolute differences implementation in FPGA hardware. Euromicro Conference, pp 183–188
5. Yufei L, Xiubo F, Qin W (2007) A high-performance low cost SAD architecture for video coding. IEEE Trans Consum Electron 53(2):535–541
6. www.compuphase.com: Visited, 18th Sep 2012
7. Nadir ND, Samir BB, Josefina Janier J (2011) Article: fast template matching method based optimized sum of absolute difference algorithm for face localization. Int J Comput Appl 18(8):30–34 (Published by Foundation of Computer Science. BibTeX)

An Integrated Method for Classification of Indus and English Document Images

A. S. Kavitha, P. Shivakumara and G. Hemantha Kumar

Abstract Unlike classification of documents with plain background and high resolution, classification of historical document, namely Indus script written on stone, wall, and palm leaves is challenging because of sources on which script is written and various handwriting, which causes noise, distortions, background variations, multisized text, and multifont. In this paper, we propose an integrated method that has two-stage algorithms to classify Indus and English from the South Indian documents. The first stage uses morphological operations and thinning on Canny of the input image to study the straightness and cursiveness of thinned components to classify the Indus document from the South Indian and English. The second stage proposes region growing and thinning to study the straightness and cursiveness of the thinned edges to classify the English from the South Indian documents. We select 100 documents for each script in total 600 documents to evaluate the performance of the method. The comparative study with existing method shows that the proposed method outperforms the existing method in terms of classification rate.

Keywords Indus document classification · Thinning · Straightness · Cursiveness · Region growing · English document classification

A. S. Kavitha (✉) · G. Hemantha Kumar
Department of Studies in Computer Science, University of Mysore,
Mysore, Karnataka, India
e-mail: kavitha_sanjay_as@yahoo.co.in

G. Hemantha Kumar
e-mail: ghk.2007@yahoo.com

P. Shivakumara
School of Computing, National University of Singapore, Singapore, Singapore
e-mail: shiva@comp.nus.edu.sg; hudempsk@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_35,

1 Introduction

It is noted that there is a great demand for the advanced technology which can retrieve the documents from multiscript and multilingual databases efficiently and accurately [1–5]. This is because of dramatic change in the development of technologies in the recent years especially devices like camera and videos to capture the document images. Since multilingual country like India that has 22 official languages, a document containing multiscripts is quite common. This makes hard for the current OCR to recognize and understand the document because the current OCR is designed to recognize the document containing only one script but not for multiscript in a single document. Therefore, to enhance the performance of the current OCR, it is necessary to identify the scripts in the document having more than two scripts before selecting an appropriate OCR [1–5]. Plenty of research has been carried out for identification of script at line, block, and word level in the past [6–8]. However, these methods consider scripts such as English, Chinese, Japanese, Arabic, Hindi, Devanagiri, Bangla, and South Indian scripts—Kannada, Tamil, Telugu, and Malayalam but not the document like Indus for script identification. The reason might be non-availability of this document as it is very old document used in the period of 2500–3000 BC and another reason might me the complexity involved in Indus document processing. These scripts have been used for communication in olden days. It is noticed that these scripts give more information for the eyes but not for ears. In addition, at that time, there was no pen to write, scripts are engraved on stones and metals. They were using some kinds of sticks with some ink to write the script. As a result, the Indus document contains variety of handwriting styles and all character of different shapes associated with animals or visual pictures. In this way, the Indus document is different from the normal documents and hence Indus document classification is challenging and it is useful for Archeological department, Karnataka, India where huge number of Indus documents is kept without labeling and translating due to non-availability of specialized experts and laborious manual work to preserve the documents permanently. According to interaction with them, classification and recognition of Indus document helps them greatly in preserving old scripts permanently and translation helps for future study.

Therefore, in this work, we consider the problem of classification of Indus script from English and South Indian documents by integrating two methods. Method-1 classifies Indus document based on morphological operation, straightness, and cursiveness of the components in the document, and Method-2 classifies English from the South Indian documents based on region growing, straightness, and cursiveness of the components in the documents. The integrated method classifies both Indus and English documents from the South Indian documents.

2 Related Literature

In document analysis, identification of scripts in a document having plain background and high resolution is a familiar problem. Plenty of methods are available in the literature for identification scripts such as English, Chinese, Arabic, Japanese, Devanagari, Bangla, and South Indian scripts (Kannada, Telugu, Tamil, and Malayalam). An overview of script identification methodologies based on structure and visual appearance is presented in [1]. It is noted from this review that the proposed methods work well for plain background and distortion-free images but not the image like Indus where we have unpredictable background and shape of the characters. The rotation invariant features for automatic script identification are proposed by Tan [2] based on Gabor filter. This work considers 6 scripts for identification. Busch et al. [3] have explored the combination of wavelet and Gabor features to identify the scripts. However, these methods expect a large number of training samples to achieve good classification rate. Shijian and Tan [4] have proposed a method for script identification in noisy and degraded document images based on document vectorization. Although the method is tolerant to various types of document degradations, it may not perform well for Indus document since it requires shape of the characters. Texture features based on Gabor filter and discrete cosine transform are used for script identification at the word level in many papers [5–7] where the methods expect high-contrast document for segmentation of words. Similarly, a study of character shape for identifying scripts is proposed in [8, 9]. These methods perform well as long as segmentation works well, and the character shape is preserved. Online script identification is addressed in [10] where the spatial and temporal information is used to recognize the words and text lines. In the context of Indian language document analysis, the automatic separation of text lines from multilingual documents by extracting features from profiles, water reservoir concepts, and contour tracing is presented [11]. The Gabor filter–based technique and direction distance histogram classifiers are proposed by Choudhary et al. [12] for Indian script identification. Chanda and Pal [13] have proposed an automatic technique for word-wise script identification of Devanagari, English, and Urdu scripts from single document. Zhou et al. [14] have proposed method based on connected component profile analysis for Bangla and English script identification. Zone-based structural features for script identification are proposed by Gopakumar et al. [15]. Aithal et al. [16] have proposed a method to identify trilingual document containing English, Hindi, and Tamil at line level based on projection profile analysis of the text lines. These methods are good for the segmented text lines but not for whole document like Indus where we cannot segment the text lines due to background complexity. In the same way, Rani et al. [17] have proposed comparative analysis of Gabor and discriminating feature extraction techniques for script identification. This work considers Punjabi and English scripts for classification, and it is shown that selecting an appropriate feature extraction and classifier improves the accuracy of the script identification. However, this work focuses on normal documents but not like Indus documents.

Padma and Vijaya [18] have proposed a method based on connected component analysis and zonalization of the text lines for identification of English, Hindi, and Telugu documents. However, this work considers synthetic images rather than real scanned images for testing the method performance. Recently, composite script identification and orientation detection for Indian text images is proposed by Ghosh and Chaudhuri [19]. This method considers eleven scripts for identification purpose. The features presented in this work are derived from the connected component analysis. These features are good only if the connected components preserve their shapes.

Based on literature review, it is observed that most of the papers used Roman and Devanagiri as the common scripts and a few papers consider other scripts for identification purpose [6]. In addition, the main focus of these methods is that the identification of scripts in documents with plain background and high contrast but not the scripts like Indus and distorted document. To the best of our knowledge, none of the papers addressed the problem of Indus script identification. Thus, there is a great demand for Indus script classification. Hence, we propose a new integrated method to classify Indus and English scripts automatically from the South Indian documents. We propose two methods that are morphological based and region growing based to solve classification problem of Indus and English from the South Indian documents. The method introduces new features to study the behavior of the components in the documents for classification.

3 Proposed Integrated Method

The reason to propose integrated method is that when we observe the Indus document shown in Fig. 1a, it is realized that identifying the script in the Indus document is hard compared to normal documents due to animals-like objects are present in the noisy background along with Indus text. It is evident from the Canny of the input image (Fig. 1a) shown in Fig. 1b that we cannot differentiate what is foreground (Indus text) and background. In order to find solution to such a complex problem, we propose integrated method which consists of two methods to divide the complexity of the problem. It is noticed from the images shown in Fig. 1c–j that the amount of cursiveness in Indus is more than the amount of cursiveness in the South Indian documents and the amount of cursiveness in the South Indian documents is more than the amount of cursiveness of English document. This observation makes us to propose a method based on straightness and cursiveness of the edge components in the Canny of the corresponding input document images. To extract cursiveness in Indus documents, we propose Method-1 based on morphological operation and thinning to estimate the amount of cursiveness in terms of percentage of the cursive components in the edge images. Method-2 based on region growing and thinning to estimate the amount of straightness in terms of percentage of straight components in the edge images extracts more straightness in the English than South Indian documents. We are

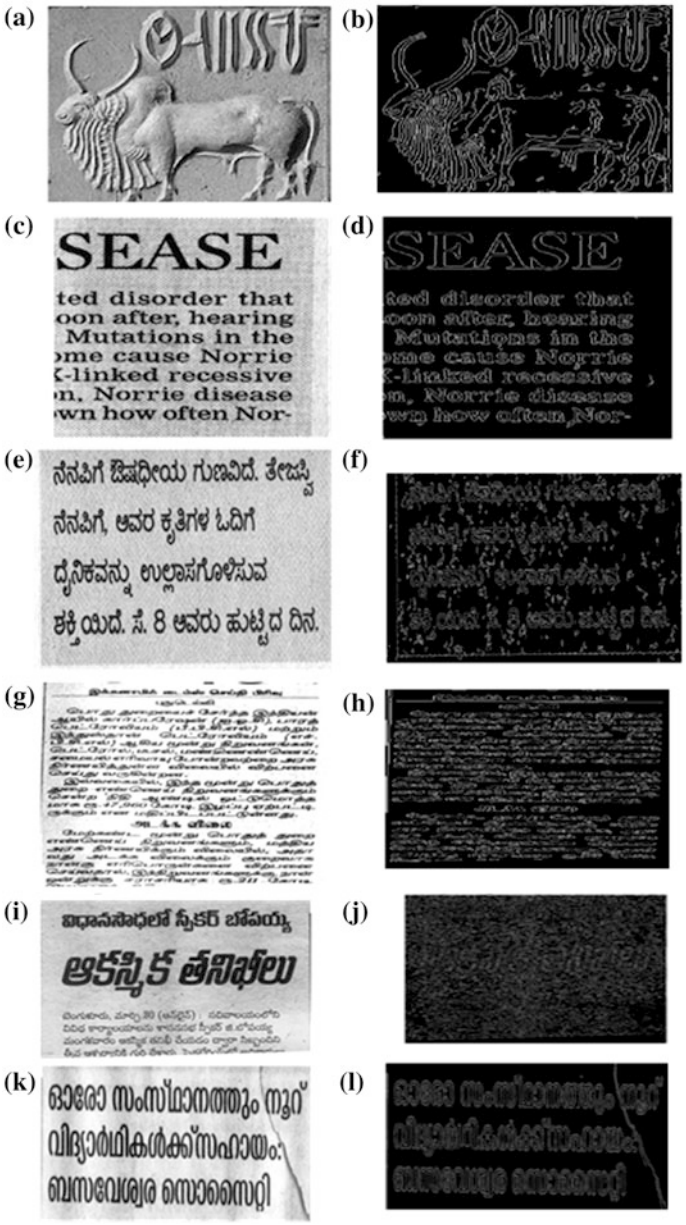


Fig. 1 Complexity involved in classifying Indus from English and South Indian documents. **a** Indus document, **b** Canny image, **c** English document, **d** Canny image, **e** Kannada document, **f** Canny image, **g** Tamil document, **h** Canny image, **i** Telugu document, **j** Canny image, **k** Malayalam document, **l** Canny image

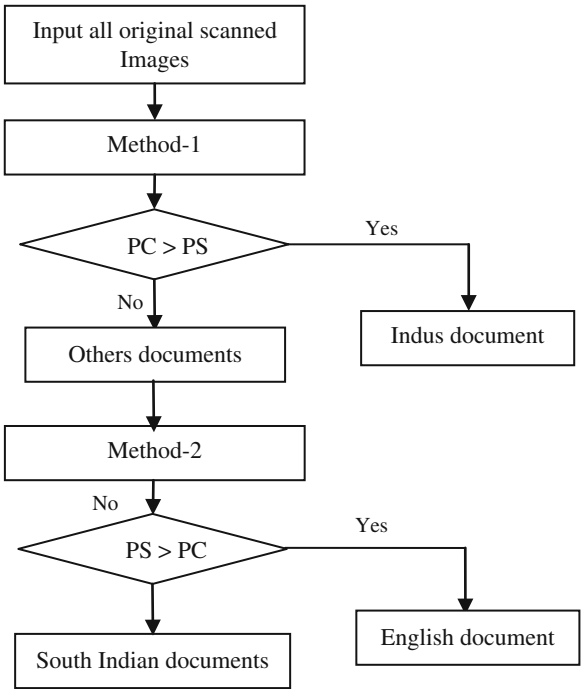


Fig. 2 Flow diagram of the integrated method

inspired from the work presented in [20] for script identification at text-line level in video by studying straightness and cursiveness of thinned text lines, where the method has been used for identifying English and Chinese documents but not like Indus documents. This splits; therefore, we propose integrated method to classify the Indus and English documents from the South Indian documents. The flow diagram of the integrated method is shown in Fig. 2 where Method-1 allows more tolerance to extract the more cursiveness in the Indus document compared to South Indian documents and Method-2 extracts the more straightness in the English document compared to South Indian documents. In Fig. 2, PC denotes percentage of cursiveness and PS denotes Percentage of Straightness.

3.1 Method-1 for Indus Document Classification

It is noted from the Edge map of Indus, English, Kannada, Tamil, Telugu, and Malayalam shown in Fig. 1 where edge components in Indus are more complex and cursive than edge components in English and South Indian scripts. This is due to the presence of animals-like objects with noisy background in case of Indus documents and due to the presence of modifiers in case of South Indian documents compared to English documents. To extract these observations and to study the local information, we analyze the feature of each edge component to study the

straightness and cursiveness of the components. The Method-1 performs dilation using morphological operation to group the edge components which are close to each other as shown in column 1 in Fig. 3 for the corresponding document images, where each word in the line is grouped in English and South Indian documents due to clear space between the text lines while edge components are grouped in irregular passion in Indus document due to the background complexity. The method converts dilated components to single-pixel-width components using thinning algorithm to study the cursiveness and straightness of the edge components in each image as shown in column 2 in Fig. 3. For each thinned-edge component, the method determines center point, say centroid (C), and it checks whether this centroid falls on its or not as shown in column 3 in Fig. 3 where the centroids are marked when they satisfy the condition of straightness property. The centroid is defined as the sum of x coordinates of edge components divided by the number of coordinates (C_x) and the sum of y coordinates of the edge divided by the number of coordinates in the edge (C_y). If the centroid (C_x, C_y) falls on its edge within certain range, then the edge components are considered as straight-edge components else cursive-edge components. Then, the method finds percentage of cursive components and straight components for each image. If the PC is more than the PS, then the whole document is considered as Indus document else other documents (English and South Indian documents). The PS/PC of component is defined as the number of straight/cursive components divided by the total number of respective components in the document multiplied by 100. Mathematically, the centroid, straightness, and cursiveness properties are defined as follows. Let $X = \{x1, x2, \ldots, xn\}$ and $Y = \{y1, y2, \ldots, yn\}$ be the sets of X and Y co-ordinates, respectively, of the edge components. The center of edge component (C_x, C_Y) is computed as defined in Eq. (1) where

$$C_x = \frac{1}{n}\sum_{i=1}^{n} xix_i \quad \text{and} \quad C_y = \frac{1}{n}\sum_{i=1}^{n} xiy_i \tag{1}$$

where n is the number of coordinates in the edge component. The condition to define the straightness (dist) is as in Eq. (2). If the $|\text{dist}| \leq \pm 1$, then the edge component is considered as a straight component else it is a cursive component.

$$|\text{dist}| = \text{SQRT}\left(\text{abs}\left(C_x - X^1\right)2 + \text{abs}\left(C_y - Y^1\right)\right) \tag{2}$$

where $(X^1, Y^1) = \left(C_x + A,\ C_y + B\right)$ and $A,\ B = \{-1 \ldots 1\}$.

It is noted from the column 3 in Fig. 3 that the marked centroids are less in Indus document compared to the centroids marked in English and South Indian documents. This shows that Indus document produce more cursiveness compared to the cursiveness in the other documents. This is because Indus document has more cursive-edge components due to background complexity and irregular-shaped characters while the edge components in English and South Indian documents have regular shape due to plain background and clear shape of the characters. Hence, the Method-1 classifies Indus document from other documents correctly.

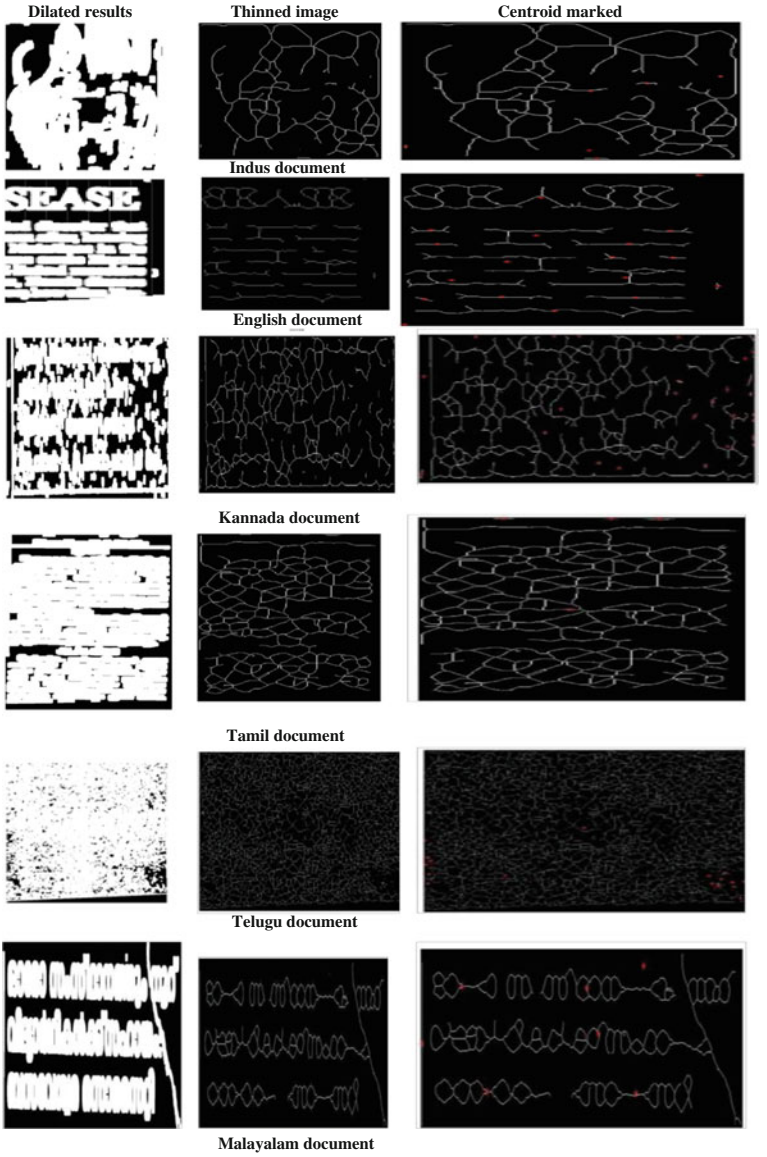


Fig. 3 Classification of Indus document by the Method-1

3.2 Method-2 for English Document Classification

We found that the Method-1 allows more tolerance to study the amount of cursiveness in the Indus documents because of dilation done by the morphological operations on edge components in the documents. This dilation operation

sometimes groups the components in the different lines when there is a less space between the text lines or if noise is present between the text lines. As a result, the Method-1 classifies English and South Indian documents into one group and Indus documents into another group. However, we notice from edge image of English and South Indian documents in Fig. 1 that we can see clear rectangular-shaped words or characters in English document as there are no modifiers unlike in South Indian documents while polygonal-shaped words in South Indian documents due to modifiers and sub-components. To extract this behavior, we propose region growing based on nearest neighbor concept to group the characters in the words in the same line without touching adjacent text lines though modifiers are present in the images. For each edge component in the document, the method fills the component and it allows the contour of the filled component to grow until it reaches nearest neighbor pixel of the component. Then, it merges the present filled component with the previous filled component to make one component. In this way, the method merges the characters in the words as one component and sometimes merges the words in the same text line as one component if there is a less space between the words. Since the space between the character and words is less than the space between the text lines, the region-growing method grows along text-line direction without touching adjacent text lines as shown in Fig. 4. Then, the method applies thinning algorithm to get single pixel width of the filed components as shown in Fig. 4 where we can notice clearly the straight-line components for English and cursive-line components for South Indian documents. We use the same steps mentioned in the Method-1 for computing the percentage of straight components and cursive components for each document. If the percentage of straight components is more than the percentage of cursive components, then the document is classified as English otherwise the document is classified as other documents.

In summary, the Method-1 classifies Indus document from the English and South Indian documents and the Method-2 classifies English document from the South Indian documents. The integration method classifies Indus and English from the South Indian documents.

4 Experimental Results and Comparative Study

We collect total 600 images (100 Indus, 100 English, 100 Kannada, 100 Tamil, 100 Telugu, and 100 Malayalam documents) for experimentation and for the purpose of evaluation. Since there is no benchmark/standard database is available, we create our own dataset by scanning the documents. The sample images from our database for different scripts are shown in Fig. 5 where one can see the complexity of the Indus document processing compared to other documents processing. To evaluate the performance of the method, we generate confusion matrix for classification of Indus and English with others documents and we consider classification rate as a performance measure. We compare classification

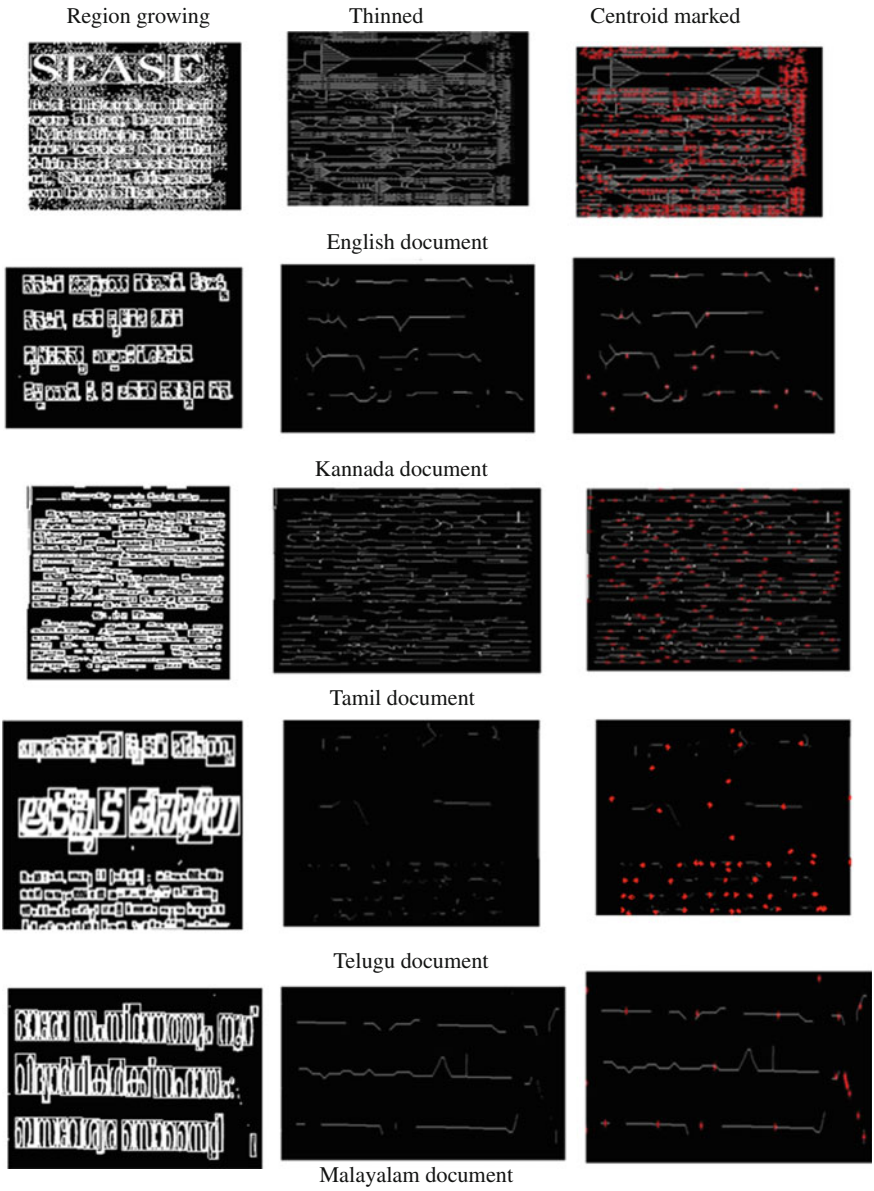


Fig. 4 Classification of English document from South Indian documents by Method-2

results with the classification results of the existing method [18] which considers English and South Indian scripts for identification to show that the proposed method is better than the existing method in terms of classification rate. The existing method extracts features based on connected component analysis and

Fig. 5 Sample documents chosen from our database

experimented on synthetic images but not on real images to identify the scripts. Since it works based on connected components analysis, it is good for the documents without having any distortions and plain background but not for the document like Indus where the complex background and irregular-shaped texts are present. Therefore, the existing method gives poor accuracy for classification of our dataset compared to the proposed method results. The reason for choosing this existing method is that this method focuses on English and South Indian scripts and it can be extended to other scripts as well. This will be discussed in the subsequent sections.

4.1 Experiment on Classification of Indus Documents

The experimental results of both the proposed integrated method and the existing methods for classification of Indus document from the other documents are reported in Table 1 where one can notice the classification rate of the proposed method is higher than the existing method for Indus to Indus and others to others. The main reason for getting poor classification rate for the existing method is that the existing method is good for synthetic image but not for real images (scanned images) where distortions and complex background exist. Note that since existing method classifies the scripts at text-line level and our work classifies at document level, we segment the text lines using projection profile analysis for the input and then we use voting criteria to classify the whole document. In addition, the existing method does not consider Indus document for experimentation; however, it is stated that the method can be used to classify the English from any other scripts. Therefore, we consider Indus document as others group and English as English group for classification as shown in Table 1.

Table 1 Confusion matrix for the proposed and existing method on Indus document classification

Proposed integrated method			Existing method		
Scripts	Indus (%)	Others (%)	Script	Indus (%)	English (%)
Indus	**67**	33	Indus	52	48
Others	35.8	**64.2**	English	66	34

Table 2 Confusion matrix for the proposed and existing method on English document classification

Proposed integrated method			Existing method		
Scripts	English (%)	Others (%)	Scripts	English (%)	Others (%)
English	**80**	20	English	34	66
Others	34.75	**65.25**	Others	73	27

4.2 Experiment on Classification of English Document

The experimental results of the proposed integrated method and the existing method are reported in Table 2 where the classification rate of the proposed method is higher than the existing method for English document classification from the South Indian documents. The reason for getting poor classification rate for the existing method is same as the above-mentioned in the previous section. The same criteria explained in previous section for conducting experiments on classification using existing method are used. It is observed from the Tables 1 and 2 that proposed method gives good results for English document classification from South Indian documents compared to Indus document from other document because of complexity involved in classifying Indus document and the proposed features are not good enough to achieve good classification rate for the Indus document classification.

5 Conclusion and Future Work

This paper presents new features based on visual clues over an edge map of Indus, English, Tamil, Telugu, Malayalam, and Kannada document images for classification of Indus and English scripts in this work. The extracted features reflect the straightness and cursiveness properties of edge components in the document images. We propose an integrated method consists of Method-1 for Indus document classification and Method-2 for English document classification from South Indian documents. Method-1 uses morphological operation for classification while the Method-2 uses region growing for classification. The experimental results and comparative study shows that the proposed integrated method outperforms the existing method in terms of classification rate. However, it is noticed from the

results that the classification rate of the proposed method for Indus document is low compared to the accuracy as in document analysis. We are planning to improve the method to achieve better accuracy and classification of South Indian scripts. The current work focuses on classification of Indus and English from South Indian documents but not classification of Kannada, Tamil, Telugu, and Malayalam.

References

1. Ghosh D, Dube T, Shivaprasad AP (2010) Script recognition-review. IEEE Trans PAMI, pp 2142–2161
2. Tan TN (1998) Rotation invariant texture features and their use in automatic script identification. IEEE Trans PAMI, pp 751–756
3. Busch A, Boles WW, Sridharan S (2005) Texture for script identification. IEEE Trans PAMI, pp 1720–1732
4. Shijian L, Tan CL (2008) Script and language identification in noisy and degraded document images. IEEE Trans PAMI, pp 14–24
5. Jaeger S, Ma H, Doermann D (2005) Identifying script on word-level with informational confidence. In: Proceedings of ICDAR, pp 416–420
6. Pati PB, Ramakrishnan AG (2008) Word level multi-script identification. Pattern Recogn Lett, pp 1218–1229
7. Chanda S, Pal S, Franke K, Pal U (2009) Two-stage approach for word-wise script identification. In: Proceedings of ICDAR, pp 926–930
8. Chanda S, Terrades OR, Pal U (2007) SVM based scheme for Thai and English script identification. In: Proceedings of ICDAR, pp 551–555
9. Li L, Tan CL (2008) Script identification of camera-based images. In: Proceedings of ICPR
10. Namboodiri AM, Jain AK (2002) On-line script recognition. In: Proceedings of ICPR, pp 736–739
11. Pal U, Sinha S, Chaudhuri BB (2003) Multi-script line identification from indian documents. In: Proceedings of ICDAR, pp 880–884
12. Choudhury S, Harit G, Madnani S, Shet RB (2000) Identification of scripts of indian languages by combining trainable classifiers. In: Proceedings of ICVGIP
13. Chanda S, Pal U (2005) English Devanagari and Urdu text identification. In: Proceedings of ICDAR, pp 538–545
14. Zhou L, Lu Y, Tan CL (2006) Bangla/English script identification based on analysis of connected component profiles. In: Proceedings of DAS, pp 243–254
15. Gopakumar R, Subbareaddy NV, Makkithaya K, Acharya UD (2010) Zone-based structural features extraction for script identification from Indian documents. In: Proceedings of ICIIS, pp 420–425
16. Aithal PK, Rajesh G, Acharya DU, Krishanamoorthi M, Subbareddy NV (2011) Script identification for a Tri-lingual document. In: Proceedings of CNC, pp 434–439
17. Rani R, Dhir R, Lehal GS (2011) Comparative analysis of Gabor and discriminating features extraction techniques for script identification. In: Proceedings of ICISILL, pp 174–179
18. Padma MC, Vijaya PA (2009) Monothetic separation of Telugu, Hindi and English text lines from multi script document. In: Proceedings of ICSMC, pp 4870–4875
19. Ghosh S, Chaudhuri BB (2011) Composite script identification and orientation detection for indian text images. In: Proceedings of ICDAR, pp 294–298
20. Phan TQ, Shivakumara P, Ding Z, Lu S, Tan CL (2011) Video script identification based on text lines. In: Proceedings of ICDAR, pp 1240–1244

Uterine Fibroid Segmentation and Measurement Based on Morphological Functions in Graphical Vision Assistant Tool

S. Prabakar, K. Porkumaran and J. Guna Sundari

Abstract Uterine fibroid is the most predominant problem among women of child-bearing age where the secretion of estrogen hormone plays significant role. The presence of fibroid can cause severe pain, infertility, and repeated miscarriages. Since the detection of fibroid and treatment is the crucial factor on women health especially in pregnancy, ultrasound (US) imaging is the most common modality for detecting fibroids. Because of the presence of speckle noise, the segmentation of fibroid from an US image is the tedious process. The proposed methodology has been used for automating this task by morphological functions available in graphical vision assistant tool. The modified morphological image cleaning (MMIC) algorithm for filtering and Canny edge detector have been utilized for fibroid segmentation and binary image morphological approaches adopted for analyzing the fibroid. The proposed algorithm has been developed, implemented, and validated in LabVIEW vision assistant toolbox. The outcomes of the proposed method have been evaluated and appreciated by experienced gynecologists and found that the manual intervention is eliminated on the investigation of diseased.

Keywords Morphology · MMIC · Segmentation · Ultrasound scanning · Uterine fibroid · Vision assistant tool

S. Prabakar (✉)
Biomedical Engineering Department, Karunya University, Coimbatore, Tamil Nadu 641114, India
e-mail: srisornaprabu@gmail.com

K. Porkumaran
Dr. NGP Institute of Technology, Coimbatore 641048, India
e-mail: porkumaran@ieee.org

J. Guna Sundari
Computer Applications Department, SREC, Coimbatore, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_36,

1 Introduction

Uterine fibroid is the most common benign (not cancerous) tumor of a woman's uterus (womb). Fibroids are common pelvic tumors of the smooth muscle that are normally found in the wall of the uterus. It is also known as uterine leiomyoma. They can develop within the uterine wall itself or attach to it. This originates from the smooth muscle layer (myometrium) and the accompanying connective tissue of the uterus. They may grow as a single tumor or in clusters. Most fibroids are asymptomatic; they can grow and cause heavy and painful menstruation, pelvic pain, urinary frequency and urgency, and, in some cases, infertility [1], so even though they are termed "benign (not cancerous) tumors," fibroids potentially can cause many health problems. Ultrasound (US) scanning is a simple, a traumatic, and safe method of examination of this disease. Generally, US is used for imaging soft tissues in organs so that the fibroids in uterus can be identified quickly, easily, and safely with stumpy cost. This paper proposed graphical-system-design-based vision techniques to segment and analyze the fibroids in US uterus images.

A computerized screening system can be used for fully automated mass screening. Such systems screen a large number of US images and segment the abnormal images, which are then further examined by a radiologist. This would save a significant amount of workload and time for radiologists, allowing them to concentrate their resources on surgery and treatment. The speckle noise present in the US images reduces the quality of the image, and the detection and analysis of the disease are the highly challenging task. Various techniques for speckle noise removal are available in the literatures [2–6]. There are many segmentation algorithms such as seed point, region-growing, active contours, and morphological operations for segmenting medical images [7]. The windows adaptive threshold and Otsu's algorithm are used for US liver image segmentation. A region-based segmentation and seed point selection methods for US images using local statistics have produced results that are less sensitive to the pixel location, and it also allowed a segmentation of the accurate homogeneous regions [8, 9]. Morphological-based segmentation has been proposed in different research works for various applications [10, 11]. The watershed-based segmentation and region merging were presented to segment tumors from US images [12]. The over-segmentation associated with watershed algorithm has been resolved by hybrid method combined with morphological functions, Gaussian, and region-growing techniques [13, 14]. Some of the automatic approaches for US image segmentation have been used Canny operators [15, 16]. Based on the various studies on US image segmentation, the mathematical morphological methodology is the most suitable approach for uterine fibroid segmentation [17, 18]. Certain investigations using morphological image cleaning algorithm (MIC) have been utilized for fluid-filled uterine fibroid feature extraction. The present work proposed modified morphological image cleaning (MMIC) algorithm, and Canny edge detector has been adopted for uterine fibroid segmentation using graphical vision assistant tool.

This paper has been organized as follows. In Sect. 2, the US image acquisition and various proposed scheme modules have been explained in detail. The results of the proposed scheme, graphical representations, and validation of the outcomes have been discussed in Sect. 3. The conclusion and future work for uterine fibroid segmentation and analysis have been described in Sect. 4.

2 Proposed Method

The proposed methods for the diseased US images have the following series of steps which have been implemented for segmentation, detection, and analysis of uterine fibroid.

Step 1: Image Acquisition through ultrasound machine and US image database creation.
Step 2: MMIC filtering algorithm used to reduce speckle noise.
Step 3: Histogram equalization method for image enhancement.
Step 4: Created closed region of fibroid using mathematical morphological operation.
Step 5: Extracting the edge data using canny edge detector.
Step 6: Creation of concrete contour of the fibroid.
Step 7: Overlapping fibroid contour on original US image.
Step 8: Measurement of fibroid area, diameter and perimeter for the diagnosis of the disease.

In the following subsections, each step will be discussed in details and all the steps are combined as a pipeline for a whole scheme.

2.1 Image Acquisition and Database

The symptoms of the uterine fibroid might cause certain inconvenience in women which needs to be diagnosed and treated. The usage of US scanners in the diagnosis has been well established because of its noninvasive nature, low cost, capability of forming real-time imaging, and continuing improvement in image quality. The B-mode imaging is the most commonly suitable modality for this problem.

The US uterine fibroid image has been acquired from different patients in different hospitals. The database contains different cases that cover different types of abnormalities like submucosal, subserosal. Every different abnormality indicator has different features in shape, brightness, size, and distribution. Each image has been accompanied with an expert radiologist and the gynecologists' diagnosis proved by a biopsy. The size of the image is 640 $\times$ 480, and it has been reduced to 350 $\times$ 350 for further processing.

2.2 Overview of the Uterine Fibroid Investigation System

The uterine fibroid investigation system has been developed by the LabVIEW vision assistant tool based on the mathematical morphological functions. The overview of the system is shown in Fig. 1. The US images from the database have been preprocessed using morphological techniques such as dilation and erosion to remove speckle noise, and the reconstructed image has been used to extract the fibroid part. The boundary descriptions are extracted by segmentation algorithm such as the mathematical analysis of the fibroid image. Then, the contour is extracted in which a boundary has been placed automatically using morphological algorithm. Then, the features of the uterine fibroid have been extracted through which the fibroid's area, diameter, compactness, shape, etc., have been measured, and the information delivered by this approach has helped the gynecologist for better diagnosis and treatment with short span of time.

2.3 Speckle Reduction

US images have some inherent problems, including speckle noise and interference. Since the cleaning process takes much time, the algorithm has been modified and applied on US images such that it has less processing to get the result. The true-color uterine fibroid image has been converted to grayscale image. Then, the structuring element has been initialized, and a series of operation such as opening–closing followed by closing–opening (OCCO) has been approached to remove the noise.

A speckle does not have regular shape. It is not appropriate to use predefined structuring elements, like disk, rectangle, hexagon, for morphological processing. Therefore, an arbitrary structuring element which resembles the speckle shape is designed. For this random speckle, samples are taken from different US images and the structuring elements were designed as shown in Fig. 2.

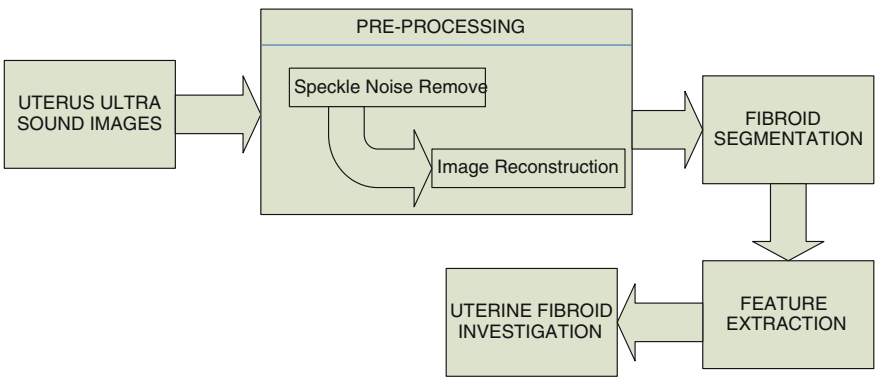


Fig. 1 Overview of the uterine fibroid investigation system

```
00000000000000
00001111000000
00111111110000
00111111111110
01111111111110
01111111111110
```

Fig. 2 Structuring element

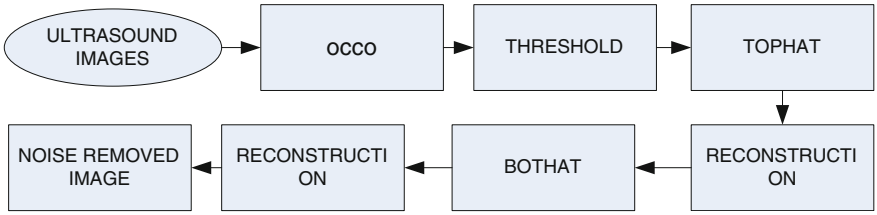


Fig. 3 Preprocessing steps

The speckle noise removal process of the fibroid image is shown in Fig. 3. First, for filtering using open–close and close–open (OCCO), the arbitrary structuring elements that resemble the shape of the speckle have been used. The Tophat of the image has been obtained by subtracting the original image from the opening of the OCCO-filtered image. The standard deviation has been utilized to acquire the binarized image using the threshold value of the Tophat fibroid image. Then, the image is reconstructed by closing the binary image. The bottom hat (Bothat) of the reconstructed image has been obtained by subtracting the original image from the closing of the OCCO-filtered image. The standard deviation has been utilized to acquire the binarized image using the threshold value of the Bothat fibroid image. Then, the image is reconstructed by opening the binary image. The input fibroid image has been added to the reconstructed Tophat image and sequentially subtracted from the reconstructed Bothat image. Finally, the speckle-noise-removed image has been obtained for further processing.

2.4 MMIC Algorithm

The following series of steps has been involved in MMIC algorithm implementation.

Step 1: Read the input image using the IMAQ Read file tool in the LabVIEW NI Vision Assistant.

Step 2: Convert the true color image into grayscale image.

Step 3: Reconstruct the image with the structuring element in 2D multi resolution tool.
Step 4: By using this structuring element a series of operation such as opening-closing followed by closing-opening is done to remove noise.
Step 5: Top hat of the image is found by subtracting the original image with opening of the OCCO filtered image.
Step 6: Find the threshold value of the top hat image and binarised it using the standard deviation.
Step 7: Then reconstruct the image by closing of the binary image.
Step 8: Bottom hat of the image is found by subtracting the original image with closing of the OCCO filtered image.
Step 9: Find the threshold value of the bottom hat image and binarised it using the standard deviation.
Step 10: Then reconstruct the image by opening of the binary image.
Step 11: Use erosion and dilation to expose the required fibroid features.
Step 12: Use the Connected Component and Canny edge detection tool to extract the fibroid from the binary image.

2.5 Morphological Operations

Our aim is to find the contour of fibroid and segment the region enclosed by the boundary. In general, mathematical morphology (MM) is the foundation of morphological image processing, which consists of a set of operators that transform images according to the above characterizations. Morphological operators often take a binary image and a structuring element as input and combine them using a set operator (intersection, union, inclusion, complement) [19]. They process objects in the input image based on characteristics of its shape, which are encoded in the structuring element. After the implementation of OCCO, the connected components have been used to extract the area of the fibroid.

3 Result and Discussion

The MM operators have been implemented using graphical vision assistant tool, and the results have been obtained as shown in Fig. 4. This proposed algorithm has been performed for 10 patient's uterine fibroid images. For every patient, minimum of four US images have been taken with various positions to obtain the consistency on the algorithm result. In all the cases, the proposed method has been performed well and delivers the 90 % accuracy on fibroid segmentation.

Fig. 4 Original uterine fibroid image, MM-based speckle-noise-removed image, binarized uterine fibroid image, connected component–based image, and fibroid image showing contour extraction using canny edge detector

The segmented fibroid portion has been used to analyze the fibroid. The representative features of the segmented image have been used to determine the mode of treatment. For instance, if the size of the fibroid is very small, it can be treated by giving medicine, whereas if the size is big, it should be removed by surgery. The shape-based features, which have been considered, are total area, perimeter, diameter, eccentricity, major axis, and minor axis. In the present work, the total area, perimeter, and diameter have been estimated. The area of the fibroid has been identified by the number of pixels inside the segmented fibroid image. This is the vital parameter used to specify the size of the fibroid. It can also be found by the formula $A = (\pi * a * b)/2$, where a and b are the horizontal and vertical diameters. Perimeter is the distance around the boundary of the region of interest. Regionprops in the mathscript of the graphical system design has been computed the perimeter by calculating the distance between each adjoining pair of pixels around the border of the segmented fibroid region. If the image contains discontiguous regions, regionprops returns unexpected results. The diameter has been measured and specified by the radiologist as horizontal and vertical diameters in centimeters. After segmenting the fibroid, the horizontal and vertical diameters have been calculated in pixels and converted to centimeters. The most significant features such as area, diameter, and perimeter have been estimated from the segmented fibroid images as shown in Table 1, and the graphical representation is shown in Figs. 5, 6. Based on the assessment of these parameters, the physicians will decide the nature of the treatment.

Table 1 Various parameters (area, diameter, and perimeter) measurement of fibroid for various patients fibroid images

Images	Area (no of on pixels)	Diameter (mm)	Perimeter (mm)
Images 1	1.7456e + 004	33.87	123.05
Images 2	2.6142e + 002	59.32	247.96
Images 3	1.9483e + 001	53.46	273.42
Images 4	1.7664e + 003	48.11	186.95
Images 5	2.4640e + 001	102.17	550.25
Images 6	2.0722e + 003	100.79	581.87
Images 7	2.0156e + 004	59.62	310.39
Images 8	1.3856e + 002	41.65	167.88
Images 9	1.2692e + 002	47.29	214.69
Images 10	2.7409e + 001	99.96	592.21

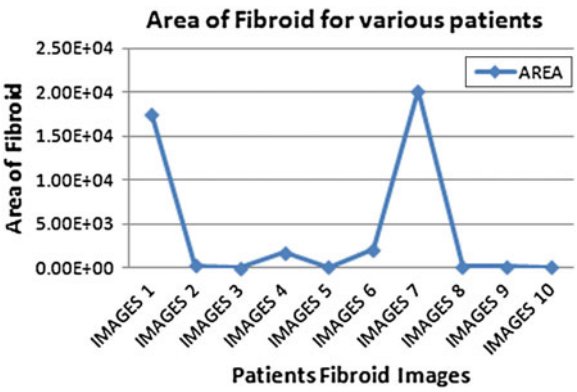


Fig. 5 Analysis of fibroid area

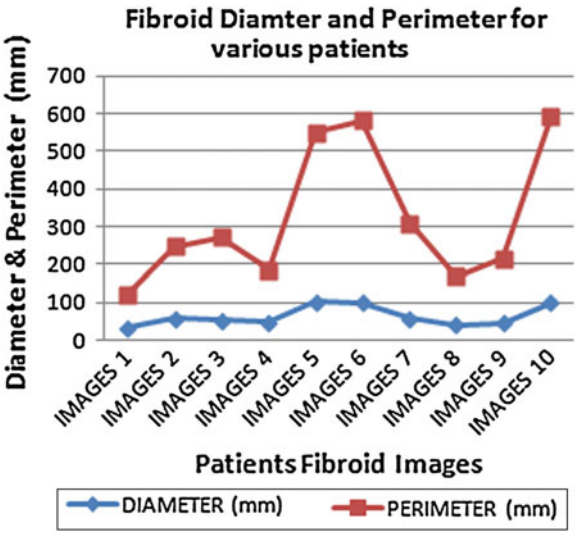


Fig. 6 Analysis of fibroid diameter and perimeter

The other parameters such as eccentricity, major axis, and minor axis are used to help the radiological planning. The eccentricity specifies the unconventional behavior of the fibroid region that has the same second central moments as the region. The eccentricity is the ratio of the distance between the region of the fibroid and its major axis length. The value is between 0 and 1. The major axis is the scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region of fibroid. The minor axis is the length (in pixels) of the minor axis of the fibroid region, which has the same, normalized second central moments as the region of fibroid. These additional features could be estimated from the segmented uterine fibroid image.

4 Conclusion and Future Work

In this paper, an algorithm based on morphological operation in graphical vision assistant tool has been presented for the extraction of uterine fibroid area. The method could extract the area of the fibroid based on Canny edge detection and morphology methods. The subjective appearance of the output image was good. The performance of this algorithm has also been found to be good since it overcomes the over-segmentation. The seed point selection to start segmentation has been eliminated in this method. This method is very much useful for radiological planning for removal of fibroid. This automatic segmentation method will reduce the time of technicians to locate the fibroid on the screen. The algorithm has been performed for 10 patient's uterine fibroid images to extract the features, and the outcomes have found to be more precise. The other features such as eccentricity, major axis, and minor axis will also be estimated with graphical vision assistant toolto provide more information about the fibroid treatment. The LabVIEW implementation of this work will lead to embedded development work, that is, by which a handheld device will be proposed to analyze the fibroid diagnosis and will provide more ergonomic feasibility on fibroid treatment.

Acknowledgments We would like to sincerely thank, Dr. N. Kalpana, Hari Scan Center, Erode, for providing the US images and validation of proposed work. Karunya University, Dr. N.G.P Institute of Technology, Coimbatore, and Kovai Medical Center Hospitals are also acknowledged for their facilitation.

References

1. Hassani N (1975) Ultrasonic evaluation of uterine fibroids -the sonic shadow sign. J Natl Med Assoc 67(4):307–311
2. Karamam M, Kutay M, Bozdagi G (1995) An adaptive speckle suppression filter for medical ultrasonic images. IEEE Trans Med, Imaging
3. Peters RA II (1995) A new algorithm for image noise reduction using mathematical morphology. IEEE Trans Image Process 4:554–568
4. Thangavel K, Manavalan R, Aroquiaraj IL (2009) Removal of speckle noise from ultrasound medical image based on special filters: comparative study. ICGST-GVIP J, 9(3):25–32, ISSN 1687-398X
5. Vanithamani R, Umamaheswari G (2010) Performance analysis of filters for speckle reduction in medical ultrasound images. Int J Comput Appl 12(6):0975–8887
6. Jeyalakshmi TR, Ramar K (2010) A modified method for speckle noise removal in ultrasound medical images. Int J Comput Electr Eng 2(1):1793–8163
7. Noble JA, Boukerroui D (2006) Ultrasound image segmentation: a survey. IEEE Trans Med Imaging 25:987–1010
8. Thakur A, Anand RS (2005) A local statistics based region growing segmentation method for ultrasound medical images. Int J Signal Process 1:141–146
9. Shan J, Cheng HD, Wang Y (2008) A novel automatic seed point selection algorithm for breast ultrasound images. In: 19th international conference on pattern recognition
10. Infantosi AFC, Luz LMS, Pereira WCA, Alvarenga AV (2008) Breast ultrasound segmentation using morphologic operators and a gaussian function constraint. In: Proceedings 20, www.springerlink.com © Springer, Berlin Heidelberg, pp 520–523

11. Nallaperumal K, Krishnaveni K, Varghese J, Saudia S, Annam S, Kumar P (2007) A novel multi-scale morphological watershed segmentation algorithm. JISE 1(2), GA, USA, ISSN:1934-5
12. Bleau A, Leon LJ (2000) Watershed-based segmentation and region merging. Comput Vision Image Underst 77:317–370
13. Levner I, Zhang H (2007) Classification-driven watershed segmentation. IEEE Trans Image Process 16:1437–1445
14. Vincent L (1992) Morphological area openings and closings for grayscale image. In: Proceedings of NATO shape in picture workshop, Driebergen, The Netherlands, pp 197–208
15. Yang X, Ding M, Lou L, Yuchi M, Qiu W, Sun Y (2011) Common carotid artery lumen segmentation in B-mode ultrasound transverse view images. I J Image Graphics Signal Process 5:15–21
16. Hiremath PS, Tegnoor JR (2010) Automatic detection of follicles in ultrasound images of ovaries using edge based method. IJCA special issue on "recent trends in image processing and pattern recognition" RTIPPR
17. Jeyalakshmi TR, Ramar K (2009) Segmentation of uterine fibroid using morphology: an automatic approach. 978-1-4244-4711-4/09/$25.00 ©2009 IEEE
18. Jeyalakshmi R, Kadarkarai R (2010) Segmentation and feature extraction of fluid- filled uterine fibroid—a knowledge-based approach. Maejo Int J Sci Technol 4(03):405–416
19. Gonzalez RC, Woods RE (2008) Digital image processing. 3rd edn Pearson Education Inc., Publishing as Prentice Hall

Efficient Algorithm for Image Compression Using DWT Techniques

C. R. Geetha, H. Basavaraju, R. C. Manjunatha, C. P. Latha and H. D. Giriprakash

Abstract This paper proposes a method for image compression using discrete wavelet transformation (DWT) techniques. One use of wavelet approximation is in data compression. Like some other transforms, wavelet transforms can be used to transform data, and then encode the transformed data, resulting in effective compression. A related use is that of smoothing/de-noising data based on wavelet coefficient thresholding, also called wavelet shrinkage. By adaptively thresholding the wavelet coefficients that correspond to undesired frequency components smoothing and/or de-noising operations can be performed. Wavelets are functions that are concentrated in time as well as in frequency around a certain point. For practical applications, we choose wavelets which correspond to a so called "multi-resolution analysis" (MRA) due to the reversibility and the efficient computation of the appropriate transform. Wavelets fulfill certain self-similarity conditions. Images are obviously two-dimensional data. To transform images we can use two-dimensional wavelets or apply the one-dimensional transform to the rows and columns of the image successively as separable two-dimensional transform.

C. R. Geetha (✉)
Jain University, Bangalore, Karnataka, India
e-mail: geethacr.gowda@gmail.com

H. Basavaraju · H. D. Giriprakash
ECE Department, Sapthagiri College of Engineering, Bangalore, Karnataka, India
e-mail: rajhunsur@yahoo.co.in

H. D. Giriprakash
e-mail: giriprk@gmail.com

R. C. Manjunatha
ECE Department, Acharya Institute of Technology, Bangalore, Karnataka, India
e-mail: manjunathrc@acharya.ac.in

C. P. Latha
ECE Department, HMSIT, Tumkur, Karnataka, India
e-mail: cplathahms.12@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_37,

In most of the applications, where wavelets are used for image processing and compression, the latter choice is taken, because of the low computational complexity of separable transforms.

Keywords Discrete wavelet transformation · Lifting scheme

1 Introduction

The discrete wavelet transform (DWT) represents the signal in dynamic sub-band decomposition. Generation of the DWT in a wavelet packet allows sub-band analysis without the constraint of dynamic decomposition. The discrete wavelet packet transform (DWPT) performs an adaptive decomposition of frequency axis. The specific decomposition will be selected according to an optimization criterion

The DWT, based on time-scale representation, provides efficient multi-resolution sub-band decomposition of signals. It has become a powerful tool for signal processing and finds numerous applications in various fields such as audio compression, pattern recognition, texture discrimination, computer graphics. Specifically, the two-dimensional DWT and its counterpart two-dimensional inverse DWT (IDWT) play a significant role in many image/video coding applications. The fundamental idea behind wavelets is to analyze according to scale. Indeed, some researchers in the wavelet field feel that, by using wavelets, one is adopting a perspective in processing data. Wavelets are functions that satisfy certain mathematical requirements and are used in representing data or other functions. When Joseph Fourier discovered that he could superpose sinus and cosines to represent other functions. However, in wavelet analysis, the scale that one use to look at data plays a special role. Wavelet algorithms process data at different scales or resolutions [1]. Fourier Transform (FT) with its fast algorithms (FFT) is an important tool for analysis and processing of many natural signals. FT has certain limitations to characterize many natural signals, which are non-stationary (e.g. speech). Though a time varying, overlapping window based FT namely short time FT (STFT) is well known for speech processing applications, a time-scale based Wavelet Transform is a powerful mathematical tool for non-stationary signals.

Wavelet Transform uses a set of damped oscillating functions known as wavelet basis. WT in its continuous (analog) form is represented as continuous wavelet transform (CWT). CWT with various deterministic or non-deterministic basis is a more effective representation of signals for analysis as well as characterization. CWT is powerful in singularity detection. A discrete and fast implementation of CWT (generally with real-valued basis) is known as the standard DWT. With standard DWT, signal has a same data size in transform domain, and therefore, it is a non-redundant transform. A very important property was multi-resolution analysis (MRA) that allows DWT to view and process [2].

The DWT is being increasingly used for image coding. This is due to the fact that DWT supports features like progressive image transmission (by quality, by resolution), ease of compressed image manipulation, region of interest coding. DWT has traditionally been implemented by convolution. Such an implementation demands both a large number of computations and a large storage features that are not desirable for either high-speed or low-power applications. Recently, a lifting-based scheme that often requires far fewer computations has been proposed for the DWT.

The main feature of the lifting-based DWT scheme is to break up the high-pass and low-pass filters into a sequence of upper and lower triangular matrices and convert the filter implementation into banded matrix multiplications. Such a scheme has several advantages, including "in-place" computation of the DWT, integer-to-integer wavelet transform (IWT), symmetric forward and inverse transform. Therefore, it comes as no surprise that lifting has been chosen in the upcoming.

2 Methods Used in this Paper

2.1 Lifting Scheme of DFT

The basic principle of the lifting scheme is to factorize the poly-phase matrix of a wavelet filter into a sequence of alternating upper and lower triangular matrices and a diagonal matrix. This leads to the wavelet implementation by means of banded matrix multiplications.

Let $\bar{h}(z)$ and $\bar{g}(z)$ be the low-pass and high-pass analysis filters, and let $h(z)$ and $g(z)$ be the low-pass and high-pass synthesis filters. The corresponding poly-phase matrices are defined as [3]

$$\bar{P}(z)\begin{bmatrix}\bar{h}_e(z) & \bar{h}_o(z)\\ \bar{g}_e(z) & \bar{g}_o(z)\end{bmatrix}P(z)\begin{bmatrix}h_e(z) & h_o(z)\\ g_e(z) & g_o(z)\end{bmatrix} \tag{2.1}$$

If $(\bar{h}, \bar{g})$ is a complementary filters pair, then it can always be factored into lifting steps as

$$\bar{P}_1(z)\begin{bmatrix}K & 0\\ 0 & 1/K\end{bmatrix}\prod_{i=1}^{m}\begin{bmatrix}1 & S_i(z)\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ \bar{t}_i(z) & 1\end{bmatrix} \tag{2.2}$$

$$\bar{P}_1(z)\begin{bmatrix}K & 0\\ 0 & 1/K\end{bmatrix}\prod_{i=1}^{m}\begin{bmatrix}1 & 0\\ \bar{t}_i(z) & 1\end{bmatrix}\begin{bmatrix}1 & \bar{S}_i(z)\\ 0 & 1\end{bmatrix} \tag{2.3}$$

where K is a constant. The two types of lifting schemes are shown in Fig. 1 Scheme 1 $P_1(z)$ which corresponds to the factorization consisting of three steps:

1. Predict step, where the even samples are multiplied by the time domain equivalent of $t(z)$ and are added to the odd samples.

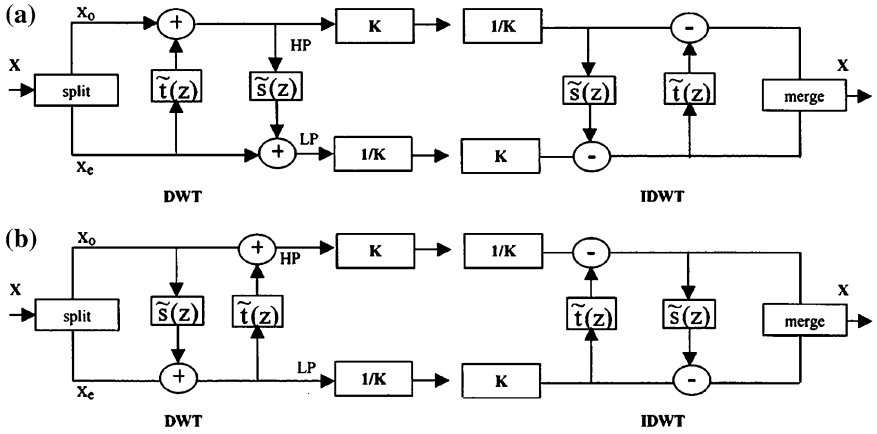


Fig. 1 Lifting schemes. **a** Scheme1. **b** Scheme2

2. Update step, where updated odd samples are multiplied by the time domain equivalent of $s(z)$ and are added to the even samples.
3. Scaling step, where the even samples are multiplied by $1/k$ and odd samples by k.

The IDWT is obtained by traversing in the reverse direction, changing the factor K to $1/K$, K factor to i/K, and reversing the signs of coefficients in $t(z)$ and $s(z)$.

In Scheme 2 which corresponds to the $\bar{P}_2(z)$ factorization, the odd samples are calculated in the first step, and the even samples are calculated in the second step. The inverse is obtained by traversing in the reverse direction [4]. The discrete wavelet transform applies several filters separately to the same signal. In contrast to that, for the lifting scheme, the signal is divided like a zipper. Then, a series of convolution accumulate operations across the divided signals are applied.

The basic idea of lifting is the following: If a pair of filters (h, g) is complementary, that is, it allows for perfect reconstruction, then for every filters, the pair (h', g) with $h'(z) = h(z) + s(z^2).g(z)$ allows for perfect reconstruction too. Of course, this is also true for every pair (h, g') of the form $g'(z) = g(z) + t(z^2).h(z)$. The converse is also true: If the filter banks (h, g) and (h', g) allow for perfect reconstruction, then there is a unique filters with $h'(z) = h(z) + s(z^2).g(z)$. Each such transform of the filter bank (or the respective operation in a wavelet transform) is called a lifting step. A sequence of lifting steps consists of alternating lifts, that is, once the low pass is fixed and the high pass is changed in the next step, the high pass is fixed and the low pass is changed. Successive steps of the same direction can be merged [3].

2.2 Discrete Wavelet Transform of One-Level DWT

The DWT [6–8] is usually computed through convolution and sub-sampling with a couple of filters to produce an approximation signal L (low pass filter result) and a detail signal H (high-pass filter result). For two-dimensional signals, there exist separable wavelets for which the computation can be decomposed into horizontal processing (on the rows) followed by vertical processing (on the columns), using the same One-dimensional filters. Figure 2 shows a one-level decomposition. Subsequent levels are obtained by iterating on the low-pass signal LL [14].

2.2.1 One-Dimensional DWT and IDWT Architecture

The multi-resolution DWT and IDWT can be viewed as cascades of several two-channel analysis and synthesis filter banks, respectively. The analysis and synthesis filter banks are shown in Fig. 3, where $H(z)$ and $G(z)$ are the analysis low-pass and high-pass filters. For perfect reconstruction, the synthesis low-pass and high-pass filters can be defined $\tilde{H}(z) = G(-z)\,\tilde{G}(z) = -H(-z)$ [5]. Many very large scale integration (VLSI) architectures have been proposed to implement DWT and IDWT, and they can be categorized into three categories, including the convolution-, lifting-, and B-spline-based architectures.

2.3 Two-Dimensional DWT and IDWT

A three-level two-dimensional IDWT computation-schedule scheme for the proposed architecture shown in Fig. 5

Fig. 2 One-level separable wavelet decomposition

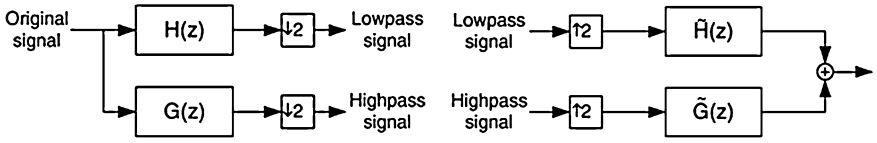


Fig. 3 Two-channel filter bank for DWT and IDWT. **a** Analysis filter bank for forward DWT. **b** Synthesis filter ban k for inverse DWT

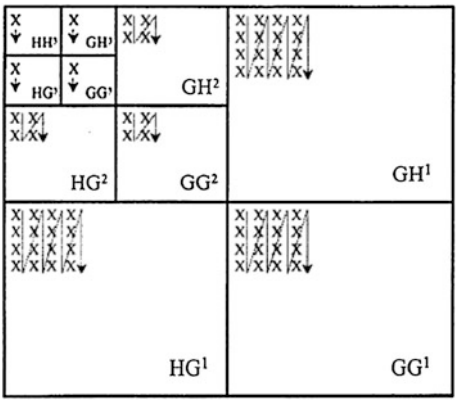


Fig. 4 Three-level computation-schedule scheme

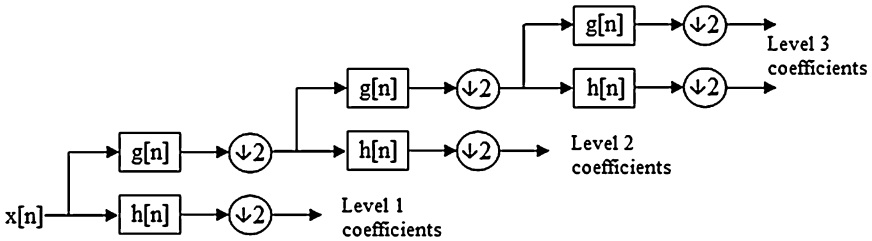


Fig. 5 Two channel filter bank at level 3

3 Experimental Analysis

Discrete wavelet transform is [9–11] used for the implementation of the image so that the image is compressed. And the inverse wavelet transform is used to get the original signal. The compression level is 3. Here, we used orthogonal filter. The input image is compressed by level 3. The below figure shows the compressed level (Fig. 6).

The Fig. 7 shows the initial tree of the image. The initial tree divides the input signal into the high-pass and low-pass signal. The branch of the high pass and low pass again divides high pass and low pass. The sub-branch also continues up to the ending of the coefficients.

The Fig. 8 shows the compression levels of the image. The compression level 3 is used.

The Fig. 9 shows the best level tree of the image. By using best level, we can get at maximum of original image. Here, the compression level 1 is used.

The Fig. 10 shows the original image and the compressed image. The number coefficients recovered is 99.74.

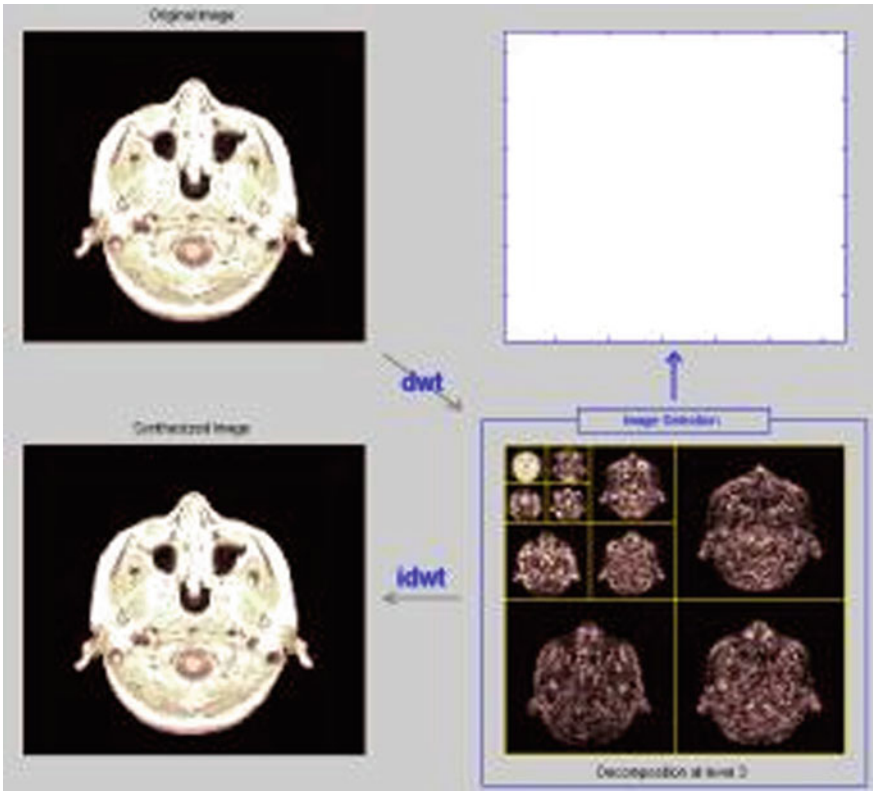


Fig. 6 The conversion of the image by using DWT and IDWT

The Fig. 11 shows the original image and the de-noised image. The stored absolute coefficients and histogram of the de-noised image.

The Fig. 12 shows the histogram of the image. The standard deviation is 2.537, median absolute deviation is 0.943, and the mean absolute deviation is 1.642.

The Fig. 13 shows the conversion of color image into gray scale image

Matlab is used to obtain the pixel values for the original image shown in Fig. 15. The size of the image is found to be of 128 × 128 pixels. A set of 4 × 4 pixels of the original image shown in Fig. 14 were given as test inputs. The original image input pixels, lifting-based DWT outputs/IDWT inputs and IDWT outputs are as shown in Table 1.

Referring to Table 1, it can be seen that the original pixels values at the input of DWT and the pixel values at the output of IDWT are same. Using IDWT output values, the original image can be reconstructed.

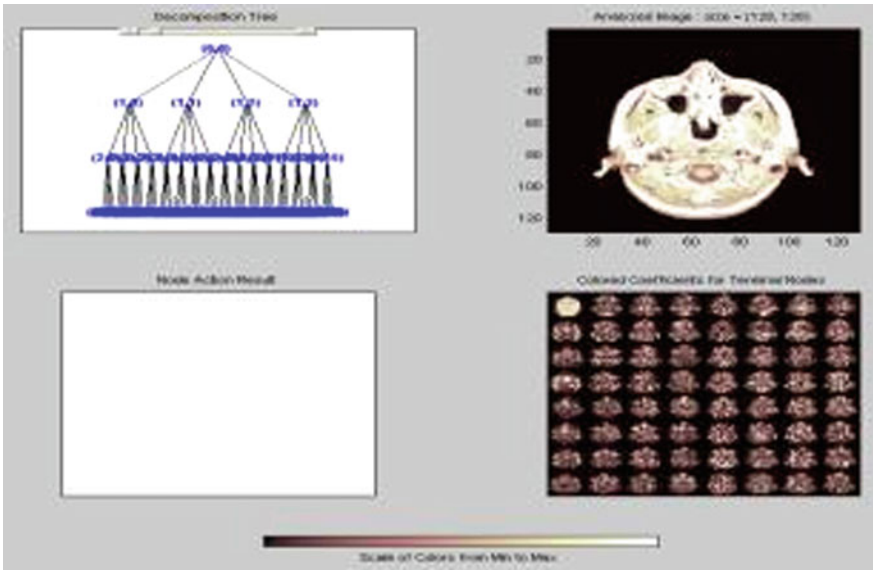


Fig. 7 Initial tree of image

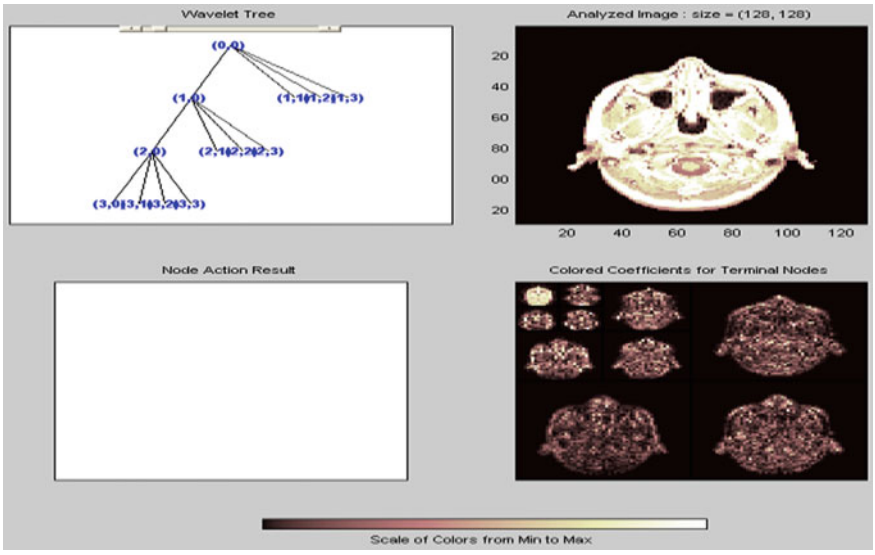


Fig. 8 Wavelet tree of image

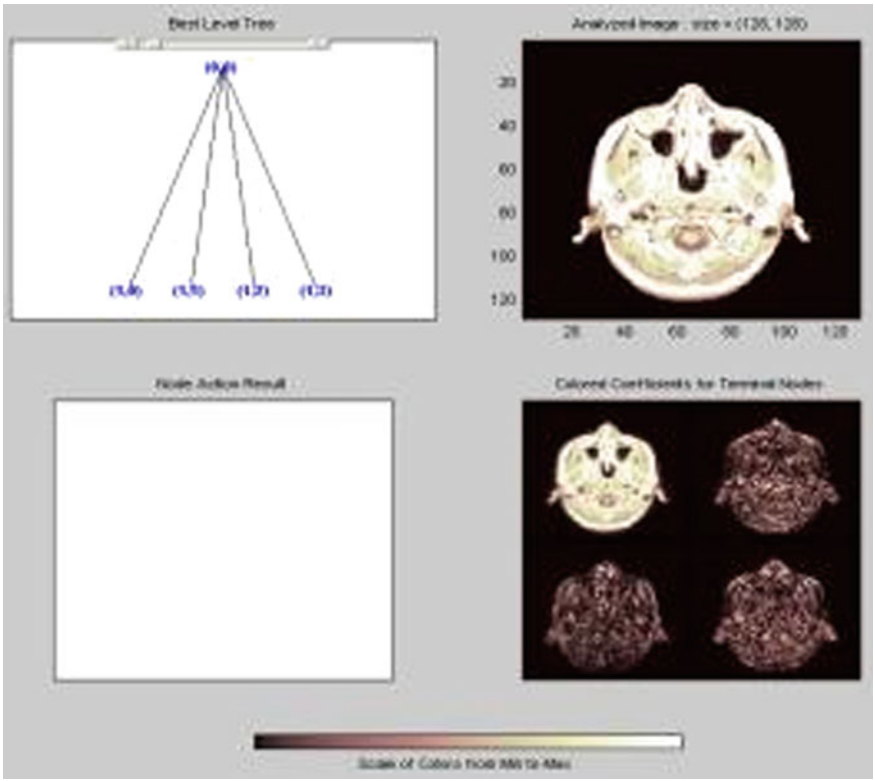


Fig. 9 Best level tree of image

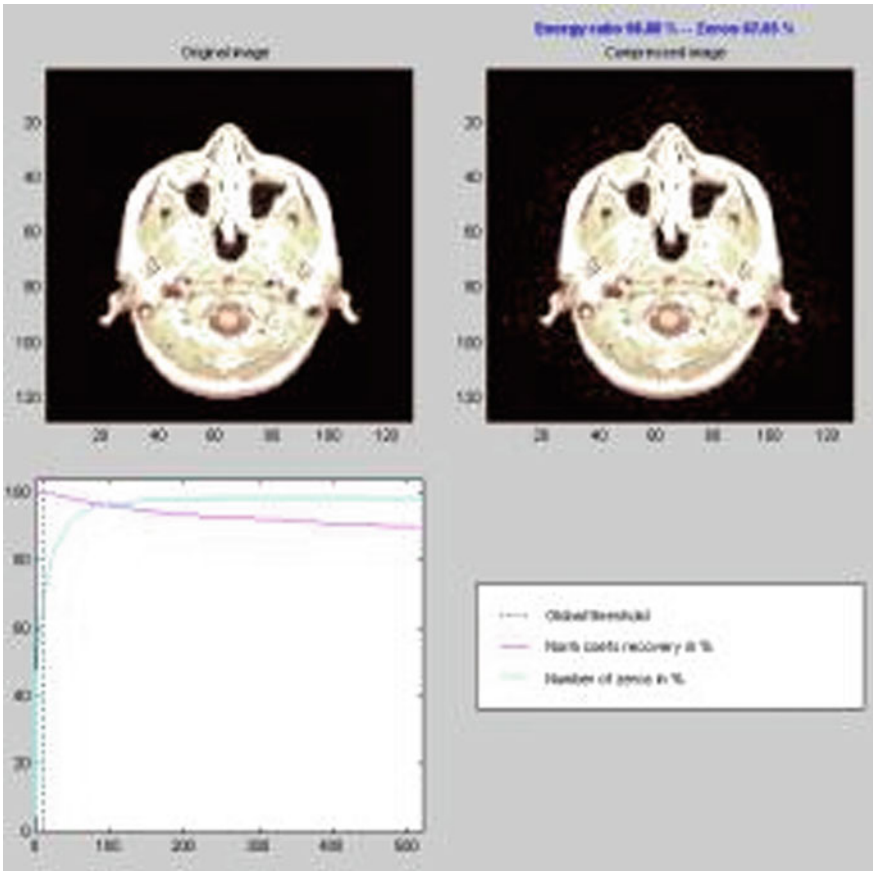


Fig. 10 Original and compressed image

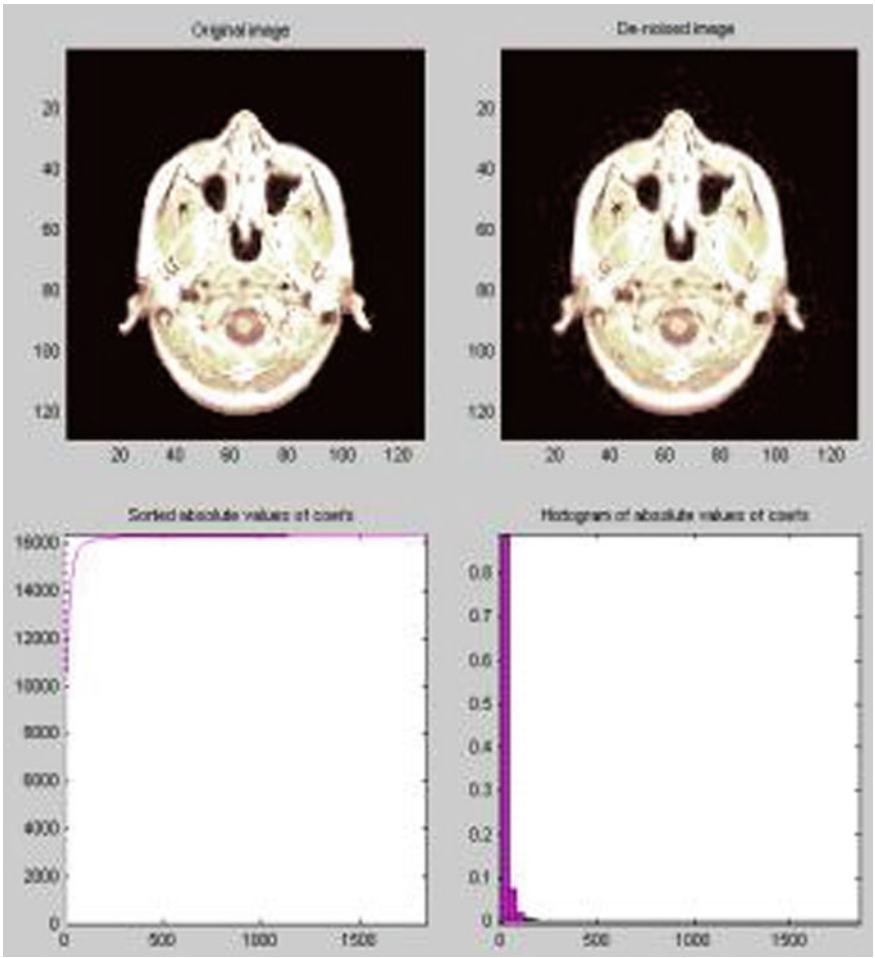


Fig. 11 De-noised image

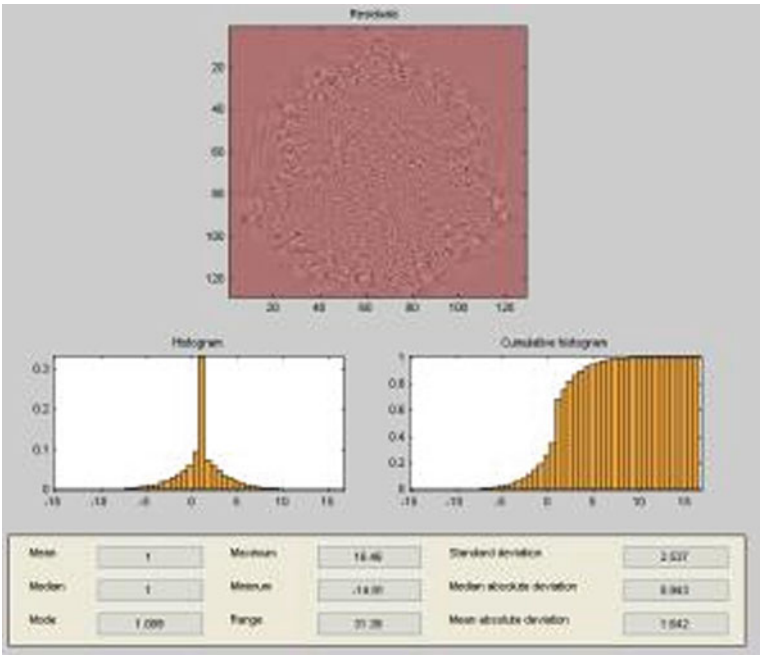


Fig. 12 Histogram of the image

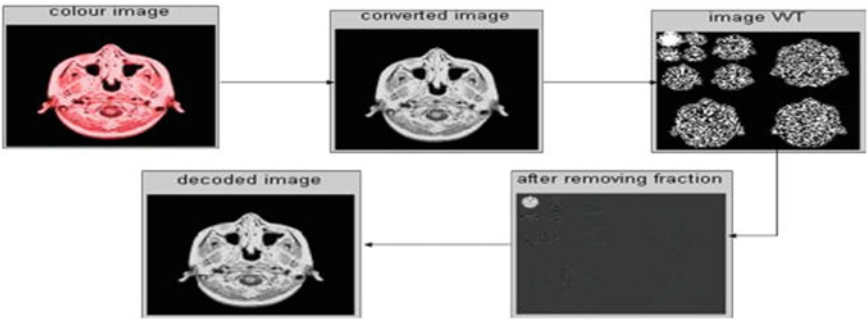


Fig. 13 Transmission of image by using DWT and IDWT

Fig. 14 Original image

Table 1 Test Results of lifting-based DWT and IDWT

Original pixel input	DWT output/IDWT input	IDWT output
0, 0, 0, 0	−65, 503, 332, 230	0, 0, 0, 0
0, 0, 0, 190	−95, 30, 283, 48	0, 0, 0, 190
0, 44, 218, 218	223, 289, −83, −102	0, 44, 218, 218
86, 218, 218, 200	0, 223, 66, −84	86, 218, 218, 200

4 Conclusions

The color image is converted into the grayscale image. The image is compressed. The image is compressed by 3 levels by using discrete wavelet transform and the compressed image is recovered by using inverse discrete wavelet transform. The original image is recovered. The peak signal-to-noise ratio is 55.62 and mean square error is 0.178. Number of coefficients recovered is 99.94 %. Standard deviation is 2.537. Median absolute deviation is 0.943. Mean absolute deviation is 1.642.

References

1. Gonzalez RD, RE Woods (2002) Digital image processing, 2nd edn, Prentice Hall of India Pvt Ltd
2. Lang R, Spray A (1995) Input buffering requirements of a systolic array for the inverse discrete wavelet transform. Proceedings international conference on application specific array processors, 24–26 July 1995, pp 166–173
3. Andra K, Chakrabarti C, Acharya T (2002) A VLSI architecture for lifting-based forward and inverse wavelet transform. IEEE Trans Signal Process 50(4):966–977

4. Acharya T, Tsai PS (2005) JPEG standard for image compression concepts, algorithms, and VLSI architectures. Wiley, New York
5. Yu C, Chen SJ (1999) Efficient VLSI architecture for 2-D inverse discrete wavelet transforms. IEEE international symposium on circuits and systems, pp 524–527, 30 May–2 June 1999
6. Merry RJE (2005) Wavelet theory and applications. A literature study, Eindhoven University of Technology, June 2005
7. Shukla PD (2003) Complex wavelet transforms and their applications for master of philosophy (m.phil.)
8. Sung TY, Hsin HC, Shieh YS, Yu CW (2006) Low-power multiplierless 2-D DWT and IDWT architectures using 4-tap Daubechies filters. 7th international conference on parallel and distributed computing, applications and technologies, pp 185–190
9. Muhit AAL, Islam MS, Othman M (2004) VLSI implementation of discrete wavelet transform (DWT) for image compression. 2nd international conference on autonomous robots and agents, pp 13–15
10. Motra AS, Bora PK, Chakrabarti I (2003) An efficient hardware implementation of DWT and IDWT. In: Conference on convergent technologies f or Asia-Pacific region, pp 95–99, 15–17 Oct 2003
11. Chang YN, Li YS (2001) Design of highly efficient VLSI architectures for 2-D DWT and 2-D IDWT. IEEE workshop on signal processing systems, pp 133–140, 26–28 Sept 2001

Identification of Common Indian Leafy Vegetables Based on Statistical Measures on Combined Color and Texture Features

Ajit Danti, Manohar Madgi and Basavaraj Anami

Abstract In this paper, identification of common Indian leafy vegetables based on statistical measures on combined color and texture features is presented. Different types of color and texture features are used to train the neural network. A BPNN-based classifier is used for identification of leafy vegetables. The identification rate is in the range 92–100 % for ten types of vegetables. Proposed method shows an improvement over the previous work. This work finds significance in various applications such as automatic vending, packing, grading of vegetables, food preparation.

Keywords Vegetable identification · Color features · Neural networks · Agricultural/horticultural produce

1 Introduction

Advancement in technologies is demanding more and more applications in computer vision and image processing. Different vision techniques are developed to meet different human needs. One of the areas is automatic grading of vegetables. Vegetables play a major role in Indian agriculture by providing food, nutritional, and economic security. There are certain computer vision systems developed for

A. Danti (✉)
J.N.N. College of Engineering, Shimoga, Karnataka 577204, India
e-mail: ajitdanti@yahoo.com

M. Madgi · B. Anami
K. L. E. Institute of Technology, Hubli, Karnataka 580030, India
e-mail: manohar.madgi@gmail.com

B. Anami
e-mail: anami_basu@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_38,

Fig. 1 Samples of common Indian leafy vegetables

agricultural/horticulture applications, namely weed identification, leaf disease identification, classification of grains and recognition of food in food industry, medical plant recognition, classification of crops and weeds. In order to automate the task of recognition, we have proposed different methodologies. Many types of leafy vegetables which are found in supermarket that are grown in India are shown in Fig. 1. The details in terms of English name, botanical name, and Indian common names are listed in Table 1. To know the state-of-the-art applications in the area of agriculture in general and vegetable identification in particular, a literature survey is carried out. The following is the gist of the survey.

Fernandez et al. [1] have developed an automated visual inspection system (AVIS) for quality control of preserved orange segments. The method adopted uses three-layer neural network perception topology.

Du and Sun [2] have developed an automated classification system of pizza sauce spread using color vision and support vector machines (SVMs). To characterize pizza sauce spread with low-dimensional color features, a sequence of image processing algorithms are developed. The image is transformed from red, green, and blue (RGB) color space to hue, saturation, and value (HSV) color space. A vector quantifier is designed to quantify the hue and saturation (HS) space to 256-dimensional space, and the quantified color features of pizza sauce spread are represented by color histogram.

Danti et al. [3] have developed a mean and range color features–based identification of common Indian leafy vegetables. The method adopted uses reduced set of color features. The leafy vegetables images have distribution of color and change in color over the image. These discriminating color features are represented by mean, variance and range values for the given image. They are extracted using equations and are used as an input to the developed neural network model, which is trained using delta learning technique. The highest identification rate is 100 %, and the lowest is 92 %. The average identification rate is 96.6 %.

Table 1 Details about the common Indian leafy vegetables

English name	Botanical name	Common name in India		
		Hindi	Kannada	Telugu
Green amaranth	*Amaranthus viridis*	Chaurai	Dhantina soppu/ rajgiri	Thota aaku
Coriander leaves	*Coriandrum sativum*	Dhania patta	Kotambari soppu	Kothimeera aaku
Curry leaves	*Murraya koenigii*	Karipatta	Karibevu soppu	Karivepaku
Dill	*Anethum graveolens* L.	Suwa	Sabsige soppu	Soya aaku
Fenugreek	*Trigonella foenumgraecum*	Methi	Menthya soppu	Menthi aaku
Sorrel leaves	*Rumex acetosa*	Meshta/ pitwaa	Pundi soppu	Gongura
Mint leaves	*Mentha spicacata*	Pudina patta	Pudina	Pudina aaku
Onion leaves	*Allium cepa*	Pyaaz	Eerulli soppu	Ullipaaya aaku
Red amaranth	*Amaranthus cruentus*	Lal chaurai	Kempu rajgiri	Koyya thota aaku
Spinach	*Spinacea oleracea*	Paalak	Palak soppu	Palak aaku

Nakano [4] have studied the application of a neural network to the color grading of apples. The whole image data collecting system was developed in this study in order to overcome unequal light intensity. This system consists of a turntable, a stage controller, and a mono-axle driver. On this system, the whole image of an apple is rolled out as one scene on a computer monitor like perspective projections. Two neural network models are proposed to classify pixels at any part of the surface of an apple into 'normal red,' 'injured red,' and 'poor red,' to grade the whole surface color of an apple into 'superior,' 'excellent,' and 'good.'

Patel et al. [5] have presented fruit detection using improved multiple features–based algorithm. It helps to locate the fruit on the tree. The algorithm is designed for calculating different weights for features like intensity, color, orientation, and edge of the input test image. The weights of different features represent the approximate locations of the fruit within an image. The detection efficiency is achieved up to 90 % for different fruit image on tree and captured at different positions.

Rocha et al. [6] have developed an automatic fruits and vegetables classification methodology from images. A unified approach combines many features and classifiers which require less training and is more adequate to establish image categories automatically using histograms, colors, and shape descriptors with an unsupervised learning method. The results show that the solution is able to reduce the classification; error is up to 15 % points with respect to the baseline.

Haralick et al. [7] have described some easily computable textural features based on gray-tone spatial dependencies. They have illustrated the application in category identification tasks of three different kinds of image data.

The literature survey has revealed that computers are applied for gradation of fruits like apple, grapes, mangoes in trading, in pattern recognition, and in

robotics. Study on leafy vegetables has not attempted. The present work of development of methodology is neural network–based model for recognition of leafy vegetables.

The paper is organized into five sections. Section 2 contains the proposed methodology and image acquisition. Section 3 describes feature extraction. Results and discussions are presented in Sects. 4 and 5 contains the conclusion of work.

2 Proposed Methodology

The proposed methodology comprises of three stages, namely image acquisition, feature extraction, and classification. The block diagram depicting these stages is given in Fig. 2.

2.1 Image Acquisition

We have considered ten different types of fresh leafy vegetable bunch of green amaranth, coriander leaves, curry leaves, dill, fenugreek, sorrel leaves, mint leaves, onion leaves, red amaranth, and spinach. These vegetables are placed against a plain white background vegetable type at a time, and images are taken. The images are acquired in the day light using a digital camera placed at a distance of 40 cm. A Sony digital camera with an image resolution of 10 mega pixels is used to capture the images. The images obtained are of size 3264 × 2448 pixels. The images are cropped and resized to 150 × 150 pixels for computational reasons.

3 Feature Extraction

Human beings identify the leafy vegetables using color, texture, and shape of the leaves. Color is the most important feature for identification of vegetables.

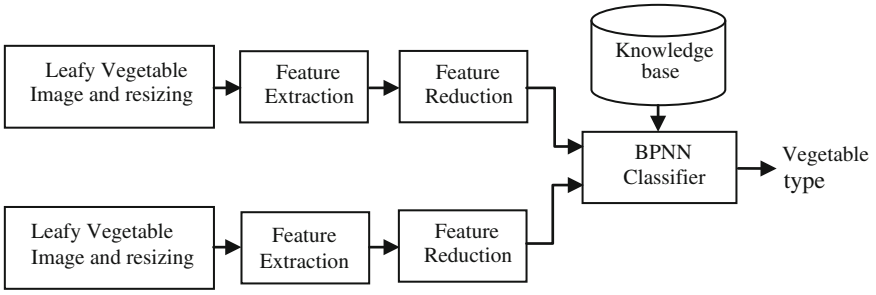


Fig. 2 Block diagram of proposed methodology

However, color is more or less green across most of the vegetables, but the shades of green color vary from vegetables to vegetables. The texture is another feature one can use to identify vegetables. In this work, we have considered combined color and texture features for identification of vegetables.

3.1 Color Feature Extraction

The values of R, G, and B color components from vegetable images are normalized so as to have the values in the range [0, 1]. The HSI color features are obtained from the RGB components.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$
$$\theta = \cos^{-1}\left\{ \frac{R - G/2 - B/2}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\} \tag{3.1}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R, G, B)] \tag{3.2}$$

$$I = \frac{1}{3}(R + G + B) \tag{3.3}$$

The leafy vegetables' images have distribution of color and change in color over the image. These discriminating features are represented by mean, variance, and range values for the given image. These 18 color features are used to characterize an image. All these 18 features are obtained by the Eqs. (3.4, 3.5 and3.6) for leafy vegetables.

$$\text{Mean } \mu = \sum_{x,y} xP(x, y) \tag{3.4}$$

$$\text{Variance } \sigma = \sum_{x,y} (x - \mu)^2 P(x, y) \tag{3.5}$$

$$\text{Range} = \max(p(x, y)) - \min(p(x, y)) \tag{3.6}$$

The mean and range features are adopted, since these satisfied the feature reduction strategy. The procedure to obtain these features is given in Algorithm 1.

Algorithm 1: Feature extraction and reduction
Input: Original 24-bit RGB color images of different vegetable samples.
Output: Reduced feature set of 12.
Start
Step 1: Input the leafy vegetable image.

Step 2: Separate R, G, and B components from the original 24-bit input color image.
Step 3: Obtain the color features hue (H), saturation (S), and intensity (I) components from RGB components.
Step 4: Extract all the 18 color features using the Eqs. (3.4, 3.5, 3.6)
Step 5: For each feature check

if (feature value >0.1) then 'It is a good feature'

Add the feature to reduced feature set.
else ignore.
Stop

3.2 Texture Feature Extraction

We have used different texture features. In our work, the gray-level co-occurrence matrix (GLCM) is used to carry out the analysis. GLCM become one of the most well known and widely used texture measures. Let a two-dimensional image $I(x, y)$, $(x = 1\ldots M, y = 1\ldots N)$, has N_g gray levels. A co-occurrence matrix depicts the joint gray-level histogram of the image or a region of the image in the form of a matrix with the dimensions of $N_g \times N_g$. The entries are the joint probability densities of pairs of gray levels that occur at pairs of points separated by the displacement vector d.

Suppose $P_d(i, j)$ denotes the cardinality of the set of pairs of pixels that have gray-level values of i and j, for a displacement vector $d = (d_x, d_y)$.

$$P_d(i, j) = |\{((r, s), (r + dx, s + dy)) : I(r, s) = i, I(r + dx, s + dy) = j\}| \tag{3.7}$$

where $(r, s) \in M \times N$ and $|.|$ is the cardinality of a set. The whole procedure of computing the co-occurrence matrix is given in the Algorithm 2.

The different types of texture features are extracted using GLCM method. The Eq. (3.8) is used to compute the co-occurrence matrix.

$$C = \frac{1}{4}(P_{0^\circ} + P_{45^\circ} + P_{90^\circ} + P_{135^\circ}) \tag{3.8}$$

$$\text{Contast}(k, n) = \sum_i \sum_j (i - j)^k P_d[i, j]^n \tag{3.9}$$

$$\text{Energy} = -\sum_i \sum_j P_d[i, j] \ln P_d[i, j] \tag{3.10}$$

Algorithm 2: Computation of co-occurrence matrix from the image *f(x, y)*
Input: Gray-level image *f(x, y)* of size M x N and with maximum gray level of L
Output: Co-occurrence matrix $P_{\varphi,d}(x, y)$, for $d = 1$ in all directions φ.
Start
Step 1: Assign $P_{\varphi,d}(x, y) = 0$, for all x, y belonging to [0, L]
Step 2: For all pixels (x_1, y_1) in the image, determine (x_2, y_2), which is at a distance d in direction φ of (0°, 45°, 90°, and 135°) and compute

$$P_{\varphi,\,d}|f(x_1, y_1), f(x_2, y_2)| = P_{\varphi,\,d}|f(x_1, y_1), f(x_2, y_2)| + 1$$

Step 3: Compute the $P_{\varphi,d} = \left(P_{0,d} + P_{45,d} + P_{90,d} + P_{135,d}\right)/4$
Stop.

We have selected only those features that have contributed to recognition of vegetables. This is how the features are reduced. The procedure adopted in obtaining the texture features is given in Algorithm 3. The Eqs. (3.7, 3.8, 3.9, 3.10) are being used in the Algorithm 3.

Algorithm 3: GLCM feature extraction
Input: RGB components of original image
Output: Texture features
Description: $P_{\varphi,d}(x, y)$ means GLCM matrices in the direction $\varphi = 0^{\circ}, 45^{\circ}, 90^{\circ}$, and 135°
Start
Step 1: For all the separated RGB components, derive the co-occurrence matrices $P_{\varphi,d}(x, y)$ for four direction ($\varphi = 0^{\circ}, 45^{\circ}, 90^{\circ}$ and 135°) and $d = 1$
Step 2: Calculate co-occurrence features like contrast, entropy, using the Eqs. (3.7, 3.8, 3.9, 3.10).
Stop

3.3 Classifier Parameters

A three-layer feedforward backpropagation artificial neural network (BPNN) is implemented using neural network toolbox of MATLAB version R2009a. The mean square error is set to 0.001. There are 24 nodes in the input layer. The hidden layer has 48 nodes. The output layer has 10 nodes, corresponding to the 10 different vegetable classes. The learning rate is set to 0.01.

4 Result and Discussion

The proposed method for leafy vegetable identification is tested on 1,000 images of leafy vegetables' bunches of 10 different types. 500 images with 50 images

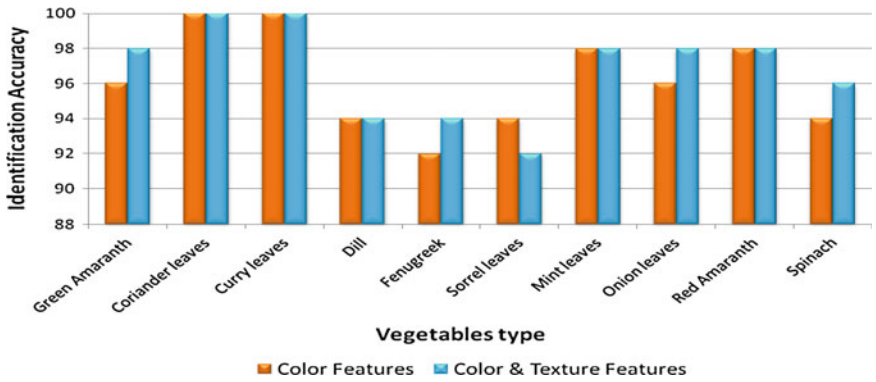


Fig. 3 Identification accuracy for different vegetables based on color and combined color and texture features

from each type are used to train the neural network classifier; the remaining 500 images are used to test the accuracy.

We have trained and tested the neural network considering all 18 color features, namely mean, variance, and range of RGB and HSI color models and 24 texture features. During experimentation, it is observed that some of the color features and texture features have very low values which do not contribute to the accuracy of the classifier. Hence, we have reduced the feature set to 12 color features and 12 texture features. These features are combined together. The identification accuracy of different types of leafy vegetables using color features and combined color and texture features is shown in Fig. 3. It is observed from the graph that the highest identification rate observed is 100 % and the lowest observed is 92 %. The average identification rate using color features is 96.6 % and using combined color and texture feature is 96.8 % which is comparable and efficient.

5 Conclusion

In this paper, we have proposed a combined color and texture feature–based classification of leafy vegetables. A BPNN classifier is used for the identification and classification. The maximum identification rate of 100 % is observed with coriander and curry leaves, and a minimum of 92 % is observed with sorrel leaves. Improvement of 0.2 % has been is observed over color features work. The work carried out has relevance to real-world applications in automatic identification of the different vegetable types, their grading, valuation, packing, sales in the supermarket, food preparation, and the like.

References

1. Fernandez C, Suardiaz J, Jimenez C, Navarao PJ, Toledo A, Iborra A (2002) Automated visual inspection system for the classification of preserved vegetables. In: Proceedings of the IEEE international symposium, vol 1, pp 265–269
2. Du CJ, Sun DW (2005) Pizza sauce spread classification using color vision and support vector machines. J Food Eng 66(2):137–145
3. Danti A, Madgi M, Anami BS (2012) Mean and range color features based identification of common Indian leafy vegetables. Int J Signal Process Image Process Pattern Recogn 5(3):151–160
4. Nakano K (1997) Application of neural networks to the color grading of apples. Comput Electron Agric 18:105–116
5. Patel HN, Jain RK, Joshi MV (2011) Fruit detection using improved multiple features based algorithm. J Comput Appl (0975–8887), 13(2):1–5
6. Rocha A, Hauagge DC, Wainer J, Goldenstein S (2010) Automatic fruit and vegetable classification from images. Comput Electron Agric 70:96–104
7. Haralick RM, Shanmugam K, Dinstein I (1973) Texture features for image classification. IEEE Trans Syst Man Cybern 6:610–621

Fast Pattern Matching Approach for Intrusion Detection Systems

M. Manjunath, K. G. Srinivasa and Anil Kumar Muppalla

Abstract Intrusion detection system (IDS) consists of set of techniques and methods for collection of packets from host system or network and analyzes those packets for anomalous content. IDSs mainly fall into two categories: signature-based IDSs and anomaly detection systems. A rule-based IDS compares the incoming packets against rule set in order to detect intrusion. A common approach followed is to build rule trees or finite automata with rule set and traverse it using a packet as input string. 30–60 % of total signature-based IDS processing time is spent on pattern matching [1]. The existing signature-based IDS cannot meet the speed demands imposed by both high network speeds and increasing number of signatures, and more CPU time is spent on searching for rules that match each packet. In this paper, we are going to present an analysis on IDS that is combined with other methods and techniques to produce greater results and hence contribute to the improvement of IDS.

Keywords SNORT · Pattern matching · Wu–Manber · Bloom filter

1 Introduction

An intrusion detection system (IDS) monitors and collects data from a target system that should be protected, processes, and correlates the gathered information, and when evidence of an intrusion is detected, it initiates responses. IDSs fall

M. Manjunath (✉) · K. G. Srinivasa · A. K. Muppalla
Machine Learning and Data Mining Laboratory, Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India
e-mail: manju.mg7@gmail.com

K. G. Srinivasa
e-mail: kgsrinivasa@gmail.com

A. K. Muppalla
e-mail: anil.kumar.848@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_39,

into two basic categories: signature-based IDSs and anomaly detection systems. Intruders have signatures, like computer viruses, that can be detected using software. We try to find data packets that contain any known intrusion-related signatures or anomalies related to Internet protocols. These signatures and rules help in finding and logging suspicious activity and thus generate alerts. Anomaly-based intrusion detection usually depends on packet anomalies present in protocol header parts. In some cases, these methods produce better results compared to signature-based IDS. Usually, an IDS captures data from the network and applies its rules to that data or detects the anomalies in it. Because of the increasing amount of traffic and threats, intrusion detection has become very resource intensive. IDS is found to consume lot of CPU time and memory due to high-speed networks and large rule sets. IDS is under the threat of allowing malicious packets to enter the network when there is a lack of CPU time. Therefore, reducing CPU time consumption in packet matching is crucial for overall intrusion detection performance. We look at implementing different pattern matching algorithms to improve the speed of intrusion detection at high speeds.

2 Related Work

In 2005, Song et al. [2] proposed a new lookup algorithm which speeds up the HASH tables where they modified bloom filters to provide exact matches. The speedup comes from cutting the number of hash collisions and using a small external memory. The main disadvantage is the need for an expensive cache like on-chip memory. Dharmapurikar et al. [3] presented a new hardware pattern matching algorithm by combining bloom filters and classic AC algorithm. They implemented bloom filters using embedded on-chip memory blocks in FPGA. Chaudhary et al. [4] proposed a new hardware parallel architecture in 2010 based on bloom filters. It can test input strings simultaneously to detect possible attacks with less delay. It computes hash functions in parallel on all test strings and can AND bit location values in the lookup array separately for different hashes of different strings.

On the other hand, software-based techniques are cheaper and easy to reconfigure. However, they are unable to match the increasing network speeds and memory constraints. Anagnostakis et al. [5] proposed a new exclusion filter in 2003 for IDS called E2xB. It is an exclusion-based filter that preprocesses the packet by using a 256-cell map to mark the existence of each character of patterns within the packet. If at least one character of a pattern is not marked, it means that this pattern does not exist in that packet. Chen et al. [6] proposed a new system by combining both fingerprinting and pattern matching techniques to accelerate pattern matching for IDS. In the programming phase of the fingerprinting, the system generates a short digest for each pattern. Then, it computes the digests of the incoming packets to match them against those of the patterns. BM algorithm is then used to perform exact pattern matching. Ramakrishnan et al. [7] proposed an

efficient and scalable system in 2010 called sigMatch to improve multi-pattern matching algorithms. It preprocesses all the patterns to organize them in a q-gram index structure called sigTree according to common subpatterns. The sigTree enhances the feature of bloom filters that results in consumption of less memory. The idea is to match all the patterns against the sigTree to discard the unmatched ones, and the remaining patterns are sent to a verification unit to check their presence.

2.1 SNORT

Snort is an open source IDS. When configured as an intrusion prevention system (IPS), it is used for monitoring and prevention of security attacks on networks. As an IDS, Snort may detect security attacks and alert the system administrator or take the designated action like logging the malicious packets. As an IPS, it modifies the iptable to drop or reject the suspected packets/stream and thus tries to prevent. Snort was originally developed by Martin Roesch, founder of Sourcefire Inc., Maryland. Snort uses a rule-driven language. It combines the benefits of signature-, protocol-, and anomaly-based inspection methods [8]. With more than three million of downloads to date, Snort is the most widely deployed intrusion detection and prevention technology worldwide. Snort is capable of signature anomaly detection as well as protocol anomaly detection. The rule set of Snort is very powerful and flexible, and it is easy to understand and construct new rules. The examples below show the rules for working with Snort in various attack scenarios:

> "alert tcp $EXTERNAL_NET any $HTTP_SERVERS 80 (msg:"WEB-IS cmd.exe access"; flags: A+; content: "cmd.exe"; nocase; classtype:web-application-attack; sid: 1002; rev:2;)"

The following rule detects any scan attempt using SYN-FIN TCP packets.

The flags keyword is used to find out which flag bits are set inside the TCP header of a packet.

> "alert tcp any -> 192.168.1.0/24 any (flags: SF; msg: "SYNC-FIN packet detected";)"

where **alert**—generates an alert using the selected alert method and then log the packet. The next field in a rule is the protocol. There are four protocols that Snort currently analyzes for suspicious behavior: TCP, UDP, ICMP, and IP. The next portion of the rule header deals with the IP addresses and port information for a given rule. The keyword any may be used to define any address. Port numbers may be specified in a number of ways, including any ports, static port definitions, ranges, and by negation. The direction operator > indicates the orientation, or direction, of the traffic that the rule applies to. The IP address and port numbers on the left side of the direction operator are considered to be the traffic coming from the source.

2.2 Pattern Matching

The pattern matching algorithms pertaining to the general keyword pattern matching problem are the ones of particular interest. In single pattern matching, given an input string P and a pattern string T, whether T appears in P? Whereas in multi-pattern matching, given an input string P and a set of pattern strings T1; T2Tm, whether any Ti appears in P? Multi-pattern matching algorithms for the basis for many IDS signature matching engines. They allow these engines to quickly search for many patterns simultaneously in input passing through such systems, but often consume most of the processing time. Thus, they should be as fast as possible to ensure system scalability into networks of ever-increasing speeds. Concurrently, they must enforce security so that they are not susceptible to algorithmic complexity attacks. Snort in earlier days used brute force pattern matching which was very slow and was place for improvement to be done. To boost performance, implementation of Boyer–Moore pattern matching algorithm was done. Later Snort's detection engine was included with "linked list of function pointers," also called a "three-dimensional linked list."

2.3 Aho–Corasick Algorithm

There are many approaches to recognizing patterns that involve using finite automata (also referred to as finite state machines). The Aho–Corasick (AC) algorithm [9] is one such classic algorithm. This algorithm also shares characteristics with the Knuth–Morris–Pratt [10] algorithm described. The idea is that a finite automaton is constructed using the set of keywords during the precomputation phase of the algorithm, and the matching involves the automaton scanning the input text string reading every character in y exactly once and taking constant time for each read of a character. It is essential to first understand finite automata theory to understand the AC algorithm's description. The notation we use for the AC automaton is a 7-tuple notation (Q, $q0$, A, δ, g, f, o) where,

- Q is a finite set of states,
- $q0$ belongs to Q is the start (initial) state,
- Q and is the set of accepting states,
- δ is the input alphabet accepted,
- g is a function from Q x δ into Q, called the good (or goto) transition function,
- f is function from Q into Q, called the fail (or failure) transition function and is a function from Q into Q, called the output function.
- O is output function, returns whether or not any state q belongs to A.

If the automaton is in a state q and reads input character (byte) a, it moves (transitions) to state $g(q, a)$ if defined; otherwise, it moves to state $f(q)$. Also if the automaton is in a state q and q belongs to the set A, then q is said to be an

Fig. 1 Aho–Corasick algorithm

```
Procedure AC(y,n,q0) state = q0
  For i is 1 to n do
    While g(state, y[i]) = fail do
      state = f(state)
    end while
    state = g(state, y[i])
    if o(state) • NULL then
      output i
    end if
  end for
end procedure
```

accepting state. Aho and Corasick's original algorithm uses a function called output to test this and furthermore returns the keyword matched at the accepting state. The AC algorithm's automaton is such that a transition into an accepting state indicates a match of one or more keywords [9]. The pseudocode for the matching phase of the algorithm is given below in Fig. 1.

2.4 Wu–Manber Algorithm

Wu and Manber created the UNIX tool agrep [11] to search for many patterns in files. They claim and demonstrate that their tool is substantially faster than both the UNIX tools egrep and fgrep. fgrep initially used the AC algorithm, and subsequently, it was upgraded to use a modified Commentz–Walter [12] algorithm. agrep is shown to match tens of thousands of patterns faster than both other tools. Require:

- $Y =$ array of n bytes representing the text input,
- $N =$ integer representing the text length,
- $q0 =$ initial state.

The users can configure the string sequences to search and match the packet is configured using the configuration interface. Wu–Manber algorithm works in two phases.

2.4.1 Preprocessing Process

WM computes the minimum length of the patterns (m) and works on blocks of size B, and WM constructs three hash tables: SHIFT, HASH, and PREFIX. The SHIFT table maps each substring of size B characters to a shift value representing the number of characters to skip on a mismatch. Let X be a block of characters, then there are couple of scenarios: If X does not appear in any pattern at all, then the shift value will be the default value of SHIFT[i] $= m-B+1$ characters. The second scenario is if X appears in one or more patterns, then the shift value will be

SHIFT[i] = $m-q$, where q is rightmost position that X occurs in any pattern. WM calculates the shift values by mapping each substring of any pattern Pi of size B to the SHIFT table. The HASH table is indexed by the same hash function used for the SHIFT table for faster access, and it aims to prevent comparing a substring to all patterns in the pattern list. For each character block with zero shift value, the HASH table lists all signatures containing that block. The PREFIX table is used to speed up the HASH table search and contains the hash values for the prefixes of the patterns in the HASH table.

2.4.2 Search Phase

In the search phase, WM hashes the first block by dividing the incoming packet in a sliding window fashion. Next, it checks the SHIFT table to find a corresponding shift value. There are two possibilities: If the shift value is greater than zero, then the sliding window is shifted by that value and a new hash value is computed for the new block. The second possibility is zero shift value indicating a match possibility, and then, both HASH and PREFIX tables are used to verify whether there is an actual match.

2.5 Bloom Filters Theory

The term "bloom filters" was first coined by Burton Bloom in 1970. The bloom filter creates a hash vector representation of strings which can easily exclude negative matches. The filter preprocesses the strings by computing k hash values ranging from 1 to m for each string. The bits corresponding to the hash values computed for all strings are set in an m-bit-long vector [13]. To search a packet for the strings, the filter is checked by computing the same k hash values on the packet in a sliding window fashion. Then, the corresponding bits of the vector are examined to determine whether the given item exists or not. Bloom filters have no false negatives, which means that if at least one bit is not set, it means that the item is not 100 % certain that it is a member of the filter. On the other hand, if all the bits are set in the vector, then the item belongs to that string set with a certain probability. Exact pattern matching must be used to verify whether the item actually belongs to the given set or not.

3 Proposed Design

By combining the bloom filters and Wu–Manber pattern matching algorithm (as shown in Fig. 2), we can achieve faster pattern matching for intrusion detection. Bloom filter can be modified to provide probable matches so that it reduces the

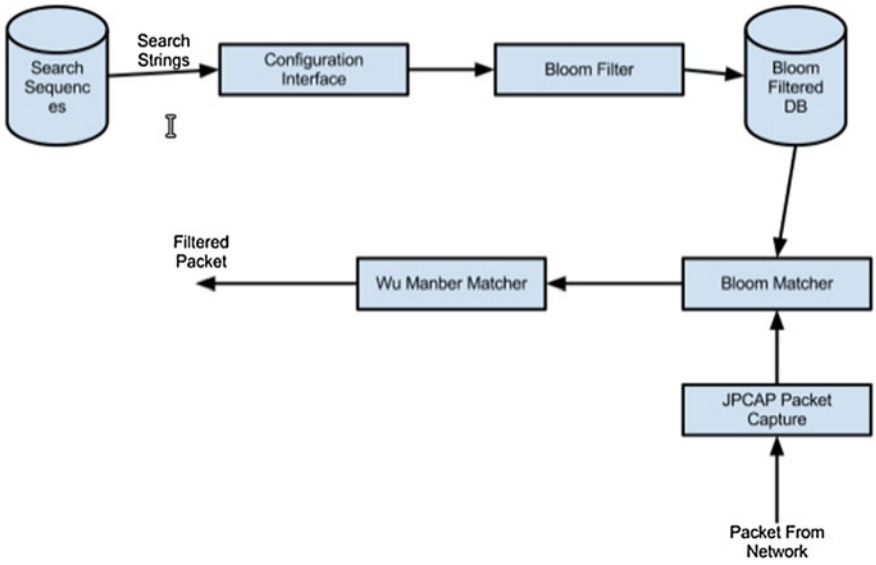


Fig. 2 Design of proposed system

number of pattern matching operations required for suspicious packets. The bloom filter can be programmed to take only prefix data from Snort signatures rule set, so that the filter processing overhead is reduced to minimum and hence speeds up the matching by skipping clean packets. The selected suspicious packets are passed to Wu–Manber algorithm so that the pattern matching is carried out on the respective packets with signature matches of Snort rule set.

4 Conclusion

Increasing connectivity and network data rates are the two buzzwords which are driving an ever-increasing computational demand for processing workloads such as pattern matching. To support the increasing demand, we intend to provide fast software-based pattern matching method to reduce the number of times to perform pattern matching. It is a new and fast software-based pattern matching algorithm to accelerate signature-based intrusion detection. It excludes clean traffic without the need to do pattern matching by using with the signature prefix. It brings out the best result when the features of both bloom matcher and the Wu–Manber matcher are applied in combination. As a future work, we are intending to probe into the disadvantages of this combination and what better can be brought out by other fast pattern matching methodologies.

Acknowledgements This project is financially supported by DRDO sponsored project titled Machine Learning Techniques for Data Mining-Based Intrusion Detection Systems (Ref. No.: ERIPR/ER/0705066/M/01/1256) to Dr. Srinivasa K G, Professor, Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India. We acknowledge Dr. T V Suresh Kumar, Dr. K Rajanikanth, Dr. D E Geetha, Mrs. Mrunalini M, and Mr. Manish Kumar for their kind support.

References

1. Aldwairi M, Alansari D (2011) Exscind: Fast pattern matching for intrusion detection using exclusion and inclusion filters. IEEE
2. Song H, Dharmapurikar S, Turner J, Lockwood J (2005) Fast hash table lookup using extended Bloom filter: an aid to network processing. In: SIGCOMM proceedings on conference on applications, technologies, architectures, and protocols for computer communications
3. Dharmapurikar S, Lockwood J (2006) Fast and scalable pattern matching for network intrusion detection systems. Sel Areas Commun IEEE J 24(10):1781–1792
4. Chaudhary D (2010) Parallel processing of bloom filter. Int J Electron Eng Res
5. Anagnostakis K, Antonatos S, Markatos E, Polychronakis M (2003) E2xB: a domain-specific string matching algorithm for intrusion detection. In: the proceedings 18th IFIP international information security conference (SEC)
6. Chen Z, Zhang Y, Chen Z, Delis A (2009) A digest and pattern matching-based intrusion detection engine. Comput J 52(6):699–723
7. Ramakrishnan K, Nikhil T, Jignesh M (2010) SigMatch: fast and scalable multi-pattern matching. In: 36th international conference on very large data bases (PVLDB)
8. http://www.sourcefire.com/security-technologies/open-source
9. Aho A, Corasick M (1975) Efficient string matching: An aid to bibliographic search. Commun ACM 18(6):333. doi:10.1145/360825.360855
10. Knuth DE, Morris JH, Pratt VR (1977) Fast pattern matching in strings. SIAM J Comput 6(2):323
11. Wu S, Manber U (1992) Agrep—A fast approximate pattern-matching tool. In: Proceedings USENIX Winter 1992 technical conference, San Francisco, pp 153–162
12. Commentz-Walter B (1979) A string matching algorithm fast on the average. In: Proceedings of the 6th colloquium on automata, languages and programming, Springer, London, UK, p 118â€.132
13. Bloom B (1979) Space/time trade-offs in hash coding with allowable errors. Commun ACM

A Comparative Analysis of Energy-Efficient Routing Protocols in Wireless Sensor Networks

K. Narendra, V. Varun and G. H. Raghunandan

Abstract A wireless sensor network (WSN) consists of spatially distributed autonomous sensors for reliable monitoring of a variety of environments for both civil and military application and to cooperatively pass their data through the network to a main location. Routing protocols have significant impact on the overall energy consumption of sensor networks. Suitable energy-efficient routing algorithms are required to the inherent characteristics of these types of networks. Due to resource limitations in wireless sensor networks, prolonging the network lifetime has been of a great interest. Most of the energy of sensor nodes is utilized for the transmission of data to the base station. Thus, it makes them to deplete their energy much faster. In this paper, we look at and compare the different communication protocols, which can have significant impact on the overall energy dissipation of these networks.

Keywords Clusters · Cluster head · Movable base station · Mobile sink · Nodes · Wireless sensor networks (WSN)

1 Introduction

The WSN is built of nodes from a few to several hundreds or even thousands, where each node is connected to one (or sometimes several) sensor. These sensor nodes have certain limits, such as low power, limited processor, limited memory, and

K. Narendra (✉) · V. Varun · G. H. Raghunandan
Department of Electronics and Communication Engineering, Global Academy of Technology, Visvesvaraya Technological University, Bangalore, India
e-mail: narenkram7@gmail.com

V. Varun
e-mail: varunv301@gmail.com

G. H. Raghunandan
e-mail: raghunandangh@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_40,

communication-restricted, and have a wide range of applications such as military, battlefield, and environment monitoring. Each sensor node is based on its decision on transmission, the information it currently has, knowledge of its computing, communication, and energy resources, and have capability to collect and route data either to other sensors or back to an external base station or stations which may be a fixed or a mobile node capable of connecting the sensor network to the Internet or an existing communication infrastructure or to the Internet where users have access to the reported data. The lifetime of a wireless sensor network is limited to the battery lifetime of the sensor nodes. Many energy-efficient protocols or schemes have been proposed for sensor networks in recent years.

One of the conventional protocols in use was direct communication protocol. Using a direct communication protocol, each sensor sends its data directly to the base station. If the base station is far away from the nodes, direct communication will require a large amount of transmit power from each node. This will quickly drain the battery of the nodes and reduce the system lifetime. However, the only reception in this protocol occurs at the base station, so if either the base station is close to the nodes, or the energy required to receive the data is large, this may be an acceptable (and possibly optimal) method of communication. The second conventional protocol we consider is a "minimum-energy" routing protocol. In these protocols, nodes route data destined ultimately for the base station through intermediate nodes. Thus, nodes act as routers for other nodes' data in addition to sensing the environment. These protocols differ in the way the routes are chosen. Some of these protocols only consider the energy of the transmitter and neglect the energy dissipation of the receivers in determining the routes. Then, there are some other protocols like LEACH, HUMS, Average Energy-based Routing Protocol and the Centrality-based Cluster approach. In this paper, we discuss in brief about these protocols and do extensive comparison between them.

Section 2 gives the classification of routing protocols. Section 3 of the paper deals with brief description of four routing protocols. Section 4 is the comparison between the protocols and briefs the advantage of the Centrality-based Cluster approach. Future research is explained in Sect. 5, and finally, Sect. 6 gives the conclusion of the paper.

2 Classification of Routing Protocols

The design space for routing algorithms for WSNs is quite large, and we can classify the routing algorithms for WSNs in many different ways. In general, routing in WSNs can be classified into flat-based routing, hierarchical-based routing, and location-based routing depending on the network structure.

The first category is the flat-based or the data-centric routing protocols. In WSNs, node-centric communication is not a commonly expected communication type. Therefore, routing protocols designed for WSNs are more data-centric or geocentric. In data-centric routing, the sink sends queries to certain regions and

waits for data from the sensors located in the selected regions. Since data are being requested through queries, attribute-based naming is necessary to specify the properties of data. Here, data are usually transmitted from every sensor node within the deployment region with significant redundancy.

The second type is the hierarchical- or cluster-based routing protocols, proposed in wireless networks. A hierarchical routing protocol is a natural approach to take for heterogeneous networks where some of the nodes are more powerful than the other nodes, that is, the higher-energy nodes can be used to process and send the information, while lower-energy nodes can be used to perform the sensing in the proximity of the target. The hierarchy does not always depend on the power of nodes. In these protocols, different nodes are grouped to form clusters and data from nodes belonging to a single cluster can be combined (aggregated).The clustering protocols have several advantages like scalable, energy efficient in finding routes, and easy to manage.

The third category is the location-based routing protocol. In this method, the nodes know where they are in a geographical region. The incoming signal strengths can estimate the distance between neighboring nodes. Relative coordinates of neighboring nodes can be obtained by exchanging such information between neighbors. An alternate for this is that the location of nodes may be available directly by communicating with a satellite, using global positioning system (GPS), if nodes are equipped with a small low power GPS receiver. Some location-based schemes demand that nodes should go to sleep if there is no activity to save energy. The more the number of sleeping nodes, the more the energy saving obtained in the network. Thus, they are used to improve the performance of routing and to provide new types of services.

3 Hierarchical-Based Routing Protocols

3.1 LEACH: Low-Energy Adaptive Clustering Hierarchy

LEACH [1] is an autonomous adaptive clustering protocol that distributes the energy load evenly among the sensors in the network using randomization. The nodes organize themselves into local clusters, with one node acting as the local base station or cluster head. In order to not drain the battery of a single sensor randomized rotation of the high-energy cluster head position such that it rotates among the various sensors. Once the cluster head has all the data from the nodes in its cluster, the cluster head node aggregates the data and then transmits the compressed data to the base station. Since it compress the amount of data being sent from the clusters to the base station, energy dissipation is reduced and system lifetime is thus enhanced.

3.2 HUMS: Half-qUadrant–Based Moving Strategy

In this method, it uses a moving strategy where a sink visits each sensor node to collect sensed data [2]. A mobile sink makes moving decision based on HUMS. The sink moves toward the nodes, which has the highest residual energy and starts collecting the data. The node and sink communication include three phases. In the first phase, sink sends position notification to the nodes and then the nodes send their respective data to the sink. In the second phase, it sends by multi-hop, and in the final phase, the sink moves to a new position based on an algorithm and the new reporting commences. By using HUMS, the sink can move on its own by adapting to the diverse node deployments. Thus, energy dissipation is checked and system lifetime is enhanced considerably. But this protocol can be applied only for an event-driven application.

3.3 An Average Energy-Based Routing Protocol for Mobile Sink

This protocol [3] is an energy-efficient cluster-based mechanism. It has been designed to improve the energy efficiency with decreasing communication. It adopts a Sink Routing Schedule Scheme where it randomly selects a cluster head whose remaining energy is above the threshold value. This protocol uses Hello Packet as an indication to the visited cluster by the mobile sink and Information Packet where the cluster head collects the information and transmits it to the sink. Since it follows a queuing procedure, the mobile sink will not be able to move to the cluster head region immediately where there is a requirement if nodes are about to die.

3.4 Centrality-Based Cluster Approach

Here, the Centrality-based Cluster Approach [4] is taken into consideration where a movable base station is used to reduce the energy consumption of the cluster heads. In this approach, a node which is at equal distances from all other nodes in the cluster is elected as the cluster head. According to the algorithm proposed, the base station is moved to a specific cluster head and reduces the CH's energy consumption in each of the rounds. Decision to which critical CH it should move is made by the fuzzy system at the base station. The fuzzy logic decision is based on three parameters:

Table 1 Comparison of Routing Protocols in WSN

Routing protocols	LEACH	HUMS	AERP	CBCA
Classification	Hierarchical	Hierarchical	Hierarchical	Hierarchical
Movable base station	No	Yes	Yes	Yes
Cluster heads	Yes	Yes	Yes	Yes
Energy consumption	Highest	Lesser than AERP	Lesser than LEACH	Lowest
Lifetime of nodes	Low	Moderate	Moderate	High
Data aggregation	Yes	Yes	Yes	Yes
Energy dissipation	High	Moderate	Moderate	Low
Variance of energy	High	Moderate	Moderate	Low

- CH residual energy.
- Number of nodes in the cluster.
- Distance from the base station to CH.

4 Comparison of Routing Protocols

In this paper, we have compared the following routing protocols according to their design characteristics:

- LEACH [1]: Low-Energy Adaptive Clustering Hierarchy.
- HUMS [2]: Half-qUadrant-based Moving Strategy (HUMS).
- AERP [3]: Average Energy Routing Protocol.
- CBCA [4]: Centrality-based Cluster Approach (Table 1).

5 Future Research

Future works in routing techniques focus on different directions. In comparison with the above-mentioned routing protocols, the variance of energy is high in the LEACH, moderate in HUMS and AERP, whereas in CBCA, it is low. Although CBCA is relatively better when compared to other routing protocols, there is a need to make the variance of energy more stable. Also, prolonging the network lifetime has been of a great interest due to resource limitations in wireless sensor networks.

There needs to be a significant research done to improve the lifetime of the wireless sensors. Works on reducing the energy consumption by the nodes need to be done in a more efficient manner.

6 Conclusion

Prolonging the network lifetime has been of a great interest due to resource limitations in wireless sensor networks. Here, we have done a comparative analysis of energy efficient routing protocols, namely LEACH, HUMS, AERP, and CBCA. By comparing, we can observe that the energy dissipation and energy consumption of the sensors are lesser in CBCA than the other routing techniques. When we take into consideration the energy variation, it is highest in LEACH, moderate in AERP and HUMS, but better stable in CBCA. As found in the comparative study based on the above parameters, lifetime of the nodes is more in CBCA. Although the CBCA routing protocol looks promising, there is a need to explore the practicality of the technique employed. We have highlighted some of the future challenges and issues in improving the energy efficiency of sensors.

References

1. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceeding of 33rd Hawaii international conference on system sciences 2000, vol 8. pp 8020
2. Sun L, Bi Y (2009) A moving strategy for mobile sinks in wireless sensor networks. IJSNet 5(3):173–184
3. Wang YH, Yu CY, Chen WT, Wang CX (2008) An average energy based routing protocol for mobile sink in wireless sensor networks. In: First IEEE international conference on Ubi-media computing. pp 44–49
4. Raghunandan GH, Metri S (2012) A novel approach to increase overall efficiency in wireless sensor networks. In: IEEE international conference on computing, electronics and electrical technologies 2012. pp 699–703

Author Biographies

K. Narendra received B.E degree in Electronics and Communication from Visvesvaraya Technological University in 2012. His area of interest includes computer networks, wireless communication, and programming languages.

V. Varun received B.E degree in Electronics and Communication from Visvesvaraya Technological University. His area of interest includes wireless communication, wireless power transmission, and computer networks.

G. H. Raghunandan received B.E degree in Electronics and Communication Engineering from Visvesvaraya Technological University. He did his M.Tech in the field of Digital Electronics and Communication from Visvesvaraya Technological University, Belgaum. He is currently working as an Assistant Professor in Department of Electronics and Communication at Global Academy of Technology, Bangalore. His areas of interest are Wireless Ad-Hoc and Sensor Networks. He has published 10 research papers at International Conferences, which includes IEEE and IISc conferences, one paper in International Journal and 18 papers in National Conferences.

Implementation of an Ethernet Bridge Using Avalon Memory-Mapped Interface

Akshay S. Bharadwaj and Deepak Mankkadan

Abstract Network bridging is a method of transferring data between two PCs or embedded boards using a bridge or switch. Here, we use the Ethernet network standard to produce a bridge. The traditional way of implementing an Ethernet bridge using an FPGA is to use a single NIOS II soft processor and define the hardware using Verilog. Here, we propose a method of bridging using two different NIOS II processors: The hardware is defined in Verilog and the protocol is implemented using C language. This gives higher speeds and reduces latency. The design has been prototyped and verified on Altera DE2-115 board which has Cyclone IV E device.

Keywords Field-programmable gate array (FPGA) · Ethernet · NIOS II · Verilog · Qsys · Triple speed ethernet media access control (TSE MAC)

1 Introduction

A field-programmable gate array (FPGA) is an integrated circuit that allows the customer to configure it as per his needs. The configuration is specified in hardware description language (HDL) which programs the logical elements as per the application. FPGAs can be used to implement any logical function that an ASIC can perform. Communication between the boards with ASIC or FPGA is the key to most of the applications. Over the years, Ethernet has become a standard for most of the board-to-board communication. Ethernet is a computer networking technology

A. S. Bharadwaj (✉) · D. Mankkadan
Design Engineer, Redmount Consultancy, Bangalore, India
e-mail: akshy.bharadwaj@gmail.com

D. Mankkadan
e-mail: deepak87_m@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_41,

for local area networks (LANs). Ethernet has largely replaced competing wired LAN technologies. The IEEE 802.3 is the standard for the Ethernet protocol. Data rates have periodically increased from the original 10 Mbps to 100 Gbps. One of the important use of Ethernet in an board is for network bridging. Network bridging can be described as the action taken by network equipment to allow two or more communication networks, or two or more network segments, creating an aggregate network. A network bridge is a network device that connects more than one network segment. In the OSI model, bridging acts in the first two layers. There are four types of network-bridging technologies: simple bridging, multiport bridging, learning or transparent bridging, and source route bridging.

2 Problem Definition

The Altera NIOS II soft processor reduces the throughput of the data transfer as it has to process the data that are coming from two different Ethernet interfaces. The data from the Ethernet interfaces (PHY) are sent to the TSE MAC which is controlled by single NIOS II processor. TSE MAC acts as the second layer above that, we have NicheStack Transmission Control Protocol/Internet Protocol (TCP/IP) stack instead of the TPC/IP suite which is the small memory footprint TCP/IP layer. This is designed to reduce the resource usage of the NIOS II processor and also provides the full feature of the TCP/IP stack. Since the NicheStack has lot of instructions, threads to execute, it consumes lot of processor resources which in turn reduces the throughput and increases the latency (processing time).

When we implement two NIOS II soft processors, we can independently control one TSE MAC using one NIOS II processor and the other TSE MAC with the second processor. In this design, we will be able to save 50 % of the processor's resources as one of the TSE MAC is being controlled by another processor. As a result, the speed is increased and so is throughput.

3 Design and Verification of Ethernet Bridge

3.1 Hardware Design

We use Qsys tool which is a part of Quartus software provided by Altera. The Qsys tool allows us to add the hardware components we require and make the necessary connections between the components. When the design is complete, the tool generates the Verilog or VHSIC hardware description language (VHDL) code and also the require test bench. The hardware design is as follows:

We add a NIOS II soft processor, [1] phase-locked loop (PLL), joint test action group/universal asynchronous receiver/transmitter (JTAG/UART), timer, and a

Parallel I/O. The clock into the PLL is given from the 50 MHz NIOS clock, and the output of the PLL is given as clock into the other components, configured at 100 MHz. The block diagram below describes the current system and the proposed system, respectively.

The Fig. 1 describes the current design for the implementation of the Ethernet bridge using a single NIOS II processor.

Figure 2 describes the proposed design for the implementation of the Ethernet bridge using a two separate NIOS II processors and two TSE MAC. We add an additional hardware i.e., [4] Mutex to share the programing memory.

All the components are connected to the NIOS II processor through an Avalon Memory-Mapped interface. Now we add an [2] SDRAM to the design and is connected to the NIOS II processor. An interrupt is connected from the processor to the JTAG/UART and the timer. To the SDRAM, both the data master and the instruction master of the NIOS II processor are connected. Now, the second processor is added along with other components mentioned before except for the SDRAM. The NIOS II processor is also connected the SDRAM. A mutex is added and connected to the data master of both the processors. A CFI Flash is added and interfaced with the first NIOS II processor. TSE MAC, SGDMA TX, SGDMA RX, and the descriptor memory are added to each of the processors. Two FIFOs are interfaced to each processor such that it reads from one processor and writes to the other.

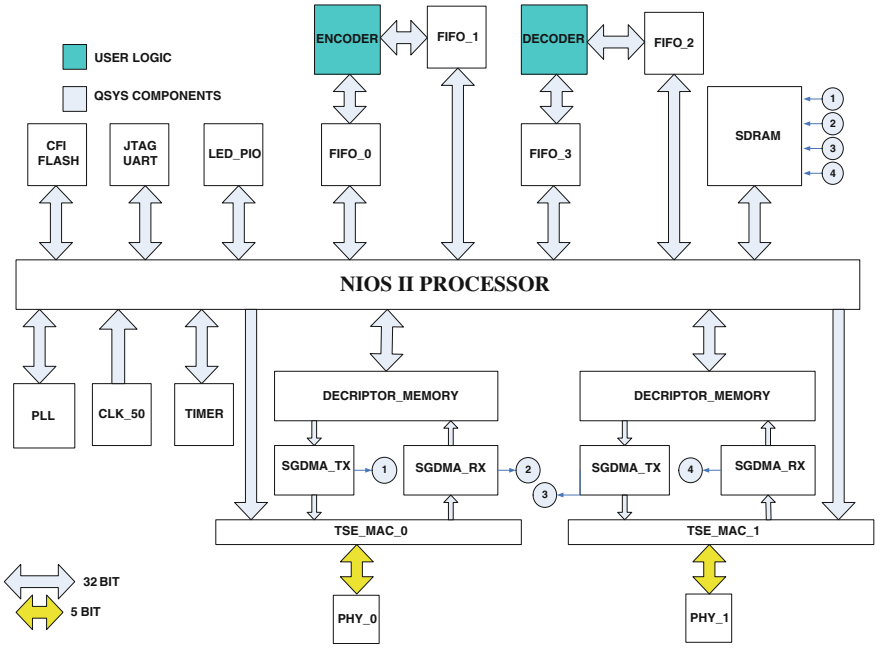


Fig. 1 Block diagram of the Ethernet Bridge in the current system

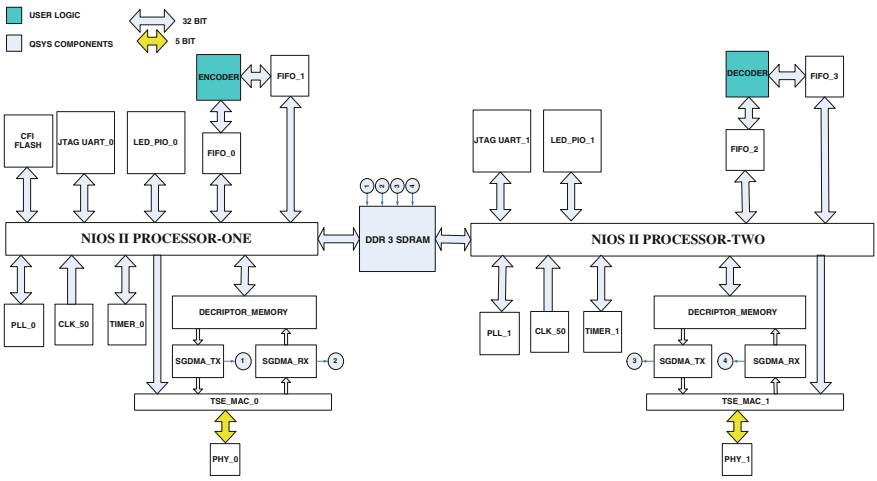


Fig. 2 Block diagram of the Ethernet Bridge in the proposed system

3.2 Software Implementation

In NIOS, we will compile both Ethernet bridge application and the [3] board specific package. Ethernet bridge application contains tasks that will initialize NicheStack and rest of the simple socket server task like led control for debugging. Board specific package consists of drivers for all the components in the system, [5, 6] NicheStack and microC/OS II. The main function is defined in iniche_init.c in Ethernet bridge application which initializes all tasks. Application calls the function netmain()m which prepares all interfaces. The netmain.c is defined in board specific package which consists of thread tk_netmain. Upon start-up, tk_netmain runs into endless loop which checks whether there is any packet arrived in rcvdq queue. If it finds queue length is not equals to zero, it calls function pktdemux(). Pktdemux first checks the interface type and protocol type. Based on this, it will call function arp_rcv() or ip_rcv(). The function ip_rcv() first compare whether packet length is less than header length. If it is less than header length, the packet is ignored. Then, it will copy the checksum to some variable and then calculate the checksum of received packet. If checksum are same, packet is accepted and then the packet is analyzed for IP version, protocol, etc. Then, the function ip_rcv_phase2() is called which will check whether the received packet is for itself or it should be forwarded. If the packet is unrecognized or TTL $= 0$, then the packet is discarded. If the destination IP address does not match its own, IP address packet needs to be forwarded. In such a case, it will check the routing table for routing and find the appropriate route to destination. When it is forwarding the packet, it checks whether packet is too big and fragmentation can be done or not and does the fragmentation. If fragmentation is needed and flag is set in packet

such that it does not allow fragmentation, then packet is discarded. After fragmentation, it will again calculate new checksum and forward the modified packet to destination.

4 Implementation and Verification

The Verilog code generated by QSYS is compiled on Quartus software and the pin assignments are made. This is compiled again and the produced SOF file is downloaded on to the DE2-115 board which has Cyclone IV E device using the Quartus programmer. The software also generates a SOPCinfo file which contains the details of the hardware, memory address, base addresses of each component. This file is used in NIOS IDE to generate the board specific package (BSP), and the bridging application code is attached to the BSP generated. The project is built and is run as hardware. The two Ethernet ports are configured to the IP address of 192.168.1.100 and 192.168.1.101. The default gateway is set at 192.168.1.1.

The board is now connected to two different PCs where the PC connected to first port has an IP Address of 192.168.1.5 and the second PC has the IP Address of 192.168.1.8. The default gate way being 192.168.1.1. Using the command prompt, the IP addresses of the port and the PC on the other end are pinged. After pinging the two ports and the PC on the other end successfully, an FTP server is run on one of the PC and FTP client is run on the other PC. A file of few Mbytes is attempted to transfer and is successfully transferred. Hence, the bridging using two NIOS II processors is implemented and verified successfully.

5 Result

When the above-explained setup is implemented on the board, file of various sizes was transferred from once PC to the other. It was found that the throughput had increased 1.6 times compared to the previous setup. While transferring a file of around 60 KB, transfer rate with the previous setup was nearly 360 MB/s. With the new setup and transferring the file of same size, the transfer rate achieved was around 620 MB/s. The transfer rates varied as the file size varied.

6 Conclusion

The implementation of the Ethernet bridge using two NIOS II processors is done successfully. The two TSE MACs are controlled by two different processors and hence the load on single processor has reduced. With this setup, the speed of

transfer has increased greatly and also the throughput. Overall performance of the system has improved and the efficiency is also higher. Ethernet bridge is prototyped and implemented on FPGA and also can be ported to ASIC.

References

1. Altera (2011) Cyclone IV E device handbook, V 1.5. Altera Corp
2. (2010) DE2-115 datasheet and user guide, V 1.2. Altera Corp
3. (2010) My first NIOS II software, V 2.0. Altera Corp
4. (2011) Creating multiprocessor NIOS II systems, V 2.0. Altera Corp
5. (2011) Using the NicheStack TCP/IP stack—NIOS II edition, V 3.0. Altera Corp
6. (2011) NIOS II software developer's handbook, V 11.0.0. Altera Corp. Ethernet and the NicheStack TCP/IP Stack—NIOS II edition, V 11.0.0, May 2011. Altera Corp

Performance Analysis of EBG Antenna for Base Station Applications

Krishnananda and T. S. Rukmini

Abstract The objective of this paper is to analyze the performance of electromagnetic band gap (EBG) antenna for base station applications through simulation. The proposed analysis is carried out using the high frequency structure simulator (HFSS). In our method, after comparing the results, we can say that the best obtained result is with respect to two layer FSS with cylinders having resonant cavity width of 10 mm. With this type of structure, we have obtained gain of 4.51 dB with bandwidth of 67.5 MHz, and voltage standing wave ratio (VSWR) is 1.38. Also, we simulated, and result is analyzed for several EBG structures at resonating frequency 2.4 GHz which is found suitable in mobile base station applications.

Keywords Electromagnetic band gap (EBG) · High frequency structure simulator (HFSS) · Frequency selective surfaces (FSS) · Microstrip · Surface waves · Resonating frequency

1 Introduction

Electromagnetic band gap (EBG) materials are metamaterials. They are artificial materials engineered to have properties that may not be found in nature. Metamaterials usually gain their properties from structure rather than composition, using small inhomogeneities to create effective macroscopic behavior. Metamaterial consists of periodic structures. EBG always referred to as photonic band gap (PBG) [1–3]. This structure is compact which has good potential to build low

Krishnananda (✉)
Electronics Department, Kuvempu University, Shivamogga, Karnataka, India
e-mail: konikrishna@yahoo.co.in

T. S. Rukmini
Electronics Department, N M I T, Bengaluru, Karnataka, India
e-mail: rukminirvce.rao@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_42,

profile and high efficiency antenna surface. The main advantage of EBG structure is their ability to suppress the surface wave current. The generation of surface waves decreases the antenna efficiency and degrades the antenna performance. Furthermore, it increases the mutual coupling of the antenna array. Two solutions to the surface wave problem are available now. One of the approaches is based on the micromachining technology in which part of the substrate beneath the radiating element is removed to realize a low efficiency dielectric constant environment for the antenna. In this case, the power loss through surface wave excitation is reduced and coupling of power to the space wave enhanced. The second technique relies on EBG Engineering. The feature of surface wave suppression helps improve antenna's performance such as increasing the antenna gain and less power wasted when radiating in backward direction [4]. Microstrip patch antennas have been an attractive choice in mobile and radio wireless communication.

The performance and operation of microstrip antenna is driven mainly by the geometry of the printed patch, and the material characteristics of the substrate onto which the antenna is printed. The microstrip antenna offers low profile, conformable to planar and non-planar surfaces, simple and inexpensive to fabricate using modern printed-circuit technology, mechanically robust when mounted on rigid surfaces and very versatile in terms of resonant frequency, polarization, patterns and impedance [5]. Major disadvantages of microstrip antenna are their low efficiency, low power, very narrow frequency bandwidth and existence of surface waves. There are numerous substrates that can be used to design microstrip antennas. The substrates usually have dielectric constants in the range of $2.2 \leq \varepsilon_r \leq 12$ for operation at frequencies ranging from 1 to 100 GHz [6]. Two types of EBG structure to be tried here: first, without slotted implementation of EBG layer and second, to analyze the effect, by drilling patches on the EBG layer [3]. EBG structure can be design by various shapes, and every shape will have different frequency band gap.

The new version 11 of high frequency structure simulator (HFSS) provides better accuracy and performance than previous versions according to Ansoft. Some of the new features in HFSS version 11 are: It is able to simulate larger structures (multiple wavelengths) in a more efficient manner, revised port solver can simulate lower frequencies, and it has a lower noise floor and genetic algorithm optimization as part of optometric. HFSS is the industry standard software for *S*-parameter and full-wave extraction and for the electromagnetic simulation of high-frequency and high-speed components. HFSS is widely used for the design of on-chip embedded passives, PCB interconnects, antennas, RF or microwave components and high-frequency IC packages.

2 Design of Rectangular Patch

Design of rectangular patch involves the calculation of the width and length of the patch when resonant frequency f_0, dielectric constant of the substrate and height of the substrate are given [2]. The design of the rectangular patch is given below.

First step of design is to find width of the patch, W. For efficient radiation, the width W is given by (1).

$$W = \frac{C}{2f_0\sqrt{\frac{(\varepsilon_r+1)}{2}}} \tag{1}$$

where c = velocity of light, f_0 = resonant frequency, ε_r = dielectric constant of the substrate. Taking $C = 3 \times 10^8$ m/s, $f_0 = 2.4$ GHz and ε_r = 4.4 for FR4, we got width of the patch as $W = 38$ mm. The dimensions of the patch along its length have now been extended on each end by a distance ΔL, and effective dielectric constant is given by (2).

$$\Delta L = 0.412\,h\frac{(\varepsilon_{\text{reff}} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{\text{reff}} - 0.258)\left(\frac{W}{h} + 0.8\right)} \tag{2}$$

where W is width of the patch, h is height of the substrate and $\varepsilon_{\text{reff}}$ is effective dielectric constant and is taken as,

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{w}\right]^{\frac{1}{2}} \tag{3}$$

Taking $\varepsilon_r = 4.4$, $h = 1.57$ mm, $W = 38$ mm, we got $\varepsilon_{\text{reff}}$ as 4.09 and hence $\Delta L = 0.725$ mm. Length of the patch, L, is taken as $L = [\{\lambda/(2\sqrt{(\varepsilon_{\text{reff}})}\} - 2\,\Delta L]$. Where λ is the wavelength and was taken as $\lambda = 125$ mm for our design. After placing the values of λ, $\varepsilon_{\text{reff}}$, and ΔL, in equation for L, the value of length of the patch resulted as 29.45 mm.

The effective length of the patch, L_{eff}, was taken as $L_{\text{eff}} = L + 2\ \Delta\text{L}$. After putting the values of L and ΔL, the effective length of the patch was $L_{\text{eff}} = 30.9$ mm. The patch should maintain $\lambda/2$ distance on the substrate, that is, 62.5 mm. The design of EBG is carried out by using (a) cylindrical patches on FSS layer, (b) cuboidal patches on FSS layer with and without hole, and (c) conical patches on FSS layer. For the design of cylindrical patches on FSS layer, we took radius = 11.032 mm, height = 0.1 mm, distance between the patches along length of the FSS layer = 22.064 mm, distance between the patches along width of the FSS layer = 23.285 mm, height of EBG substrate is taken as 5 mm and the distance between the EBG substrate and microstrip patch substrate = 5 mm. For cuboidal patches on FSS layer without hole, we took length = 22.064 mm, width = 23.285 mm, height = 0.1 mm, distance between the patches along length of the FSS layer = 22.064 mm, distance between the patches along width of the FSS layer = 23.285 mm, height of EBG substrate = 5 mm and distance between the EBG substrate and microstrip patch substrate = 5 mm. For cuboidal patches on FSS layer with hole, we took hole radius 1 = 2.758 mm, hole radius 2 = 5.516 mm. And for conical patches on FSS layer design, we had radius = 11.032 mm, height 1 = 0.1 mm, height 2 = 1 mm, height 3 = 2 mm, height 4 = 4 mm, distance between the patches along length of the FSS

layer = 22.064 mm, distance between the patches along width of the FSS layer = 23.285 mm, height of EBG substrate = 5 mm and distance between the EBG substrate and microstrip patch substrate = 5 mm. In HFSS software tool, the above data are uploaded and run for the simulation result. Layer 1, 2 and 3 for all the three kinds of EBG patches can be used. Figures 1, 2, and 3 show the three kinds of EBG design.

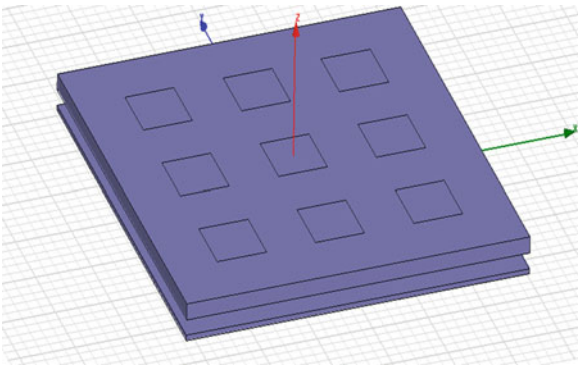


Fig. 1 Designed one layer EBG with FSS consisting of *cuboids*

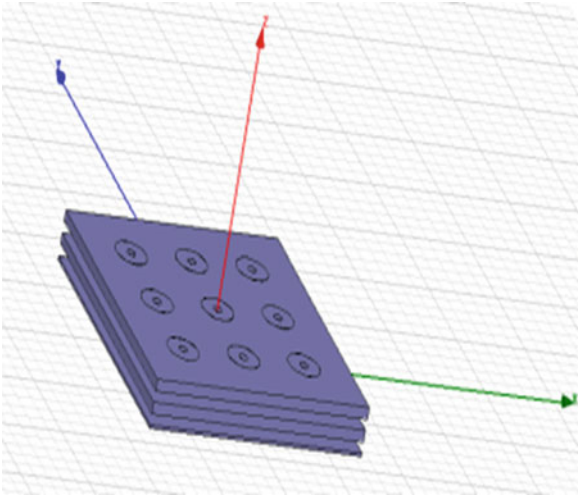


Fig. 2 Designed two layer EBG with FSS consisting of *slotted cylinders*

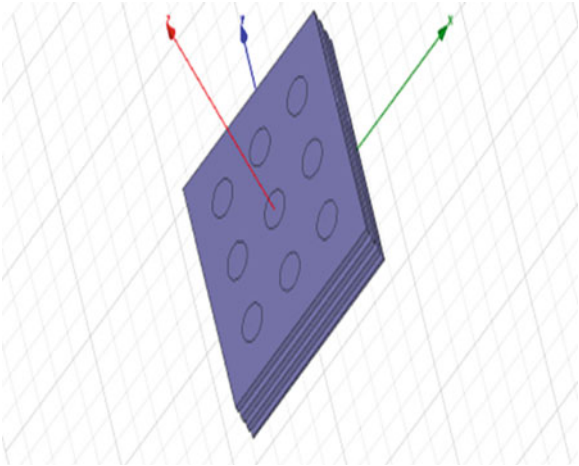


Fig. 3 Three layer EBG with FSS consisting of cones with height 0.1 mm

3 Result

After simulation, we got some results at resonating frequency 2.4 GHz which is designed for mobile base station applications. The simulation results are represented in Table 1. Types of EBG structures listed are rectangular patch without EBG, one layer FSS with cylinders, one layer FSS with cuboids, one layer FSS with cones (height 0.1 mm) and one layer FSS with cones (height 1 mm). By changing the height of the cones from 0.1 to 1 mm, it is found that there is not much change in the values of return loss, gain and voltage standing wave ratio (VSWR). The other structures are one layer FSS with cones (height 2 mm), one layer FSS with cones (height 3 mm), one layer FSS with slotted (2.758 mm hole) cylinders, one layer FSS with slotted (2.758 mm hole) cuboids, one layer FSS with slotted (5.516 mm hole) cuboids, two layer FSS with cylinders (resonant cavity Height 2.5 mm), two layer FSS with cylinders (resonant cavity height is 10 mm), two layer FSS with cuboids (resonant cavity with height 5 mm), three layer FSS with cylinders(resonant cavity height 5 mm) and three layer FSS with cones having height of 0.1 mm. Few simulation results are shown in Figs. 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15.

Hence, it is realized that antenna without EBG provides gain of 2.83 dB and bandwidth of 60 MHz with VSWR of 1.1. After putting EBG layer over rectangular patch, we have obtained improved results depending upon EBG layer implemented. After comparing the results, we can say that the best obtained result is with respect to two layer FSS with cylinders having resonant cavity width of 10 mm. The resulted gain is 4.51 dB at the bandwidth of 67.5 MHz, and VSWR is 1.38.

Table 1 Simulation result

Sl. No.	Types of EBG structure	Lower frequency (in GHz)	Upper frequency (in GHz)	Band width (in MHz)	Return loss (in dB)	Gain (in dB)	VSWR
1	Without EBG (rectangular patch shape)	2.2975	2.3575	60	−28.6	2.83	1.1
2	One layer FSS with cylinders	2.2675	2.32	52.5	−12.93	5.22	1.6
3	One layer FSS with cuboids	2.2675	2.2975	30	−10.317	4.75	1.84
4	One layer FSS with cones (height 0.1 mm)	2.2675	2.32	52.5	−12.98	5.30	1.6
5	One layer FSS with cones (height 1 mm)	2.2675	2.32	52.5	−12.94	5.20	1.57
6	One layer FSS with cones (height 2 mm)	2.2675	2.32	52.5	−12.88	5.20	1.64
7	One layer FSS with cones (height 3 mm)	2.2675	2.32	52.5	−12.88	5.13	1.6
8	One layer FSS with slotted cylinders (2.758 mm hole)	2.2675	2.32	52.5	−12.99	5.24	1.6
9	One layer FSS with slotted cuboids (2.758 mm hole)	2.275	2.30	25	−10.39	5.03	1.84
10	One layer FSS with slotted cuboids (5.516 mm hole)	2.27	2.30	25	−10.21	4.89	1.9
11	Two layer FSS with cylinders (resonant cavity height 2.5 mm)	2.2525	2.3125	60	−13.48	4.98	1.54
12	Two layer FSS with cylinders (resonant cavity height 10 mm)	2.3125	2.38	67.5	−16.34	4.51	1.38
13	Two layer FSS with cuboids (resonant cavity height 5 mm)	2.3350	2.3725	37.5	−10.66	3.77	1.825
14	Three layer FSS with cylinders (resonant cavity height 5 mm)	2.350	2.3650	60	−14.1994	2.294	1.48
15	Three layer FSS with cones (height 0.1 mm)	2.3050	2.3650	60	−14.15	2.97	1.6

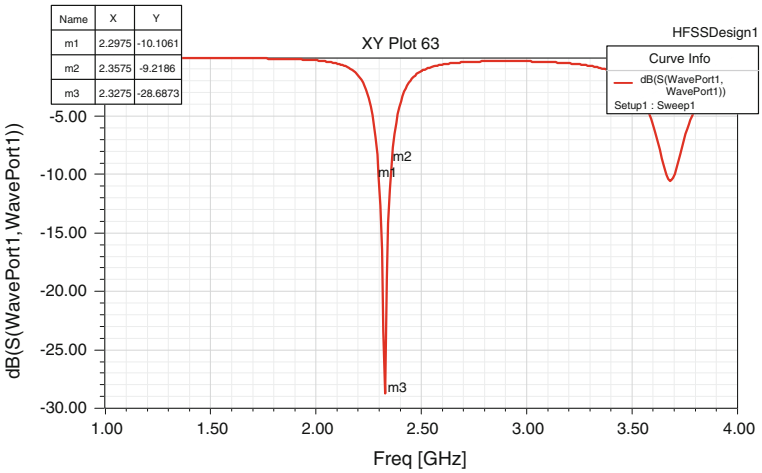


Fig. 4 Frequency versus return loss plot for a rectangular patch without EBG

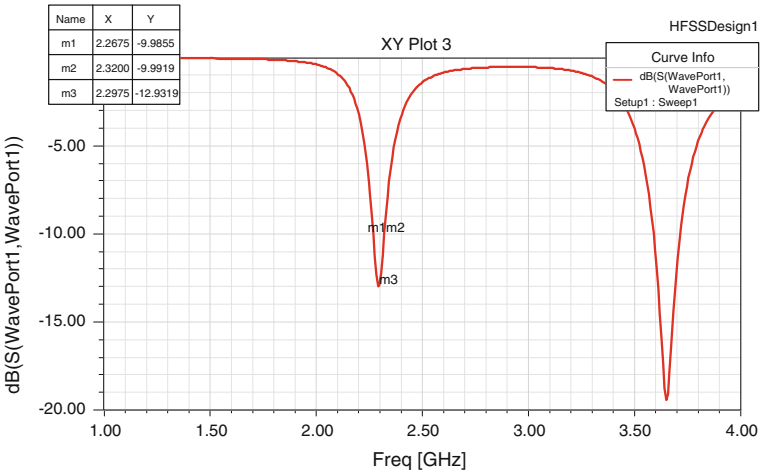


Fig. 5 Frequency versus return loss plot for a one layer EBG with FSS consisting of *cuboids*

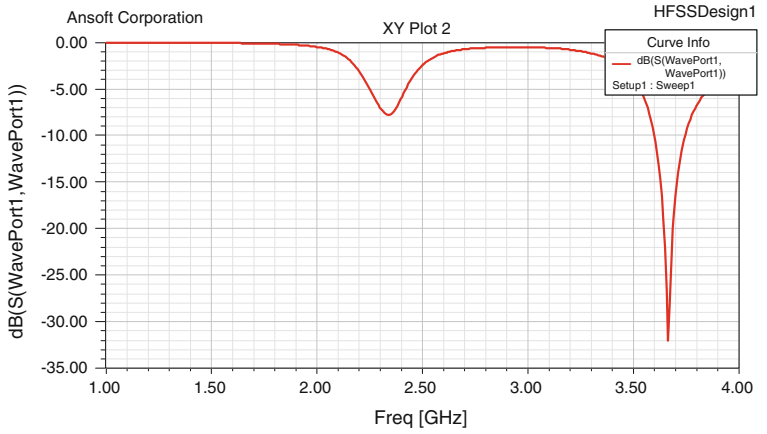


Fig. 6 Frequency versus return loss plot for two layer EBG with FSS consisting of *slotted cylinders*

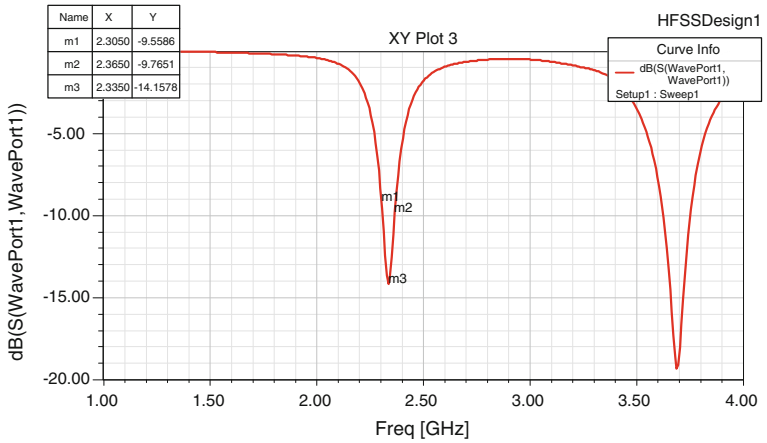


Fig. 7 Frequency versus return loss plot for three layer EBG with FSS consisting of *cones* with height 0.1 mm

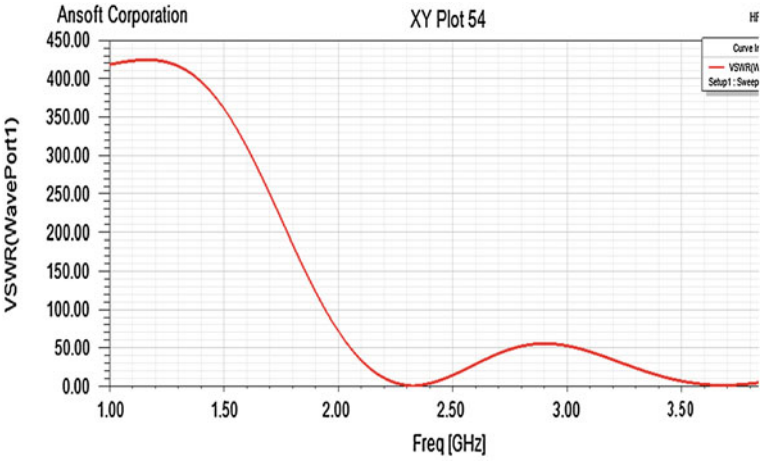


Fig. 8 Frequency versus VSWR for a rectangular patch without EBG

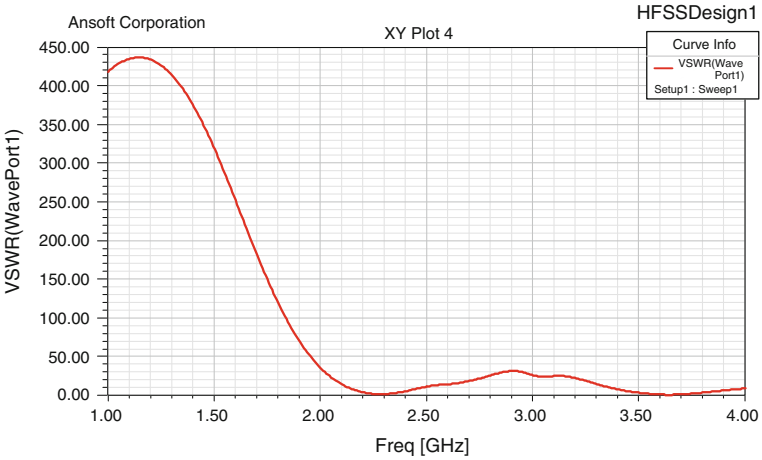


Fig. 9 Frequency versus VSWRplot for a one layer EBG with FSS consisting of *cuboids*

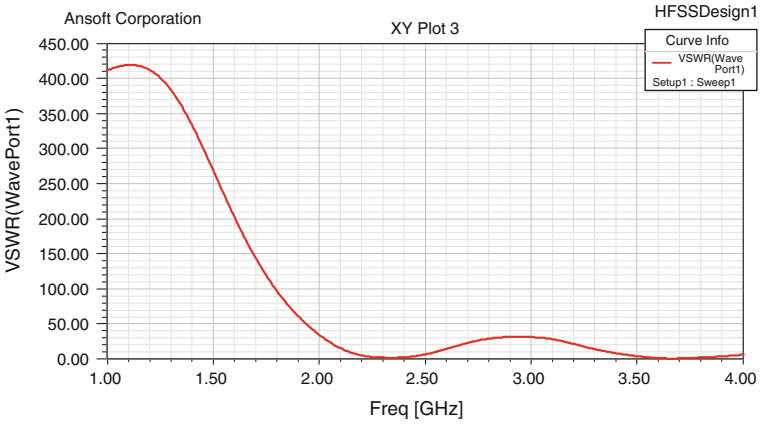


Fig. 10 Frequency versus VSWR plot for two layer EBG with FSS consisting of *slotted cylinders*

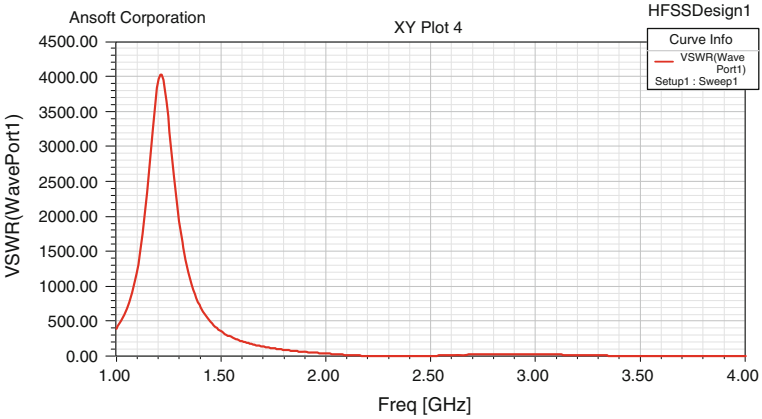


Fig. 11 Frequency versus VSWR plot for three layer EBG with FSS consisting of *cones* with height 0.1 mm

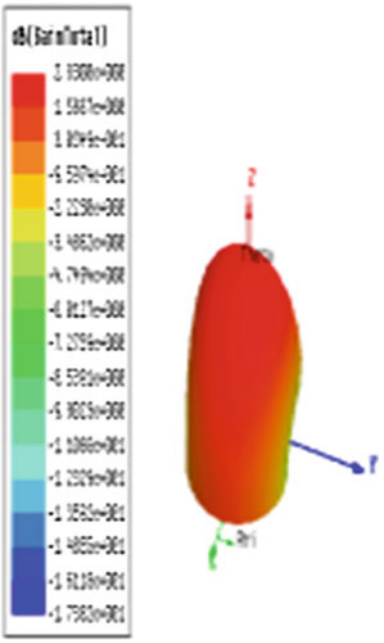


Fig. 12 Gain plot for a rectangular patch without EBG

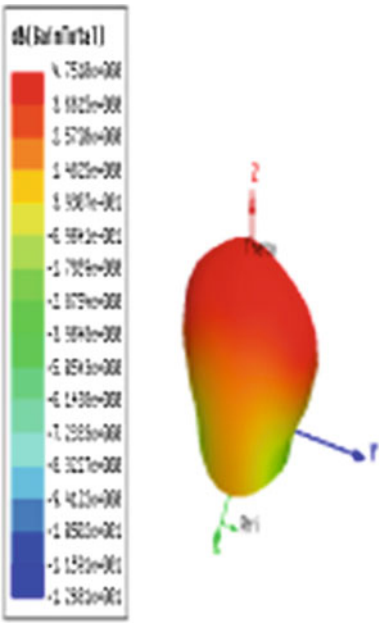


Fig. 13 Gain plot for a one layer EBG with FSS consisting of *cuboids*

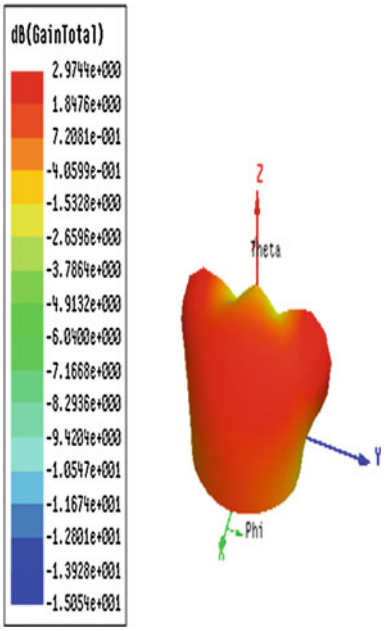


Fig. 14 Gain plot for a two layer EBG with FSS consisting of *slotted cylinder*

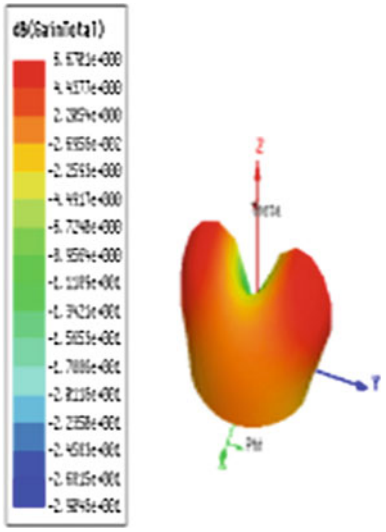


Fig. 15 Gain plot for three layer with FSS consisting of *cones* with height 0.1 mm

4 Conclusion

The performance analysis of EBG antenna at a frequency of 2.4 GHz which is suitable in mobile communications has been presented. The utilization of the HFSS software involved in the process helps minimize the processing time for the calculation and the simulation of the design. Many types of EBG structures with different number of layer have been implemented. Integration of the EBG structure with the microstrip array antenna improved the bandwidth and the gain. As an overall conclusion, all the planned works and the objectives of this paper have been successfully implemented, and the improvement of the antenna performances with bandwidth and gain is analyzed after designing different EBG structures.

5 Future Work

Further works can be taken up in order to optimize the design of the EBG structure. Some suggestions for future work are (a) the multi band EBG structure may be design to cover the multiband frequency range, (b) miniaturization of the EBG structure to achieve the compactness of the antenna, and (c) other applications of EBG especially with filter design may be done.

References

1. Pozar DM (1992) Microstrip antennas. IEEE Proc 80:79–81
2. Derneryd AG (1978) A theoretical investigation of the rectangular microstrip antenna element. IEEE Trans Antennas Propag AP-26(4):532–535
3. Rahim MKA (2008) EBG design for microwave devices. Research votes no: 79017
4. Ge Y, Esselle KP, Hao Y (2007) Design of low-profile high gain EBG resonator antennas using genetic algorithm. IEEE Antennas Wirel Propag Lett 6
5. Lee Y, Park W, Yeo J, Mittra R (2006) Directivity enhancement of printed antennas using a class of metamaterial superstrates. Electromagnetics 26:203–218
6. Diblanc M, Rodes E, Arnaud E, Thevenot M,Monediere T, Jecko B (2005) Circularly polarized metallic EBG antenna. IEEE Microwave Wirel Compon Lett 1–3

Web Images Evaluations Based on Visual Content

K. K. Umesh and Suresha

Abstract A Web user can find images by visiting Web sites or by querying Web search engines. We believe that the images obtained from a few search engines are inadequate. Today, images in Web are represented by using keywords, or summary of Web page is used to express the meaning of a particular image. The use of search engines will provide a set of results that correspond to the topic, and a user has to browse through every result to find relevant image about the topic which is time consuming and laborious. Hence, we have collected Web images and evaluated the collected images for constructing Web image database using visual content of images.

Keywords Image collection · Hierarchical clustering · Image selection · Precision

1 Introduction

Due to rapid development of Web and its related technologies, Web has become popular with all sections of the society. The Web is used by different users to make them available globally. It is the source of many research activities. An interesting area is to retrieve images from Web on diverse topics.

Text-based information retrieval system like Google [1], AltaVista [2], etc., is deal only with textual information. Content-based image retrieval system like

K. K. Umesh (✉)
Department of Information Science and Engineering, S J College of Engineering, Mysore, India
e-mail: Umeshkatte@gmail.com

Suresha
Department of Studies in Computer Science, Manasagangothri, University of Mysore, Mysore, India
e-mail: sureshabm@yahoo.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_43,

QBIC [3], VisualSeek [4], etc., is focus on visual content of an image without considering textual content.

There have been many studies on the search engines to measure the overlap and ranking difference of leading search engines [5]. Generally, Web users think that all search engines are similar in functions, provide similar results, and index all contents available in the Web. The results obtained based on studies clearly contradict users thinking [6]. The result of these studies indicates that the search engines are different from searching and retrieving procedures, and they provide different search results.

In this paper, we focus on collecting images from different search engines and provide the comparison results of manual evaluation and system evaluation. In Sect. 2, we discuss the overview of proposed methodology. The experimental results are discussed in Sect. 3. Conclusion is discussed in Sect. 4.

2 Methodology

We collected nearly 5,000 color images of 10 classes from different commercial search engine such as Google [1], Yahoo, Bing, Exalead, and Ask image search on different images like Apple, Kingfisher, Cow, Beach, Elephant, Crane, Fox, Cock, Peacock, and Yak as part of our dataset. The Web images collected were of different sizes. The Web images were grouped into a set of classes. We extracted images from search engines using keywords. The strength of keyword relevancy is the absolute value of the extracted top 100 images determined using Eq. 1.

$$\text{Precision} = N_{RI}/(N_{RI} + N_{NR}) \tag{1}$$

where N_{RI} and N_{NR} are the number of relevant images and the number of irrelevant images to their keywords, respectively.

In a few Web images, we found that images are irrelevant based on human observation, and such images are manually grouped into irrelevant images. We also found some of the images contained only cartoons, freehand diagrams from the Google search engine, it is difficult to judge whether these images are of relevant or irrelevant; such kind of images grouped into irrelevant, since we focus only on *.JPEG/*.JPG (joint photographic experts group) for collecting Web images.

2.1 Image Collection

We have built two classes of images, namely relevant class and irrelevant class. We collected images from 5 search engines for 10 categories of images like Apple, Kingfisher, etc. In every category, around 100 images are collected and categorized into relevant class and irrelevant class. The total number of images in both

the classes is around 5,000. The average precision of collected images is 0.58 %. The measure of relevancy is very difficult to measure manually in certain cases. We observed that some relevant images are misclassified into the irrelevant image class and similarly irrelevant images are classified into relevant class, since human intervention for classifying images from large number of images is laborious and time consuming.

In the next section, we have discussed about image selection based on visual content of images to minimize the human interaction to classify Web images as relevant class or irrelevant class.

2.2 Image Selection

Users' perception varies from person to person, and human evaluation is not sufficient to categorize images into relevant and irrelevant classes [7]. Hence, we have analyzed the content of images from relevant set and cluster similar images into a few groups. The selection of images from the cluster is based on the image features. The algorithm is as follows:

1. For each of the collected images, we extract a color histogram as image features.
2. Convert RGB color space into HSV color space, quantize the HSV color space into 256 bins of each color space, and make a color distribution histogram represented in a 768-dimensional vector of each image [8].
3. Compute the distance between pair of images in group A, and the distance represents the degree of dissimilarity between the two images that is calculated based on their image features. The distance between ith and jth images is computed by the Eq. 2.

$$d_{ij} = \left(h_i - h_j\right)^t A \left(h_i - h_j\right) \tag{2}$$

In condition that

$$A = \left[a_{ij}\right] \tag{3}$$

$$a_{ij} = \begin{cases} 1, & (i = j) \\ e^{-cd_{ij}/d_{\max}}, & (i \neq j) \end{cases} \tag{4}$$

$$d_{\max} = \max_{ij}(d_{ij}) \tag{5}$$

where h_i and h_j represent ith and jth image's image feature vector, respectively, c is a constant, and d_{ij} is the Euclid distance between ith and jth bin of each color histogram.

4. Based on the distance between images, images in group A are clustered by the cluster analysis method. We adopt a hierarchical clustering method out of the existing clustering methods, because a hierarchical clustering method requires the minimum distance as the predetermined constant and does not require the number of clusters in advance [9]. The system repeats merging clusters, until all distances between clusters are more than the threshold. For using this method, we have to define an inter-cluster distance. Here, we use the farthest neighbor method (FN) as one of hierarchical clustering methods. In the FN clustering method, we define the distance between clusters as the largest distance between two images belonging to mutually different clusters. Images that have many similar images make bigger clusters, while unique images make smaller clusters.
5. Here, we observe that in many cases images in the larger clusters as relevant ones and images in the smaller clusters as irrelevant ones. Then, the system throws away small clusters which have fewer images than a certain threshold value.

In the experiments, we set 5 % of the total number of images in group A as the threshold. It stores all images in the remaining clusters as output images. In addition, we compute an average feature vector of each cluster for all the remaining clusters.

6. To select images from group B, distance is computed between a feature vector of group B image and an average vector of each remaining cluster of group A. If an image in group B has less distance than a pre-determined threshold from an average feature of remaining clusters of group A, it is selected and considered as a relevant image. By the processing described above, finally we obtained relevant images related to the given keywords as the final results of the image collection are used further in our experiments.

3 Experiments and Results

3.1 Evaluation of Gathered Images

In the collection of Web images, the relevant and irrelevant images are grouped by the subjective evaluation. For example, a keyword ‘Apple’ was input to the Google image search engine, and from the result returned by the search engine, we have selected the top 100 images out of thousands of images to evaluate the precision of keyword search with the search engine. As per the literature, most of the search engine users are interested to see the results of starting pages as they are highly relevant to the users’ request. For example, Google has its own search engine

ranking procedures (SERPs) for finding images based on keyword [10]. In addition, we employed a few search engines to collect images related to Apple, Kingfisher, Cow, Beach, Elephant, Crane, Fox, Cock, Peacock, and Yak. We evaluated the fetched images based on class. We found that 59 images were relevant image (class A) and 41 were irrelevant images (class B) for an Apple class. To evaluate the precision of relevancy, we used Eq. 1.

In case of 'Apple', 190 images were decided to be class A out of 500 images and 310 images in class B, respectively. The total number of images of class A and class B were 500. This proved that image selection by keyword evaluation across the search engine employed in these systems worked effectively. Since the search engines are used their own SERPs [10], there is no common procedure to rank the images; hence, the obtained results varies from one search engine to another. During the collection of Web images, we have nearly collected 5000 images in five search engines for experimentation. The class A contains 2916 images, and class B contains 2,084 images.

3.2 Evaluation of Selection of Images

In the selection part, we selected images from group A using image features. Image features are cluster into groups. Images are selected from group B by computing the distance between irrelevant images and group A images. For example, in case of class 'Apple', 12 clusters were made by the image-feature-based clustering. Since the number of images in group A was 190, out of 190 we have selected 178 images from 12 clusters.

During the experiments, to select the number of relevant cluster, we have analyzed the cluster using dendrogram to know the distribution of images in hierarchical cluster. The most difficult part of cluster analysis is fixing the threshold due to heterogeneous nature of Web images. Hence, we consider a threshold of 5 % of total number of images to get the satisfactory results. For example, 32 'Apple' images were selected from 310 irrelevant 'Apple' images. Finally, the total number of output images for 'Apple' was 210, and their precision is 0.42 %.

In each experiment, the number of final output images is the total number of collected images. The precision of selection is almost equal to the human evaluation. In some of the cases, the precision of selection is improved compared to the precision of collected images; especially, where the images include the objects like 'Cow' and 'Cow Boy', it is difficult to judge and distinguish whether relevant or irrelevant. The color is very essential to distinguish their objects from the images. The precision of manual evaluation is almost equal to our method, which implies that our method is effective for image gathering from WWW (Table 1).

Table 1 Number of images selected from collection of images

Class	Images in group A		Images in group B		Total (A+B) Selected	Precision	
	Collected	Selected	Collected	Selected		Collected (A)	Selected
Apple	190	178	310	32	210	0.38	0.42
Kingfisher	347	327	153	22	349	0.69	0.70
Cow	372	356	128	34	390	0.74	0.78
Beach	391	388	109	04	392	0.78	0.78
Elephant	427	422	73	10	432	0.85	0.86
Crane	157	147	34	24	171	0.31	0.34
Fox	245	232	25	09	241	0.49	0.48
Cock	93	76	407	23	99	0.19	0.19
Peacock	368	310	132	28	338	0.74	0.68
Yak	326	297	174	49	346	0.65	0.69
Total/Average	2,916	2,733	1,545	235	2,968	58.2	59.2

3.3 Discussion on the Results of the Experiments

The precision of collected images was more than 58 % for most of keywords. This indicates that the effectiveness of keyword is achieved 58 % for 10 classes of image. In some of the cases, it achieves more than 85 % and below 25 % in few cases. This shows that most of the commercial search engine was analyzed and ranked images based on keywords. The keywords decide the rank of images. This implies that many search engines ignore the visual content of images. Since file names of most of images include 'Elephant' category, but in the case of 'Crane' class, the crane keywords appear in Web pages which contain the information about tools and machines. So that it is less relevant to the 'Bird'.

For total images which were the sum of group A and group B images, the precisions were from 67 to 82 %. The precisions of selected images for all the keywords were superior to the precisions of collected images. Particularly, in the case of images which have salient colors such as a yellow 'cow' with brown grass, white 'Mountain' with blue sky, and a red 'apple,' adoption color histogram as a image feature worked effectively and improved precisions. On the other hand, for certain cases like cock class, various color clothes and variable color face due to illumination change, the ratio of improvement was relatively small, indicating that using only using color histogram as a image feature is sufficient.

4 Conclusion

The existing Web image search engines are designed to look for an image based on keywords. A few images obtained from search engines are not relevant. The use of a single popular search engine provides an ambiguous view of the document and does not represent the entire Web community. The relevancies of search engines

differ from one another in terms of their design and implementation to search and retrieve similar images.

In this paper, we have conducted experiments to evaluate collected images from search engines. The collected images are evaluated manually and also evaluated using visual features with automated evaluation system.

To automate selection of images from the collection of images, we use only a color histogram as an image feature. The results of manual evaluation and the automated evaluation have been discussed. The precision of group A (relevant) images from automated evaluation system is increased 1.00 % than the precision of manual evaluation system. We observed and verified that the group A and group B images are classified properly in the corresponding groups. Based on the results, our proposed approach has obtained better results with image features than the manual evaluation.

References

1. http://www.google.co.in
2. http://www.altavista.com/
3. Flicker M, Sawhney H, Niblack W, Ashley J, Huang Q, Dom B et al (1995) Query by image and video content: the system. IEEE Comput 28(9):23–32
4. Yani K (2003) Image collector II: a system for generic more than one thousand images from the Web for one keyword. In: Proceedings of IEEE international conference on multimedia and expo, vol 1, pp 785–788
5. Faulring WA, Rosenholtz R, Morrison J, Pirolli P (2001) Using thumbnails to search the Web. In: Proceeding of the SIGCHI conference on human factors in computing systems, pp 198–205
6. Customer Experience Rankings for North America Search Web Sites: Survey: Google, Yahoo Still Favorites in North America study by Keynote.com
7. Candan KS, Figueiredo MAT, Jain AK, Zhang HJ (2001) Image classification for content-based indexing. IEEE Trans Image Process 10(1):117–130
8. Foley JD, van Dam A, Feiner SK, Hughes JF (1990) Computer graphics: principles and practice, 2nd edn. Addison-Wesley, Reading, MA
9. Krishnamachari S, Abdel-Mottaleb M (1999) Hierarchical clustering algorithm for fast image retrieval. In: Proceedings of SPIE conference on storage and retrieval for image and video, pp 427–435
10. http://www.convertingcopy.com/google-image-ranking-how-to-get-an-image-to-rank-in-google/

Robust Speech Recognition Using Wavelet Domain Front End and Hidden Markov Models

Rajeswari, N. N. S. S. R. K. Prasad and V. Sathyanarayana

Abstract This paper presents a method to address the issue of noise robustness using wavelet domain in the front end of an automatic speech recognition (ASR) system, which combines speech enhancement and the feature extraction. The proposed method includes a time-adapted hybrid wavelet domain speech enhancement using Teager energy operators (TEO) and dynamic perceptual wavelet packet (PWP) features applied to a hidden Markov model (HMM)–based classifier. The experiments are performed using the HTK toolkit for speaker-independent database which are trained in a clean environment and later tested in the presence of AWGN. It has been seen from the experimental results that the proposed method has a better recognition rate than the most popular MFCC-based feature vectors and HMM-based ASR in noisy environment.

Keywords Automatic speech recognition · Perceptual wavelet packet · Hidden Markov model · Mel-frequency cepstral coefficients · Teager energy operator · Additive white gaussian noise

Rajeswari (✉)
Department of Electronics and Communication Enginering, Acharya Institute of Technology, Bangalore, India
e-mail: rajeswari@acharya.ac.in

N. N. S. S. R. K.Prasad
ADA, Ministry of Defence, Goverment of India, Bangalore, India
e-mail: nnssrkprasad2007@gmail.com

V. Sathyanarayana
Department of DSP, CET, Jain University, Bangalore, India
e-mail: satyaec49@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_44,

1 Introduction

An ASR system finds large applications demanding for real-time environments which are embedded with high ambient noise levels. The performance of an ASR is acceptable in clean environments; however, the system performance degrades in the presence of noise. Thus, there is a strong need for noise robustness to be considered [1–3].

Much of the research in speech recognition aims at first robust feature extraction and the second being building a robust classifier [4]. The most popular MFCC features based on short-time Fourier transform (STFT) and power spectrum estimation do not give a good representation of noisy speech. The features based on the STFT produce uniform resolution over the time–frequency plane. Due to this, it is difficult to detect sudden bursts in a slowly varying signal or the highly non-stationary parts of the speech signal. The recent approach of wavelet packets which segment the frequency axis and makes uniform translation in time is been proposed. Wavelet coefficients provide flexible and efficient manipulation of a speech signal localized in the time–frequency plane which is an alternative to MFCC [5–7]. The perceptual wavelet filter bank is built to approximate the critical band responses of the human ear. Wavelet packets decompose the data evenly into all bins, but PWPs decompose only critical bins [8].

In this paper, we propose a wavelet domain front end for an ASR which combines speech enhancement and the feature extraction. The proposed method includes a time-adapted wavelet domain hybrid speech enhancement using Teager energy operators (TEO) and dynamic perceptual wavelet packet (PWP) features applied to a hidden Markov model (HMM)–based classifier.

The rest of the paper is organized as follows. Section 2 introduces a block diagram of the proposed wavelet domain front end of the ASR and provides detailed description of each constituting part. Section 3 describes the recognizer and the toolkit used. Section 4 evaluates the performance of the proposed system under different levels of noise. The conclusion is presented in Sect. 5.

2 Proposed Wavelet Domain Front End of ASR

Figure 1 describes the proposed noise robust wavelet domain front end of an ASR. The noisy speech input wave files are sampled at 16 kHz and segmented into frames of 24-ms duration with frame shift interval of 10-ms overlap.

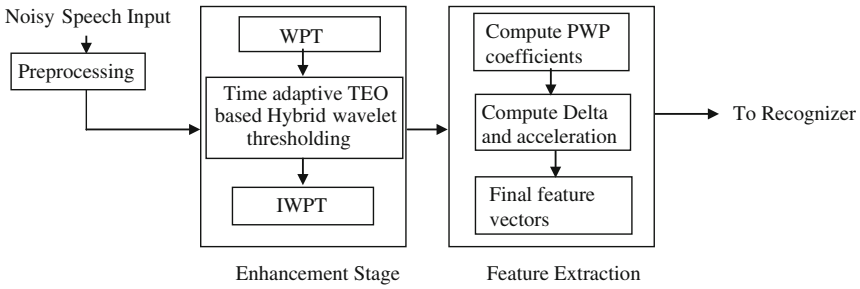


Fig. 1 Block diagram of proposed wavelet domain front end of ASR

2.1 Speech Enhancement

In the proposed method, wavelet packet transform (WPT) is applied to each input frame. The coefficients obtained are then subjected to Teager energy approximation [9, 10], where the threshold is adapted with respect to the voiced/unvoiced segments of the speech data. A hybrid thresholding process is adopted which is a compromise for the conventional hard and soft thresholding in preserving both the edges and reducing the noise.

2.1.1 Wavelet Packet analysis

For a j-level WP transform, the noisy speech signal $y[n]$ with frame length N is decomposed into 2^j sub-bands. The mth WP coefficient of the kth sub-band is expressed as follows:

$$W_{k,m}^{J} = \mathrm{WPT}\{y(n), j\} \tag{1}$$

where $n = 1,\ldots N$, $m = 1,\ldots, N/2^j$ and $k = 1,\ldots 2^j$.

2.1.2 Teager Energy Operator on Wavelet Coefficients

– Teager energy approximation for each WPT sub-band is computed

$$\mathrm{TEO_{i,k}} = \mathrm{Y_{i,k}^2} - \mathrm{Y_{i,k-1}Y_{i,k+1}} \tag{2}$$

– TEO coefficients are smoothened in order to reduce the sensitivity to noise

$$M_{i,k} = \mathrm{TEO}_{i,k} * H_p \tag{3}$$

– Normalize the TEO coefficients

$$M'_{i,k} = \left[\frac{M_{i,k}}{\max\left(M_{i,k}\right)} \right] \tag{4}$$

– Time-scale adaptive threshold based on Bayes shrink for each sub-band k is computed.

$$\lambda_{i,k} = \lambda_i (1 - M'_{i,k}) \tag{5}$$

2.1.3 Denoising by Thresholding

Denoising using wavelet packet coefficients is performed by thresholding; that is, the coefficients which fall below the specific value are shrunk and the later retained. Different thresholding techniques have been proposed. However, there are two popular thresholding functions used in the speech enhancement systems which are the hard and the soft thresholding functions [11–13].

Hard thresholding is given by

$$T_s(\lambda, w_k) = \begin{cases} w_k & \text{if } |w_k| > \lambda \\ 0 & \text{if } |w_k| \leq \lambda \end{cases} \tag{6}$$

Soft thresholding is given by

$$T_s(\lambda, w_k) = \begin{cases} \text{sgn}(w_k)|w_k| - \lambda & \text{if } |w_k| > \lambda \\ 0 & \text{if } |w_k| \leq \lambda \end{cases} \tag{7}$$

where w_k represents wavelet coefficients and λ the threshold value.

However, hard thresholding is best in preserving edges but worst in denoising, while soft thresholding is best in reducing noise but worst in preserving edges. In order to have a general case of both reducing noise and preserving edges, a hybrid thresholding is used.

Hybrid thresholding is given as follows:

$$T_s(\lambda, w_k) = \begin{cases} w_k * \frac{|w_k|^\alpha - \lambda^\alpha}{|w_k|^\alpha} & \text{if } |w_k| > \lambda \\ 0 & \text{if } |w_k| \leq \lambda \end{cases} \tag{8}$$

With careful tuning of parameter α for a particular signal, one can achieve best denoising effect within thresholding framework.

The enhanced speech is then reconstructed using the inverse WP transform

$$x'(n) = \text{WPT}^{-1}\left\{W'^{J}_{K}, j\right\} \tag{9}$$

2.2 Dynamic Perceptual Wavelet Packet Feature Extraction

Wavelet coefficients provide flexible and efficient manipulation of a speech signal localized in the time–frequency plane [5–7]. The perceptual wavelet filter bank is built to approximate the critical band responses of the human ear. Wavelet packets decompose the data evenly into all bins but PWPs decompose only critical bins [5, 8, 14]. The size of the decomposition tree is directly related to the number of critical bins. The decomposition is implemented by an efficient 7-level tree structure, depicted in Fig. 2. The PWP transform is used to decompose $nx(n)$ into several frequency bands that approximate the critical bands. The terminal nodes of the tree represent a non-uniform filter bank.

The PWP coefficients for the sub-bands are generated as follows:

$$wj, i(k) = \text{pwpt}\,(nx\,(n)) \tag{10}$$

where $n = 1, 2, 3,\ldots L$ (L is the frame length),

$j = 0, 1, 2,\ldots, 7$ (j is the number of levels),
$i = 1, 2, 3,\ldots\ (2^j - 1)$ (i is the sub-band index in each level of j).

The static PWP coefficients are made more robust by computing the delta and the acceleration coefficients.

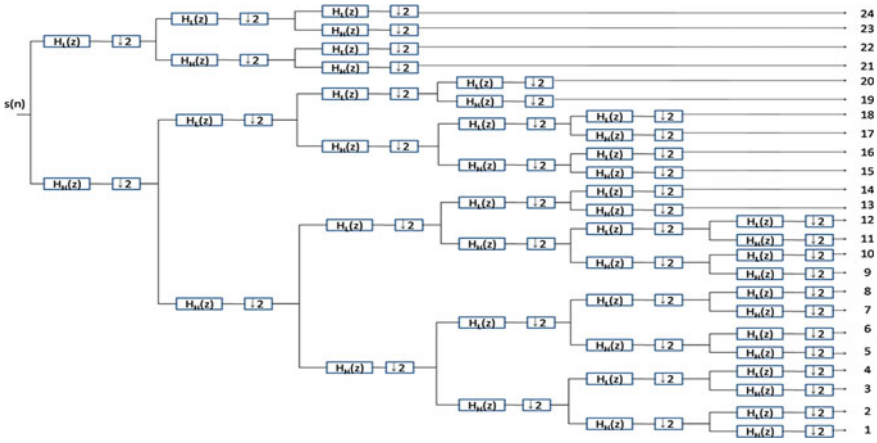


Fig. 2 Tree structure of PWPT

3 Speech Recognition

HMMs are the most successful statistical models for classification of speech. HMM is a stochastic approach which models the given problem as 'doubly stochastic process' [15–17].

A N-state HMM is defined by the parameter set $\lambda = \{\pi_i, a_{ij}, b_i(x), i,j = 1,\ldots N\}$, where

π_i initial state probability for state i,

a_{ij} transition probability from state i to state j,

$b_i(x)$ state observation probability density function (pdf) that is usually modeled by a mixture of Gaussian densities.

A five-state continuous density HMM is used as the statistical model for classification of speech signals. The hidden Markov model toolkit (HTK) is a portable toolkit for building and manipulating HMMs optimized for speech recognition [18].

4 Experimental Setup and Results

To evaluate the performance of the proposed method, recognition experiments were carried out using TIMIT database. The data were digitized with sampling rate of 16 kHz and 16 bits/sample quantization. This database consists of 200 training sentences from 6 male and female speakers and testing sentences randomly picked from training data. To simulate various noisy conditions, the testing sentences were corrupted by the additive white Gaussian noise with SNR conditions from 40 to 0 dB. The baseline recognition system was implemented on HTK toolkit with continuous density HMM models.

In the proposed feature extraction method, the perceptual wavelet features are extracted using MATLAB and written into HTK format using htkwrite function available from voicebox MATLAB package. The 24 perceptual wavelet filter banks are constructed by trial and error. The proposed tree shows excellent results using the Daubechies 45 prototype. Using a hamming windowed analysis frame length of 20 ms, the resulted 13-dimensional features plus their delta and delta–delta features, in other word, totally 39-dimensional features were used for speech recognition.

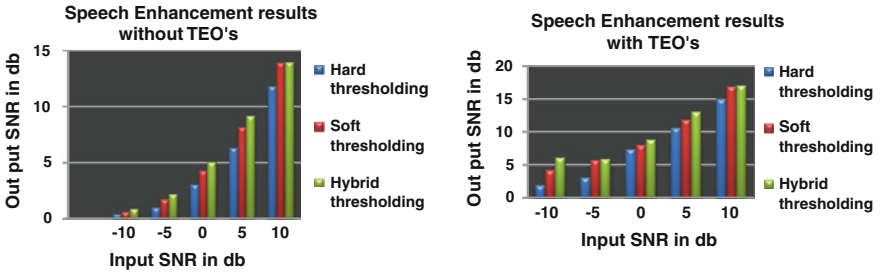
Table 1 shows that compared to the standard wavelet thresholding method, time-adapted wavelet-based hybrid thresholding using the TEOs as proposed in this paper has a better enhancement effect for SNR levels ranging from −10 to +10 db. The proposed wavelet domain front end features with HMM classifier–based ASR performs better recognition when compared to the conventional ASR with MFCC features and HMM as shown in Table 2. The PWPs capable of decomposing the critical bins during feature extraction supports in providing robust features for the ASR.

Table 1 Input/output SNR values obtained with different thresholding methods

Input SNR/output SNR (in db)		−10	−5	0	5	10
Hard thresholding	Without TEO	0.4184	1.2036	3.0537	6.2870	11.7655
	With TEO	1.9318	3.0627	7.3341	10.6041	14.9781
Soft thresholding	Without TEO	0.6155	1.7538	4.2496	8.1132	13.8283
	With TEO	4.2299	5.7624	8.0517	11.8721	16.8563
Hybrid thresholding	Without TEO	0.8724	2.1756	5.0328	9.1108	13.8886
	With TEO	6.0823	5.9035	8.8096	13.0061	17.0388

Table 2 Comparison of recognition accuracy for different features at various noise levels

Noise level	MFCC_E_D_A	MWAVELET_E_D_A
0 db	20.33	21
5 db	41.23	49.36
10 db	68.66	71.26
20 db	79.88	82.46
25 db	82.63	86.43
30 db	85.55	89.65
35 db	89.66	90.35
40 db	91.56	93.24



5 Conclusions

A wavelet domain front end for an ASR which combines both the enhancement and the feature extraction for different noise levels has been presented. The proposed time-adapted wavelet-based hybrid thresholding using the TEOs outperforms the conventional wavelet-based denoising schemes, as shown in Table 1. The dynamic PWP features, which decomposes only the critical sub-bands applied to a HMM–based classifier along with the proposed enhancement algorithm, recognize better than the conventional MFCC features with HMM as shown in Table 2. Further the ASR can be improved with a hybrid classifier into context and for larger database.

Acknowledgments We would like to express our sincere thanks to Aeronautical Development Agency, Ministry of Defence, DRDO, Bangalore, India, for supporting to do our research work.

References

1. Gong Y (1995) Speech recognition in noisy environments: a survey. Speech Commun 16(3):261–291
2. Rabiner L, Juang BH (1996) Fundamentals of speech recognition, vol 103 Prentice Hall Englewood Cliffs, New Jersey
3. O'Shaughnessy D (2001) Speech communication: human and machine. IEEE Press
4. Yusnita MA (2011) Phoneme-based or isolated-word modelling speech recognition system. In: IEEE 7th international colloquium on signal processing and its applications, pp 304–309
5. Jiang H et al (2003) Feature extraction using wavelet packet strategy. In: Proceedings of 42nd IEEE conference on decision and control, pp 4517–4520
6. Jie Y, Zhenli W (2009) Noise robust speech recognition by combining speech enhancement in the wavelet domain and Lin-log RASTA. ISECS 415–418
7. Gupta M, Gilbert A (2002) Robust speech recognition using wavelet coefficient features. IEEE Trans Speech Audio Process 445–448
8. Bourouba H et al (2008) Robust speech recognition using perceptual wavelet de-nosing and Mel-frequency product spectrum Cepstral coefficient features. Proc Informatica 32:283–288
9. Hesham T (2004) A time-space adapted wavelet de-noising algorithm for robust ASR in low SNR environments. IEEE Trans Speech Audio Process 1:311–314
10. Bahoura M, Rouat J (2001) Wavelet speech enhancement based on the Teager energy operator. Signal Process Lett IEEE 8(1):10–12
11. Chang S, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression. IEEE Trans Image Process 9(9):1532–1546
12. Donoho DL, Johnstone IM (1995) De-noising by soft-thresholding. IEEE Trans Inf Theory 41(3):613–627
13. Donoho DM, Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 90(432):1200–1224
14. Mallat S (2001) A wavelet tour of signal processing. Academic Press, London
15. Jisn H et al (2006) Large margin HMMs for speech recognition. IEEE Trans Speech Audio Lang Process 14(5):1584–1595
16. Vaseghi SV, Milner BP (1997) Noise compensation methods for HMM speech recognition in adverse environments. IEEE Trans Speech Audio Process 5(1):11–21
17. Mark J, Gales F, Young SJ (1996) Robust continuous speech recognition using parallel model combination. IEEE Trans Speech Audio Process 4(5):352–359
18. Young S (2009) The HTK book. Version 3.4. Cambridge University Engineering Department. Cambridge, UK

Combining Ensemble of Classifiers Using Voting-Based Rule to Predict Radiological Ratings for Lung Nodule Malignancy

K. Vinay, Ashok Rao and G. Hemanthakumar

Abstract In this paper, we are proposing new ensemble strategy for classification of lung nodules based on their malignancy ratings. The procedure we followed is simpler. In the first step, we construct different homogenous ensemble models such as bagged decision tree (BaDT), boosted decision tree (BoBT), and random subspace–based decision tree (RSSDT). In the next step, we combine previously constructed models with voting scheme to yield ensemble of homogenous ensemble of classifiers. We also examine the behavior of our method for heterogeneity in the system. This is done by constructing ensemble of heterogeneous ensemble of classifiers. For this, we have also considered bagged KNN (BaKNN), boosted KNN (BoKNN), bagged PART (BaPART), and boosted PART classifier (BoPART). The results we are obtaining from our strategy are significant compared to homogenous ensemble model.

Keywords Ensemble of classifier · Bagging · Boosting · Random subspace · Decision tree · KNN · PART

K. Vinay (✉) · G. Hemanthakumar
DoS in Computer Science, University of Mysore, Mysore, India
e-mail: gotovinni@gmail.com

G. Hemanthakumar
e-mail: ghk.2007@yahoo.com

A. Rao
Freelance Academician, 165, 11th main, S.Puram, Mysore, India
e-mail: ashokrao.mys@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_45,

1 Introduction

Since medical domain often encounters with problem of imbalance in input data, lung image database consortium (LIDC) [1] is one such example for imbalanced data. The description about the LIDC dataset is given in next section. We often see similar sort of imbalanced dataset in medical domain, and hence, machine learning techniques are very much required to efficiently solve such difficulties. Single classifier model often fails to estimate correct label and many time it will end up in bias and overfitting problems. Since the data are class imbalanced, there will be problem of overfitting of classifier toward the class which has more number of samples. This problem of overfitting can be overcome using machine learning techniques.

Machine learning techniques include ensemble of classifier methods, which use boosting, bagging, random subspace, random forest, ensemble selection, and many others. Since these methods deal with data resampling, subsampling, random subspace evaluating, fusion and selection of attributes and instances approaches. Therefore, one can expect good and supplementary results when compared to single classifier model.

2 Literature Review

In this paper, we have given a brief review about various attempts which are made using machine learning techniques for the problem of classification. Ekrain et.al [2] investigated several approaches to combine delineated boundaries and ratings from multiple observers, and they have used P-map analysis with union, intersection, and threshold probability to combine the boundary reading and claimed that threshold probability approach provides good level of agreement. Ebadollahi et al. [3] proposed a framework that uses semantic methods to describe visual abnormalities and exchange knowledge with medical domain. Nakumura et al. [4] worked on simulating the radiologists' perception of diagnostic characteristic rating such as shape, margin, irregularity, spiculation, lobulation, texture on a scale of 1–6, and they extracted various statistical and geometric image features including Fourier and radial gradient indices and correlated these features with the radiologists' ratings.

Ensembles are particularly suitable for datasets with a large dimensionality of the feature set and create diversity while retaining the number of instances for training. It has also been shown in literature that ensembles work well for fMRI data [5]. Traditionally, linear or homogenous base classifiers have been used in numerous instances. However, in this work, we consider a set of seven heterogeneous base classifiers viz., naive Bayes, k-nearest neighbor, J48 decision trees, random forest, support vector machine (SVM), radial basis function (RBF) neural network, and multilayer perceptrons (MLP) in the ensemble. These use different methods such as voting, stacking, and AdaBoost with diverse ensemble creation by

oppositional relabeling of artificial training examples (DECORATE) as a meta-learner. These exhibit better performance over linear or homogenous ensemble-based classifier model. Significant work toward designing panel of expert machine learning classifier which automates the radiologist work of predicting nodule ratings is done by Dmitriy and Raichu [6], and they have proposed active, decorate, a new meta-learning strategy for ensemble of classifier domain.

3 Methods and Materials

LIDC [1] provides lung CT image data which are publically available through National Cancer Institute's Imaging Archive (Web site—http://ncia.nci.nih.gov). This dataset consists of image data, radiologist's nodule outline details, and radiologist subjective characteristic ratings. The LIDC dataset currently contains complete thoracic CT scans of 399 patients acquired over different periods of time. LIDC data download comes with DICOM image and the nodule information in the XML file. This has information regarding the spatial location information about three types of lesions, and they are nodules < 3 mm, nodules > 3 mm, and non-nodules > 3 mm in maximum diameter as marked by panel of 4 expert radiologists. For any lesion greater than 3 mm in diameter, XML file contains spatial coordinates of nodule's outline. Since the number of radiologist in LIDC panel is 4, it is obvious that each nodule > 3 mm has 4 nodule outlines. Moreover, any radiologist who identifies a nodule > 3 mm also provides subjective ratings for 9 nodule characteristics, viz., lobulation, internal structure, calcification, subtlety, spiculation, margin, sphericity, texture, and malignancy.

In this work, we are considering subset of data from LIDC and the overview of dataset is given in Table 1.

3.1 Image Feature Extraction

In this work, we consider same set of features which we have calculated in our earlier work [7, 8]. Our feature set includes 55, two-dimensional, low-level image features grouped into four categories: size feature, shape feature, intensity feature, and

Table 1 Overview of dataset

Dataset and feature extraction details	
No. of cases considered	124
No. of instances	14,956
No. of nodules	4,532
No. of low-level image features extracted	55
No. of radiologist characteristic ratings considered	08
Total no. of features used in the work	63

Table 2 Low-level image features

Size feature	Shape feature	Intensity feature
Area	Circularity	MinIntensity
Convex area	Roughness	MaxIntensity
Perimeter	Elongation	MeanIntensity
Convex perimeter	Compactness	SDIntensity
EquivDiameter	Eccentricity	
MajorAxisLength	Solidity	
MinorAxisLength	Extent	

Texture features

24 Gabor features are mean and standard deviation of 12 different Gabor response images at orientation = 0, 45, 90, 135 and time frequency = 0.3, 0.4, 0.5

13 Haralick features calculated from co-occurrence matrices: energy, correlation, inertia, entropy, inverse difference moment, sum average, sum variance, sum entropy, difference, average, difference variance, difference entropy, information measure of correlation 1, information measure of correlation 2

texture feature. Further, we consider the nodule which has the largest area and the image features are extracted only for this largest nodule. Table 2 gives the details about low-level image features that we have used in this work.

4 Proposed Method

In this work, we are trying to boost the performance of ensemble of classifier response by using voting method. Voting method is a class for combining the classifier responses, and it is originally given by Ludmila and Kuncheva [9, 10]. Different combinations for probability estimates such as average of probabilities, product of probabilities, majority voting, minimum of probabilities, maximum of probabilities, and median rule for classification are available. Our procedure in this work is of twofold. In stage once we have constructed the ensemble of classifier using bagging, boosting, and random subspace approaches. In the later stage, we use voting principle to further classification task. In voting method, we have used majority vote scheme to estimate the probability.

Further, we have trained classifiers such as k-nearest neighbor and PART, a rule-based classifier to construct the ensembles. The idea behind choosing these classifiers is to create the heterogeneous ensemble of classifier and to compare between homogenous ensembles of classifier. The brief description and algorithm for bagging, boosting, and random subspace what we have used in this work are given in subsequent sections.

4.1 Bagging

Breiman [11] introduced the concept of bagging. The term bagging is acronym for bootstrap aggregating. The steps used in bagging are as follows: The committee of classifier is built on bootstrap replicates of the training set. The classifier outputs are later combined by voting principle. Only labeled training set is afforded in bagging. Sample with or without replacement of original data points to create a new training set of certain length is carried out. To make use of variations in training set, the base classifier should be unstable, that is, small changes in training set should lead to large changes in classifier output.

4.2 Boosting

Boosting algorithm allocates weights to a set of strategies used to predict the outcome of a certain event. Boosting method was originally inspired by an online learning algorithm called *HEGDGE(β)*. The weight of strategy S_i can be interpreted as the probability that S_i is the best predicting strategy in the group. The distribution is updated online after each new outcome. Strategies with the correct prediction receive more weight, while the weights of the strategies with incorrect predictions are reduced.

In our work, we have used AdaBoost. M1 algorithm to carry out experiments. AdaBoost was proposed initially for two classes and then extended for multiple classes. AdaBoost. M1 is the most straightforward multiclass extension of AdaBoost. We have constructed ensemble of boosted J48 decision tree and ensemble of boosted KNN classifier using AdaBoost. M1 method.

4.3 Random Subspace

Choosing random subsets of features is termed the random subspace method [12]. Each classifier in the ensemble is built upon a randomly chosen subset of features. The random subspace method has been found to work well when there is redundant information which is distributed across all the features rather than concentrated in a subset of them. The random subspace method we have considered in this work is given by [13]. This method constructs decision tree–based classifier that maintains the highest accuracy on training data and improves on generalization accuracy as it grows in complexity. The decisions of individual trees are combined by averaging the conditional probability of each class at the leaves. In our work, we have used J48 decision tree as base classifier to construct ensemble of J48 tree using random subspace method.

5 Results and Discussion

In this work, we have experimented with reference to seven performance metrics: accuracy (ACC), root mean square error (RMSE), true-positive rate (TPR), false-positive rate (FPR), F-measure, area under curve (AUC), and kappa statistics. These metrics are chosen because we have categorized the behavior of performance into four groups: threshold metric, probabilistic metric, rank metric, and agreement metric. We have fixed threshold to 0. The classifier which performs above the threshold (>0.5) is considered to be good performer, and the classifier which performs below the threshold (<0.5) is regarded as under performer. RMSE is used as probability metric. Probability metric is minimized when the predicted value for each case is equal to true conditional probability. Lower the RMSE value, better the performer. TPR, FPR, and AUC are used as a rank metric, and these metrics measure how well the positive cases and negative cases are ordered and viewed. Kappa statistics is used as agreement measures, which in turn reflect how well model agrees between the expert prediction and machine prediction. The kappa interpretation scale has been given in Table 3.

We have carried out experiments in three stages. In first stage, we have trained independent classifier ensembles, that is, bagged J48 tree, bagged KNN, bagged PART, boosted J48 tree, BoKNN, boosted PART, and random subspace with J48 as base classifier. The environment and parameters for classifier model are given below.

J48 decision tree: tree split criteria = binary split, confidence factor = 0.25, minimum number of objects = 2.

KNN classifier: K = 5, Linear search technique with Euclidean distance measure is used as nearest neighbor search algorithm.

PART classifier: With binary split and confidence factor = 0.25, minimum number of objects = 2.

Bagging: Bag percent size = 100 and number of iterations = 50.

AdaBoost. M1: Number of iterations = 50; we used reweighting instead of resampling for building new training set, and weight threshold for pruning is set to 100.

Random Subspace: Number of iteration = 50 and subspace size = 50 %.

With the above parameter setup, we experimented and tabulated the results in Tables 4 and 5. We have considered J48 decision tree with bagging, boosting, and

Table 3 Kappa statistics interpretation scale

K-value	Strength of agreement
<0	Poor
0–0.2	Slight
0.21–0.4	Fair
0.41–0.6	Moderate
0.61–0.8	Substantial
0.81–1	Almost perfect

Table 4 Results from experiments using ensemble of homogenous classifiers

Classifier model	ACC	RMSE	TPR	FPR	F-measure	AUC	Kappa	Mean
BaDT	**0.84**	**0.21**	**0.85**	0.01	**0.90**	**0.98**	**0.80**	**0.65**
BaKNN	0.67	0.29	0.74	0.02	0.79	0.94	0.58	0.57
BaPART	0.82	0.23	0.83	**0.00**	0.89	**0.98**	0.78	0.64
BoDT	**0.84**	0.24	0.84	0.00	**0.90**	**0.98**	**0.80**	**0.65**
BoKNN	0.66	0.32	0.77	0.03	0.78	0.90	0.57	0.57
BoPART	0.83	0.25	0.84	0.00	**0.90**	**0.98**	0.79	**0.65**
RSSDT	0.82	0.25	0.82	0.01	0.88	**0.98**	0.77	0.64

Table 5 Results from the experiments using voting classifier

Classifier Model	ACC	RMSE	TPR	FPR	F-measure	AUC	Kappa	Mean
Voting (ensemble of homogenous classifiers)	1.00	0.00	1.00	0.00	1.00	1.00	1.00	0.71
Voting (ensemble of heterogeneous classifiers)	0.99	0.01	1.00	0.00	1.00	1.00	1.00	0.71

random subspace in first set. This is because we require noticing the homogenous nature of ensemble of classifier. Later, we have included BaKNN, BaPART, BoKNN, and PART to create heterogeneous ensemble of classifier.

5.1 Ensemble of Homogenous Classifiers' Response

When we have trained the model with bagging and boosting methods, in the most of scenario, irrespective of the base classifiers chosen, the ensemble methods yielded similar results. The maximum performances are highlighted in table. When we compare the average of all seven metrics, we obtained 0.65 for BaDT, BaPART, and BoPART.

5.2 On Combining Classifier Models with Voting

When the responses from previously built ensemble classifier model are used with voting principle, we have noticed the significant improvement in all the categories of performance metric. In case of voting with homogenous ensemble of classier, the voting seems to work much better and we have achieved very good results, whereas voting with heterogeneous ensemble of classifier is also performed much better compared to previously built ensemble. But homogenous ensemble of voting-based model has shown small improvement over heterogeneous based

ensemble in probabilistic metric. When we consider average of performance, then both the models are equally performing.

6 Conclusion

In this paper, we have proposed new strategy which can be regarded as ensemble of ensemble-based classifier. Our main objective is to identify the significance of outer level meta-learning on good performing ensemble of classifiers. Our result showed that using voting class as combination rile, one can still improve the results of ensemble of classifiers. In the field of machine learning, it has been claimed that ensemble of classifier performs better when compared to single classifier model. Our method has given sufficient proof that using higher-level combination rule on good performing classifier, it is possible to obtain much significant and supplementary results.

References

1. Lung Image Database Consortium (LIDC) available at http://ncia.nci.nih.gov
2. Varutbangkul E, Mitrovic V, Aichu D, Furst J (2008) Combining boundaries abd rating from multiple observers for predicting lung nodule characteristics. In: IEEE international conference on biocomputing, bioinformatics and biomedical technologies, pp 82–87
3. Ebadollahi S, Johnson DE (2008) Diao M. Retrieving clinical cases through a concept space representation of text and images. SPIE Med, Imaging Symp
4. Nakumura K, Yoshida H, Engelmann R. MacMahon H, Kasturagawa S, Ishida T, et al (2000) Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks. Radiology, pp 823–30
5. Kuncheva LI, Rodr′ıguez JJ, Plumpton CO, Linden DEJ, Johnston SJ (2010) Random subspace ensembles for fMRI classification. IEEE Trans Med Imaging 29(2):531–542
6. Zinovev D, Raicu D, Furst J, Armato SG (2009) Predicting radiological panel opinions using a panel of machine learning classifiers. Algorithms 2:1473–1502. doi:10.3390/a2041473
7. Vinay K, Rao A, Hemantha Kumar G Comparative study on performance of single classifier with ensemble of classifiers in predicting radiological experts ratings on lung nodules. Indian international conference on artificial intelligence (IICAI-11). ISBN: 978-0-9727412-8-6, pp 393–403
8. Vinay K, Rao A, Hemantha Kumar G Computerized analysis of classification of lung nodules and comparison between homogeneous and heterogeneous ensemble of classifier model. 2011 third national conference on computer vision, pattern recognition, image processing and graphics, 978-0-7695-4599-8/11, IEEE doi:10.1109/NCVPRIPG.2011.56, pp 231–234
9. Kuncheva LI (2004) Combining pattern classifiers: methods and algorithms. Wiley, Inc, Chichester
10. J. Kittler, M. Hatef, Robert P.W. Duin, J. Matas (1998). On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence. 20(3):226–239
11. L. Breiman. Bagging predictors. Technical Report 421, Department of Statistics, University of California, Berkeley, (1994)

12. Freund Y, Schapire RE (1997) A decision–theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55(1):119–139
13. Ho TK (1998) The random space method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20(8):832–844

A Secured Authentication System Using an Effective Keystroke Dynamics

G. Jagadamba, S. P. Sharmila and Thejas Gouda

Abstract In the field of computer security, most promising field is securing the data by allowing ease access to authorized users. The biometric techniques like face recognition, voice recognition and digital signatures provide good authentication security. The keystroke dynamics is defined to be a low cost, strong behavioral biometric-based authentication system, based on consistent typing rhythm patterns at a keyboard terminal, which will be individually unique. This paper exhibits an effective, efficient and robust user authentication. Authentication system is based on effective Adaptive Learning Classification (ALC) algorithm, where a self-threshold for each user was decided based on user input. Training and testing data lead to an average false reject rate of 10.00 % and the average false accept rate of 0.0025 %.

Keywords Biometrics · Keystroke dynamics · Rejection rate · Acceptance rate

1 Introduction

The increasing use of automated information systems together with our pervasive use of computers has greatly simplified our lives, while making us immensely dependent on computers and digital networks. For any information system to serve

G. Jagadamba (✉) · T. Gouda
Department of Information Science and Engineering, Siddaganga Institute of Technology, Tumkur-03, India
e-mail: jagadambasu@gmail.com

T. Gouda
e-mail: thejas777777@gmail.com

S. P. Sharmila
Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur-03, India
e-mail: sharmila.h.shukthij@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_46,

its purpose, information should be protected from unauthorized access, disclosure, disruption, modification, perusal, inspection, recording or destruction. Information security includes five security characters [1] like confidentiality, integrity, availability, authenticity and accountability, which can be mutually exclusive.

"User authentication is an assurance that communicating entity is the one claimed." User authentication is the basis for most types of access control and for user accountability. It can be categorized into three classes [2–5], namely,

Knowledge based: such as password or secret information.

Object or Possession: such as smart card, passport and driver's licenses.

Biometric based: such as fingerprint, iris and voice.

Biometrics user authentications are the technologies that are measurable on physiological or behavioral characteristics [6]. A behavioral biometric recognition system can run in two different modes; identification and verification. Here, automated approaches will be defined for an identity for each user (human or process). The verification process realizes the data capture and the feature extraction and the comparison with the biometric model. Two commonly used standards in biometrics according to [1] are False Rejection Rate (FRR) and False Acceptance rate (FAR), and these should be as less as possible for any good biometric system. Keystroke rhythm is a natural choice for computer security, which is an art of recognizing an individual based on typing patterns. It includes several different measurements like latency between consecutive keystrokes, duration of the keystroke and hold-time [6], which can be detected when the user presses the keys in the keyboard. In this paper, we propose an efficient safeguard system based on authenticating a user to access computers by recognizing certain unique and habitual patterns with a user's typing rhythm. Rest of the paper is organized as follows: Sect. 2 highlights related work, Sect. 3 focuses on keystroke latency metrics and implementation of user authentication system, Sect. 4 includes testing results, and Sect. 5 concludes the proposed work.

2 Related Work

The keystroke rhythms of the user are measured to develop a unique biometric template of the user's typing pattern for future authentication. Raw measurements available from each keyboard can be recorded to determine dwell time (the time a key is pressed) or flight time (the time between key down and the next key down and the time between key up and the next key up). After recording, data are processed through the algorithm, which serves the primary pattern for future comparison and analysis.

Different methodologies have been adopted depending upon dwell time and flight time. A summary of existing methodologies with remarks has been collected from [2, 6–10] are listed here (Table 1).

Table 1 Different keystroke methodologies used for training and testing

Sl. No	Methodology	Remarks
1	Hold key timings	Total time periods and pressure were measured using Euclidian distance measure between two vectors of typed characters, stored as template
2	Ant colony optimization technique	Mean and standard deviation was used to extract the features from the keystroke duration, latency and digraph
3	Keystroke timings	Successive keystrokes were recorded and used for authentication and result with FAR 4 % for seven users
4	Standard and measure	The mean, standard deviation of keystroke latencies and digraph between reference profile and test data are compared. A result of 17 % FAR and 30 % FRR was obtained
5	Pattern recognition and neural network	Fuzzy ARTMAP, RBFN and LVQ neural network paradigms were used. BPSTF, potential function and Bayes' rule algorithms gave moderate performance
6	Bio password	This technique is fast, accurate, transparent and scalable
7	Parallel decision trees (DTs) on keystroke patterns	By this method the average false reject rate was 9.62 % and the average false accept rate was 0.88 %
8	Telling hUman and bot apart (TUBA)	Monitoring user's keystroke-dynamics patterns and identifying intruders. The bot-generated keystroke sequences are detected with high true positive rates 93 %.
9	Virtual key force	This method improves the accuracy by 90.4 % and reduces the training and testing time within 0.025 s

3 User Authentication Using Keystroke Dynamics

In this model, we adopt a static keystroke authentication methodology, which is perfectly suitable to authenticate an individual by asking user to type his own password. Before login to his computer session, typing rhythm is verified with the prototype stored in the server database. Changing the password implies to enroll again, because the methods are not able to work with a different password. When we build a model, we take the latencies between adjacent keystrokes among several samples of a user stored in the database and then compute a vector of means and standard deviations for the latencies between each pair of keystrokes using an adaptive learning classification (ALC) algorithm. The vector of means and standard deviations then represents the user's profile and used to classify the users as authenticator or intruder.

3.1 Keystroke Latency Metrics

In our approach, we monitor all the key events that the user types which includes alphanumeric values including shift, backspace and caps lock. Typing one key triggers a pair of key events: press and release, which we call a keystroke. The Press Release-latencies and Release Press-latencies are grouped, respectively, in three different ways: bigram, trigram and word-gram. A key event can be a bigram event, a trigram event or a word-gram event [2]. For our proposed system, we considered the word-gram and bigram event.

3.2 Implementation

The system architecture consists of client, server and database. The client acts like an end user with various types of activities. Client performs activities like login to system, transactions, view to his data. Server includes the functions like registering the server with Remote Method Invocation (RMI) registry, establishing the communication to client, processing of all clients requests, running keystroke procedure to check for valid user, sending the successful key typing latencies to database and provides application services to authenticated clients, if any variation a third-level security policies will be defined. Where third-level policies may include sending random code to user's mobile, user's email account or check security answer sent by client and proceed further processing. Database is responsible for storing all the relevant keystroke latency data and also user information. It will receive the queries from the server process, execute them and return the corresponding results back to the server process. A database server is used to store various information required by the server process. Information like user details, commodity prices, details of transactions, incoming and outgoing messages is stored on the database in different tables. The secret information of a user, that is, password will be stored in an encrypted form.

3.3 Adaptive Learning Classification Algorithm

When the user signs up, user typing samples are captured without his knowledge for first five consecutive times. After sample collection process, the keystroke biometrics authentication mechanism comes into operation. Whenever the user tries to login to the system, the average of the recent five successful transactions typing frequency is considered to cross check with the current typing frequency. In addition, a threshold value is provided for the respective users, where the threshold is the standard deviation of the samples. If the typing frequency of the users is between the above or below the average threshold, then the user is authenticated;

otherwise a user is unauthenticated and third-level security policies will be presented. The following steps are followed in ALC algorithm to make the decision authenticator or intruder.

ALC Algorithm

1. Initialization: Capturing the values, x = d, from username = 'Specific User'.
2. Template Design: Mean and standard deviation is calculated for the specific user, and the corresponding threshold is calculated.

$$\mu = \frac{\sum_{i=1}^{n} x_i}{N}$$

$$\sigma = \sqrt{\frac{\sum_i (x_i - \mu)^2}{N}}$$

$$\text{Upper Threshold} = (\mu + \sigma)$$

$$\text{Lower Threshold} = (\mu - \sigma)$$

3. Analysis: if (Lower Threshold $\leq \mu \leq$ Upper Threshold)-authenticator, else intruder.
4. Decision: if (Authenticator) Service is continued else (intruder) third-level security policies will be introduced.

4 Testing Results

Testing was carried out in password string latency, that is, flight time between the first and the last key pressed and password character latency, that is, flight time between the successive pair of keys pressed. During the experiment, the delete operations and irrelevant shift key pressed were eliminated from the data. A user is able to login successfully only when successful match is made between stored genuine template and live sample falls below the specified threshold. The false match is registered when a live sample exceeds the specified threshold, then the user will not be authenticated. In the testing procedure of string latency, four participants were allowed to login to their own account and data collected where analyzed. The Fig. 1 shows the frequency variations for 10 attempts, and Fig. 2 shows the experimental results where failure rates of FRR falls around 10–40 %. In the same way, character latencies where collected for 4 participants by allowing them to login to a particular account and frequency variations are depicted in Fig. 3. The success rate of FAR falls between 0 and 20 % and success rate of FAR is almost 0.0025 %, which is shown in Fig. 4. All the tests are done for an Internet banking application, and the response time was observed to be in terms of 2–5 ms

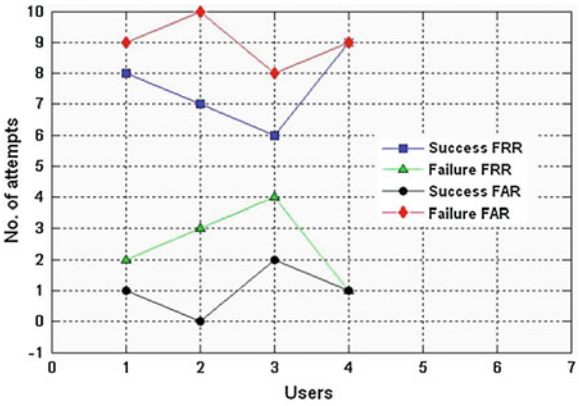


Fig. 1 Frequency variations of string latency

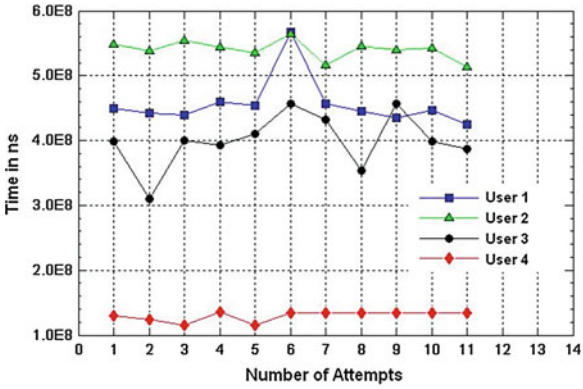


Fig. 2 FRR and FAR for string latency

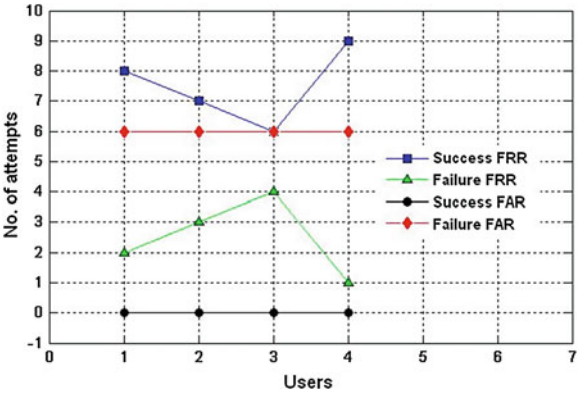


Fig. 3 Frequency variations of character latency

with the Internet connectivity of 10.0 Mbps, with some 50 % of traffic in the channel in client–server architecture. The third-party policy verification response times are observed to be almost 2 ms for the designed system.

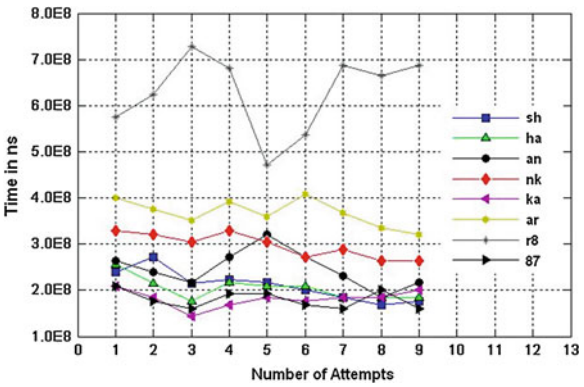


Fig. 4 FRR and FAR for character latency

5 Conclusion

We have presented a framework for defining a quick and secure authentication approach and do not need any special hardware except keyboard for the applications in the highly connected world of tomorrow. This approach has no effect on someone's privacy, since the behavior of an individual is considered. As individual behavior cannot be copied, it becomes an identification token for verification. Our keystroke biometrics exhibits an inexpensive user authentication system for applications like Internet banking. This can be easily adoptable for all online commerce. Performance can be improved by minimizing the FAR by adapting relative keystrokes with good fault tolerance methods.

References

1. Giot R, Dorizziy B, Rosenberger C (2011) Analysis of template update strategies for keystroke dynamics. In: Proceedings of IEEE symposium series on computational intelligence, vol 1. Paris, pp 21–28
2. Shanmugapriya D, Padmavathi G (2011) An efficient feature selection technique for user authentication using keystroke dynamics. Proc Int J Comput Sci Netw Secur 11(10):191–195
3. Bishop M (2003) Computer security: art and science. In: Proceedings of 51th symposium on networked and distributed system security, Addison-Wesley Professional. Boston, MA, ISBN 0-201-44099-7, p 123130
4. Sukhai NB (2004) Access control & biometrics. In: Proceedings of ACM 1st annual conference on information security curriculum development. New York, pp 124–127
5. Williams JM (2002) Biometrics or biohazards? In: Proceedings of ACM workshop new security paradigms. New York, pp 97–107
6. Karnan M, Krishnaraj N (2010) Bio password-keystroke dynamic approach to secure mobile devices. In: Proceedings of conference IEEE-2010, international conference on computational intelligence and computing research, doi:10.1109/ICCIC
7. O'Gorman L (2003) Comparing passwords, tokens, and biometrics for user authentication. Proc IEEE 91(12):2019–2040

8. Karnan M, Akila M, Kalamani A (2009) Feature subset selection in keystroke dynamics using ant colony optimization. Proc J Eng Technol Res 1(5):072–080
9. Sheng Y, Phoha VV, Rovnyak SM (2005) A parallel decision tree-based method for user authentication based on keystroke patterns. In: Proc IEEE Trans Syst Man Cybern Part B: Cybern 35(4):826–833
10. Stefan D, Shu X, Yao D (2011) Robustness of keystroke-dynamic based biometrics against synthetic forgeries. Elsevier ltd, doi:10.1016/j.cose.1011.10.001

High-Speed Low-Power VLSI Architecture for SPST-Equipped Booth Multiplier Using Modified Carry Look Ahead Adder

Pratima S. Mudhenagudi, Narayan V. Sugur, Saroja V. Siddamal and R. M. Banakar

Abstract Nowadays, various techniques have been developed for reducing the power consumption of VLSI designs, such as pipelining and parallel processing, reducing the dynamic power, voltage scaling, clock gating, etc. To increase the processing speed of the silicon IC, logic gates are made using CNT FETs and designed using the VLSI technology. Lowering down the power consumption and enhancing the processing speed of IC designs are undoubtedly the two important design challenges in designing ICs. The objective of this paper is to provide high-speed and low-power multiplier. In this paper, a VLSI designed low-power; high-speed multiplier is proposed using the SPST approach. This multiplier is designed by applying the spurious power suppression technique (SPST) on booth multiplier. To add partial products, Modified CLA adder and SPST adder are used. The proposed architecture is synthesized. In Xilinx RTL, chip X-C6SLX45T3FGG484 Spartan 3 series is selected for benchmarking. The timing report shows that to perform 16-bit multiplier, the minimum period required for array multiplier is 21.068 ns, and SPST-equipped booth multiplier requires 14.886 ns. The proposed multiplier requires 13.263 ns. An improvement of 37.04 % is achieved in speed when compared to the proposed multiplier with array multiplier, and an improvement of 10.9 % is achieved in speed when compared to SPST-equipped booth multiplier adder. The proposed multiplier implementation with AND gates has an extremely high flexibility on adjusting the data asserting time. This facilitates the robustness of SPST can attain 30 % speed improvement.

P. S. Mudhenagudi (✉) · N. V. Sugur · S. V. Siddamal · R. M. Banakar
B V Bhoomaraddi College of Engineering and Technology, Hubli, India
e-mail: smpratima@gmail.com

N. V. Sugur
e-mail: narayansugur@yahoo.in

S. V. Siddamal
e-mail: sarojavs@bvb.edu

R. M. Banakar
e-mail: banakar@bvb.edu

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_47,

Keywords Booth encoder circuit · SPST technique · Detection unit · MCLA adder · SPST adder

1 Introduction

Fast multipliers are essential parts of digital signal processing systems. The speed of multiply operation is of great importance in digital signal processing as well as in the general-purpose processors today, especially since the media processing took off. Multiplication can be considered as a series of repeated additions. The number to be added is the multiplicand, the number of times that it is added is the multiplier, and the result is the product. Each step of addition generates a partial product. This repeated addition method that is suggested by the arithmetic definition is slow that it is almost always replaced by an algorithm that makes use of positional representation.

The basic multiplication principle is twofold, that is, evaluation of partial products and accumulation of the shifted partial products. It is performed by the successive additions of the columns of the shifted partial product matrix. The multiplier is successfully shifted and gates the appropriate bit of the multiplicand. The delayed gated instance of the multiplicand must all be in the same column of the shifted partial product matrix. They are then added to form the product bit for the particular form. Multiplication is therefore a multi-operand operation [1–3].

A system's performance is generally determined by the performance of the multiplier because the multiplier is generally the slowest clement in the system. Furthermore, it is generally the most area-consuming. Hence, optimizing the speed of the multiplier is one of the major design issues.

This paper presents a low-power booth multiplier design with spurious power suppression technique (SPST) and modified CLA adder (MCLA) in which modified Booth multiplier improves speed of multipliers and SPST adder will reduce the power consumption in addition process by minimizing the switching activity [1, 3, 4]. The rest of the paper is organized as follows. In Sect. 2, spurious power suppression technique is explained. Section 3 describes the architecture design of 16 bit SPST booth multiplier. In Sect. 4, MCLA adder and, in Sect. 5, SPST adder are discussed. Section 6 pave a way to the implementation and results followed by conclusion in Sect. 7.

2 Spurious Power Suppression Technique

The SPST uses a detection logic circuit to detect the effective data range of arithmetic units, for example, adders or multipliers [1]. When a portion of data does not affect the final computing results, this portion is latched to avoid useless

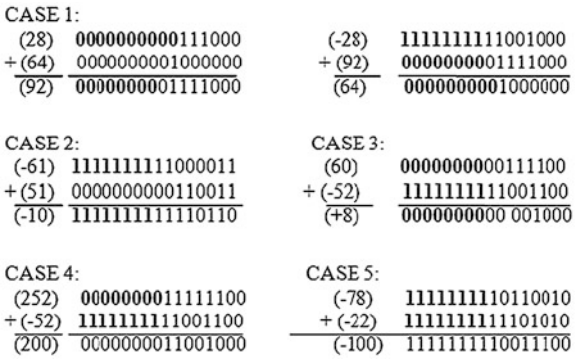


Fig. 1 Examples for SPST addition

data transitions occurring inside the arithmetic units. This data-controlling unit brings evident power reduction [3, 5].

To illustrate the SPST, five cases of a 16-bit addition are explained as shown in Fig. 1. The 1st case illustrates a transient state in which the spurious transitions of carry signals occur in the MSP though the final result of the MSP is unchanged. The 2nd and 3rd cases describe the situations of one negative operand adding another positive operand without and with carry from LSP, respectively. Moreover, the 4th and 5th cases, respectively, demonstrate the conditions of two negative operands addition without and with carry-in from LSP. In those cases, the results of the MSP are predictable; therefore, the computations in the MSP are useless and can be neglected.

Eliminating those spurious computations will not only save the power consumed inside the SPST adder/sub tractor but also decrease the glitching noises which will affect the next arithmetic circuits [1, 3, 4].

3 Proposed SPST-Equipped Booth Multiplier

In Fig. 2, the architecture SPST-equipped Booth multiplier is illustrated. Multiplicand is fed to partial product generator block where five candidates of partial products are generated, that is, $\{-2A, -A, 0, A, 2A\}$. The advantage of this method is halving the number of partial products.

The set of partial product will then be passed through the latches in the next stages. These latches will check for spurious activity by simple ANDing operation. To check for the spurious activity, two additional signals, that is, close1 and close2 are obtained by ANDing the most significant bits of multiplier bits as shown in Eqs. 3 and 6.

To generate the partial product, Radix-4 encoding scheme is used as shown in Table 1. Using partial product generator circuit, these partial products can be obtained as shown in Fig. 3. These partial products are selected using multiplexer and added finally to the adder block.

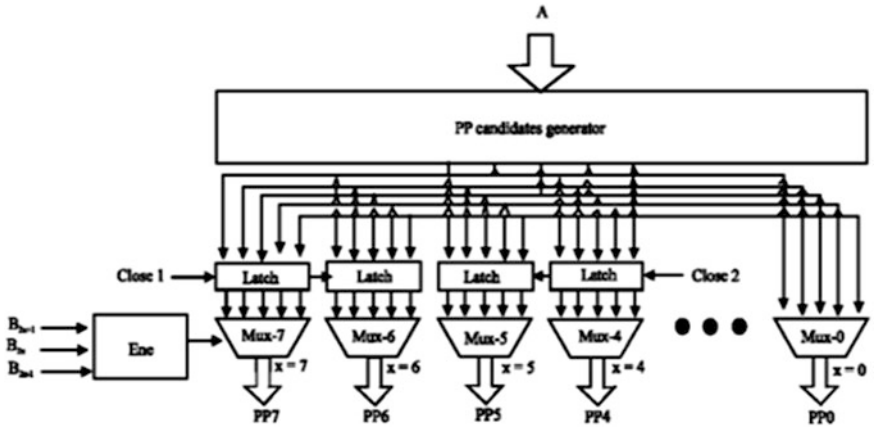


Fig. 2 Booth encoder circuit

Table 1 Radix-4 encoding scheme

B_{n+2}	B_{n+1}	B_n	Partial product(PP)
0	0	0	$0 \times A$
0	0	1	$1 \times A$
0	1	0	$1 \times A$
0	1	1	$2 \times A$
1	0	0	$-2 \times A$
1	0	1	$-1 \times A$
1	1	0	$-1 \times A$
1	1	1	$-0 \times A$

$$A_{MSP} = A[15:0] \quad ; \quad B_{MSP} = B[15:0]; \tag{1}$$

$$A_{and} = A\ [15] * A\ [14] * \ldots A\ [0]; \tag{2}$$

$$B_{and} = B\ [15] * B\ [14] * \ldots B[0]; \tag{3}$$

$$A_{nor} = \overline{A[15] + A[14] + \cdots A[0]}; \tag{4}$$

$$B_{nor} = \overline{B[15] + B[14] + \cdots B[0]}; \tag{5}$$

$$\text{Close} = \overline{(A_{and} + A_{nor}) * (B_{and} + B_{nor})}; \tag{6}$$

The partial products from the booth encoder circuit will be fed to the adder block. Since there are more chances of spurious activity in LSP part, SPST adder are used. In MSP part, MCLA adders are used as shown in Fig. 4.

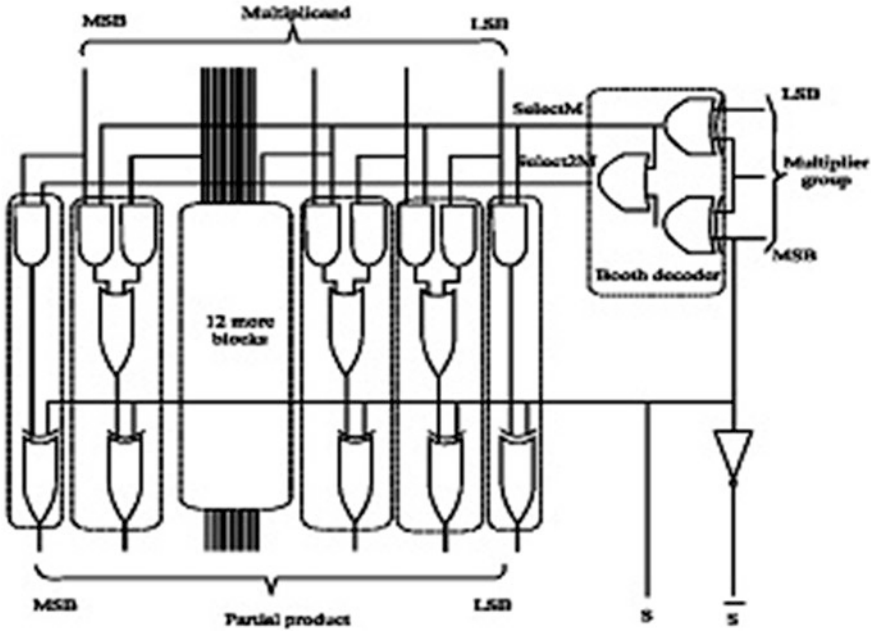


Fig. 3 Partial product generator

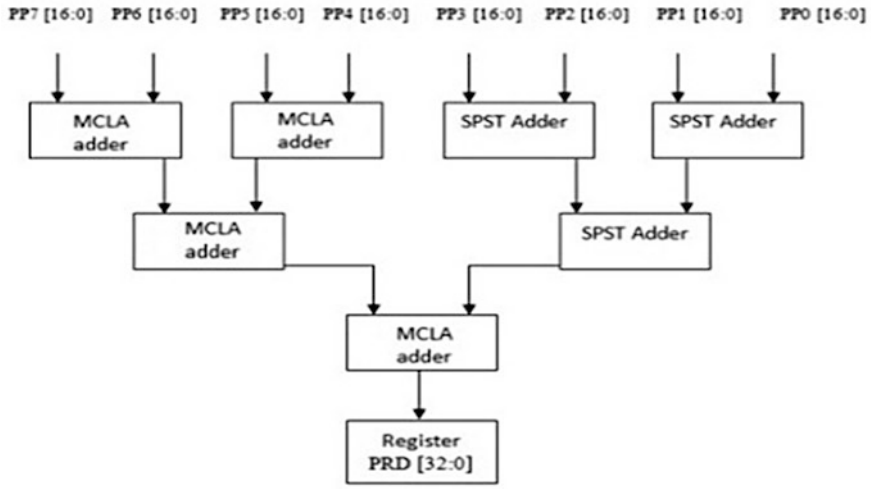


Fig. 4 Adder block

4 Modified CLA Adder

The main concept of MCLA is to use NAND gates to replace AND and NOT gates of CLA adder. The design of MCLA adder consist of two parts,

- Arithmetic adder circuit
- Carry look ahead circuit

In MCLA, a new full adder called metamorphosis of PFA is used. Here, the generator g_i is produced by using NAND gate. The carries of the next stage are also generated using NAND gate and is written as

$$c_{i+1} = \overline{\overline{g}_i(\overline{p_i c_i})} \tag{7}$$

Where C_{i+1} is the next stage carry, C_i is the previous carry, g_i is the generate logic, and p_i is propagate logic. In metamorphosis PFA, the signal $\bar{g}_i$ is implemented with NAND gate. It is faster than the g_i of PFA implemented with AND gates. The above-generated carries are implemented in carry look ahead circuit. The AND gate is used to generate g bit. This can be simplified using NAND gate and a NOT gate to replace AND gate of the previous level. With this logic one gate, delay is increased due to one NOT gate. To reduce this in MCLA, NOT gate of previous stage is canceled with NOT gate of present stage. With the cancellation of NOT gate, one gate delay is reduced at every stage. This again adds to the improvements of the speed.

5 16-Bit SPST Adder

The 16-bit SPST adder used in the adder block is divided into MSP and LSP at the place between 8th and 9th bit as shown in Fig. 4. Latches implemented by simple AND gates are used to control the input data of the MSP. When the MSP part contains continuous zeros or ones, MSP part calculation is eliminated. Otherwise, it will act as a normal 16-bit adder as shown in Fig. 5.

6 Implementation and Results of SPST-Equipped Booth Multiplier Using MCLA Adder

To prove the architectural concept after functional validation, the proposed architecture is synthesized. In Xilinx RTL, chip XC6SLX45T3FGG484 spartan 3 series is selected for benchmarking. The tool does a high level of optimization and generates net list.

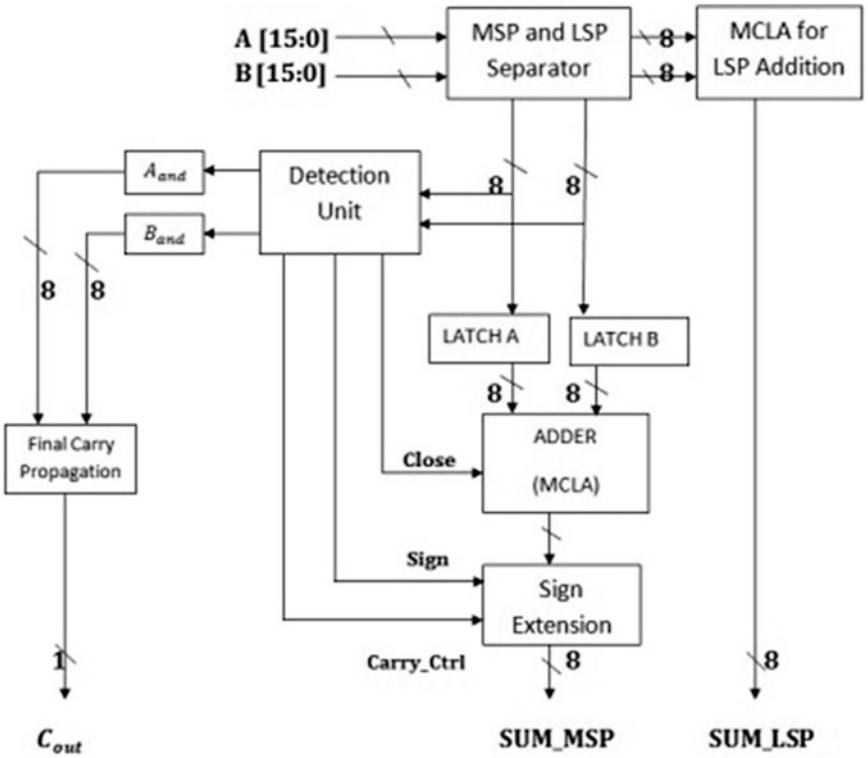


Fig. 5 16-bit SPST adder

Table 2 shows the performance analysis of the proposed multiplier when compared to other multiplier. The timing report shows that to perform 16-bit addition, minimum period required for array multiplier is 21.068 ns and SPST-equipped booth multiplier requires 14.886 ns. The proposed adder requires 13.263 ns. An improvement of 37.04 % is achieved in speed when compared to proposed multiplier with array multiplier, and an improvement of 10.9 % is achieved in speed when compared with SPST-equipped booth multiplier. Figures 6 and 7 show the gate utilization graph of the proposed multiplier. The proposed multiplier is implemented using basic gates like AND, Inverter, and OR. It uses 2 input, 3 input, and 4 input LUT's as shown in Gate Utilization Graph-I.

Table 2 Performance analysis of SPST adder using MCLA

Multiplier type	Array multiplier	SPST-equipped booth multiplier	Proposed multiplier	Percentage improvements	
Delay	21.068 ns	14.886 ns	13.263 ns	1 and 3 37.04	2 and 3 10.9

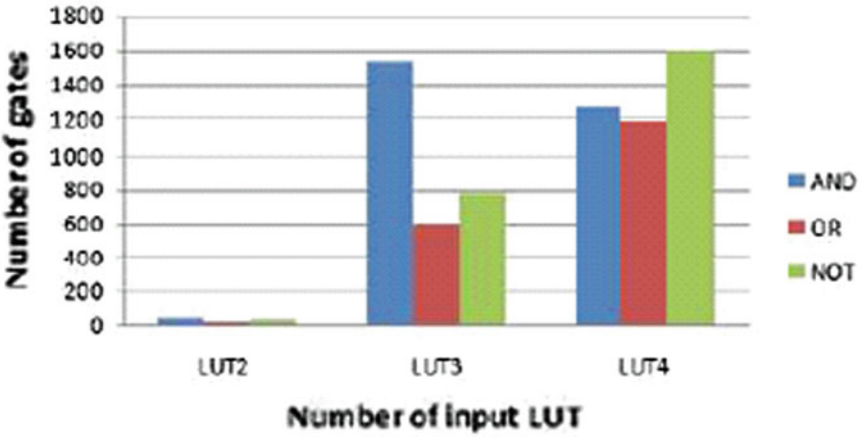


Fig. 6 Gate utilization graph-I

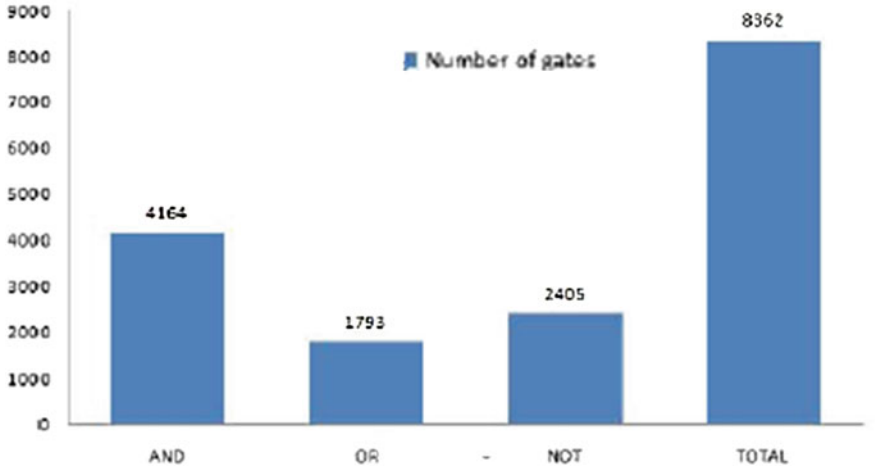


Fig. 7 Gate utilization graph-II

A total number of 8362 gates, of which 4164 are AND gates, 1793 are OR gates, 2405 are NOT gates used to implement the proposed multiplier, are as shown in Gate Utilization graph-II.

7 Conclusion

The proposed architecture is synthesized in Xilinx RTL; chip XC6SLX45T3FGG484 spartan 3 series is selected for benchmarking. The timing report shows that to perform 16-bit multiplication, the minimum period required for array multiplier is 21.068 ns and SPST-equipped booth multiplier requires 14.886 ns. The proposed adder requires 13.263 ns. An improvement of 37.04 % is achieved in speed when compared to array multiplier, and an improvement of 10.9 % is achieved in speed when compared to SPST-equipped booth multiplier.

References

1. Marimuthu CN, Thangaraj P (2008) Low power high performance multiplier. ICGST PDCS, vol 8(1), Dec 2008
2. Benini CL, Micheli GD, Macii A, Macii E, Poncino M, Scarsi R (2000) Glitching power minimization by selective gate freezing. IEEE Trans Very Large Scale Integr Syst 8(3):287–297
3. Praveen Kumar M (2011) A spurious-power suppression technique for multimedia/DSP applications. Int J Adv Eng Sci Technol 11(1):035–051c
4. Chen H, Chao KC, Guo JI, Wang JS, Chu YS (2005) An efficient spurious power suppression technique (SPST) and its applications on MPEG-4 AVC/H.264 transform coding design. In: Proceedings of IEEE international symposium low power electron, pp 155–160, Dec 2005
5. Lakshmi Narayanan G, Venkataramani B (2005) Optimization techniques for FPGA-based wave pipelined DSP blocks, IEEE Trans Very Large Scale Integr Syst 13(7):783–792

Energy Efficient ECC Encryption Using ECDH

Ravi Kishore Kodali and N. V. S. Narasimha Sarma

Abstract Elliptic curve cryptography (ECC) provides a secure means of exchanging keys among communicating hosts using the Diffie–Hellman (DH) key exchange algorithm. This work presents an implementation of ECC encryption making use of the DH key exchange algorithm. Encryption and decryption of text messages have also been attempted. In ECC, we normally start with mapping a character of message to an affine point on the elliptic curve, which is called encoding. A comparison of the proposed algorithm and Koblitz's method shows that the proposed algorithm is as secure as Koblitz's encoding and has less computational complexity due to the elimination of encoding, thereby improving energy efficiency of the crypto-system to be used in resource constrained applications, such as wireless sensor networks (WSNs). It is almost infeasible to attempt a brute force attack. The security strength of the algorithm is proportional to key length. As the key length increases, the data that can be sent at a time also increase.

Keywords ECC · Decoding · Koblitz encoding · EC-DH

1 Introduction

Elliptic curve cryptography (ECC) is ideal for environments such as cellular phones and smart cards [1], wireless sensor networks (WSN's) [2], etc. Moreover, because of the apparent hardness of the underlying elliptic curve discrete logarithm problem (ECDLP) [3] and [4], ECC systems are also well suited for applications requiring security, which need to last longer. ECC offers performance

R. K. Kodali (✉) · N. V. S. Narasimha Sarma
Department of Electronics and Communications Engineering, National Institute of Technology, Warangal, Andhra Pradesh, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_48, © Springer India 2014

advantage at higher security levels [5]. Most of the products and standards that use public key cryptography for encryption make use of the RSA algorithm today. The principal advantage of the ECC compared to the RSA is that it offers better security at reduced key sizes, thereby reducing processing overhead [6, 7]. ECC makes use of elliptic curves (EC's), in which the variables and coefficients are all restricted to elements of a finite field. Each user taking part in public key cryptography uses a pair of keys: a public key and a private key [8]. Only that particular user knows the private key, whereas the public keys are distributed among all the users intending to communicate. Some public key algorithms may require a set of predefined constants to be known by all the devices taking part in the communication. In ECC, these predefined constants are also called domain parameters.

An understanding of the ECC needs mathematical background on EC's [3].

The general cubic equation of EC's is given by the Eq. (1).

$$y^2 + axy + by = x^3 + cx^2 + dx + e \tag{1}$$

But for our purpose, it is sufficient to limit the equation to the form

$$y^2 = x^3 + ax + b, \tag{2}$$

where $4a^3 + 27b^2 = 0$.

Let E_P (a, b) consisting of all the points (x, y) that satisfy the Eq. (2) together with element at infinity O. A group can be defined based on the set E_P (a, b) for specific values of a and b. The heart of ECC is discrete logarithm problem ECDLP that can be stated as it is computationally infeasible to find a value k, such that, $Q = kP$, where P and Q are known points on the curve. But it should be relatively easy to find Q where k and P are known.

The first part of algorithm is to generate public and private keys by both the parties participating in the communication. Both the users should select a random base/generator public point, G, on the curve, whose order is a prime P. Each user generates a random secret integer less than the order of G. Public key of a particular user is the scalar multiplication of his secret integer and the generator point.

Next, phase is sharing a key using Diffie–Hellman algorithm that provides secure key exchange. Finally, the message is hidden into this key and thus the message is encrypted.

2 Algorithm Description

The equation defining an elliptic curve over Galvano field GF(p) [3] is as follows:

$$(y^2 = x^3 + ax + b) \bmod p, \tag{3}$$

where $4a_3 + 27b_2 = 0$, (x, y) ϵ GF (p) and a, b are integers $< p$

An elliptic curve, E over GF (p), consists of the solutions (x, y) defined by Eq. (3), along with an additional element called O, which is the point of EC at infinity. The basic elliptic curve operations are point addition and point doubling. ECC primitives require scalar point multiplication. Say, given a point P (x, y) on an EC, one needs to compute kP, where k is a positive integer. This is achieved by a series of doubling and addition of P.

Example:

If $k = 23$, then $kP = 23 \times P = 2(2(2(2P) + P) + P) + P$. Therefore, we apply doubling and addition depending on a sequence of operations determined for k. Public keys are generated using scalar multiplication of private keys with generator point, which applies series of doubling and addition operations.

2.1 Notations

1. G is Generator point with prime order m [9].
2. Key pair {P, n}, where P is Public key = n * G and n is Private key < m.
3. User Alice is U_{Alice} and User Bob is U_{Bob}.
4. $\oplus$ is X-OR operation.
5. L is the no. of bits in the message to be sent at a time = $[\log_2(\text{message}) + 1]$,

$V^{1} = V \bmod 2^{L}$ (first L bits)

Algorithm 1 ALGORITHM WITH SIMPLE HIDING

1. Key Generation
   - A. Begin
   - B. Initiate the connection between users U_{Alice} and U_{Bob}
   - C. $U_{Alice} = (P_{Alice}, n_{Alice})$ is the key pair for U_{Alice}. $U_{Bob} = (P_{Bob}, n_{Bob})$ is the key pair for U_{Bob}.
   - D. U_{Bob} sends the point P_{Bob} to U_{Alice}. Similarly U_{Alice} sends P_{Alice} to U_{Bob}.
2. Encryption : Hiding
   - E. U_{Alice} computes the point $n_{Alice}(P_{Bob})$. Let this be $S = (x_s, y_s)$. U_{Alice} calculates Cipher $= x_s^{1} \oplus$ Message. U_{Alice} sends cipher to Bob
3. Decryption
   - F. U_{Bob} computes the point $n_{Bob} \times P_{Alice}$ which is same as the point $S = (x_s, y_s)$. Finally U_{Bob} decrypts the message using Message $= x^{1}{}_{s} \oplus$ Cipher.

2.2 Correctness

$$P_{Bob} = n_{Bob} \times G \tag{4}$$

$$P_{Alice} = n_{Alice} \times G \tag{5}$$

$$n_{Alice} \times P_{Bob} = n_{Alice} \times n_{Bob} \times G = n_{Bob} \times P_{Alice} = S \tag{6}$$

$$\text{Cipher} = X_S^1 \oplus \text{Message} \tag{7}$$

$$X_S^1 \oplus \text{Cipher} = X_S^1 \oplus X_S^1 \oplus \text{Message} \tag{8}$$

Hiding message in the point plays an important role in security of algorithm.

3 Implementation of the Proposed Algorithm

Once the defining of EC is known, we can select a base point G. G has (x, y) coordinates that satisfy the Eq. (3). The base point has the smallest x, y value that satisfy the elliptic curve EC. The ECC method requires that we select a random integer k, which needs to be kept secret. Then, kG is evaluated, by a series of addition and doubling operations, as discussed. For the purpose of

---

Algorithm 2 ALGORITHM WITH COMPLEX HIDING

1. Key Generation
   - A. Begin
   - B. Initiate the connection between users U_{Alice} and U_{Bob}
   - C. $U_{Alice} = (P_{Alice}, n_{Alice})$ is the key pair for U_{Alice}. $U_{Bob} = (P_{Bob}, n_{Bob})$ is the key pair for U_{Bob}.
   - D. U_{Bob} sends the point P_{Bob} to U_{Alice}. Similarly, U_{Alice} sends P_{Alice} to U_{Bob}.
2. Encryption : Hiding
   - E. U_{Alice} computes the point $n_{Alice1}(P_{Bob1})$ Let it be $S_1 = (x_{s1}, y_{s1})$ $n_{Alice1}(P_{Bob2})$ and let it be $S_2 = (x_{s2}, y_{s2})$
   - F. U_{Alice} calculates the distance between S_1 and S_2. Let it be D.
   - G. U_{Alice} calculates $V_1 = x_{s1}$ and $V_2 = x_{s2}$. Cipher $= V_1 \oplus V_2 \oplus$ Message.
   - H. U_{Alice} sends cipher to U_{Bob}.
   - I. For the next session, Alice changes the values to $V_1 = V_1 + D$, $V_2 = V_2 + D$.
3. Decryption
   - J. U_{Bob} computes the point and S_2.
   - K. U_{Bob} calculates $V_1 = x_{s1}$ and $V_2 = x_{s2}$.
   - L. Finally, U_{Bob} decrypts the message using Message $= V_1 \oplus V_2 \oplus$ Cipher.
   - M. For the next session, Bob changes the values to $V_1 = V_1 + D$, $V_2 = V_2 + D$.

---

this discussion, we shall call the source as host Alice, and the destination as host Bob. We select the private key of the host Bob, n_{Bob}. Bob's public key is computed using $P_{Bob} = n_{Bob} \times G$. Similarly Alice computes her public key $P_{Alice} = n_{Alice} \times G$ and these two public keys are exchanged and a secret point, S, is computed by both.

3.1 Simple Hiding

Suppose Alice wants to encrypt and transmit a character to Bob, she does the following: Assume that host Alice wants to transmit the character C. Then, the ASCII value of the character C is used to modify the secret point, thereby hiding the message somehow into the secret point. The hiding process is the x-or operation between the first n-bits of the x-coordinate of a secret point, S, and the message to be transmitted, where n is the number of bits in the message. The EC is

$$y^2 \bmod 487 = (x^3 - 5x + 25) \bmod 487 \tag{9}$$

The base point, G, is selected as (2, 5). The base point implies that it has the smallest x, y co-ordinates satisfying the EC equation (9). The order of G is $m = 163$. $n_{Alice} = 62$ and $n_{Bob} = 71$. $P_{Bob} = 71(2, 5) = (302, 57)$. The secret point is $S = n_{Alice}\ P_{Bob} = 62(302, 57) = (11, 422)$. L = 8-bits in message. Therefore, $X_s^1 = 11 = 00001011$.

Encryption: Plaintext is C, whose ASCII value is 67.

Therefore, message = 67 = 01000011.

Ci p h e r = $X_s^1 \oplus$ Message = 01001000 = 72 = H

Decryption: Cipher = H = 01001000.

Message = $X_s^1 \oplus$ Cipher = 01000011 = C.

Table 1 illustrates the method for a series of characters "SEND".

3.2 Complex Hiding

The Elliptic Curve is

$$y^2 \bmod 487 = (x^3 - 5x + 25) \bmod 487 \tag{10}$$

G = (2, 5). Suppose $n_{Alice1} = 63$, $n_{Alice2} = 93$, $n_{Bob1} = 71$, $n_{Bob2} = 53$. Therefore, $P_{Alice1} = 63 \times G = (139, 347)$, $P_{Alice2} = 93 \times G = (486, 121)$,

$P_{Bob1} = 71 \times G = (302, 57)$ and $P_{Bob2} = 53 \times G = (176, 76)$. According to the algorithm, $S_1 = (x_{s1}, y_{s1}) = (11, 422)$ and $S_2 = (x_{s2}, y_{s2}) = (152, 126)$.

Distance between these two points is calculated as

$$D = ((152 - 11)^2 + (422 - 126)^2) = 327 = 101000111D^1 = 71 = 01000111$$

Table 1 Encryption with simple hiding

Message and its ASCII	Cipher and its ASCII
S(83)	X (88)
E(69)	S(78)
N(78)	S(69)
D(68)	S(79)

Table 2 Encryption with complex hiding

Message and its ASCII	Cipher (8-bit)
S(83)	192
E(69)	200
N (78)	241
D(68)	233

$$V_1 = x_{s1} = 11 = 00001011 V_2 = x_{s2} = 152 = 10011000$$

Encryption: Plaintext is C, whose ASCII value is 67.
Therefore, Message = 67 = 01000011

$$\text{Cipher} = V_{11} \oplus V_{21} \oplus \text{Message} = 11010000 = 208$$

Decryption: Cipher = 11010000

$$\text{Message} = V_{11} \oplus V_{21} \oplus \text{Cipher} = 01000011 = \text{C}$$

For the next session, $V_1 = V_1 + D = 11 + 71 = 82 = 01010010$

$$V_2 = V_2 + D = 152 + 71 = 223 = 11011111.$$

A sample sequence of characters "SEND" is to be sent. Table 2 illustrates the method for a series of characters "SEND". Each character encryption requires two 8-bit addition and two 8-bit X-OR operations. The time elapsed for encryption of single character or 8-bit data, when implemented in MATLAB, is 0.000008 s. Therefore, the CPU time elapsed for encryption of the message SEND 4 * (0.000008) = 0.000032 s.

The complex hiding method of encryption typically consumes more power when compared with simple hiding method, but less than any of the other ECC encryption techniques, which generally use encoding process. An example of such algorithm is ECC encryption with Koblitz method of encoding [10]. Complex hiding of message provides the algorithm with efficient security. It is almost infeasible to attempt brute force attack. Since encryption involves XOR operations, therefore it consumes less computational time and power with efficient security.

4 Comparison with Koblitz's Algorithm

4.1 Time and Power Consumption

ECC encryption with Koblitz's method of encoding is one of the best encryption algorithms that provide reliable security. Such algorithms consume additional power for encoding and encryption, whereas our process does not consume so much power. The plot given in the Fig. 1 shows the time consumed for encryption for various lengths of primes.

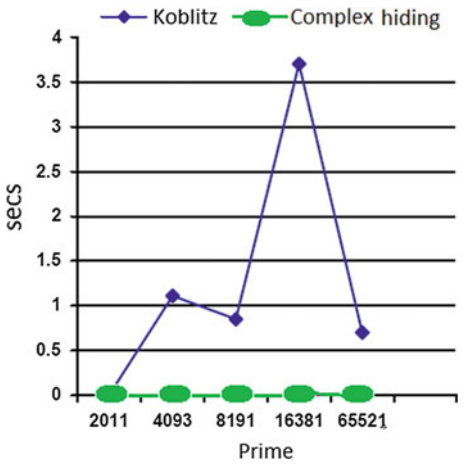


Fig. 1 Comparison between Koblitz's encoding [10] and complex hiding algorithm

4.2 Security

Though simple hiding is less secure when compared with complex process but for appropriate prime lengths, the breaking process with brute force attack consumes hundreds of years for completion. It is almost infeasible to attempt brute force attack on complex hiding process. Breaking of this algorithm requires the knowledge of three parameters V_1, V_2, and D.

Assumption of either of the V_1 and V_2 values requires breaking of ECDLP, which is almost infeasible for higher lengths keys. Since D is a step function, tracing points back from D(if known) is also a difficult task. Also, the change in values V_1 and V_2 after each session adds more security for the message. Realizing these above parameters from any of the known ciphers is impossible.

5 Conclusions

A plaintext message SEND is taken for implementing the algorithm proposed in this paper. Each character in the message is represented by its ASCII value, and then, encryption is carried out using both the algorithms. The execution time for encoding and decoding functions is no more required. The execution time taken for encryption is constant for different values of a, b, P for particular key length. Range of message bits that can be sent per single cipher is equal to the key length. As the key length increases, the security also increases exponentially due to the complexity of ECDLP. Since this algorithm consumes less energy for encryption, it can be used in WSN's. As a result, the life of each sensor node can be extended in addition to providing security.

References

1. Lauter K (2004) The advantages of elliptic curve cryptography for wireless security. IEEE Wirel Commun 11(1):62–67
2. Tian X, Wong DS, Zhu RW (2005) Analysis and improvement of an authenticated key exchange protocol for sensor networks. IEEE Commun Lett 9(11):970–972
3. Hankerson D, Menezes AJ, Vanstone S (2004) Guide to elliptic curve cryptography. Springer, New York
4. Stallings W (2003) Network and internetwork security: principles and practice. Prentice-Hall, Inc., Englewood Cliffs
5. Hasan MA (2001) Power analysis attacks and algorithmic approaches to their countermeasures for koblitz curve cryptosystems. IEEE Trans Comput 50(10):1071–1083
6. Qiu Q, Xiong Q (2004) Research on elliptic curve cryptography. In: Proceedings of the 8th international conference on computer supported cooperative work in design, vol 2, pp 698–701
7. Khabbazian M, Gulliver TA, Bhargava VK (2007) Double point compression with applications to speeding up random point multiplication. IEEE Trans Comput 56(3):305–313
8. Longa P, Miri A (2008) Fast and flexible elliptic curve point arithmetic over prime fields. IEEE Trans Comput 57(3):289–302
9. Ramasamy RR, Prabakar MA, Devi MI, Suguna M (2009) Knapsack based ECC encryption and decryption. Int J Netw Secur 9(3):218–226
10. Padma Bh, Chandravathi D, Roja PP (2010) Encoding and decoding of a message in the implementation of elliptic curve cryptography using Koblitzs method. Int J Comput Sci Eng 2:1904–1907

Adaptive Support Weight-Based Stereo Correspondence Algorithm for Face Images

C. J. Prabhakar and K. Jyothi

Abstract The stereo correspondences of human faces are often very difficult to achieve because of uniform texture, slow changes in depth, and occlusion. In this paper, we introduce an adaptive weight-based stereo correspondence method for face images. We estimate the support weights of the pixels in a given support window based on a color similarity and proximity to reduce the fattening effect. Further, the reference image is segmented using mean-shift segmentation method, and then, self-adaptability measure is used to estimate the initial disparity using correlation-based SSD method. Dissimilarity at a given pixel is then computed using the initial disparity and support weights of both support windows. Finally, the correspondence selected by the winner takes all the method. The experiments are carried out on stereo images of face database and Middlebury database. The experimental results show that the proposed algorithm produces a smooth disparity map while preserving sharp depth discontinuities accurately.

Keywords Initial disparity · Stereo matching · Adaptive support weight

1 Introduction

Stereo reconstruction is one of the most extensively research topics in computer vision for estimation of depth from digital images. It has two steps—correspondence matching and depth estimation. Stereo correspondence is a process of finding reliable matches and extracting dense map. Stereo reconstruction has

C. J. Prabhakar (✉)
Department of Studies and Research in Computer Science, Kuvempu University, Shankaraghatta, Karnataka, India
e-mail: psajjan@yahoo.com

K. Jyothi
Department of IS&E, J.N.N College of Engineering, Shimoga, Karnataka, India
e-mail: jyothi6575@rediffmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_49,

potential uses in many applications such as autonomous navigation, 3D reconstruction, and object recognition. Scharstein and Szeliski [1] have provided a survey of stereo correspondence algorithms, and taxonomy was given based on matching cost, aggregation, and optimization. In general, two main categories can be distinguished: local algorithms and global algorithms. The local algorithms estimate the disparity value at a given point only based on some correlation of color or intensity values within a support region around the point. Local-based methods assume all pixels in the support window have similar depth. This assumption is violated in depth discontinuities, and the best correlation matches for each pixel. This sometime resulted in wrong disparity estimation for image points having ambiguity of depth. So that texture less regions, repetitive textured regions, and regions with depth discontinuity fail to match correctly and this results in the "foreground-fattening" phenomenon. To obtain accurate results at depth discontinuities and also in homogeneous regions, an appropriate support weight should be assigned for each pixel adaptively. Some of the methods [2–6] will try to assign appropriate support weights to the pixels in a support window while fixing the shape and size of a local support window. Xu et al. [2] also presented an algorithm that estimates adaptive support weights by radial computation techniques. They used the certainties of the initial disparity distribution to determine support weights. These methods, however, are dependent on the initial disparity estimation, which may be erroneous. Yoon and Kweon [6] presented a new cost aggregation scheme that uses a fixed-size support window with per pixel varying weight. The weight is computed based on the color similarity and geometric distance to the center pixel of interest. Very strong results were obtained without any global optimization. Conventional adaptive-window cost aggregation techniques focus on varying the size, shape, and position of the support window, whereas the adaptive weight method [6] uses a large fixed-size support window and assigns a support weight to each pixel in the window.

In this paper, we introduce stereo correspondence method for face images to get accurate results at depth discontinuities as well as in homogeneous regions. We compute the support weights of the pixels in a given support window using color similarity and spatial proximity. The proposed method is composed of three parts: adaptive support weight computation, dissimilarity computation based on the support weights, and finally, disparity selection using winner takes all algorithm.

The remaining sections of the paper are organized as follows. The proposed adaptive weight-based stereo matching method is discussed in sect. 2. The experimental results are presented in section sect. 3. Finally, the conclusion is given in section sect. 4.

2 Stereo Matching

In the proposed stereo matching technique, stereo image pair is input to the system. The images are first rectified using uncalibrated rectification method [7] to suppress the vertical displacement. The support weights of the pixels are

estimated in a given support window based on a color similarity and proximity to reduce the fattening effect. The reference image is segmented using mean-shift segmentation method [8]. The self-adapting dissimilarity measure [9] is used to estimate the initial disparity, and then, dissimilarity is computed using the raw matching costs and support weights of both support windows.

2.1 Adaptive Support Weight Estimation

The difference between pixel colors is measured in the CIELab color space, mainly because it strongly correlates with human color discrimination and has perceptually meaningful measure of color similarity [6]. The proximity weights are assigned to the pixels based on their color difference and proximity. When Δc_{pq} represents the Euclidean distance between two colors, $c_p = [L_p, a_p, b_p]$ and $c_q = [L_q, a_q, b_q]$ in the CIELab color space. The color similarity of p and q is defined as

$$f_s(\Delta C_{\mathrm{pq}}) = \exp\left(-\Delta C_{\mathrm{pq}}/\gamma_c\right). \tag{1}$$

The spatial proximity is the Euclidean distance between p and q is expressed as

$$f_p(\Delta g_{\mathrm{pq}}) = \exp\left(-\Delta g_{\mathrm{pq}}/\gamma_p\right). \tag{2}$$

The support weight of a pixel is expressed as

$$w(p,q) = \exp\left(-\left(\frac{\Delta C_{\mathrm{pq}}}{\gamma_c} + \frac{\Delta g_{\mathrm{pq}}}{\gamma_p}\right)\right), \tag{3}$$

where γ_c and γ_p are user-defined parameter.

2.2 Estimation of Initial Disparity

To estimate the initial disparity of reference image, the rectified reference image is first divided into regions of homogeneous color using mean-shift segmentation [8]. Then, disparity map is estimated by considering sum of square difference between intensity and gradient of each pixel of stereo pairs. The dissimilarity measure based on the gradient fields and intensities of the images yields good accuracy in the resulting disparity map when compared to classical intensity-based method. The main reason for this accuracy enhancement is the better robustness of the gradient against differences in sampling and local brightness changes between cameras. That is defined as

$$SSD(q, q_d) = \sum\nolimits_{q,q_d \in W(x,y)} (b_1(q) - b_2(q_d))^2 \tag{4}$$

$$Grad\ (q, q_d) = \sum_{q,q_d \in W_x(x,y)} |\nabla_x b_1(q) - \nabla_x b_2(q_d)|^2 + \sum_{q,q_d \in W_y(x,y)} \left|\nabla_y b_1(q) - \nabla_y b_2(q_d)\right|.^2, \tag{5}$$

where W is 11 $\times$ 11 window. The left to right cross-checking is used for occlusion handling and to find reliable correspondences. The resulting dissimilarity measure is given by

$$C(q, q_d) = \mathrm{SSD}(q, q_d) + \mathrm{Grad}(q, q_d). \tag{6}$$

2.3 Adaptive Support Weight Stereo Matching

Adaptive support weight window-based stereo matching applies weights to each of the pixels. The dissimilarity between pixels is then measured by aggregating the raw matching costs with the support weights of both the reference and target support windows. Here, we consider support weight of reference window and target window to minimize errors in the dissimilarity measure. The disparity between pixel q and q_d can be expressed as

$$D(q, q_d) = \sum w(p, q) w(p\prime, q_d) C(q, q_d). \tag{7}$$

After the dissimilarity computation, winner takes all optimization, which is used to choose the disparity with lowest matching cost. In our method, instead of assigning disparity value to each pixel, the disparity plane is estimated by using disparities of each segment. The estimated disparity plane is assigned to each segment using least-square plane fitting method.

3 Experimental Results

In order to evaluate the performance of the proposed stereo correspondence algorithm, we have carried out experiments using the test images of stereo face database [10] and Middlebury stereo database [11]. The proposed algorithm is evaluated using a root-mean-square (RMS) error. The RMS error of the proposed method is compared with standard correlation-based self-adaptive dissimilarity algorithm (CSAD).

3.1 Experiments on Stereo Face Images

To evaluate the performance of the proposed stereo correspondence algorithm, we have carried out experiments on test images of stereo face database [10]. It consists of stereo pairs of 70 subjects (35 males, 35 females) of size 676 × 516, recorded from six different viewpoints. The next five viewpoints range from a frontal to a profile view with respect to the viewing direction of the first camera, in equal steps of 1/8 radians. The intensities on the facial part of the image vary considerably over the different viewpoints.

The stereo images are first rectified using uncalibrated rectification [7] to suppress the vertical displacement. The rectified reference image is segmented into homogeneous region using mean-shift segmentation method. The result obtained after applying mean-shift color segmentation method to the reference is shown in Fig. 1a. The result of our method on face images is compared with the result of well-known CSAD [9]. The disparity map obtained from our proposed method is shown in Fig. 1b, and the Fig. 1c shows the disparity map produced by CSAD. The results show that fattening effect in disparity map obtained by our algorithm is lesser than the result of segment-based self-adaptive dissimilarity algorithm.

3.2 Experiments on Stereo Images of Middlebury Database

To validate the result of our algorithm with ground truth, experiments are performed on some of the stereo test images from the Middlebury stereo database [11]. We have used stereo images viz., Teddy and Cones (First Column of Fig. 2) of Middlebury database. The second and third column of Fig. 2 show the corresponding image ground truth and result of proposed method.

RMS method is a quantitative way to estimate the quality of the computed correspondences. The RMS error is calculated between the computed depth map $d_C(x, y)$ and the ground truth map $d_T(x, y)$, that is,

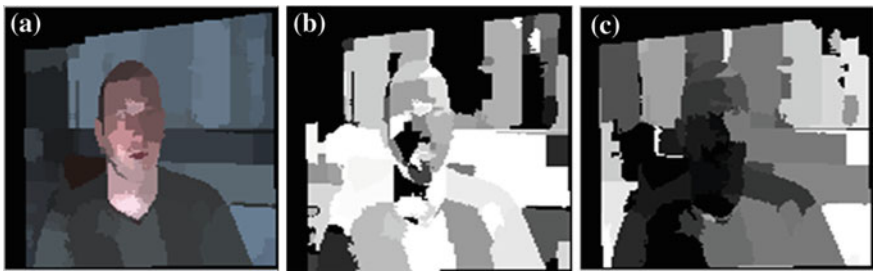


Fig. 1 **a** Homogeneous region obtained by color segmentation. **b** Disparity map of proposed method. **c** CSAD method

Fig. 2 Results on Middlebury datasets. From *top to bottom* teddy, cones. From *left to right* reference images, ground truth disparities, the results of the proposed algorithm

Table 1 Root-mean-square errors

Methods	Teddy	Cones	Aloe
Proposed method	0.0539	0.0529	0.0419
CSAD	0.0557	0.0552	0.0422

$$R = \left(\frac{1}{N} \sum_{(x,y)} |d_C(x,y) - d_T(x,y)|^2 \right)^{\frac{1}{2}}, \tag{8}$$

The Table 1 shows the RMS error estimated for results obtained by proposed method and CSAD algorithm. The RMS error of CSAD algorithm is more than RMS error of proposed method.

4 Conclusion

This paper introduces a novel area-based stereo correspondence algorithm based on adaptive support weight strategy, which deploys segmentation information in order to increase the reliability of the matches. The support weights of the pixels are calculated using a color similarity and spatial proximity. The reference image

is segmented using mean-shift segmentation method, and then, self-adaptability measure is used to estimate the initial disparity using correlation-based SSD method. The dissimilarity is then computed using the raw matching costs and estimated support weights of both support windows. The RMS error of results obtained by our proposed method is less compared to CSAD. The experimental results show that the proposed technique produces a smooth disparity map with reduced fattening effects.

References

1. Scharstein D, Szeliski R (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int J Comput Vis 7(1–3):7–42
2. Xu Y, Wang D, Feng T, Shum HY (2002) Stereo computation using radial adaptive windows. In: Proceedings of international conference pattern recognition, vol 3, pp 595–598
3. Yoon KJ, Kweon IS (2006) Adaptive support-weight approach for correspondence search. IEEE Trans PAMIl 28(4):650–656
4. Hosni A, Bleyer M, Gelautz M, Rhemann C (2009) Local stereo matching using geodesic support weights. In: 16th IEEE international conference on image processing, Nov 2009, pp 2093–2096
5. Gu Z, Su X, Liu Y, Zhang Q (2008) Local stereo matching with adaptive support-weight, rank transform and disparity calibration. Pattern Recogn Lett 29(9):1230–1235
6. Yoon KJ, Kweon IS (2005) Locally adaptive support-weight approach for visual correspondence search. In: Proceedings of conference on computer vision and pattern recognition (CVPR), pp 924–931
7. Fusiello A, Irsara L (2008) Quasi-euclidean uncalibrated Epipolar rectification, ICPR, pp 1–4
8. Comaniciu D, Meer P (2002) Mean shift a robust approach toward feature space analysis. IEEE PAMI 24:603–619
9. Klaus A, Sormann M, Karner K (2006) Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: Proceedings of the IEEE international conference on pattern recognition (ICPR'06), vol 3, pp 15–18
10. http://cvlab.epfl.ch/data/strecham VS/2008
11. Middlebury database. http://vision.middlebury.edu/stereo/data/, 2010

Adaptive Traffic Load Sharing for GSM Network

K. R. Sudhindra and V. Sridhar

Abstract TCH (traffic channel) congestion is one of the key performance indicators (KPI), which influence the network performance and customer satisfaction in GSM (global system for mobile communication) network. The increase in TCH congestion in the network results in large number of TCH blocking which greatly affects the subscriber satisfaction and revenue of the service provider. The available congestion relief methodologies such as cell splitting, aggressive frequency reuse (AFR) pattern, microcells, and expanding frequency band require time and investment for implementation. An adaptive traffic load sharing model for GSM network is proposed which tunes the radio parameters such as receive signal level access minimum (RXLEV_ACCESS_MIN), cell reselection offset (CRO), and antenna tilt for balancing the traffic between congested and its non-congested neighboring cells. The adaptive tuning of these parameters based on traffic condition would help to reduce the turnaround time in resolving congestion issues.

Keywords GSM · TCH congestion · Traffic balancing

1 Introduction

GSM is the most widely accepted cellular technology in the world. Once the network is designed and made operational, the performance is monitored continuously and compared with key performance indicators (KPI). Traffic channel

K. R. Sudhindra (✉)
Idea Cellular Ltd, Bangalore, India
e-mail: sudhindra_kr@rediffmail.com

V. Sridhar
P.E.S. College of Engineering, Mandya, India
e-mail: venusridhar@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_50,

(TCH) congestion is one of the KPI which affects the performance of the live GSM network. The increase in subscriber base and attractive marketing strategies will result in high TCH congestion and demands proportional capacity expansions. There are many capacity expansion solutions available in the literature most of which requires additional expenditure.

The cell splitting, AFR, microcell, and expanding band require an additional investment on the network and are beneficial only if increase in traffic remain constant. In the live network, the traffic depends on subscriber's behavior and can be classified as user traffic behaviors and user movement behaviors. The fluctuation factor (generally 1.05–1.1) is used to measure the degree of traffic deviation on the network and has to be considered during capacity expansion. Channel configuration is calculated by frequency amount rather than linear configuration based on requirements, which leads to channel wastage. For example, if only 9 TCHs are required for a cell after the calculation, 2 transceivers (TRX) must be configured. According to the research result, the nonlinear configuration reduces the network utilization by 20–25 %, with 20 % for provincial capitals, 25 % for common cities, and 30 % for backward districts with many single carrier cells. Service provider has to take appropriate decision during the transition state of network traffic from lower to higher. In order to avoid additional investment of TRX or base stations, an adaptive traffic load sharing model to control TCH congestion is proposed. The model depends on the parameters such as RXLEV_ACCESS_MIN, CRO, and antenna tilt which are tuned dynamically to balance the traffic between congested cell and its non-congested nearest cell.

2 Literature Survey

Continuous monitoring of TCH congestion and its optimization will reduce the number of call blockings in the network and improves subscriber satisfaction. The common TCH congestion relief methods found in literature are cell splitting, AFR, microcells, and expansion of frequency band, half rate, and assignment to another cell [1].

2.1 Cell Splitting

During the initial phase of GSM network construction, the coverage was given importance with large antenna height, large inter site distance, and cell radius. With the increase in subscriber base, the original cells can be split into smaller cells with lesser coverage area. The radius of the cell coverage can be reduced by minimizing the space between stations and lowering antenna height or increasing down tilt. The cell splitting method creates a space for adding additional base stations and frequencies which helps in expanding network capacity. The space

between macrocellular base stations should be at least 400 m. In medium cities, the height is recommended to be 25 m. If cell splitting is performed, a new coverage and frequency plan will be required to ensure inter cell interference is kept to a minimum [2]. The cost of installing new hardware (Transceiver and BTS) and devising new frequency and coverage plan is considerable. As per the results from the live network, the ratio of traffic in the busiest hour to the quietest hour is almost equal to 20:1 [3]. Therefore, it is vital for an operator to ensure that their resources (both hardware and spectrum) are utilized to their full potential.

2.2 Aggressive Frequency Reuse Pattern

Common AFR patterns are multiple reuse pattern (MRP), 1×3 (1×1) multiplexing technology and concentric cell. A solution for providing high capacity in GSM using existing macrocells was discussed in [4] by using MRP in combination with frequency hopping. The use of frequency hopping will result in frequency diversity and interference diversity. The frequency diversity balances the quality between slow and fast fading and an approximate coverage gain equal to fast fading margin can be achieved. Interference diversity shares the interference between different users often referred as interference averaging and helps in improving receiver performance [5]. Field trails in live GSM network indicate that it is possible to implement an average frequency reuse of 7.5 without degrading the required network quality [6].

2.3 Microcells

Microcells are realized as an isolated single cell rather than clusters to accommodate teletraffic hot spots. The microcellular base-station antennas are usually located at some 6–9 m height and attached to walls or mounted on the roofs of low buildings. Macrocellular sectors will oversail the microcellular clusters, providing the coverage in microcellular radio dead spots and assisting handovers when candidate microcells have no free channels. There are two approaches for assigning frequencies to the microcells; reuse the same frequencies between microcells and the macrocells or partition the available frequency band such that macrocells and microcells have their own unique set of frequencies. The spectral efficiency of both frequency-shared and frequency-partitioned arrangements for microcells under different cluster size was evaluated by Robincombs et al. [7] who found that partitioned arrangement was found to be two to three times more efficient. A 10-30 percent of capacity improvement has been shown in [8] by deployment of microcells in hot spots.

2.4 Expansion of Frequency Band

The GSM 900 MHz band is apparently insufficient to meet the capacity requirements of heavy-traffic area in spite of using capacity expansion techniques such as carrier expansion, cell splitting, and AFR. The use of dual band network will be an alternative solution for capacity expansion. The GSM900 network and the GSM1800 network are the same in network structure, voice coding, modulation technology, and signaling procedures, but are different in frequency range, radio propagation feature, feeder loss, and coverage range. The frequency resource of 1,800 MHz is abundant than 900 MHz which reduces the capacity pressure effectively. GSM 900 MHz can be configured to guarantee the network function (coverage) and 1800 MHz is used for extracting actual traffic. The dual band network of the same manufacturer can use the co-BCCH (Broadcast control Channel) technology to save one CCCH (Common Control Channel). The radio parameters which define the relationship between the GSM900 system and the GSM1800 system (for example, networking mode, traffic flow direction, handover relationship, and priority need to be properly adjusted in order to improve the dual band network quality and capacity.

2.5 Half Rate

TCH in GSM can carry user speech or data which can be either in full-rate (TCH/F-13 kbits/s) or half rate (TCH/H-5.6kbits/s) channel mode. When TCH/H is in use, one timeslot may be shared by two connections, thus doubling the number of connections that can be potentially be handled by a TRX, and at the same time, the interference generated in the system would be halved for the same number of connections. TCH/H channels are interleaved in 4 bursts instead of 8 as the TCH/F, but the distance between consecutive bursts is 16 burst periods instead of 8 [8]. GSM half rate (HR) offers enhanced capacity over the air interface corresponding to the proportion of mobiles within a coverage area that support half rate technology. The use of HR has an associated degradation in speech quality since HR speech codec has a worse error-free MOS (mean opinion score). The market perception of degradation in speech quality by HR codec was the reason for less deployment HR codec in the GSM network.

2.6 Cell Load Share

M/s Ericson has come up with an optional solution called cell load sharing (CLS) feature in its base station controller software which is designed to redistribute traffic load between neighboring cells. CLS offers a possibility for offloading of “highly loaded” cell before congestion. This is done by triggering congestion-based

handovers to neighboring cells (if the neighboring cell allows incoming CLS handovers) when the channel utilization within the cell reaches a certain definable percentage. Traffic distribution density is never uniform. The most effective way to successfully use congestion relief features like CLS is to identify high congestion density sites within a BSC (Base Station Controller) and if there are uncongested and reliable (good RF transition at the cell boundary) first-tier neighbors in the near vicinity, then make use of it. If the transition boundary between the source and target cells is weak, then excessive usage of CLS results in increased no of dropped calls and "ping pong" handovers. The efficiency limits are defined by setting range of key parameters, actual network capacity and size and number of congested and uncongested cells [9].

3 Methodology

An adaptive traffic load sharing model to control TCH congestion is designed which works on the radio parameters such as RXLEV_ACCESS_MIN, CRO and antenna tilt which primarily balances the traffic and maintain congestion levels of the network within the desired limit [10, 11].

The block diagram of the proposed model is shown in Fig. 1. The model proceeds by measuring the TCH congestion rate of the GSM network with the current traffic load and compares it with the desired targeted value (benchmark value of TCH congestion rate is <0.5 %). The differential value between the current and target performance is fed to the congestion control module, which is central unit for suggesting improved parameter value. In an event where congestion rate is greater than desired value, then measures are suggested for reducing the variance by the module. If the current congestion values are better than the desired values, then congestion control module exits without taking any action. This loop is reiterated until the optimal value of TCH congestion rate is met.

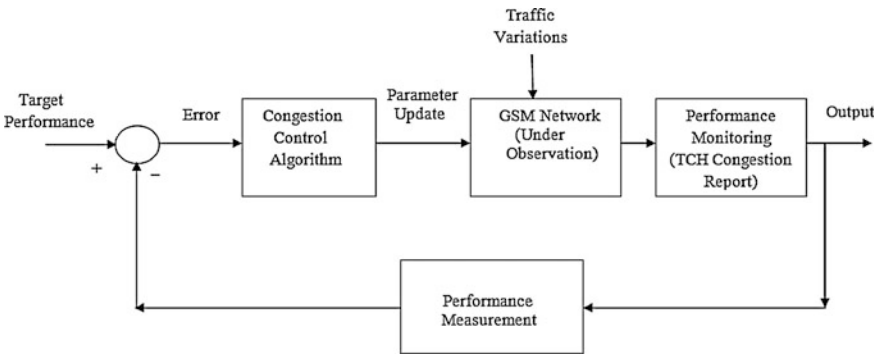


Fig. 1 Block diagram of proposed model of congestion relief

Table 1 Description of simulation

Test case	Description
Proposed adaptive model to control congestion where load balancing between BS3 and BS1 is simulated.	*BS1* RX_Lev_ACCMIN is varied between −95 and −104 dBm CRO is varied between 0 and 20 dBm, and antenna tilt is varied between 7° and 4° (decreasing tilt) *BS3* RX_Lev_ACCMIN is varied between −95 and −86 dBm and antenna tilt is varied between 7° and 10° (increasing tilt) *BS2, BS4, BS5, BS6, BS7* keeping default values of RX_Lev_ACCMIN (-95 dBm), CRO (0 dB), and antenna tilt (7°)

3.1 Simulation Environment

GSM scenario is created using Qualnet 5.0 simulator with total number of "7" base stations (BS1 to BS7) of omni directional cells and "61" mobile stations. Both BSs and MSs are randomly spread across 3x3 sqKm area on digital synthetic terrain. Total number of channels used is "14" and all the BSs are connected to a single MSC (Mobile Switching Center). Simulation is conducted on Pentium IV computer with Microsoft Windows XP Professional version 2002.

The proposed model of congestion relief rely on values of RXLEV_ACCESS_MIN, CRO and antenna tilt parameters to balance the traffic between congested and its non-congested neighboring cells. The description of the simulation model is shown in Table 1.

4 Results and Discussion

Simulation result of proposed model is shown in Fig. 2, where x-axis represents simulation time in seconds and y-axis represents the total number of MSs latched across different BSs. Closed loop adaptive congestion control mechanism is simulated where RX_Lev_ACCMIN, CRO, and antenna tilt are varied sequentially till it reaches defined threshold specified in Table 1. The simulation considers BSs as congested when number of MSs latched to it is more than 30 else considers as non-congested. The congestion control algorithm tunes the selected parameters of congested BS till defined threshold values are reached or the number of MSs reduces to 50 % of its initial value.

From simulation result, it is found that BS-3 is congested during beginning of the simulation whose traffic is offloaded to the neighboring base station BS-1. The offloading of traffic has happened through tuning of selected parameters at different time as shown in Table 2. At simulation time of 480 s, the number of MSs latched to BS-3 reduces to 15 after which parameter tuning stops resulting in

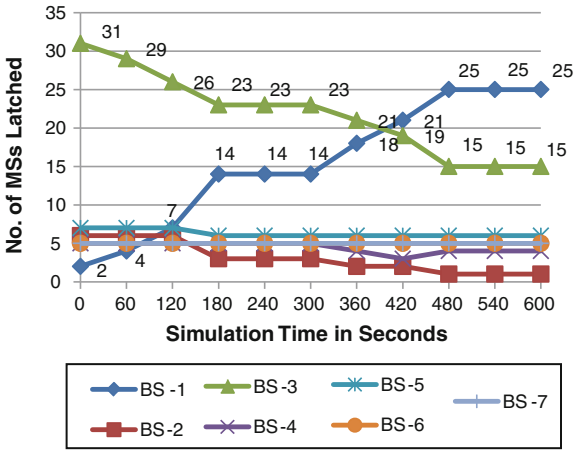


Fig. 2 Simulation result of proposed model

Table 2 Simulation results in tabular form

	RX_Lev_ACCMIN in dBm		CRO		Antenna tilt in degrees		No. of MSs latched						
Time in secs	BS1	B S3	BS1	BS3	BS1	BS3	BS1	BS2	BS3	BS4	BS5	BS6	BS7
0	−95	−95	0	0	7	7	2	6	31	5	7	5	5
60	−98	−92	0	0	7	7	4	6	29	5	7	5	5
120	−101	−89	0	0	7	7	7	6	26	5	7	5	5
180	−104	−86	0	0	7	7	14	3	23	5	6	5	5
240	−104	−86	10	0	7	7	14	3	23	5	6	5	5
300	−104	−86	20	0	7	7	14	3	23	5	6	5	5
360	−104	−86	20	0	8	6	18	2	21	4	6	5	5
420	−104	−86	20	0	9	5	21	2	19	3	6	5	5
480	−104	−86	20	0	10	4	25	1	15	4	6	5	5
540	−104	−86	20	0	10	4	25	1	15	4	6	5	5
600	−104	−86	20	0	10	4	25	1	15	4	6	5	5

unchanged status as indicated by horizontal lines. The threshold values of parameters are chosen based on results discussed in [12].

5 Conclusion

There are radio parameters defined in GSM tuning of which could beat congestion in cost-effective manner. However, lack of experts and case studies on handling radio parameters makes operator to keep default values in order to avoid any adverse effect on entire network. We have identified the key radio parameters

which contribute for traffic balancing and proposed a novel method of adaptive traffic load sharing for GSM network. The simulation results reveal that RX_Lev_ACCMIN, CRO, and antenna tilt could be sequentially tuned for off-loading traffic from congested cell to its non-congested cell. The optimum performance of the model depends upon the threshold settings of the chosen parameters in the algorithm. Excessive study on these parameter settings is required before it is practically implemented. Practical implementation of this model would certainly reduce the turnaround time for resolving momentary congestion issues.

Acknowledgments The authors would like to thank Nihon Communication Solutions (P) Ltd to have made possible to set up simulation using Qualnet 5.0.

References

1. Nasrin W, Rajib MMI (2009) An analytical approach to enhance the capacity of GSM frequency hopping networks with intelligent underlay-overlay. J Commun 4(6):414–422
2. Spilling AG, Nix AR (2000) Performance enhancement in cellular networks with dynamic cell sizing. IEEE PIMRC 2:1589–1593
3. Lam D, Cox DC, Widom J (1997) Teletraffic modeling for personal communication services. IEEE Commun Mag 35(2):79–87
4. Olofsson H et al (1995) Interference diversity as means for increased capacity in GSM. In: Proceedings of 1st EPMCC, Italy, pp 97–102
5. Kolonits A (1997) Evaluating the potential of multiple re-use patterns for optimizing existing network capacity. IIR maximizing capacity workshop, June 1997, London
6. Engstrom S, Johansson T et al (1998) Multiple reuse patterns for frequency planning in GSM networks. IEEE Commun Mag 3:2004–2008
7. Steele R, Williams J, Chandler D, Dehghan S, Collard A (1995) Teletraffic performance of GSM900/DCS1800 in street microcells. IEEE Commun Mag 33:102–108
8. Mullner R, Ball CF et al (2005) Dynamic half-rate allocation for adaptive multi-rate speech codecs in GERAN radio networks. In: Proceedings of 14th Ist Mobile and Wireless Communications Summit, Dresdon
9. Novak V (2002) Cell load sharing feature and traffic optimization in GSM network. Telekomunikacijie 47:94–98
10. ETSI, TS/SMG-020508QR, V5.1.0 (1996) GSM technical specification: digital cellular telecommunication system (Phase 2+); Radio subsytem link control (GSM 05.08)
11. http://docs.commscope.com/Public/electrical_and_mechanical_downtilt_effect_on_pattern_performance.pdf
12. Sudhindra KR, Sridhar V (2012) Congestion relief for GSM using parameter tuning. International conference on computer science and application, Cochin, ICCA 2012

Psychoacoustic Model-1 Implementation for MPEG Audio Encoder Using Wavelet Packet Decomposition

B. K. Jagadeesh and B. Siva Kumar

Abstract Audio compression is the lossy compression technique of converting audio signal into an efficiently encoded representation that can later be decoded to produce a close approximation of the original signal. In MPEG 1, psychoacoustic model plays a crucial role in audio compression. In audio compression technology, compression will be achieved by exploiting the auditory masking characteristic of human ear. By applying psychoacoustic principles, it is possible to analyze the signal and computation of noise. Here, signal masking is a function of frequency. In this work, we are investigating implementation of psychoacoustics model-1 for MPEG-1 using wavelet packet decomposition to compress high-quality audio signal. Most psychoacoustic models for coding applications use a uniform spectral decomposition to approximate the frequency selectivity of the human auditory system. For implementing this algorithm, a design of psychoacoustic model was developed following the model used in the standard MPEG-1 audio. In this work, we have proposed psychoacoustic model-1 implementation based on wavelet packet decomposition instead of conventional Fast Fourier transform (FFT).

Keywords ATH · DWT · FFT · SMR · MMT · Psychoacoustic

1 Introduction

Data compression refers to the way of reducing data size without affecting the quality of the data, Audio compression form of data compression, and to get compressed audio, different compression methods have been designed and

B. K. Jagadeesh (✉)
ECE Department, HKBK College of Engineering, Bangalore, Karnataka, India
e-mail: jbkanade_99@yahoo.com

B. Siva Kumar
Department of Telecommunication, Doctor Ambedkar Institute of Technology, Bangalore, Karnataka, India
e-mail: sivabs2000@yahoo.co.uk

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_51,

implemented. This varies from simple technique to most advance and complex one that takes sensitivity of the human ear. In the process of audio compression, perceptual limitation of human ear is exploited. Limitation in human hearing leads to eliminate some sound information that is not perceived in the original signal. Ear is an organ of large sensibility and can present a high resolution and a great dynamic range of the signal. If the filtering operation is not proper, then it will lead to loss of quality of the aural sound. MPEG hybrid data compression (involves both lossy and lossless compression) technique algorithm achieves compression by exploiting the perceptual limitation of the human ear. By applying audio compression algorithms, it is possible to get compact digital representations of audio signals for the purpose of efficient transmission without impairing the quality at the receiving end. The main purpose of the audio coding is to represent the audio signal with a minimum number of bits while achieving transparent signal reproduction.

The absolute threshold of hearing (ATH) characterizes the amount of energy required in a pure tone so that it can be detected by a listener in a noiseless environment. The absolute threshold is typically expressed in terms of dB sound pressure level (SPL). The frequency dependence of this threshold was quantified as early as 1940, when Fletcher reported test results for a range of listeners that were generated in a National Institutes of Health study of typical American hearing acuity. The quiet (absolute) threshold is well approximated by the nonlinear function [1]. Shaping the coding distortion spectrum by applying ATH is first step toward the perceptual coding. Considered on its own, however, the absolute threshold is of limited value in the coding context. The detection threshold for spectrally complex quantization noise is a modified version of the absolute threshold, with its shape determined by the stimuli present at any given time. Since stimuli are in general time varying, the detection threshold is also a time-varying function of the input signal. Auditory masking is a psychoacoustical phenomenon in which a weak signal is masked in the presence of a stronger signal, the stronger signal is called masker, and the signal which is masked by stronger signal is called maskee. Exploiting this (masker tone masking maskee tone) phenomenon in perceptual audio compression is achieved so that the original audio signal is treated as a masker for distortions introduced by lossy data.

The psychoacoustic model used in the perceptual audio coder is based on the Psychoacoustic Model-1 from the MPEG-1 Audio Standard. The MPEG-1 Audio Standard describes two different psychoacoustic models, that is, psychoacoustic models-1 and psychoacoustic models-2, the first being computationally simpler and suitable for coding at higher bit rates and the second being more complex but also more reliable at lower bit rates due to the complexity associated with the construction of a psychoacoustic model-2. In this work, we have used psychoacoustic models-1 for our implementation.

The majority of MPEG coders applies a psychoacoustic model for coding and uses the filter bank to approximate the frequency selectivity of the human auditory system. Figure 1a and b shows a diagram of the structure of a generic perceptual audio coder. Figure 1a shows the structure of the encoder, which has three main

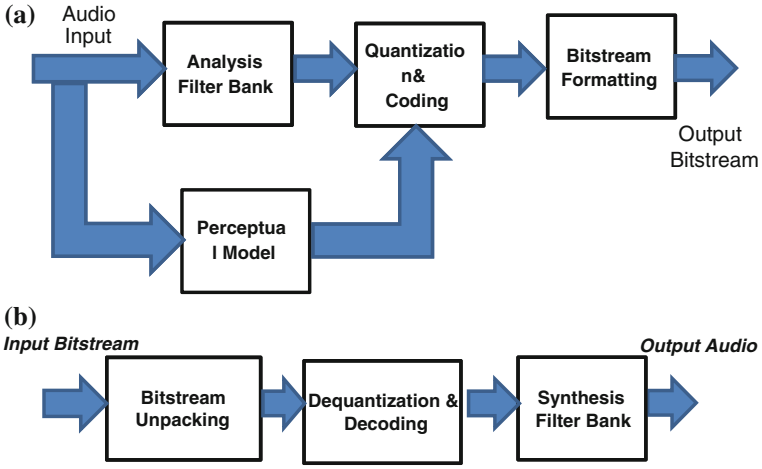


Fig. 1 **a** Encoder. **b** Decoder

stages and a fourth is bit stream formatting stage, and Fig. 1b shows the decoder, which has three stages. Decoder operates on the encoded input audio signal (bitstream) and outputs the reconstructed (original) signal.

Three stages are there in the encoder, namely the signal analysis, Quantization and encoding, and bitstream formatting. The three stages of the decoder correspond to the signal synthesis, de-quantization and decoding, and bitstream extraction stages, respectively. The extra stage in the encoder is the psychoacoustic model, which is not required in the decoder since the information is implicitly encoded as side-information. This means that perceptual coders are asymmetrical in that the encoder has a greater computational requirement than the decoder, which actually can be desirable in certain applications where one\server "encodes the signal for many clients." The discrete wavelet transform can conveniently decompose the signal into an auditory critical band-like partition [2, 3]. Signal decomposition into critical bands resulting from wavelet analysis needs to satisfy the spectral resolution requirements of the human auditory system.

2 Wavelet Packet Decomposition

Wavelet packet decomposition [4] provides a solution that makes possible for a finer and adjustable resolution of frequencies at high frequencies. This makes adaptation to particular signals [5]. Psychoacoustic model achieves an improved decomposition of the signal into critical bands using the discrete wavelet packet decomposition transform (DWPT).This results in a spectral partition which approximates the critical band distribution much closer than before. Furthermore, the Masking thresholds are computed entirely in the wavelet domain to get

approximation of critical bands using wavelets analysis and should meet the spectral resolution requirements. The wavelet basis also plays important role to satisfy temporal resolution of the signal. The typical range is less than 10 ms at high frequencies to 100 ms [6] at low frequencies. Individual masking threshold will be calculated for each component; then, deduction of tonal and nontonal components from the spectrum that has resulted from wavelet decomposition will give the global masking threshold [4]. The continuous wavelets transform (CWT) of signal x relative to the basic wavelet is given by:

$$W_{\psi}x(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t)\psi^{*}\left(\frac{t-b}{a}\right) \tag{1}$$

where a,b (a,b $\in \mathbb{R}$, a $\neq$ 0) are, respectively, the translation and scale parameters. Furthermore, ψ(t − b/a) represents the wavelet basis functions that are derived from a single mother wavelet function, ψ(t), through dilations a and translations b. The wavelet basis functions represent an Orthonormal basis to the space of L^2(R) such that,

$$L^2(R) = \text{span}\ \{\psi ab\ (t);\ a \in R +\ ,\ b \in R\}$$

If the basic wavelet satisfies the admissibility condition, then the wavelet reconstruction formula is:

$$x(t) = \iint_{R} W_{\psi}x(a,b)\psi_{a,b}(t)\frac{\text{dadb}}{a^2} \tag{2}$$

The standard DWT involves a dyadic tree structure in which the low-channel side is successively split down to a certain depth. Wavelet decomposition is a wavelet transform in which the signal is passed through more number of filters. The detailed coefficients will be obtained from the right-leaf node of each level and the approximation coefficients will be obtained from the left-leaf node at the lowest level. Figure 2 illustrates DWPT where the nodes represent the wavelet coefficients (at various decomposition levels).

Wavelet packet decomposition with depth one splits the signal into high-pass and low-pass bands. Depth two splits the low-pass spectrum from depth one. In each stage wavelet decomposition splits low-pass spectrum from previous stage, this yields an octave band-pass filter bank wherein sampling rate of each sub-band is proportional to its bandwidth. Wavlet analysis is efficient because of portions of the frequency toward the low frequency; the psychoacoustic model is based on many studies of human perception. Studies have proven that the average human does not hear all frequencies as same.

While choosing specific wavelet decomposition, the author has considered some restrictions to create orthogonal translates and dilates of the wavelet (the same number of coefficients than the scaling functions) and to ensure regularity (fast decay of coefficients controlled by choosing wavelets with large number of

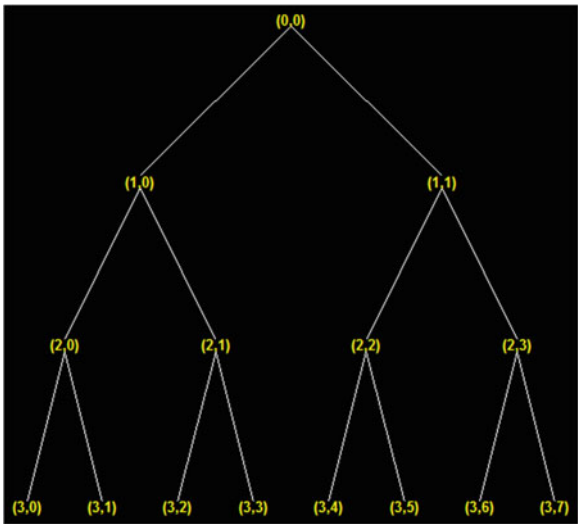


Fig. 2 Wavelet packet decomposition with depth three

vanishing moments). In this work, author used orthogonal family of wavelets with name "daubechies" and Haar. In this work, author has used daubechies "DB10" family wavelet

3 Proposed design of psychoacoustic model-1

Figure 3 shows the proposed design of psychoacoustic model-1 for MPEG audio using wavelet packet decomposition.

The signal spectrum will be obtained by applying wavelet packet decomposition instead of FFT and whose connections are selected in such a way that subbands correspond to the best possible one to the critical bands.

4 Implementation of Proposed Design of Psychoacoustic Model-1 with Wavelet Packet Decomposition

The detailed design of the proposed method is as shown in Fig. 4. The implementation steps as below

Step 1: Framing

The "wav" file is actually an uncompressed audio signal, the sampling rate chosen here is 44,100 Hz. Divide the audio signal into different frames each frame of size 2048.

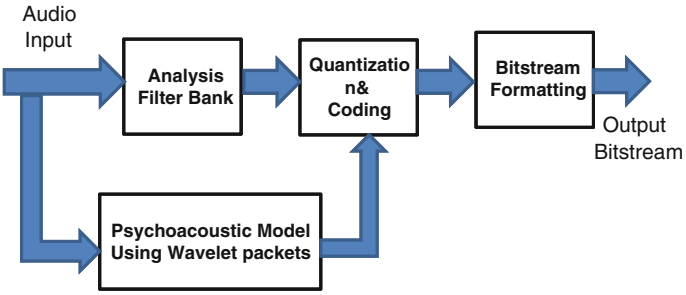


Fig. 3 Design of psychoacoustic model with wavelet packet decomposition

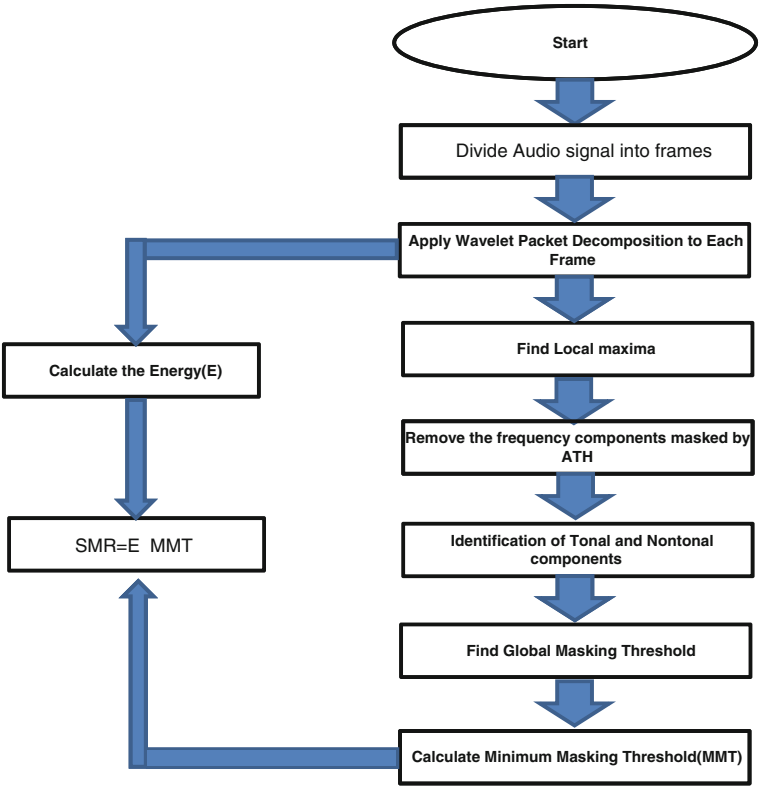


Fig. 4 Psychoacoustic model with wavelet packet decomposition

Step 2: Apply Wavelet Packet decomposition to each Frame and Calculation of Energy

Wavelet packet decomposition is a wavelet transform in which the signal is passed through more number of filters. In this work, author used orthogonal family

of wavelets with name "daubechies" and "Haar" wavelets. There are so many types of "daubechies" wavelet. This work uses "db10" Fig. 2 shows the wavelet packet decomposition. Applying wavelet packet of depth level five will give total of 32 sub-bands that replicate the critical bands in human auditory system, Subsequently calculating the energy $E_{subband}$ for each sub-band.

Step 3: Local maxima

After wavelet decomposition, we are concerned with finding

$$X(i) > X(i \pm 1)$$

Step 4: Elimination of frequency components masked by ATH

The ATH gives the information about the amount of energy needed in a pure tone such that it can be detected by a healthy listener in noiseless environment [1]. The local maxima calculated in step 3 helps in determining those coefficients of an audio signal which are irrelevant for the healthy listener. The aim using ATH is to retain only those coefficients which lie above the curve as in the Fig. 4. The quiet (absolute) threshold is well approximated [1] by the nonlinear function.

$$T_q(f) = 3.64(f/1,000)^{-0.8} - 6.5e^{-0.6(f/1,000-3.3)2} + 10^{-3}(f/1,000)^4 (\text{dB SPL})$$

which is representative of a young listener with acute hearing [1] as shown in the Fig. 5.

Step 5: Localization of tonal and nontonal components

After all the estimations and normalizations done, next step is identifying the tonal and nontonal masking components from local maxima that has been obtained. If frequency bin (component) exceeds neighboring components within a

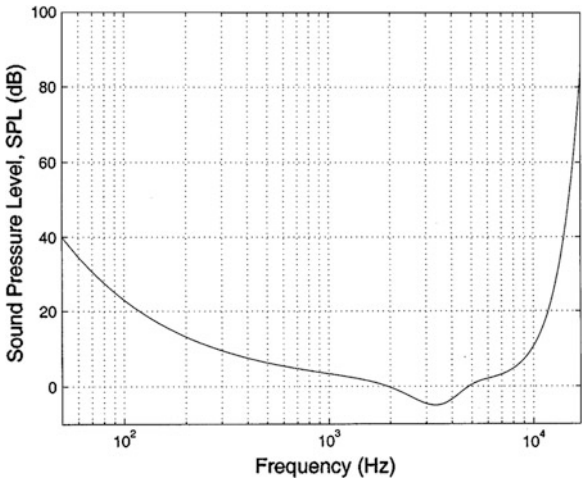


Fig. 5 Absolute threshold of hearing in quiet [1]

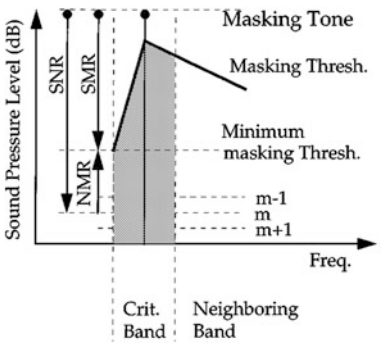


Fig. 6 Schematic representation of simultaneous masking [1]

bark distance by at least 7 db then it will treated as "tonal" otherwise it will be considered as "nontonal". Notions of critical bandwidth and simultaneous masking in the audio coding context give rise to some convenient terminology illustrated in Fig. 6. More detail is provided in [1].

Step 6: Individual masking threshold

Every tonal and nontonal component has its effect on the neighboring coefficients. The effect can be analyzed as its individual Masking threshold. This can be analyzed by the spreading function; the spreading function [1] is given by

$$\mathrm{SF_{db}(x)} = 15.81 + 7.5(\mathrm{x} + 0.474) - 17.5\sqrt{1 + (\mathrm{x} + 0.474)^2}\mathrm{db}$$

where unit of x is Bark and $\mathrm{SF_{db}(x)}$ is expressed in dB.

Critical band analysis is done, and the spread of masking has been accounted for; masking thresholds in perceptual coders are often established by the decibel relations [1]

$$\mathrm{TH}_N = \mathrm{E}_T - 14.5 - \mathrm{B} \text{ and } \mathrm{TH}_T = \mathrm{E}_N - \mathrm{K}$$

where $\mathrm{TH_N}$ and $\mathrm{TH_T}$ noise and tone masking thresholds, respectively. Figure 6 shows the schematic representation of simultaneous masking. Where $\mathrm{TH_N}$ and $\mathrm{TH_T}$ noise and tone masking thresholds, respectively. Figure 6 shows the schematic representation of simultaneous masking.

Step 7: Global masking threshold

Global masking threshold is additive in nature. It is calculated by adding the effect of all the thresholds calculated above. An adjustable resolution of frequencies at high frequencies can be achieved through wavelet decomposition. It gives a structure that allows adaptation to particular signal

Step 8: Minimum masking threshold (*MMT*)

The minimum of the global masking threshold for each sub-band is calculated and is referred as MMT.

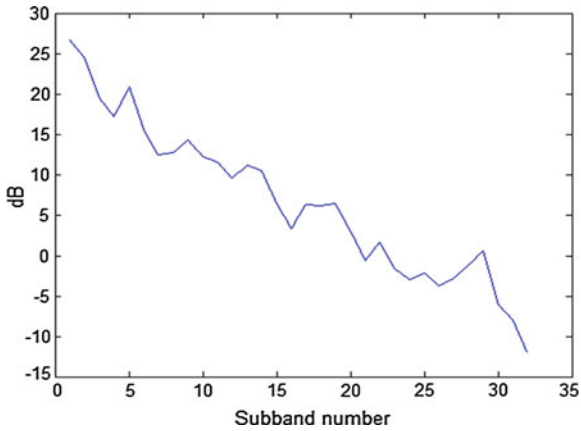


Fig. 7 SMR for different sub-bands (FFT)

Step 9: Calculation of Signal to mask ratio (*SMR*)

The SMR will be calculated by the equation

$$\text{SMR} = \text{E}_{\text{subband}} - \text{MMT}$$

Where $\text{E}_{\text{subband}}$ is the energy in the each sub-band and MMT is MMT for that particular sub-band. Based on SMR, bit allocation is done by bit allocation block for each sub-band. More the SMR requires, more bits need to be allocated for that particular sub-band.

5 Experimental Results

In this work, we have implemented Psychoacoustic model based on the standard MPEG audio using wavelet packet decomposition instead of FFT. This new psychoacoustic model has been implemented in reference coder based on the

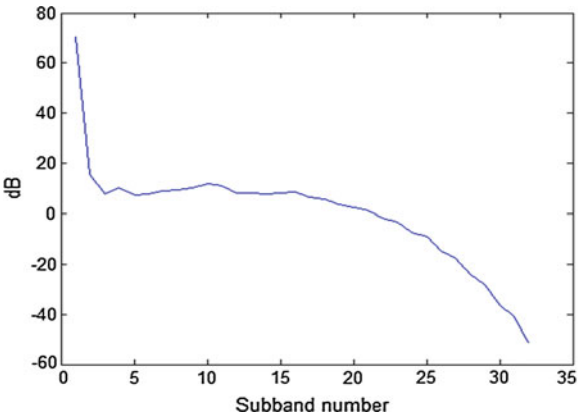


Fig. 8 SMR for different sub-bands (db10-Wavelet packet)

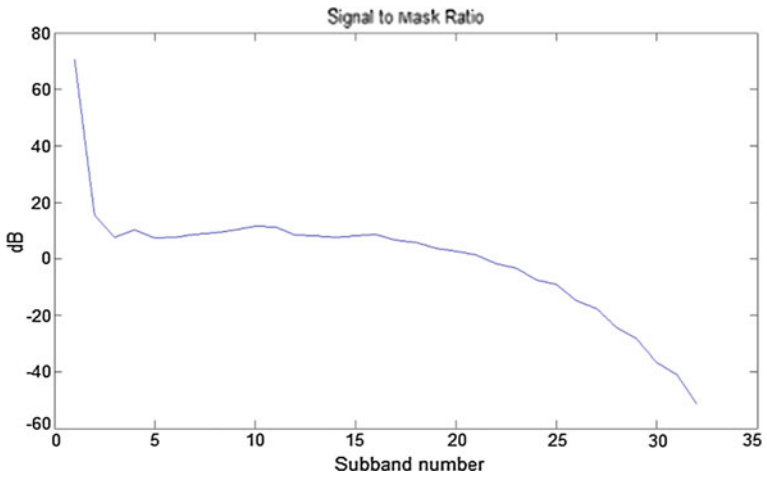


Fig. 9 SMR for different sub-bands (Haar wavelet)

standard MPEG-1 audio. Figures 7 and 8 show the SMR plot using FFT method and proposed wavelet packet (daubechies) method, respectively. From Fig. 8, it is clear that SMR obtained for sub-band 1 is more compared to the results from FFT implementation, and for other sub-bands from 2 to 32, the SMR is significantly low compared to FFT method. Figure 9 shows the SMR plot using Haar wavelets

6 Conclusion and Future Work

The psychoacoustic model-1 based on wavelet packet decomposition takes an account for the critical bands and also takes an account of the masking phenomenon. The specialty of the proposed model is that it gives an analysis by wavelet packet decomposition on the frequency bands that gives the closer approximation of the critical bands of the ear. Implementation results from this paper can be integrated with the other blocks of the MPEG audio codecs [2] to get overall compressed audio bitstream.

References

1. Painter T, Spanias A (2000) Perceptual coding of digital audio. In: Proceedings of IEEE transactions signal, vol 88, pp 51–513
2. Information technology-Coding of moving pictures and associated audio for digital storage media at up to about 1, 5 Mbit/s-Part3: Audio (1999)
3. Zurera MR, Ferreras FL, Amores MPJ, Basc on SM, Reyes NR (2001) A new algorithm for translating psycho-acoustic information to the wavelet domain. Sig Process 81(3):519–531

4. Burrus CS, Gopinath RA, Guo H (1998) Introduction to wavelets and wavelets transforms: a primer. Prentice Hall, Englewood Cliffs, Upper Sadd le River, NJ (USA)
5. Wickerhauser MV (1994) Adapted wavelet analysis from theory to software. Wellesley, Assachusetts
6. Zurera MR, Ferreras FL, Amores MPJ, Basc′ on SM, Reyes NR (2001) A new algorithm for translating psycho-acoustic information to the wavelet domain. Sig Process 81(3):519–531
7. Sinha D, Tewfik A (1993) Low bit rate transparent audio compression using adapted wavelets. IEEE Trans Sig Proc 3463–3479
8. Bosi M, Goldberg RE (2003) Introduction to digital audio coding and standards. Kluwer Academic Publishers, New York
9. Lincoln B (1998) An experimental high fidelity perceptual audio coder. Project in MUS420 Win97, March 1998
10. Johnston JD (1988) Transform coding of audio signals using perceptual noise criteria. IEEE J Sel Areas Commun 6(2):314–323
11. Carnero B, Drygajlo A (1999) Perceptual speech coding and enhancement using frame-synchronized fast wavelet packet transform algorithms. IEEE Trans Sig Process 47(6):1622–1635
12. Strang G, Nguyen T (1996) Wavelets and filter banks. Wellesley-Cambridge Press, Wellesley
13. Chui CK, Wang Jz (1992) On compactly supported spline wavelets and a duality principle. Rans Amer Mathematic Soc 330:903–915
14. Ogden RT (1997) Essential wavelets for statistical applications and data analysis. Birkhauser, Boston
15. Meyer Y (1989) Orthonormal wavelets. in wavelets, Proc. Int. Conf. France, Springer, Berlin

Recommending Music by Combining Content-Based and Collaborative Filtering with User Preferences

S. Sneha, D. S. Jayalakshmi, J. Shruthi and Uttarika Ratnakar Shetty

Abstract Currently available music recommender systems either use content-based or collaborative filtering, due to which the system either suffers from repetitive recommendations or takes a lot of user interaction to give good recommendations. We present a novel approach, which uses features from content-based and collaborative filtering to recommend suitable tracks from a collection of songs. Audio signal processing techniques are used to find the musical features that define the similarity of musical pieces. The K-nearest neighbor classifier (KNN) is employed on these set of features to find the most similar musical pieces. Collaborative filtering is done on the musical pieces by assigning ranks and calculating frequencies of the piece being played. Furthermore, user-defined rules regarding the user's personal preferences are applied on the set of songs obtained using the filtering techniques. The result is a unique algorithm which eliminates the present need for a user to wait for the computer to learn his musical inclinations or having to subject himself to listening to the same songs multiple times.

Keywords Music recommender systems · Content-based filtering · Collaborative filtering · Hybrid music recommenders · Information retrieval

S. Sneha (✉) · D. S. Jayalakshmi · J. Shruthi · U. R. Shetty
M.S. Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: sneha4ssn@gmail.com

D. S. Jayalakshmi
e-mail: jayalakshmids@msrit.edu

J. Shruthi
e-mail: shruthiford1@gmail.com

U. R. Shetty
e-mail: uttarika.shetty@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_52,

1 Introduction

Current online music recommendation services, such as Last.fm and Pandora, recommend music to a user based on the user's interests. Most sites use either collaborative or content-based approaches for recommendations: both of which have specific advantages and disadvantages [1, 2]. Collaborative recommendations suffer from a few issues; if a song has not yet been rated by a user, it cannot be recommended. They depend on an adequate number of ratings by a user for the system to accurately assess the user's interests; even if a track is supposed to be similar to a user's taste, the user may not always enjoy the recommended music [3]. Content-based recommendations eliminate these problems, but the process of analyzing only music data to make recommendations is not as accurate as current collaborative methods [4]. They also assume that the user will know what music he/she wants to search for. Therefore, users may never find music they have not heard of before. This paper centers on the possibility that combining these two methods into one approach will improve the recommendation results because it will have the advantages of both methods.

The music recommender system detailed in this paper is based on an initial music library of about 3,500 songs. A hybrid music recommender system is built combining user preferences with content-based and collaborative filtering. The results obtained indicate that the pieces of music recommended are similar, familiar, popular, and serendipitous to the user all the while not taking a performance hit due to the large amount of calculations done.

2 Related Work

Kim et al. [4] presented a content-based system using acoustic features and a dynamic K-means clustering algorithm. Cano et al. [5] presented a "content-based system" wherein descriptions related to instrumentation, rhythm, and harmony are extracted from music signal using similarity metrics. Tzanetakis and Cook [6] presented a way to extract features from music and ascertain musical genre automatically. Logan [7] presented that recommendation can be made based solely on acoustics from groups of similar songs or "song sets." Chen and Chen [8] presents a collaborative recommendation system. He groups pieces of music which are similar to each other and users who have similar preference. Yapriady [9] shows a way to do collaborative filtering combined with demographic data.

2.1 Hybrid Music Recommender Systems

Research has demonstrated that a hybrid approach, combining collaborative filtering and content-based filtering could be more effective in some cases [4].

Hybrid approaches can be implemented in several ways: by making content-based and collaborative filtering–based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice versa); or by unifying the approaches into one model [10].

2.2 Audio Features for Music Analysis

The characterization of music and speech signals can be done with good accuracy on the basis of different types of features such as pitch, timbral features, music content, music tempo, and rhythm features [5]. The following specific features are used to represent timbral texture in our system.

- ***Zero Crossing Ratio*** counts the number of times that the sign of the signal amplitude changes in the time domain in one frame. Zero crossing ratio, Z_t, is defined

$$Z_t = \frac{1}{2}\sum_{n=1}^{N} |\text{sign}(x[n]) - \text{sign}(x[n-1])| \tag{1}$$

  where sign function is 1 for positive arguments and zero for negative arguments, and $x[n]$ is the time domain signal for frame t. For single-voiced signals, zero crossings are used to make a rough estimation of the fundamental frequency [8]. For complex signals, it provides a measure of the noisiness of the signal. For example, heavy metal music due to guitar distortion and lots of drums will tend to have much higher zero crossing values than classical music.

- ***Spectral centroid*** is defined as the center of gravity of the magnitude spectrum of the STFT

$$C_t = \frac{\sum_{n=1}^{N} M_t[n] * n}{\sum_{n=1}^{N} M_t[n]} \tag{2}$$

  where $M_t[n]$ is the magnitude of the Fourier transform at frame t and frequency bin n. The centroid is a measure of spectral shape, and higher centroid values correspond to brighter textures with more high frequencies [8].

- ***Spectral Roll-off*** is defined as the frequency below which 85 % of the magnitude distribution of the spectrum is concentrated. The equation for roll-off is

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 * \sum_{n=1}^{N} M_t[n] \tag{3}$$

  If R_t is the largest value of n for which this equation is satisfied, then this frequency R_t is the roll-off. It shows how much of the signal's energy is concentrated in the lower frequencies [5].

- ***Spectral flux*** is defined as the squared difference between the normalized magnitudes of successive spectral distributions

$$F_t = \sum_{n=1}^{N} (N_t[n] - N_{t-1}[n])^2 \quad (4)$$

where $Nt[n]$ and $Nt - 1[n]$ are the normalized magnitude of the Fourier transform at the current frame t and the previous frame $t - 1$, respectively. The spectral flux is a measure of the amount of local spectral change. Flux is an important feature in the separation of music from speech and an important perceptual attribute in the characterization of musical instrument timbre [5].

- ***Fraction of low-energy window*** is defined as the percentage of analysis windows that have less RMS energy than the average RMS energy across the texture window. It can indicate how much of a signal section quiet relative to the rest of the signal section is. For example, vocal music with silences will have large low-energy value, while continuous strings will have small low-energy values [11].

3 Proposed Hybrid Music Recommender System

We propose a hybrid music recommender system based on cascade approach [10]. Figure 1 shows a system designed in such a way wherein a user and an administrator can access it on two different interfaces. There are four main modules: creation of account by the user, the logging in of the user, the recommendation module, and the module for the administrator. The database contains data regarding the output of the classifier, the details regarding the user preferences, the acoustic features, and the ID3 tags of the various songs in the database, and a collaborative module which contains data about the songs various users have played.

3.1 Algorithm

Songs similar to every other song with respect to the audio feature set are determined using the K-nearest neighbor classifier (KNN). Collaborative filtering is done by assigning ranks to the songs played based on ratings given by the user as well as probabilistically determining the user's preference to the song by the amount of time he spends on listening to the song. Initially, the user can choose to give any preferences to a specific artist or genre of music which would make the recommendations more prolific. The algorithm performs content-based and collaborative filtering and finally applies user preference rules. Various iterations of the above set of algorithms are applied to achieve the most optimal output.

Content-based Filtering Content-based filtering is done by applying the KNN on all of the acoustic features that have been extracted from the various pieces of

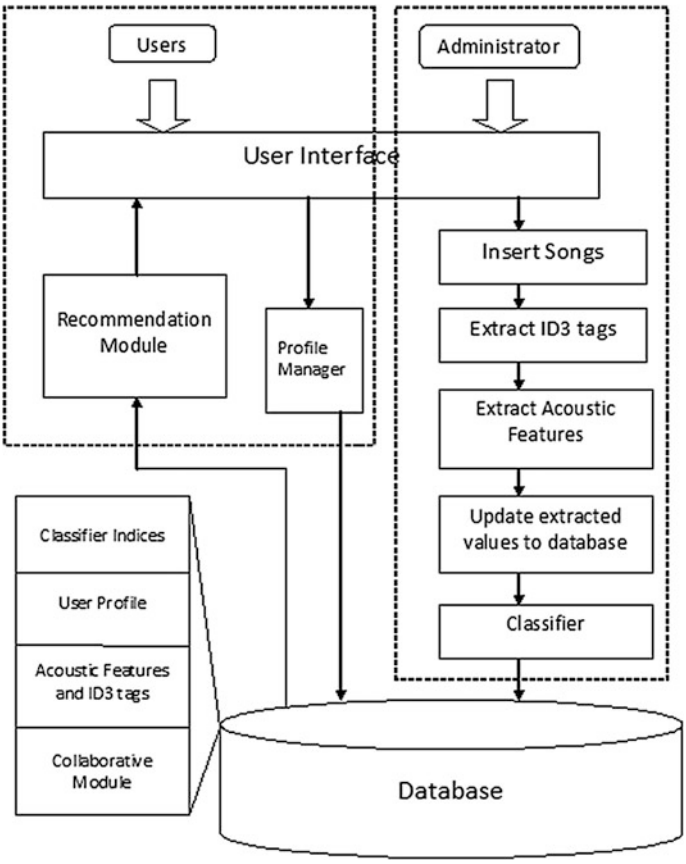


Fig. 1 Hybrid music recommender system design

music, and the ten songs which are the most similar to each particular song are determined. The value of "k" used here is 10, chosen to leverage the situation of "cold play" wherein a first time user has not provided many preferences and still wants quite a few recommendations from the system.

Collaborative Filtering Collaborative filtering is done on the initial data set to obtain a set of 10 songs. The following criteria contribute in identifying and determining the final result set.

1. *User Rating* A user can rate a recommended song discretely from 1 to 5. If the user does not rate the song, the algorithm rates the song for the user assuming that the amount of time the user spends in listening to the song recommended is directly proportional to how amicable he finds the song. Therefore, the calculated rating would be:

$$\text{Rating} = \frac{(\text{Number of seconds the user has played the song})}{\text{Total Duration of the song in seconds}} * 5 \qquad (5)$$

2. *Frequency of the song being played* The number of times the user plays a specific song as well as the number of times the song is played by multiple users is calculated.
3. *User Interactions* If a user explicitly specifies his interest in a specific artist or genre while using the system or prior to using the system, it is converted into a rule and is used for generating of the result set. When a user specifies that he is open to new genres a rule for his serendipitous inclination is created.

Collaborative Filtering Algorithm.

1. Apply the rules specified by the user regarding his preferred artists and genres on all the songs in the data set.
2. From the resulting set choose the songs that have been played by the other users.
3. From the updated result sets choose the songs that have the average rating of 3.5 or above on a scale of 1–5.
4. From this final result set, randomly choose 10 songs.
5. If the user has specified that he would like to be adventurous with the types of songs he listens to, randomly remove a song from the data set and add a random song from the entire data set.

Determination of the Optimal Result Set The final rule applied on both of the result sets is checking whether the user previously listened to any of the songs in the result set and has rated it low. If the user has rated a song low, then that song is removed from the result set. Finally, five songs from the result sets of both the filtering techniques are randomly chosen and the result sets are unified. The user is ultimately given 10 optimal recommendations.

4 Results

Two experimental systems were built; a hybrid system using the algorithm specified in this paper, and the other with a purely content-based filtering approach. Fifty users were asked to test both the recommender systems and rate the two systems on the specified criteria, namely type of recommendations made, performance of the system, new artists' discovery and randomness of new recommendation iterations.

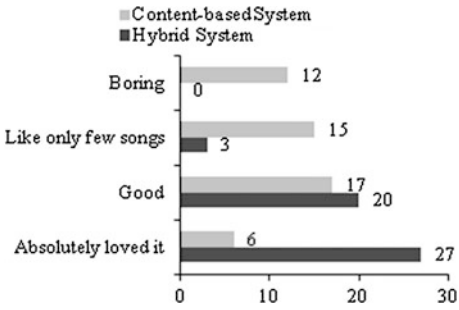


Fig. 2 Comparsion on type of recommendations made with respect to different artists and genres

1. **Type of recommendations made**

Figure 2 shows that the hybrid system has performed better due to the collaborative module present in the system mainly because it introduced an element of surprise. Users got cross-genre recommendations from the content-based module, and familiar songs from the collaborative module consisting of songs that have been consistently rated high by other users. Users also liked the fact that they discovered new songs because the system matched user preferences of multiple users.

2. **Performance of the system**

Figure 3 shows that users found the content-based system faster in generating a new list of recommendations. This is because the hybrid system does computations on the songs retrieved from the collaborative table. While the content-based system requires 1 ms to compute a new list of recommendations, the hybrid system requires 1.5 ms, which can be termed negligible with respect to the overall user experience.

3. **New artists' discovery**

Figure 4 shows that the users who used both the systems felt that they discovered more songs using the hybrid recommender system rather than the content-based one. While both the systems enable the users to discover many artists because the

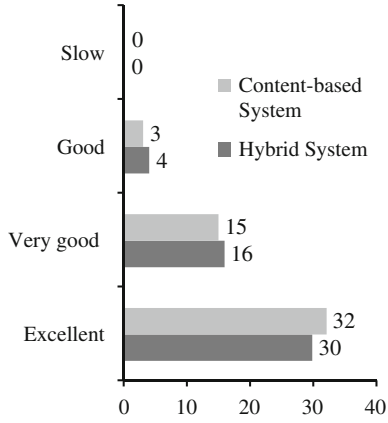


Fig. 3 Performance of the system

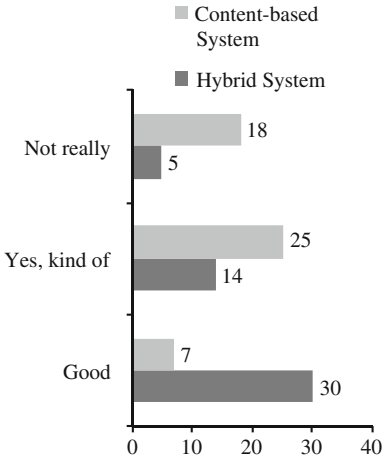


Fig. 4 New artists' discovery

content-based module recommends similar songs based on acoustic features across genres, the hybrid system's advantage is in the collaborative module.

4. **Randomness of new recommendation iterations**

Figure 5 shows that 68 % of the users who used both systems felt that the hybrid system was more random, while a mere 16 % of the users felt that the content-based system was very random. Randomness quantifies how different and unique the recommendations are. In a content-based system, if a user were to go on requesting for recommendations when a particular song was being played, he would get the same list of songs as before, since the ten most similar songs to a particular song would always remain the same and no further processing is done

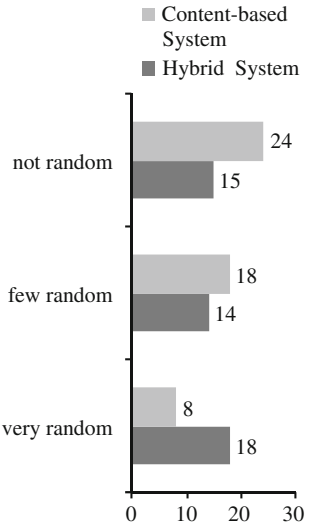


Fig. 5 Randomness of new recommendation iterations

with respect to user ranks or rules. On the other hand, while using the hybrid system, if the user were to repeatedly request for new recommendations, he would get five different songs each time. This happens due to the collaborative module where a constant shuffling and randomizing of songs is done to ensure that the user discovers new songs each time.

5 Conclusions

The system built using the algorithm specified provides better results in terms of accuracy and user satisfaction which further reinforces that a hybrid approach to recommender systems overshadows both the content-based approach and the collaborative approach. Future enhancements can be done in the algorithmic level by using a different mix of audio features and changing the classifier used to classify the songs. Other classifiers like naïve Bayesian classifier, support vector machines can be used. The system can also be implemented on a distributed environment.

References

1. Song Y, Dixon S, Pearce M (2012) A survey of music recommendation systems and future perspectives. 9th international symposium on computer music modelling and retrieval (CMMR 2012), 19–22 June 2012
2. Burke R (2002) Hybrid recommender systems: survey and experiments. User Model User-Adap Inter 12(4):331–370
3. Hijikata Y, Iwahama K, Takegawa K, Nishida S (2006) Content-based music filtering system with editable user profile, SAC, pp 1–6
4. Kim DM, Kim K, Park KH, Lee JH, Lee KM (2007) A music recommendation system with a dynamic K-means clustering algorithm. IEEE sixth international conference on machine learning and applications, pp 1–5
5. Cano P, Koppenberger M, Wack N (2005) Content-based music audio recommendation. In: Proceedings of ACM international conference on multimedia, pp 211–212
6. Tzanetakis G, Cook P (2002) Musical genre classification of audio signals. IEEE Trans Speech Audio Process 293–302
7. Logan B (2004) Music recommendation from song sets. In: Proceedings of ISMIR, pp 425–428
8. Chen H, Chen A (2005) A music recommendation system based on music data grouping and user interests. J Intell Inf Syst 24(2/3):113–132
9. Yapriady B, Uitdenbogerd AL (2005) Combining demographic data with collaborative filtering for automatic music recommendation. Lecture notes in computer science. Springer, pp 201–207
10. Yoshii K Studies on hybrid music recommendation using timbral and rhythmic features. Ph.D. thesis at http://winnie.kuis.kyoto-u.ac.jp/members/yoshii/d-thesis-yoshii.pdf
11. Subramanian H (2004) Audio signal classification. EE Dept, IIT Bombay, pp 1–5, Nov 2004
12. Fomby TB (2008) K-nearest neighbours algorithm: prediction and classification. Lecture notes in Southern Methodist University, Dallas, TX, Feb 2008, pp 1–5

Performance of Sphere Decoder for MIMO System Using LLL Algorithm

Suneeta V. Budihal, Rashmi Hiremath and R. M. Banakar

Abstract Presently, lattice reduction aided detectors are proposed for multiple-input multiple-output (MIMO) systems to achieve performance like maximum likelihood receivers and yet with complexity similar to linear receiver. In MIMO systems, sphere decoding (SD) achieves performance equivalent to full search maximum likelihood decoding with reduced complexity. One of key issues in SD is the selection of an initial radius of the search hypersphere. This paper presents a deterministic radius selection algorithm using the Babai estimate. The traditional LLL reduction algorithm that was used for reducing real lattice bases is used here. The numerical show that the lattice reduction is aided. SD algorithm achieved more success rate here. There is a reduction in the radius of the sphere decoder by 64 % for the same parameter. This reduction in initial radius reduces the number of iterations required to decode the received bits.

Keywords Sphere decoder · Lattice basis reduction · MIMO

1 Introduction

The applications in communications may be modeled as integer least squares problems. The goal is to find the solution to the integer least squares problem, which could be the encoded integer vector in these applications. From the lattice

S. V. Budihal (✉) · R. Hiremath · R. M. Banakar
B.V. Bhoomaraddi College of Engineering and Technology, Hubli, India
e-mail: suneeta_vb@bvb.edu

R. Hiremath
e-mail: rashmi.v.hiremath@gmail.com

R. M. Banakar
e-mail: banakar@bvb.edu

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_53,

space point of view, finding the solution to an integer least squares problem is equivalent to finding the closest lattice point to a given point. Sphere decoding (SD) is often used to the searching of the closest lattice point. Originally, SD, is developed by Finke and Phost in [1] in 1985, enumerates all lattice points in a hypersphere centered at a given vector. He proposed that SD can be viewed as tree search.

In [2], the problem of selecting initial radius is carried by introducing natural candidate, which is the covering radius of the lattice, defined to be the smallest radius of spheres centered at the lattice points that cover the entire space.

In [3], selection of radius d of the hypersphere is discussed. If d is too large, the hypersphere contains too many lattice points, then the search complexity may be exponential to d; if d is too small, on the other hand, the hypersphere may contain no lattice points. There are no general guidelines for selecting an appropriate d. It is application dependent.

In this paper, the issue of selecting initial radius is presented. First, we present a deterministic radius selection algorithm using the Babai estimate [4]. By selecting initial radius, if it is too large, we obtain too many points and search remains exponential in size; where as if is too small, we obtain no points inside the sphere. Then, we apply LLL algorithm to Babai estimate to reduce the complexity.

2 System Model

Consider multiple-input multiple-output (MIMO) systems with M transmit antennas and N receive antennas. The discrete time received signal can be expressed as

$$y = Hx + n$$

where H denotes the channel matrix, x denotes the vector of transmitted symbols, n is the vector of independent and identically distributed noise, and y is the vector of received symbols.

SD is a method for solving the integer least squares problem:

$$\min_{x \in Z^m} \|Hx - y\|_2^2$$

where $y \in R^m$ and $H \in R^{n \times m}$. Note that while x is an integer vector, both the matrix H and vector y are real. As the standard way of solving least squares problems [5], assuming the matrix H is of full column rank, H is first reduced into an upper triangular matrix using orthogonal transformations, such as the householder transformation, to obtain the QR decomposition:

$$H = Q \begin{bmatrix} R \\ 0 \end{bmatrix}$$

where $Q \in R^{n\times n}$ is orthogonal, and $R \in R^{m\times m}$ is upper triangular. Partitioning $Q = [Q1\ Q2]$, where $Q1$ is $n \times m$, and $Q2$ is $n \times (n - m)$, we get

$$\begin{aligned}\|Hx - y\|_2^2 &= \left\| Q\begin{bmatrix} R \\ 0 \end{bmatrix} x - y \right\|_2^2 \\ &= \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} x - Q^T y \right\|_2^2 \\ &= \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} x - \begin{bmatrix} {Q_1}^T \\ {Q_2}^T \end{bmatrix} y \right\|_2^2 \\ &= \left\| Rx - Q_1^T y \right\|_2^2 + \left\| Q_2^T y \right\|_2^2\end{aligned}$$

Note that the second term in the above equation is independent of x. Denoting $\hat{y}$ as $Q_1^T y$, the integer least squares problem is then reduced to following triangular integer least squares problem:

$$\min_{x \in Z^m} \|Rx - \hat{y}\|_2^2$$

SD solves the above triangular integer least squares problem arising from communication applications. It searches a solution in a predetermined hypersphere centered at $\hat{y}$.

Organization of paper: In Sect. 3, SD is described. In Sect. 4, adding LLL algorithm to SD is presented. Section 5 summarizes the results.

3 Sphere Decoder

SD, introduced originally by Finke and Pohst in [1] in 1985, enumerates all lattice points in a hypersphere centered at a given vector. It searches a lattice point in a hypersphere of radius r and centered at $\hat{y}$, that is closest, in Euclidean distance, to the center. Therefore, by restricting the search area, it can reduce the computational complexity of solving the triangular integer least squares problem. Figure 1 illustrates a hypersphere centered at a vector represented by the hollow point.

The sphere radius of d and centered at y can be defined as

$$X = \{x, \|Hx - y\| \le d\}$$

whose condition is equivalent to

$$\|Rx - \hat{y}\|_2^2 \le d^2$$

where $\hat{d}^2 = d^2 - \left\| Q_2^T y \right\|_2^2$.

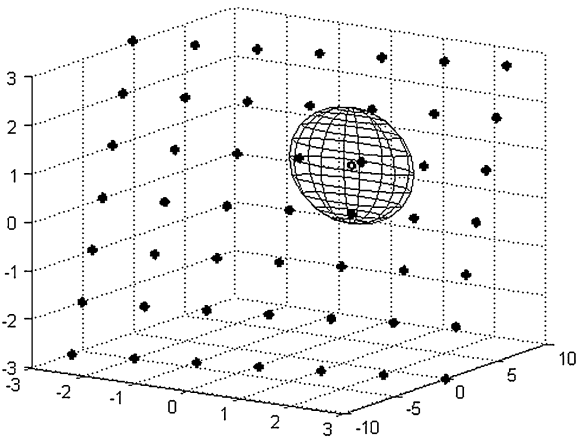


Fig. 1 Geometric interpretation of hypersphere in a lattice space

Since R is upper triangular, so rewriting the above condition, in entrywise as

$$\hat{d}^2 \geq \left(\sum_{i=1}^{m} \left(\sum_{j=1}^{m} r_{i,j} x_j - \hat{y}_i \right) \right)^2$$

where $r_{i,j}, j > i$ denotes the (i, j)th entry of R. The above equation is expanded to get the equations.

$$\hat{d}^2 \geq \left(\hat{y}_m - r_{m,m} x_m \right)^2 + \left(\hat{y}_{m-1} - r_{m-1,m} x_m - r_{m-1,m-1} x_{m-1} \right)^2 + \cdots$$

The first term in the right side of above equation depends only on the mth entry x_m of lattice point s, the second term depends on the entries x_m and x_{m-1}, and so on. By solving, we get

$$\left\lceil \frac{-\hat{d} + \hat{y}_m}{r_{m,m}} \right\rceil \leq x_m \leq \left\lfloor \frac{\hat{d} + \hat{y}_m}{r_{m,m}} \right\rfloor$$

Following the above procedure, we can obtain the intervals for x_{m-1}, x_{m-2}, and so on until we reach x_1. Then, we are able to determine all the lattice points in the hypersphere of radius.

Sphere decoding algorithm

Inputs: R, where R is upper triangular.

$\hat{y}$, where $\hat{y}$ is the y reduced by the QR decomposition.

d, radius of sphere

Output: x or null.

Step 1: set k = m, $d_m^2 = d^2 - \left\| Q_2^T y \right\|^2$, $\hat{y}_{m|m+1} = y_m$.

Step 2: (Bounds for x_k), set $UB_{x(k)} = \frac{d_k + \hat{y}_{k|k+1}}{r_{k,k}}$,

$$LB_{x(k)} = \frac{-d_k + \hat{y}_{k|k+1}}{r_{k,k}}$$

Step 3: (Increase x_k) $x_k = x_k + 1$

If $LB_{x(k)} \leq UB_{x(k)}$, go to step 5; else go to step 4.

Step 4: (Increase k) k = k+1

If k = m+1, terminate algorithm; else go to step 3.

Step 5: (Decrease k) if k = 1, go to Step 6; else

$$\text{k} = \text{k}-1,\ \hat{y}_{k|k-1} = \hat{y}_k - \sum_{j=k+1}^{m} r_{k,j}x_j,$$

$$d_k^2 = d_{k+1}^2 - \left(\hat{y}_{k+1|k+2} - r_{k+1,k+1}x_{k+1}\right)^2$$

and go to step 2.

Step 6: solution found. Save x and its distance from y,

$d_m^2 = d_1^2 + \left(y_1 - r_{1,1}x_1\right)^2$ and go to step 3.

3.1 Radius Selection Using Babai's Algorithm

In [6], Qiao proposed a deterministic method for selecting an initial hypersphere radius. This method is designed for communication applications. First, we find the real solution for the triangular system $Rx = \hat{y}$, which is the real least squares solution. Then, we round the entries of s to their nearest integers to obtain the lattice point:

$$\hat{x} = [x] \in Z^m$$

The radius is set to the distance, in Euclidean sense is given by

$$\hat{d} = \|R\hat{x} - \hat{y}\|_2$$

Radius selection algorithm using Babai estimate is as follows:

Inputs: R, where R is upper triangular.
$\hat{y}$, where $\hat{y}$ is the y reduced by the QR decomposition.

Output: the radius $\hat{d}$ of the search sphere.

Step 1: Solve $x \in R^m$ for R$x = \hat{y}$
Step 2: Round $\hat{x} = [x] \in Z^m$
Step 3: Set $\hat{d} = R\hat{x} - \hat{y}_2$

The radius computed by this algorithm ensures that the search sphere contains at least one lattice point.
Therefore, the integer least squares solution lies in the sphere if the radius is computed exactly.

4 Applying LLL Algorithm to the Sphere Decoder

The LLL algorithm is originated from Lenstra, Lenstra, and Lovasz. It is widely used by researchers as a preprocessor to solve the integer least squares problem.

LLL algorithm

1. [Q, R] ← qr(H)
2. $1 \leftarrow 2$
3. $H_l \leftarrow H_l - \left\lceil \frac{R_{l-1,l}}{R_{l-1,l-1}} \right\rfloor H_{l-1}$
4. if $\delta \left|R_{l-1,l-1}\right|^2 > \left|R_{l,l}\right|^2 + \left|R_{l-1,l}\right|^2$ then
5. $H_{l-1} \leftarrow H_l$
6. $l \leftarrow max(l-1, 2)$
7. else

for k = l−1 *to* 1

8. $H_l \leftarrow H_l - \left\lceil \frac{R_{k,l}}{R_{k,k}} \right\rfloor H_k$
9. $l \leftarrow l + 1$
10. end if
11. [Q, R] ← qr(H)

LLL algorithm is able to reduce the computational complexity of SD in two ways. First, it can reduce the initial radius of the hypersphere by reducing the norm of *R*. Second, since SD is a depth-first search algorithm for the lattice points inside a hypersphere, the LLL algorithm can reduce the total number of search paths.

5 Results

Let

$$\mathrm{H} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 0 & -2 \\ -1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \mathrm{y} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

After QR decomposition,

$$\mathrm{R} \approx \begin{bmatrix} -2.4495 & 0.4082 & -0.4082 \\ 0 & 1.3540 & 1.6002 \\ 0 & 0 & 1.8091 \end{bmatrix} \quad \text{and} \quad \hat{y} = \begin{bmatrix} 1.2247 \\ 0.3693 \\ -0.6030 \end{bmatrix}$$

The computed real solution for the triangular system $Rx = \hat{y}$,

$$x \approx \begin{bmatrix} -0.3333 \\ -0.6667 \\ -0.3333 \end{bmatrix}$$

Rounding the entries of x to their nearest integers,

$$x \approx \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Table 1 shows the computation of initial radius of sphere decoder without LLL and with LLL.

Results from Babai estimate: Initial radius to be 0.31 and maximum to be 0.6030.

Results from preprocessing LLL algorithm: Initial radius to be 0.11 and maximum to be 0.8781.

Figure 2 shows comparison of number of iterations of sphere decoder without LLL and with LLL algorithm. Considering the different sizes of channel matrix H, numbers of iterations are calculated. Iterations refer to complexity of the system.

Table 1 Computation of initial radius of sphere decoder without LLL and with LLL

Radius selection algorithm	Considering above numerical example
Using Babai estimate (without LLL)	$0.31 \leq d \leq 0.6030$
Applying LLL algorithm to Babai estimate	$0.11 \leq d \leq 0.8781$

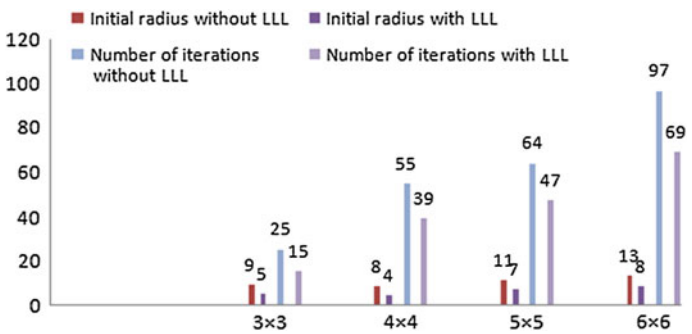


Fig. 2 Performance of sphere decoder with the application of LLL algorithm

In this experiment, a number of iterations are reduced with the application of LLL algorithm.

6 Conclusion

In this paper, sphere decoder is proposed and the issue of selecting search radius in is discussed. First, radius selection algorithm using Babai estimate is presented. Then extending the sphere decoder by LLL algorithm to it. Numericals show that SD algorithm with LLL algorithm, achieved more success rate compare to algorithm without LLL. There is a reduction in the radius of the sphere decoder by 64 % for the same parameter. As there is reduction in the initial radius of sphere decoder, number of iterations required to decode the received bits are also reduced. Hence, complexity is reduced.

References

1. Fincke U, Pohst M (1985) Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. Math Comput 44(170):463–471
2. Hassibi B, Vikalo H (2005) On the sphere-decoding algorithm I. expected complexity. IEEE Trans Signal Process 53(8)
3. Zhao F, Qiao S (2009) Radius selection algorithms for sphere decoding, Canada
4. Qiao S (2008) Integer least squares: Sphere decoding and the LLL algorithm. In: Proceedings of C3S2E—08, ACM international conference proceedings series, May 2008, pp 23–28
5. Golub GH, Van Loan CF (1996) Matrix computations, 3rd edn. Johns Hopkins University Press, Baltimore
6. Grotschel M, Lovasz L, Schriver A (1993) Geometric algorithms and combinatorial optimization, 2nd edn. Springer, New York

Reduced Bit Rate Using Image Inpainting

Veeramma Yatnalli and K. L. Sudha

Abstract In wireless communication, the data need to be efficiently compressed by utilizing the channel bandwidth and thus to increase the data bit rate. We can enhance the compression performance by not sending the significant portion of the image. This can be done by intelligently choosing the blocks from the image with non-relevant information in it. By removing these blocks, size of the image is reduced and the amount of bits needed to transmit the image can be significantly reduced. The removed block's information can be recovered by using the correlation between the blocks and its neighbors. This is done automatically and in a fast way with the boundary information of the lost blocks, thereby allowing to simultaneously fill-in numerous regions containing completely different structures and surrounding backgrounds. This paper addresses the issue of performing inpainting on the decompressed image to fill-in the missing information. The effectiveness of this approach is demonstrated with various compression factors.

Keywords Compression · Inpainting · Bits per pixel

1 Introduction

Data compression is a science of representing information in compact form. Despite rapid progress in mass-storage density, processor speeds and digital communication system performance demand for data storage capacity and data-transmission bandwidth continues to outstrip the capabilities of available

V. Yatnalli (✉)
JSS Academy of Technical Education, Bangalore, India
e-mail: veeramma71@gmail.com

K. L. Sudha
Dayananda Sagar College of Engineering, Bangalore, India
e-mail: klsudha1@rediffmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_54,

technologies. Image compression is an important component of the solutions available for creating image file sizes of manageable and transmittable dimensions.

Image data compression becomes still more important because of the fact that the transfer of uncompressed graphical data require far more bandwidth and data transfer rate. JPEG has led to the popularity of the compressed form of image data. This has brought on the proliferation of image retrieval techniques in the compressed domain.

Many attempts have been made to develop new compression techniques by eliminating the redundancies within images to achieve better coding performance. Among these, image inpainting is a very promising technique to be utilized in image compression, aiming at perceptual quality instead of pixel-wise fidelity. Since inpainting faithfully reproduces lost information, we can afford to voluntarily remove some blocks containing smooth information even prior to compression. This means that only those blocks that can be easily restored are chosen and erased. In this way, the information related to redundancy is removed prior to the transmission. By removing these blocks, the size of the image can be reduced. This reduced size image efficiently utilizes the bandwidth during transmission. The compressed image is decompressed back, and image's removed blocks are recovered by image inpainting. Thereby, automating the process and reducing the interaction are required by the user. The only interaction required by the user is the selection of the region of the image to be removed.

2 Related Work

Many new compression techniques have been developed by utilizing different features within images to achieve high coding performance. Among those, an image inpainting is a practical approach to be employed in image compression. Inpainting was first introduced into digital image processing by Bertalmio et al. [1], as a process to restore damaged regions of an image so that the image appears normal to the not familiar observers. In this technique, the information is gathered over the boundary of the inpainting domain. Then, image smoothness, estimated by Laplacian operator, is propagated along isophote (i.e., a line whose points have the same gray value) directions with anisotropic diffusion. The region to be removed can be marked by the user depending upon on the subjective selection of the area. In practice, the area can be selected by using any image manipulation software or by free-hand selection and is given a different color. The user-selected area is usually called as the mask or the region to be inpainted. The missing information in the image is then filled by propagating the pixels inward from the boundary of the missing region. This application is useful in object removal [2]. Subsequently, several mathematical models of inpainting have been presented [3, 4] to adequately reconstruct regions that contain both structure and texture. Results of image inpainting [5, 6] also demonstrate that it can naturally recover structural regions as well as texture regions.

To employ inpainting in compression, the approach is to directly drop some homogeneous regions in the encoding but fill them up in the decoding by inpainting, as proposed in [5]. By not sending the significant portion of the image, amount of bits needed to transmit the image can be significantly reduced. We can think of the transmitter intentionally not sending certain block of image and the receiver filling in the missing region upon reception of the incomplete picture.

At the encoder, the image is analyzed at block level and then divided into featured and non-featured blocks. For the majority of featured blocks, assistant parameters instead of pixel values are coded and transmitted [7]. A few featured blocks (selected as exemplars) as well as the non-featured blocks are coded using traditional DCT-based methods. At the decoder side, dropped regions are recovered by PAI [8], relying on both delivered parameters and reserved regions. Due to the assistant parameters, the reliability of inpainting is greatly enhanced. Since large regions can be dropped during encoding, the scheme achieves a higher coding performance when compared with traditional image compression methods at similar visual quality.

In our proposed approach, we have used the combination of DCT and image inpainting to improve the compression performance. For this method to be successful, we have chosen the blocks that can be easily restored. Since inpainting is employed in compression, it is desirable to drop as many regions as possible to save coding bits while still maintaining good visual quality. The results obtained from the approach for the improvement of the compression are compared with the results obtained using DCT alone.

3 Proposed Algorithm

The algorithm used for image compression using inpainting consist of the following steps:

1. Select the blocks to be removed.
2. Apply DCT compression.
3. Fill-in removed blocks using image inpainting technique.

At the encoder side, steps 1 and 2 are followed.
At the decoder side, step 3 is followed. The steps are summarized in Fig. 1.
The three algorithms will be described in the following order.

1. *Block Removal*: Our goal is to remove as many blocks as possible and still recover the image with good perceptual quality. Remove the blocks in smooth area so that they can be restored back using structure inpainting. Such blocks are not noticeable even if we simply fill-in DC value. We can apply structure inpainting if a block falls into one of the two cases, when a block does not contain strong edges and when it is not composed of fine repetitive patterns. Sharp (strong) edges are critical when human recognize the shape of an object.

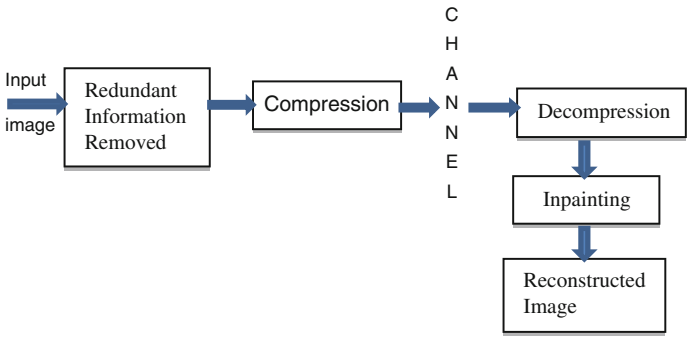


Fig. 1 Compression achieved through DCT and inpainting

2. *Discrete Cosine Transform*: The widely used JPEG image compression standard use Discrete Cosine Transform (DCT). It is a widely used and robust method for image compression [9]. It has excellent compaction for highly correlated data. DCT has fixed basis images, and it gives good compromise during a step called quantization where part of compression actually occurs. The less important frequencies are discarded, hence, the use of term lossy. The following is the general overview of the compression using DCT.

a. The input image is broken into 8 × 8 blocks.
b. Working from left to right, top to bottom, Forward Discrete cosine transform (FDCT) is applied to each block by multiplying with DCT matrix T (shown below) on the left and transpose of DCT matrix on its right.
c. Each block is compressed through quantization.
d. The array of compressed blocks that constitute the image is stored in a drastically reduced amount of space.
e. The image is reconstructed by a process that uses Inverse Discrete Cosine Transform (IDCT).

The formula given below is used to for the computation of FDCT and IDCT coefficients in each sub-block:

$$T_{i,j} = \begin{cases} \frac{1}{\sqrt{N}} & \text{if} \quad i = 0 \\ \sqrt{\frac{2}{N}} \cos\left[\frac{(2j+1)i\pi}{2N}\right] & \text{if} \quad i > 0 \end{cases} \tag{1}$$

The formulae used to compute FDCT and IDCT are as shown, respectively;

$$F = T \times X \times T$$

Here, F is the 8 × 8 sub-blocks containing DCT coefficients, and X is the 8 × 8 sub-block of the original image matrix.

$$A = T' \times S \times T$$

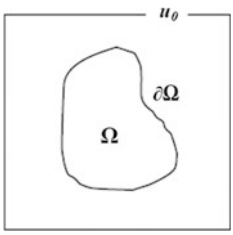


Fig. 2 The image, the region Ω to be inpainted, and its boundary $\partial\Omega$

Here, A is the sub-block containing IDCT coefficients, and S is the sub-block containing quantized coefficients.

3. *Digital Inpainting*: Digital inpainting helps to perform inpainting digitally through image processing in some sense. Thereby, automating the process and reducing the interaction required by the user. The only interaction required by the user is the selection of the region of the image to be removed [10]. The user can select an area through a free-hand selection or polygon selection. In practice, the area is selected using any image manipulation software and is given a different color. After the user selects the regions to be removed, the remaining portion of the image is transmitted. The inpainting algorithm automatically fills-in these regions with information surrounding them. The fill-in is done in such a way that isophote lines arriving at the regions' boundaries are propagated inside. This is automatically done (and in a fast way), thereby allowing to simultaneously fill-in numerous regions containing completely different structures and surrounding backgrounds.

Now, let Ω denote the set of pixels (the region) of the image to be inpainted. Let $\partial\Omega$ denote the pixel boundary of Ω so that $\partial\Omega \subset \Omega$ as shown in Fig. 2.

Partial differential equations (PDEs) are used for a large variety of image processing tasks, and recently, they have been proposed for so-called inpainting techniques, which use PDE-based interpolation methods to fill-in missing image data from a given inpainting mask. In inpainting, we find an information about the region to be redefined. This information is obtained from the regions surrounding the missing area. The following steps describe the general solution to the problem:

Step 1: SPECIFY Ω
Step 2: $\partial\Omega$ = THE BOUNDARY OF Ω (Region outside Missing Area)
Step 3: INPAINT ALL PIXELS IN Ω BASED ON INFORMATION in $\partial\Omega$

Let us consider a 4×4 sub-block $f(x, y)$ from the received image as shown below.

$$f(x,y) = \begin{matrix} z_1 & z_2 & z_3 & z_4 \\ z_5 & z_6 & z_7 & z_8 \\ z_9 & z_{10} & z_{11} & z_{12} \\ z_{13} & z_{14} & z_{15} & z_{16} \end{matrix}$$

The image has the pixel values as shown below.

$$f(x,y) = \begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 2 & 4 \\ 2 & 2 & 2 & 3 \\ 4 & 3 & 2 & 1 \end{matrix}$$

In this image, a redundant information characterized by uniformly distributed gray level is removed as shown below.

$$f(x,y) = \begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 3 \\ 2 & 0 & 0 & 3 \\ 4 & 3 & 2 & 1 \end{matrix}$$

The information about the pixels z_6, z_7, z_{10} and z_{11} has been totally lost. If the symbol ? stands for the corresponding gray scale value of the pixel unknown, these pixels are redefined as NaN (Not a Number) elements. Points for which the values are not fixed (i.e., missing data) are termed as interior points. The elements z_6, z_7, z_{10} and z_{11} are interior points.

Then the matrix Mask representing the above image is given by:

$$\text{Mask} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & ? & ? & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In an image, the pixels for which there exist a fixed value (i.e., outside the missing domain) are termed as boundary points. The set of boundary points provides a Dirichlet boundary condition. The problem of finding harmonic function subject to its boundary conditions is called the Dirichlet problem. The Dirichlet problem can be stated as follows:

> Given a function f that has values everywhere on the boundary of a region in R^n, is there a unique continuous function u twice continuously differentiable in the interior and continuous on the boundary, such that u is harmonic in the interior and u = f on the boundary. This requirement is called the Dirichlet boundary condition.

Finding a harmonic function that satisfies the boundary conditions may be viewed as a method for finding values on the interior of the volume that interpolate between the boundary values in the smoothest possible fashion. Continuous function twice continuously differentiable in the interior is approximated in terms of finite difference.

$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$, where the function $f(x,y)$ describes a two-dimensional matrix. The discrete form of Laplacian operator in x and y direction is defined in Eq. 2 as,

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \quad (2)$$

The matrix form of equation 2 is as shown in Fig.3 as matrix D and is done for all unknown points.

$$\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$$

To find the values of interior points, the solution of equations is obtained for X as

$$D \cdot X = R$$

D represents connection of an interior point to the neighbors and to itself. The interior points are connected to themselves with '4', and they are connected to neighboring points with '−1'.

Where X is n × m × 1 row matrix whose elements are taken column-wise from X, and each row corresponds to one pixel.

R in the above equation is obtained from the pixel values at the boundary of the inpainting domain. Therefore, we can write the linear equation form using,

$$D \cdot X = R \text{ as } X = R \backslash D$$

The set of linear equations are solved to finally obtain the values for NaN elements as shown below,

1	2	3	4
1	2	2.62	3
2	2.37	2.50	3
3	3	2	1

4 Performance Evaluation

Measurement of Image Quality:

The design of an imaging system should begin with an analysis of the physical characteristics of the originals and the means through which the images may be generated. A detailed examination of some of the originals may be necessary to determine the level of detail within the original that might be meaningful for a researcher or scholar. Generally, image quality is assessed from quality assessment parameters. The two commonly used measures for quantifying the error between images are mean square error (MSE) and peak signal-to-noise ratio (PSNR).

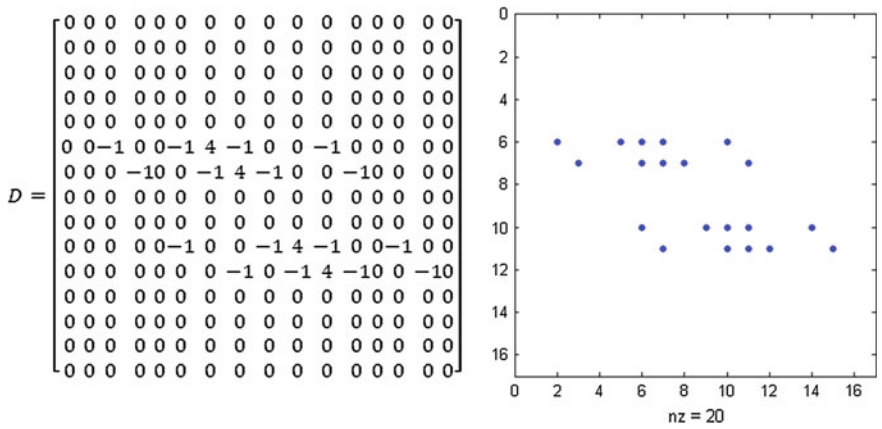


Fig. 3 Sparse matrix of sum of row and column partials of matrix **D**

Mean Square Error: MSE between the original image $\boldsymbol{I}$ and the reconstructed image $\boldsymbol{I}'$ is given by

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I(x,y) - I'(x,y)]^2 \tag{3}$$

It is very useful measure as it gives an average value of energy lost in the lossy compression of the original image I. A human observing two images affected by the same type of degradation will generally judge the one with smaller MSE closer to the original. A very small MSE indicates that the image is very close to original.

Peak Signal-to-Noise Ratio (PSNR): The PSNR between two 8-bit images (in dB) can be obtained using the formula, which tends to be cited more often, since it is a logarithmic measure, and our brains seem to respond logarithmically to intensity. Increasing PSNR represents increasing fidelity of compression. Generally, when the PSNR is 40 dB or larger, the two images are virtually distinguishable by human observers.

$$PSNR = 10\log_{10}\left[\frac{255^2}{MSE}\right] \tag{4}$$

A very logical way of measuring how well a compression algorithm compresses a given set of data is to look at the ratio of the number of bits required to represent the data before compression to the number of bits required to represent the data after compression. This ratio is called as *compression ratio*.

Fig. 4 **a** Original image. **b** Redundant information removed. **c** Inpainted image for compression ratio of 2. **d** Inpainted image for compression ratio of 4 (error found at *right top*). **e** Inpainted image for compression ratio of 4 and corrected with one iteration. **f** Inpainted image for compression ratio of 8 and corrected with one iteration. **g** Inpainted image for compression ratio of 16 and corrected with one iteration. **h** Inpainted image for compression ratio of 32 and corrected with one iteration

5 Results

The sample image of cameraman is used for the simulation.

An increase in the compression performance can be obtained by removing the non-significant information in the image prior to compression and then performing image inpainting at the receiver. The combination of DCT and inpainting is applied to cameraman.tif image (512 × 512) with various compression ratios.

In the first step of our approach, we have removed the blocks in smooth area so that they can be restored back using structure inpainting. It is shown in Fig. 4b wherein the blocks in smooth area are removed so that they can be effectively recovered back using inpainting at the decoder side. The amount of compression achieved by the combination of DCT and inpainting is better than that of using DCT only. This is indicated in Table 1 which illustrates that the amount of compression achieved is more than the compression value specified.

Figure 4a through Fig. 4h shows the results for various compression ratios obtained with the combination of inpainting and DCT. These figures shows the results obtained with various compression ratios applied to cameraman image.

Figure 4c is an image constructed with a compression factor of 2 which demonstrates that the image is recovered without any error. However, to reconstruct the image with a compression factor of 4 and higher value, we notice that a

Table 1 Quality assessment table for an image using DCT with inpainting

Compression ratio	DCT with inpainting				
	Compression achieved with DCT and inpainting	Mean square error (MSE)	Peak signal-to-noise ratio (PSNR)	Percentage increase in compression	Bits per pixel
2:1	2.1305	2.1817	44.7428	6.5251	3.76
4:1	4.2338	15.2252	36.3052	5.8458	1.89
8:1	8.5115	82.0470	28.9902	6.3937	0.94
16:1	16.9997	151.5088	26.3264	6.2881	0.47
32:1	33.9466	315.3520	23.1428	6.0831	0.24

kind of distortion appearing in the missing region as shown in Fig. 4d due to higher compression factors. The error is removed by using an iteration of inpainting applied to such images.

The re-inpainted images are as shown in the figures from Fig. 4e through Fig. 4h.

In Table 1, the quality assessment parameters like MSE and PSNR are tabulated with different compression ratios. Table 1 illustrates the various quality assessment parameters implemented using DCT with inpainting. From Table 1, we observe that the bits per pixel required for the proposed image compression method is less in comparison with the number of bits required to represent each pixel in the compressed image using DCT only.

6 Conclusion

A technique for the filling in of missing blocks in wireless transmission of compressed images is discussed. Here, we are dealing with filling in of missing blocks which are intentionally removed before the transmission over the channel. This approach could be utilized to reconstruct the incomplete image even though a large amount of redundant information is removed prior to transmission. As long as the features in the image are not completely lost, they can be satisfactorily reconstructed using computationally efficient image inpainting algorithms. This eliminates the need for retransmission of lost blocks. When the image resolution is increased, the quality of reconstruction improves and a retransmission request is rarely required, resulting in a better effective data-transmission rate. The future work in this field is to improve the compression, quality of the image, and to utilize the bandwidth efficiently. If the textured blocks are also removed prior to transmission, we can achieve better compression performance with the texture synthesis algorithms

References

1. Bertalmio M, Sapiro G, Caselles V, Ballester C (2000) Image inpainting. In: Proceedings of SIGGRAPH 2000, July 2000, pp 417–424
2. Criminisi A, Pérez P, Toyama K (2003) Object removal by exemplar based Inpainting. In: Proceedings conference computer vision and pattern recognition, Madison, WI, June 2003
3. Chan TF, Shen J, Vese L (2002) Variational PDE Models in Image Processing. UCLA Computational and Applied Mathematics Reports, pp 02–61, Dec 2002
4. Ballester C, Bertalmio M, Caselles V, Sapiro G, Verdera J (2001) Filling-in by joint interpolation of vector fields and gray levels. IEEE Trans Image Process 10:1200–1211
5. Rane SD, Sapiro G, Bertalmio M (2003) Structure and texture filling-in of missing image blocks in wireless transmission and compression applications. IEEE Trans Image Process 12(3):296–303
6. Bertalmio M, Vese L, Sapiro G, Osher S (2003) Simultaneous structure and texture image inpainting. IEEE Trans Image Process 12:882–889
7. Caia JF, Ji H, Shang F, Shenb Z (2010) Inpainting for Compressed Images. 11 Jan 2010 submitted to Elsevier
8. Xiong Z, Sun X, Wu F (2010) Block-based image compression with parameter-assistant inpainting. IEEE Trans Image Process 19(6), June 2010
9. Wallace GK (1991) The JPEG still picture compression standard. Commun ACM 34(4):30–44
10. Adolfson S (2004) Algorithms regarding automatic retouching of user selected regions in digital images. Master's Degree Project, Royal Institute of Technology, Stockholm

Biometric Vault Scheme Using Data Hiding and Standard Encryption

Manoj Krishnaswamy and G. Hemantha Kumar

Abstract In this paper, we propose a scheme of a biometric vault for enhancing the security of biometric systems. Many biometric systems are using traditional methods which are wide open to attacks. With increasing use of biometrics, there is a real threat for the conventional systems using databases, which store images of users in raw and unaltered form. If compromised, it is not only irrevocable, but can also be misused for cross-matching across different databases. By using passwords for user and biometric vault authentication, we can achieve a better level of security for the biometric system. Data hiding is performed by generating a key file and interjecting it to cancellable biometric data. For verifying data integrity, MD5 hashes are stored using standard AES-256 encryption cipher. Also to show that existing biometric matching systems will have acceptable alteration to be incorporated.

Keywords Cancellable biometrics · Biometric vault · Template security · Encryption · Data hiding

1 Introduction

Many citizens are concerned that their privacy will be invaded. Some fear that it could lead to a "total surveillance society", with the government and other authorities having the ability to know where you are and what you are doing at all times. This is not to be an underestimated concept as history has shown that states have typically abused such access before [1]. It is possible that data obtained during biometric enrolment may be used in ways for which the enrolled individual has not consented.

M. Krishnaswamy (✉) · G. Hemantha Kumar
Department of Studies in Computer Science, University of Mysore, Mysore, India
e-mail: manojkrishnaswamy@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_55,

One advantage of passwords over biometrics is that they can be reissued. If a token or a password is lost or stolen, it can be cancelled and replaced by a newer version. This is not naturally available in biometrics. If someone's face is compromised from a database, they cannot cancel or reissue it. Cancellable biometrics is a way in which we incorporate protection and the replacement features into biometrics. It was first proposed by Ratha et al. [2].

Several methods for generating cancellable biometrics have been proposed. The first fingerprint-based cancellable biometric system was designed and developed by Tulyakov et al. [3]. Essentially, cancellable biometrics performs a distortion of the biometric image or features before matching. The variability in the distortion parameters provides the cancellable nature of the scheme. Some of the proposed techniques operate using their own recognition engines, such as Teoh et al. [4] and Savvides et al. [5], whereas other methods, such as Dabbah et al. [6], take the advantage of the advancement of the well-established biometric research for their recognition front-end to conduct recognition. Although this increases the restrictions on the protection system, it makes the cancellable templates more accessible for available biometric technologies. By our earlier work [7], we have shown that by using a different Gaussian filter, cancellable templates can be generated without adversely affecting the system while maintaining matching accuracy for face biometric systems.

A cryptographic hash function is a hash function, that is, an algorithm that takes an arbitrary block of data and returns a fixed-size bit string, the (cryptographic) hash value, such that an (accidental or intentional) change in the data will (with very high probability) change the hash value. The data to be encoded are often called the "message", and the hash value is sometimes called message digest or simply digest. The MD5, message digest algorithm, is a widely used cryptographic hash function that produces a 128-bit (16-byte) hash value, specified in RFC 1321 [8]. MD5 has been employed in a wide variety of security applications and is also commonly used to check data integrity. An MD5 hash is typically expressed as a hexadecimal number, 32 digits long.

The Advanced Encryption Standard (AES) is a specification for the encryption of electronic data established by the US National Institute of Standards and Technology (NIST) in 2001 [9]. The algorithm described by AES is a symmetric key algorithm, meaning the same key is used for both encrypting and decrypting the data. AES is based on a design principle known as a substitution–permutation network and is fast in both software and hardware. We have used AES to generate key size of 256 bits (AES-256).

Biometric cryptosystems, as discussed by Dodis et al. [10], Hao et al. [11], Nandakumar et al. [12], Sutcu et al. [13], use techniques that associate an external key with a user's biometric to obtain helper data. The helper data should not reveal any significant information about the template or the key, and at the same time, it can be used to recover the key when the original biometric is presented. These systems are cumbersome to decode and verify for large datasets. Also the fuzzy data could overwrite existing data and lead to lesser matching accuracy. The concept of data hiding in digital watermarks has been discussed by Hsu and

Wu [14]. Encryption algorithm to secure the image using fingerprint and password has been discussed by Kaur et al. [15], which involves more time-consuming methods. Those methods embed secret data inside the biometric image itself.

We have approached the system in a different manner, mainly where we embed the transformed biometric image data inside an encrypted file. We also store the information which helps to check for data integrity and provides a check against tampering.

2 Proposed Method

Figure 1 shows the generic overview of the proposed method where the biometric vault holds the transformed templates and additional data. For enrolment process which is shown in Fig. 2, we start off with the usual biometric scan of the user and use its cancellable template, as demonstrated in our previous work on face biometrics [7]. A key file is generated by a polynomial function, which uses the user password. The MD5 hash digest is calculated for the cancellable template and is encrypted by using standard AES-256 bit cipher. The key file and the cancellable template are processed in their binary form, where we interject bits from cancellable template into specific locations in the key file. We thus stream stitch the files to obtain the transformed template. The MD5 hash digest is calculated for the transformed template and encrypted using AES-256 bit cipher and stored in the vault tagged with the associated user-id, along with the previously calculated and encrypted MD5 hash digest of the cancellable template. So for a particular

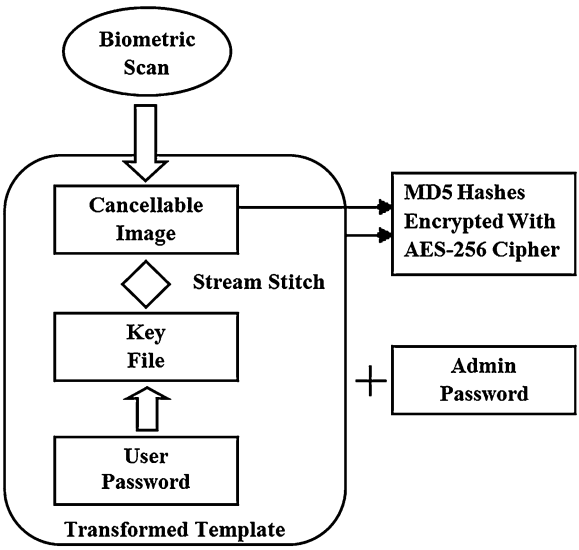


Fig. 1 Generic overview of the proposed method

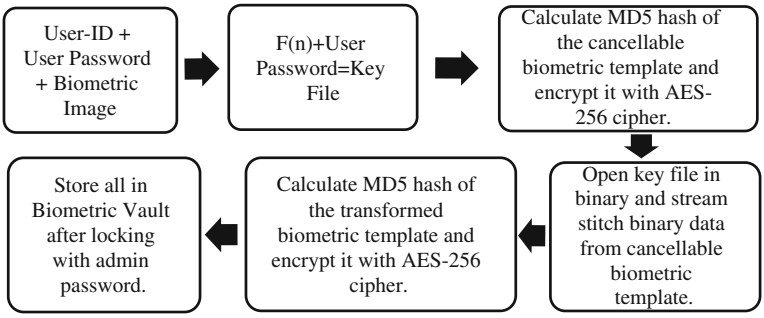


Fig. 2 Block diagram of proposed method for enrolment

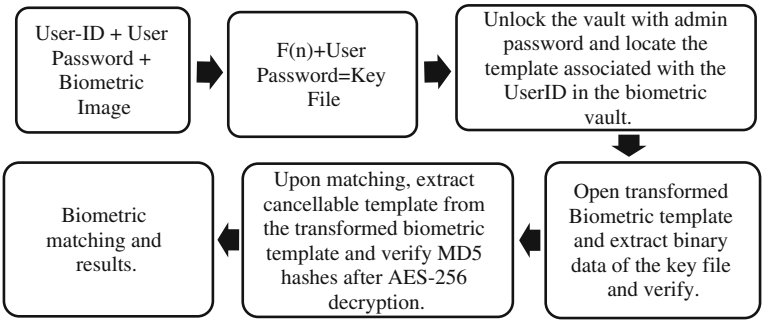


Fig. 3 Block diagram of proposed method for verification

user, the vault will hold the stream-stitched data and the encrypted hash values. The biometric vault is then locked with an admin password to further enhance security.

For verification which is shown in Fig. 3, we have the user-id, user biometric scan and user password. The admin password unlocks the biometric vault, and the same function, as used in enrolment, is applied with the user password to generate a temporary key file. The associated user contents from the vault are then opened for processing. The stored stream-stitched data are then used to extract the key file and verified with the temporary key file. Upon matching, the cancellable template is extracted and the MD5 hashes are verified after AES-256 cipher decryption. This process allows not only for matching user password but also for checking the data integrity of the stored cancellable templates. The extracted cancellable template and the biometric scan are then sent for biometric analysis and verification. Both the proposed enrolment and verification processes are more clearly elaborated below.

2.1 Enrolment

Step 1: User provides user-id, password and biometric scan.
Step 2: Apply function $F(n) = D_1W + D_2X + D_3Y + D_4Z$, where D_1, D_2, D_3, D_4 correspond to the 4-digit user password. W, X, Y, Z can be any value, which is used to generate the key file of a particular length, here in our case 512 K bits of key file data.
Step 3: Calculate and store the MD5 hash value of the cancellable biometric template after encrypting with user password by using AES-256 cipher.
Step 4: Perform stream-stitching by opening the key file in binary stream and append the binary data from cancellable biometric template. At every 256 bits of the key file, insert 128 bits of the cancellable biometric template until the bits of cancellable template are exhausted. This process of stream-stitching or interjecting the bits at defined locations can also be varied if needed, that is, instead of inserting at 256th bit, any other position can be chosen.
Step 5: Calculate and store the MD5 hash value of the transformed (stream-stitched) biometric template after encrypting with user password by using AES-256 cipher.
Step 6: Store all the data after locking with admin password in the biometric vault.

2.2 Verification

Step 1: User provides user-id, password and biometric scan.
Step 2: Apply the same function used in enrolment to generate a temporary key file. $F(n) = D_1W + D_2X + D_3Y + D_4Z$, where D_1, D_2, D_3, D_4 correspond to the 4-digit user password.
Step 3: Locate the transformed (stream-stitched) template associated with the user-id in the biometric vault after unlocking with admin password.
Step 4: Open transformed template from the vault and extract binary data of the key file and verify with the temporary key file as obtained in step 2. If not equal, then the process ends here due to password mismatch or a tampered key file.
Step 5: Extract the cancellable template from the transformed (stream-stitched) template and calculate its MD5 hash value. Verify the MD5 hash of the reconstructed template and the stored MD5 hash (decrypted with AES-256 cipher) from the vault associated with the user-id. This helps to check against tampering of the stored biometric data and also to verify the extracted data.
Step 6: Send the input biometric scan and the extracted cancellable template for biometric analysis and verification.

3 Results and Discussion

Figure 4 is a sample face image from the ATT dataset [16], which is unaltered. Figure 5 is the cancellable face image produced after a different Gaussian filter, derived from our earlier work on cancellable face biometrics [7].

A portion of the binary representation of the stream-stitched data is shown in Fig. 6. The shaded grey area (5th row from above) shows the interjected bits of the cancellable biometric template, with rest of bits are from the generated key file.

Table 1 shows the MD5 hash values, with and without AES-256 encryption, for the cancellable template (Fig. 5) and its corresponding transformed (stream-stitched) template (Fig. 6). The AES-256 cipher uses the key (user password = 1234) to encrypt or decrypt the MD5 hashes. The decrypted MD5 hash value is same as for the unencrypted MD5 hash, in case of the vault contents being secure.

The novelty of our proposed scheme is the way the key file and the cancellable template are stream-stitched in binary form. Also, we have proposed to use AES-256 cipher and MD5 hashing method. The MD5 hash values provide a check

Fig. 4 Sample face image from ATT dataset

Fig. 5 Cancellable face template

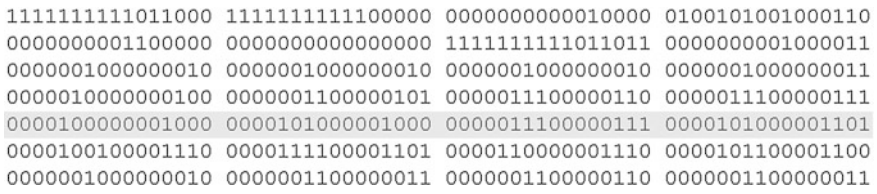


Fig. 6 Portion of key file in binary representation stream-stitched with biometric image

against the data integrity of the cancellable biometric template and the transformed biometric template, thereby providing dual functionality. By varying the parameters of the function F(n), different key file data can be produced, in case of systems requiring the same user in many biometric systems. Several existing methods are susceptible to error bits among other problems. Error bits are non-existent in our proposed system, and cancellable biometric template data are kept intact.

The actual data of the cancellable biometric template are opened in binary stream along with encrypted key file, and the bits are stream-stitched as explained earlier. This approach allows the use of standard encryption methods (in our case AES-256), which are proven to be secure and computationally fast. Since the actual data of the cancellable biometric template can be extracted from the known locations of the key file, it can be easily replaced in case of compromise. Also in that case, due to the non-invertible nature of the cancellable biometric template, cross-matching across different biometric systems can be prevented. Hence, even though raw biometric scans can also be used in the proposed scheme, we have emphasized the usage of cancellable templates.

The user-id is needed for accessing the template directly, to aid in speed of execution. If user-id is not used, the contents of the vault have to be decrypted as many times as the stored number of templates in the vault. Hence, speed of execution is reduced in such a case. So to further enhance the level of security, we have proposed the use of another password (admin password) which is known only by the concerned authority. By further locking the entire vault, after data hiding, with the admin password, we can provide an additional layer of security for the biometric vault.

In case only the user password is compromised or forgotten, a new user password can be generated and F(n) reapplied to generate a new key file and stream-stitched at the known location of the interjection bits. Here, the admin password provides verification to interject data. Also, reenrolment will not be required. But in case of compromising only the admin password, it can be changed easily. Hence, it would be even better if the admin password is changed periodically with an acceptable time frame. To break into the system, the user password, admin password, and the locations of the data-hidden bits of the cancellable template all have to be compromised. Even in such a case, the proposed scheme provides revocability and disposability at various levels and a new set can be generated without too much hassle and worrying about lost information.

Table 1 MD5 hashes of cancellable biometric template and transformed template

	Cancellable template	Transformed template
MD5 Hash unencrypted:	A240E8DC57BBE7AC9543D9B68F8EFA8D	E2E64D98E73A3D9CE71F09B322398D92
MD5 Hash encrypted with AES-256 bit (key = 1,234) in HEX representation	1B98CD90466F48C7CEB48A315B42062066 959CF189DDB1C7D795E816BCF2C05 C0B5A399E114513FB834D6DDC14D5 A31588EC9DAC39DC6A5A4FFD6038 F063340FE5C0AF7130665D500B789 60346995888	CEDD65CD9219ADCFB5F6D6640E234EE 3AF4AE829E3B2B47ECC0096180F8E75 D8D5F53998D64EC82BD6CE960E7AB80 31A5FB6C10AF32D54EF3884F3CDB56130 AE5752913344265034FB50570516D5185C

The proposed biometric vault system is feasible for ATM-like systems and others where user information is provided along with password and biometric scan. Due to the scheme of the vault, the speed of execution is greatly enhanced in systems which have thousands of users. Any biometric type (face, fingerprint, iris, etc.), single- or multi-modal biometric template, can be used in the proposed scheme.

4 Conclusion

We have proposed a biometric vault system with main considerations on security and integrity of the stored information. Our proposed scheme of the biometric vault seems novel and unique. The proposed system can be used in systems requiring fast verification at multiple levels with some drawbacks as discussed. The proposed scheme focuses on securing the user's cancellable biometric templates by data hiding using binary stream-stitching of the templates with the helper data derived from user inputs. By using MD5 hashes to verify vault contents against tampering, we have used the same information also to ensure that the right information is extracted for verification and without any error bits. Also, by using proven and standard cryptographic method (AES-256) for data encryption, the vault is further enhanced. The proposed scheme which uses the process of stream-stitching, MD5 hashing algorithm, and AES-256 cipher method are computationally fast. Hence, it can be incorporated into existing systems without too much hassle.

References

1. Civil Liberties and Facial Recognition Software. About.com, The New York Times Company. pp. *pp*. 2, Archived from the original on 2006-03-01, Retrieved 2007-09-17 (2007)
2. Ratha NK, Connell JH, Bolle RM (2001) Enhancing security and privacy in biometrics-based authentication systems. IBM Syst J 40:614–634
3. Tulyakov S, Farooq F, Govindaraju V (2005) Symmetric hash functions for fingerprint minutiae. In: Proceedings of int'l workshop pattern recognition for crime prevention, security, and surveillance, pp 30–38
4. Teoh ABJ, Goh A, Ngo DCL (2006) Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. IEEE Trans Pattern Anal Mach Intell 28:1892–1901
5. Savvides M, Kumar BVKV, Khosla PK (2004) Corefaces- robust shift invariant PCA based correlation filter for illumination tolerant face recognition. Paper presented at IEEE computer society conference on computer vision and pattern recognition (CVPR'04)
6. Dabbah MA, Woo WL, Dlay SS (2007) Secure authentication for face recognition. Paper presented at IEEE symposium on computational intelligence in image and signal processing, CIISP 2007
7. Hemantha Kumar G, Krishnaswamy M (2011) Cancellable face biometrics using image blurring. Int J Mach Intell 3(4/5):272
8. RFC 1321 section 3.4, Step 4. Process Message in 16-Word Blocks, p 5

9. Announcing the advanced encryption standard (AES) United States national institute of standards and technology (NIST) (2001) 26 Nov 2001, Retrieved 2 Oct (2012)
10. Dodis Y, Ostrovsky R, Reyzin L, Smith A (2006) Fuzzy extractors: how to generate strong keys from biometrics and other noisy data. Tech Rep 235, Cryptology ePrint Archive
11. Hao F, Anderson R, Daugman J (2006) Combining crypto with biometrics effectively. IEEE Trans Comput 55(9):1081–1088
12. Nandakumar K, Jain AK, Pankanti S (2007) Fingerprint-based fuzzy vault implementation and performance. IEEE Trans Inform Forensics Secur 2(4):744–757
13. Sutcu Y, Li Q, Memon N (2007) Protecting biometric templates with sketch: theory and practice. IEEE Trans Inform Forensics Secur 2(3):503–512
14. Hsu CT, Wu JL (1999) Hidden digital watermarks in images. IEEE Trans Image Process 8(1):58–68
15. Kaur M, Dr. Sofat S, Saraswat D (2010) Template and database security in biometrics systems: a challenging task. Int J Comput Appl 4(5):0975–8887, July 2010
16. AT&T Laboratories Cambridge, The Database of Faces. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

Voltage Stabilization Through Reactive Power Injection at Secondary Terminals of Distribution Level Feeders—Using Thyristor-Switched Capacitor

B. Preethi and B. V. Sumangala

Abstract This paper presents the analysis of the FACTS device—thyristor-switched capacitor (TSC) connected on the secondary terminals of the distribution feeder (transformer) in a power system network for effective voltage stability and reactive power control. The FACTS device is advantageous over the mechanical switch operated capacitors in the aspects of faster control, improved system performance, and system voltage stability. The power system network simulations are carried out, and the model of the TSC is realized for a 220 kV transmission line system using Simulink in MATLAB.

1 Introduction

The expansion in power transmission system has taken place not only due to the increase in generation and loads but also due to the extensive interconnection among different power utilities. The major factor for system interconnections is to economize the generation reserves, to minimize generation cost, and also to achieve the reliability of supply. The demand for electricity increases which is proportional to growth in population and creation of special economic zones, urbanization, and industrialization. Setting up of a new generation station is not feasible because of economic and environmental considerations. So the challenge of meeting the demand for electrical energy can be achieved by improvising the system performance of existing power system network. This in turn is achieved by

B. Preethi (✉)
Department of EEE, Global Academy of Technology, Bangalore, India
e-mail: preeths_b@yahoo.com

B. V. Sumangala
Department of EEE, Dr. Ambedkar Institute of Technology, Bangalore, India
e-mail: sumangala_bv@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_56,

minimizing power losses in the transmission and distribution system, effective reactive power control, and tapping alternative generation resources and by implementation of energy management systems.

Except for the limited number of HVDC links in power system, the vast majority of transmission lines are AC. The AC transmission network requires dynamic reactive power control to maintain satisfactory voltage profile under varying load conditions and transient disturbances. In case of mismatch in the reactive power balance in the system, this can result in voltage instability and collapse.

As a commitment by the regulation of the Electricity Act, it is required to maintain rated constant voltage, constant frequency power supply to the consumers with system security. The consistency of these parameters mainly depends on the characteristics of the load, majority of which are inductive in nature. This characteristic of inductive loads reflects with dip in voltage. Also there is reduction in operating power factor, which further results in increased power losses. This in turn results in degradation in the power system performance and also the consumer getting penalized for low power factor.

In the conventional systems, load conditions vary throughout the day. There is either overvoltage due to lightly loaded condition or overloaded condition during peak hours of load consumption, which results in unbalanced load. The unbalanced load in each phase of the load feeder results in frequent failure of distribution transformers. These voltage variations also affect the power flow performance of the grid. In order to stabilize the load end voltage to the rated value, effective reactive power management system has to be incorporated at the secondary terminals of the distribution transformer.

Reactive power flow involves generation of additional reactive power. Reactive power transmission over long distances results in voltage drops and losses. Hence, reactive power transmission is not only uneconomical but also ineffective from the control point of view. The balance between reactive power generation and demand has to be maintained on a regional basis within the area of operation concerned.

Reactive power demand varies based on the load variations. When the demand is low, there will be excess reactive power, that is, capacitive reactive power available and the voltage magnitude would increase. So it is necessary to connect parallel reactors for consuming the additional capacitive reactive power of the lines or else to change the on-load transformer tap changer settings. During peak loads, there is large voltage drops and also the inductive reactive power of the lines adds to the reactive power flow. Hence, additional capacitive reactive power has to be injected into the system to meet additional reactive power requirements.

Reactive power (var) compensation or control [1] is an essential part in a power system to minimize power transmission losses, to maximize power transmission capability, to stabilize the power system, and to maintain the supply voltage. Reactive power compensation of AC lines using fixed series capacitors can solve some of the problems associated with AC networks [2, 3]. However, the slow nature of control using mechanical switches (circuit breakers) and limits on the

frequency of switching imply that faster dynamic controls are required to overcome the problems of AC transmission networks.

Flexible Alternating Current Transmission Systems (FACTS) [4, 5] are new devices emanating from recent innovative technologies that are capable of altering voltage, phase angle, and/or impedance at particular points in power systems. Their fast response offers a high potential for power system stability enhancement apart from steady-state flow control. FACTS involves the application of high-power electronic controllers in AC transmission networks which enable fast and reliable control of power flows and voltages. The objectives are as follows:

1. Regulation and enhancement of power flow by reducing transmission losses in prescribed transmission routes.
2. Improve stability with fast-acting voltage regulation.
3. Secure loading of lines nearer to their thermal limits.
4. Damping of oscillations which can threaten security or limit the usable line capacity.
5. Higher transient stability limit.

There are several FACTS controllers which have been developed or proposed. These include the following:

1. Static var compensator (SVC).
2. Thyristor-controlled series capacitor (TCSC).
3. Thyristor-controlled phase angle regulator (TCPAR).
4. Static synchronous compensator (STATCOM) or static condenser (STATCON).
5. Unified power flow controller (UPFC).

2 Static Var Compensator

SVC [6, 7] are devices which control the voltage at their point of connection to the power system by adjusting their susceptance to compensate for reactive power deficiencies. SVC provides fast-acting dynamic reactive compensation for voltage support during contingency events. SVC also dampens power swings and reduces system losses by optimized reactive power control.

These compensators comprise of either thyristor-switched capacitors (TSCs) and/or thyristor-controlled reactor (TCR) with fixed (permanently connected) power factor correction capacitors, which also provide, when combined with appropriate tuning reactors, harmonic filtering. SVC schemes on the secondary side of a coupling transformer provide voltage support, increased transient stability, and improved damping.

Functionally, the thyristor-controlled SVCs are controllable reactive impedances. Essentially, the thyristor valves control the current drawn by the capacitor and reactor banks and thereby the reactive power they generate for or absorb from

the AC system. Thus, an SVC in addition with the necessary coupling transformer requires fully rated capacitors and reactors to generate or absorb reactive power and an electronic switching circuit (thyristor valves) accounting for a large system.

2.1 SVC: Characteristics and Operation

A shunt connected static var generator or absorber whose output is adjusted to exchange capacitive or inductive current so as to maintain or control specific parameters of the electrical system, typically the bus voltage by means of providing required reactive compensation.

In such installations, AC capacitors are used as the var source and shunt-connected phase-controlled reactors are used to absorb vars [4, 6]. The net reactive power is the difference between the two. A conventional SVC employs TSCs with valves rated at two times the peak line voltage seen by the capacitors and phase-controlled reactors with valves rated at the peak line voltage for the reactors (Fig. 1).

The characteristic of the SVC [8] is as shown in Fig. 2. The normal characteristics represented by load line A will intersect the SVC characteristics at reference voltage. The load line B intersects the SVC characteristics at voltage V2. Since voltage V2 is above the reference voltage, it indicates inductive reactive power requirement. The current I1 corresponding to voltage V2, hence, is in the inductive region, and it requires operation of thyristor-controlled reactor. The load line C intersects the

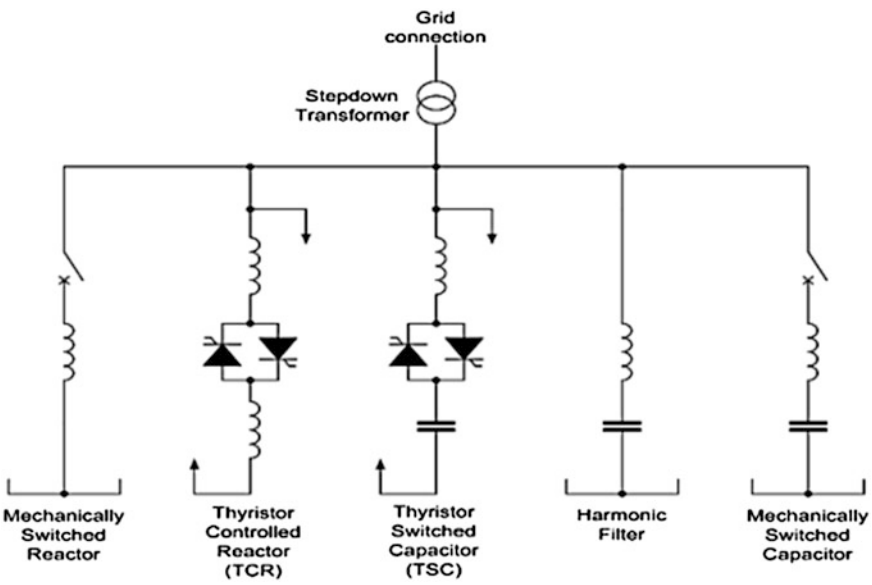


Fig. 1 Static var compensator

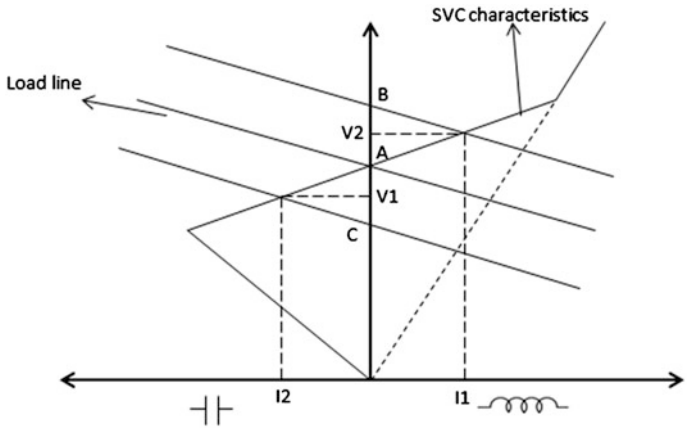


Fig. 2 Composite SVC characteristics

SVC characteristics at voltage V1. As it can be observed that the voltage V1 is below the reference voltage, and hence, capacitive reactive power is required. The current I2 is in the capacitive region of operation.

2.2 Thyristor-Switched Capacitor

The representation of thyristor-switched capacitor is shown in Fig 3.

The three thyristor valves are connected in delta in order to eliminate zero sequence triple-n harmonics (3rd, 9th…). The harmonics remain trapped inside the delta and thus reduce harmonic injection into the power system. Reactor L is used to limit the surge current when thyristor is ON under abnormal operating conditions.

Advantages of TSC [9] over mechanically switched capacitors.

1. TSC provides instantaneous response to the changes in the system parameters.
2. Faster control is achieved by the switching of thyristor-operated devices.
3. TSCs are more reliable compared to the mechanically switched capacitors.
4. Permanent connection of fixed capacitors in the network draws unnecessary charging current.

3 Simulation Details

The system under consideration for simulation is given in Fig. 4. The simulations are carried out using Simulink of MATLAB software.

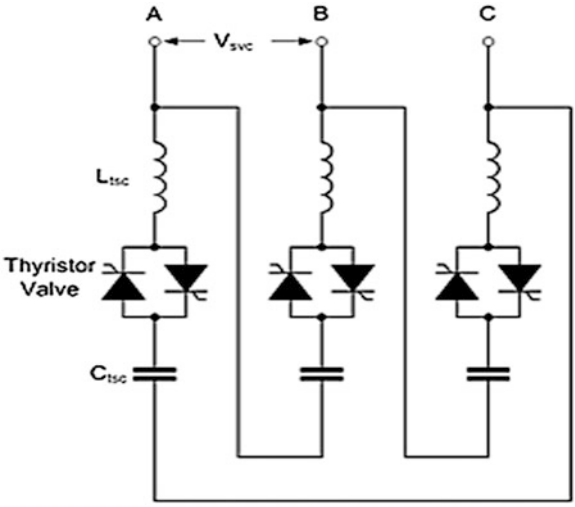


Fig. 3 Thyristor-switched capacitor

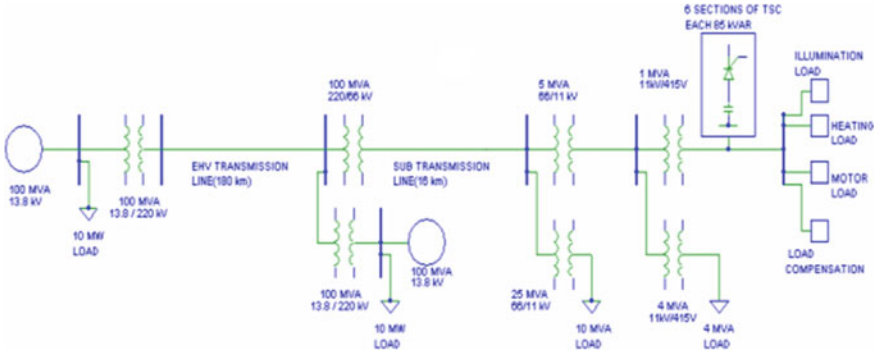


Fig. 4 Radial transmission line power system network

With reference to the Fig. 4, the loads connected at the secondary side of distribution transformer represent the illumination, heating, motor, and capacitive load. The capacitive load represents the load compensation provided at the consumer end (e.g., industrial customers). The load compensation at the consumer end is not ensured always, the reason being failure of the reactive compensation provided at consumer end or absence of the compensation due to lack of knowledge. The results are obtained in the absence of this load compensation, which brings out the significance of the proposed TSC controller. The TSC controller connected at the secondary side of the distribution transformer stabilizes the load end terminal voltage at rated value irrespective of the presence of load compensation at consumer end which will be referred to as load compensation (existing). TSC is symbolically represented in simulation as two antiparallel thyristors and a capacitor connected in series in each phase.

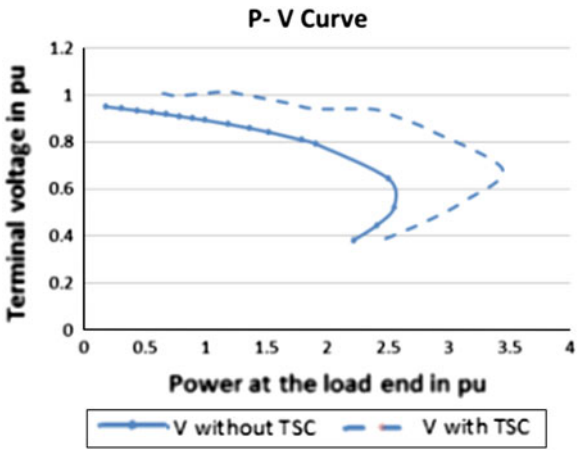


Fig. 5 P–V curve

3.1 Voltage Profile

The P–V curve with TSC shows a significant improvement in the tune of nearly 40 % of the power handling capability (nose point) of the transmission line. With the incorporation of the TSC, it is shifted further, which shows the improvement in the power handling capability [7] of the system (Fig. 5).

3.2 Operation of TSC Controller

Power factor is taken as the reference for improvement in voltage profile. The load end power factor is compared with a reference value (in the simulation taken as 0.98), and the TSCs are switched ON in steps till the load end power factor reaches the desired, that is, the reference value.

A daily load cycle consisting of residential and commercial load is considered as shown in Fig. 6.

4 Results: Voltage Profile

The simulation results for voltage profile without the TSC and with TSC connected along with load compensation are shown. The profile also compares the results where TSC is connected and load compensation (existing) is not provided which may be practically possible due to the failure of the compensation provided or absence of the compensation at consumer end.

There is good improvement in the tune of 0.9 pu in the voltage profile with the load compensation and with TSC connected (Fig. 7).

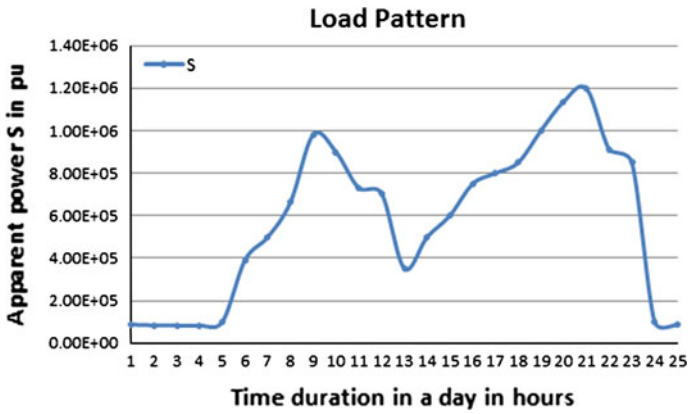


Fig. 6 Daily load cycle

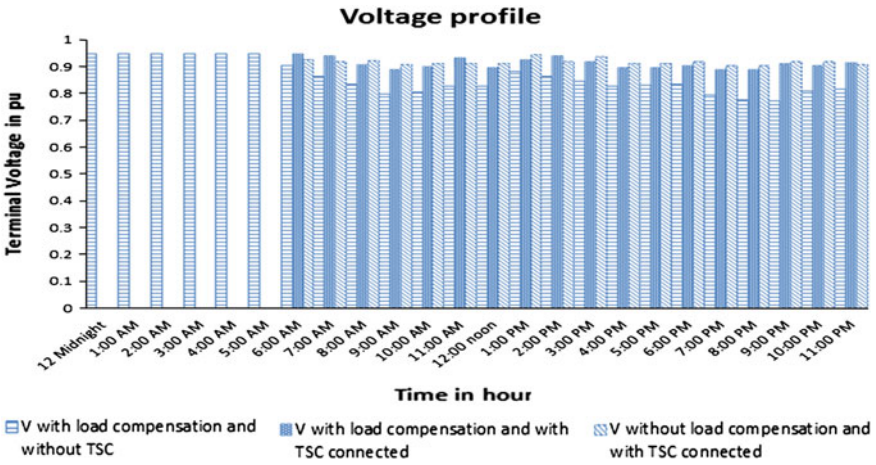


Fig. 7 Voltage profile

In case of without load compensation (existing) and with TSC connected, a further improvement in a value of around 0.93 pu is observed during the peak loads. Hence, the voltage profile is maintained nearly at 1 pu, and the improvement in the profile is nearly 20 %.

4.1 Details of Power Delivery

It can be inferred from the Table (1) that the apparent power deliverable at a particular load by the distribution transformer increases as the compensation is provided at the secondary terminals of the transformer. This is due to reason being

Table 1 Apparent power delivered at load end of transformer

Time	Apparent power S in pu	S with load compensation and TSC not connected	S with load compensation and with TSC connected	S without load compensation and with TSC connected
12 Midnight	0.100	0.0868		
3:00 A.M.	0.084	0.0682		
6:00 A.M.	0.393	0.3882	0.5799	0.5730
8:00 A.M.	0.665	0.5217	0.7394	0.9548
9:00 A.M.	0.982	0.7691	1.0110	1.2110
10:00 A.M.	0.900	0.5926	0.9998	1.1600
1:00 P.M.	0.352	0.2782	0.3793	0.6023
4:00 P.M.	0.750	0.5372	0.7713	0.9761
6:00 P.M.	0.850	0.7098	1.0080	1.2180
7:00 P.M.	1.000	0.7861	1.1490	1.3380
8:00 P.M.	1.137	0.7510	1.3260	1.5030
9:00 P.M.	1.198	0.9797	1.6000	1.7680
10:00 P.M.	0.912	0.6761	1.0920	1.3020
11:00 P.M.	0.850	0.5926	1.0380	1.0130

improvement in power factor with the incorporation of the reactive compensation and reduction in losses. Thus, the power deliverable by the transformer increases, which in turn increases system's efficiency.

5 Conclusion

A real 220 kV system simulated using MATLAB—Simulink. With the incorporation of TSC at the distribution transformer end, the following are the conclusions:

1. The voltage is maintained at constant rated value—almost 1 pu, thus achieving voltage stability with a variation of ±2 %.
2. With the improvement in power factor, power losses are reduced in turn improving the system's efficiency.
3. Fast and automatic control is achieved with the use of TSC compared to mechanical switches.

References

1. Blajszezak G (2007) Reactive power control in transmission network as a tool for reliable supply. In: Proceedings of 9th international conference, electrical power quality and utilisation, Barcelona, 9–11 Oct 2007
2. Su Y, Cheng S, Wen J (2006) Reactive power generation management for the improvement of power system voltage stability margin. In: Proceedings of IEEE, 1-4244-0332-4/06

3. Lazar ID (1980) Electrical systems analysis and design for industrial plants. McGraw-Hill, New York
4. Padiyar KR (2007) FACTS controllers in power transmission and distribution. 1st edn, New Age International Publication, June 2007
5. Hingorani NG Gyugyi L (2000) Understanding FACTS: concepts and technology of flexible AC transmission systems. In: Proceedings of IEEE Press 2000
6. Chopade P, Dr. Bikdash M, Dr. Kateeb I, Dr. Kelkar AD (2011) Reactive power management and voltage control of large transmission system using SVC (Static VAR Compensator). In: Proceedings of IEEE, 978-1-61284-738-2/11
7. Kundur P (2008) Power system stability and control. Tata McGraw-Hill, New York
8. Katira MJ, Porate K.B. (2009) Computer simulation of 132/11 KV distribution substation using static var compensator (SVC) for voltage enhancement—a case study. In: Proceedings of IEEE, 2009 978-0-7695-3884-6/09
9. Gelen A, Yalcinoz T (2009) The behavior of thyristor switched capacitor (TSC) installed in an infinite bus system. In: Proceedings of IEEE, 978-1-4244-3861-7/09

Exposing Digital Forgery in Video by Mean Frame Comparison Techniques

Govindraj B. Chittapur, S. Murali, H. S. Prabhakara and Basavaraj S. Anami

Abstract Nowadays, videos are core part of live entertainment in television and movies, and they are breathing of real entertainment world. People believe movies and video snaps in everywhere of digital media. Trusting such unbelievable video may create sensation over the news and gossip media world. Identifying such telecasted and forecasted video's truthfulness is challenging in multimedia forensic and video forensic area. In this regard, we propose a methodology which detects the forgery region frames from given input video by using mean frame comparison techniques.

Keywords Digital forgery · Video analysis · Video forensic · Video trustfulness · Mean frame

1 Introduction to Image and Video Forgery

Today, we are seeing doctored images and videos regularly. While these images might tarnish the public opinion of a celebrity, cases involving manipulated images with more serious implications have arisen in science and law. The art of making an image/video forgery is as old as photography itself. In its early

S. Murali · H. S. Prabhakara
Maharaja Institute of Technoloy, Mysore, India
e-mail: nymurali@yahoo.com

H. S. Prabhakara
e-mail: prabhassan@gmail.com

G. B. Chittapur (✉)
Basaveshwar Engineering College, Bagalkot, India
e-mail: govindraj_bec@yahoo.com

B. S. Anami
KLE Institute of Technology, Hubli, India
e-mail: anami_bassu@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_57,

years, photography quickly became the chosen method for making portraits, and portrait photographers learned that they could improve sales by retouching their photographs to please the purchaser [1]. Though photograph and video manipulation has become more common in the age of digital cameras and image editing software, it actually dates back almost as far as the invention of photography. For recent years, an exhaustive inventory of every photograph manipulation in video would be nearly impossible, so we focus here on the instances that have been most controversial or notorious, or ones that raise the most interesting ethical questions [2]. The photographers of the era also experimented with video composting, that is, combining multiple images into one and creating doctored videos.

Digital images/video offer many attributes for a tamper detection algorithm to take advantage of, specifically the color and brightness of individual pixels as well as an image's resolution and format. These properties allow for analysis and comparison between the fundamentals of digital forgeries in an effort to develop an algorithm for detecting image tampering in a video.

Two types of video forensics scheme are widely used for image/video forgery detection: active schemes and passive schemes. With the active schemes, the tampered region can be extracted using a watermark. However, this scheme must have source files to embed the watermark first; otherwise, the detection process will fail [3]. On contrary, the passive schemes extract some intrinsic fingerprint traces of image/video to detect the tampered regions.

In this paper, we are proposing passive forgery detection method in a digital video based on the statistical property of mean comparison. We propose to analyze the temporal difference of each frame set in given input video and successfully detect the tampered region by comparing each frame by frame residential to locate the tampered regions of video.

2 Related Work on Video Forgery

Currently, most acquisition and manipulation tools use the JPEG and MPEG standard for image and video compression. As a result, one of the standard approaches is to use the blocking fingerprints introduced by MPEG compression, as reliable indicators of possible image tampering. Not only do these inconsistencies help determine possible forgery, but they can also be used to detect the anomalies in adding or removing frames in video sequence, tampered in region forgery or mask the content in objects in frames videos that method of forgery was used. Many passive schemes have been developed based on these in-painting JPEG images [4], histogram equalization based contrast enhancement techniques [5], blue screen special effects in videos [6], Markov model on motion residue video [7], Ballistics motion [8], non-sampled contourlet transformation and gradients information [9], advanced statistical and adaptive threshold [10], correlation noise residue [11], double MPEG compression [12], detection duplication [13], detecting double quantization [14], photograph image forgery techniques [15], and

luminance level techniques [16], anti-forensic techniques for frame add/delete in MPEG video. Despite the good detection results that this algorithm gets, it is described as a very complicated and high time processing algorithm.

3 Methodology and Implementation of Video Forgery Detection

When a video sequence is captured, there is typically a great deal of redundancy between each frame of video. An MPEG or video compression technique exploits this redundancy by predicting certain frames in the video sequence from others, then by encoding the residual error between the predicted frame and the actual frame. Because the prediction error can be compressed at a higher rate than a frame in its entirety, this leads to a more efficient compression scheme. Performing compression in this manner has its drawbacks, because error introduced from one frame will propagate to all frames predicted from it. To prevent error propagation, the video sequence is divided into segments, where each segment is referred to as a group of pictures (gp), during MPEG and other video compression. Frame prediction is performed within each segment, but never across segments, thus preventing decoding errors in one frame from spreading throughout video sequence. Within each group of picture, frames are divided into three types: intra-frames (I-frames), predicted frames (P-frames), and bidirectional frames (B-frames). Each gp begins with an I-frame, followed by a number of P-frames and B-frames. No prediction is performed when encoding I-frames; therefore, each I-frame is encoded and decoded independently. During encoding, each I-frame is compressed through a lossy process similar to JPEG compression. P-frames are predicatively encoded through a process known as motion estimation.

A predicted version of the current P-frame is obtained by first segmenting the frame into 16×16 pixel blocks known as macroblocks, then searching the previous P- or I-frame, known as the anchor frame, for the macroblock that best matches each macroblock in the current P-frame. The locations of these macroblocks in the anchor frame are stored, along with how far each macroblock must be displaced to create the predicted frame. These displacements are referred to as motion vectors. The residual error between the predicted frame and the current frame, known as the prediction error, is then compressed using the same JPEG like process that I-frames undergo.

3.1 Mean Frame Comparison Technique

In the given set of tested video, the content of pictures is read as well as the content of source of original video where in mean compare operation is performed on every

frame of the video, and the mean comparison is performed in frame-by-frame fashion, that is, each and every pixel undergoes comparison with pixel-by-pixel comparison and extracting the mean of each frame in frame set. On comparison of frames, if the mean value of frames in frame set value is equal, then that area is masked in black. If there is any difference in the frame value between these groups of pictures, then that area/region can be identified to be tampered. The result viewed will show a black masked area where there is no difference between these group of picture frames, that is, no difference between the pixel value of original and forged frames in videos. And the area where there is pixel frame difference will be highlighted which was a tampered area is been suspected; mathematical model for mean comparison is given as: Let $f(x)$ is the original group of picture frame set, and let $f(y)$ is the forged group of picture frame set by extracting the features of each frames in forged video set. On using mean comparisons, we obtained as

$$f(z)'' = f(x) \sim f(y) \tag{1}$$

where $f(z)''$ is the suspected and identified forged region of picture frames in digital videos where $f(x)$ and $f(y)$ are original and forged feature set videos.

4 Results and Discussion

Videos are captured from Sony digital camera, and results are implemented in MATLAB 9. We consider 100 different video galleries set collected from various source. We tested on all categories of indoor scene and outdoor scene videos and completely succeeded in detecting forgery region from digital video.

Figure 1 explains sample group of pictures extracted from original video sources. We have forged some frames using video editing software and created second group of testing images shown in Fig. 2. Figure 3 demonstrated sample group of picture by executing mean frame comparison algorithm, and result of algorithm and results are compared in Fig. 4, where we succeeded to identify the forgery group of pictures from input video.

Fig. 1 Frame set of original video

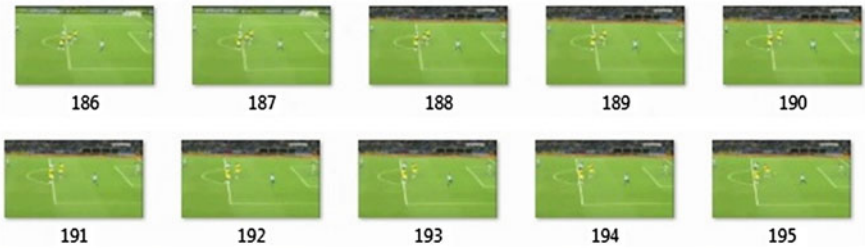


Fig. 2 Frame set forged video

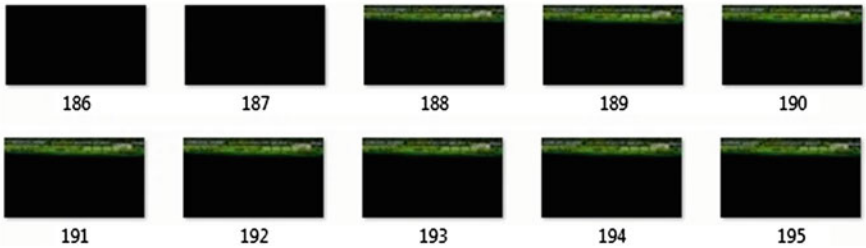


Fig. 3 Resultant frame set forgery region detection

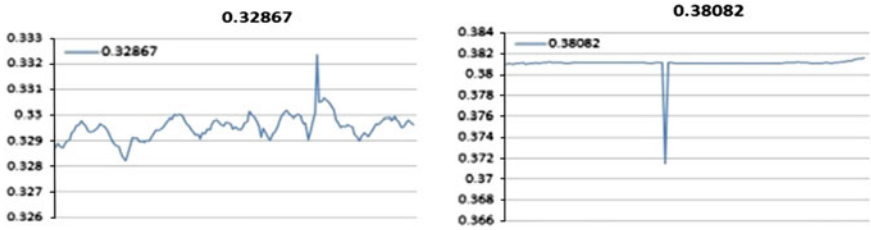


Fig. 4 Resultant mean difference from forgery region from frame set

5 Conclusion

This paper outlined the methodology used in detecting digital video tampering from group of picture frames with known and unknown origin. A wide range of techniques available in detecting the video tampering in digital world is very difficult to analyze and test the knowledge of forgery level in publishing videos in digital media world. The detection system proposed mean frame comparison technique includes methods based on mean and pixel comparison of each frames in video data frame set used with unknown image source. An experiment testing this method has been set up that will help in determining the accuracy and correctness

of the proposed tamper detection techniques, as well as when each fails. It has been conjectured that these methods will help in detecting various types of image forgeries, but one has to acknowledge that no "silver-bullet" exists to account for every type of video forgery.

References

1. Murali S, Anami BS, Chittapur GB (2012) Detection of photo-forgery detection techniques. Int J Mach Intell 4(1). ISSN: 09752927 EISSN: 09759166: 2012
2. Stamm MC, Ray Liu KJ (2011) Anti-forensics for frame deletion/addition in MPEG video. IEEE international conference on acoustics, speech and signal processing. doi:10.1109/ICASSP.2011.5946872
3. Ng TT, Chang SF, Lin C-Y, Sun Q (2006) Passive-blind image forensics. In: Zeng W, Yu H, Lin C-Y (eds.) Multimedia security technologies for digital rights. Elsevier, Amsterdam
4. Zhao YQ et al (2012) Tampered region detection of inpainting JPEG images. Optik Int J Light Electron Opt. http://dx.doi.org/10.1016/j.ijleo.2012.08.018
5. Soong-Der C et al (2012) A new image quality measure for assessment of histogram equalization-based contrast enhancement techniques. Int J Digital Signal Process 22:640–647. http://dx.doi.org/10.1016/j.dsp.2012.04.002
6. Xu J et al (2012) Detection of blue screen special effects in videos. Phys Procedia 33:1316–1322. doi:10.1016/j.phpro.2012.05.217
7. Kancherla K et al (2012) Novel blind video forgery detection using markov models on motion residue. In: Intelligent information and database systems lecture notes in computer science, vol 7198, pp 308–315
8. Farid H et al. (2012) Exposing digital forgeries in ballistic motion. IEEE Trans Inf Forensics Sec 7(1), feb 2012. doi:0.1109/TIFS.2011.2165843
9. Chen R et al (2012) Video forgery detection based on Non-sub sampled contourlet transformation and gradient information. Informational Technol J. ISSN: 1812-5638. doi:10.2923/itj.2012
10. Christodoulou L et al. (2011) Advanced statistical and adaptive threshold techniques for moving object detection and segmentation. 17th international conference on digital signal processing. doi:10.1109/ICDSP.2011.6004875
11. Hsu CC, Hung TY, Lin CW, Hsu CT (2008) Video forgery detection using correlation of noise residue, Department of Electrical Engineering, Tiwan
12. Wong W, Farid H (2006) Exposing digital forgeries in video by detecting double MPEG compression. In: Proceeding of MM& SEC 2006, ACM Publications ACM 1595934936/06/0009
13. Wong W, Farid H (2007) Exposing digital forgeries in video by detecting duplication. In: Proceeding of MM& SEC 2007, ACM 9781595938572/07/0009
14. Wong W, Farid H (2009) Exposing digital forgeries in video by detecting double quantization. In: Proceeding of MM& SEC 2009, ACM 978-1-59593-857-2/07/0009
15. Murali S, Chittapur GB, Anami BS (2012) Detection Of digital photo image forgery. IEEE international conference on advanced communication and control technology. doi:10.1109/ICACCCT.2012.6320754
16. Murali S, Anami BS, Chittapur GB (2011) Detection of copy-create image forgery using luminance level techniques. Third national conference on computer vision, pattern recognition, image processing and graphics. doi:10.1109/NCVPRIPG.2011.53

Mobile WiMAX Performance Improvement Using MRRC Scheme for Real-Time Application

Arathi R. Shankar, Adarsh Pattar and V. Sambasiva Rao

Abstract IEEE 802.16e (or) Mobile WiMAX is the network which is designed with the quality of service in mind. In this paper, an attempt has been made to qualitatively analyze the performance factors of Mobile WiMAX under flat-fading Rayleigh channel using MRRC techniques. The physical layer of Mobile WiMAX is simulated by using maximal ratio combining diversity techniques. The physical layer performance parameters have been measured and compared by increasing the number of diversity branches and by employing different modulation techniques. Proposed diversity scheme is compared with respect to throughput and Bit error rate by employing different modulation schemes for a real-time application.

Keywords MAC layer · OFDMA · Physical Layer · BER · MRRC · Throughput · SNR · Diversity

1 Introduction To Mobile Wimax

Mobile WiMAX is a broadband wireless solution that enables convergence of mobile and fixed broadband networks through a common wide area broadband radio access technology and flexible network architecture [1, 2]. The Mobile

A. R. Shankar (✉)
Department of Electronics and Communication, BMS College of Engineering, Bangalore, India
e-mail: arathi.rshekhar@gmail.com

A. Pattar
BMS College of Engineering, Bangalore, India
e-mail: techie04@gmail.com

V. Sambasiva Rao
Department of Electronics, PESIT, Bangalore, India
e-mail: vsrao@pes.edu

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_58,

WiMAX Air Interface adopts Orthogonal Frequency Division Multiple Access (OFDMA) [2] is introduced in the IEEE 802.16e. Amendments support scalable channel bandwidths from 1.25 to 20 MHz. Three main receiver diversity models are selective diversity, equal gain combining, and maximal ratio combining.

2 Proposed Mobile WiMAX END-to-END Physical Layer Model

The model shown in Fig. 1 is divided into 3 parts

1. Mobile WiMAX transmitter
2. Receiver diversity model
3. Mobile WiMAX receiver.

In this section, we are going to discuss about Mobile WiMAX transmitter and receiver blocks, the receiver diversity models. The recorded analog audio data are digitized through PCM encoder. Pulse-code modulation (PCM) encoder modifies the input analog audio pulses by PAM to create a completely digital signal. To do so, PCM first quantizes the PAM pulses. The results of quantized levels are presented into its 7 binary bits. The eighth bit indicates the sign. The processed input binary data stream is ensured against transmission errors with forward error correction codes (FEC) and interleaver [3–5]. The encoded bits from the output of

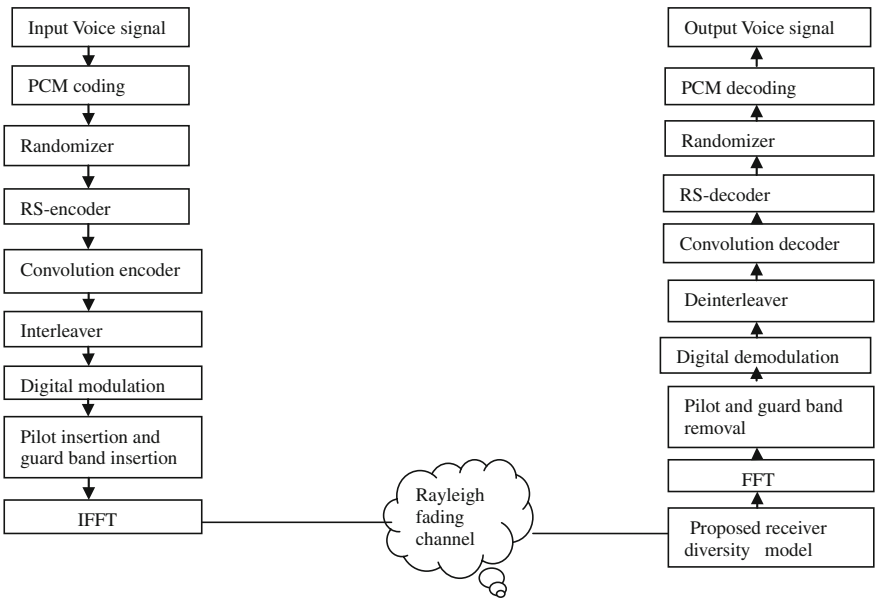


Fig. 1 Proposed end-to-end model

Table 1 OFDM parameters for simulation

Parameters	Value
FFT size	256
Number of used subcarriers	192
Number of guard bands	56
Cyclic prefix	¼
Channel bandwidth	5 MHz
Subcarrier frequency spacing	10.094 kHz
OFDM symbol duration	102.9 μs

convolution encoder are interleaved prior to converted into each of the either eight complex modulation symbols in BPSK, QPSK, 4QAM, 8PSK, 16PSK, and 16QAM modulation schemes and fed to a multicarrier modulator for transmission. In the modulator, the digitally modulated information symbols are transmitted in parallel on subcarriers through the implementation as an inverse discrete Fourier transform (IDFT) on a block of information symbols followed by an analog-to-digital converter (ADC). To mitigate the effects of inter-symbol interference (ISI) caused by channel time spread, each block of IDFT coefficients is typically preceded by a cyclic prefix. After Ifft operation, data are forwarded to proposed diversity model. After processing, the data is sent to the WiMAX receiver at the receiver side and it is demodulated, de-interleaved and then PCM decoded.

The Table 1 will give OFDM parameters used in our simulation model:

3 Simulation Results and Discussion

In this paper, we will simulate and analyze the Receiver Diversity models are simulated and analyzed in flat fading Rayleigh channel, which we discussed in last paper and taking recorded voice signal as real-time input data. A simulation model is prepared for Mobile WiMAX physical layer with using MATLAB 7.11.b (2010a), and output simulation results are as given below.

We are using 3 modulation schemes with simulation parameters listed in Tables 2, 3, and 4 for 3 proposed diversity models [6]. In rest of the paper we will mention the table numbers of the above listed tables.

Table 2 Simulation parameters for BPSK scheme

BPSK Modulation scheme	
Parameters	Value
Size of sub frame	88 bits
Convolution code rate	½
Bandwidth	5 MHz
Sampling frequency	4.04 kHz
FFT size	256

Table 3 Simulation parameters for QPSK scheme

QPSK Modulation scheme	
Parameters	Value
Size of sub frame	184 bits
Convolution code rate	2/3 with RS encoding of n = 40 and k=36
Bandwidth	5 MHz
Sampling frequency	4.04 kHz

Table 4 Simulation parameters for QAM-16

Qam-16 Modulation Scheme	
Parameters	Value
Size of sub frame	280 bits
Convolution code rate	2/3 with RS encoding of n = 64 and k = 48
Bandwidth	5 MHz
Sampling frequency	4.04 kHz

3.1 Maximal Ratio Combining Diversity Model

Here, simulation results of maximal ratio combining diversity model for 3 different modulation schemes are presented. The Fig. 2 shows input [7–9] real-time signal used for simulation.

Case 1: BPSK Modulation scheme (Figs. 3 and 4):

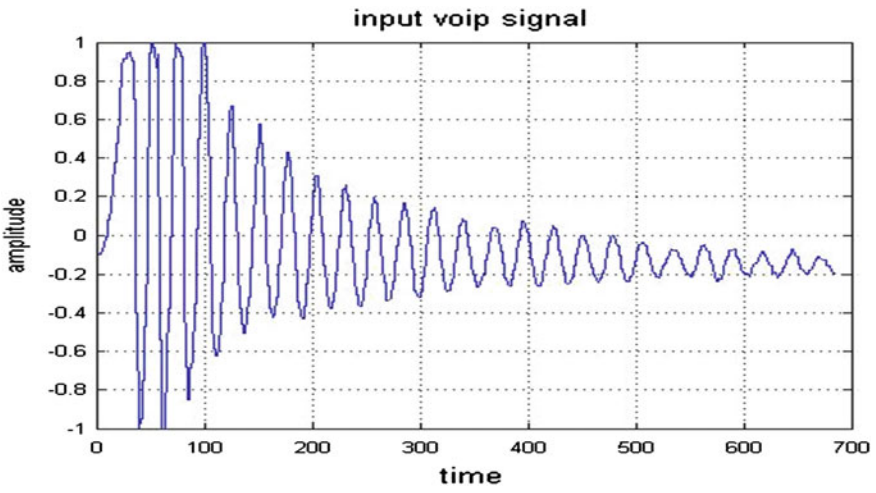


Fig. 2 Input voice signal

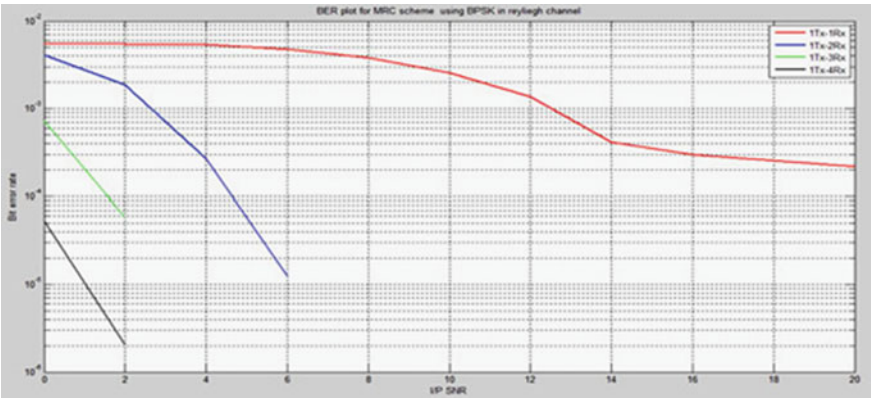


Fig. 3 BER for BPSK under MRC

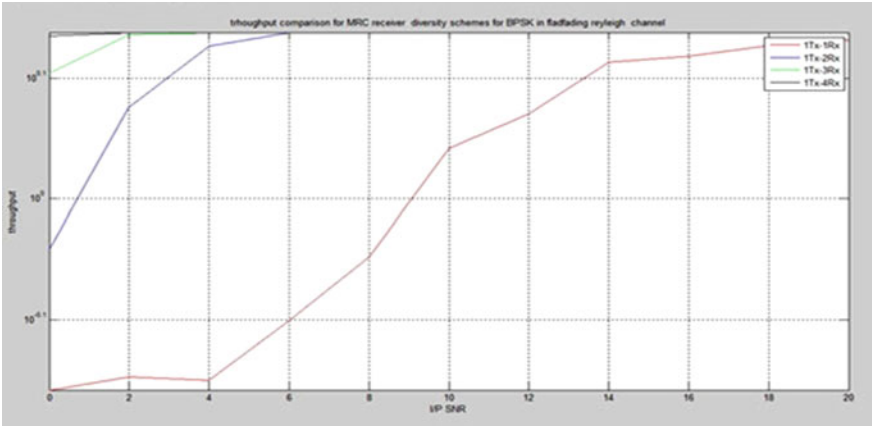


Fig. 4 Throughput for BPSK under MRC

The above three Figs. 2, 3, 4 results show that performance of MRC will become equivalent to AWGN [10–13] when number of branches increased as the gain in effective SNR is high and is directly proportional to number of diversity branches.

Case 2: QPSK Modulation scheme (Figs. 5 and 6):

The above 3 figures shows BER, throughput for QPSK scheme. The throughput and capacity are increased by a factor of 2.

Case 3: QAM-16 Modulation scheme (Figs. 7 and 8):

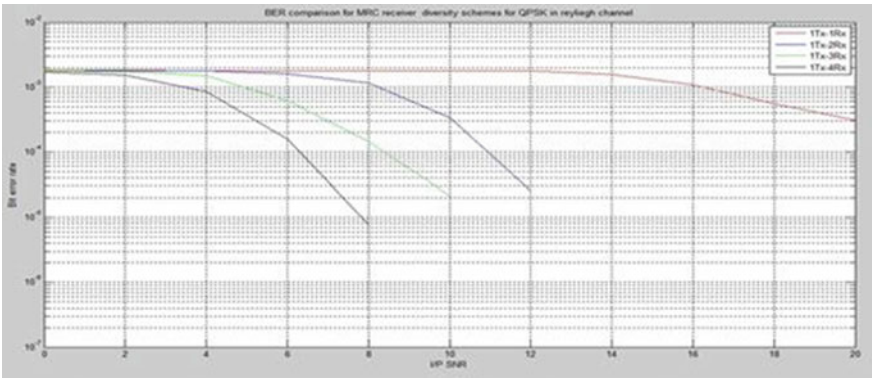


Fig. 5 BER for QPSK in MRC

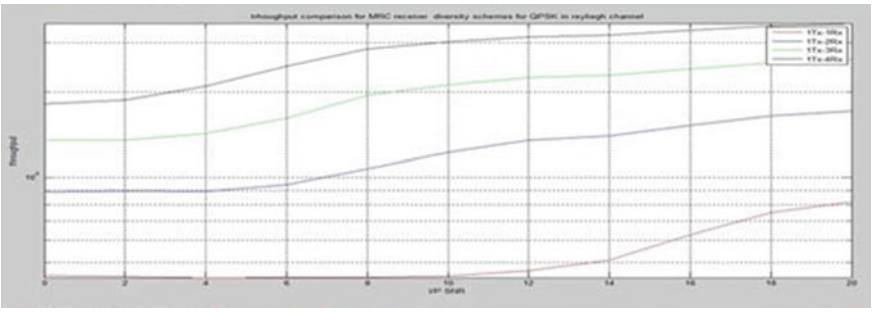


Fig. 6 Throughput comparison for QPSK under MRC

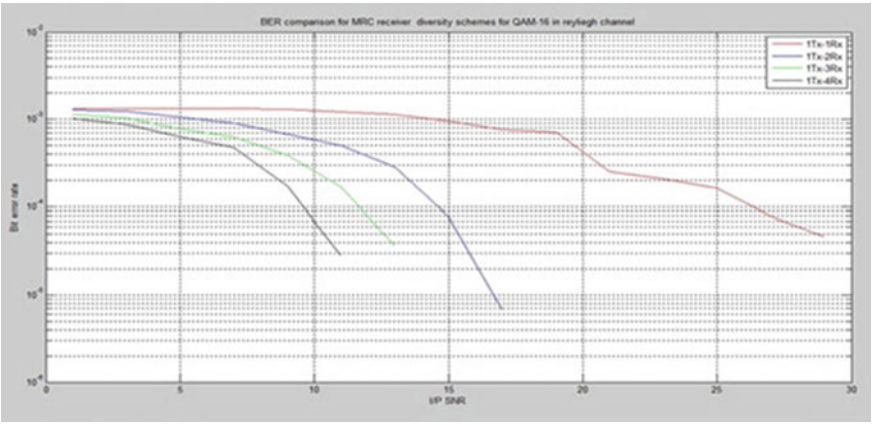


Fig. 7 BER comparison for QAM16 under MRC

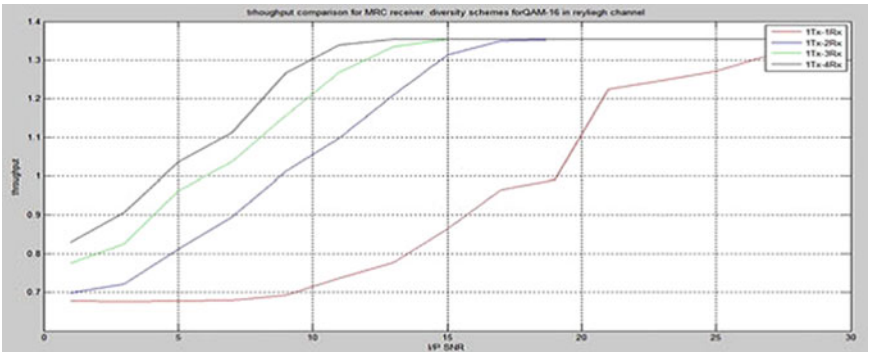


Fig. 8 Throughput comparison of QAM16 under MRC

The above figures show that even at higher end modulation of QAM-16, the system performance is excellent and BER is reduced to zero by taking very less input SNR value of 12, which is an ideal SNR value for QAM-16bit under AWGN channel [14]. The throughput and capacity were increased by a factor of 4.

4 Conclusion

This work is focused on enhancing physical layer performance of Mobile WiMAX by maximal ratio combining diversity to eliminate ISI and multipath fading problem and to improve cross-layer parameters from physical layer to MAC layer.

After simulating three models for same voice under three modulation scheme, we concluded that the performance maximal ratio combining scheme performance is nearly equal to AWGN channel performance in all the three modulation schemes.

Under higher end modulation schemes, the channel coefficients become inconsistent and we will use channel inversion technique at the receiver side to equalize the channel coefficients and combat the inconsistency. We suggest that mapping of QoS parameters and a service class mapping have to be performed between physical and MAC Layer to improve an end-to-end QoS performance.

References

1. Arafat O, Dimyati K (2010) Performance parameter of mobile WiMAX: a study on the physical layer of mobile WiMAX under different communication channels & modulation technique. In: Computer engineering and applications, international conference on 2010, vol 2 pp 533–537, second international conference on computer engineering and applications, 2010

2. IEEE 802.16e-2005 (2006) IEEE standard for local and metropolitan area networks, part 16, air interface for fixed and mobile broadband wireless access systems. IEEE Press, New York
3. Proakis JG (1995) Digital communications. McGraw-Hill, New York
4. Gravano S (2001) Introduction to error control codes. Oxford University Press, Oxford
5. Tse D, Viswanathan P (2005) Fundamentals of wireless communication. Cambridge University Press, Cambridge
6. Alamouti SM (1998) A simple transmit diversity technique for wireless communications. IEEE J Sel Areas Commun 16(8):1451–1458
7. Islam AMA, Kader BMA (2011) BER performance analysis of a real data communication through WiMAX-PHY layer over an AWGN and fading channels. Int J Electr Comput Sci IJECS-IJENS 10(4):13–16
8. Telagarapu P, Naidu GBSR, Chiranjeevi K (2011) Analysis of coding techniques in WiMAX. Int J Comput Appl 22(3):0975–8887
9. Sanghoi P, Kansal L (2012) Analysis of WiMAX physical layer using spatial multiplexing. Int J Wire Mob Netw (IJWMN) 4(2). doi:10.5121/ijwmn.2012.4207103
10. Khan MN, Ghauri S (2008) The WiMAX 802.16e physical layer model. Int Conf Wirel, Mob Multimedia Netw 2:117–120. ISBN:978-0-86341-887-7
11. Mountassir J, Balta H, Oltean M, Kovaci M, Isar A (2011) A physical layer simulator for WiMAX in rayleigh fading channel. In: 6th IEEE international symposium on applied computational intelligence and informatics, pp 281–281, 2011
12. Zerrouki H, Feham M (2010) High throughput of WiMAX MIMO-OFDM Including Adaptive Modulation and Coding. Int J Comput Sci Inf Secur 7(1):86. ISSN:19475500, DOAJ
13. Chen Y, Tellambura C (2005) Performance analysis of three-branch selection combining over arbitrarily correlated rayleigh-fading channels. IEEE Trans Wire Commun 4(3):861
14. Dietze K, Dietrich CB, Stutzman WL (2002) Analysis of a two-branch maximal ratio and selection diversity system with unequal SNRs and correlated inputs for a rayleigh fading channel. IEEE Trans Wire Commun 1(2):274–281

Feature Extraction Using Shearlet Cone Representation for Face Recognition

Ajit Danti and K. M. Poornima

Abstract In this paper, a new feature extraction technique is proposed. In the proposed method, features are extracted by representing the face image in the form of horizontal and vertical cones using Shearlet transform. Then, principal component analysis (PCA) is used to reduce the dimension of feature set and obtain dominant features. k-nearest neighbor (k-NN) classifier is used to classify the test face image. Here, k-NN is used as a method for classifying faces based on closest training examples in the feature matrix. The test face is classified by a majority vote of its neighbors, with the face being assigned to the class most common among its k-nearest neighbors. The experiments for the proposed method were conducted on ORL database, Essex Grimace database, and Faces database, and the comparative results are presented. The Shearlet transform with PCA technique significantly reduced the recognition time.

Keywords Face recognition · Feature extraction · Shearlet transform · k-NN classifier · PCA

1 Introduction

Human face recognition is a difficult problem in computer vision [1]. Face recognition is challenging because it is a real-world problem. The human face is a complex natural object that tends not to have easily identified edges and features. Because of this, it is difficult to develop a mathematical model of the face that can

A. Danti · K. M. Poornima (✉)
JNN College of Engineering, Shimoga 577204, Karnataka, India
e-mail: kmpoornima@sify.com

A. Danti
e-mail: ajitdanti@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_59,

be used as prior knowledge when analyzing a particular image. Recently, face recognition is attracting much attention in society of network, multimedia, and information access. Areas such as network security, content indexing and retrieval, and video compression benefit from face recognition technology because people are the center of attention in lot of video [2].

Multiresolution analysis tools like wavelets have been found very useful in analyzing the image contents; hence, they widely used in image processing, pattern recognition, and computer vision. Over the years, other multiresolution techniques such as contourlets, ridgelets, and curvelets were developed. Shearlet transform is a recent addition. It has been used in image denoising [3], edge analysis [4], and image separation [5], but not much work has been done to solve pattern recognition problems.

In our earlier work [6], we used Shearlet transform to extract the features for face recognition. The extracted features were large in size, which increases the processing time. Increase in the processing time decreases the performance of the overall system. Better performance can be obtained if the processing time is reduced, which can be achieved by reducing the extracted feature vector without losing the important or dominant features. This reduction in features is achieved through principal component analysis (PCA). The proposed method has been evaluated by conducting different experiments. Initially, only one image per subject has been used in the training phase; later, multiple sample images have been used to increase the recognition rate.

2 Proposed Method

The basis of standard eigenface technique [7] is called PCA and is widely used in face recognition. Through row or column concatenation, each image is converted into a vector and its dimension is reduced by employing PCA. The problem with this method is high computation complexity and to minimize this, PCA is applied on Shearlet transform. Next section gives the brief description of Shearlets.

2.1 Shearlet Transform

Shearlet systems are systems of functions generated by one single generator with parabolic scaling, shearing, and translation operators applied to it, and in much the same way, wavelet systems are dyadic scaling and translations of a single function, but including a precise control of directionality. Of the many directional representation systems proposed in the last decade, Shearlets are among the most versatile and successful systems. Because they exhibit list of desirable properties, Shearlet systems can be generated by one function, they provide precise resolution of wavefront sets, they allow compactly supported analyzing elements, they are

associated with fast decomposition algorithms, and they provide a unified treatment of the continuum and the digital realm.

Shearlet systems are based on three parameters: $a > 0$ being the scale parameter measuring the resolution, $s \in R$ being the shear parameter measuring the directionality, and $t \in R^2$ being the translation parameter measuring the position. To achieve optimal sparsity, Shearlets are scaled according to a parabolic scaling law encoded in the parabolic scaling matrix A_a, $a > 0$, and exhibit directionality by parameterizing slope encoded in the shear matrix S_s, $s \in R$, defined by Eq. (1).

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \quad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \tag{1}$$

where A_a denotes a parabolic scaling matrix and S_s denotes a shear matrix.

The Shearlet system associated with a Shearlet $\Psi \in L^2(R^2)$ is given by Eq. (2)

$$\Psi_{\mathrm{jkm}}(x) = 2^{-\frac{3j}{2}}\, \Psi(S_{-k}A_{4-j}x - m) : j,k \in Z, m \in Z^2 \tag{2}$$

where j provides the scale, k and m detect the direction and position of singularities, respectively. The associated Shearlet transform SH_Ψ is then defined on $L^2(R^2)$ by Eq. (3)

$$SH_\Psi f(j,k,m) = \langle f, \Psi_{\mathrm{jkm}} \rangle \tag{3}$$

To provide an equal treatment of the x- and y-axis, the frequency plane is split into four cones C_{11}–C_{22}, defined in Eq. (4) and it is represented in Fig. 1. The resultant discrete Shearlet system is denoted by Eq. (5).

$$C_l = \left\{ \begin{array}{l} \{(\xi_1\xi_2) \in R^2 : \xi_1 \geq 1, |\xi_1/\xi_2| \geq 1\} : l = 21, \\ \{(\xi_1\xi_2) \in R^2 : \xi_2 \geq 1, |\xi_1/\xi_2| \leq 1\} : l = 11, \\ \{(\xi_1\xi_2) \in R^2 : \xi_1 \leq -1, |\xi_1/\xi_2| \geq 1\} : l = 22, \\ \{(\xi_1\xi_2) \in R^2 : \xi_2 \leq -1, |\xi_1/\xi_2| \leq 1\} : l = 12. \end{array} \right\} \tag{4}$$

$$\{\sigma_\eta : \eta \in N_0 \times \{-2^j \ldots 2^j\} \times Z^2 \times \{11, 12, 21, 22\}\} \tag{5}$$

where η = (j, s, m, ι) is index scale, orientation, position, and cone.

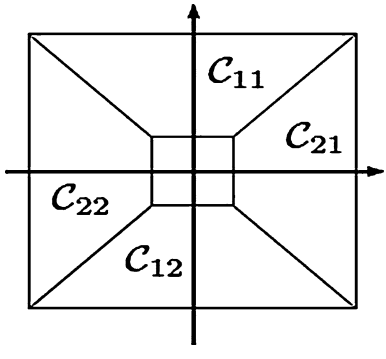


Fig. 1 Cones C_{11}–C_{22}

The Shearlets for C_{11}, C_{12}, C_{21}, and C_{22} are defined likewise by symmetry. $L^2(R^2)$ is the choice of a two-parameter dilation group, where one parameter ensures the multiscale property [8], whereas the second parameter provides a means to detect directions. The definition shows that Shearlets live on anisotropic regions of width 2^{-2j} and length 2^{-j} at various orientations.

2.2 Shearlet-Based Feature Extraction

Principal component analysis has been successfully applied on wavelet decomposed images in [9, 10]. Though wavelets are suitable for detecting point singularities in images but they fail to represent curved discontinuities. Curvelets were developed by Candes and Donoho in 2004, which provides optimally sparse approximations of anisotropic features. But it has two drawbacks: firstly, the curvelet system is not singly generated, that is, it is not derived from the action of countably many operators applied to a single or finite set of generating functions; secondly, its construction involves rotations and these operators do not preserve the digital lattice which prevents a direct transition from the continuum to the digital setting. Later, contourlets were introduced in 2005 by Do and Vetterli, but a proper continuum theory is missing in this approach [11].

Shearlets were introduced by Guo et al. in 2005 and are derived from a single or finite set of generators and ensures a unified treatment of the continuum and digital world because the shear matrix preserves the integer lattice. It has some unique properties: it has a single or finite set of generating functions, it provides optimally sparse representation for a large class of multidimensional data, it is possible to use compactly supported analyzing functions, it has fast algorithmic implementations, and it allows a unified treatment of the continuum and digital realms. Shearlets allow optimal encoding of several classes of multivariate data through its ability to sparsely represent anisotropic features and hence Shearlets emerged as the new tool for processing massive and higher dimensional data.

The proposed face recognition system is divided into two phases: training phase and classification phase. In training phase, Shearlet transformation is used for feature extraction. Each face image is decomposed along horizontal and vertical direction and cone-like representations are obtained. The decomposition is based mainly on the number of levels, the scale parameter, and the shear parameter. Here, the number of levels taken is 5; therefore, the scale parameter varies from 0 to 4, and the shear parameter varies from $2^{(-j/2)}$ to $2^{(j/2)}$. For each face image, we obtained five horizontal and five vertical cones. Features are extracted from these horizontal and vertical cones. Obtained features are larger in size. Principal component analysis (PCA) is employed in order to reduce the dimension without losing the important or dominant features. In classification stage, test images are subjected to the same operations and are transformed to the same PCA representational basis. k-nearest neighbor (k-NN) classifier is used to perform the recognition task.

Fig. 2 Sample images of ORL database

3 Experiments

3.1 Datasets

In order to test the efficiency of the proposed approach, we have conducted the experiments on ORL database [12], Essex Grimace database [13], and Faces94 [14] database. The ORL database contains only grayscale face images. These images are of size 92 × 112. The database contains 10 different images each for 40 distinct persons. The face images were taken at different times by varying the lighting, facial expression, and facial details. All the images were taken against a dark homogeneous background. Sample images of this dataset are shown in Fig. 2.

Essex Grimace database consists of a sequence of 20 images each for 18 individuals consisting of male and female subjects. During the sequence, the subjects move their head and make grimaces which get more extreme toward the end of the sequence. Images are taken against a plain background, with very little variation in illumination. The images are in ".jpg" format and of size 180 × 200. Sample images of this database are shown in Fig. 3.

Faces94 database contains images of 153 individuals, 20 images per person. The subjects sit at fixed distance from the camera and are asked to speak, while a sequence of images is taken. Faces of this database show considerable expression changes but very minor variations in lighting or head position. Sample images shown in Fig. 4.

Fig. 3 Sample images of Essex Grimace database

Fig. 4 Sample images of Faces94 database

3.2 Results and Discussion

For training purpose, initially single image of each person is considered from ORL database and the remaining set is used for testing. Later, the number of training images is increased to 5 and the remaining images are used for testing.

In the training phase, the original face images are decomposed up to 5 levels by using Shearlet transform, by varying the scale parameter from 0 to 4 and shear parameter from $2^{(-j/2)}$ to $2^{(j/2)}$. Thus, 10 coefficients are obtained, five for horizontal cone and five for vertical cone and are concatenated to form the meaningful feature vector. Since the feature vector is large in size, to reduce the size, principal component analysis is employed. After applying the PCA, feature vector size is reduced from 29,664 × 200 to 200 × 200. In the testing phase, same steps are repeated to obtain Shearlets coefficients and are transformed into PCA representational basis. Then, k-NN classifier is used to perform the recognition task. Series of experiments are conducted by considering k parameter value as 3 and 5. When only one face image per class has been used in the training phase, the recognition rate was less. Later, the number of training images is increased to five. As the number of trained images increased, the recognition rate also increased. We achieved a recognition rate of 88.5 % with 5 training images and $k = 3$. The recognition result is given in Table 1.

Same experiments are repeated for Essex Grimace database and Faces94 database; for Essex Grimace database, we obtained a recognition rate of 92 % with 5 training images and $k = 3$, and for Faces94 database, a recognition rate of 90.5 % with 5 training images and $k = 3$ is obtained. The results are given in Tables 2 and 3, respectively.

By verifying the above results, we concluded that the proposed method gives better recognition rate for the Essex Grimace database with five training images per class and $k = 3$ when compared with other two databases, because variation in intensity level is less in Essex Grimace database.

3.3 Comparative Study

In the previous section, different results of Shearlet-based PCA technique have been presented. In order to show the capability of the proposed method, we have compared it against well-known existing methods such as eigenface, wavelet-based PCA, and curvelet-based PCA. Experiments are conducted by considering

Table 1 Recognition rate versus number of training images (ORL)

Trained images in each class	1	2	3	4	5
Recognition rate when $k = 3$	68.5	73.0	81.0	83.0	88.5
Recognition rate when $k = 5$	63.0	71.5	80.0	81.5	86.0

Table 2 Recognition rate versus number of training images (Essex Grimace)

Trained images in each class	1	2	3	4	5
Recognition rate when k = 3	72.0	77.5	84.5	88.0	92.0
Recognition rate when k = 5	70.0	73.5	80.0	83.0	87.0

Table 3 Recognition rate versus number of training images (Faces94)

Trained images in each class	1	2	3	4	5
Recognition rate when k = 3	70.0	74.5	83.0	87.5	90.5
Recognition rate when k = 5	68.0	72.5	78.0	81.5	87.0

Table 4 Comparative study

Recognition method	Recognition rate (%)		
	ORL	Essex Grimace	Faces94
Eigenface	72.5	69.5	71.0
Waveletface + PCA	81.0	83.5	82.5
Curveletface + PCA	83.0	89.0	86.0
Proposed method	88.5	92.0	90.5

same databases (ORL, Essex Grimace, and Faces94), the number of training and testing face images for each class as 5 with k = 3 and k = 5. For wavelet-based PCA method, 3-level wavelet decomposition was performed by using Haar Wavelet. The best 60 eigenvectors were selected for both wavelet- and curvelet-based PCAs. The results are presented in Table 4.

4 Conclusions

A novel feature extraction method for face recognition based on Shearlet transformation is proposed. Shearlets are faster in extracting the features than any other transformations available. Here, face images are represented as cones and features are extracted. PCA is employed to reduce the size of the feature set. k-nearest neighbor (k-NN) is used for the classification. The proposed approach is experimented on ORL database, Essex Grimace database, Faces94 database, and the comparative results are presented. Other dimensionality reduction techniques such as ICA and LDA can also be employed to reduce the feature vector. Further, the performance can be improved by employing better classifiers.

References

1. Lin SH (2000) An introduction to face recognition technology. IC media corporation, informing science special issue on multimedia informing technologies-part 2, 3(1)
2. Zhao W, Chellappa R, Rosenfeld A, Phillips PJ (2003) Face recognition: a literature survey. ACM Comput Surv 35:399–458
3. Easley GR, Labate D, Colonna F (2009) Shearlet-based total variation diffusion for denoising. IEEE Trans Image Process 18(2):260–268
4. Yi S, Labate D, Easley GR, Krim H (2009) A shearlet approach to edge analysis and detection. IEEE Trans Image Process 18(5):929–941
5. Kutyniok G, Lim WQ (2012) Image separation using wavelets and shearlets. Curves and Surfaces (Avignon, France, 2010), Lecture Notes in Computer Science 6920, Springer, Berlin, pp 416–430
6. Danti A, Poornima KM (2012) Face recognition using Shearlets. In: 7th IEEE international conference on industrial and information systems, 1–6 Aug 2012
7. Turk M, Pentland A (1991) Face Recognition using eigenfaces. In: Proceedings computer vision and pattern recognition, pp 586–591, 1991
8. Guo K, Labate D, Lim W, Weiss G, Wilson E (2006) Wavelets with composite dilations and their MRA properties. Appl Comput Harmon Anal 20(2):202–236
9. Feng GC, Yuen PC, Dai DQ (2000) Human Face Recognition using PCA on Wavelet Subband. J Electron Imaging 9(2):226–233
10. Manjunath BS, Chellappa R, van der Malsburg C (1992) A feature based approach to face recognition. In: Proceedings IEEE computer society conference on computer vision and pattern recognition, pp 373–378, 1992
11. Kutyniok G, Labate D (2012) Shearlets: multiscale analysis for multivariate data. Springer, Berlin
12. http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces
13. http://cswww.essex.ac.uk/mv/allfaces/grimace.zip
14. http://cswww.essex.ac.uk/mv/allfaces/faces94.html

Combined SLM and Companding for PAPR Reduction in OFDM

Kiran Sonna and M. N. Suma

Abstract In this paper, peak-to-average ratio reduction of OFDM system is discussed. Selective mapping (SLM) technique with RS and BCH coding and different phase sequences, with and without companding, for PAPR reduction in OFDM system is compared. Simulation results show better performance in terms of PAPR reduction by employing coding and phase sequences with companding.

Keywords PAPR · Companding · Phase sequences · SLM

1 Introduction

The advantage of orthogonal frequency division multiplexing (OFDM) over single-carrier schemes in wireless communications is its ability to manage with severe channel interference and fading due to multipath conditions with simple equalization. One of the issues in transmission with OFDM that needs to be addressed is high peak-to-average power ratio (PAPR). OFDM is a composite signal; high peaks in OFDM result from the superposition of a large number of statistically independent subchannels which can constructively sum up to high peaks.

Therefore, PAPR of OFDM is proportional to the number of subcarriers used. With large PAPR, the ADC and DAC implementation is difficult and demands linear RF amplifiers that are less power efficient.

K. Sonna (✉) · M. N. Suma
Electronics and Communication Engineering, BMS College of Engineering, Bangalore, India
e-mail: sonnakiran3@gmail.com

M. N. Suma
e-mail: suma.bms@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_60,

The criterion of the PAPR reduction is to find the approach that can reduce PAPR largely, and at the same time, it can keep the good BER performance, low implementation complexity, minimum side information to be transmitted, and no additional power requirement. Among the methods used to reduce PAPR, selective mapping (SLM) with coding [1] is very efficient probabilistic approach for reducing the PAPR. Further companding added will aid to reduce peak without affecting performance. In this paper, the PAPR performance based on coding with SLM and different phase sequence matrices is studied, along with addition of μ law companding.

2 Selective Mapping Technique and Coding

PAPR reduction with technique is done by either inserting redundant bits [2] or by use of phase sequences [3]. In the method based on redundant bits, few dummy sequences are inserted along with the information bits in order to reduce the PAPR. These dummy sequences do not carry any information but used for increasing the average power, which in turn reduce PAPR value. Amount of redundant bits added are removed at the receiving end. With use of phase sequence, output of each block is multiplied with different phase sequences and the PAPR of each combination is determined. Among this available combination, one with the lowest PAPR is transmitted. In SLM scheme, purpose is to make the occurrence of the peaks less frequent by introducing redundancy is not to eliminate the peaks, but to reduce the peak. Further, this scheme can handle any number of subcarriers. BCH codes form a large class of binary multiple random error-correcting codes. A Reed–Solomon code is a special case of a BCH code in which the length of the code is one less than the size of the field over which the symbols are defined. From the literature, it is found that there are similarities between BCH and RS codes in terms of encoding, decoding scheme, and their error-correcting capabilities. In this paper, we compare RS and BCH coding with SLM to improve PAPR performance [4,5].

3 Companding

Companding technique is introduced along with SLM to reduce the PAPR of OFDM signal. The OFDM signal after IFFT is companded at the transmitter. At the receiver end, received signal is expanded.

4 Phase Sequences

4.1 Riemann Sequence

The Riemann matrix is defined by $A = B\ (2{:}\ n + 1, 2{:}\ n + 1)$, where $B\ (i, j) = i-1$ if i divides j, and $B\ (i, j) = -1$ otherwise.

The Riemann matrix has the following properties:

1. Each Eigen value $e\ (i)$ satisfies abs $(e(i)) <= m-1/m$, where $m = n + 1$. With $i \leq e\ (i) \leq i + 1$ with at most m-sqrt (m) exceptions.
2. All integers in the interval $(m/3, m/2]$ are Eigen values.

4.2 Hadamard Matrix

This technique focuses on relationship between correlation properties of OFDM input sequence and PAPR probability. Hadamard matrix-based scheme reduces the occurrence of the high peaks compared to the conventional OFDM system. It is found that the PAPR can be reduced to about 2 dB in 16 QAM OFDM systems, without any power increase and side information and with little increase in system complexity.

4.3 Cauchy's Matrix

It is an n-by-n matrix and is generated by $C\ (i, j) = 1/(x\ (i) + y(j))$. Arguments x and y are vectors of length n. If we pass in scalars for x and y, they are interpreted as vectors 1: x and 1: y

4.4 Binomial Matrix

It is an n-by-n matrix, with integer entries such that $A\text{^}2 = 2\text{^}\ (n-1)*$ eye (n). Thus,

$B = A*2\text{^}\ ((1-n)/2)$ is involuntary, that is, $B\text{^}2 =$ eye (n).

5 Simulation Model and Results

Simulation model for transmitter and receiver with companding is shown in Figs. 1 and 2 for SLM with BCH and RS coding. Simulation model based on independent phase sequences with companding technique is shown in Fig. 3.

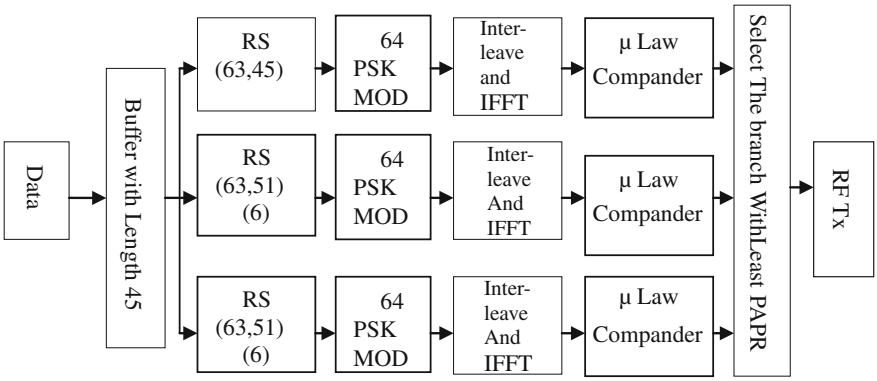


Fig. 1 SLM with RS encoding and companding

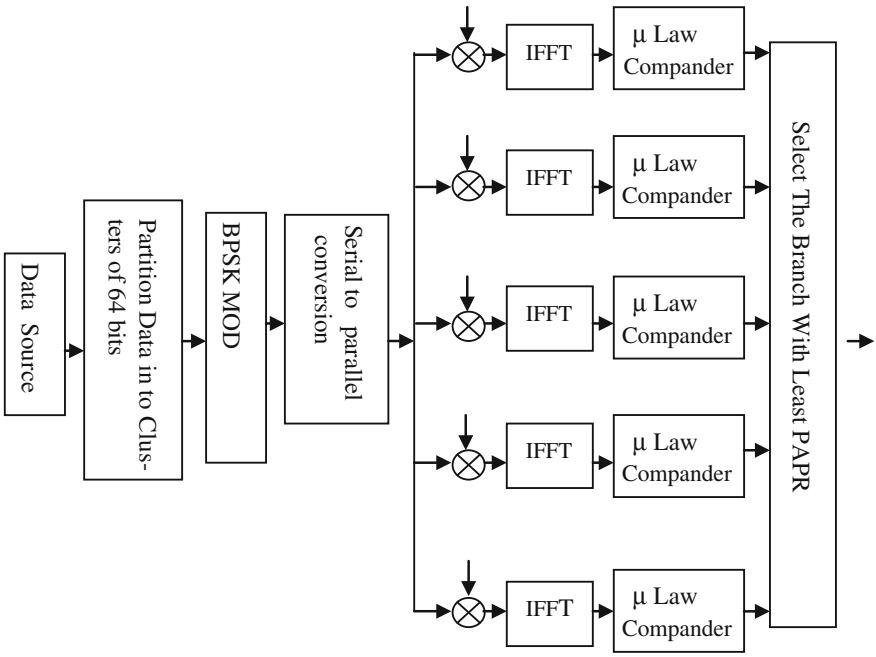


Fig. 2 SLM with companding

Simulation model remains same for all phase sequences used. Figure 4 shows CCDF plot for comparison of BCH and RS codes with and without companding. Results of PAPR obtained are indicated in Tables 1, 2, 3, and 4.

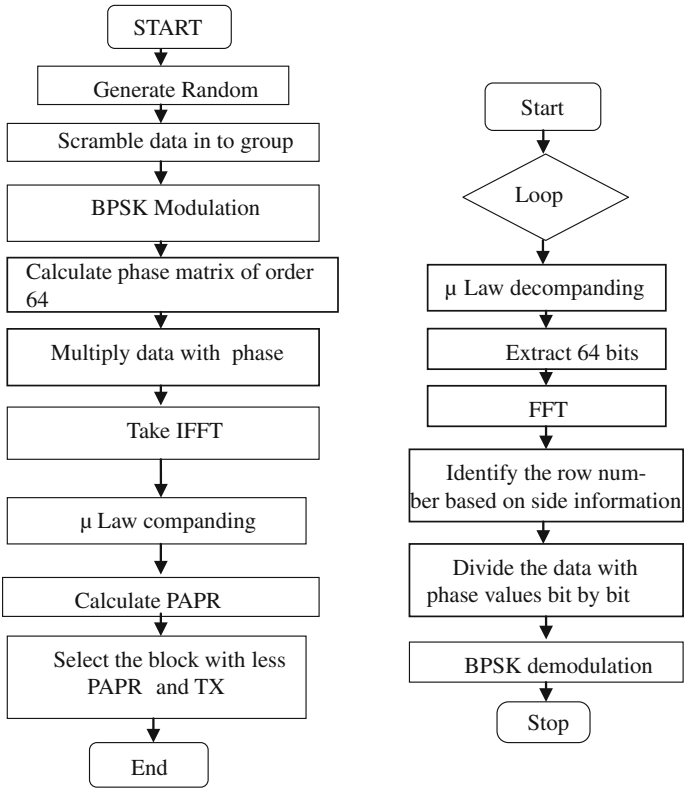


Fig. 3 Flow chart for the model

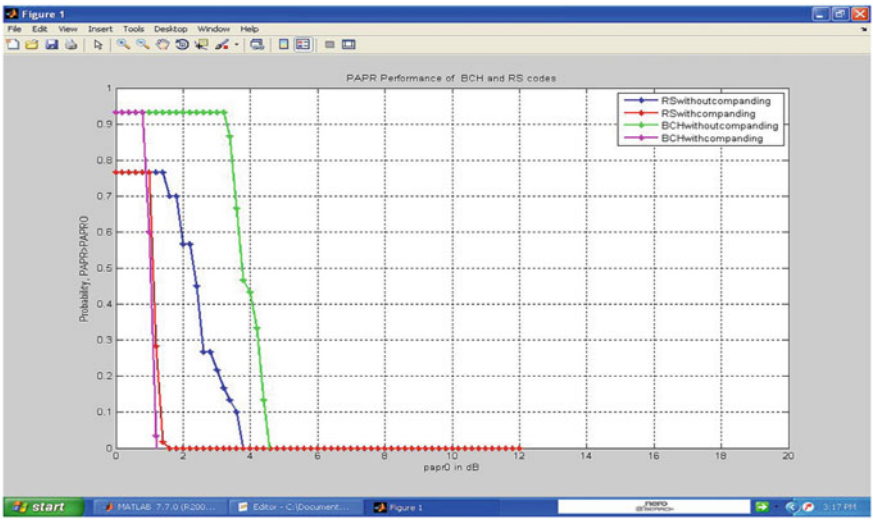


Fig. 4 CCDF Comparison of BCH with RS Schemes with and without companding

Table 1 Simulation Parameters

Simulation parameters	Type/Value
Coding scheme used	BCH,RS
Modulation scheme	BPSK
Total number of subcarriers	64,128
Compression parameter	μ Law with $\mu = 100$

Table 2 PAPR for SLM with RS—with and without companding

Number of subcarriers	Companding applied	PAPR in db
64	No	3.2
64	Yes	0.7
128	No	4.2
128	Yes	1.08

Table 3 PAPR for SLM with BCH—with and without companding

Number of subcarriers	Companding applied	PAPR in dB
64	No	3.8
64	Yes	0.94
128	No	4.73
128	Yes	1.30

Table 4 PAPR with and without companding for different phase sequences

		PAPR in dB		
Companding	Riemann	Hadamard	Cauchy's	Binomial
No	1.25	4.5	2.25	5.5
Yes	1.07	2.4	1.62	1.68

6 Conclusion

In this paper, PAPR reduction based on SLM and coding technique is discussed. The advantages offered by SLM technique are utilized to reduce the higher PAPR with BCH and RS codes as channel encoding technique which not only facilitate in the reduction of PAPR but also in the error correction. The simulated results showed that RS code performs better in PAPR reduction compared with BCH code.

Further with different phase sequences and companding, simulation results show better performance in PAPR with companding.

References

1. Zolghadrasli A, Ghamat MH (2007) A new blind PAPR reduction and error correction method in OFDM system using BCH codes. In: IEEE international conference on signal processing and communications, 2007. ICSPC 2007, pp 1051–1054, 24–27 Nov 2007
2. Han SH, Lee JH (2005) An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. Wireless Commun IEEE 12(2):56–65
3. Baxley RJ, Zhou T (2004) Power saving analysis of peak-to-average power ratio reduction in OFDM. IEEE Trans Consum Electron 50(3):792798
4. Vimal SP, Kumar KS (2011) A new SLM technique for PAPR reduction in OFDM systems. Eur J Sci Res 65(2):221–230. ISSN 1450-216X
5. Han SH, Lee JH (2005) An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. Mag IEEE Wireless Commun 12:56–65

Color Image Digital Watermarking

Neeraj Bhargava and Manish Mathuria

Abstract Due to the rapid development of digital technologies, the digital watermarking techniques have been widely used in research to solve some important issues of digital world. One of the most important issues is to apply digital watermarking on color images for copyright protection and content authentication. DWT is very popular for digital watermarking of gray image, but this research paper is the result of experiment of DWT on color images. Basically, a color image consists of R, G, and B channels, if image is in RGB color space. These color channels are matrices, which facilitate us to work on any channel independently of others. In this research paper, R-channel (or red component) is used to analyze the result. The experiment results show that the proposed algorithm keeps the quality of image well and is robust enough to common image processing operations.

Keywords Color image · Watermark · DWT · Red component · Haar wavelet

1 Introduction

Internet is a collection of digital data (digital image, video, and audio). Digital watermarking is common solution to ownership verification of digital data. These days color images can easily be clicked and shared via various technologies. The color image gives more information about the object compared to gray image. The

N. Bhargava (✉)
Department of Computer Science, School of Engineering and System Sciences, MDS University, Ajmer, Rajasthan, India
e-mail: drneerajbhargava@yahoo.co.in

M. Mathuria
Government Engineering College, Ajmer, Rajasthan, India
e-mail: manishmathuria@outlook.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_61,

astronomical and medical images are easily available over the Internet to share knowledge with people, but they all should be secured for identification of the owner. The main objective of digital watermarking is to provide image authentication to protect intellectual property rights.

Digital watermarking is the concept of embedding secret information in the raw data, which can be later extracted to confirm the identity of the owner. Actually, secret information is known as watermark, and it can be a serial number or image. Watermark can be visible or invisible; visible watermarks are usually logos, inlaid with image such that partially visible in the background of image. But invisible watermarks are secret information such that it is hidden inside the image and it can only be detected by an authorized party [1].

Digital watermarking techniques embed watermark invisibly and are known as 'Invisible Watermarking.' There are two very important requirements of invisible watermarking. They are

- Robustness: The watermarked image should be resistant to piracy attack and image processing operations.
- Security: The invisible watermark should be undetectable by unauthorized user. As well as retrieval of invisible watermark by authorized party should be readable [2].

This research paper is an effort to establish invisible watermarking and its requirements.

Basically, any transform technique is applied on the original image to change the frequency and then secret information (i.e., watermark) is embedded inside it. The two-dimensional discrete Fourier or cosine transform uses forward discrete transform function which is completely a function of the spatial frequency and does not include any information about the pixel. All basic functions are sinusoids; instead, wavelet transform uses other kinds of basic function, that is, Haar, Daubechies. These basic functions of wavelet transform are termed as mother wavelet. The mother wavelet that can be represented in terms of low- and high-pass filter coefficients is merely applicable for the transformation. It is very important for researcher to select appropriate mother wavelet. This research paper is the experimental result of considering the Haar wavelet as the mother wavelet, to achieve discrete wavelet transform (DWT) [3] (Fig. 1).

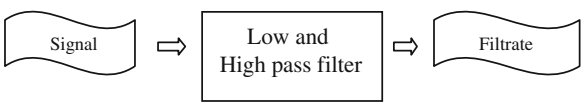


Fig. 1 The basic filtering approach of DWT [4]

2 Proposed Work

A good quality digital watermarking of color images is principal objective of this research paper. After reviewing much of the literature, it is analyzed that still there is a requirement for best digital watermarking approach for color images. Most of the published research papers deduce only about the gray image digital watermarking. The color image digital watermarking is slightly researched and published. This research paper proposed a new approach to impress invisible digital watermark inside the color image as well as maintain the quality of image. And extraction of watermark is also accomplished with the care of watermark to be recognizable.

2.1 Color Component Decomposition

All digital images are nothing but a digital rectangular matrix. MATLAB is computational software which is used to compute matrix manipulations. It provides some advance packages and tools for image processing. Initially, an original image is provided as input to the watermarking algorithm. Naturally, all the physical objects are colored, and all the colors can be derived from the three basic colors, that is, red, green, and blue. These basic colors are known as primary colors, and the color space is called RGB color space. The RGB produces the combination of colors when it composes into a single image. Similarly, red (R), green (G), and blue (B) components can be decomposed from the color image. MATLAB allows us to separate RGB components from the input color image using simple matrix operation. Basically, separate components are individual matrices. It allows us to work on any component separately. Algorithm of this research paper uses red (R) component to hide secret information [4, 5].

2.2 DWT Using Haar Wavelet

Basically, digital wavelet transform (DWT) requires down-sampling, up-sampling, low-pass filtering, and high-pass filtering. These all four operations are applied on the one dimension which is a row or column but one at a time.

- Down-sampling: It is a process to skip every alternate sample of input matrix.
- Up-sampling: It is just a reverse process of down-sampling. It inserts single zero between the samples.
- Filtering: A specific discrete filter has specific impulse response (i.e., low pass or high pass). And impulse response depends on the wavelet type selection. The impulse responses are also known as Wavelet Filter Coefficients. In the filtering

process, convolution operation applied on digital signal and digital filter impulse response [4].

MATLAB provides an inbuilt function named 'dwt2' for DWT. It takes two inputs by default; the first is the name of input image, and the second is the name of mother wavelet to be used. And it produces four output arguments as follows: 'A' (approximated component), 'H' (horizontal component), 'V' (vertical component), and 'D' (diagonal component).

2.3 Embedding Watermark Inside Image

The collection of detailed components (A, H, V, and D) by DWT operation allows algorithm to select appropriate component for further decomposition. But this research paper only deals with one-level decomposition of digital image and uses A (approximated component) to perform watermarking.

The watermarking is nothing but an addition and multiplication of frequency components of original image with watermark. After calculation operation, the new matrix is generated with new frequency values.

2.4 Inverse Discrete Wavelet Transform

Finally, to get resultant image (i.e., watermarked image) after embedding process, algorithm performs Inverse Discrete Wavelet Transform (IDWT). It can be accomplished by MATLAB's inbuilt function 'idwt2' for two-dimensional transformations. This function requires five inputs; the first four are 'A', 'H', 'V', and 'D'; the fifth input is the name of mother wavelet [6, 7, 8].

3 Experiment Data

3.1 Selection of Input Original Images

This proposed algorithm is tested using four different color images. All these images are different from each other by objects and colors. To analyze the effect of variation in different color intensities, experiment is applied on four distinct images. It can take input image of any size (Fig. 2).

Fig. 2 The original images for watermarking. **a** Computer. **b** Flower. **c** Gate. **d** Guitar

Fig. 3 The watermark image

3.2 Selection of Input Watermark Image

The experiment is performed by using watermark as binary image. It can be of any size because the developed program can easily convert it to appropriate size (Fig. 3).

Fig. 4 The output images. **a** wComputer. **b** wFlower. **c** wGate. **d** wGuitar

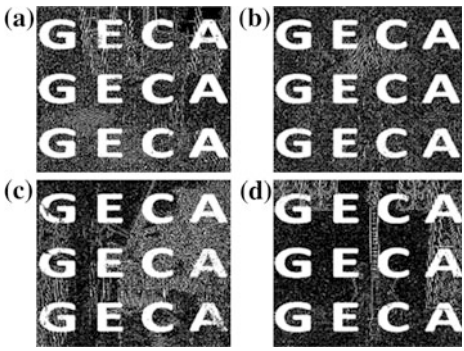


Fig. 5 The recovered images. **a** rw_Computer. **b** rw_Flower. **c** rw_Gate. **d** rw_Guitar

3.3 Output Images After Watermark Embedding

Figure 4

3.4 Recovered Watermark

Figure 5

4 Results

The peak signal-to-noise ratio (PSNR) is most commonly employed to check the quality of the image. Typically, PSNR values which lie between 30 and 50 represent good quality of the image. In which, higher PSNR value is better. The correlation coefficient of the two images of identical size tells how much similar are the two images. If the value of correlation coefficient is near to 1, then the two images are very similar. This value is 1 for same images (Table 1).

Table 1 PSNR and correlation coefficients

Sr. No.	Original image	Watermarked image	PSNR	Correlation coefficients
1.	Computer.jpg	wComputer.jpg	38	0.9968
2.	Flower.jpg	wFlower.jpg	41	0.9976
3.	Gate.jpg	wGate.jpg	36	0.9949
4.	Guitar.jpg	wGuitar.jpg	40	0.9975

5 Conclusion

The color image watermarking is must to secure digital images. After applying the watermarking on the images, the user can easily share it. Result shows that the image quality is not affected by watermarking operation. R, G, and B channels can be used separately for embedding the watermark. Red color component is very effective because invisible watermark is not detectable easily. This work can further be implemented or researched for many image applications.

References

1. Khatkale PB, Jadhav KP, Khasne MV (2012) Digital watermarking technique for authentication of color image. Int J Emerg Technol Adv Eng 2(7), ISSN 2250–2459, July 2012
2. Samiappan V, Ammasai K (2012) Robust digital image watermarking for color images. Eur J Sci Res 76(1):117–126, ISSN1450–216X, Euro Jounals Publishing, Inc. 2012
3. Weiss LG (1994) Wavelet and wideband correlation processing. IEEE Signal Process Mag 11(1):13–32 Jan 1994
4. Nuruzzaman M (2005) Digital image fundamentals in MATLAB. Book, ISBN 1-4208-6965-5 (sc)
5. Gunjal BL (2011) Wavelet based color image watermarking scheme giving high robustness and exact correlation. Int J Emerg Trands Eng Technol I(1), ISSN 2248-9592
6. Anuradha RPS (2012) DWT based watermarking algorithm using haar wavelet. Int J Electron Comput Sci Eng 1(1), ISSN 2277-1956
7. Singla K, Kaur S (2011) Invisible digital watermarking for color images. IJAEST 10(2):270–274
8. Yusof Y, Khalifa OO (2007) Digital watermarking for digital images using wavelet transform. In: Proceeding of IEEE-2007

Damped Sinusoidal Basis Functions–Based Transformation for Image Compression

Tera Hindu and Suyakanth V. Gangashetty

Abstract Normally, the discrete cosine transformation (DCT) and the discrete wavelet transformations (DWT) are used for the purpose of image compression. Some of the limitations of the wavelet functions are the choice of wavelet function, wavelet order, the selection of threshold value, and setting the optimal level of decomposition. In Fourier transform, the basis functions are pure sinusoids. In the Bessel transformation, the basis functions are damped sinusoids. The application of 2-dimensional Bessel transform (BT) for image compression is an experiment and has been applied ahead of wavelet transform, because of its variable structure in single decomposition can achieve variable amount of compression. The performance of image compression using BT is compared to that of wavelet transform and discrete cosine transform (DCT). The results are viewed in terms of peak signal to noise ratio (PSNR) and mean square error (MSE) for various compression ratios (CRs).

Keywords Discrete cosine transform · Discrete wavelet transform · Bessel transformation · Compression ratio

1 Introduction

Storing an image is not always a problem because now computers are capable enough to handle large amounts of data. However, it is often desirable to use the limited resources more efficiently. The electronic gadgets like digital cameras have

T. Hindu (✉) · S. V. Gangashetty
International Institute of Information Technology (IIIT), Gachibowli, Hyderabad, Andhra Pradesh 500032, India
e-mail: hindutera10@gmail.com

S. V. Gangashetty
e-mail: svg@iiit.ac.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_62,

totally insufficient amount of memory where the importance of the compression of image is still greatly felt. The image compression techniques are mainly of two types: I. transform based (DCT, Fourier transform, wavelet transform) and II. nontransform based (pulse code modulation, differential pulse code modulation) [1]. In transform-based compression techniques, compression can be achieved by transforming the image data by projecting it on orthogonal basis functions and then encoding the resulted coefficients [1]. Transform-based image compression techniques are widely used. In recent years, the discrete wavelet transform (DWT) is used as a powerful technique for image compression, especially for images on higher resolutions [2]. The advantage of using DWT is that the entire image is transformed and compressed as a single data object rather than block by block as in discrete cosine transform (DCT)–based system, which allows for a uniform distribution of compression error across the entire image [3]. No image compression technique can be expected to work well for all class of images, and for different applications, suitable image compression techniques are selected.

The main idea of image compression is to minimize the correlation among the adjacent pixels [4], as this property is satisfied by the Bessel basis functions which inspired us to work on applying BT toward image compression. BT has been used as a technique for feature extraction of speech signals used for various speech processing applications [5].

The paper is organized as follows: Section 2 gives the background of image compression based on DCT and DWT. Section 3 describes the formulation of 2-dimensional Bessel transformation (BT). The band limit structure and procedure how the BT is used for image compression are described in Sect. 4. The decomposed images of DCT, DWT, and band-limited images of BT are compared in Sect. 5. Section 6 gives the summary of the studies on areal compression and conclusions from our studies.

2 Background

In this section, we discuss DCT- and DWT-based compression techniques in brief.

2.1 DCT for Image Compression

DCT-based compression technique is a simple and effective way to achieve image compression [6]. The compression through DCT is done by making image into $N \times N$ blocks, and each block is DCT transformed independently to give $N \times N$ DCT coefficients per block [7–9]. Most of the DCT coefficients in each block are near to zero. Now, the compression is achieved by setting those coefficients to zero, and the remaining coefficients are represented with reduced precision that is determined by quantizer scale [10]. The quantized coefficients are encoded for

transmission. The primary limitation of DCT-based compression system is the block-based segmentation [11]. The image quality is poorer for high compression ratios (CRs) [12].

2.2 DWT for Image Compression

The DWT is developed as a standard tool in image compression because of its data reduction capability [13]. The wavelet decomposition consists of consecutive operation of row-wise and column-wise of input image data followed by down-sampling to obtain wavelet coefficients and also called sub-bands [14]. At one-level decomposition of input data DWT comprises four sub-bands LL (low/low), LH, HL, and HH (high/high) [15]. There are three types of detail images for each resolution: HL contains horizontal components of image, LH gives vertical components of image, and similarly HH gives diagonal components of image [15]. The LL sub-band is used for further levels of decomposition. After decomposing the image for certain levels, the compression is achieved by making the wavelet coefficients to zero whose magnitude is less than threshold. The LL sub-band at final level is not involved in compression. Thus, a typical 2-D DWT is used for image compression [10]. The image compression based on DWT can provide better image quality than DCT, especially on higher CRs [12]. The choice of wavelet function, wavelet order, and setting the optimal level of decomposition are the limitations [16].

3 Bessel Transform

As BT for 1-D signal presented by authors in [17], we map it to 2-D. A finite-duration signal $x(t_1, t_2)$ in the interval $0 < t_1 < a_1,\ 0 < t_2 < a_2$, is represented by using the zero-order Bessel functions of the first kind in an infinite Bessel series as

$$x(t_1, t_2) = \sum_{m_1=1}^{a_1} \sum_{m_2=1}^{a_2} c(m_1, m_2) J_0\left(\frac{\alpha_{m_1}}{a_1} t_1\right) J_0\left(\frac{\alpha_{m_2}}{a_2} t_2\right) \tag{1}$$

where $\alpha_{m_1}, \alpha_{m_2}$ are the roots of $J_0(t_1) = 0$ and $J_0(t_2) = 0$ and $m_1 = \{1, 2, 3 \ldots\ldots a_1\}$, $m_2 = \{1, 2, 3 \ldots\ldots a_2\}$.

The coefficients $c(m_1, m_2)$ are given as:

$$c(m_1, m_2) = \frac{4 \int_0^{a_1} \int_0^{a_2} t_1 t_2 x(t_1, t_2) J_0\left(\frac{\alpha_{m_1}}{a_1} t_1\right) J_0\left(\frac{\alpha_{m_2}}{a_2} t_2\right) \mathrm{d}t_1 \mathrm{d}t_2}{(a_1 a_2)^2 [J_1(\alpha_{m_1}) J_1(\alpha_{m_2})]^2} \tag{2}$$

where $J_1(\alpha_{m_1})$ and $J_1(\alpha_{m_2})$ are first-order Bessel functions.

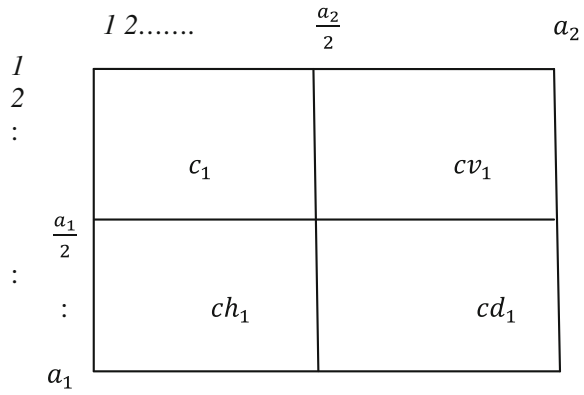


Fig. 1 2^{-1} BT-band structure

The Bessel coefficients $c(m_1, m_2)$ are unique for a given signal $x(t_1, t_2)$ [18]. The intervals between successive zero crossings of the Bessel function $J_0(t)$ increase slowly with time and approaches to π in the limit. However, unlike the sinusoidal basis functions in the Fourier series, the Bessel functions decay over time. This feature of the Bessel functions makes the Bessel series expansion suitable for the analysis of nonstationary signals [19–21].

4 BT for Image Compression

4.1 Band-Limiting Structure of BT

As BT is flexible for signal band limiting [21], the input image is Bessel transformed, and the four band-limited images are constructed using the range of Bessel coefficients selected shown in Fig. 1, where $c(m_1, m_2)$ is assumed to be size that is equal to that of input image signal, $c_1 = c\left(1 : \frac{a_1}{2},\ 1 : \frac{a_2}{2}\right)$, $cv_1 = c\left(1 : \frac{a_1}{2}, \frac{a_2}{2} : a_2\right)$, $ch_1 = c\left(\frac{a_1}{2} : a_1,\ 1 : \frac{a_2}{2}\right)$ and $cd_1 = c\left(\frac{a_1}{2} : a_1,\ \frac{a_2}{2} : a_2\right)$. We name the structure of decomposition in Fig. 1 to 2^{-1} BT-band structure.

This structure of decomposition in BT is equivalent to that of one-level decomposition using DWT. The structure of image bands can be varied by choosing the different range of Bessel coefficients; for example, the decomposition shown in Fig. 2 gives 2^{-2} BT-band structure.

4.2 Formulation of Band-Limiting Structure

We assume that band-limited images $\{sv_1, sv_2, \ldots..\}$, $\{sh_1, sh_2, \ldots..\}$, and $\{sd_1, sd_2, \ldots..\}$, are obtained from $\{cv_1, cv_2, \ldots..\}$, $\{ch_1, ch_2, \ldots..\}$, and $\{cd_1, cd_2, \ldots..\}$. The band-limited images obtained for 2^{-1} BT-band structure

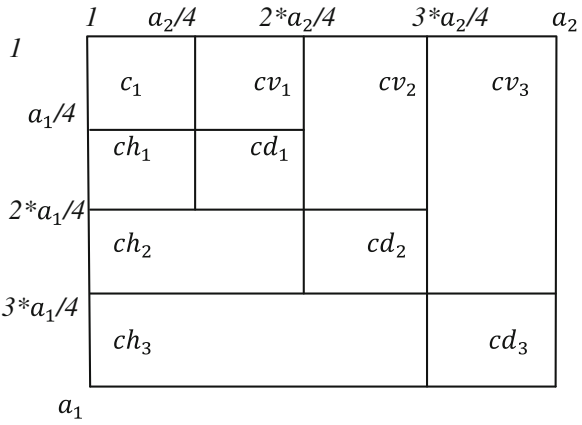


Fig. 2 2^{-2} BT-band structure

are sv_1, sh_1 and sd_1 whose dimensions (b_1, b_2) are half of the dimensions of original image and are formulated as follows:

$$sh_1(t_1, t_2) = \sum_{m_1=a_1/2}^{a_1} \sum_{m_2=1}^{a_2/2} ch_1(m_1, m_2) J_0\left(\frac{\alpha_{m_1}}{b_1} t_1\right) J_0\left(\frac{\alpha_{m_2}}{b_2} t_2\right) \tag{3}$$

where α_{m_1}, α_{m_2} are the roots of $J_0(t_1) = 0$ and $J_0(t_2) = 0$, $m_1 = \{a_1/2 \ldots a_1\}$, $m_2 = \{1, 2 \ldots a_2/2\}$, $b_1 = \frac{a_1}{2}$ and $b_2 = \frac{a_2}{2}$.

$$sv_1(t_1, t_2) = \sum_{m_1=1}^{a_1/2} \sum_{m_2=a_2/2}^{a_2} cv_1(m_1, m_2) J_0\left(\frac{\alpha_{m_1}}{b_1} t_1\right) J_0\left(\frac{\alpha_{m_2}}{b_2} t_2\right)$$

where $m_1 = \{1, 2 \ldots a_1/2\}$, $m_2 = \{a_2/2 \ldots a_2\}$, $b_1 = \frac{a_1}{2}$ and $b_2 = \frac{a_2}{2}$.

Similarly, sd_1 is given as

$$sd_1(t_1, t_2) = \sum_{m_1=a_1/2}^{a_1} \sum_{m_2=a_2/2}^{a_2} cd_1(m_1, m_2) J_0\left(\frac{\alpha_{m_1}}{b_1} t_1\right) J_0\left(\frac{\alpha_{m_2}}{b_2} t_2\right) \tag{5}$$

where $m_1 = \{a_1/2 \ldots a_1\}$, $m_2 = \{a_2/2 \ldots a_2\}$, $b_1 = \frac{a_1}{2}$ and $b_2 = \frac{a_2}{2}$.

Similarly, the formulations for different BT-band structures are made by changing the values of m_1, m_2, b_1 and b_2.

4.3 Procedure of Image Compression

The basic steps for implementing the image compression using BT is outlined in following steps, without taking the influence of quantizer and entropy encoder into account, in order to differentiate the compression performance of BT to DWT.

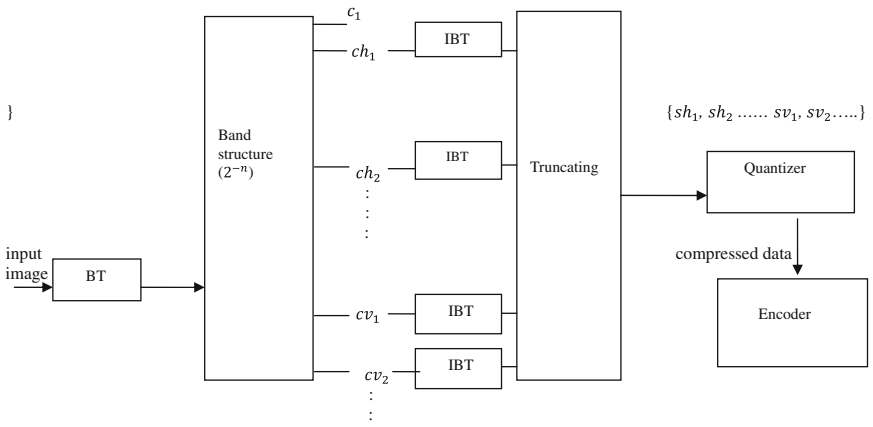


Fig. 3 Block diagram of transform coding system for compressed data

1. BT is applied on image, and band structure is selected.
2. By keeping c_1 set of coefficients intact, the band-limited images are obtained for $\{cv_1, cv_2, \ldots..\}$, $\{ch_1, ch_2, \ldots..\}$, and $\{cd_1, cd_2, \ldots..\}$ set of coefficients.
3. The threshold should be set, and pixel values of those band-limited images below the threshold are truncated to zero.
4. The Bessel coefficients are calculated from the manipulated band-limited images, and they are concatenated with c_1 to form similar band structure.
5. The image is reconstructed by applying inverse BT.

Steps 1, 2, and 3 are described in Fig. 3, and steps 4 and 5 are described in Fig. 4.

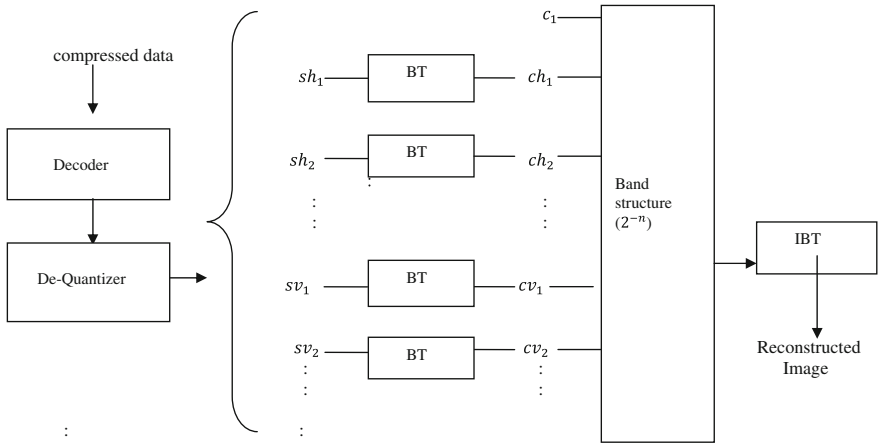


Fig. 4 Block diagram of reconstructing image from compressed data

5 Results

We perform experiment on an aerial view image [22], which consists of large number of edges as shown in Fig. 5. The image quality is evaluated in terms of PSNR and MSE calculated for reconstructed signal.

$$\mathrm{MSE} = \frac{1}{a_1 a_2} \sum_{j=1}^{a_1} \sum_{i=1}^{a_2} \left(x_{ij} - \hat{x}_{ij}\right)^2 \tag{6}$$

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255}{\mathrm{MSE}}\right) \tag{7}$$

MSE represents the difference between original image x and reconstructed image $\hat{x}$, where a_1 and a_2 are dimensions of x and $\hat{x}$. The CR is defined as the ratio of number of nonzero elements in the original image to the number of nonzero elements of compressed image [23].

Now, image compression is performed on aerial view image. DCT with block size 8×8 is used for compression, and values of PSNR and MSE are listed in Table 1. Values listed in Tables 2 and 3 illustrate the performance of DWT and BT for CRs (2.4, 3.2, 10, and 20), respectively. The reconstructed images of compressed data based on DCT are shown in Fig. 6. By observing the values listed in Table 2, we notice that the third-level decomposition is optimal level for

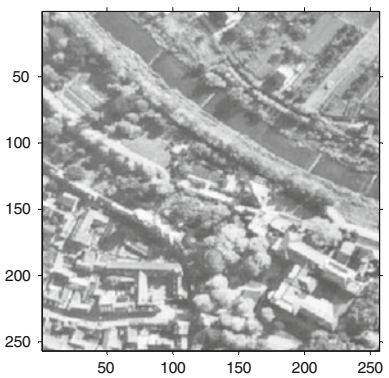


Fig. 5 Original aerial view image of 256 × 256 pixel

Table 1 PSNR in decibels and MSE results of CRs (2.4, 3.2, 10, and 20) for DCT-based compression

CR	PSNR	MSE
2.4	55.72	0.00068
3.2	52.56	0.0014
10	36.75	0.0538
20	32.33	0.1503

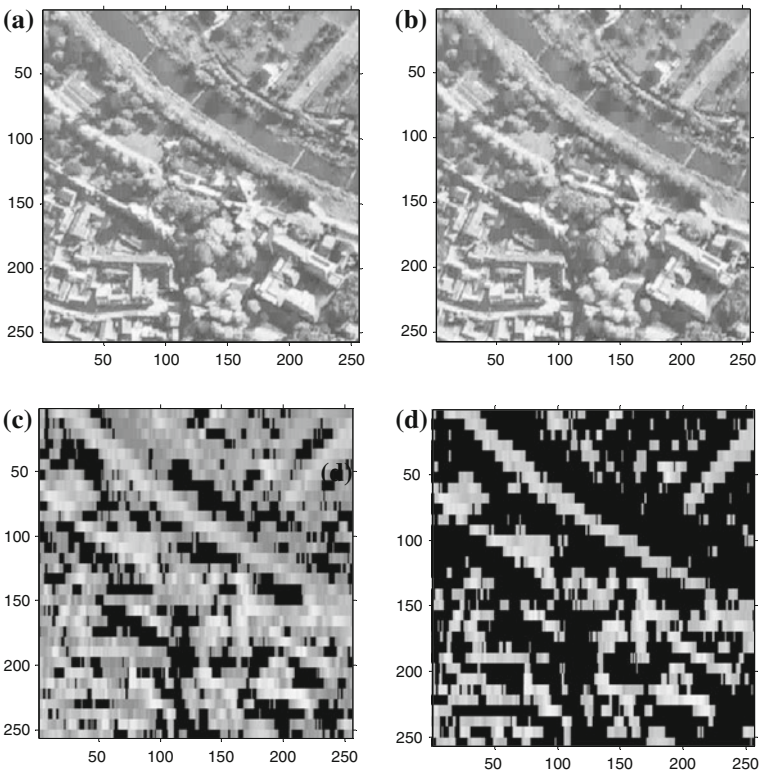


Fig. 6 Reconstructed images of DCT compressed data with CRs. **a** 2.4, **b** 3.2, **c** 10, and **d** 20

Table 2 PSNR in decibels and MSE results of CRs (2.4, 3.2, 10, and 20) for DWT-based compression

Number of levels of decomposition	CR							
	20		10		3.2		2.4	
	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
1	–	–	–	–	53.05	0.0013	56.87	0.00052
2	–	–	47.59	0.0044	55.79	0.00067	58.69	0.00034
3	45.97	0.0064	48.44	0.0036	55.95	0.00064	59.05	0.00032
4	46.20	0.0061	48.70	0.0034	55.95	0.00064	59.01	0.00032
5	46.27	0.0060	48.63	0.0035	55.97	0.00065	59.04	0.00032
6	46.25	0.0060	48.62	0.0035	55.94	0.00064	59.00	0.00032

achieving CR 2.4, and fifth-level decomposition is optimum level for achieving CRs 3.2, 10, and 20. The reconstructed images of compressed data for optimum decomposition levels of DWT are shown in Fig. 7.

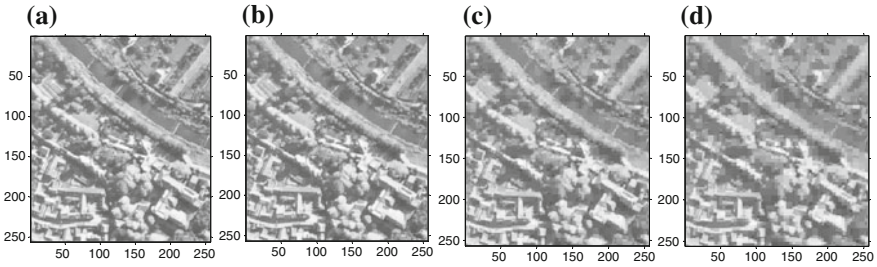


Fig. 7 Reconstructed images with CRs: **a** 2.4 obtained for third-level decomposition. **b** 3.2 obtained for fifth-level decomposition. **c** 10 obtained for fifth-level decomposition, and **d** 20 obtained for fifth-level decomposition of DWT

Table 3 PSNR & DPSNR in decibels and MSE and DMSE results of CRs (2.4, 3.2, 10, and 20) for BT-based compression

FBT-band structure	CR							
	20		10		3.2		2.4	
	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE
2^{-1}	–	–	–	–	55.05	0.000798	58.31	0.00037
2^{-2}	–	–	49.61	0.0028	54.88	0.00082	55.95	0.00064
2^{-3}	46.7	0.0054	48.21	0.0038	50.15	0.0025	50.5	0.0023
2^{-4}	43.04	0.0126	43.37	0.0117	43.94	0.0103	43.99	0.0102
2^{-5}	39.17	0.0308	39.2	0.039	39.53	0.0284	39.58	0.0283

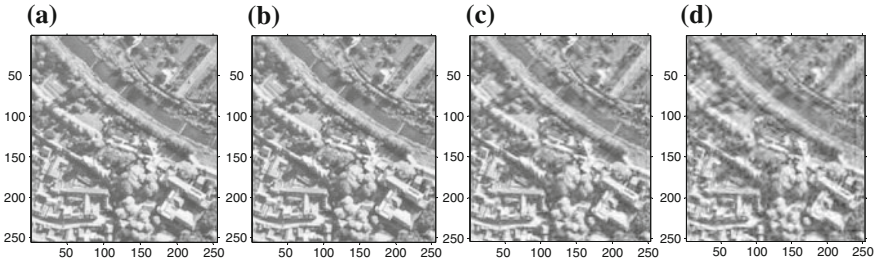


Fig. 8 Reconstructed images for maximum PSNR with CRs. **a** 2.4 obtained for 2^{-4} BT-band structure. **b** 3.2 obtained for 2^{-3} BT-band structure. **c** 10 obtained for 2^{-2} BT-band structure, and **d** 20 obtained for 2^{-2} BT-band structure

The values listed in Table 3 indicate that the maximum PSNR values obtained for CRs 2.4, 3.2, 10, and 20 when 2^{-1}, 2^{-1}, 2^{-2}, and 2^{-3} BT-band structures are used. The reconstructed images based on BT with maximum PSNR are shown in Fig. 8.

6 Summary and Conclusions

On comparison of the values of Tables 1, 2, and 3, it is evident that PSNR values in case of BT are better than PSNR values in case of DWT and DCT. The image is preserved more in BT than in DWT and DCT for higher CRs which is evident by comparing the images of Figs. 8–7. Thus, the compression based on Bessel transformation can replace DWT and DCT. The studies show that the damped sinusoidal basis functions–based transformation, namely Bessel transformation, can play significant role in image compression.

References

1. Averbunch A, Lazar D, Israeli M (1996) Image compression using wavelet transform and multi resolution decomposition. IEEE Trans Image Process 5(1):4–15
2. Hilton ML, Jawerth BD, Sengupta A (1994) Compressing still and moving images with wavelets. Multimedia Syst 2(3):218–227
3. Xiang Z, Ramchandran K, Orchard MT, Zhang YQ (1999) A comparative study of DCT- and wavelet-based image coding. IEEE Trans Circ Syst Video Technol 9:692–695
4. Rabbani M, Jones PW (1991) Digital image compression techniques. SPIE Optical Engineering Press, Bellingham
5. Pachori RB, Sircar P (2006) Speech analysis using fourier-bessel expansion and discrete energy separation algorithm. IEEE Proc Digit Signal Process Workshop, pp 423–428
6. Grgic S, Mrak M, Sprljan N (1999) DCT lab: didactic tool for the study of image compression (in Croatian). Department of radio communication microwave engineering, University of Zagreb Internal Pub
7. Ahmed N, Natarajan T, Rao KR (1974) Discrete cosine transform. IEEE Trans Comput C-23:90–93
8. Pennebaker WB, Mitchell JL (1993) JPEG still image data compression standards. Van Nostrand Reinhold, USA
9. Rao KR, Yip P (1990) Discrete cosine transforms-algorithms, advantages, applications. Academic Press, NewYork
10. Grgic S, Grgic M, Zovko-Cihlar B (2001) Performance analysis of image compression using wavelets. IEEE Trans Ind Electron 48(3):682–695
11. Xiong Z, Guleryuz O, Orchard MT (1996) A DCT-based embedded image coder. IEEE Sign Process Lett 3:289–290
12. Bauer S, Zovko-Cihlar B, Grgic M (1996) The influence of impairments from digital compression of video signal on perceived picture quality. In: Proceedings of 3rd international workshop image and signal processing, Manchester, UK, pp 245–248
13. Antonini M, Barlaud M, Mathieu P, Daubechies I (1992) Image coding using wavelet transform. IEEE Trans Image Process 1(2):205–220
14. Donoho DL (1995) De-noising by soft-thresholding. IEEE Trans Inf Theory 41(3):613–627
15. Grgic S, Kers K, Grbic M (1999) Image compression using wavelets. Int Symp Ind Electron 1:99–104
16. Mandal MK, Panchanathan S, Aboulnasr T (1996) Choice of wavelets for image compression. Lect Notes Comput Sci 1133:239–249
17. Schroeder J (1993) Signal processing via fourier-bessel series expansion. Digit Signal Process 3:112–124

18. Chen CS, Gopalan K, Mitra P (1985) Speech signal analysis and synthesis via fourier-bessel representation. IEEE Int Conf Acoust Speech Signal Process 2:497–500
19. Gurgen FS, Chen CS (1990) Speech enhancement by fourier-bessel coefficients of speech and noise. IEEE Proc Comput Speech Vision 137(5):290–294
20. Pachori RB, Sircar P (2008) EEG signals analysis using FB expansion and second-order TVAR. Signal Process 88:415–420
21. Pachori RB, Sircar P (2010) Analysis of multicomponent AM-FM signals using FB-DESA method. Digit Signal Process 20:42–62
22. The USC-SIPI Image Database (2010) Signal and Image Processing Institute, University of Southern California, Miscellaneous, vol. 3. Available: http://sipi.usc.edu/database/index.html . Accessed 10 Apr 2010]
23. Ragab AS, Mohamed ASA, Hamid MS (1998) Efficiency of analytical transforms for image compression. IEEE conf Radio Sci, pp B16/1-B1610

Energy Model for the Configured MSP430F1612 on a TELOSB Mote with the Help of Contiki

K. Senthil Babu, Darshan Virupaksha, Shachi P. Mudgal and C. Nagaraja

Abstract Wireless sensor networks (WSN) are attracting a wide range of application because of its exponential growth in its performance. However, there are certain drawbacks with respect to the power available in the node. In this paper, we present a hardware configuration of TelosB mote with the help of Contiki OS which improves the performance of the mote by supporting with additional inbuilt flash memory. The paper includes the energy calculation of the new hardware configured. The existing hardware MSP430F1611 provides 48 kB of flash memory which is replaced by MSP430F1612, and Contiki is one such OS which is specifically designed for WSN. In order to provide more flexibility to the application developer, requires of Contiki on a TelosB mote. Contiki support for this modified TELOSB is not available, thereby making an attempt to understand Contiki and port to the modified TELOSB.

Keywords Contiki · Energy model · MSP430F1612

K. Senthil Babu (✉) · C. Nagaraja
SK University Anantapur, Anantapur, Andhra Pradesh, India
e-mail: ksenthilbabu80@gmail.com

C. Nagaraja
e-mail: c_nagaraja@yahoo.co.uk

D. Virupaksha · S. P. Mudgal
AIEMS Bangalore, Bangalore, Karnataka, India
e-mail: darshan.virupaksha@gmail.com

S. P. Mudgal
e-mail: shachimdgl@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_63,

1 Introduction

Advances in wireless communication technology enable the deployment of networks of small sensors. These sensor networks have applications in military monitoring, health, industrial control, weather monitoring, commodity tracking, home control, etc. A critical constraint on sensors networks is that sensor nodes employ batteries. A second constraint is that sensors will be deployed unattended and in large numbers, so that it will be difficult to change or recharge batteries in the sensors. Therefore, all systems, processes and communication protocols for sensors and sensor networks must minimize power consumption.

The existing research on energy consumption of sensors is usually based on either theoretical models or computer simulations. Some studies measure the power consumption of sensors using an oscilloscope to determine power consumption in each of the several states; however, tests were conducted over short time intervals and with no statistical validation. Moreover, the small physical size and low per-device cost limit the complexity of the system. Typical sensor devices are equipped with 8-bit microcontrollers, code memory on the order of 100 kB, and less than 20 kB of RAM. Moore's law predicts that these devices can be made significantly smaller and less expensive in the future. While this means that sensor networks can be deployed to greater extents, it does not necessarily imply that the resources will be less constrained.

The characteristics of WSN impose additional challenges on OS design for WSN. Application programmers can invoke different OS services through system calls. An OS multiplexes system resources in two ways, that is, in time and in space [1]. Considering the resource constraints of typical sensor nodes in a WSN, a new approach is required for OS design in WSN. For a designer of an operating system for sensor nodes, the challenge lies in finding lightweight mechanisms and abstractions that provide a rich enough execution environment while staying within the limitations of the constrained devices

2 Contiki: Architecture, Power Save Mode, Program Model

Contiki [2, 3] is a small highly portable multitasking computer operating system developed for use on a number of memory-constrained networked systems ranging from 8-bit computers to embedded systems on microcontrollers, including sensor network motes. A typical Contiki configuration needs 2 kB of RAM and 40 kB of ROM. Contiki provides IP communication for both IPv4 and IPv6. Contiki provides dynamic loading and unloading of individual programs and services.

The Contiki OS follows the modular architecture. Contiki combines the benefits of both event-driven systems and preemptible thread, thus following the hybrid system. In sensor networks, being able to power down the node when the network

is inactive is the best way to reduce energy consumption. Power conservation mechanisms depend on both the applications [4] and the network protocols. Contiki supports preemptive multithreading model. Multithreading is implemented as a library on top of the event-driven kernel [5]. The library can be linked with applications that require multithreading. The Contiki multithreading library is divided into two parts: a platform-independent part and a platform-specific part. The platform-independent part interfaces with event kernel and the platform-specific part of the library implements stack switching and preemption primitives. For multithreading, Contiki uses protothreads [5]. Protothreads are designed for severely memory constraint devices because they are stack-less and lightweight. The main features of protothreads are as follows: very small memory overhead (only two bytes per protothread), no extra stack for a thread, and highly portable.

Contiki supports a rich set of communication protocols. In Contiki, an application can use both versions of IP, that is, IPv4 and IPv6. Contiki provides an implementation of *u*IP, a TCP/IP protocol stack for small 8-bit microcontrollers. The *u*IP implementation has the minimum set of features needed for a full TCP/IP stack. *u*IP supports TCP, UDP, ICMP, and IP protocols. Contiki provides another lightweight layered protocol stack, called Rime, for network-based communication. Rime provides single hop unicast, single hop broadcast, and multi-hop communication support.

3 Simulators

A big issue for people who are working in wireless sensor network-related areas is lack of hardware, and also there is need for simulators for code testing before fusing the code into the hardware. Different simulators are available for testing such as MSPsim—provides a strong platform for testing and debugging (all Contiki codes for TELOSB can be tested on MSPsim), COOJA—also available to test network algorithms designed and implemented in Contiki. Porting [6] is the process of adapting software so that an executable program can be created for a computing environment that is different from the one for which it was originally designed (e.g., different CPU, operating system, or third party library).

4 Energy Models

4.1 The Classical Energy Consumption Model

Heinzelman et al. [7] proposed an energy consumption model for sensors based on the observation that the energy consumption would likely be dominated by the data communications subsystem. Table 1 reproduces their model. The model

Table 1 Power consumption of different states

Radio mode	Energy consumption
Tx/Rx electronics	50 nJ/bit
Tx amplifier	100 pJ/bit/m^2
Idle	40 nJ/bit
Sleep	0

considers a low power consumption radio that was slightly better than some standard definitions, like Bluetooth. The model provides a commonly used starting point; however, the model has not been verified against the behavior of a physical radio in a wireless sensor network. When computing node energy consumption, the CPU and the sensors are consumers that may or may not be neglected, depending on the nature of the application.

4.2 Power State Model

In the power state model [8], the power levels are measured depending on the state and the measurements of a real wireless card on laboratory equipment.

The measured values only apply for the exact model of card. It does not take into account transient consumption from mode switches, power consumption of host (i.e., from packet processing) and usually assumes fixed output power (not normally used with transmission power control). Power consumption during different state is shown in Fig. 1.

5 Modified TELOSB

TELOSB [9] is developed at UC Berkeley and is widely used in research activities. TELOSB is based on 8 MHz, 16-bit RISC TI MSP430 processor with 10 kB RAM, 16 kB configuration EEPROM, and 48 kB flash memory. It draws 1.8 mA

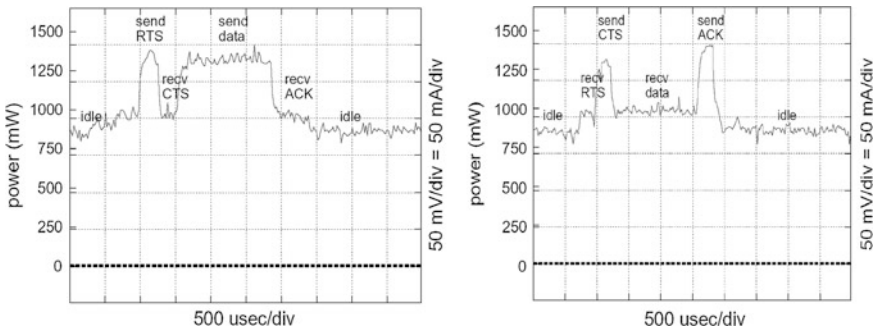


Fig. 1 Different transmission and receiver states

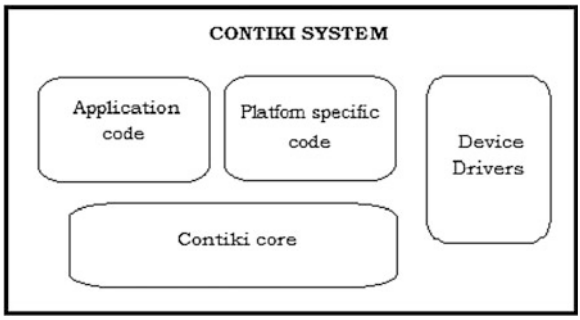


Fig. 2 Contiki tree

in active mode and 5.1 μA in sleep mode. It is an IEEE 802.15.4 compliant having a common CC2420 radio chip.

Many encryption and other security-based algorithms have large code size, requiring more flash memory. Hence, to overcome the issue of insufficient memory in the microcontroller flash, the microcontroller in the hardware was replaced with another MSP430 series-based controller, that is, **MSP430F1612**

5.1 Porting Contiki to Modified TELOSB

In order to increase flexibility to the embedded programmer, we have replaced the microcontroller with the one with a large size flash memory.

Contiki tree can be broadly segmented into three layers as shown in Fig. 2. In order to port Contiki to modified TELOSB, the platform-specific layer (*target folder)* and the lowest layer, that is, the core/CPU layer, must be studied and modified. The target folder consists of platform-specific files. These files act as bridge between the application program and the low-level device drivers of the hardware.

6 Implementation

As has already been mentioned in the previous section, the Contiki build system has been designed to

(a) make it easy to recompile applications for different platforms
(b) keep application code out of the Contiki directories

The general Contiki directory structure is shown below:

apps/architecture-independent applications
core/system source code
cpu/CPU-specific

doc/documentation
examples/example project directories
platform/platform-specific code
tools/software for building Contiki, sending files

After the Contiki core was ported to TELOSB [9], it had to be tested whether or not it was working properly. The simplest way to determine this was to write a simple API to switch the LEDs on and off. Referring to the circuit diagram of TELOSB, we see that the red, green, and blue correspond to pins P5.4, P5.5, and P5.6, respectively, that is, pins 4, 5, and 6 on port 5. A test code for radio function was written which basically transmits a packet and one of the LED blinks when the packet is transmitted. A simple procedure for energy consumption is also studied on the configured hardware using a power state model. The hardware setup is shown and the power consumed at various states is studied.

7 Experimental Result

The port presented several challenges which were tackled sequentially. The first challenge is having to work with a very large and diverse code base. The next challenge faced was of editing files containing the low-level implementation and debugging weird behavior of the motes.

Probably one of the most difficult obstacles was figuring out what exactly is responsible for various bugs in the system. Most often, the case was that the problem did not lie in just one source file, but in several. A Contiki kernel containing a "blinking lights" process was successfully fused to a TELOSB mote and successfully tested. The test code was subjected to various changes in terms of blink rate for individual LEDs as well (Fig. 3).

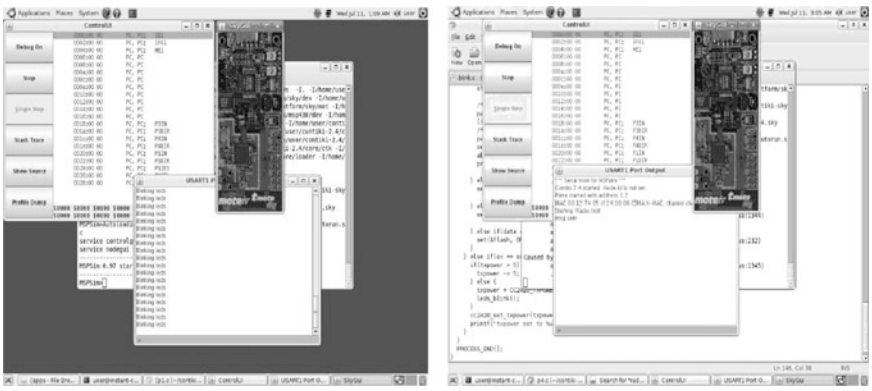


Fig. 3 Result of LED blink and result of radio-test code

The radio chip CC2420 [6] port in the Contiki kernel had to be tested; hence, we fused a radio-test code, which was programmed to transmit a single data packet. There were few issues found to radio packet transmission was put on infinite loop. The issues are yet to be solved. In the end, it can be concluded that a minimal port has been achieved keeping in mind the integration issues between the new microcontroller and the radio chip CC2420. The experimental setup consists of two laptops and two motes, with one being the transmitter and other being the receiver as shown in Fig. 4.

The current waveforms were obtained to verify power consumption characteristics of the hardware during transmission and reception as shown in Fig. 5.

In the process of transmission, we deliberately insert delay to ensure that oscillator output to be stable. During this period, the microcontroller is kept in low power mode to minimize power consumption. The most important factor for reducing power consumption is using the MSP430's clock system to maximize the time in LPM3. LPM3 power consumption is less than 2 μA typical with both a real-time clock function and all interrupts active. A 32-kHz watch crystal is used for the ACLK and the CPU is clocked from the DCO (normally off) which has a

Fig. 4 Transmitter and receiver experimental setup

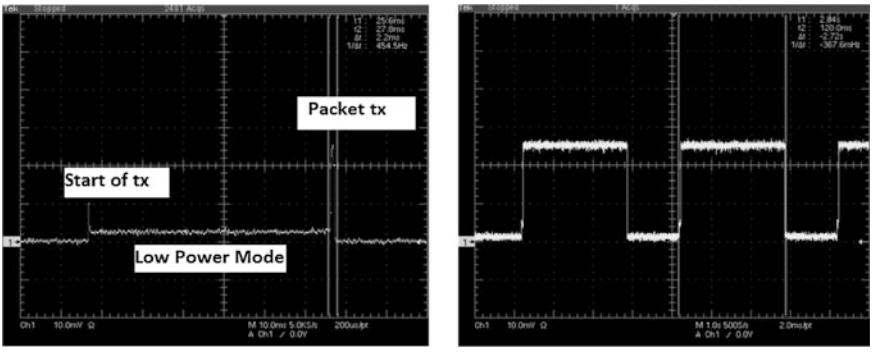


Fig. 5 Transmission and reception waveform

6-μs wakeup. The reception code uses onboard LEDs to indicate the reception of the packets. A LED blink will indicate that a packet is successfully received. The above waveform represents the power consumption of the hardware during reception. The power consumption is relatively high because of the use of LEDs.

8 Future Enhancements

The results presented in this paper are only initial results testing the basic functionalities of the Contiki system. We see at least three possible directions for future work, as described below.

Fully functional port

Limited Contiki port achieved has to be made fully functional where all the functionalities of the system are tested exhaustively.

Implementation and testing of unique LOW POWER IP STACK

Contiki as an operating system upholds a State-of-the-Art Networking Stack. Contiki has been at the forefront of exploring networking such as porting the first and smallest IPv6 compatible networking stack, while allowing easy access to develop or modify the stack at your own will. Thereby, constant development of the routing protocols in IP domain can be used on wireless sensor networks (WSN). Implementation and testing of these protocols on a Contiki kernel for the new hardware is a huge challenge and quite rewarding as well.

Energy Model and Energy measurements

The above results of the amount of energy consumed are a basic method to study the amount of power consumed at various states. Contiki uses the most intuitive way of controlling multiple tasks through protothreads, a very lightweight form of POSIX inspired threading, analogous to coroutines or Python generators. Simply define a protothread, and you are set to start your programming.

Contiki provides two intuitive ways of controlling the threads: events and timers. Events can initiate an appropriate response such as processing a data package when it is received by the radio. Timers allow you to schedule actions as required, such as reading some sensor data. Any lightweight programs make use of these intuitive structures which allow Contiki to save energy during idle periods. Developing energy models which take cognizance of the idle state of the nodes is vital for research and development activities in the field of sensor networks.

Contiki is field supporting a lot of R&D activities. Hence, there is wide range of scope for the improvement of Contiki as an OS to blend with the hardware and function efficiently.

9 Conclusion

The main contribution of this paper is to study the power consumed by the configured hardware at different stages and to identify the different situations in which the power is consumed. The modification brought out in this paper was the change in microcontroller. The porting is on the whole successful but integration issues with the radio chip are yet to be solved. At the moment, this port is not available in online repository of Contiki, so this can be integrated in future versions.

The Contiki OS was proved to be portable. The difficulty in porting lies in understanding the Contiki structure and flow of execution. The paper also tries to give a brief introduction on the key requirements of WSN, need for an operating system for efficient working WSN.

References

1. Levis P, Madden S, Gay D, Polastre J, Szewczyk R, Woo A, Brewer E, Culler D (2004) The emergence of networking abstractions and techniques in TinyOS. In: Proceedings of the NSDI
2. Dunkles A, Gronvall B, Voigt T (2004) Contiki-a lightweight and flexible operating system for tiny networked sensors
3. http://en.wikipedia.org/wiki/Contiki
4. Raghunathan V, Schurgers C, Park S, Srivastava M (2002) Energy aware wireless microsensor networks
5. Dunkels A, Schmidt O, Voigt T, Ali M (2006) Protothreads: simplifying event-driven programming of memory-constrained embedded systems. Swedish Institute of Computer Science, Box 1263, SE-16429 Kista, Sweden. TU Delft, Mekelweg 4, 2628 CD Delft, The Netherlands adam@sics.se, oliver@jantzer-schmidt.de, thiemo@sics.se, m.ali@tudelft.nl
6. http://en.wikipedia.org/wiki/Porting
7. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In IEEE Hawaii international conference on systems sciences
8. www.cs.jhu.edu/ ~ dholmer/600.647/energy-models
9. Stan A (2007) Porting the Core of the Contiki operating system to the TelosB and MicaZ platforms. Computer Science, International University Bremen, Campus Ring, Bremen, Germany, vol 1, p 28759

Design and Implementation of an Advanced Rectifier Stage Topology with Maximum Power Point Tracking for Hybrid Energy Systems

C. Kathirvel and K. Porkumaran

Abstract With the concern regarding the state of our deteriorating planet, environment-friendly solutions are becoming more prominent. This paper presents a new system configuration of the front-end rectifier stage consisting of Cuk–SEPIC fused converter for a hybrid wind/photovoltaic energy system. The configuration allows the two sources to supply the load separately or simultaneously depending on the availability of the energy sources. With this converter integration, additional input filters are not necessary to filter out high-frequency harmonics. The fused multi-input rectifier stage also allows maximum power point tracking (MPPT) to be used to extract maximum power from the wind and sun when it is available. A microcontroller-based control method is used for the MPPT of the energy system. For the reduction in tracking time, fuzzy logic control algorithm is implemented with the embedded microcontroller for improved performance. Operational analysis of the proposed system will be discussed in this paper. Simulation results are given to highlight the merits of the proposed circuit.

Keywords Renewable energy · Hybrid energy system · Rectifier stage · Fuzzy logic · MPPT

C. Kathirvel (✉)
Research Scholar, Department of Electrical and Electronics Engineering, Sri Ramakrishna Engineering College, Anna University, Chennai, India
e-mail: ckadhir@rediffmail.com

K. Porkumaran
Dr. N.G.P. Institute of Technology, Anna University, Chennai, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_64,

1 Introduction

The increasing concern of global warming and the depletion of fossil fuel reserves force to look at sustainable energy solutions to preserve the earth for the future generations. Other than hydro power, photovoltaic and wind energy holds the most potential to meet our energy demands. The wind energy is capable of supplying large amounts of power but its presence is highly unpredictable. The common inherent drawback of wind and photovoltaic systems is the intermittent natures that make them unreliable. By combining these two intermittent sources and by incorporating maximum power point tracking (MPPT) algorithms, the system's power transfer efficiency and reliability can be improved for the system. Several hybrid wind/PV power systems with MPPT control have been proposed and discussed in works [1, 3, 5]. Most of the systems in literature use a separate DC/DC boost converter connected in parallel in the rectifier stage to perform the MPPT control for each of the renewable energy sources [6].

A multi-input structure has been suggested by Chen et al. [3] and combines the sources from the DC end while still achieving MPPT for each renewable source. The structure proposed by Chen et al. [3] is a fusion of the buck and buck-boost converter. The systems in literature require passive input filters to remove the high-frequency current harmonics injected into wind turbine generators. In this paper, an alternative multi-input rectifier structure is proposed for hybrid wind/solar energy systems. The proposed design is a fusion of the Cuk and SEPIC converters. The features of the proposed topology are as follows: (1) it eliminates the need for separate input filters for PFC; (2) it can support step-up/step-down operations for each renewable source; (3) MPPT can be realized for each source; (4) individual and simultaneous operation is supported. Simulation results are provided to verify with the feasibility of the proposed system [7].

The circuit schematic of the proposed rectifier stage for a hybrid wind/PV system under consideration is given in Fig. 1.

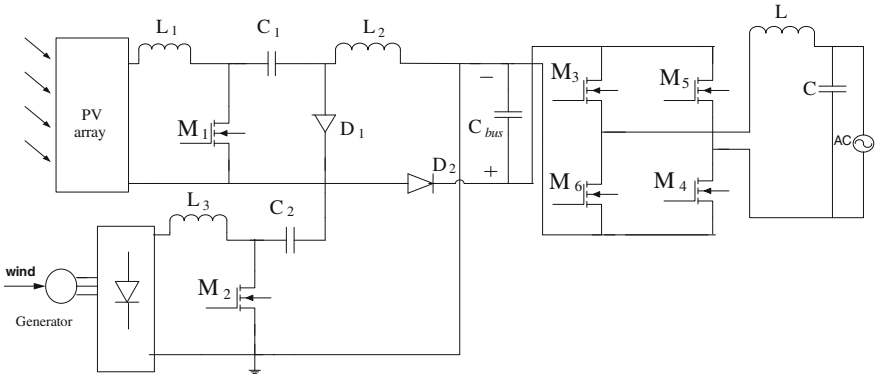


Fig. 1 Circuit schematic

2 Proposed Multi-Input Rectifier Stage

A system diagram of the proposed rectifier stage of a hybrid energy system is shown in Fig. 1, where one of the inputs is connected to the output of the PV array and the other input connected to the output of a generator. The fusion of the two converters is achieved by reconfiguring the two existing diodes from each converter and the shared utilization of the Cuk output inductor by the SEPIC converter. This configuration allows each converter to operate normally individually in the event that one source is unavailable. In the case when only the wind source is available, $D1$ turns off and $D2$ turns on; the proposed circuit becomes a SEPIC converter, and the input to output voltage relationship is given by (1). On the other hand, if only the PV source is available, then $D2$ turns off and $D1$ will always be on and the circuit becomes a Cuk converter. The input to output voltage relationship is given by (2). In both cases, both converters have step-up/step-down capability, which provide more design flexibility in the system if duty ratio control is utilized to perform MPPT control [9].

$$\frac{V_{\mathrm{dc}}}{V_w} = \frac{d_2}{1 - d_2} \tag{1}$$

$$\frac{V_{\mathrm{dc}}}{V_{pv}} = \frac{d_1}{1 - d_1} \tag{2}$$

Figure 2 illustrates the various switching states of the proposed converter. If the turn on duration of $M1$ is longer than $M2$, then the switching states will be state I, II, IV. Similarly, the switching states will be state I, III, IV [9]. The mathematical expression that relates the total output voltage and the two input sources is given by Eq. (3).

$$V_{\mathrm{dc}} = \frac{d_1}{1 - d_1} V_{PV} + \frac{d_2}{1 - d_2} V_W \tag{3}$$

It is observed that V_{dc} is simply the sum of the two output voltages of the Cuk and SEPIC converters. This further implies that V_{dc} can be controlled by d_1 and d_2 individually or simultaneously [2, 4].

The various switching states of the proposed converter are given in Fig. 2.

The voltage stress for the two switches is given by Eqs. (4) and (5), respectively. As for the current stress, it is observed from Fig. 3 that the peak current always occurs at the end of the on-time of the MOSFET. Both the Cuk and SEPIC MOSFET current consist of both the input current and the capacitors (C_1 or C_2) current. L_{eq1} and L_{eq2}, given by (6) and (7), represent the equivalent inductance of Cuk and SEPIC converters, respectively.

$$V_{\mathrm{ds1}} = V_{pv}\left(1 + \frac{d_1}{1 - d_1}\right) \tag{4}$$

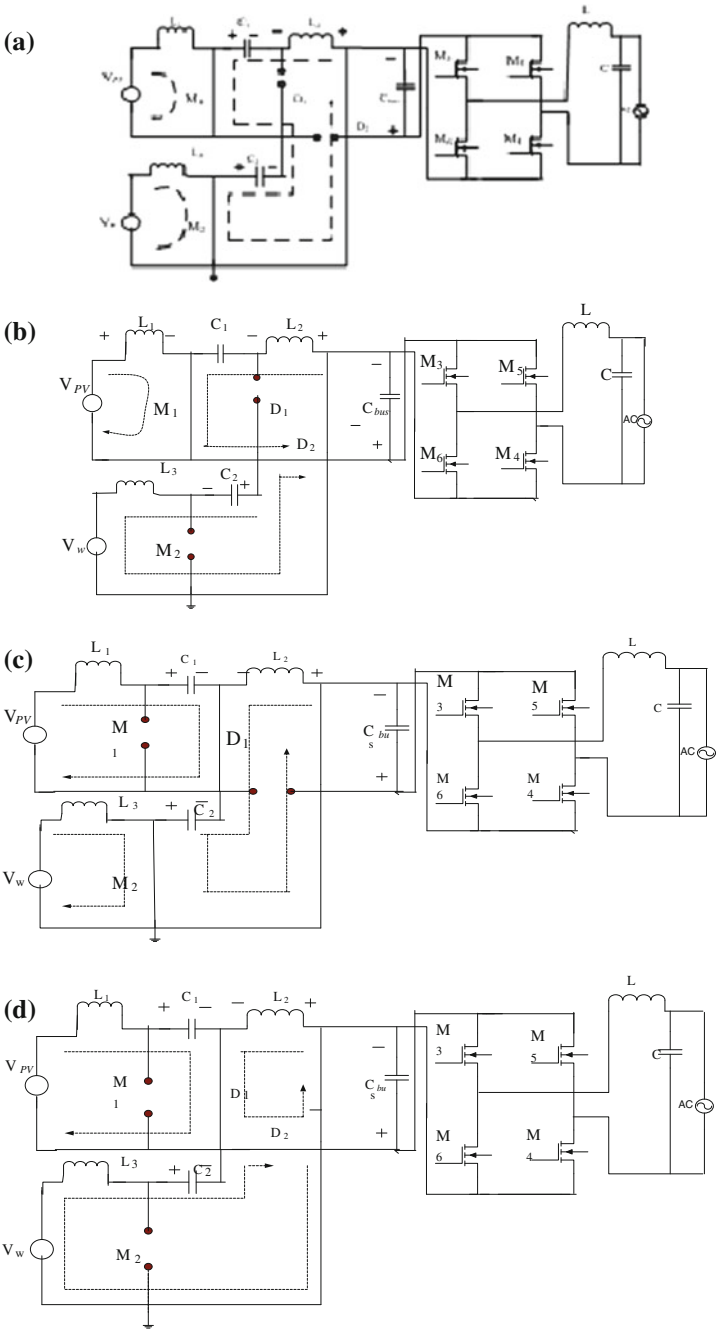


Fig. 2 **a–d** Switching states within a switching cycle

$$V_{\mathrm{ds2}} = V_{pv}\left(1 + \frac{d_2}{1 - d_2}\right) \tag{5}$$

$$L_{\mathrm{eq1}} = \frac{L_1 L_2}{L_1 + L_2} \tag{6}$$

$$L_{\mathrm{eq2}} = \frac{L_3 L_2}{L_3 + L_2} \tag{7}$$

The DC bus voltage, VDC, will be regulated by the DC/AC inverter with sinusoidal pulse width modulation (SPWM) control to achieve the input–output power-flow balance. The DC/AC inverter will inject a sinusoidal current into the ac mains. The SPWM gate signals of switches M3 through M6 for producing sinusoidal AC current is generated by the microcontroller where the amplitude of the AC current is determined by the error signal of the measured DC bus voltage V_{DC} and the reference one, $V_{DC,ref}$. If the measured DC bus voltage is less than the reference value, then the amplitude of the AC output current will be decreased in order to increase the DC bus voltage. On the contrary, if the DC bus voltage is higher than the reference one, then the amplitude of the AC output current will be decreased. On the other point of view, the DC bus voltage is regulated by the DC/AC inverter and the input–output power balance can be achieved.

3 MPPT Control of Proposed Circuit

A common inherent drawback of wind and PV systems is the intermittent nature of their energy sources. These drawbacks tend to make these renewable systems inefficient. By incorporating MPPT algorithms, the systems' power transfer efficiency can be improved significantly [8, 9]. Over the last decade, microcontrollers have made using fuzzy logic control popular for MPPT. For the reduction in tracking time, fuzzy logic control algorithm is implemented with the embedded microcontroller for improved performance compared to conventional techniques. Fuzzy logic control generally consists of three stages: fuzzification, rule base table lookup, and defuzzification. During fuzzification, numerical input variables are converted into linguistic variables based on a membership function. The inputs to a

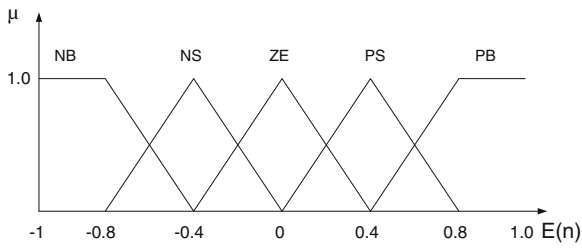


Fig. 3 Relationship between linguistics variable and error

MPPT fuzzy logic controller are usually an error E and a change of error ΔE as given in (8) and (9), respectively.

$$E(n) = \frac{P(n) - P(n-1)}{V(n) - V(n-1)} \tag{8}$$

$$\Delta E(n) = E(n) - E(n-1) \tag{9}$$

E and ΔE are calculated and converted into the linguistic variables during fuzzification. Linguistic variables are non-precise variables that often convey a surprising amount of information. To simplify the control calculation, the values of error E and change of error ΔE can be normalized with (10) before fuzzification process:

$$Y_s = \begin{cases} -1, Y < -Y^* \\ \frac{Y}{Y^*}, Y < Y^* \\ 1, Y > Y^* \end{cases} \tag{10}$$

where $Y^* = Y_{\max}$, so the scopes of error E and change of error ΔE will be $[-1, 1]$. Figure 3 shows the relations between measured error and the linguistic term, such as positive small, positive medium, and positive big. At some point, the error is positive small, and at some point the error is positive big, and the space between positive big and positive small indicates an error that is, to some degree, a bit of both. The horizontal axis in the following graph shows the measured or crisp value of error. The vertical axis describes the degree to which a linguistic variable fits with the crisp measured data.

The MPPT algorithm employed based on voltage and power feedback control approach, with the step size being dependant on the slope of the power versus voltage (P–V) curve as shown in Fig. 4. The slope of the PV array power curve is zero at the maximum power point (MPP), increasing on the left of the MPP and decreasing on the right-hand side of the MPP.

The basic equations of this method are as follows:

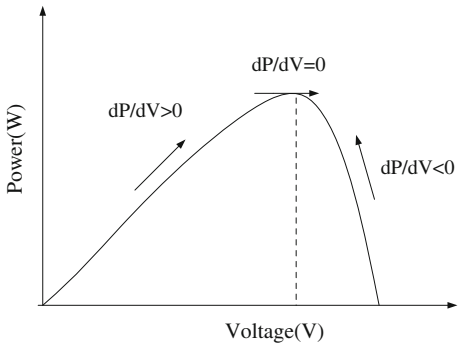


Fig. 4 *P–V* characteristic of a *PV* module

$$\frac{dP}{dV} = 0, \quad \text{At MPP} \tag{11}$$

$$\frac{dP}{dV} > 0, \quad \text{Operation point on the left of MPP} \tag{12}$$

$$\frac{dP}{dV} < 0, \quad \text{Operation point on the right of MPP} \tag{13}$$

where P and V are the PV array output power and voltage, respectively, power point. The designed controller regulates the converter output voltage by varying duty cycle of the PWM signal using MPP algorithm, and it maximizes the output power extracted from photovoltaic array. The fuzzy logic controller output is typically a change in duty ratio ΔD of the power converter. The linguistic variables assigned to ΔD for the different combinations of E and ΔE as shown in Table 1. If, for example, the operating point is far to the left of the MPP as shown in Fig. 4, that is E is PB, and ΔE is ZE, then we want to largely increase the duty ratio, that is ΔD should be PB to reach the MPP. The final step in the fuzzy logic controller is to combine the fuzzy output into a crisp systems output. The result of the defuzzification has to be a numeric value which determines the change of duty cycle of the PWM signal used to drive the MOSFET. There are various methods to calculate the crisp output of the system.

4 Simulation and Results

In this section, simulation results from MATLAB Simulink version R2011a is given to verify that the proposed multi-input rectifier stage can support individual as well as simultaneous operation. The simulation is done, for example when only solar is available, when only wind is available, and when both the energy sources are available. The rectifier DC output voltage is inverted through a full-bridge PWM inverter which uses SPWM.

Circuit parameters are given as follows:

The PV panel voltage, $V_{pv} = 60$ V
The rectified output voltage of $V_W = 280$

Table 1 Fuzzy rule base table

ΔEE	NB	NS	ZE	PS	PB
NB	ZE	ZE	NB	NB	NB
NS	ZE	ZE	NS	NS	NS
ZE	NS	ZE	ZE	ZE	PS
PS	PS	PS	PS	ZE	ZE
PB	PB	PB	PB	ZE	ZE

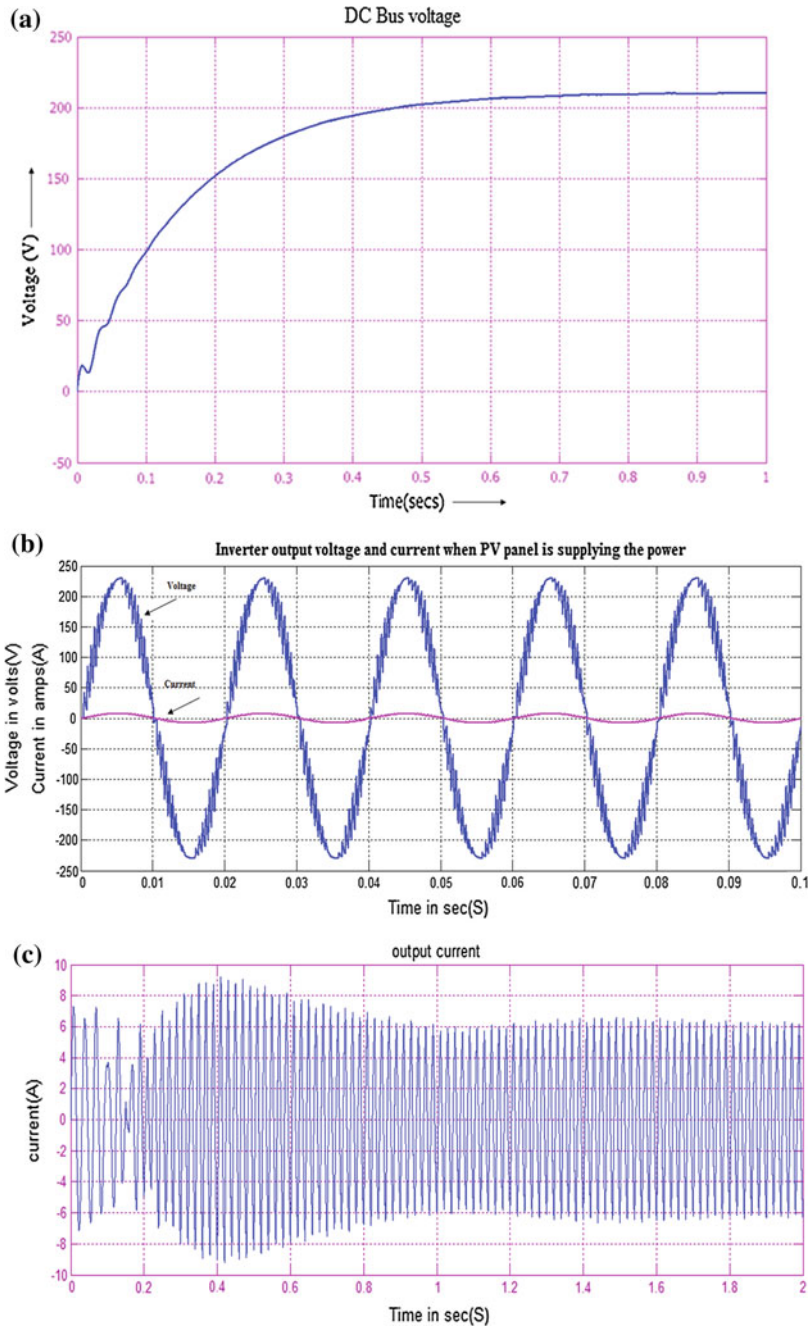


Fig. 5 Waveforms of the proposed system when only *PV* panel is supplying the power. **a** DC bus voltage waveform. **b** Inverter output voltage and current waveform. **c** Expanded AC output current waveform

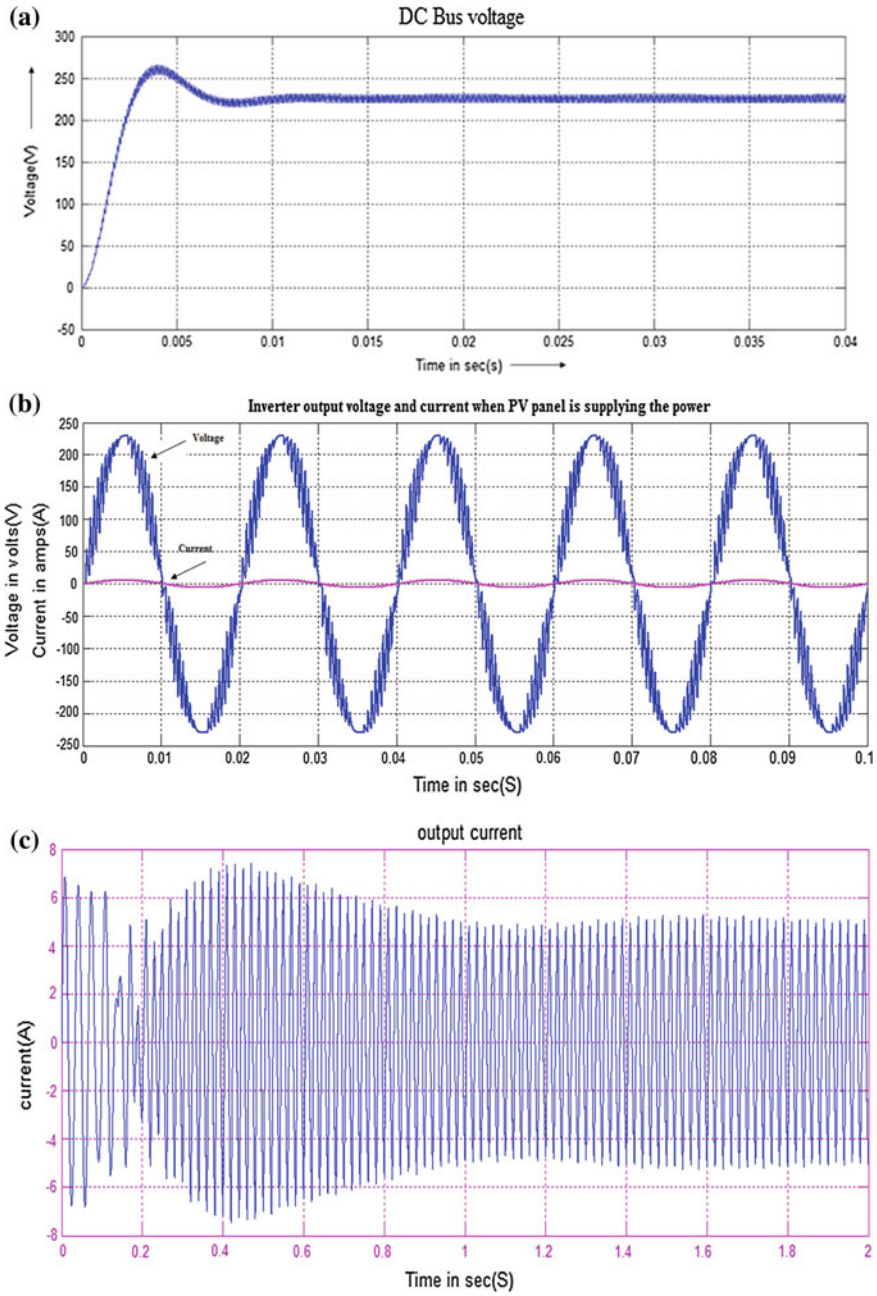


Fig. 6 Waveforms of the proposed system when only wind turbine is supplying the power. **a** DC bus voltage waveform. **b** Inverter output voltage and current waveform. **c** Expanded AC output current waveform

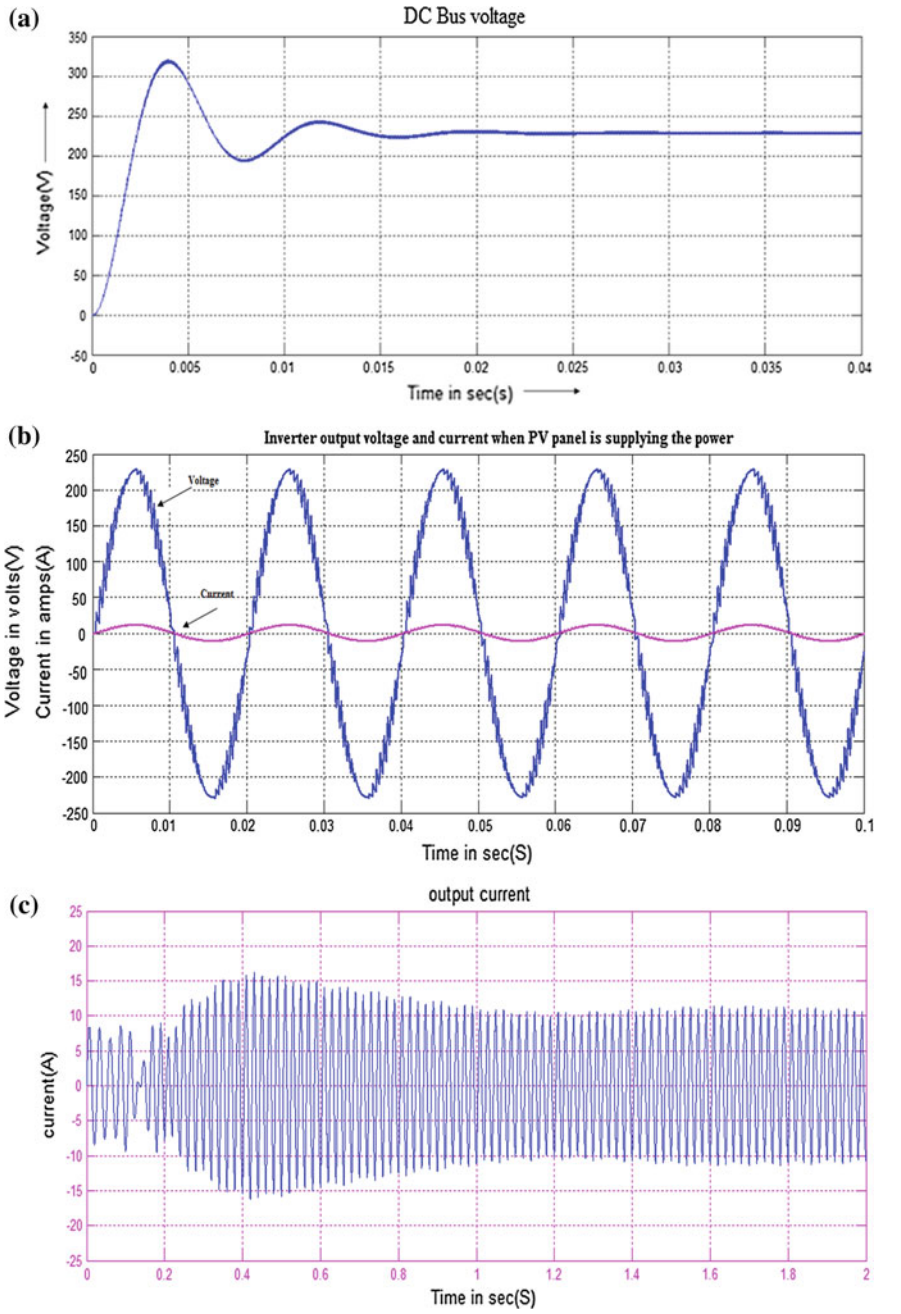


Fig..7 Waveforms of the proposed system when both *PV* panel and wind turbine is supplying the power. **a** DC bus voltage waveform. **b** Inverter output voltage and current waveform. **c** Expanded AC output current waveform

The wind generator
Switching frequency of M_1 and M_2 = 10 kHz
Switching frequency of M_3–M_6 = 5 kHz

The inductor and capacitor values of the Cuk and SEPIC converter are given below:

Inductor L_1 = 30 mH
Inductor L_2 = 6 mH
Inductor L_3 = 6 mH
Capacitor C_1 = 330 μH
Duty cycle D_{M1} = 0.79
Duty cycle D_{M2} = 0.48
Filter inductance = 10 mH
Filter capacitance = 6.3 nF

The Simulink block of the rectifier topology with MPPT consists of a PV array model, wind turbine with a PMSG, Cuk and SEPIC fused rectifier stage, the MPPT block, single-phase full-bridge PWM inverter. The gate pulse of inverter switches are controlled by SPWM.

Figure 5a shows the waveforms of the DC bus voltage when only the PV array is connected to the proposed rectifier topology. The AC output voltage and current waveforms of the proposed system are shown in Fig. 5b, and the expanded current waveform is shown in Fig. 5c. The output current obtained is 6 A.

Figure 6a shows the waveforms of the DC bus voltage when only wind turbine is connected to the proposed rectifier topology. The AC output voltage and current waveforms of the proposed system are shown in Fig. 6b, and the expanded current waveform is shown in Fig. 6c. In this case, the output current obtained is 4.7 A which is less than that obtained when *PV* array is connected to the system.

Figure 7a shows the waveforms of the DC bus voltage when both PV panel and wind turbine are connected to the proposed rectifier topology. The AC output voltage and current waveforms of the proposed system are shown in Fig. 7b, and the expanded current waveform is shown in Fig. 7c. In this case, the output current obtained is 9.4 A which is more than that obtained when either *PV* panel or wind turbine is connected to the system.

5 Conclusion

An advanced rectifier stage topology was designed and simulated using MATLAB Simulink version R2011a, and the results have been presented to demonstrate the proposed rectifier stage topology. The simulation results shows that individual or simultaneous operation is possible to give a constant output voltage, but slight variations in power. In this proposed system, additional input filters are not necessary to filter out high-frequency harmonics. Both renewable sources can be

stepped up/down (supports wide ranges of *PV* and wind input). Individual and simultaneous operation is supported. The reduced number components reduce the complexity of the system and hence the cost.

References

1. Ahmed NA, Miyatake M, Al-Othman AK (2008) Power fluctuations suppression of stand-alone hybrid generation combining solar photovoltaic/wind turbine and fuel cell systems. Proc Energy Convers Manage 49:2711–2719
2. Meza C, Martínez J, Guinjoan F (2005) Boost-Buck inverter variable structure control for grid-connected photovoltaic systems. Power Electron Spec Conf 4:1995–2000
3. Chen YM, Liu YC, Hung SC, Cheng CS (2007) Multi-input inverter for grid-connected hybrid PV/Wind power system. IEEE Trans Power Electron 22
4. Hua CC, Tsai BC (2010) Design of a wide input range DC/DC converter based on SEPIC topology for fuel cell power conversion. Int Power Electron Conf 22(2):535–542
5. Das D, Esmaili R, Xu L, Nichols D (2005) An optimal design of a grid connected Hybrid/Photovoltaic/Fuel cell system for distributed energy production. In: Proceedings of the IEEE industrial electronics conference, Nov 2005, pp 2499–2504
6. Simonetti DSL (1997) The discontinuous conduction mode Sepic and Cuk power factor preregulators: analysis and design. IEEE Trans Ind Electron 44(5)
7. Walker GR, Sernia PC (2004) Cascaded DC–DC converter connection of photovoltaic modules. IEEE Trans Power Electron 19(4)
8. Hsiao Y-T, Chen C-H (2002) Maximum power tracking for photovoltaic power system. In: Conference record of the 37th IAS annual meeting industrial application conference, pp 1035–1040
9. Hui J, Bakshai A, Jain PK (2010) A hybrid wind-solar energy system: a new rectifier stage topology. Applied power electronics conference and exposition (APEC), pp 155–161

Novel Boundary Evolving Algorithm Using Iterative Curve-Fitting Technique to Trim Down the Image Inpainting Complexity Under Extreme Conditions

A. Bindu and C. N. Ravi Kumar

Abstract Face recognition for establishment of unique identification of an individual confronts many hurdles in the form of occlusions, variations in intensity across the face, damage to the facial features during accidents, etc. It becomes a prime necessity to reconstruct the facial features to extract the correct identity. Image inpainting is an imminent and fascinating area of research predominantly used for reconstructing the obscured precincts of an image or for intelligently adding up of a new component into the image. Proficient algorithms have been put forth for the successful accomplishment of the task of image inpainting. The efficiency is well pronounced if the regions to be reconstructed are well within the boundary regions of the objects to be reconstructed. The bottleneck arises when the enclosure to be reconstructed falls on the boundary, wherein it becomes very intricate for the image inpainting algorithms to perform efficiently, mainly due to the reason that the bounds for the object are not known definitely. Such circumstances give rise to smudged image bounds which are not acceptable and also not admissible, thereby resulting in ineffectual object (face) recognition. Through this paper, a humble effort has been put forth to set the boundary and to trim down the complexities involved while inpainting under acute conditions using the affirmative features provided by nonlinear curve-fitting techniques.

Keywords Curve Fitting · Inpainting · Occlusion · Face Recognition · Regression Analysis

A. Bindu (✉) · C. N. Ravi Kumar
Sri Jayachamarajendra College of Engineering, Mysore, India
e-mail: bindukiranmys@gmail.com

C. N. Ravi Kumar
e-mail: kumarcnr123@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_65,

1 Introduction

Face forms the most easily accessible biometric among the widely accepted biometrics. Face forms the prime password for efficient access of individual identity and authentication. Accomplishment of error-free recognition with precision needs to be provided with the complete facial information, which is practically far from truth in a real-time environment. The hindrances for efficient face detection and recognition may be posed due to occlusion or due to drastic intensity variations across the face or the prime facial features may be marred due to an accident. Many algorithms have been proposed to address facial inpainting. But, the elaborate literature survey reveals that the existing techniques do not promise the efficient face inpainting when the occlusion is present across the boundary of the face in which case it becomes very complicated for successful accomplishment of reconstruction.

One of the prime tasks in pattern recognition and computer vision is fitting of geometric primitives to a set of points. With introspection into the study of the human face structure, it is evidently stipulated that the human face can be approximated by an ellipse. Through the proposed algorithm, primitive models are used to fit the perspective projection of the circle to the contour points of the partially visible face, thereby paving way for the reduction and simplification of data and, consequently, enhancing faster and simpler processing. The proposed boundary evolution algorithm is a preliminary step toward validation of the image inpainting techniques.

2 Literature Survey

Reconstruction of face through surgical methods dates back to about 800 BC—Sushruta, known to be the father of surgery, is quoted to have used reconstructive surgical techniques and made significant contributions to the field of plastic and cataract surgery *[Wikipedia]*. Research work related to the geographical boundaries description for surgical face reconstruction was initiated and proposed by Penn [1]. In the present computerized age, where automation of tasks by machines takes the lead, the research work initiated by Bertalmio et al. [2] was a step toward paving a new opening for image disocclusion using the technique similar to those used by artists to retouch the damaged paintings. This new technique came to be widely known as image inpainting. Image inpainting was oriented toward filling the depleted image region or the hole the structure elements or isophotes present surrounding the hole. Many inpainting algorithms proposed by scientists to achieve reconstruction were localized toward removal of small scratches in the image. Exemplar-based inpainting proposed by Crimnisi et al. [3] came as a fruitful solution for inpainting large depleted regions. The exemplar technique exhibited commendable efficiency in propagating both texture and structure

information. The realm of inpainting techniques was extended to 3D surfaces by Han et al. [4] and Jin et al. [5].

The literature review towards facial inpainting till date is as follows: The experimental results discussed by Hwang et al. [6] and Mo et al. [7] predominantly rely on retrieving the occluded portion of the face from the reference face image database. Hosoi T et al. [13] came up with a fast inpainting technique, which was based on the subspace method similarity. The major drawback posed by the facial inpainting approaches is that the drastic illumination variability in real-life images will not faithfully conform to the specific illumination and photographic conditions of the target image. Zhuang et al. [8] proposed a facial inpainting technique which follows the principle of occluded face recognition by classifying a query image into one or more categories in a database and extrapolate the principle for searching of faces with missing areas with the aid of prior probabilistic distribution of facial structural information. Drawback with the [8] [14] approach is that the occlusion considered is very small and such small occlusions scarcely occur in real life, and it also assumes normal and uniform distribution of luminance across the face. Literature review of the related work to date percolates to the fact that the work done till date concentrates on feature occlusions limited to either the eyes or mouth regions with a prior presumption that the data availability is limited only to the frontal face and paves way for the evident fact that the contributions toward the evolution of the contour of the occluded face region are still unexplored! In a real-life situation, the occlusion cannot be restricted to be oriented only towards the eye or the mouth regions. For example, if a person has met with an accident and has lost the right half of his face, it becomes necessary to evolve the face contour before the application of any reconstruction techniques. The proposed Boundary Evolution Algorithm is a humble step toward efficient accomplishment of occluded face contour evolution.

3 Ellipse Fitting

Extensive analysis of face images with reference to colossal collection of popular face databases could be evidently stipulated that the human face can be approximated by an ellipse.

An ellipse represented in the two-dimensional plane can be identified by 5 degrees of freedom defining its position, orientation, shape, and scale. An ellipse in general is expressed parametrically as the path of a point $(X(t), Y(t))$ which corresponds to the coordinates of the contour points of the visible face region, where

$$X(t) = Xc + a\cos t\cos\varphi - b\sin t\sin\varphi \tag{1a}$$

and

$$Y(t) = Yc + a\cos t\sin\varphi + b\sin t\cos\varphi, \tag{1b}$$

the parameter 't' varies from 0 to 2π to complete an entire loop. (Xc, Yc) is the center of the ellipse, and 'φ' forms the angle between the X-axis and the major axis of the ellipse [9]. The sphere of analytical geometry defines the ellipse by a set of points (X, Y) in the Cartesian plane which satisfies the implicit equation $AX^2 + BXY + CY^2 + DX + EY + F = 0$ provided $(B^2 - 4AC) < 0$. If $a > b$, ellipse in the form of canonical implicit equation is represented as

$$x^2/a^2 + y^2/b^2 = 1 \tag{2}$$

where 'x' and 'y' are the point coordinates in the canonical system, and (x_c, y_c) is the coordinates of the center of the ellipse. Ellipse whose center lies on a point other than the origin of the coordinate system is represented by

$$(X - X_c)^2/a^2 + (Y - Y_c)^2/b^2 = 1 \tag{3}$$

For an ellipse in canonical form, we have

$$Y = +- b\sqrt{\left(1 - (x/a)^2\right)} \tag{4}$$

Reference [9], Eq. 4 is viewed synonymously with the Eq. 5. In the proposed method, the five degrees of freedom have been considered as the set of dependent parameters. The notion of fitting of geometric primitives to a set of points [10] has been exploited in the proposed algorithm of boundary evolution. The use of primitive models allows reduction and simplification of data and, consequently, faster and simpler processing.

4 Regression Analysis

The proposed boundary evolution algorithm prediction plays a very important role in evolving the missing contour of the occluded face region. Prediction forms the prime solicitude of regression analysis, as a result of which regression analysis has been effectively and successfully exploited in the current algorithm. Regression analysis—a efficient statistical approach—is used to forecast the changes in the dependent variable reflected by the corresponding changes in one or more independent variables [11].

Regression models represented mathematically as

$$Y \approx f(X, \beta) \tag{5}$$

which involves the unknown parameters, β; the independent variables, X; and the dependent variable, Y. In various spheres of applications, different terminologies are used in place of dependent and independent variables. In the proposed algorithm, the coordinates of the visible contour of the detected face regions are considered as the independent variables. The evolved parameters of the ellipse

(degrees of freedom of ellipse) viz. foci, ellipse center, major axis length, minor axis length, and the eccentricity are considered as the dependent variables.

Least squares specify the overall minimization of the sum of the squares of the errors made in solving every single equation encountered in each and every iteration.

5 Proposed Algorithm

The proposed boundary evolution technique initially involves the detection of valid face regions by using skin illumination compensation algorithm [12] and edge detection by using Canny edge detection operator. Following that, the boundary evolution algorithm exploits the strategies stipulated by the **G**auss–**N**ewton algorithm, which is used to solve nonlinear least squares problem. The Gauss–Newton algorithm used to minimize a sum of squared function values. The proposed work of research employs the nonlinear least squares as the Gauss–Newton algorithm has the advantage that second-order derivatives, which can be challenging to compute, are not required [9]. Given m functions $r_1, \ldots, r_m$ (the different equations involving the dependent and independent variables as in Eqs. 4 and 5) of n variables $\beta = (\beta_1, \ldots, \beta_n)$, with $m \geq n$, the Gauss–Newton algorithm finds the minimum of the sum of squares.

In the boundary evolving algorithm, 'n' corresponds to the major axis length represented by 'a' and the minor axis length represented by 'b,' with reference to the Eq. 4. All the points on the boundary of the partially occluded face region are not taken into consideration, but they are sampled at periodic intervals on the boundary of the detected (to be reconstructed) face region. The sampling is carried out on all the points of the boundary corresponding to the region of interest with the intention of the computational complexity reduction.

$$S(\beta) = \sum_{i=1}^{m} r_i^2(\beta) \tag{6}$$

Starting with an initial guess $\beta^{(0)}$ for the minimum, the Gauss–Newton method proceeds by the iterations. $\beta^{s+1} = \beta^s + \Delta$, where Δ is a small step.

$$S(\beta^{(s)} + \Delta) = S(\beta^{(s)}) + \left[\frac{\partial S}{\partial \beta_i}\right]\Delta + \frac{1}{2}\Delta^T\left[\frac{\partial^2 S(\beta)}{\partial \beta_i \partial \beta_j}\right]\Delta \tag{7}$$

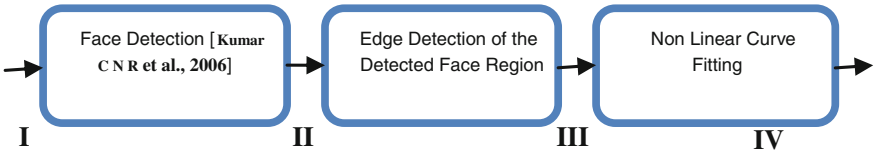
The Jacobian matrix is defined as

$$J_r(\beta) = \frac{\partial r_i}{\partial \beta_j}\bigg|_{\beta} \tag{8}$$

where $\left[\frac{\partial S}{\partial \beta_i}\right]$ in Eq. 7 could be replaced with $\mathbf{J}_\mathbf{r}^T\mathbf{r}$, and the Hessian Matrix in the right-hand side of the expression can be approximated with $\mathbf{J}_\mathbf{r}^T\mathbf{J}_\mathbf{r}$. Thus, replacing the expression assumes the following form: [9].

$$\mathrm{S}\left(\beta^{(s)}+\Delta\right) \approx S\left(\beta^{(s)}\right)+\mathbf{J}_\mathbf{r}^T\mathbf{r}\Delta+\frac{1}{2}\Delta^T\mathbf{J}_\mathbf{r}^T\mathbf{J}_\mathbf{r}\Delta.$$

In the current problem of ellipse fitting, the prime goal is to deduce the ellipse parameters viz. the major axis, the minor axis, the ellipse center, and the orientation of the ellipse, that is, the β parameters such that the model function $y = f(x, \beta)$ fits best the data points (xi, yi) retrieved and sampled from the detected edges of the region of interest (subsequently minimizing the errors or variations in fit at each iteration). The residual, which is the difference between the ideal fit and the nonlinear least squares fit, is to be found which is a good result in comparison with that of the ground truth image, in terms of the complexity involved and the efficiency with reference to the closeness between the actual boundary of the region of interest and the output of the nonlinear curve-fitting method.



Block Diagram 1. The major constituent steps adopted in accomplishing the purpose of the proposed work of research. **I** Input Color Image; **II** Detected Face after the application of Skin Illumination Compensation Algorithm [7]; **III** Edge Detected version of the Detected Face Region rendered by Canny Edge Detection technique; **IV** Expected Region of Interest with constricted valid bounds by the application of Non Linear Curve Fitting Technique.

6 Data Set

The database is compiled with due consideration given to the frontal face being occluded with different degrees. The database can be visualized into 2 broad categories: the first category consists of personally captured images and the second category consists of images collected from the World Wide Web (WWW). The first category has been built by considering 10 different variants per subject, with reference to the degree of occlusions and variations in the illumination conditions. Similar variants are collected for 100 different subjects with background variations being uniform, totaling the section 1 content to 1,000 samples. For each of the 100 subjects, one of the samples collected is an ideal one considered as the ground

Fig. 1 Occluded face boundary reconstruction

truth for the respective subject (without any occlusion). The second section in the data set involves 50 random images collected from the WWW. The frontal face contours in the web images are considered as the ground truth, and the different degrees of partial s are artificially induced (10 different degrees of occlusions).

7 Experiments and Results

The proposed algorithm for boundary evolution automatically selects and assigns the data set samples (both the first and the second category of the database considered together) randomly as training and testing categories during each and every iteration (for e.g., 25 % training and 75 % testing, 40 % training and 60 % testing and the like). The resulting evolved boundary of the partially occluded face is found to be at par with that of the corresponding ground truth when the occlusion degree is approximately equal to or less than 40%. Fig.1 depicts occluded boundary reconstruction for single face and the Fig.2 portrays the results of the proposed Algorithm for Multiple faces with varying illumination conditions.

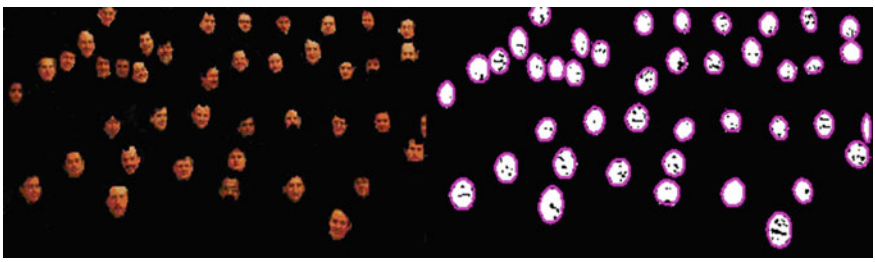


Fig. 2 Reconstructed face boundaries under varying illumination conditions

Table 1 Accuracy of the proposed algorithm in evolving occluded face contour

Sl.No	Percentage of occlusion	Success rate	
		Gradual intensity variations across the face (%)	Drastic intensity variations across the face (%)
1	0–25	>99	96–100
2	25–50	Around 95	>90
3	50–75	Around 85	Around 75

8 Conclusion and Future Work

The prime purpose of the proposed work is to set boundary limits for partially occluded facial images for successful accomplishment of the task of image inpainting, which in turn would lead to the efficient face detection and recognition. Through this boundary evolving algorithm, we have put forth a humble effort to trim down the image inpainting complexity under exorbitant conditions by setting the limits for the inpainting algorithms when the occlusions on the region of interest are toward the boundary. The boundary evolving algorithm makes use of the nonlinear regression and has resulted to be near efficient in detecting the contours of the occluded face with reference to the experimental results as depicted in (Table 1). The future enhancement would be oriented toward solving the contour evolution drawbacks present in the current boundary evolving algorithm with respect to occlusions covering beyond 50 % of the facial region.

References

1. Penn JG (1976) Geographical boundaries of facial reconstruction. S Afr Med J 50:1468
2. Bertalmio M, Sapiro G, Caselles V, Ballester C (2000) Image inpainting. In: Proceedings of computer graphics (SIGGRAPH 2000), pp 417–424
3. Criminisi A, Perez P, Toyama V (2003) Object removal by exemplar based inpainting.In: Proceedings of IEEE conference computer vision pattern recognition, vol 2. pp 721–728, Jun 2003
4. Han F, Zhu SC (2003) Bayesian reconstruction of 3d shapes and scenes from a single image. In: Proceedings of IEEE international conference computer vision, 2003
5. Jin H, Soatto S, Yezzi AJ (2003) Multi-view stereo beyond lambert. In: Proceedings of IEEE conference computer vision pattern recognition, 2003
6. Hwang BW, Lee SW (2003) Reconstruction of partially damaged face images based on a morphable face model. IEEE Trans Pattern Anal Mach Intell 25(3):365–372. doi: 10.1109 ITPAMI,1182099
7. Mo ZY, Lewis JP, Neumann U (2004) Face inpainting with local linear representations. In: Proceedings of british machine vision conference. London, U.K
8. Zhuang YT, Wang YS, Shih TK, Tang NC (2009) Patch-guided facial image inpainting by shape propagation. J Zhejiang Univ Sci A, 10(2):232–238
9. Wikipedia:http://en.wikipedia.org/wiki/Ellipse,http://en.wikipedia.org/wiki/Regression_analysis
10. http://en.wikipedia.org/wiki/Least_squares

11. http://www.businessdictionary.com/definition/regression- analysis-RA.html
12. Kumar CNR, Bindu A (2006) An efficient skin illumination compensation model for efficient face detection. In: Proceedings of IEEE industrial electronics, IECON 2006, 32nd annual conference, ISSN: 1553-572X 3444–3449
13. Hosoi T, Kobayashi K, Ito K, Aoki T (2011) Fast image inpainting using similarity of subspace method. In: Proceedings of 18th IEEE international conference on image processing
14. Bindu A, RaviKumar CN (2011) Novel bound setting algorithm for occluded region reconstruction for reducing the inpainting complexity under extreme conditions. Int J Comput Appl 16(5), (0975–8887)

ICA- and PCA-Based Face Recognition Systems—A Survey

S. Vasudha

Abstract Face recognition is one of the important applications of image processing, and it has gained significant attention in wide range of law enforcement areas. The availability of feasible technologies after two decades of research is also one of the causes to gain much importance. Although the existing automated machine recognition systems have certain level of maturity, their accomplishments are limited due to real-time challenges. For example, face recognition for the images which are acquired in high contrast with different levels of illumination is a critical problem. Various applications in defense and commercial areas demand real-time and high-level precision face recognition systems. In turn, accuracy involves many floating point operations which will be costly as well as complex in terms of implementation. The biggest challenge for present face recognition systems exists in meeting the capabilities of human perception system. This paper presents the comparative study of two subspace projection techniques within the framework of baseline face recognition system. The primary technique taken here is principal component analysis (PCA), one of the well-recognized projection techniques. The second technique is independent component analysis (ICA), a scheme that produces spatially localized and statistically independent basis vectors. The database used for comparison is FERET database. The outcome of this comparative study aids to understand that, with an appropriate distance metric, PCA performs better than ICA on an automated human face recognition task.

Keywords Principal component analysis · Independent component analysis · FERET database

S. Vasudha (✉)
Research Scholar, Department of Electrical Engineering, Visvesvaraya Technological University, Belgaum, Karnataka 590018, India
e-mail: vasudha.dsce@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_66,

1 Introduction

As one of the most sought after applications of image processing, face recognition has received significant attention, especially during the past two decades. Automatic human face recognition has received significant attention in the recent years, in the field of pattern recognition, computer vision and machine learning, computer graphics groups, owing to the high demand for advanced security and authentication needs. This common interest among various research groups working in diverse fields is motivated by remarkable ability to recognize faces and the fact that this human activity is a primary concern both in everyday life and in cyber space. There are many other commercial and forensic applications that require the use of face recognition and verification technologies such as automated crowd surveillance in airports, public events, access control in an organization, hospitals, mugshot identification in issuing driver licenses, design of human computer interface (HCI), and content-based image database management. Currently, image-based face recognition techniques can be mainly categorized into two groups based on the face representations that is used viz., appearance-based which uses holistic texture features and model-based which employs face and texture of the face along with 3D depth information.

Even though the concept of recognizing a person from facial features is perceptive, a biometric face recognition procedure makes human recognition a more automated and computerized process. Facial recognition technique is different from other biometrics because it can be used for surveillance purposes. For instance, in public places, this can be used from safety authorities to locate certain individuals such as criminals and missing children. In addition, since faces can be captured from distance, facial recognition can be done without any physical contact. This feature gives face recognition technique an added ability.

Study of fundamental theory of multivariate analysis involving a mathematical method to transform number of correlated variables into number of uncorrelated variables has helps to improve the performance of a face recognition system. The principal component analysis (PCA) ultimately reduces the number of effective variables used for classification which are compared with some statistical method and understanding of these concepts help in improving the performance of automated face recognition system.

2 Review of Recent PCA and ICA Works

PCA can be defined in general as a mathematical method that converts a set of observations of correlated variables into a set of values of linearly uncorrelated variables using an orthogonal transformation. This method is applied for feature extraction in face image projection techniques. In PCA, face image matrix needs to be transformed into a vector prior to feature extraction, but in two-dimensional

PCA (2DPCA) is based on 2D matrices and hence face image matrix transformation is not required and image covariance matrix is constructed directly using the original image matrices. In 2DPCA, the size of the face image covariance matrix is much smaller compared to PCA. Therefore, 2DPCA has two important advantages over PCA. Primary advantage is accuracy because evaluation complexity of covariance matrix is small, second advantage is time required to calculate corresponding eigenvectors which is less, and hence, speed is improved and contributed to the performance improvement of a face recognition system.

3 Problem Formulation

In the course of comparison of PCA versus ICA, deduce an appropriate distance metric, such that PCA performs better than ICA on an automated human face recognition task. The database used for comparison is FERET database which is a collection of face images in a controlled collection.

4 Face Recognition System

A general problem statement of machine recognition of faces can be formulated as 'Given still or video image of a scene, identify or verify one or more persons in the scene using a stored database of faces.'

A generic solution to the problem stated above involves segmentation of faces—face detection from cluttered scenes as a first step following with feature extraction from the face regions and recognition/verification. A generic face recognition flowchart is depicted in Fig. 1.

5 Mathematical Modeling of PCA

Assume that train image (IM) is stored in a vector of size 'N.'

$$\mathrm{IM}^j = \left[\mathrm{IM}_1^j, \mathrm{IM}_2^j, \ldots, \mathrm{IM}_N^j\right]^T. \tag{1}$$

These images are mean centered using the following formula.

$$\overline{\mathrm{IM}}^j = \mathrm{IM}^{\mathrm{j}} - \mathrm{Mean} \tag{2}$$

$$\mathrm{Mean} = \frac{1}{n}\sum\nolimits_{j=1}^{n} \mathrm{IM}^j. \tag{3}$$

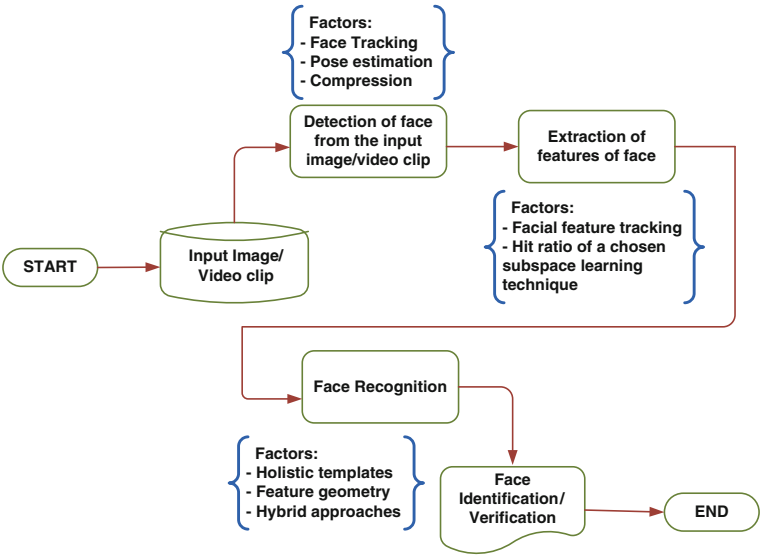


Fig. 1 Face recognition system

IM (mean centered) vectors are combined side-by-side to create a data matrix of size $N \times Q$ (where Q is the number of images).

$$\overline{Y} = \left[\overline{\mathrm{IM}}^1 | \overline{\mathrm{IM}}^2 | \ldots | \overline{\mathrm{IM}}^Q\right]. \tag{4}$$

The calculation of covariance matrix is done by multiplying data matrix 'Y' by its transpose matrix.

$$\alpha = \overline{Y} \times \overline{Y}^T. \tag{5}$$

The covariance matrix 'α' contains 'Q' eigenvectors associated with nonzero eigenvalues, assuming $Q < N$. The eigenvectors are sorted in descending order according to their respective associated eigenvalues. The eigenvector associated with the largest eigenvalue is the eigenvector that finds the greatest variance in the images until the smallest eigenvalue is associated with the eigenvector that finds the least variance in the images. The matrix of eigenvectors is the eigenspace 'E,' where each column of 'E' is an eigenvector.

$$E = [e_1 | e_2 | \ldots e_Q]. \tag{6}$$

Calculation of dot product of the image with each of the ordered eigenvectors gives the projection of training images into eigenspace.

$$\widetilde{\mathrm{IM}}^j = E^T \cdot \overline{\mathrm{IM}}^j. \tag{7}$$

Test image(TM) is also projected onto eigenspace.

$$\widetilde{\mathrm{TM}}^j = E^T \cdot \overline{\mathrm{TM}}^j. \tag{8}$$

The images can be compared using similarity measures such as city block norm, Euclidean norm, and the least similarity measure calculated from train and test projections is termed as matching images.

6 Mathematical Modeling of ICA

Consider there are 'n' random variables which are modeled as combination of 'n' source variables.

$$y_i = b_{i1} r_1 + b_{i2} r_2 + \cdots + b_{in} r_n. \tag{9}$$

where, $b_{i1} r_1$represents the mixing coefficients.

The above equation can be interpreted as $x = Wr$, where 'W' is the 'B − 1' is the de-mixing matrix and 'x' denotes the independent components.

$$x = [x_1, x_2, \ldots, x_n]^T. \tag{10}$$

Using maximum likelihood, non-gaussianity maximization, the estimation of de-mixing matrix is accomplished.

7 FERET Database

The FERET face recognition database is a set of face images collected by NIST from 1993 to 1997. Each image contains a single face. Prior to processing the faces are registered to each other, and the backgrounds are eliminated. Sample face images from FERET database are shown in Fig. 2.

8 Comparison of PCA and ICA

The systems used for comparison of ICA with PCA use a subspace projection technique to compute a set of subspace basis vectors. Successively, compress a gallery of stored images by projecting them onto the basis vectors. New images are matched to stored images by projecting them onto the basis vectors and matching their projections to the nearest projected image. Object recognition system used in this study contains three steps and operates on registered face images of training, gallery, and probe sets. The first step computes the subspace basis vectors from the training images. The second step projects the gallery images into the subspace. The third step projects a probe image into the subspace and retrieves the closest

Fig. 2 Sample face images from FERET database

galley image to the probe image as measured in subspace. Depending on the application, the training and gallery image sets be the same. However, the probe set is always dislodged from the gallery and training sets.

Face recognition is a significant biometric application, the comparison of PCA and ICA on FERET dataset gains prominent importance. Before the full FERET database was available, there were claims by other researchers that ICA performs better than or equals to PCA on face images [1–4]. This paper contradicts this claim.

The nature of the recognition task effects the evaluation. In particular, face recognition is more 'holistic' in the sense that global properties of the face (e.g., face color, width, length) are significant. Such features may be more easily captured by PCA than ICA, since ICA basis vectors are more spatially localized than their PCA counterparts.

9 Results

The comparison results of PCA and ICA on the FERET data set are given in Table 1. The algorithms are compared by measuring the probability of matching of a probe image with the gallery image in subspace with three different distance measures for PCA and cosine angle for ICA. It presents results from a larger study with 215 gallery images and a total of 340 probe images. The measures used for PCA are L1 norm (city block norm) which sums up the absolute difference in between pixels, L2 norm (Euclidean norm) which sums up squared distance

Table 1 Results of PCA and ICA simulations

	PCA		ICA		PCA		ICA		PCA		ICA		PCA		ICA
	FAR (Face Acceptance Ratio)								FRR (Face Rejection Ratio)						
FERET Dataset	L1	L2	Cosine	FERET Dataset	L1	L2	Cosine	FERET Dataset	L1	L2	Cosine	FERET Dataset	L1	L2	Cosine
fafb	79.98	71.9	77.65	fafc	23.45	15.87	18.11	fafb	20.02	28.1	22.35	fafc	76.55	84.13	81.89
Duplicate II	35.87	35.67	32.44	Duplicate II	19.67	5.89	8.09	Duplicate II	64.13	64.33	67.56	Duplicate II	80.33	94.11	91.91

between pixels, and cosine similarity measures, and ICA is implemented from cosine similarity measure. The threshold values fixed for the distance measures are in line with confidence interval $[X - 4\ \%(X1)\ \ X + 4\ \%(X1)]$, $X1 = \text{L1, L2,}$ cosine.

10 Conclusions

The result in Table 1 helps to understand that PCA performs better than ICA for every probe set when L1 norm is used. Using these outputs, the previous claims can be contradicted. This concludes that the proper choice of distance measure can maximize the performance of each technique.

References

1. Bartlett MS, Lades HM, Sejnowski TJ (1998) Independent component representations for face recognition. SPIE symposium on electronic imaging: science and technology; conference on human vision and electronic imaging III
2. Liu C, Wechsler H (1999) Comparative assessment of independent component analysis (ICA) for face recognition. International conference on audio and video based biometric person authentication
3. Moghaddam B (1999) Principal manifolds and bayesian subspaces for visual recognition. International conference on computer vision, pp 1131–1136
4. Yuen PC, Lai JH (2000) Independent component analysis of face images. IEEE workshop on biologically motivated computer vision

Relationship Between View Angle and Distance Between Parallel Edges with Vanishing Point in 2D Images

S. Murali, R. Deepu and Vikram Raju

Abstract A vanishing point is one of the possibly several points in a 2D image where lines that are parallel in the 3D source converge. Vanishing point provides a lot of information in most of the applications like distance of objects in 3D Scene. This point depends on camera viewing angle as well as distance between the parallel edges in a scene. In order to find the relationship between the vanishing points with view angle and distance between parallel edges, various data were collected by having an experimental setup by varying the distance between two parallel sticks in an image and also by varying the angle of the camera with the parallel sticks. Different data mining techniques were used, and the results were found in two phases. The results of regression model for the relationship between the angle of the camera and the X and Y coordinates of the vanishing point have been analyzed using neural network model, which gave 100 % validity on the model.

Keywords Vanishing point · Regression model · Analysis of variance · Image analysis · Angle detection · Neural networks · Multilayer perceptron

S. Murali (✉) · R. Deepu · V. Raju
Department of Computer Science and Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India
e-mail: murali@mitmysore.in

R. Deepu
e-mail: rdeepusingh@mitmysore.in

V. Raju
e-mail: vikramraju17@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_67,

1 Introduction

Vanishing points are useful in applications such as road detection in unmanned cars, object detection in an image, and training canines. In automated vehicles driving on the roads, it can be used to monitor and control the vehicle to be kept in the specified lane of the road when it deviates by tracing the vanishing point. It can also be used in embedded system applications to assist the visually impaired to direct them on a pathway. The method to find the vanishing point works with simple 8-bit gray scale image. The fundamental visual element is the trend line or "directionality" of the image data, in the neighborhood of a pixel. This directionality is measured as the magnitude of the convolution of a pixel's neighborhood with a Gabor filter. Gabor filters are directional wavelet-type filters or masks. They consist of a plane wave, in any direction in the image plane, modulated by a Gaussian around a point. The Gaussian localizes the response and guarantees that the convolution integral converges, and the plane wave affords a nonzero convolution over discontinuities that are perpendicular to the wave's direction. In other directions, the periodicity of the wave brings the convolution close to zero [1]. The phase of the wave, relative to a feature at the origin, is best accounted for by a sine/cosine combination of filters: cosine responds to even-parity features, sine responds to odd-parity features, and the other responses are vector combinations of these two. We are interested only in the magnitude of the response. The common formulas for the Gabor filter are as follows:

$$G_{\cos}(x,y) = e^{\frac{(-x'^2+\gamma^2 y'^2)}{2\sigma^2}} \cos(\omega x')$$

$$G_{\sin}(x,y) = e^{\frac{(-x'^2+\gamma^2 y'^2)}{2\sigma^2}} \sin(\omega x')$$

where the wave propagates along the x coordinate, which is rotated in some arbitrary direction relative to the image coordinates x and y; ω is the circular frequency of the wave, $\omega = 2\pi / \lambda$. Parameter σ is the half-width of the Gaussian and defines roughly the size scale on which discontinuities are detected. Parameter γ is the eccentricity of the Gaussian: if $\gamma \neq 1$, the Gaussian is elongated in the direction of the wave, or perpendicular to it. Repeated convolutions of the same mask with every pixel's neighborhood can be accelerated by convolving in the Fourier space, where the convolution of two functions is expressed as the product of their Fourier transforms.

Intuitively, one can think of the same mask being translated across the image. In Fourier space, translations become multiplications by a phase factor, and so it is possible to transform the image and the mask once, multiply the transforms, and then perform the inverse transform to obtain the value of the convolution at every pixel of the image at once. In the program used, the image is convolved with Gabor filters in a (large) number of evenly spaced directions, from zero to 180°, and the direction with the strongest response is retained as the local direction field (trend line) of the image. Naturally, this calculation is by far the most computationally

intensive part of the algorithm [2]. A well-known Fast Fourier Transform library (FFTW) is used.

Researchers have proposed various methods to determine the vanishing point of 2D images. Forstner [3] presented a method wherein the vanishing points were estimated from line segments and their rotation matrix, using spherically normalized homogenous coordinates and allowing vanishing points to be at infinity. The proposed method had minimal representation for uncertainty of homogenous coordinates of 2D points and 2D lines and rotations to avoid use of singular covariance matrices of observed line segments. A more efficient method was proposed by Antolovic et al. [2], where line segments were filtered using Gabor filters and having a convolution close to zero. The output of Gabor filters was processed to Fast Fourier Transformations. There are many methods to determine the vanishing point, but none to relate the camera angle with the vanishing point [4]. In this study, a model will be formed trying to bring out a relationship between the angle of camera and the position of the vanishing points.

2 Methodology

Numerous photographs were taken by altering the distance between the sticks and the angle between the camera and the horizontal. Here, the angle (α) gives the angle between the camera axis and the horizontal. Using the vanishing point method described by Danko Antolovic, Alex Leykin, and Steven D. Johnson, the vanishing point was identified for the various data set images. The camera angle α was varied from 60° to 120° with increments of 10°, for every alteration in the distance between the sticks from 25 to 115 cm apart with increment of 10 cm for each trial. The angle was varied only from 60° to 120°, since any angle beyond that took the two parallel sticks under study out of the focus area of the camera. The vanishing point in this case is characterized by the two parallel sticks. The resultant data were used to build a relationship between them. A schematic of the experiment is shown in Fig. 1.

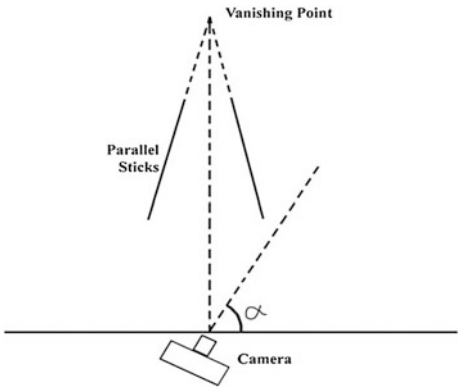


Fig. 1 Schematic of the experimental setup

3 Phase-I

3.1 General Regression Analysis

A general regression analysis is carried out for the data set for various angles (α) and the X and Y coordinates of the vanishing point determined for the various images. Table 1 shows the initial results.

From the Table 1, we get the regression coefficients for the obtained regression model. The model equation is defined by:

$$\alpha = 231.480 - 0.056X - 0.369Y$$

Here, by obtaining the X and Y coordinates of the vanishing point, the camera angle α can be determined. The corresponding P value for the constant, X and Y coordinates is 0, thus having a strong significance in the model.

4 Phase-II

4.1 Neural Networks and Multilayer Perceptron

A multilayer perceptron (MLP) is a feed forward artificial neural network model that maps a set of input data onto a set of appropriate output. MLP utilizes a supervised learning technique called back propagation for training the network [5]. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable.

Input Layer—a vector of predictor variable values $(X_1 \ldots X_p)$ is presented to the input layer. The input layer standardizes these values so that the range of each variable is -1 to 1, using the equation mentioned below. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the bias that is fed to each of the hidden layers, and this bias is multiplied by a weight and added to the sum going into the neuron. In input layer, distance will be used and a small bias is also incorporated.

Hidden Layer—arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight (W_{ji}), and the resulting weighted values are

Table 1 Regression coefficients

Term	Coefficient	SE coefficient	T	P
Constant	231.480	14.8529	15.5848	0.0
X	−0.056	0.0122	−4.5650	0.0
Y	−0.369	0.0369	−10.0217	0.0

added together producing a combined value U_j. The weighted sum (U_j) is fed into a transfer function, σ, which outputs a value H_j. The outputs from the hidden layer are distributed to the output layer. In this case, since all data are equally important, all have same weights.

Output Layer—arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight (W_{kj}), and the resulting weighted values are added together producing a combined value V_j. The weighted sum (V_j) is fed into a transfer function, σ, which outputs a value Y_k. The Y values are the outputs of the network. The X and Y coordinates of the vanishing point are determined.

Missing value handling for this experiment consists of user- and system-missing values that are treated as missing. Statistics are based on cases with valid data for all variables used by the procedure. Weight handling is not applicable in this experiment, since all coordinates and angles and distance between sticks are equally important and significant to the study.

4.2 Multilayer Perceptron: Angle α Versus X *Coordinate*

A total of 92 of 130 observations were used to build the models which were later used to predict the remaining observations (Table 2).

The input layer consists of two covariates, the distance between the sticks in the experiment and the X coordinate of the vanishing point. It thus consists of two units, excluding the bias unit in the setup. The rescaling method used for the covariates was standardized. The dependent and independent variable were standardized using

$$\frac{X - \text{Mean}}{\text{Sample Size}}$$

Only one hidden layer is used with three units in the hidden layer excluding the bias unit. The activation function used was the hyperbolic tangent function. The output layer depends only on the X coordinates with only one unit in it. A standardized rescaling method was used for the scale dependents. An identity activation function was used in this case. To incorporate the error, a sum of squares error function was incorporated (Fig. 2, Table 3).

Table 2 Case processing summary

	N	Percent
Sample		
Training	92	70.8
Testing	38	29.2
Valid	130	100.0
Excluded	0	
Total	130	

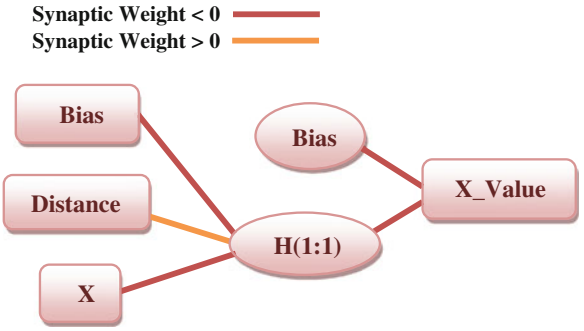


Fig. 2 Multilayer perceptron model for angle α versus X coordinate

Table 3 Model summary

Training	Sum of squares error	33.206
	Relative error	0.730
	Stopping rule used	One consecutive step(s) with no decrease in error (error computations are based on the testing sample)
Testing	Sum of squares error	8.478
	Relative error	0.510

SSE is the error function that the network tries to minimize during training. Note that the sums of squares and all following error values are computed for the rescaled values of the dependent variables. The relative error for each scale-dependent variable is the ratio of the sum of squares error for the dependent variable to the sum of squares error for the "null" model, in which the mean value of the dependent variable is used as the predicted value for each case (Tables 4, 5).

Summation of importance is 1. Hence, the importance of each of the independent variables in predicting the dependent can be understood. From the Fig. 3, we can clearly understand the importance of the X coordinates of the vanishing point for various camera angles α, and has a 100 % importance level, while the

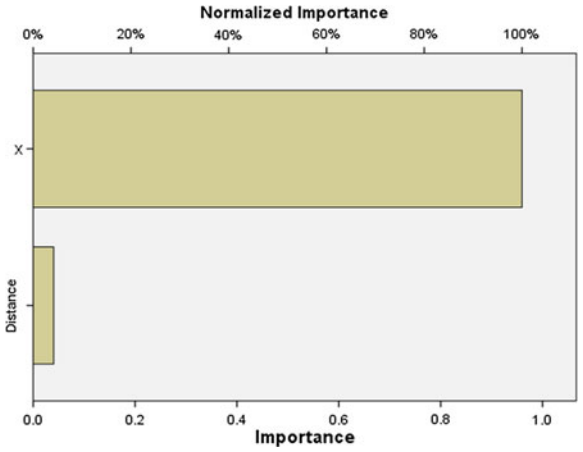
Table 4 Parameter estimates

Predictor	Predicted	
	Hidden layer 1 H(1:1)	Output layer X coordinates
Input layer		
(Bias)	−0.110	
Distance	0.038	
X	−0.821	
Hidden layer 1		
(Bias)		−0.137
H(1:1)		−0.928

Table 5 Independent variable importance

	Importance	Normalized importance (%)
Distance	0.040	4.2
X	0.960	100.0

Fig. 3 Variable importance



distance between the sticks does not account much for the model. This agrees with the first regression equation initially derived.

4.3 Multilayer Perceptron: Angle α Versus Y *Coordinate*

A total of 86 of 130 observations were used to build the model which was later used to predict 44 of the remaining observations (Table 6).

The input layer consists of two covariates: the distance between the sticks in the experiment and the *Y* coordinate of the vanishing point. It thus consists of two units, excluding the bias unit in the setup. The rescaling method used for the covariates was standardized. The dependent and independent variable were standardized using

Table 6 Case processing summary

	N	Percentage
Sample		
Training	86	66.2
Testing	44	33.8
Valid	130	100.0
Excluded	0	
Total	130	

Table 7 Model summary

Training	Sum of squares Error	22.622
	Relative error	0.532
	Stopping rule used	One consecutive step(s) with no decrease in error (error computations are based on the testing sample)
Testing	Sum of squares error	10.772
	Relative error	0.571

$$\frac{Y - \text{Mean}}{\text{Sample Size}}$$

Only one hidden layer was used with three units in the hidden layer excluding the bias unit. The activation function used was the hyperbolic tangent function. The output layer depended only on the Y coordinates with only one unit in it. A standardized rescaling method was used for the scale dependents. An identity activation function was used in this case. To incorporate the error, a sum of squares error function was incorporated (Table 7, Fig. 4).

Similar to the case of that of the X coordinates, even the Y coordinates follows the same model (Tables 8, 9).

Summation of importance is 1. Hence, the importance of each of the independent variables in predicting the dependent can be understood.

From the Fig. 5, we can clearly understand the importance of the Y coordinates of the vanishing point for various camera angles α, and has a 100 % importance level, while the distance between the sticks does not account much for the model. This agrees with the first regression equation initially derived.

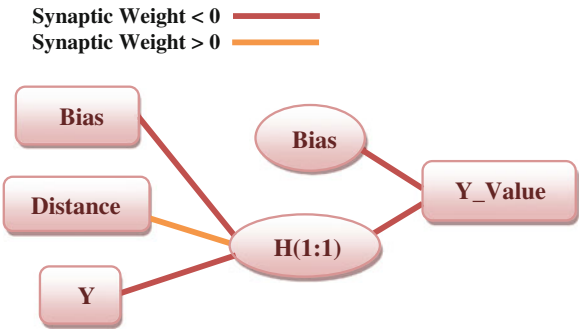


Fig. 4 Multilayer perceptron model for angle α versus Y coordinate

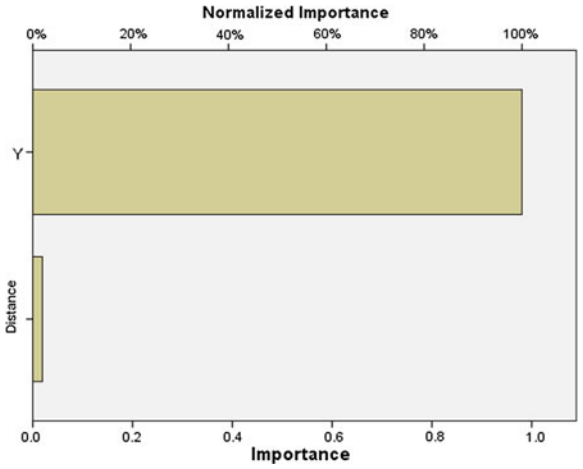
Table 8 Parameter Estimates

Predictor	Predicted	
	Hidden layer 1 H(1:1)	Output layer *Y* coordinates
Input layer		
(Bias)	0.534	
Distance	−0.019	
Y	0.896	
Hidden layer 1		
(Bias)		−0.272
H(1:1)		1.128

Table 9 Independent variable importance

	Importance	Normalized importance (%)
Distance	0.019	2.0
Y	0.981	100.0

Fig. 5 Variable importance



5 Conclusion

From the two models, the regression model and the neural network model, we can easily find a relationship between the camera angle and the vanishing point. The predicted values from the neural network model were verified by finding the vanishing point of the specified images manually, which agree with the prescribed model. The values also agree with the regression equation thus formed with an error of ±5 % collectively. Using this information and the given vanishing point, various applications can be handled like road detection in unmanned cars, object

detection in an image, as a tool to assist the visually challenged, and training canines are a few to name.

References

1. Lee T (1996) Image representation using 2D Gabor wavelets. IEEE Trans Pattern Anal Mach Intell 18(10):959–971
2. Antolovic D, Leykin A, Johnson SD (2005) Vanishing point: a visual road-detection program for a DARPA grand challenge vehicle
3. Forstner W (2010) Optimal vanishing point detection and rotation estimation of single images from a Legoland scene
4. Aguilera DG, Gomez Lahoz J, Finat Codes J (2005) A new method for vanishing points detection in 3D reconstruction from a single view
5. Nazzal M, El-Emary IM, Najim SA (2008) Multilayer perceptron neural network (MLPs) for analyzing the properties of Jordan Oil Shale Jamal

Airborne Character Recognition System with Image Processing Techniques

Nithin Chandra Bharadwaj, G. Praveen , S. Naveen Kumar and Sundari Tribhuvanam

Abstract This paper proposes a novel "air-writing" character recognition system (ABCRS) based on optical detection of red light. The user can write a character in the air with a red light-emitting device. The trajectories of the light can be captured and detected by a camera during the writing process, and then, a character reconstruction algorithm is employed to convert them into a 2-D image as a virtual character trajectory after a series of pre-processing steps including background analysis and noise reduction. Input images are pre-processed, and the image is done with a set of image processing techniques. Input database consists of upper-case English alpha numeric characters. Template-matching algorithms are developed to match input images with the standard database. The feature extraction and template matching using cross-correlation (CC) and performance index (PI) are considered to recognise the characters. Experiment results show that Airborne Character Recognition System (ABCRS) able to recognise digits and English alphabetical characters with an encouraging accuracy with an efficiency of 86 %. For the acquired database, it is found that cross-correlation method gives better results.

Keywords Airborne character recognition system (ABCRS) · Character mapping · Filtering · Thresholding · Template matching · Feature extraction · Cross-correlation · Performance index

N. C. Bharadwaj (✉) · G. Praveen · S. Naveen Kumar · S. Tribhuvanam
ECE Department, Atria Institute of Technology, Bangalore, Karnataka, India
e-mail: nithin_bharadwaj2000@yahoo.com

G. Praveen
e-mail: govindaraju.praveen@yahoo.co.in

S. Naveen Kumar
e-mail: navin6pd@gmail.com

S. Tribhuvanam
e-mail: stribhuvanam@yahoo.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_68,

1 Introduction

There are many technological breakthroughs since the invention of the digital computer, but they still receive input from traditional low-bandwidth devices like keyboard or mouse. These devices are inconvenient and unnatural for providing three- or higher-degrees-of-freedom inputs. This present work deals with human–computer interaction (HCI) to develop a machine can understand audio, visual or touch-based information [1–3].

Until a few decades ago, research in the field of optical character recognition (OCR) was limited to document images acquired with flatbed desktop scanners. The usability of such systems is limited due to the large and non-portable scanners and the need of a computing system [4]. The shot speed of a scanner is slower than that of a digital camera. Recently, with the advancement of processing speed and internal memory of hand-held mobile devices such as high-end cell phones, smart phones, iPhones, iPods having built-in digital cameras, the current research is based on running OCR applications on such devices for having real-time results [5, 6].

The processing speed and memory size of hand-held devices are not yet sufficient enough so as to run desktop-based OCR algorithms that are computationally expensive and require high amount of memory. The processing speeds of mobile devices with built-in camera vary from few MHz to GHz. The handset "Nokia 6600" with an in-built VGA camera contains an ARM9 32-bit RISC CPU having a processing speed of 104 MHz [2, 7]. The "Nokia N95 8 GB" has a Dual ARM-11 332 MHz processors and the latest HTC ONE X is armed with a quad-core NVIDIA TEGRA 3 processor clocked at 1.5 GHz with 2 GB RAM [3, 8]. Processing speed of other mobile phones and PDAs is usually in between them. The internal memory of mobile devices is less than that of desktop computers [6]. Therefore, it was an immense need to design computationally efficient and lightweight OCR algorithms for hand-held mobile devices and smart phones. However, computing under hand-held devices involves a number of challenges. Because of the non-contact nature of digital cameras attached to hand-held devices, acquired images very often suffer from skew and perspective distortion. In addition to that, manual involvement in the capturing process, uneven and insufficient illumination and unavailability of sophisticated focusing system yields poor-quality images [1, 2, 9, 10].

In this paper, we propose a novel application based on vision HCI: a virtual airborne character recognition system (ABCRS) based on optical detection of red light. The user can write characters virtually with a red light-emitting device in the air in a wireless fashion by the movement of the hand [1]. This method employs recognition of isolated and combinational handwritten characters in a noiseless environment by a digital camera. The basic principle is to identify specific terminologies in each character and extend the same to a set of characters in order to achieve accurate results with very low complexity algorithms. A character reconstruction algorithm is employed to convert them to an image on a 2-D plane

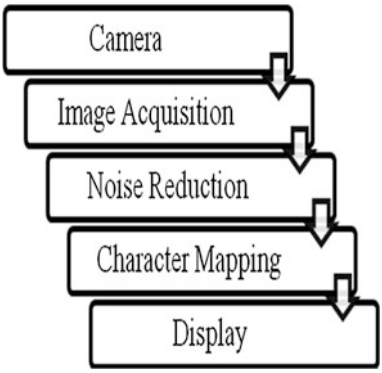


Fig. 1 Flow of ABCRS

after series of pre-processing and noise reduction [1, 2, 4]. These are recorded and stored as images which form the database for this work. Finally, using the character recognition algorithm, the virtual characters are recognised [6, 9]. This work proposes two methods for evaluating the template matching–cross-correlation and performance index (PI). This work can implemented with other languages also. The proposed system finds its application during presentations in the conferences, class rooms and lecture halls.

Figure 1 shows the flow of the proposed work. This has three main modules: image acquisition, noise reduction and character mapping. Different image processing techniques are applied for image enhancement, noise suppression, edge detection and character reconstruction and recognition. Finally, this work is evaluated with parameters—performance index and cross-correlation coefficient.

2 Method and Experimental Setup

2.1 Creation of the Database

The data acquisition was manually done in an environment with controlled illumination to reduce the effect of noise. A standard 12 MP iBall Web camera with a frame rate of 15 frames per second and 42 frames per trigger in RGB colour space with no backlight compensation is employed. These frames are added to form a single image. The flow chart of the data acquisition is shown in the Fig. 2. The database consists of 26 classes of upper-case English characters (A–Z) and 10 classes of digits (0–9) with 5 samples each. Input images have JPEG format at acquisition stage. Figure 3 shows the frames acquired in the dark room with minimum noise for digit 2 by the iBall camera. Figures 4, 5 show some of the two-dimensional images of the database by laser tracking and video capturing technique, respectively.

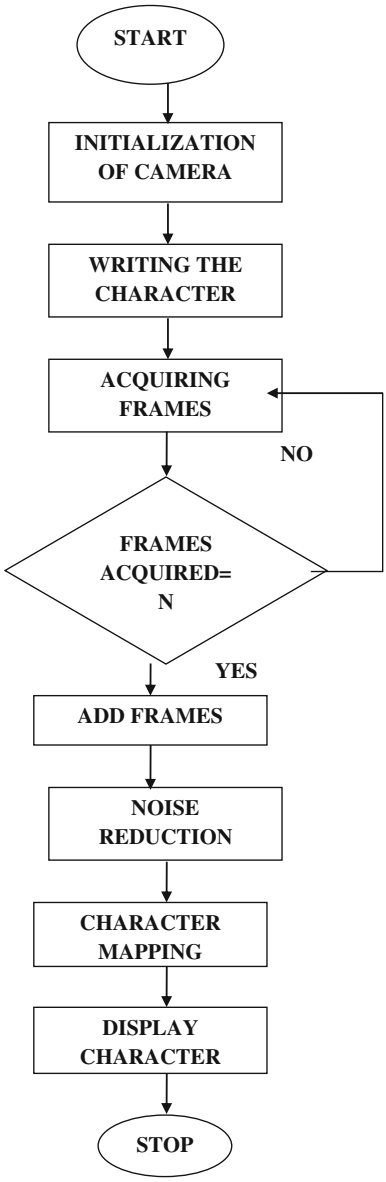


Fig. 2 Flow chart of data acquisition

2.2 Image Enhancement

Noise is added at data acquisition stage as the environment is not ideal. Images are prone to uni- and bipolar random noise and Gaussian noise [2–4]. Intensity saturation is observed during RGB to grey conversion. The images are pre-processed

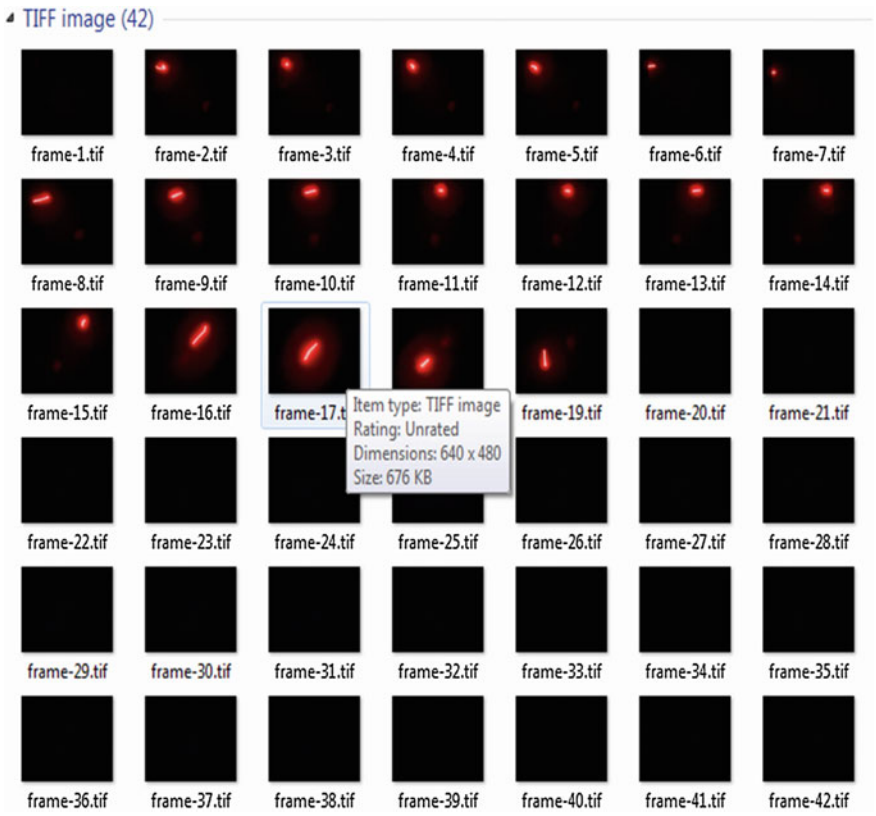


Fig. 3 Acquired frames for digit 2

by thresholding, binarisation and filtering [11, 12, 13]. The Figures 6, 7 show the algorithms for spatially filtered output image Out_Imag (i, j) and the binarisation.

Median filtering is done to reduce blurring as it is less sensitive than the mean to extreme values (called *outliers*). Median filter is able to remove these outliers without reducing the sharpness of the image. Weiner filtering is considered to reduce Gaussian noise [4, 11, 13]. The output of these processing steps with different window sizes is shown in Fig. 8. Sample database of noise-reduced output is shown in the Fig. 9. Figure 10 shows the corresponding binarised images. Images are enhanced with histogram normalisation [11, 12]. Figure 12 shows the acquired image and its histogram.

3 Character Mapping

In this work, the image of the written character may be sensed off line by moving the laser pointer on a dark background which is recorded as a video and then converted into a single image. Alternatively, the movements of the laser pointer

Fig. 4 Input image database by laser tracking

Fig. 5 Input image database by video frames

may be sensed "online", by moving the laser pointer on the computer screen surface called "laser tracking" [4]. Character image, obtained from any one of these methods, is post-processed by converting them into binary images as shown in Fig. 11.

3.1 Image Segmentation

Contour-based approach is used for segmenting the image for edge detection of the characters. Discontinuities were detected using gradient operators. Sobel masks of size 3 × 3 are considered both in x and y direction and shown in the Fig. 13. These kernels are designed to respond maximally to edges running vertically and

Fig. 6 Flow chart of the thresholding

```
for Image size=m*n, window=N
   Start= (N+1)/2
for Start< i <(m-Start+1)
    for Start< j <(n-Start+1)
            limit= (N-1)/2 , sum=0
            for –limit< k<+limit
                for –limit< l<+limit
sum=sum+Out_Imag(i+k,j+l)
                endfor
            endfor
                    if sum <N*N-1,
             Out_Imag(i,j)=0(black))
             else  (Out_Imag(i,j)=1(white)
            endif
    endfor
endfor
endfor
```

Fig. 7 Binarisation algorithm

```
begin
  for all pixels(x, y) in a TR
      if intensity(x, y) < (Gmax +G min) /2, then
                                x
          mark (x, y) as foreground  else

            if no. of foreground neighbors > 4,
then
                    mark (x, y) as foreground
                else
                    mark (x, y) as background
            end if
      end if
  end for
```

Fig. 8 Noise-reduced image with **a** 3 × 3 window, **b** 5 × 5 window and **c** 7 × 7 window

horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations [11, 12, 14]. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation x and y according to the Eq. 1.

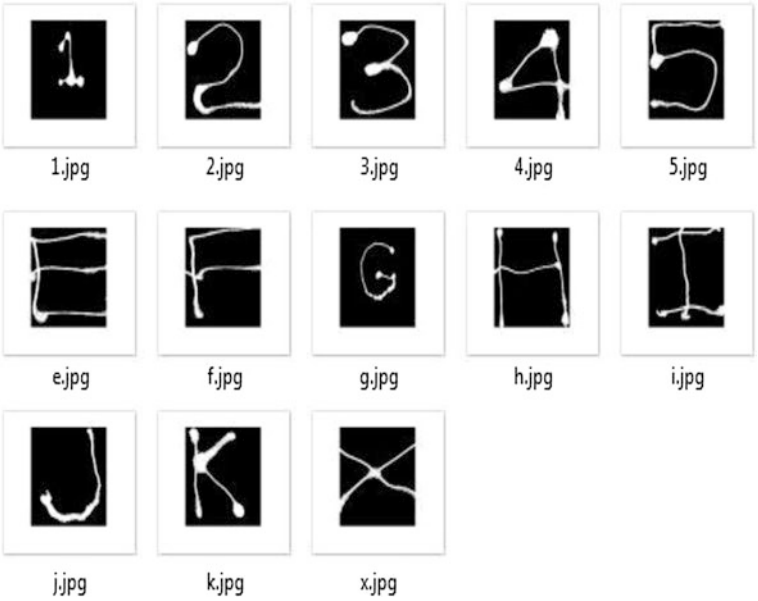


Fig. 9 Noise-reduced images

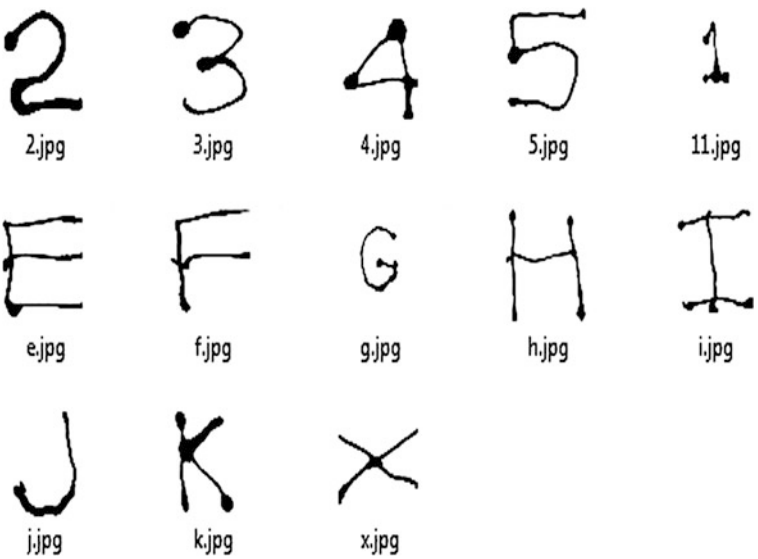


Fig. 10 Binary image data

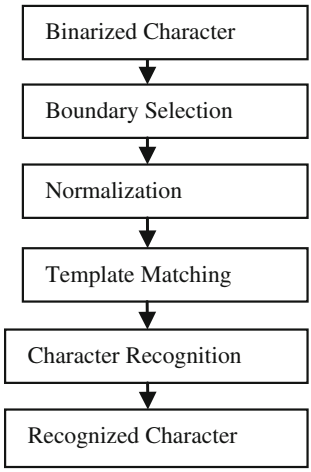


Fig. 11 Character recognition module

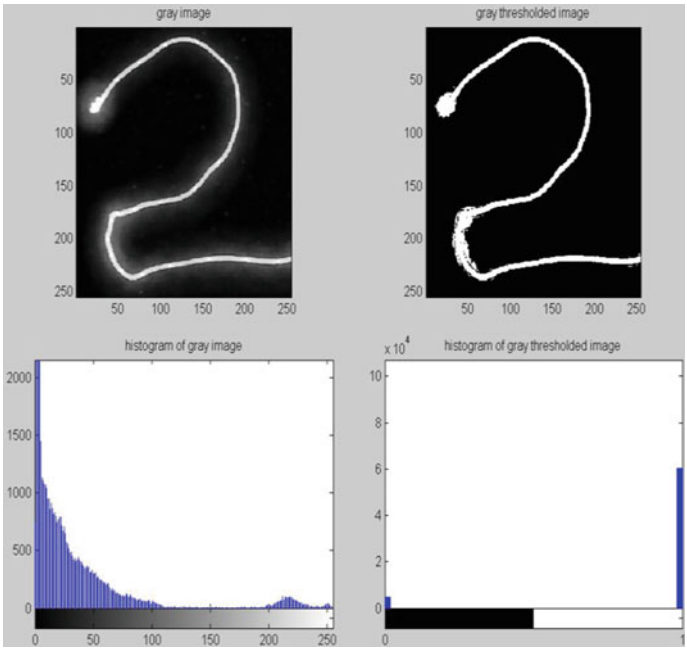


Fig. 12 Input images and its histogram

$$\begin{aligned} Gx &= (z7 + 2z8 + z9) - (z1 + 2z4 + z3) \\ Gy &= (z3 + 2z6 + z9) - (z1 + 2z4 + z7) \end{aligned} \tag{1}$$

These are combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient [15, 16]. The gradient magnitude and angle of orientation of the edge (q) (relative to the pixel grid) giving rise to the spatial gradient is given by Eqs. 2, 3[11, 15, 16]. The output of these processing is shown in Fig. 13 (Fig. 14).

$$|G| = \sqrt{(Gx^2 + Gy^2)} \tag{2}$$

$$q = \tan^{-1}(Gy/Gx) \tag{3}$$

3.2 Feature Extraction

The image is cropped and converted into binary images, and connected components are labelled and grouped together to find degree of matching with the standard character template. The normalised cross-correlation coefficient (r) computed according to the Eq. 4 [8, 11, 13, 16].

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right)\left(\sum_m \sum_n (B_{mn} - \bar{B})^2\right)}} \tag{4}$$

The performance index (ρ_i) is computed for the same input database according to the Eq. 5 where A is one of the character templates and B is handwritten character image

Fig. 13 Masks used by Sobel operator

Z1	Z2	Z3
Z4	Z5	Z6
Z7	Z8	Z9
3×3 image		
-1	-2	-1
0	0	0
1	2	1
Gx		
-1	0	+1
-2	0	2
-1	0	+1
Gy		

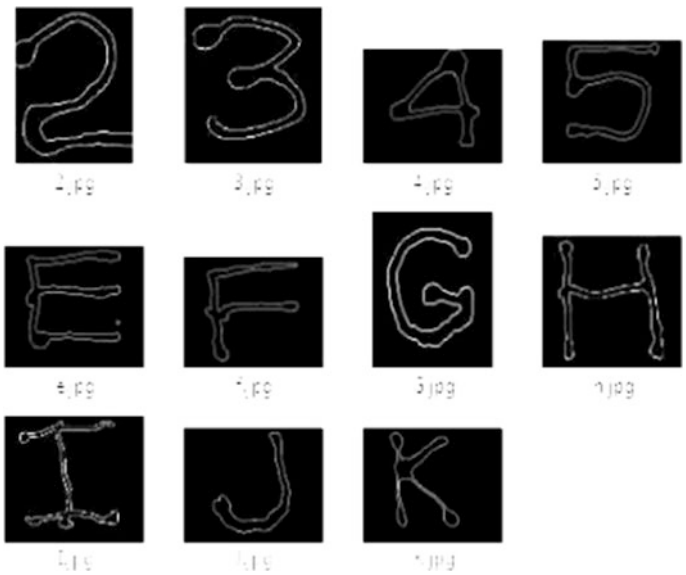


Fig. 14 Edge detection of various characters

$$\rho_i(x\,,\,y) = \sum_{x=1}^{42} \sum_{y=1}^{24} |B(x,y) - A(x,y)| \tag{5}$$

4 Results and Discussions

The database was generated in a dark room of laboratory standard. The air written–processed character is matched with the standard template, and output is the display of the character. The final result for character-2 is shown in the Fig. 15 and the output database is indicated in the Fig. 16. The Table 1 indicates the normalised cross-correlation of the input characters with the standard templates and the corresponding values of the performance index. In case of the normalised cross-correlation, positive coefficient indicates the extent of matching and is high for the similar characters. The negative coefficient is the result of mismatch. Lowest value of performance index gives the best match.

For template matcing, the size of the character acquired must not be less than 42 × 42 pixels. The processing time was 3.806446 s with a Pentium IV PC with 512 MB RAM and can be reduced to 2.4 s with a 1 GB RAM. In this work, the characters written within a time frame of 5 s are considered. Simple words of three

Fig. 15 Processed final character and the output

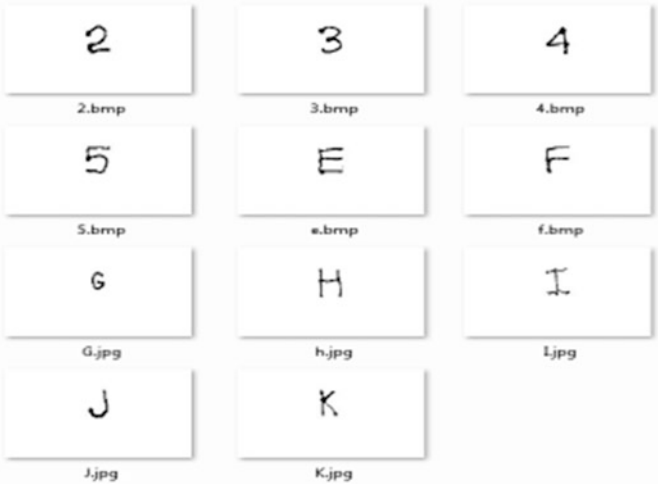


Fig. 16 The final output database

characters are recognised with this implementation with the additional segmentation algorithms. Hence, work is limited to single character identification.

The cross-correlation method does not have any restriction on window size while implementing the noise reduction algorithm. But for PI method, the window is 31×31 for better results. The result indicates that the PI value remains same for multiple matches. This leads into mismatch indicating that template matching based on cross-correlation is better than the PI method. Experimental results indicates that the efficiency of 2-dimensional cross-correlation 86 % and is better than PI method.

Table 1 Experimental results

Acquired Char → / Standard template ↓										
	CC	**PI**	**CC**	**PI**	**CC**	**PI**	**CC**	**PI**	**CC**	**PI**
	-0.2424	635	-0.0952	568	0.3635	340	0.2161	401	-0.2243	616
	0.2580	328	0.2308	323	-0.2246	555	0.0056	482	0.3963	323
	0.0041	523	0.0531	508	-0.0617	550	0.1116	455	0.5944	206
	0.1976	380	0.1312	403	-0.2885	617	0.0258	482	0.4150	301
	**0.5416**	**216**	0.2267	355	-0.3318	633	0.1294	430	0.1615	427
	0.2086	400	0.1827	415	-0.0734	537	0.2106	398	0.1869	411
	-0.0291	516	0.0663	471	-0.0727	537	-0.0378	522	**0.6968**	**153**
	0.1849	390	-0.0391	485	-0.0917	527	0.0449	474	-0.0060	509
	0.2215	434	0.0725	517	-0.1865	621	-0.0328	532	-0.0999	553
	-0.0305	547	0.0947	500	-0.0532	552	-0.3765	699	0.2114	398
	0.0531	464	-0.0155	491	0.1279	431	**0.5493**	**226**	-0.0427	527
	0.3329	377	0.1602	472	-0.3661	704	0.2872	373	0.1128	448
	0.0049	428	0.0868	369	-0.1115	493	-0.0027	482	-0.0385	525
	0.0360	440	-0.1182	491	-0.0662	497	0.2952	348	0.0618	477
	0.0361	490	0.0243	497	-0.1011	555	-0.0704	540	0.6011	201
	0.1216	433	0.2877	356	-0.0283	508	0.1635	419	0.1861	412
	-0.0829	550	0.1618	439	0.0688	475	0.0475	482	0.4543	275
	0.1252	413	0.2926	330	0.0305	464	0.4264	285	0.1422	436
	0.0588	461	0.0427	464	-0.0518	516	-0.0526	525	0.4758	266
	0.0444	524	0.0495	537	-0.0531	563	-0.0860	560	0.0051	501
	0.1243	432	0.0478	465	-0.2320	607	-0.1145	558	0.3155	347
	-0.0830	571	0.1785	462	0.0246	516	-0.1678	597	-0.1209	564
	-0.1167	544	-0.0484	505	0.1103	439	0.0320	484	-0.0148	513
	-0.1426	586	0.0792	487	0.1480	443	0.1608	428	-0.1105	559
	0.0031	555	0.0764	542	-0.0137	554	-0.1283	583	-0.0733	538
	0.1573	422	0.3642	329	-0.1884	589	-0.1970	600	0.0503	479
	-0.0888	538	-0.1690	569	0.0976	449	-0.0740	538	-0.0132	511
	0.0757	453	**0.4146**	**296**	-0.1841	580	-0.1208	559	0.1807	414
	-0.0179	510	0.1878	415	-0.1231	561	-0.2258	616	0.4270	289
	-0.3463	661	-0.0644	522	**0.4667**	**270**	0.0321	485	-0.0669	538
	0.1509	402	-0.1526	529	-0.0198	487	-0.1031	544	0.3528	331
	-0.0400	484	-0.0124	457	0.0144	465	-0.0102	496	0.5616	231
	-0.0277	545	0.2858	414	-0.0839	568	-0.2813	653	0.0402	484
	0.0644	432	0.1174	395	-0.0264	481	-0.1164	546	0.4049	309
	0.0074	463	0.0473	432	-0.1004	520	-0.0773	529	0.4606	280
	0.0012	471	-0.0004	460	-0.0026	478	-0.0774	531	0.5135	252

CC : Cross-correlation
PI : Performance Index

5 Conclusion

In this paper, we have made an attempt to take human–computer interface to a new level by virtual character recognition. This paper presents a novel airborne character recognition system based on optical detection of red light. With this system, users can input characters virtually in the air just using a simple light-emitting device, without any additional device such as keyboard, touch screen or digital pen. This system is much smaller, cheaper and more convenient. This system can be extended for real time. Results can be improved with ideal data acquisition environment.

Noise reduction algorithms yield different outputs for different inputs since each user has a unique way of writing. Successful character mapping is based on the geometry of the written characters by individual users. In this work, acquisition and mapping of the following have been accomplished:

- Straight line characters—E, F, H, I, L, T
- Slanted characters—A, K, M, N, V, W, X, Y, Z
- Curved characters—B, C, D, G, J, O, P, Q, R, S, U
- Words such as ON, NO and ONE.

Acknowledgments The authors acknowledge the management, principal and staff of Atria Institute of Technology, Bangalore, India, to carryout this work in the institute campus.

References

1. Ding K, Jin L, Yan H (2008) A virtual character recognition system based on optical detection of red light and its embedded implementation. College of electronic information South China University of technology Guangzhou Guangdong, China. 978-1-4244-2342-2. APCCAS, IEEE, pp 930–935
2. Faruk Mollah A, Majumder N, Basu S, Nasipuri M (2011) Design of an optical character recognition system for camera-based handheld devices. IJCSI Int J Comput Sci Issues 8(4)(1):1694–0814, 283–288
3. Pal A, Singh D (2010) Handwritten English character recognition using neural network. Int J Comput Sci Commun 1(2):141–144
4. Wilfong G, Sinden F, Ruedisueli L (1996) Online recognition of hand written symbols. IEEE Trans Pattern Anal Mach Intell 18(9)
5. Karmokar BC, Mahmud MAP, Siddiquee MK, Nafi KW, Kar TS (2012) Touchless written English characters recognition using neural network. Int J Comput Organ Trends 2(3):80–84 ISSN: 2249–2593
6. Optical character recognition: the association for automatic identification and data capture technologies, Pittsburgh, USA, Sept 2000, pp 1–10
7. http://www.gsmarena.com/nokia_6600-454.php
8. http://netnlap.com/web-camera/299-iball-face2face-c80-8mp-webcamera-webcam- night vision-.html
9. Xiong Y, Huo Q, Chan C (2001) A discrete stochastic model for the offline recognition of Chinese characters. IEEE Trans Pattern Anal Mach Intell 23(7):774–782

10. Jain AK, Duin RPW, Mao J (2000) Statistical pattern recognition: a review. IEEE Trans Pattern Anal Mach Intell 22(1):4–37
11. Gonzalez RC, Woods RE (2002) Digital image processing, 2nd edn, Pearson-Prentice Hall, New Jersey
12. Pratt WK (2009) Digital image processing, 4th edn, Wiley Publications, New York
13. Sonka M, Hlavac V, Boyle R (2008) Digital image processing and computer vision. Cengage Learning, India edn
14. www.imageprocessingplace.com/DIP-3E/dip3e_main_page.htm
15. Pal NR, Pal SK (1993) A review on image segmentation techniques. Pattern Recogn 26(9):1277–1294
16. Sanjeev Kunte R, Samuel RDS (2007) A simple and efficient optical character recognition system for basic symbols in printed Kannada text. *sadhanā*, vol 32. Part 5, pp 521–533

Author Biographies

Nithin Chandra Bharadwaj N is a graduate in Electronics and Communication. His domains of interest are Signal Processing and image processing, Embedded system design and VLSI.

Praveen G is a graduate in Electronics and Communication. His areas of research are Signal Processing and image processing, hardware circuit design.

Sundari Tribhuvanam is a faculty of Electronics and Communication Department, Atria Institute of Technology, Bangalore, India. Her areas of research and teaching are Biomedical Engineering, Digital Signal Processing, Digital Image Processing, Signals and Systems, Biomedical Signal and Image Analysis, Computer Vision, Medical Imaging and Computer-aided diagnosis.

An Intelligent Clustered Cooperative Transmission Protocol for Wireless Sensor Networks

K. R. Jnana Swaroop and G. Padmaja Devi

Abstract In wireless sensor networks, nodes have limited energy resources and, consequently, energy conservation is critical for extending the lifetime of wireless networks. Cooperation-based routing algorithm is implemented by finding a shortest path route first and then improving the route using cooperative communication. In cooperative networks, transmitting and receiving nodes recruit neighbouring nodes to assist in communication. The end-to-end robustness of the protocol to data-packet loss, along with the trade-off between energy consumption and error rates, is analysed. The simulation results are used to compare the energy savings and the end-to-end robustness of the protocol with two non-cooperative schemes. The comparison results show that when the nodes are placed on a grid and as compared to the disjoint-paths scheme, cooperative protocol reduces the probability of failure to deliver a packet to destination by a factor of up to 100. Similarly, compared to the one-path scheme, probability of failure to deliver a packet to destination is up to 10,000. The energy saving of cooperative protocol over disjoint-paths scheme is up to 85 % and over the one-path scheme is up to 95 % for large values of intercluster distances. The reduction in error rate and the energy savings translate into increased lifetime of cooperative sensor networks.

K. R. J. Swaroop (✉) · G. P. Devi
Department of Electronics and Communication Engineering,
Malnad college of Engineering, Hassan, India
e-mail: jnanaswaroop@gmail.com

G. P. Devi
e-mail: padmajadevig@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_69,

ACRONYMS AND ABBREVIATIONS

AODV	Ad-hoc On Demand Distance Vector Routing
ARQ	Automatic repeat request
BER	Bit error rate
CAN	Cooperation Along Non-cooperative path
CSI	Channel state information
DSDV	Destination-Sequenced Distance-Vector Routing
DSR	Dynamic Source Routing
FEC	Forward error correction
LAN	Local Area Network
MIMO	Multiple input-multiple output
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
MATLAB	MATrix LABoratory
MISO	Multiple Input Single Output
ML	Maximum likelihood
MRC	Maximal ratio combining
NAV	Network Allocation Vector
PHY	Physical Layer
QoS	Quality of Service
RERR	Route error
RERP	Route replay
RREQ	Route request
SNR	Signal-to-noise ratio
VAA	Virtual antenna arrays
WSN	Wireless Sensor Network

1 Introduction

In wireless sensor network (WSN) nodes are battery powered which are deployed to perform a specific task for a long period of time. Sensor node lifetime shows a strong dependence on the battery lifetime. In a multihop WSN, each node plays a dual role as data sender and data router. In wireless sensor networks, nodes have limited energy resources and, consequently, protocols designed for sensor networks should be energy-efficient. One recent technology that allows energy saving is clustered cooperative transmission. In clustered cooperative transmission, multiple nodes simultaneously receive, decode, and retransmit data packets. Typically, sensors are grouped into clusters, and each cluster has one cluster head. Thus, a cooperative communication model with multiple nodes on both ends of a hop and with each data packet being transmitted only once per hop is considered.

In this model of clustered cooperative transmission, the classical route from a source node to a sink node is replaced with a multihop cooperative path, and the classical point-to-point communication is replaced with many-to-many cooperative communication. Each hop on this path represents communication from many geographically close nodes, called a sending cluster, to another cluster of nodes, termed a receiving cluster. The nodes in each cluster cooperate in transmission of packets, which propagate along the path from one cluster to the next as shown in Fig. 1.

In this model of cooperative transmission, every node in the receiving cluster receives data from every node in the sending cluster. Sending nodes are synchronized, and the power level of the received signal at a receiving node is the sum of all the signal powers coming from all the sender nodes. This reduces the likelihood of a packet being received in error. It is assumed that some mechanism for error detection is incorporated into the packet format, so a node that receives an error packet will not transmit on the next hop in the path.

2 Clustered Cooperative Transmission Protocol

In wireless networks, generally, only one best route is chosen for forwarding data from a source terminal to a destination terminal. Clustered cooperative transmission protocol consists of two phases. In the routing phase, the initial path between the source and the sink nodes is discovered as an underlying one-node-thick path. Then, the path undergoes a thickening process in the recruiting-and-transmitting phase. In the recruiting-and-transmitting phase, the nodes on the initial path become cluster heads, which recruit additional adjacent nodes from their neighbourhood [1]. Due to the fact that the cluster heads recruit nodes from their

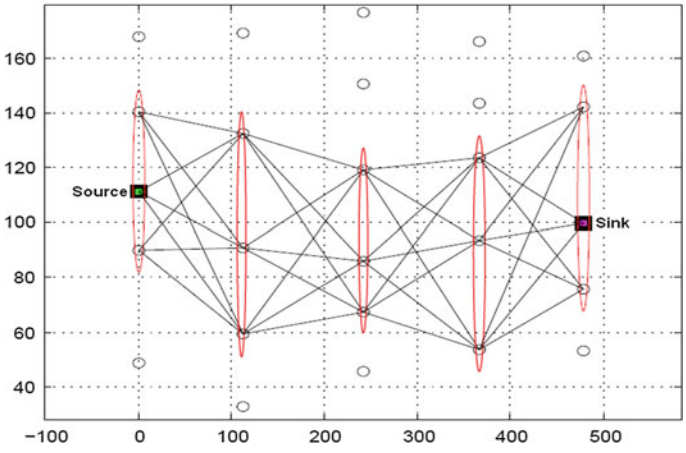


Fig. 1 Clustered cooperative transmission protocol

immediate neighbourhood, the intercluster distances are significantly larger than the distances between nodes in the same cluster.

When a packet is received by a cluster head of the receiving cluster, the cluster head initiates recruiting by the next node on the one-node-thick path. Once this recruiting is completed and the receiving cluster is established, the packet is transmitted from the sending cluster to the newly established receiving cluster [2]. The equivalent of the one-node-thick path is called as 'non-cooperative path' [3] between the source and the sink nodes and is found first.

2.1 Operation of the Recruiting-and-transmitting Phase

Figure 2a–f demonstrates the operation of the 'recruiting-and-transmitting' phase. In the current hop, node 2 is the sending cluster head and has a packet to be sent to node 5. Node 2 sends a request-to-recruit (RR) packet to node 5 (Fig. 2a), causing node 5 (cluster head) to start the formation of the receiving cluster. The CF packet contains the waiting-time-to-send and the transmission power level P_t. The transmission power level is the total transmission power (a protocol-selectable parameter) divided by the number of the nodes in the sending cluster. In this example, the value of P_t is divided by 3 (nodes 1–2–3 are cooperating in sending). After the waiting-time-to-send expires, sending cluster nodes 1–2–3 send the data packet to the receiving cluster nodes 4–5–6 (Fig. 2f). If a data packet was not received at the receiving cluster head node or was received in error, the packet is deemed lost, and the whole recruit-and-transmit phase will restart again [4].

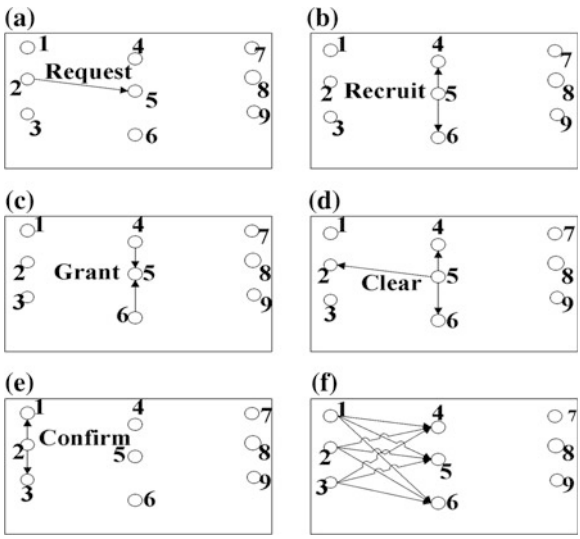


Fig. 2 Recruiting-and-transmitting phase operation: (**a**) request-to-recruit (RR) packet; (**b**) recruit (REC) packet; (**c**) grant (GR) packet; (**d**) clear (CL) packet; (**e**) confirm (CF) packet; (**f**) transmission of the data packet

A timer is associated with every exchange of control packets, so that if a critical control packet is lost, the recruit-and-transmit phase will restart again.

2.2 Calculating the Cost of Links

The cost of a link from node i to node j, $C_{i,j}$, is calculated by node i as:

$$C_{i,j} = \frac{(e_{i,j})}{R_i / R_{\text{avg}}} \tag{1}$$

where $e_{i,j}$ is the energy cost of the link, R_i is the residual battery energy of node i, and R_{avg} is the average residual battery energy of the neighbours of node i [5].

2.3 Assumptions of the Clustered Cooperative Model

Our model of cooperative communication assumes m transmitters located in the sending cluster and a single receiver located in the receiving cluster [1]. Under these circumstances, and with known signal-to-noise ratio (SNR) at the receiver of SNR, the probability of an error at the receiver is given by [6]:

$$P(\text{error}) = f(\text{SNR}, m) = \left(1 + \frac{\text{SNR}}{2}\right)^{-m} \tag{2}$$

In this model, it is assumed that the power attenuation due to distance is governed by the $d_{i,j}^{\gamma}$ law [7], where $d_{i,j}$ is the distance between node i and node j and γ is the attenuation exponent. In particular, let P_η be the noise power at the receiver and P_t be the transmitter transmission power measured at nominal distance equal to 1. When a packet is transmitted from node i to node j, the SNR measured at the receiver is computed as:

$$\text{SNR} = \frac{\left(\frac{P_t}{d_{i,j}^{\gamma}}\right)}{P_\eta} = \frac{P_t}{d_{i,j}^{(\gamma)} \cdot P_\eta} \tag{3}$$

The bit error probability is then determined by (2). It is also assumed that for a packet to be successfully received, all the bits in the packet must be successfully received [8]. Basically, it is assumed that nodes are positioned in a grid placement. It is also assumed that the horizontal and vertical distances in the grid are d and that every cluster consists of m nodes. Thus, depending on the choice of the receiving cluster, intercluster distance can be d, 2d, or 3d. To generalize, it is assumed that the intercluster distance is βd.

3 Packet Failure Probability Analysis

The failure probability that a packet does not reach the sink due to reception error(s) along the path is computed, and then the failure probability of clustered cooperative transmission protocol is compared to the failure probability using the disjoint-paths scheme and the one-path scheme.

3.1 Clustered Cooperative Transmission Protocol

Let the nodes in the cluster be indexed from 1 to m. Let p_I^J be the success probability that nodes with binary representation $I = t_m \ldots t_2 t_1$ transmit a data packet of length L bits to nodes with binary representation $J = r_{\mathrm{m}} \ldots r_2 r_1$ across a single hop, and let SNR_j be the SNR of the received signal at node j. Then

$$\mathrm{BER} = f\left(\mathrm{SNR}_j, \sum_{i=1}^{m} r_i\right) \tag{4}$$

$$p_I^J = \prod_{j=1}^{m} \left[\left(1 - r_j\right)\left(1 - (1 - \mathrm{BER})^L\right) + r_j (1 - \mathrm{BER})^L\right] \tag{5}$$

Let $p_I^J(k)$ be the probability that a packet reaches the kth hop to nodes with binary representation J, given that at least one copy reaches hop k-1, BER is the bit error rate at the receiving cluster, and then the success probability of data-packet transmission from source cluster to destination cluster after h hops is computed by:

$$p_I^J(k) = \prod_{k=1}^{h} p_I^J(k-1) \left\{ \prod_{j=1}^{m} \left[\left(1 - r_j\right)\left(1 - (1 - \mathrm{BER})^L\right) + r_j (1 - \mathrm{BER})^L\right] \right\} \tag{6}$$

It is assumed at the sending cluster that the reception pattern of nodes in a receiving cluster by a binary representation $r_m r_{m\text{-}1} \ldots r_3\ r_2\ r_1 = 1\ 1 \ldots\ 1\ 1\ 1$(all 1's). Thus, the above equation becomes

$$p_I^J(k) = \prod_{k=1}^{h} p_I^J(k-1) \cdot \left\{ \prod_{j=1}^{m} \left[r_j (1 - \mathrm{BER})^L\right] \right\} \tag{7}$$

Now, let $\mathrm{F}_{\mathrm{coop}}^k$ be the probability of failure of a packet to reach any node by the kth hop by the clustered cooperative transmission protocol, then

$$F_{\mathrm{Coop}}^k = 1 - p_I^J(k) \tag{8}$$

3.2 Disjoint-Paths Scheme

The probability of a bit error at a receiver is computed as $f\left(\frac{P_t}{P_\eta d^\gamma \beta^\gamma}, 1\right)$. Then, the probability that a packet of length L bits successfully reaches the sink over one path of h hops is $\left(1 - f\left(\frac{P_t}{P_\eta d^\gamma \beta^\gamma}, 1\right)\right)^{Lk}$. Let F^k_{Disj} be the probability of failure of a packet to reach any node by the *kth* hop of the disjoint-paths scheme, then

$$F^{\text{k}}_{\text{dis}j} = \left(1 - \left(1 - f\left(\frac{P_t}{P_\eta d^\gamma \beta^\gamma}, 1\right)\right)^{Lk}\right)^m \tag{9}$$

3.3 One-Path Scheme

The analysis in this case is similar to the disjoint-paths case, but with one path only and each node transmitting with power of $\sum_{j=1}^{m} P_t(j)$, where $P_t(j)$ is the transmission power of the jth node. Let F^k_{One} be the probability of failure of a packet to reach the kth node of the one-path scheme, then

$$F^k_{\text{One}} = 1 - \left(1 - f\left(\frac{mP_t}{P_\eta d^\gamma \beta^\gamma}, 1\right)\right)^{Lk} \tag{10}$$

It is assumed that $P_\eta = 10^{-7}$ mW, which corresponds to thermal noise across a 10 kΩ resistance in room temperature and a 1 MHz bandwidth. It is also assumed that $d = 20$ m.

4 Energy Consumption Analysis

In this section, the energy consumption of cooperative protocol is compared with two non-cooperative protocols: one-path scheme and disjoint-paths scheme for k hops. To make the comparison of energy consumption of any two schemes meaningful, the failure probability needs to be kept equal for the compared schemes. Since the probability of bit error is a function of the SNR of the received signal, the failure probability is taken as p_f.

For every value of the failure probability p_f, the needed transmission power of a single node P_t from (9) to (11) is calculated. The transmission power level is the total transmission power (a protocol-selectable parameter) divided by the number of the nodes in the sending cluster. It is assumed that the power consumption for the cooperative protocol is $m^2 \cdot P_t$, since there are m transmissions per hop, with

each transmission being of the type m-to-1. It is assumed that the power consumption is $m \cdot P_t$ for the disjoint-paths scheme and P_t for the one-path protocol.

Finally, the transmission power needed for cooperative protocol and the transmission power of the disjoint-paths and the one-path scheme to achieve the same probability of success ($1\text{-}p_f$) are plotted to indicate energy saving achieved by cooperative transmission scheme [1].

5 Simulation Results

Matlab is used as the simulation package to compare the three protocols in the multiple-hop scenario. There are two sets of experiments. In the first set, nodes are positioned on a grid to obtain simulation results. In the second set of experiments, nodes are randomly placed for a more realistic scenario.

In the first set of experiments, there are 10×25 nodes placed in a grid format of 10 rows and 25 columns. The distance between a node and its closest neighbour in the grid is $d = 20$ m. The sink's column varies depending on the parameter β. It is assumed that the channel bandwidth is 1 Mb/s, the length of the data packets is 1,024 bits, γ is 3, and the maximum waiting time T is 1.5 ms.

In the second set of experiments, nodes are randomly placed for a more realistic scenario, in which 250 nodes are randomly placed in a 900×400 m^2 area and the transmission range per node is 150 m. The source and the sink nodes are picked randomly. The metrics that are used for evaluation are the failure probability, the energy consumption, and throughput upper-bound capacity.

5.1 Packet Failure Probability Analysis

For different values of the following parameters, the number of hops in a route k, the transmission power measured at nominal distance P_t, the number of cooperative nodes m, and intercluster distance β (Fig. 3, 4, 5) are plotted for the failure probability that a packet does not reach the sink due to reception error(s) along the path. It is assumed that $P_\eta = 10^{-7}$ mW, which corresponds to thermal noise across a 10 kΩ resistance in room temperature and a 1 MHz bandwidth. It is also assumed that $d = 20$ m. For equations from (8) to (10), the failure probability of clustered cooperative transmission protocol ('Coop') is plotted in solid lines, the disjoint-paths ('Disj') failure probability is plotted in dashed lines, and the one-path ('One') failure probability is plotted in dashed-dotted lines.

In Fig. 3, the failure probability versus the packet transmission power P_t is plotted, for $k = 5$ and for $m = 3$, $\beta = 3$ and $\gamma = 3$. The failure probability of clustered cooperative transmission protocol is lower compared to the three other protocols for a given k. The failure probability of clustered cooperative

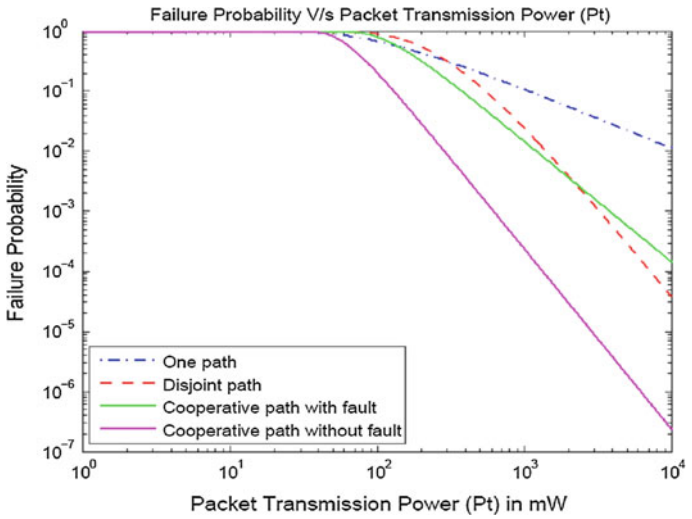


Fig. 3 Failure probability versus packet transmission power P_t ($k = 5$, $m = 3$, and $\gamma = 3$)

transmission protocol is lower compared to the other protocols by a factor of up to 100 compared to the disjoint-paths scheme, and by a factor of up to 10,000 compared to the one-path scheme, depending on the values of considered parameters.

In Fig. 4 the failure probability versus β for a fixed transmission power level is plotted, where $k = 5$, $m = 3$, and $\gamma = 3$. The failure probability increases as β increases since, as β increases, the received power decreases and the probability

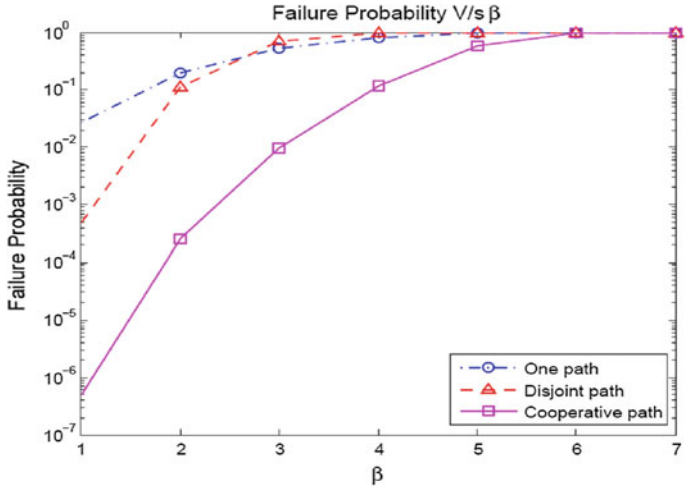


Fig. 4 Failure probability versus β ($h = 5$, $m = 3$, and $\gamma = 3$)

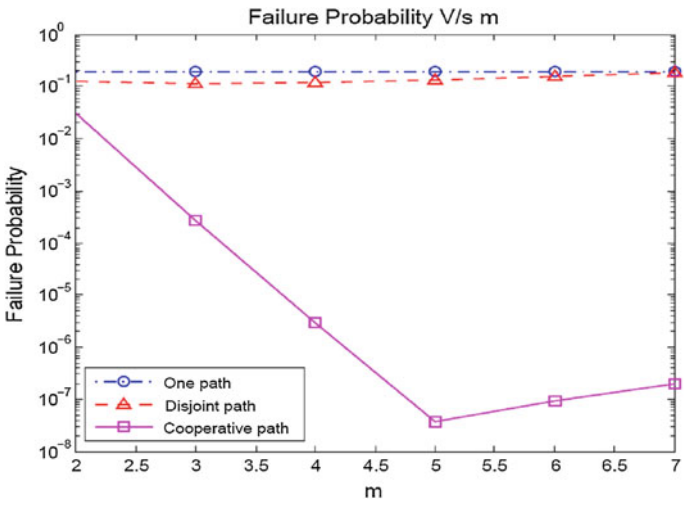


Fig. 5 Failure probability versus m ($mP_t = 1$ W, $\gamma = 3$, and $\beta = 3$)

failure increases. The transmission power level of a node is the total transmission power of a cluster (a protocol-selectable parameter) divided by the number of the nodes in the sending cluster.

In Fig. 5 the failure probability versus m is plotted for k = 5, where the total transmission power $mP_t = 1$ W is fixed. The failure probability of cooperative protocol decreases with increase in the number of nodes until $m = 5$, and then the failure probability increases. Thus, there is an optimal value of m, which minimizes the failure probability. When m increases, there are more receiving nodes, and the probability of receiving more copies of the packet increases.

On the other hand, increasing m decreases the transmission power per node P_t as the total transmission power per cluster is fixed. This results in smaller received power level at each receiver node and hence smaller probability of correctly receiving the packet.

5.2 Energy Consumption Analysis

To compare the energy consumption of cooperative protocol to the one-path and the disjoint-paths scheme for k hops, the failure probability (p_f) needs to be kept equal for the compared schemes. The energy consumption measures the sum of the energy used for all control and data-packet transmissions.

In Fig. 6, the packet transmission power needed by cooperative protocol, the one-path scheme, and the disjoint-paths scheme for $k = 5$ hops is plotted for $p_f = 0.1$ and then β is varied for different values of m.

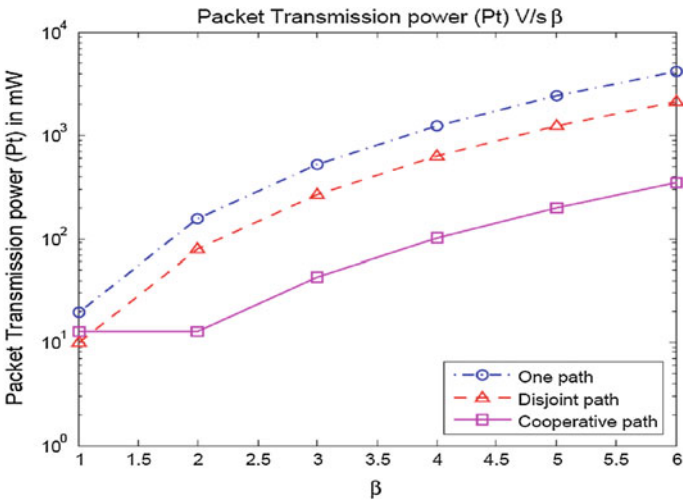


Fig. 6 Packet transmission power versus β ($k = 5$, $m = 4$, and $\gamma = 3$)

Clustered cooperative transmission protocol has larger energy consumption than the disjoint-paths scheme with $\beta = 1$. The reason for that is the high overhead of control packets when the number of cooperative nodes is large and β is small. These extra control packet transmissions affect the performance for small values of β. However, for large values of β, the cooperation among a large number of nodes improves the performance, and the transmission can be successful over larger distances between the sending and the receiving clusters. Similarly, when m is small, regardless of β, there will be no energy saving for clustered cooperative transmission protocol over the disjoint-paths scheme.

This clustered cooperative transmission protocol can save up to 85 % of energy over the disjoint-paths scheme and up to 95 % of energy over the one-path scheme for large values of β. The amount of savings increases as the failure probability decreases and as β increases.

6 Conclusion

Performance of clustered cooperative transmission protocol is evaluated, where nodes in a sending cluster are synchronized to communicate a packet to nodes in a receiving cluster. When the nodes are placed on a grid and as compared to the disjoint-paths scheme, cooperative protocol reduces the probability of failure to deliver a packet to destination by a factor of up to 100, depending on the values of considered parameters. Similarly, compared to the one-path scheme, probability of failure to deliver a packet to destination is up to 10,000.

In cooperative protocol, when nodes are positioned on a grid, the energy saving of cooperative protocol over disjoint-paths scheme is up to 85 % and over the one-path scheme is up to 95 % for large values of intercluster distances. Overall, the study demonstrates that the energy savings of this proposed protocol, relative to the other schemes, do not substantially decrease even when the data-packet loss is observed.

References

1. Elhawary M, Haas ZJ (2011) Energy-efficient protocol for cooperative networks. IEEE/ACM Trans Networking 19(2)
2. Laneman J, Tse D, Wornell G (2004) Cooperative diversity in wireless networks: Efficient protocols and outage behavior. IEEE Trans Inf Theory 50(12):3062–3080
3. Khandani A, Abounadi J, Modiano E, Zheng L (2007) Cooperative routing in static wireless networks. IEEE Trans Commun 55(11):2185–2192
4. Stefanov A, Erkip E (2002) Cooperative information transmission in wireless networks. In: Proceedings of Asian–Eur. Workshop Inf. Theory, Breisach, Germany, pp 90–93
5. Hen I (2006) MIMO architecture for wireless communication. Intel Technol J 10(2):157–165
6. Yuan Y, Chen M, Kwon T (2006) A novel cluster-based cooperative MIMO scheme for multi-hop wireless sensor networks. EURASIP J Wireless Commun Netw 2006(2):38
7. An Introduction to Matlab, Guide 7. Durham University homepage on Resources—ITS (2009) [Online]. Available: http://www.dur.ac.uk/resources/its/info/guides/7Matlab.pdf
8. Perkins C, Royer E (1999) Ad-hoc on-demand distance vector routing, pp. 90–100. In Proceedings of IEEE WMCSA, New Orleans, LA, Feb. 1999

An Embedded Web Server for Industrial Automation

Sahana K. Adyanthaya and P. A. Vijaya

Abstract In the present IT era, there is a great need of fully automated industrial system. With the growing popularity of Internet and the development of embedded technology, Web technology has been extended to the development and application of embedded system. In the future pervasive computing environment, people would be surrounded by a great multitude of microcomputers that would be embedded in home appliances, factory devices, and so on. It is a better choice for these devices to be embedded a Web server. An embedded Web server (EWS) is a microcontroller that contains an Internet software suite as well as application code for monitoring and controlling systems. The main aim of the project discussed in this paper is to use the personal computer (PC) with the Ethernet card to send the messages to remote locations which is received by the 32-bit ARM processor in order to control the devices in the real-time environment. By using this project, it is possible to change the status (ON/OFF) of any appliance connected to ARM processor board via Internet. The messages from the PC can also be sent to any remote locations in the world using this technology. The codes for the implementation of this project have been written in embedded C language as it is easy to understand and is suitable for the embedded system design. This project has been designed for monitoring and controlling the status of two devices. The client can monitor and control the status (ON/OFF) of the devices via Internet from any corner of the world.

Keywords Embedded Web server · ARM · Web · TCP/IP protocol · Ethernet · Embedded system

S. K. Adyanthaya (✉)
Department of E & C Engineering, Alva's Institute of Engineering & Technology, Moodbidri, Karnataka, India
e-mail: sahana.adyanthaya@gmail.com

P. A. Vijaya
Department of E & C Engineering, Malnad College of Engineering, Hassan, Karnataka, India
e-mail: pavmkv@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_70,

1 Introduction

The embedded systems are those which are not programmed directly by the user and they are used in various fields from power plant to control systems for a car and home appliances. Embedded systems usually are specialized hardware–software systems, with built-in operational environment, closed architecture, specific interfaces, and complex organization of an external and internal data flow. With the growing popularity of the World Wide Web (WWW), it is obvious that its underlying technologies are useful not only for browsing the Web but for much more than that [1]. Important application area for embedded systems includes control systems for complex industrial equipments. The basic function of an industrial embedded system is to sense and accumulate data about the state of a device and to control its operation. Embedded system should also provide a device's state information to operators and receive instructions from the operators. In recent times, there is a growing need for accessing remote devices by the user in his/her absence and an embedded Web server (EWS) can serve this purpose. The EWS is a microcontroller that contains an Internet software suite as well as application code for monitoring and controlling systems. The EWS developed in this project allows the embedded devices to be connected to the Internet and also enables the users to access, control, and manage the devices using a standard browser over the Internet without any restriction of time and space. This EWS could be beneficial for industrial automation, home automation, remote data acquisition systems, hospitals, and more.

2 System Design

The main purpose of the EWS is to allow the client to monitor and control the status (ON/OFF) of the devices interfaced to the 16/32-bit ARM processor via Internet. The hardware platform of the system uses 32-bit ARM processor. The software is written in embedded C language. In recent times, the core of the embedded system is either microprocessor unit or microcontroller unit to control and monitor the devices. In view of limitations of standalone, embedded systems for controlling and operating from the place of its installation, interconnection of same with TCP/IP networks can provide a broad interactive use for monitoring and controlling from the different client installations. From the very first usually RS232, RS485 or CAN was adopted, but of late, these manners are too limited in distance as more and more devices are not only gaining embedded systems but are also are being attached to a network [2]. The system discussed in this paper makes use of the TCP/IP (Ethernet) module with the known IP address and is programmed using the ARM board, and it is connected to the Ethernet card of the PC in the real-time environment such that the user can sit anywhere in the world to

access this hardware interfaced to the ARM board. The embedded devices are connected to the EWS by means of a serial interface. The EWS is connected to the router/modem by means of an Ethernet cable. The EWS has its own IP address. Upon the client request, the EWS collects the device's state information and sends them to the client who may be present locally or at a remote location. Also the EWS receives the commands from the users and sends them to the devices. The overview of the system design is shown in Fig. 1.

3 The Hardware Design of the System

The complete structure of the EWS for industrial automation is as shown in Fig. 2. The EWS performs the serial to Ethernet conversion and vice versa. The 32 bit ARM processor is programmed to receive the requests from a local/remote client to either send the status (ON/OFF) of the devices interfaced to the processor or to change the status of the devices interfaced to it. The messages sent from a remote client will be displayed on the LCD interfaced to the processor. The Ethernet controller directly supports the Ethernet service and RJ45 communication.

4 The Software Design of the System

The system software of the embedded controller has to monitor the status of the device interfaced to it. Also the Web server should allow the client to monitor and change the status of the devices interfaced to the controller from any remote location via Internet. Basically, the function of the controlling software is to

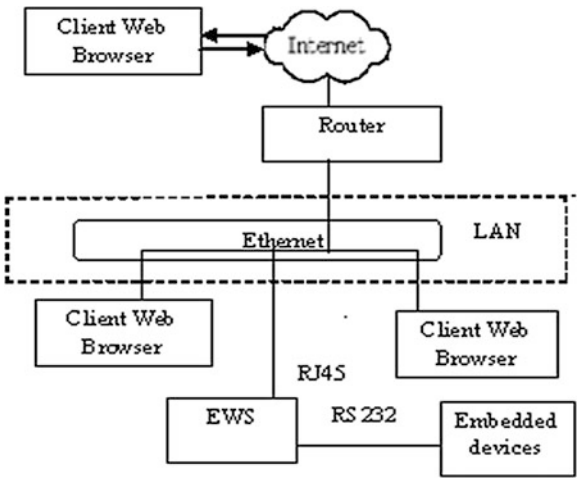


Fig. 1 Overview of the system design

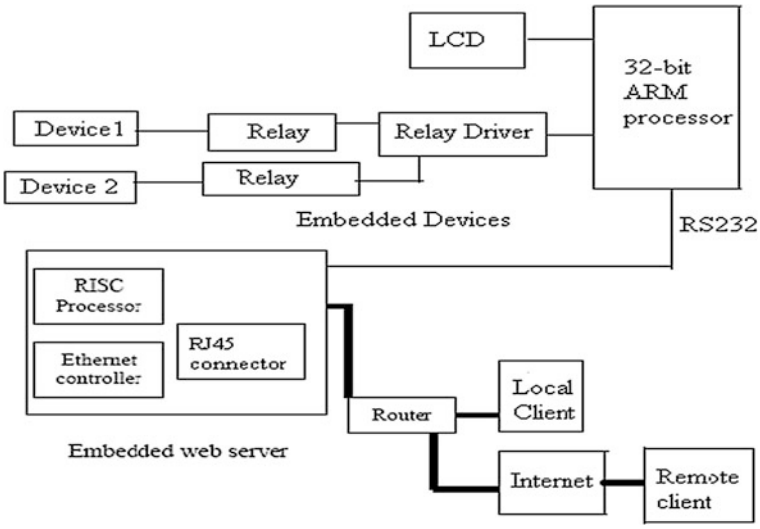


Fig. 2 Complete structure of the EWS for industrial automation

coordinate the execution of various functions with respect to the requests made by the user after gaining access to the system. While the user is still connected, the software reads and interprets the different requests to generate corresponding control signals for the correct operation of the entire system. Embedded C language is used to develop the system software.

4.1 Embedded C language

The embedded C programming language is a general purpose programming language that provides code efficiency, elements of structured programming, and a rich set of operators. Its generality combined with its absence of restrictions makes embedded C a convenient and effective programming solution for a wide variety of software tasks. Many applications can be solved more easily and efficiently with embedded C than with more specialized languages.

4.2 The Lightweight TCP/IP Stack

The Lightweight IP (lwIP) stack is an open-source implementation of the TCP/IP stack developed specifically to reduce resource usage while maintaining a full-scale TCP/IP stack for embedded systems. The lwIP makes it possible to connect

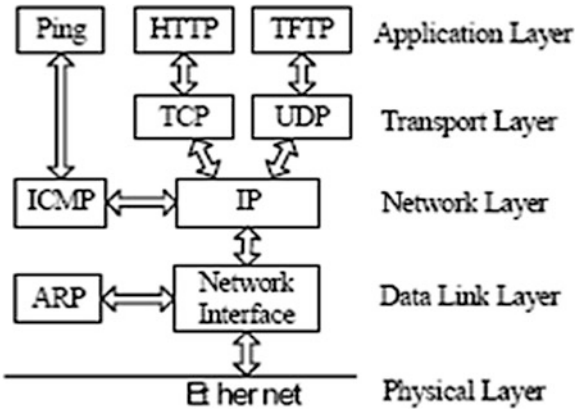


Fig. 3 Lightweight TCP/IP protocol stack

the system to a local intranet or the Internet. The lwIP stack is more suitable for the embedded systems because of small data and code size requirements. lwIP can be used with OS or without OS. The core of lwIP consists of the actual implementations of the IP, ICMP, UDP, and TCP protocols, as well as the support functions such as buffer and memory management [3, 4]. This lightweight TCP/IP protocol stack provides the function of the EWS and enables the embedded devices to have the Internet interface by the way of embedding the EWS in the devices. The basic TCP/IP protocols necessary to implement the EWS are shown in Fig. 3.

4.3 Flowchart for the EWS

The flowchart for the EWS for industrial automation is as shown in Fig. 4.

5 Experimental Setup and Results

The experimental setup of the EWS is shown in Fig. 5. The target board is connected to the Ethernet controller of the network. One end of the power cable is connected to the power socket of the ARM board and the other end is connected to the USB socket of the PC. The Ethernet module is powered on by providing 5 V supply. The relay section is powered on by providing 12 V supply.

The EWS is first tested by interfacing it with the PC using RS232 cable. To test the EWS, hyperterminal is used. The connection is established via the suitable communication port as shown in the Fig. 6. Figure 7 shows the connection settings of the hyperterminal. The settings of 9,600 bits per second with 8 data bits, no parity, 1 stop bit, and no flow control are selected (Fig. 8).

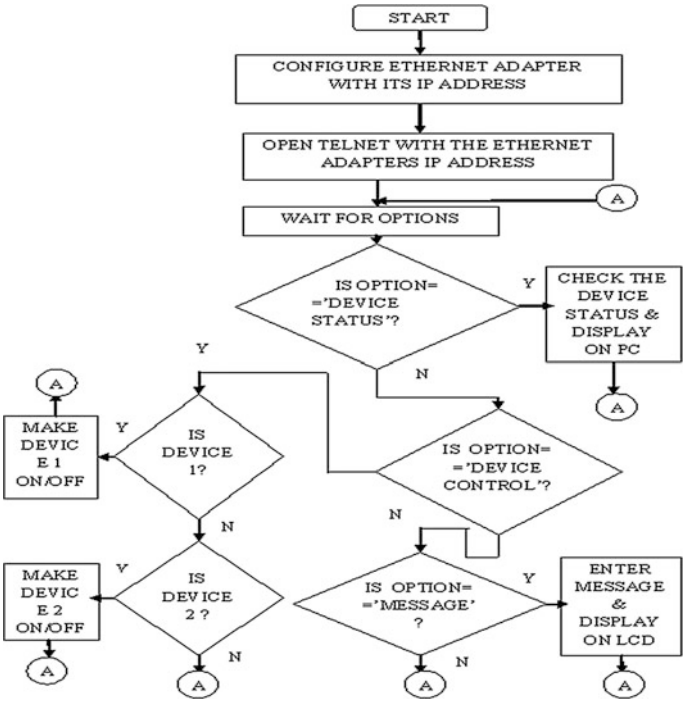


Fig. 4 Flowchart for the EWS for industrial automation

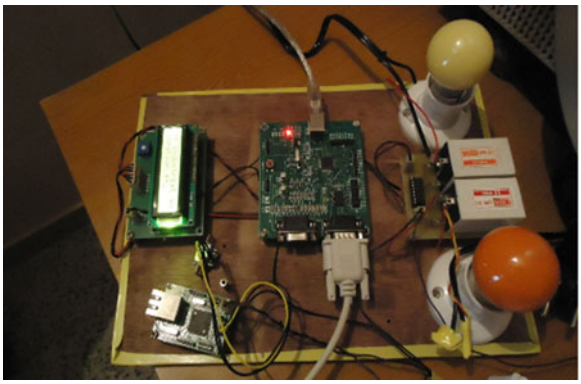


Fig. 5 Snapshot showing the experimental setup of the EWS

After testing the EWS using RS232 connection, it is then connected over the Internet. To do this, the null modem cable from the serial port on the embedded device board is connected to the serial port on the EWS and the RJ-45 socket on the EWS is connected to the router/modem using Ethernet cable. Team viewer

Fig. 6 Snapshot showing the Com port selection

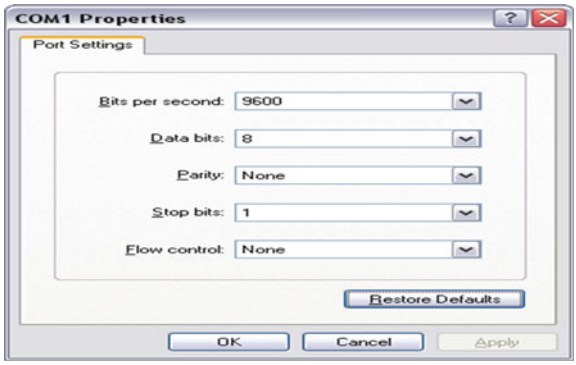


Fig. 7 Snapshot showing the connection settings of the hyperterminal

software should be installed in both the client and the server PCs. The protocol to be selected in the hyperterminal window is TCP/IP, and the IP address of the EWS is 192.168.1.11. Figure 9 shows the snapshot of the setting up of the TCP/IP connection and entering the IP address. Snapshot of the device being controlled by a remote client is shown in Fig. 10.

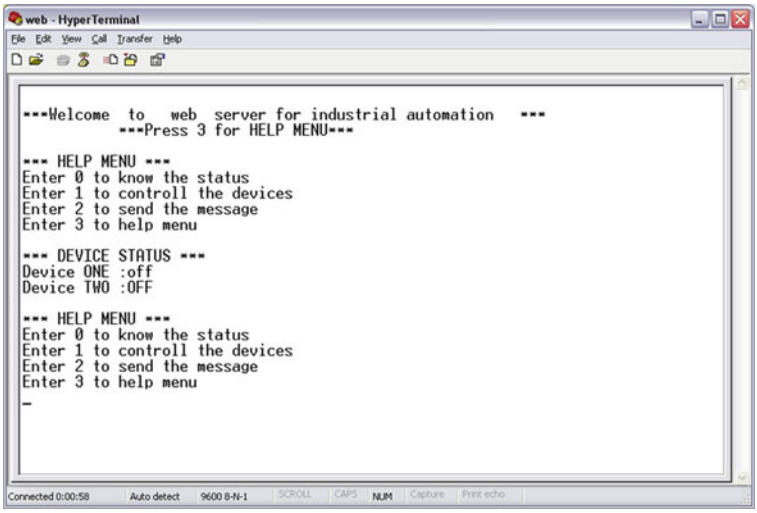
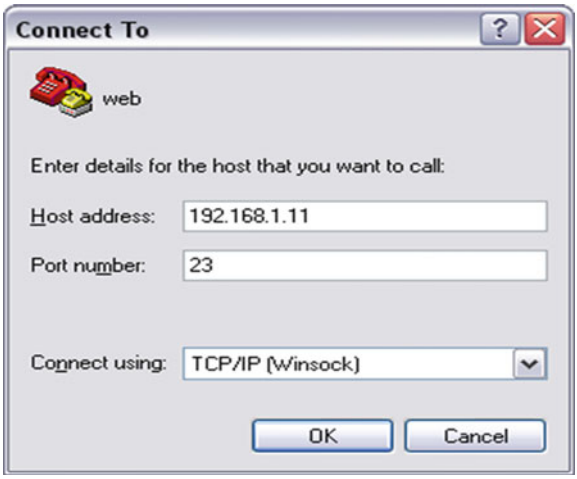


Fig. 8 Snapshot of the connection window displaying the device status

Fig. 9 Snapshot of the setting up of the TCP/IP connection and entering the IP address



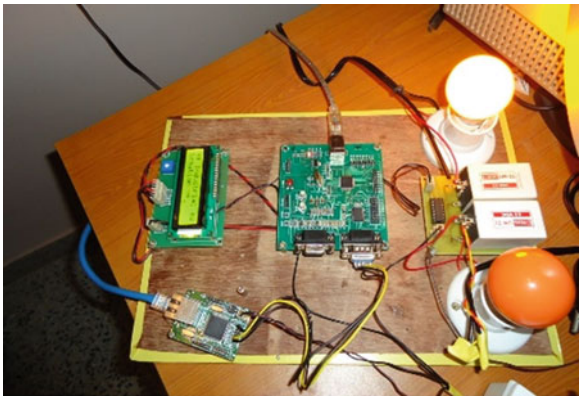


Fig. 10 Snapshot of the device being controlled by a remote client

6 Conclusions

The developed EWS makes use of the TCP/IP (Ethernet) module with the known IP address and is programmed using the ARM board. It is connected to the Ethernet card of the PC in the real-time environment such that the user can sit anywhere in the world to access this hardware interfaced to the ARM board. Two devices have been interfaced to the ARM board. The developed EWS has been able to monitor and control the status (ON/OFF) of the devices connected to it. Application area of this technology includes telecommunication devices, industrial controllers, measuring instruments, hospitals, factory automation, home automation, and consumer electronics.

References

1. Choi MJ, Ju HT, Cha HJ, Kim SH, JW-K Hong (2000) An efficient embedded web server for web-based network element management. Network operations and management symposium, pp 187–200, Honolulu, HI,10–14 Apr 2000
2. Lin T, Zhao H, Wang J, Han G, Wang J (2004) An embedded web server for equipment. In: IEEE proceedings of the 7th international symposium on parallel architectures, algorithms and networks (ISPAN'04), pp 345–350, 10–12 May 2004
3. Mei-qiong Z, Chang-peng J (2008) Research and implementation of embedded web server. International conference on multimedia and information technology, pp 123–125, Feb 2008
4. Liu L (2010) Research on technology of embedded web server application. 2nd IEEE International conference on information management & engineering (ICIME), pp 187–189

High-Performance Service Scheduling Strategy for Cloud Environment

Y. Shobha and M. Dakshayini

Abstract Cloud computing, as an emerging computing model, provides storage, network and computing resources as services to the requested users on demand over the Internet. Efficient and convenient utilization of cloud resources to achieve high throughput and quality of service (QoS) is a major challenge. Hence, in order to achieve good QoS and high performance of the cloud, this paper has proposed a data center-based cloud architecture and an efficient service scheduling strategy. This is done by monitoring the changing resource and virtual machine (VM) status with dynamic VM configuration. The main targets of the proposed work are to achieve high throughput with good QoS. The simulation results demonstrate that it achieves the objectives.

Keywords Resource allocation · Service scheduling · Virtual machine provisioning · QoS · Performance

1 Introduction

Cloud computing is emerging as a new and efficient way of providing computing resources to all the communities including research [1] and business [2]. Cloud computing resources such as infrastructure, platforms and software applications are made accessible over the Internet to remote users with services with scalability and reliability of a data center.

Y. Shobha (✉)
Research scholar in CMJ University, Department of CSE, BIT, Bengaluru, India
e-mail: shobha.bitcse@gmail.com

M. Dakshayini
Department of ISE, BMSCE, Bengaluru, India
e-mail: dakshayini.ise@bmsce.ac.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_71,

Cloud computing is characterized as making services available on demand. Cloud management is responsible for all the resources used by all the applications deployed in the cloud [3]. Cloud depends on the virtualization of resources to provide management combined with separation of users. Although cloud computing has been extensively adopted by the industry and other communities, still there are several research issues to be addressed like resource scheduling, resource management, security, and fault tolerance.

2 Related Work

Resource scheduling and management, fault tolerance, and QoS are the major research areas in the cloud computing environment. These parameters play critical role in cloud and grid computing. Resource scheduling and management system are responsible for selecting and allocating the required appropriate set of resources to the user's job and managing them. Today, many research works have been done on resource scheduling and management in cloud and grid computing.

Armstrong and Agarwal [1] have developed a virtual machine resource manager for distributed compute clouds. Calhirors et al. [4] have presented an adaptive provisioning mechanism for the delivery of resources of Software as a Service (SaaS) applications. Fan and Zhao, Daksha [5] have developed a service-oriented storage resource architecture. Li and Li [6] have presented an efficient cloud resource provisioning approach. Li and Guo [2] have proposed a model for resource scheduling in cloud computing based on stochastic integer programming technique, but none of these papers have considered the concept of admission control and deadline-based service.

In this paper, we are proposing an efficient service scheduling policy and the supporting cloud architecture to achieve high throughput and optimization.

The organization of rest of the paper is as follows: In Sect. 3, we present a model of the problem, and Sect. 4 describes the proposed approach and algorithm in detail. In Sect. 5, we present a simulation model and the results. Finally, in Sect. 6, we conclude the paper and refer to further work.

3 The Optimization Problem

3.1 System Model

The architecture of the cloud environment considered in this work is depicted in Fig. 1. This cloud environment consists of a cluster of N data centers (DC) DC_1, DC_2,…,DC_n. Structure of data center is shown in Fig. 2. The virtual machines (VMs) are instantiated at the DC based on the requirement specifications of the

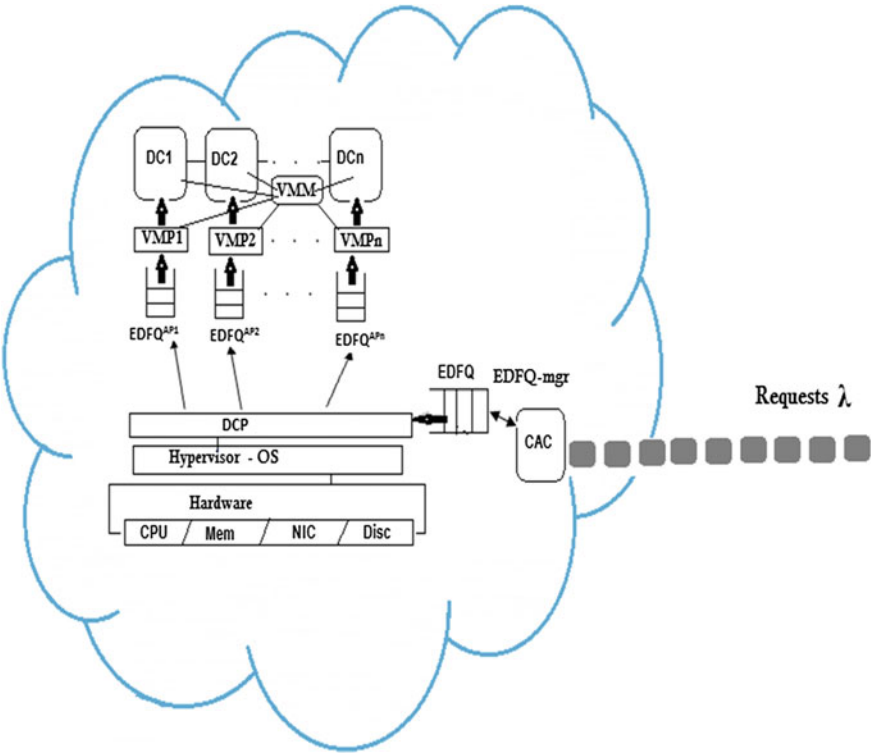


Fig. 1 Data center-based cloud architecture

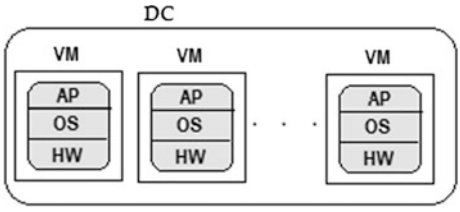


Fig. 2 Structure of data center

user. A DC that runs the specific applications and the VM that processes the user job takes care of the software and data requirements for the execution of that job. A software agent, Resource Status Monitoring Agent (RSMA), runs through these DCs to monitor and report the status of resources and VMs of DCs to virtual machine manager (VMM).

With this system model, user's requests arrive according to Poisson process with the mean arrival rate λ.

The total service time (Tst_i) of ith user request (Usr-Req^i) in the cloud is as follows:

$$Tst_i = Tt_i^{\text{EDFQ}} + Tt_i^{\text{EDFQ}^{\text{AP}}} + \mu_i$$

where

Tst_i	Total time a Usr-Req^i spends in the cloud to get serviced
Tt_i^{EDFQ}	The total time a Usr-Req^i waits in the main Early Deadline First Queue (EDFQ)
$Tt_i^{\text{EDFQ}^{\text{AP}}}$	The total time a Usr-Req^i waits in the Application-specific Early Deadline First Queue (EDFQ$^{\text{AP}}$)
μ_i	The total processing time taken by the Usr-Req^i

The maximum delay a Usr-Req can bear in the cloud to get serviced is specified as deadline D_L by the user in service-level agreement (SLA). To meet the QoS requirement of the user request,

$$Tst_i \leq D_L^i$$

The other optimization variable considered is throughput of the cloud system, $C_{\text{thr}} = \frac{\text{TNR}_{\text{ser}}}{\text{TNR}_{\text{arv}}}$. Cloud system throughput is the ratio of the number of requests served (NR_{ser}) to the total number of requests arrived (NR_{Arv}) at the system.

Our optimization model is to maximize the overall performance of the cloud system C_{thr} by providing the best QoS.

That is,

- **Maximize** Throughput of the cloud

$$C_{\text{thr}} = \frac{\text{TNRser}}{\text{TNRarv}} \quad \text{Subject to} \quad Tst_i \leq D_L^i$$

4 Proposed Approach and Algorithm

4.1 Architecture

The cloud environment considered in this work is shown in Fig. 1. This system consists of various mechanisms as explained below.

Cloud Admission Control (CAC)

CAC is an important component at which all the requests from various users arrive. Decision of accepting or rejecting the user request is made by CAC by communicating with DCP and VMM. Once the Usr-Req is accepted, CAC forward it to EDFQ manager (EDFQ-mgr). EDFQ-mgr inserts this Usr-Req into the main EDFQ (implemented using linked list) based on the D_L specified in its SLA.

Hence, all the Usr-Reqs in EDFQ are stored along with their SLA and sorted on their D_L.

Data Center Provisioner (DCP)

DCP reads the Usr-Reqs from the main EDFQ one after the other and inserts them into appropriate $EDFQ^{AP}$ based on the D_L of the Usr-Reqs. Hence, all the Usr-Reqs in each $EDFQ^{AP}$ are also stored along with their SLA and sorted on their D_L.

Virtual Machine Provisioner (VMP)

VMP reads Usr-Reqs from its $EDFQ^{AP}$ and checks for the availability of the required resources. If resources required to process the job are available, then VMP configures the VM for their requirement and provides it to the Usr-Req and initiates the processing.

Resource Status Monitoring Agent (RSMA)

A software agent named RSMA keeps running across DCs for continuous monitoring of the status of VMs (failure, extra VM assigned, shortage of any resource) that are executing the user jobs. If RSMA comes across any such fault status with any VM in any DC, it reports to the virtual machine manager (VMM). This report is used to reallocate the resources dynamically, when shortage of resources is noticed. When the processing of the user job is successfully completed, VM allotted is freed back to VM pool of DC^j and processed job is given to the user.

Virtual Machine Manager (VMM)

If processing of the user job is not successful due to any type of fault, RSMA immediately reports to VMM. VMM checks for the availability of the required resources in the same DC or in other DCs.

If the required resources are available, configure and allocate the necessary VM and then migrate the job and initiate the continuation of job processing.

If the required resources are not available, VMM checks with the report of extra resources allocated to user jobs. Using the dynamic resource reallocation technique, the extra resource allocated is deallocated. With such deallocated resources, necessary VM is reconfigured and assigned to continue the processing of the stopped job by the corresponding DC and then initiate the continuation of the processing of the job.

If the necessary VM is not obtained with the dynamic resource reallocation also, then the job is inserted back to $EDFQ^{AP}$ that may result in D_L Miss for that job which is very minimal in our work.

4.2 Algorithm

Any user request Usr-Req that arrives at the cloud undergoes the following mechanisms.

Admission Control

```
Cloud admission controller CAC contacts the VMM
  If (VM_req-cinfg is available to serve the Usr-Req^i)
      EDFQ-mgr  ←  Usr-Req^i
    INSERT (Usr-Req^i+ SLA^i, EDFQ, D_L)
  else
     perform dynamic reallocation
```

Dynamic VMP Algorithm

```
For i = 1 -  QS_EDFQ
     INSERT (Usr-Job^i + SLA^i, EDFQ^APi, D_L)
End
For j = 1 to N
  For i = 1 to QS_EDFQAPj
      Usr-Job^i = Config-VM_req
       initiate the processing.
  End
End
```

5 Simulation Model and Performance Evaluation

5.1 Simulation Model

The cloud environment considered in our simulation model consists of a single cloud with a group of 5 DCs with parameters shown in the Table 1. We assume that there is sufficient bandwidth in the cloud network.

Simulation has been conducted for 5 h.

Table 1 Simulation parameters

Notation	System parameters	Default values
N	No. of DCs	5
M	No. of application based sub queues	5
QS_{EDFQ}	Size of main queue EDFQ	100 user reqs
QS_{EDFQAP}	Size of application based sub queue EDFQ	20 user reqs
λ	Mean requests arrival rate at the cloud	58 reqs/h

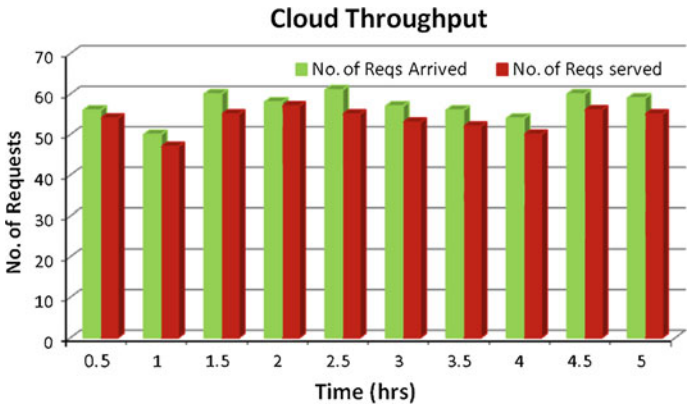


Fig. 3 Cloud throughput

5.2 Performance Evaluation

High Cloud Throughput With the proposed strategy, Usr-Reqs are queued up in the main queue EDFQ based on their negotiated deadline. All these jobs are again distributed appropriately among a set of N second-level application-based EDFQs. This is done by DCP based on their job type and resource requirements, and then these jobs get serviced by the corresponding data centers. Hence, by distributing the jobs among N data centers based on their application requirement, more number of jobs are serviced at the faster rate. This strategy with the proposed cloud architecture has achieved very high (92 %) throughput of the cloud with guaranteed QoS as shown in Fig. 3.

6 Conclusion

In this paper, an efficient service scheduling strategy is proposed to achieve high performance of the cloud system. This strategy with the proposed cloud architecture has achieved very high (92 %) performance with guaranteed QoS. Distribution of jobs among N data centers based on their application requirement has made the system to service more number of jobs at the faster rate. In our future work, we are planning to extend this model to implement distributed cloud system to improve the performance of the cloud system.

References

1. Armstrong P, Agarwal A (2010) Cloud scheduler: a resource manager for distributed compute clouds. arXiv:1007.0050vl [cs.DC] 30 Jun 2010
2. Li Q, Guo Y (2010) Optimization of resource scheduling in cloud computing. 12th ISSNASC, pp 315–320. doi: 10.1109/SYNASC.2010.8
3. Dakshayini M, Guruprasad HS (2011) An optimal model for priority based service scheduling policy for cloud computing environment. Int J Comput Appl (IJCA) (0975–8887) 32(9): 23–29, ISBN: 978-93-80865-22-4
4. Calhirors RN, Ranjan R, Buyya R (2011) Virtual machine provisioning based on analytical performance and QoS in cloud computing environment. In: International conference on parallel processing 2011, ISBN: 978-0-7695-4510-3, DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/ICPP.2011.17
5. Fan Z, Zhao X (2010) Service-oriented storage resource architecture for cloud computing. IEEE, pp 400–405
6. Li C, Li LY (2012) Optimal resource provisioning for Cloud Computing environment. J Supercomput. doi: 10.1007/s11227- 012-0775-9, April 2012
7. Ericson K, Pallickara S (2010) Survey of storage and fault tolerance strategies used in cloud computing, chap 6. In: Furht B, Escalante A (eds) Handbook of cloud computing, Springer, New York

Embedding Patient Information in Medical Images Augmented with Compression by Rotation Technique

C. Nagaraju, S. Parthasarathy and M. B. Subramanya

Abstract This paper presents embedding of patient information inside the medical image for transmission, storage, and telemedicine applications. New efficient compression technique is implemented to achieve high data rate in transmission over a communication channel. There are numerous different techniques present for image compression. Many current compression schemes provide a very high compression rate but with loss of quality in the compressed image. In several applications of medicine, maintaining high-quality image is a main criterion in diagnosing a patient. This paper discusses a hybrid model of compression in maintaining the quality of the medical images, and patient information is recoverable without distortion. The patient information is encrypted before interleaving with medical images to ensure greater security. The proposed algorithm is applied for medical images like CT, MRI, and angiogram. The patient information is embedded inside the medical images before the transmission image is compressed using rotation technique. This paper presents a method to enhance medical data security, confidentiality, quality of the image. The performance is quantified through statistical parameters like MSE and PSNR.

Keywords Compression · Embedding · Mean square error (MSE) · Patient information · Peak signal-to-noise ratio (PSNR)

C. Nagaraju (✉) · S. Parthasarathy · M. B. Subramanya
BGSIT, Professor, PESCE, Research Scholar, IIT Roorkee, Roorkee, India
e-mail: nagaraj.bgs@gmail.com

S. Parthasarathy
e-mail: vsarathypartha@yahoo.com

M. B. Subramanya
e-mail: panilmysster@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_72,

1 Introduction

Medical imaging techniques provide human body pictures in digital form with prohibitive data; hence, compression is necessary for storage and communication purposes. Medical images are significant in their characteristics and purposes. Medical images are generally acquired from image processing techniques, such as computed tomography (CT), magnetic resonance (MRI), ultrasound (US), X-ray, and positron emission tomography (PET). In practice, medical image compression must result in lossless image or due to presence minor loss the diagnosis may wrong. To address this problem in the area of image compression, many efficient techniques with significant characteristics and features have recently been developed for both lossy and lossless compression [1, 2].

From the last few years, various works have been done by International Telecommunication Union and International Organization for Standardization (ISO) to ascertain an international standard for compressing grayscale and color images. The result of their work is JPEG, the Joint Photographic Experts Group. JPEG is mainly for compression of continuous time (multilevel) digital images. After examining a number of coding methodologies, the JPEG members selected discrete cosine transform (DCT) as the best and JPEG was selected as Draft International Standard (DIS) in 1991 and an International Standard (IS) in 1992 [1, 3].

The main advantage of image compression is for transmission and storage. The image compression techniques are used to obtain minimized bit rate with respect to some distortion constraint. To achieve minimized bit rate in image compression, correlation of pixel intensities is most widely used. The most standard compression methods are predictive or waveform coding and transform coding. In predictive coding or waveform coding, pixel correlation can be reduced directly in the spatial domain [4, 5]. By using transform coding, intensity value can be represented by few parameters by transforming into another domain. All of the above methods use mean square error (MSE) to make mathematical analysis more easy and efficient.

Many researchers have attempted to apply digital image compression methods to medical image. The initial emphasis was on information preserving methods. Scan pixel difference was researched by Takaya et al. in [6]. Assche et al. [7] explored the inter-frame redundancy in Rajendra Acharya [8]. Linear predictive coding schemes were investigated. The experiments on lossless compression techniques have all resulted in low compression rate. Transform coding schemes such as principal component analysis (PCA) and discrete cosine transform (DCT) were applied in Refs. [9, 10] to get better rates. In order to achieve higher compression rates without detracting from quality, region of interest-based methods were investigated in the recent years. In Hu et al. [9], an ROI-DCT algorithm is implemented where more DCT coefficients are utilized in ROI. Cosman et al. [11] used a sub-band compression scheme in Refs. [12] and [13] for application to mammography. 3-D wavelet compression was investigated.

Another important application for encoding is transmitting binary data through the Internet because the binary data may be corrupted if not encoded as different systems interpret the data differently. Hence, data are encoded before sending and decode it upon arrival as it assures that the data are not corrupted while communication.

Encoding and encryption are two different security features used in image processing techniques. In encoding, data are easily converted from one representation to another where only specific algorithm is needed, and then compared with encryption; no key is required. For example, a thief is standing in front of house with the door bearing more than dozen knobs but no lock. Now the thief has find to the right knob to enter inside the house. Similarly, when storing or transmitting sensitive information should always use encryption where as encoding algorithms can be implemented with many different encryption algorithms (ciphers), perhaps even more than the former. It is worth noting that ciphers typically have very short life spans, and while popular ciphers in use today have withstood rigorous attacks [14], it is likely that will not always be the same.

The present paper is organized in four sections. Section 2 describes a detailed methodology adopted by the authors for interleaving of text in images. Later medical image is compressed by performing rotation method. Section 3 presents results and discussion with visual snapshots. Section 4 draws conclusions of present technique to embed text in images.

2 Methodology

Figure 1 indicates the steps involved in interleaving an image with electronic patient record (EPR) before transmission and retrieval of the same at the receiving end of the system. The proposed scheme involves reading EPR as text file, encryption of the text file, and interleaving the encrypted text in images followed by retrieving the interleaved text file from images.

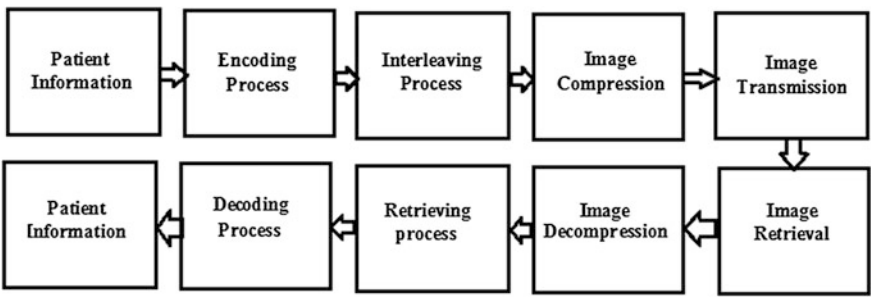


Fig. 1 Block diagram representation of trans-receiver system

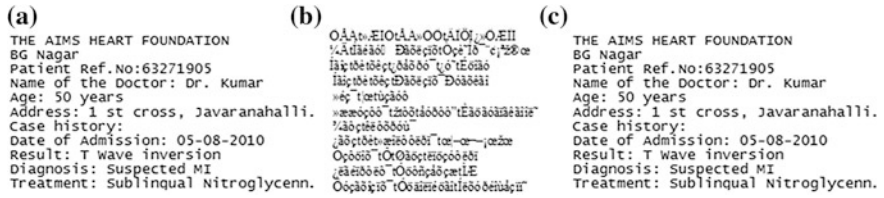


Fig. 2 Encoded and decoded patient information **a** Original Patient Information **b** Encoded Patient Information **c** Retrieved Patient Information

2.1 Encoding

In the proposed scheme, patient information within a text document is represented using ASCII or Unicode. In order to embed patient information in the medical image, text document is represented in binary pattern. The patient information in the text file is as shown in Fig. 2a. The patient information is encoded using

$$T_e = (\log(T_o \times 2) \times 100) - 300 \tag{1}$$

where T_e is the encrypted text and T_o is the ASCII or Unicode of the original text. The encrypted information (T_e) is stored as an integer. The ASCII or Unicode of the encrypted text shown in Fig. 2b.

2.2 Interleaving Process

The binary data obtained from the ASCII or Unicode of the text file is swapped with the least significant bit (LSB) of the grayscale image bit by bit. Eight bits of each ASCII or Unicode thus replace eight consecutive pixels of LSBs of the image [8]. This cycle of interleaving of ASCII or Unicode in consecutive pixels is repeated to include all the characters in the text file. The ASCII or Unicode of the encrypted text shown in Fig. 2b are broken into bits and interleaved into the pixels of desired medical images of the patient using

$$\text{image slice } (n) = \text{rem}\left(\frac{\text{image pixel or quotient of image pixel}}{2^n}\right). \tag{2}$$

2.3 Decoding

Equation 3 is used for the retrieval of patient information from interleaved image.

$$T_o = \exp\left((T_e + 300)/100 - \log(2)\right) \tag{3}$$

Even when T_e is rounded off to the nearest integer, exact data are reconstructed from the Eq. 3.

2.4 Image Compression

Compressing an image and compressing raw binary data are significantly different. General purpose compression programs can be used to compress images, but the result is less than optimal. Some of the finer details in the image can be sacrificed for the sake of saving a little more bandwidth or storage space.

In case of lossless compression, decompressed image will be an exact replica of the original image. Most of the application an approximation of the original image is enough, as long as the error between the original and the compressed image is tolerable. In medical image, information about the patient is predominant in high-frequency region. In the present methodology, image compression has done using a new technique based on pixel rotation which will not affect the high-frequency component of the medical image, but there will be a slight changes in low-frequency region of medical image, which will not lead to wrong diagnosing the

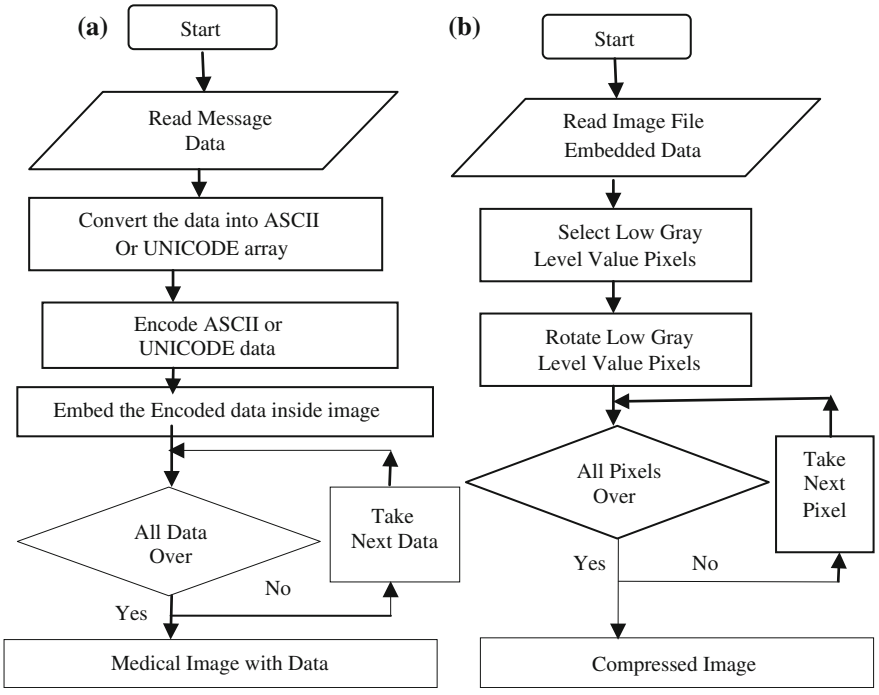


Fig. 3 **a** Embedding data in medical image, **b** Compression using rotation

patient. The flow diagrams of entire process of embedding the information inside the medical image and compression of the image are shown in Fig. 3.

2.5 Error Metrics

Mean square error (MSE) and the peak signal-to-noise ratio (PSNR) are measured using original image and compressed. Cumulative squared error between the original and compressed image will be MSE, and PSNR is a measure of the peak error. The mathematical formulae for the two are:

$$\text{MSE} = \frac{1}{XY}\sum_{I=1}^{X}\sum_{J=1}^{Y}[I(x,y) - I'(x,y)]^2 \tag{4}$$

$$\text{PSNR} = 20 \times \log_{10}\left(255/\sqrt{\text{MSE}}\right) \tag{5}$$

where $I(x,y)$ is the original image, $I'(x,y)$ is the approximated version (which is actually the decompressed image) and X and Y are the dimensions of the images. Here, the 'signal' is the original image, and the 'noise' is the error in reconstruction. Compression scheme with lower MSE (and a high PSNR) is recognized as the better one.

3 Results and Discussion

The measurable statistical parameters are PSNR and MSE. These are the important quantitative assessment parameters adopted in the image community to study the quality of interleaved images. The proposed method is applied on more than 32 different images with different sizes. Table 1 depicts the statistical parameter

Table 1 MSE and PSNR

Modality	Size	MSE	PSNR
Angiogram	256 × 256	5.4994	40.7277
	303 × 256	5.1629	41.0019
	256 × 256	2.8234	43.6231
CT	209 × 256	1.9533	45.2632
	256 × 256	3.9559	42.1584
	256 × 256	4.4407	41.6563
MRI	256 × 256	5.5829	40.6622
	210 × 256	2.4230	44.2873
	199 × 256	6.1113	40.2695

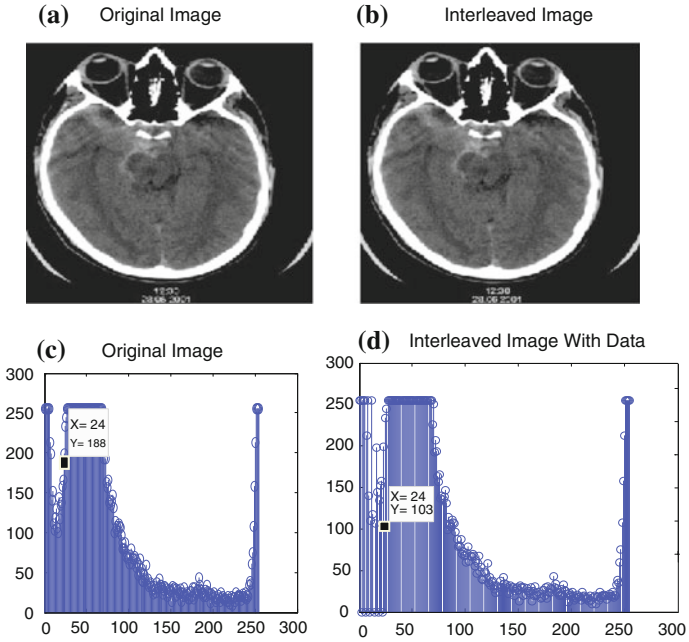


Fig. 4 **a** Original image, **b** Interleaved image, **c** Original image histogram, **d** Interleaved image histogram

obtained for interleaved image. From Fig. 4c, d, and Fig. 5c, d, it concludes no changes in the visual quality of original, interleaved, and reconstructed Image.

The image quality is not degraded on the account of the fact that the change in the LSB of a pixel changes its brightness by one part in 256. The text can be interleaved into all the pixels of LSB plane of the image. Figure 4a, b shows variation in the pixel values between the original image and interleaved image, indicating the presence of patient information in interleaved images. The shape of the histograms signifies that the distribution of pixels remain same before interleaving and after interleaving in high-frequency region.

Histogram of original image and compressed image shown in Fig. 5c, d indicates that no significant variation in high-frequency region and it will not affect to readout the patient condition based on that image.

From the histogram of reconstructed image, it concludes that no variations in high-frequency region in comparisons with the original image which indicates that after compression also it will not affect in diagnosing a patient (Fig. 6a–d).

According to [15], an improvement in algorithm is recognized by a high PSNR or a lower MSE. In agreement with this, the results of the proposed systems with high PSNR prove that reconstructed image will match the original image. Similarly, according to the report of [16], a PSNR value in the range 20–40 indicates

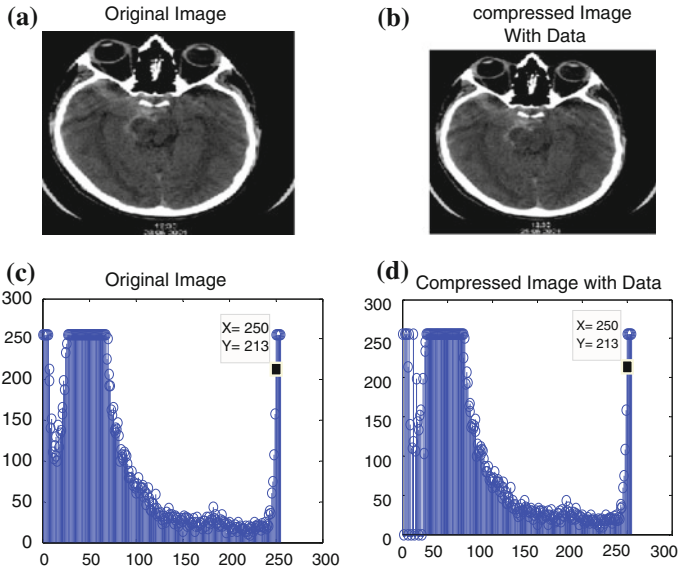


Fig. 5 **a** Original image, **b** Compressed image, **c** Original image histogram, **d** Compressed image histogram

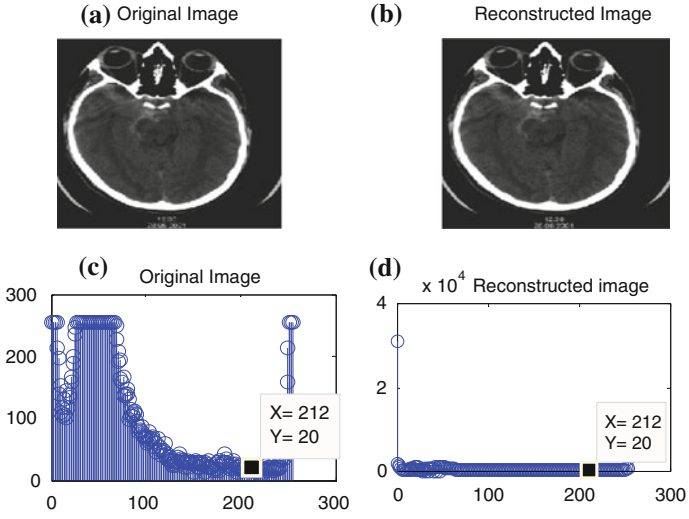


Fig. 6 **a** Original image, **b** Reconstructed image, **c** Original image histogram, **d** Reconstructed image histogram

that the resultant image is a very good match to the original image. In accordance with this report, the results shown in Table 1 of the proposed algorithm produce PSNR values in the range 35–45 dB proving that the proposed algorithms do not

degrade the images with the patient information inside the compressed medical image.

4 Conclusion

This paper has presented an efficient technique of interleaving patient information such as text documents in medical image for efficient storage and transmission. Text files are encrypted before the transmission to provide additional security. The technique is tested for different images; the MSE was found to be less than 12 and PSNR is between 35 and 45 dB. Transmission of data through a channel and retrieval of stored data from memory are prone to some kind of errors. Future work includes the effect of noise on interleaved data using standard simulation techniques and encoding the data using suitable error correction techniques to reduce the impact of these errors significantly.

References

1. Wallace GK (1991) The JPEG still picture compression standard. Commun ACM 34(4):30–44
2. Carpentieri B, Weinberger MJ, Seroussi G (2000) Lossless compression of continuous-tone images. Proc IEEE 88(11):1797–1807
3. Shapiro JM (1993) Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans Signal Process 41:3445–3462
4. ISO/IEC IS 10918 (1991) Digital compression and coding of continuous tone still images
5. Vlaciu A, Lungu S, Crisan N, Persa S (1995) New compression techniques for storage and transmission of 2-D and 3-D medical images. In: Advanced image and video communications and storage technologies, 2451, pp 370–377, Amsterdam, Netherlands, March 1995
6. Takaya K, Tannous CG (1995) Information preserved guided scan pixel difference coding for medical images. WESCANEX 95. Communications, power, and computing. Conference proceedings, IEEE, 1, pp 238–243
7. Van Assche S, De Rycke D, Philips W, Lemahieu I (2000) Exploiting interframe redundancies in the lossless compression of 3D medical images. In: Data compression conference, p 575
8. Rajendra Acharya U, Deepthi Anand, Subbanna Bhat P, and Niranjan UC (2001) Compact storage of medical images with patient information communication. IEEE Trans Inf Tech Biomedicine, 5(4), 320–323
9. Hu J-H, Wang Y, Cahill PT (1997) Multispectral code excited linear prediction coding and its application in magnetic resonance images. IEEE Trans Image Process, 6(11), 1555–1566
10. Taur JS, Tao CW (1996) Medical image compression using principal component analysis. International conference on image processing, 1, 2, 903–906
11. Cosman PC, Gray RM, Olshen RA (1994) Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy. Proc IEEE 82(6):919–932
12. Yoshioka M, Omatu S (1996) Image Compression by nonlinear principal component analysis. IEEE conference on emerging technologies and factory automation, EFTA '96, 2, pp 704–706

13. Gray RM, Olshen RA, Ikeda D, Cosman PC, Perlmutter S, Nash C, Perlmutter K (1995) Evaluating quality and utility in digital mammography. In: Proceedings of international conference on image processing 2:5–8
14. Baskurt A, Peyrin F, Benoit-Cattin H, Goutte R (1993) Coding of medical images using 3D wavelet decompositions IEEE international conference on acoustics, speech, and signal processing, ICASSP-93, 5, pp 562–565
15. Schneier, M. and Abdel-Mottaleb, M. (1996) Exploiting the JPEG compression scheme for image retrieval. IEEE Trans.Pattern Anal. Mach. Intell., Vol.18, No. 8, pp 849–853
16. Zhang, Y. (1998) Space-Filling Curve Ordered Dither, Elsevier, Computer & Graphics Journal, Vol. 22, No. 4, pp 559-563

Development of Morphological Stemmer, Analyzer and Generator for Kannada Nouns

M. C. Padma and R. J. Prathibha

Abstract Kannada is an inflectional and morphologically rich language. Developing a well-fledged morphological stemmer, analyzer, and generator (MSAG) model for Kannada is a challenging task. In any natural language processing applications, there is a great demand for MSAG model. In this paper, a hybrid MSAG model is proposed as a part of machine translation system from Kannada to any other language. This model is developed using suffix-stripping, rule-based, and paradigm-based approaches. The performance of this model is tested against a set of nouns randomly taken from a well-known dictionary called "***Kannada Rathna Kosha***" [1].

Keywords Natural language processing · Machine translation · Paradigms · Suffix stripping · Morphological analyzer · Morphological generator

1 Introduction

Natural language processing (NLP) is a field of computer science and linguistics concerned with the interaction between computer and natural language [2]. The aim of NLP is to study the problems in the automatic generation and understanding of natural languages. Machine translation (MT) is a subfield of NLP that translates given sentences from one natural language called source language into another

M. C. Padma (✉)
Department of Computer Science and Engineering, P E S College of Engineering, Mandya, Karnataka 571401, India
e-mail: padmapes@gmail.com

R. J. Prathibha
Deptartment of Information Science & Engineering, S J College of Engineering, Mysore, Karnataka 570006, India
e-mail: prathirj@yahoo.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_73,

natural language called target language. As India is a multilingual country, having 18 constitutional languages, MT plays a very significant role. It is necessary to break the language barrier among Indian languages and enable easier communication among people of different states [3]. In the Indian context, there is a great demand for developing automated tools for machine translation system (MTS). Indian languages are agglutinative, inflectional, derivational, and morphologically rich. We find many challenges in the design of MTS for Indian languages. Some of the issues related with MT are morphological structure, grammar, morphological analysis, syntactic analysis, semantic analysis, etc., of natural languages.

In this paper, we have considered Kannada as the source language which is the official language of Karnataka, a south Indian state. In MTS, there are three phases, namely source analysis, semantic analysis, and target language generation. In source analysis phase, the input sentence is analyzed and checked for its syntactical (grammatical) correctness. The source analysis phase requires some of the modules like stemmer, morphological analyzer, parts of speech (PoS) tagger, chunker, syntactic analyzer. Morphological analyzer (MA) and morphological generator (MG) are very much essential modules in NLP applications like MTS, spell checker, PoS taggers, information retrieval system (IRS), etc. This paper mainly focuses on morphological stemmer, analyzer, and generator model for Kannada nouns.

Currently, several stemmers, MAs, and MGs are available for English language and also for many Indian languages. Some standard stemmers which are generally used for English language [4] are not suitable for Indian languages, because Indian languages are morphologically rich languages. This paper presents MSAG model for Kannada nouns using hybrid approach. Main objective of this model is to get the morphological information like root, number, case, and gender of Kannada nouns.

This paper is organized as follows. Section 2 presents the related work on stemmers, MAs, and MGs for various languages. In Sect. 3, the proposed methods and modules of MSAG model are described. Experimental results and discussion are given in Sect. 4. Conclusion and future work are presented in Sect. 5.

2 Previous Work

The literature shows that several stemmers are available for both Indian and non-Indian languages. Most of them are suitable for information retrieval system [4]. Some of the Indian stemmers designed for information retrieval system are as follows: A lightweight Stemmer for Hindi [5], Assas-Band, an Affix-Exception-List Based Urdu Stemmer [6], Design and Development of a Stemmer for Punjabi [7], and Hybrid Stemmer for Gujarati [8]. There are few stemmers for MTSs [9], they split the given word into stem and suffixes, but they do not provide morphological information about the given word. Morphological stemmers (MS) are very essential in MTS to split the given word into stem and suffix based on its inflection and also to provide morphological information like case, number, and its category.

From the literature survey, it has been observed that since last two decades, the development of morphological analyzer and generator (MAG) has been successfully done for languages like English, European, and Chinese languages using various methods [10]. Menon et al. [11] have proposed a Tamil MAG using finite-state transducer (FST) and rule-based approaches. It uses lexicon and orthographic rules for nouns, verbs, and adjectives. Anand Kumar et al. [12–14] have proposed MG for Tamil language which is based on suffix stripping and FST methodologies to handle nouns and verbs. Sequence labeling MAG [14] uses corpus-based approach, where training and testing of Tamil words are carried out by support vector machine and machine learning algorithm. This system handles compound words, transitive words, and proper nouns of Tamil language. An implementation of APERTIUM MAG for Tamil attempted by Parameshwari [15] uses paradigm-based approach.

Telugu MG proposed by Sribadri Narayanan et al. [16] does not require any orthographic and morph-tactics rules. It automatically extracts the suffix list and concatenates the lemma and morphological features to generate the word. A novel approach proposed by Sunitha et al. [17] uses an unsupervised stemmer for improving the performance of existing rule-based Telugu MA. In this work, it requires small or no dictionary. The decomposition of a word into stems and suffixes and re-analyzation of unrecognized word are carried out repeatedly. The major demerit of this module is that it generates many words that are not valid.

Malayalam–Tamil machine translation attempted by Jayan et al. [18] has MA for Malayalam, MG for Tamil, bilingual dictionary of Malayalam and Tamil. It uses suffix-stripping algorithm in MA and suffix joining approach in morphological generator, Vishal Goyal et al. [19] have proposed MA and MG for Hindi using paradigm-based approach. It uses databases where all possible word forms of all root words are stored. Though the search time is very less, it takes more disk space. Hindi derivational MA proposed by Kanuparthi et al. [20] applies derivational rules on derivational suffixes and finds the majority properties to analyze Hindi words.

Vikram et al. [21] have developed a prototype of MA for Kannada language based on FST to handle nouns and verbs. Kannada MAG proposed by Shambhavi et al. [22] uses rule based with paradigm approach. It uses trie structure for storage of suffixes and root words; hence, it consumes more memory. Veerappan et al. [23] have proposed a MAG for Kannada which uses rule-based FST to analyze and generate nouns, verbs, and adjectives.

It is observed from the survey that majority of work on the development of MAG model has been done for Tamil language; hardly few attempts have been made for Kannada language. Most of the competent approaches in MA and MG are FST, rule based, and paradigm based. However, the rules and paradigms are language dependents, because each language has its own morphological structure and linguistic rules.

In this paper, the methods adopted to develop hybrid MSAG model are (a) suffix-stripping method, which is used in morphological stemmer module to remove inflectional suffixes present in the given word. (b) Paradigm-based method, which is used while designing suffix list table. (c) Rule-based method, which is used in creation of noun-base monolingual dictionary.

3 Proposed Model

The proposed MSAG model has three modules and two databases. The three modules are morphological stemmer, analyzer, and generator. The two databases are noun-suffix table and noun-base dictionary. The steps involved in constructing databases are explained below.

3.1 Design and Classification of Noun Paradigms Using Paradigm-Based Approach

Kannada is an inflectional language; various declension words are formed by adding suffixes to the root in a series. In Kannada language, there are several suffixes that are used with root word to form inflectional noun forms or declension forms. In this proposed work, these suffixes are considered as paradigms. Design of paradigms should cover all types of inflections. The biggest challenge is the grouping of nouns in such a way that the members of the same group have similar inflections. The noun-suffix table is constructed to store paradigms in a matrix format. It has 20 rows to represent morpho-lexical forms (MLF) and 14 columns to represent paradigms. A noun in Kannada can belong to any one MLF among 20 varieties. The nouns with same ending and gender need not follow the same MLF. For example, the nouns *guru* (teacher) and *magu* (child) belong to the same gender (neuter) with the same ending (*u karantha*), which follow different MLF for their inflectional forms. Few entries of noun-suffix table are given in Table 1.

3.2 Creation of Noun-Base Dictionary Using Rule-Based Approach

Noun-base dictionary is a monolingual dictionary which is constructed for the proposed MSAG module, using a well-known Kannada dictionary called "***Kannada Rathna Kosha***" [1]. There are three columns in noun-base dictionary, namely (a) base, which represents the base of the noun. (b) Category code, in which the first digit represents ending (*antha*) of base like *u karantha, a karantha,*

Table 1 Noun-suffix table

	P1	P2	P3	P4	P5	::	P14
MLF 1	anu	aru	anannu	arannu	aninda	::	are
MLF 2	yu	gaLu	yannu	gaLannu	yinda		gale
::	::	::	::	::	::	::	::
MLF20	u	ugaLu	Annu	ugaLannu	ininda	..	ugaLe

Table 2 Noun-base dictionary

Base	Category code	Modifier code
guru	u005	00
angi	i202	d1
huduga	a004	d1

e karantha, etc., second digit represents gender, and last two digits represent the MLF index number to which the base belongs. (c) Modifier code, which represents the rule, which is applied on base to derive stem. The rules are like addition of some characters to base, deletion of some characters from the base or no modification on the base, to derive the root. Some of the entries of noun-base dictionary are given in Table 2.

3.3 Methodologies Used in MSAG Model

In this proposed work, the MSAG model has three modules, namely (a) morphological stemmer, (b) morphological analyzer, and (c) morphological generator. The methodologies used in these modules are explained below.

(a) **Morphological Stemmer:** In this module, suffix-stripping algorithm is used to split the given word into stem and suffix. This algorithm does not rely on lookup table where all declension forms are stored for all words; instead, it splits the word into stem and suffix based on its inflection. It searches for inflectional suffixes in noun-suffix table. As the noun-suffix table contains 20 rows and 14 columns covering all suffixes, the size of the file is too small. Hence, the searching process is relatively fast. This module provides all possible stems, suffixes, position of the suffixes in MLF list and MLF index number for the given word. The same suffix can appear in different MLF index list with the same position or different position. Example: for the given input "***huduganu***," the output from morphological stemmer is given in Table 3.

Table 3 Output of morphological stemmer module for input "*huduganu*"

Stem	Suffix	MLF index number	Position of suffix in MLF list
hudugan	*u*	3	1
hudug	*anu*	4	1
hudug	*anu*	5	1
hudugan	*u*	9	1
hudugan	*u*	10	1
hudu	*ganu*	15	1

3.3.1 Algorithm for Morphological Stemmer

Purpose: To split word into stem and suffix.
Input: word w.
Output: List of stems and suffixes.

Steps:

1. Split w from right to left by comparing suffix paradigms in noun-suffix Table.
2. If the split string matches with paradigms

Then, return (stem, suffix, MLF index, position number of the suffix in MLF list)

3. Repeat 1 and 2 until all paradigms of 20 MLFs are checked

(b) **Morphological Analyzer:** The MA module reads the output of the morphological stemmer and looks for the stem in noun-base dictionary. Example: for the given input stem "*hudug*", if the stem is substring of base from the beginning, it gets the corresponding modifier code "d1". The modifier code is used as a rule to derive the stem from the base. Here, "d1" interprets, delete one character from base "*huduga*" to derive stem; hence, "*hudug*" is the stem. If the derived stem and the input stem are same; it gets the corresponding category code (a004; the category code of "*huduga*"). It extracts MLF index number from category code and compares with the input MLF index number, if both of them are same, then the stem is considered as valid stem.

For the given input word "*huduganu*", there are 6 lines in the output from morphological stemmer, among which second line is the valid stem and suffix, depending on the stem and MLF index number from noun-base dictionary. The output of MA module is given below.

Root/stem: "*hudug*", **Gender:** masculine, **Case:** Nominative **Number:** Singular **Ending**: "*u karantha*"

3.3.2 Algorithm for Morphological Analyzer

Purpose: To get morphological features of the input word.
Input: MLF index number, stem, suffix, position number of suffix in MLF list.
Output: Morphological features of the word w.

Steps:

1. In the noun-base dictionary, if the input stem is a substring of base from the beginning, then derive stem from base using modifier code
2. If input stem matches with derived stem, then extract (ending, gender, MLF index number from category code) and return (root, gender, case, number, ending).

(c) **Morphological Generator:** The MG gets base word of a noun, morphological features like gender, case, and number as input from the user, and then, it searches for base word in base-noun dictionary. If it is found, it extracts stem from the base by using rule-based approach which is explained in MA module. It compares input gender with the gender in category code. If it matches then it extracts MLF index number from category code. It uses suffix list table to get the corresponding suffix, based on the MLF index number, case, and number. Finally, it concatenates the stem with suffix to generate the output word.

Input: [Base] "*pustaka*" (book), **[Gender]** Neuter, **[Case]** Nominative, **[Number]** Plural
Output: "*pustakagaLu*"

3.3.3 Algorithm for Morphological Generator

Purpose: To generate a word for given base and morphological feature.
Input: Base b, morphological feature.
Output: Word w.

Steps:

1. If input base b equals to base in noun-base dictionary then return (category code and modifier code)
2. Extract MLF index number, ending, gender from category code
3. Derive stem from base using modifier code
4. Check for MLF index number in noun-suffix table
5. If (found) then return (suffix which is corresponding to morphological feature)
6. Concatenate stem with suffix to form w

4 Experimental Results and Discussion

Each module of MSAG model is tested individually and obtained the expected results. The input is given in English transliterated form. This model is tested for distinct nouns that are taken from a well-known Kannada dictionary "***Kannada Rathna Kosha***" [1].

4.1 Experimental Result of Morphological Stemmer Module

The input which is tested with this module is "***huduganu angadiyinda hannannu tandu maneyalli mithrana ammaLige kaiyinna kottanu***." Here, we have considered a set of words involving all possible cases and numbers of Kannada nouns.

Table 4 Result of morphological stemmer module

hudugan u 3 1	*tand u 3 1*	*ammaLi ge 2 7*
hudug anu 4 1	*tand u 9 1*	*ammaL ige 3 7*
hudug anu 5 1	*tand u 10 1*	*ammaLig e 3 13*
hudugan u 9 1	*tan du 14 1*	*ammaLi ge 6 7*
hudugan u 10 1		*ammaLi ge 7 7*
hudu ganu 15 1	*mane yalli 2 11*	*ammaL ige 9 7*
	mane yalli 6 11	*ammaLig e 9 13*
angadi yinda 2 5	*mane yalli 7 11*	*ammaL ige 10 7*
angadi yinda 6 5		*ammaLig e 10 13*
angadi yinda 7 5	*mithr ana 4 9*	*amm aLige 12 7*
angad iyinda 13 5	*mithr ana 5 9*	*ammaL ige 13 7*
		amm aLige 18 7
hannann u 3 1	*kaiy inna 11 9*	*ammaL ige 19 7*
hann annu 3 3		
hannann u 9 1		*kottan u 3 1*
hannann u 10 1		*kott anu 4 1*
hann annu 19 3		*kott anu 5 1*
		kottan u 9 1
		kottan u 10 1

The tokenizer splits the given sentence into words by considering space as the delimiter. The morphological stemmer splits each word based on the paradigms available in noun-suffix table. It provides all possible stems and suffixes along with MLF index number, position of noun-suffix in the suffix table. The result of morphological stemmer is given in Table 4.

4.2 Experimental Result of Morphological Analyzer Module

MA module takes the output of stemmer module as input and searches for stem in noun-base dictionary. If the stem and MLF index number of morphological stemmer output match with substring of base and MLF index of noun-base dictionary, then it displays the detailed morphological information of the given word. The result of morphological analyzer is given in Table 5.

There is a spelling mistake in the word "*kaiyinna*". The words "*tandu*" and "*kottanu*" are not in the noun-base dictionary. Hence, these words are unknown to this module. The reason for **unknown** result is either the unavailability of given input word in noun-base dictionary or due to spelling mistake in the input word.

4.3 Experimental Result of Morphological Generator Module

The input to this module is base and morphological feature like gender, case, and number. The gender of the base is also considered here, because in Kannada,

Table 5 Output of morphological analyzer

Base	Root/stem	Gender	Case	Number	Ending
huduga	*hudug*	Masculine	Nominative	Singular	*u karantha*
angadi	*angad*	Neuter	Instrumental	Singular	*ikarantha*
hannu	*hann*	Neuter	Accusative	Singular	*u karantha*
tandu unknown word					
mane	*man*	Neuter	Locative	Singular	*e karantha*
mithra	*mithr*	Masculine	Dative	Singular	*a karantha*
amma	*amm*	Feminine	Genitive	Singular	*a karantha*
kaiyinna unknown word					
kottanu unknown word					

single base with different genders is available. For example, "*surya*" can be a masculine gender or neutral gender depending on the context of usage.
Input: [Base] "*mara*", [Gender] Neuter, [Case] Instrumental, [Number] Plural
Output: "*maragaLinda*"

There are 8 cases for Kannada nouns, but we have considered only 7 cases in noun-suffix table because, in Kannada, the declension form of noun in instrumental case (*Thruthiyavibhakti*) and ablative case (*Panchamivibhakti*) is the same [24].

The proposed model is compared with the existing models [21–23], and it is given in Table 6.

Table 6 Comparision of the proposed model with the existing models

Model	No. of nouns tested	Methodologies used	Results	Remarks
Model 1 [21]	500	Finite-state transducer	Root, suffix, and case information of given noun	No gender information is available
Model 2 [22]	3250	Paradigm approach	Root, gender, case, and number (singular/plural) of the given noun	Uses trie structure for storage of suffix and root words; hence, it consumes more memory
Model 3 [23]	20,000	Rule-based approach	Root and number information.	No gender, case information is available
Proposed model	1,50,000	Hybrid approach consisting of suffix-stripping, rule-based, and paradigm-based approaches	Base, root, case, number, gender, and ending of the given noun	Provides morphological information of the given noun in detail.

5 Conclusion and Future Work

The MSAG model for Kannada noun is discussed in this paper. This is a very significant model of NLP for highly inflectional languages. Development of MSAG for Kannada language is not an easy task as the language is agglutinative, inflective, derivational, and morphologically rich. This model is developed using suffix-stripping, rule-based, and paradigm-based approaches, and it handles all nouns that are available in "***Kannada Rathna Kosha***" [1]. The performance of this model is satisfactory because the noun suffixes are grouped and stored in noun-suffix table, based on inflectional forms. This is a preprocessing model for NLP applications like machine translation system, spell checkers, etc., In future, this model can be enhanced for compound words, verbs, adverbs, adjectives, proper nouns, and *samasapada*.

References

1. Nayak HM (ed) Kannada Rathna Kosha, Kannada Sahithya Parishath, Kannada Abhivruddhi pradhikara, Bengaluru (1994)
2. Bharati A, Chaitanya V, Sangal R (1995) Natural language processing: a paninian perspective. Prentice-Hall of India, New Delhi
3. Bharati A, Chaitanya V, AP Kulkarni, Sangal R (2001) Anusaaraka: overcoming the language barrier in India, [To appear in anuvad: approaches to translation, Rukmini Bhaya Nair, Sage, New Delhi. 2001]
4. Porter MF (1980) An algorithm for suffix stripping. 14(3):130–137. http://www.cs.odu.edu/~jbollen/IR04/readings5.pdf
5. Ramanathan A, Rao DD (2003) A light weight stemmer for Hindi. 10th conference of the European chapter of the association for computational linguistics on computational linguistics for South Asian Languages
6. Akram QA, Naseer A, Hussain S (2009) Assas-bond, an affix-exception-list based urdu stemmer, ACL-IJCNLP, pp 40–47, Suntec, Singapore
7. Kumar D, Rana P (2010) Design and development of stemmer for Panjabi. Int J Comput Appl (0975-8887) 11(12):18–23
8. Patel P, Popat K, Bhattacharya P (2010) Hybrid stemmer for Gujarati. WSSANLP, 23rd International conference on computational linguistics (COLING), Beijing, pp 51–55
9. Ramachandran VA, Krishnamurthi I (2012) An iterative stemmer for Tamil language, ACIIDS, Part III, LNAI 7198, pp 197–205, Springer, Heidelberg
10. Antony PJ, Soman KP (2012) Computational morphology and natural language parsing for Indian languages: a literature survey, Int J Comput Sci Eng Technol 3(4):136–146, ISSN: 2229-3345
11. Menon AG, Saravanan S, Loganathan R, Soman K (2010) Amrita morph analyzer and generator for Tamil: a rule based approach, Tamil international conference, Coimbatore, India
12. Anand Kumar M, Dhanalakshmi VV, Rajendran S (2010) A novel algorithm for Tamil morphological generator
13. Anand Kumar M, Dhanalakshmi VV, Rajendran S (2010) A novel data driven algorithm for Tamil morphological generator. Int J Comput Appl (0975–8887) 6(12):52–56

14. Anand Kumar M, Dhanalakshmi VV, Soman KP, Rajendran S (2010) A sequence labeling approach to morphological analyzer for Tamil language. Int J Comput Sci Eng 02(06):1944–1951
15. Parameshwari K (2011) An implementation of APERTIUM morphological analyzer and generator for tamil, language in India www.languageinindia.com, special volume: problems of parsing in Indian languages
16. Sribadri Narayanan R, Saravan S, Soman KP (2011) Data driven suffix list and concatenation algorithm for Telugu morphological generator. Int J Eng Sci Technol 3(8):6712–6717, ISSN: 0975-5462
17. Sunitha KVN, Kalyani N (2009) A novel approach to improve rule based Telugu morphological analyzer. 978-1-4244-5612-3/09$26 © 2009 IEEE, pp 1649–1652
18. Jayan JP, Rajeev RR, Rajendran S (2011) Morphological analyzer and morphological generator for Malayalam–Tamil machine translation. Int J Comput Appl (0975-8887) 13(8)
19. Goyal V, Singh Lehal G (2008) Hindi morphological analyzer and generator, 978-0-7695-3267-7/08 $25.00 © 2008 IEEE DOI 10.1109/ICETET.2008.11
20. Kanuparthi N, Inumella A, Sharma DM (2012) Hindi derivational morphological analyzer. In: Proceedings of the twelfth meeting of the special interest group on computational morphology and phonology (SIGMORPHON2012), pp 10–16, Montreal, Canada, © 2012 Association for Computational Linguistics
21. Vikram TN, Urs SR (2007) Developed a prototype morphological analyzer for South Indian language of Kannada, ICADL, LNCS 4822, pp 109–116, ©Springer, Berlin
22. Shambhavi BR, Ramakanth Kumar P, Srividya K, Jyothi BJ, Kundargi S, Varsha Shastri G (2011) Kannada morphological analyser and generator using trie, Int J Comput Sci Netw Secur 11(1):112
23. Veerappan R, Antony PJ, Saravanan S (2011) A rule based Kannada morphological analyzer and generator using finite state transducer. Int J Comput Appl (0975-8887) 27(10)
24. Srikanthaiah TN (2012) Kannada madhyama vyakarana. Geetha Book House, Mysore

Quadtree Based Feature Extraction Technique for Recognizing Handwritten Kannada Characters

M. C. Padma and Saleem Pasha

Abstract Automatic recognition of handwritten characters is a challenging task and has much scope in the field of document image analysis. In this paper, a quadtree-based feature extraction technique for recognizing handwritten Kannada characters is presented. The proposed technique uses the concept of quadtree to partition the input image up to second level. Features are extracted from each partitioned image. Training data set of 3,600 samples is used to obtain consistent feature values. K-nearest neighbor classifier is used to classify the characters based on the feature values. The proposed method is tested on a data set of 1,200 samples, and at present, it shows an overall accuracy of 85.43 %.

Keywords Document image analysis · Optical character recognition (OCR) · Preprocessing · Feature extraction · Quadtree · Classifier · K-nearest neighbor

1 Introduction

Document image analysis is an emerging research area in the field of pattern recognition and image processing since many decades. One of the objectives of document image analysis is to recognize the text and graphics components in images of printed or handwritten documents and to extract the intended information [1]. The recognition of the text present in a document is achieved by

M. C. Padma (✉)
Department of Computer Science and Engineering, PES College of Engineering, Mandya, Karnataka 571401, India
e-mail: padmapes@gmail.com

S. Pasha
Department of Information Science and Engineering, PES College of Engineering, Mandya, Karnataka 571401, India
e-mail: saleempashapes@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_74,

optical character recognition (OCR), which converts scanned document images of machine-printed or handwritten text into a computer-editable format. In recent years, OCR has been greatly developed because of the prevalence of Internet and multimedia techniques.

In most of the real-life applications, majority of the documents received at various offices are filled by humans for specific purposes. In such scenarios, the documents may contain portions of printed text as well as handwritten text. Generally, the OCR system that is designed for printed documents cannot be used to process handwritten documents. Hence, there is a great demand to develop an OCR system that can process the documents containing handwritten text. Recognizing handwritten characters is very complex due to various factors like variation in writing style, mood of each individual, size of characters, quality of pen, aging of documents, quality of paper, color of ink. Because of these reasons, handwritten character recognition is still a challenging task and has much scope in the field of document image analysis. Handwritten character recognition has received more attention in academic and production fields. The recognition system can be either online or offline. Online handwritten characters are obtained by using cameras or by writing the characters on a sensitive surface such as digital tablet PCs. Offline handwritten characters are obtained from scanned images of handwritten text. In this paper, an attempt is made to develop a system to recognize offline handwritten characters. Recognizing handwritten characters is essential for various applications such as library automation, automatic postal address readers, automatic data entry from paper documents to computer, reading aid for the visually impaired. Development of an efficient and robust OCR system involves several stages such as preprocessing, feature extraction, and classification. However, in this paper, techniques for extracting the features from handwritten Kannada characters which could be further used in the development of an OCR system are presented.

The rest of the paper is organized as follows. Section 2 briefs about the previous work carried out on OCR systems. Section 3 gives brief introduction of the Kannada script. Necessary preprocessing steps are briefed out in Sect. 4. In Sect. 5, complete description of the proposed model is presented. Conclusion is given in Sect. 6.

2 Literature Survey

In this section, the literature relevant to the proposed research is highlighted. According to the literature survey, it is noticed that many OCR systems are developed for printed documents, but less work is reported for handwritten documents of Indian scripts. It is also observed from the literature that many successful attempts are reported for both printed and handwritten documents for other than Indian scripts. OCR systems for printed and handwritten numerals for Indian

scripts are also designed. Sufficient amount of work has been carried out for the development of OCR systems for printed documents of Kannada script.

Some considerable amount of work has been reported on handwritten documents having Kannada script. Sangame et al. [2] have proposed an OCR system for recognition of isolated handwritten Kannada vowels by extracting invariant moment features from zoned images, and k-nearest neighbor classifier was used for classification. Their work is limited only to vowels with an accuracy of 85.53 %. Also if the same features are applied for complete character set, the recognition rate may still go down. Thungamani et al. [3] have designed an OCR system for handwritten Kannada text recognition using features extracted from Zernike moments and using support vector machine as classifier. The feature extraction techniques proposed in [2] and [3] use global features which may not yield consistent feature values and hence result in less recognition rate. Niranjan et al. [4] have designed a system for recognition of handwritten Kannada characters using Fisher linear discriminant analysis. In paper [4], only unconstrained handwritten characters are considered for recognition. Manjunath Aradhya et al. [5] have proposed a probabilistic neural network–based approach for handwritten character recognition with an accuracy of 88.64 %. They have used Fourier transform and principal component analysis for feature extraction, but Fourier transform technique is computationally expensive for images containing large data sets. In [6], a handwritten Kannada (numerals and vowels) and English character recognition system based on spatial features is presented. They have considered only vowels and achieved an accuracy of 90.1 %. Murthy and Ramakrishnan et al. [7] have proposed a novel dexterous technique for fast and accurate recognition of the handwritten Kannada and Tamil characters. They have considered online characters, but all characters are not recognizable when the data are collected at the word level. An interesting work on handwritten Kannada Kagunita recognition was described by Leena et al. [8] using moment features. They have considered only compound characters known as Kagunita to derive moment features from directional and cut images. In [9], a handwritten Kannada word recognizer with unrestricted vocabulary using statistical dynamic space warping classifier was developed. They have considered online words for recognition. In general, it is observed from the above-discussed work that the OCR systems developed for handwritten documents containing Kannada script are limited to either vowels or characters or words and some works have considered online characters or words. The limitations observed in the literature have led to develop a suitable method for extracting efficient features from handwritten Kannada characters. Hence, in this paper, quadtree-based technique is proposed for extracting the features.

3 Kannada Script

Kannada is the official language of Karnataka, a south Indian state. Kannada characters were developed from Kadamba and Chalukya scripts, descendents of Brahmi. Sample handwritten Kannada characters are shown in Fig. 1.

Modern Kannada has 51 base characters called Varnamala [10]. There are 16 vowels and 35 consonants and also 10 different numerals. Each of these can modify a primary consonant to form a compound character. Thus, a compound character consists of consonant–vowel and consonant–consonant–vowel combinations. The set of all such combinations together with the base characters are known as aksharas.

4 Preprocessing

Preprocessing is an important step which transforms the raw image into a processed image that is suitable for better feature extraction. The transformation is achieved using several preprocessing steps such as skew detection and correction, binarization, crop to edge, noise removal, thinning [11, 12]. In this paper, skew detection and correction are performed using Fourier transform [13]. During crop to edge, extra space in the selected character is removed. A global thresholding approach is used for binarization, which converts a grayscale image into binary image. Median filtering is used for noise removal, which is a nonlinear operation often used in image processing to reduce 'salt and pepper' noise. A median filter is more effective when the goal is to simultaneously reduce noise and preserve edges. Thinning also known as skeletonization is performed to make the image crisper by reducing the binary-valued image regions to lines that approximate the skeletons of the region. All these preprocessing steps are implemented using built-in functions available in MATLAB [14]. Figure 2 shows the result after applying

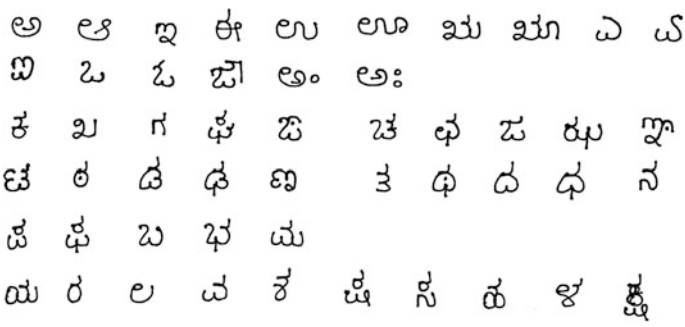


Fig. 1 Sample of handwritten Kannada characters

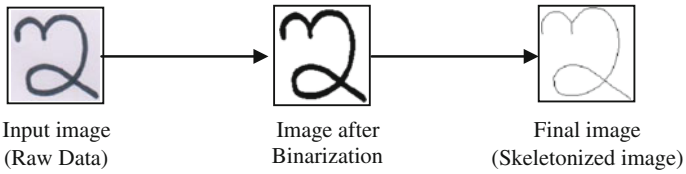


Fig. 2 Result of preprocessing steps shown for the character '2'

preprocessing steps for the character '2'. Thus, the preprocessed image is prepared ready for further processing such as feature extraction.

5 Proposed Model

The important stage of handwriting recognition scheme is the selection of good feature set, which is reasonably invariant with respect to shape variations caused by various writing styles. Different feature extraction methods are reported in [15] for different representations of the characters such as solid binary characters, character contours, thinned characters, or grayscale subimages of each individual character.

A quadtree is a tree data structure in which each internal node has exactly four subnodes [16]. Quadtree is most often used to partition a two-dimensional space by recursively subdividing it into four quadrants. Example of quadtree is shown in Fig. 3.

In this paper, the concept of quadtree is used to partition the input image, up to second level. Each partitioned image can be called zone. Different zoning approaches for character recognition are found in the literature. Sharma et al. [17] have proposed zone-based hybrid feature extraction model for handwritten English alphabets recognition using Euler number. Luciane Y. Hirabara et al. [18] have developed a dynamic zoning selection system for handwritten English character recognition. However, the major problem related to zoning mechanisms is to choose the best zoning scheme to solve the recognition problem of different classes of characters. Hence, in this paper, an attempt is made to propose quadtree-based technique to extract the features.

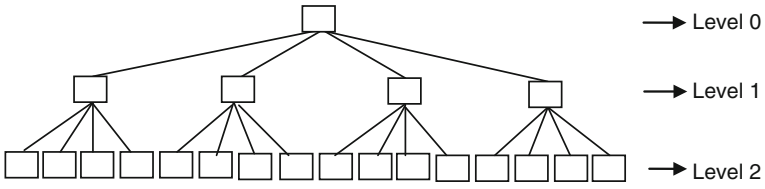


Fig. 3 Example of quadtree

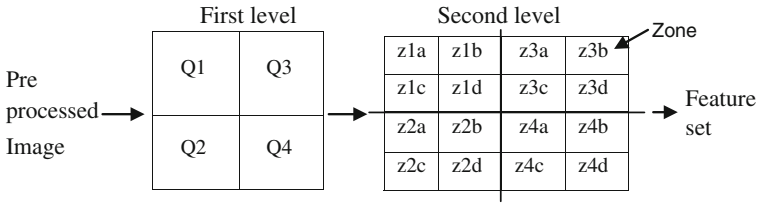


Fig. 4 Quadtree-based technique for feature extraction

5.1 Quadtree-Based Technique

In this technique, the preprocessed image is divided into four quadrants Q1, Q2, Q3, and Q4. Each quadrant is further divided into four quadrants/partitions, and each partition is called zone. For instance, a quadrant Q1 contains four zones such as Z1a, Z1b, Z1c, and Z1d as represented in Fig. 4. A set of 7 features which are explained in Sect. 5.2 are extracted from each zone. Average features of four zones Z1a, Z1b, Z1c, and Z1d are taken from quadrant Q1 to obtain 7 features. This process is repeated for remaining quadrants Q2, Q3, and Q4. Finally, 28 features are obtained from four quadrants.

5.2 Feature Extraction

Every character of Kannada script has distinct characteristic features due to the presence of different segments such as horizontal line, vertical line, left curve, right curve, upward curve, downward curve. The presence of such segments inspired us to extract these features from individual characters. Each zone of the partitioned input image may contain different segments such as horizontal line, vertical line, upward curve, left curve, downward curve, right curve. From the experimentation, it is decided to use a 3×3 mask that is applied on each zone to determine the presence of a particular segment type (horizontal line, vertical line, left curve, right curve, upward curve, downward curve). Then, the number of the presence of segment type is computed from each zone. The length of each segment type is also computed which is the total number of pixels present in that segment. These segment types are explained as features (Feature 1 to Feature 6) below.

Feature 1: **Horizontal line**: The occurrences of the components whose number of black pixels along a single row with the column number being changed are found, and these components are stored in the attribute horizontal line (black pixels having the value 0's correspond to object, and white pixels having value 1's correspond to background). The feature horizontal line is computed using the Eq. (1):

$$\text{Number of horizontal lines} \times \text{Total length of horizontal lines} \tag{1}$$

Feature 2: **Vertical line**: The occurrences of the components whose number of black pixels along a single column with the row number being changed are found, and these components are stored in the attribute vertical line. The feature vertical line is computed using the Eq. (2) below:

$$\text{Number of vertical lines} \times \text{Total length of vertical lines} \tag{2}$$

Feature 3: **Upward curve**: A component is said to have the shape of the upward curve, if the pixel values of the components are in the sequence (i, j), $(i-1, j+1)$, $(i-2, j+2)$, …,$(i-m, j+m)$, $(i-m+1, j+m+1)$, $(i-m+2, j+m+2)$, …, $(i-m+n, j+m+n)$, where $m = n$. This sequence is stored in the attribute upward curve and computed using the Eq. (3). Figure 5a shows the extraction of upward curve from each zone.

$$\text{Number of upward curves} \times \text{Total length of upward curves} \tag{3}$$

Feature 4: **Left curve**: A component is said to have the shape of the left curve, if the pixel values of the components are in the sequence (i, j), $(i+1, j-1)$, $(i+2, j-2)$, …, $(i+m, j-m)$, $(i+m+1, j-m+1)$, $(i+m+2, j-m+2)$, …, $(i+m+n, j-m+n)$, where $m = n$. This sequence is stored in the attribute left curve and computed using the Eq. (4). Figure 5b shows the extraction of left curve from each zone.

$$\text{Number of left curves} \times \text{Total length of left curves} \tag{4}$$

Feature 5: **Downward curve**: The approach followed in extracting upward curve is used to determine downward curve. The obtained sequence is stored in the attribute downward curve and computed using the Eq. (5).

$$\text{Number of downward curves} \times \text{Total length of downward curves} \tag{5}$$

Feature 6: **Right curve**: The approach followed in extracting left curve is used to determine right curve. The obtained sequence is stored in the attribute right curve and computed using the Eq. (6).

$$\text{Number of right curves} \times \text{Total length of right curves} \tag{6}$$

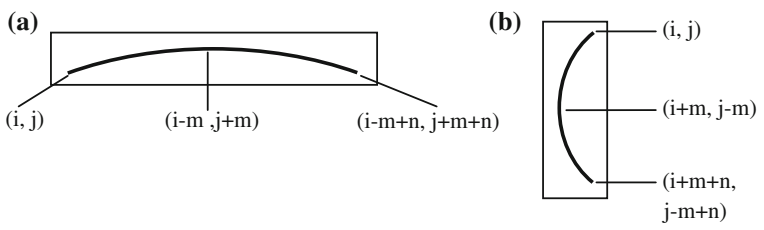


Fig. 5 Upward and left curves, **a** upward curve and **b** left curve

Feature 7: **Normalized area of zone**: This feature is computed for each zone with respect to skeleton of the image. It is needed because characters have different densities at particular zones. This feature is computed using the Eq. 7,

$$\text{Normalized area of zone} = \frac{\text{Number of pixels in each zone}}{\text{Total number of pixels in skeleton}} \tag{7}$$

The algorithm for quadtree-based feature extraction technique is explained below.

Algorithm: Quadtree-based feature extraction technique
Input: Image containing handwritten Kannada characters
Output: 28 Features from the entire image
Algorithm Begins

Step 1: The input image is divided into four quadrants Q1, Q2, Q3, and Q4.
Step 2: Each quadrant is further divided into four partitions called zones. 7 features are extracted from each zone. From four zones, a total of 28 features are extracted. Finally, average of 28 features is taken to obtain 7 features from each quadrant.
Step 3: Step 2 is repeated for all the four quadrants Q1, Q2, Q3, and Q4 to obtain 28 features from the entire image.

Algorithm Ends

5.3 Classification

Nearest neighbor classifier is an effective technique for classification problems in which the pattern classes exhibit a reasonably limited degree of variability. It is an instance-based learning or lazy learning, where the function is only approximated locally and all computation is deferred until classification. In *k*-nearest neighbor classifier, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k*-nearest neighbors (*k* is a positive integer, typically small). In this paper, *k*-nearest neighbor classifier is used for recognition of handwritten Kannada characters with different values of $k = 1$, 2, 3. Better accuracy is achieved for recognition of handwritten Kannada characters using k-nearest neighbor classifier with the value of $k = 3$. K-nearest neighbor classifier is used for classifying total samples of 1,200 handwritten Kannada characters.

5.4 Experimental Results and Discussion

The proposed model is implemented using MATLAB in Windows 7 platform. As there is no standard data set for handwritten Kannada characters, we have created our own data set for the experimentation. The handwritten Kannada characters (vowels and consonants) considered are independent of font size and style. These characters are created with no restriction on the pen, paper, ink color, ink flow, size, etc. We have collected samples of handwritten Kannada characters from different writers belonging to different age groups. The sampled images were scanned with a resolution of 300 dpi. A total of 3,600 training samples and 1,200 testing samples are used in the proposed model. The recognition of handwritten Kannada characters is achieved using k-nearest neighbor classifier with different values of $k = 1, 2, 3$. The performance of k-nearest neighbor classifier was better when the value of $k = 3$ is used. Recognition rate for different characters is given in Table 1.

Few misclassified characters obtained during experimentation are tabulated in Table 2. Misclassification occurs due to close resemblance of one character over the other. Because of these misclassified characters, we have achieved an overall accuracy of 85.43 %. The proposed method is compared with the existing methods [2, 5, 6], and it is given in Table 3.

Table 1 Recognition accuracy for handwritten Kannada characters

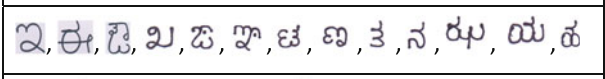
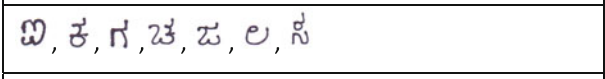
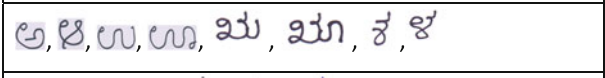
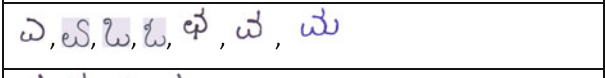
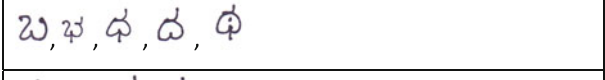
Train samples : 3600, Test samples: 1200	
Handwritten Kannada Characters	Accuracy in %
	94.00
	90.00
	90.00
	86.00
	84.00
	78.00
	76.00
Average Accuracy	85.43

Table 2 Few misclassified characters

Sl. No.	Characters confused with other characters	Sl. No.	Characters confused with other characters
1		7	
2		8	
3		9	
4		10	
5		11	
6		12	

Table 3 Comparison of the proposed method with the existing methods

Methods	Sample size	Feature extraction technique	Classifier technique	Accuracy (%)	Remarks
Method 1 [2]	1,625	Invariant moment features	k-nearest neighbor	85.53	Method is tested only on vowels
Method 2 [5]	5,000	Fourier transform and principal component analysis	Probabilistic neural network	68.89	Method is tested for both vowels and consonants. This method is computationally expensive for images containing large data sets
Method 3 [6]	1,400	Directional spatial features	k-nearest neighbor	90.1	Method is tested only on vowels
Proposed method	4,800	Quadtree-based technique	k-nearest neighbor	85.43	Proposed method is tested for both vowels and consonants. Better accuracy is obtained, when compared to existing methods

6 Conclusion

In this paper, a quadtree-based feature extraction technique for recognizing handwritten Kannada characters is presented. The method is simple and found to be effective as the features are extracted from the partitioned image. The experimental results illustrate the performance of the proposed method. Comparative

analysis also shows that the features used in the proposed method give better result compared to the existing methods. Our future work is to focus on handwritten Kannada compound characters.

References

1. Kasturi R, O' Gorman L, Govindaraju V (2002) Document image analysis: a primer. Sadhana 27(1):3–22
2. Sangame SK, Ramteke RJ, Benne R (2009) Recognition of isolated handwritten Kannada vowels. Adv Comput Res 1(2):52–55
3. Thungamani M, Ramakhanth Kumar P, Prasanna K, Rao SK (2011) Off-line handwritten Kannada text recognition using support vector machines and Zernike moments. Int J Comput Sci Network Secur 11(7):128–135
4. Niranjan SK, Kumar V, Hemanth Kumar G, Manjunath Aradhya VN (2008) FLD based unconstrained handwritten Kannada character recognition, 2nd international conference on future generation communication and networking symposia, pp 7–10
5. Manjunath Aradhya VN, Niranjan SK, Hemanth Kumar G (2010) Probabilistic neural network based approach for handwritten character recognition, IJCCT, vol 1, pp 9–13
6. Dhandra BV, Hangarge M, Mukarambi G (2010) Spatial features for handwritten Kannada and English character recognition, IJCA, special issue on recent trends in image processing and pattern recognition, pp 146–151
7. Murthy VN, Ramakrishnan AG (2011) Choice of classifiers in hierarchical recognition of online handwritten Kannada and Tamil aksharas. J Univ Comput Sci 17(1):94–106
8. Ragha LR, Sasikumar M (2011) Feature analysis for handwritten Kannada Kagunita recognition. Int J Comput Theor Eng 3(1):94–103
9. Kunwar R, Shashikiran K, Ramakrishnan AG (2010) Online handwritten Kannada word recognizer with unrestricted vocabulary, 12th international conference on frontiers in handwriting recognition, pp 611–616
10. Indira K, Selvi SS (2009) Kannada character recognition system: a review. InterJRI Sci Technol 1(2):31–42
11. Jomy J, Pramod KV, Kannan B (2011) Handwritten character recognition of south Indian scripts: a review, national conference on Indian language computing, Kochi, pp 1–6
12. Rahiman MA, Rajasree MS (2009) A detailed study and analysis of OCR research in south Indian scripts, international conference on advances in recent technologies in communication and computing, pp 31–38
13. Postl W (1986) Detection of liner oblique structure and skew scan in digitized documents. In: proceeding of international conference on pattern recognition, pp 687–689
14. Gonzalez RC, Woods RE, Eddins SL (2011) Digital image processing using Matlab, 2nd edn, McGraw Hill, New York
15. Trier OD, Jain AK, Taxt J (1996) Feature extraction methods for character recognition: a survey. Pattern Recogn 29(4):641–662
16. Finkel RA, Bentley JL (1974) Quad trees: a data structure for retrieval on composite keys, Acta Informatica, vol 4(1). Verlag, pp 1–9
17. Sharma OP, Ghose MK, Shah KB (2012) An improved zone based hybrid feature extraction model for handwritten alphabets recognition using euler number, vol 2, ISSN: 2231–2307, IJSCE
18. Hirabara LY, Aires SBK, Freitas COA, Britto AS Jr, Sabourin R (2011) Dynamic zoning selection for handwritten character recognition, Verlag, Berlin Heidelberg, pp 507–514

Effect of Contribution Links on Choosing Hard Goals in GORE Using AHP and TOPSIS

S. Vinay, Shridhar Aithal and G. Sudhakara

Abstract Decision support system in requirements engineering plays an important role in software development life cycle. The relationship between functional and non-functional requirements is the key in resolving conflicts in requirements gathering phase. In this paper, we discuss the effect of nonlinearity rating while converting the contribution links to quantitative values. We use our goal-oriented requirements engineering (GORE) method to identify the soft goals. Analytic hierarchy process (AHP) is then used to prioritize the soft goals. The output of AHP is used as input to technique for order of preference by similarity to ideal solution (TOPSIS) which produces a metric which decides the best alternative among the candidates. We illustrate the process and analyze the nonlinearity effect for meeting scheduler application.

Keywords Goal-oriented requirements engineering · Analytic hierarchy process · TOPSIS · Nonlinearity · Hard goals · Soft goals

1 Introduction

It is well acknowledged in software engineering that while functional requirements are important, eliciting and capturing the non-functional requirements (NFR) during the requirements engineering phase becomes even more important [1].

S. Vinay (✉)
PESITM, Shivamogga, India
e-mail: vinaymanyan@gmail.com

S. Aithal
TAPMI, Manipal, India
e-mail: saithal@tapmi.edu.in

G. Sudhakara
MIT, Manipal, India
e-mail: sudhakaraadiga@yahoo.co.in

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_75,

Goal-oriented requirements engineering (GORE) approaches make a good attempt to address the essential quality characteristics which are commonly known as non-functional requirements [2, 3]. While it is important to gather the NFR, it is also equally important how each of NFR is affecting the overall system goals. In this regard, decision support systems play a major roll and in this context, one needs to rely on the tools in this domain.

Gunther Rahe in his paper [4] highlights the importance of decision support system (DSS) in software engineering. Decisions are the driving engines for all stages of software development and evolution. Decisions can be related to methods, tools, and techniques. Decisions are aimed at answering the questions 'How?' 'How good?' 'When?' 'Why?' and 'Where?' The objective is to have a sound methodology which provides rationale for the decision arrived at. The importance of decision-making techniques is also addressed in Refs. [5–7]. The importance of stakeholders in decision making is done in Price and Cybulski [8]. Architecture decision making is closely linked to requirements engineering and the aspects related to this are addressed in Refs. [9] and [10]. A survey of various requirements prioritization techniques is undertaken in Hasan et al. [11]. Decision making in software engineering is extremely challenging because of a dynamically changing environment, conflicting stakeholder objectives, constraints, coupled with a high degree of uncertainty, and vagueness of the available information.

The analytic hierarchy process (AHP) is based on the experience gained by its developer, Saaty [12], while directing research projects in the US Arms Control and Disarmament Agency. The simplicity and power of the AHP has led to its widespread use across multiple domains in every part of the world.

Technique for order of preference by similarity to ideal solution (TOPSIS) method is presented in Chen and Hwang [13], with reference to Hwang and Yoon [14]. The basic principle is that the chosen alternative should have the shortest distance from the ideal solution and the farthest distance from the negative-ideal solution. Combining our GORE method with AHP and TOPSIS in decision making provides adequate rationale for the decision arrived at.

In our earlier work in Vinay et al. [15], we had proposed an approach to combine GORE with AHP by considering the case study of choosing a CMS. For the same case study, we combined GORE with AHP and TOPSIS in Vinay et al. [16]. In this paper, we apply the method for a meeting scheduler where the best alternative which fulfills the soft goals is chosen. We use our goal-oriented requirements engineering (GORE) method to identify the hard goals (functional requirements) and soft goals (non-functional requirements). Analytic hierarchy process (AHP) is then used to prioritize the soft goals. The output of GORE and AHP is combined as input to TOPSIS in order to produce a metric which decides the best alternative among the three. In Sects. 2 and 3, we discuss our proposed approach in brief and highlight how our GORE approach is used for identifying soft goals (non-functional requirements), their contribution links to each of the alternative. Section 4 discusses how AHP is used to prioritize the soft goals. Combining the output of GORE and TOPSIS is discussed in Sect. 5. The results and the effectiveness of the proposed method are discussed in Sect. 6.

Table 1 Steps involved in our approach

Step	Description	Method
1	Identify the hard goals	Apply our GORE method to identify hard goals with AND/OR decompositions
2	Identify soft goals (non-functional requirements)	Apply our GORE method to identify soft goals with AND/OR decompositions
3	Identify contribution links	Apply our GORE method to identify contribution links between soft goals and each alternative
4	Calculate priority of soft goals	Use AHP to calculate priority of soft goals. If it is a multi-level hierarchy, we need to proceed from root to leaf level soft goals in calculating priorities of soft goals using AHP. Local weights are calculated among same level goals using AHP. Global weights are the product of all local weights proceeding from leaf to root
5	Evaluation of each alternative using TOPSIS	Convert contribution link to numeric values. Apply these numeric values for TOPSIS method while comparing each alternative for the identified criteria
6	Ranking of alternatives	The alternatives are ranked in terms of increasing order of their values

2 Proposed Approach

The major steps involved in our proposed approach are shown in Table 1:

3 Identifying Soft Goals and Contribution Links Using GORE Methodology

Some of the prominent GORE methods include: keep all objectives satisfied (KAOS) method [17, 18], Tropos [19], NFR framework [2, 20]. Our GORE method retains the generic features of other techniques in extracting hard and soft goals, decomposing the goals and refining them. In addition to these, we give more emphasis on identifying the contribution links. Many methods have taken this approach of representation. In our approach, we propose the following links: ++, +, −−, −, ?. In our GORE method, these contribution links exist from leaf level hard goals to leaf level soft goals. The symbols and their meanings are shown in first two columns Table 2. The significance of the values mentioned in the third column is discussed in Sect. 5. The method of identifying these contribution links which could be either quantitative or qualitative can be found in our earlier work [21].

Meeting scheduler case study is chosen for analyzing our method, and the goal graph for the same is shown in Fig. 1. Hard goals are represented by ovals and soft goals are represented by cloud symbol. The contribution links from hard goals to soft goals are also shown in Fig. 1. Hard goals and soft goals are identified as part of GORE method. Two of the soft goals are further decomposed into sub-goals.

Table 2 Meaning of contribution links

Symbol	Meaning	Value
++	A hard goal requirement fully supports the soft goal	9
+	A hard goal requirement partially supports the soft goal	7
?	Indicates neutral or no dependency between hard and soft goal	5
−	A hard goal requirement minimally supports the soft goal	3
−−	A hard goal requirement is not supported by the soft goal	1

For example, good-quality schedule has 2 sub-goals namely minimal conflicts and good participation.

The number of alternatives which can be derived from the goal graph in Fig. 1 which fulfills the root goal schedule meeting is 6. They are mentioned in Table 3. The task is to choose an alternative which fulfills the soft goal criteria indicated by contribution links to hard goals.

4 Prioritizing Soft Goals Using AHP

We apply step 4 of our approach which involves prioritizing the soft goals. We make use of AHPs pair-wise comparison technique for this purpose. In case of a multi-level hierarchy, this is done by applying AHP starting from root level goals to leaf level goals. Local weights are calculated among same level goals using

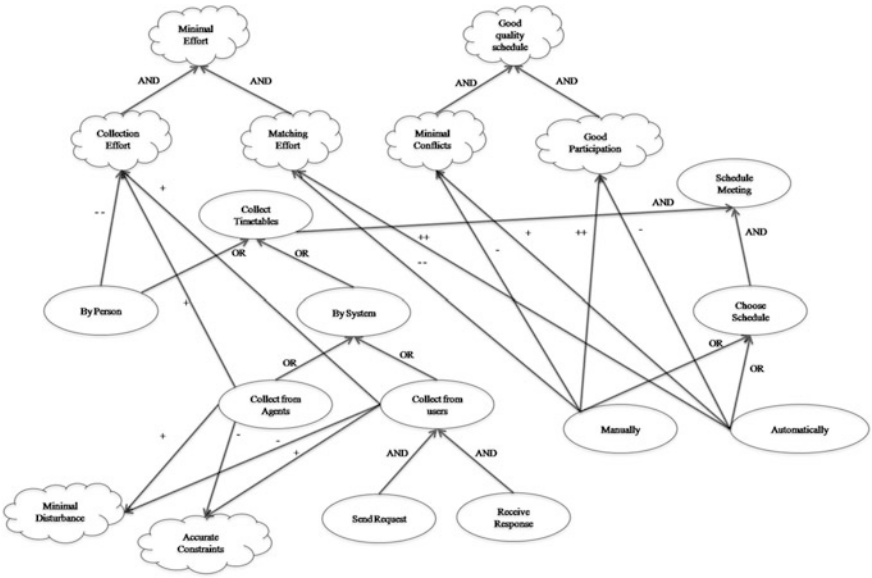


Fig. 1 Meeting scheduler goal graph

Table 3 Alternatives for meeting the root goal requirement

Alternatives	Description
A1	Collect timetables by person—choose schedule manually
A2	Collect timetables by system (agents)—choose schedule manually
A3	Collect timetables by system (users)—choose schedule manually
A4	Collect timetables by person—choose schedule automatically
A5	Collect timetables by system (agents)—choose schedule automatically
A6	Collect timetables by system (users)—choose schedule automatically

Table 4 Calculation of local and global weights

Soft goal (L1)	Local weight	Sub-goal (L2)	Local weight	Global weight
1. Minimal effort	0.23	1.1 Collection effort	0.45	0.1035
		1.2 Matching effort	0.55	0.1265
2. Good-quality schedule	0.56	2.1 Minimal conflicts	0.36	0.2016
		2.2 Good participation	0.64	0.3584
3. Minimal disturbances	0.11			0.11
4. Accurate constraints	0.10			0.10

AHP. Global weights are the product of all local weights proceeding from leaf to root. The result of this step is shown in Table 4.

5 Combining GORE and TOPSIS

The basic steps involved in combining GORE with TOPSIS and AHP involves the following.

(i) Convert contribution link identified previously to numerical values: We make use of a step proposed in quality function deployment (QFD) technique for this purpose. Quality function deployment (QFD) is a method to transform user demands into design quality, to deploy the functions forming quality, and to deploy methods for achieving the design quality into subsystems and component parts, and ultimately to specific elements of the manufacturing process [22]. Many QFD techniques [23–25] employ a 4-point or 5-point scale to convert the qualitative links to quantitative value. To our approach, we consider a 4-point scale for converting the contribution links to a numerical value. Column three of Table 2 specifies the value which we have chosen in our approach. For example, a ++ link from manually to good participation is given a number 9.

Table 5 Input to TOPSIS

Soft goals identifier/alternative	1.1	1.2	2.1	2.2	3	4
A1	1	1	3	9	5	5
A2	7	1	3	9	7	3
A3	7	1	3	9	3	7
A4	1	9	7	3	5	5
A5	7	9	7	3	7	3
A6	7	9	7	3	3	7

(ii) Apply TOPSIS technique: The six alternatives, the numeric values of contribution links, and the 6 prioritized soft goals are given as input to TOPSIS technique. TOPSIS procedure has been explained in Chen and Hwang [13]. The ideal solution is composed of all best values attainable of criteria, whereas negative-ideal solution is made up of all worst values attainable of criteria. The chosen alternative should be as close to the ideal solution as possible and as far from the negative-ideal solution as possible. The major steps involve:

(1) Calculate the normalized decision matrix.
(2) Calculate the weighted normalized decision matrix.
(3) Determine the ideal and negative-ideal solution
(4) Calculate the separation measures, using the n-dimensional Euclidean distance. The separation of each alternative from the ideal solution and the separation from the negative-ideal solution are calculated.
(5) Calculate the relative closeness to the ideal solution.
(6) Rank the preference order.

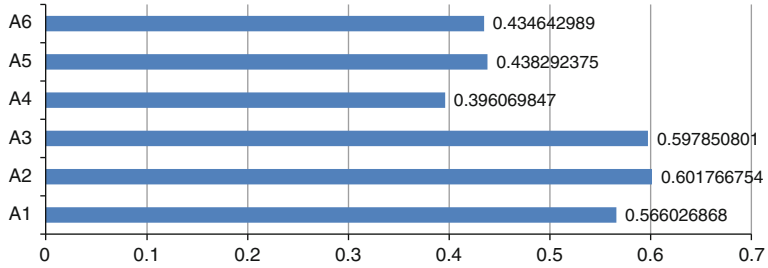
The values given as input to TOPSIS are shown in Table 5.

6 Results and Discussion

The output of TOPSIS is shown in Graph 1. It shows that alternative A2 (collect timetables by system (agents)—choose schedule manually) has the highest value followed by A3 and A1. One can infer from the graph that A4–A6 are least preferred which implies choosing a schedule manually is better than doing it automatically.

The steps proposed for integrating AHP and TOPSIS into the proposed method have been illustrated. Here are the following observations about the results obtained. The success of this approach depends on the following factors:

(i). Clear identification of hard goals and soft goals.
(ii). The identification of qualitative and quantitative criteria is an important step in the process since they directly contribute to the identification of contribution links. These contribution links play a major role in choosing the best alternative.



Graph. 1 Output of TOPSIS

(iii). Appropriate stakeholders' needs to be identified in prioritizing soft goals.

AHP is a technique for computing priorities which is widely used in many domains. The data entered by a stakeholder are checked for consistency by measuring Consistency Index or Ratio (CI/CR). If CI $>$ 0.1, then it implies that the priorities given by a stakeholder are not consistent. In the existing GORE literature, there exists technique which makes use of formal techniques [18] in choosing the best alternative. They make use of temporal logic and label propagation algorithms. Our approach differs in adopting a quantitative way of evaluating the alternative using AHP. We believe that a quantitative approach will enable all stakeholders to play an active role in decision making.

One of the existing methods which make use of a quantitative approach for decision support system in GORE is attributed goal-oriented requirements analysis (AGORA) [26]. One of the future works suggested by the authors specifies the use of AHP to assign the values subjectively. We have used AHP and TOPSIS which are widely used for decision support system in software development.

We have analyzed the effect on the output for both linear and nonlinear scale when converting contribution links to a numerical value. The proposed approach gives the correct output for linear scale only.

6.1 Validation of the Proposed Method

The approach to validate the proposed method of combining AHP, TOPSIS, and GORE is done by considering the opinion of the experts. Problem description of all systems has been given to 6 subject experts who have to recommend the best alternative to be chosen among the three. Experts chosen are from industry with a minimum experience of 5 years. The participants are required to provide rationale for the decisions which they arrive at. The experts will use their experience in deciding the alternative and are not aware about our proposed method. The comparison of the decision taken by experts with the results obtained by our approach is shown in Table 6.

Table 6 Comparison of our method with experts

Subject	Ranking of alternatives (Top 3)	Chosen alternative
Our method	A2, A3, A1	A2
Expert 1	A2, A3, A1	A2
Expert 2	A3, A2, A1	A3
Expert 3	A2, A3, A1	A2
Expert 4	A2, A3, A1	A2
Expert 5	A1, A3, A2	A1
Expert 6	A2, A1, A3	A2

Results shown in Table 5 clearly convey that A2 is chosen as the best alternative by four experts as well as our approach. Three experts have the same ranking of alternatives as our method. The experts have arrived at a decision based on their experience, importance of soft goals, and in consultation with other senior colleagues. We can clearly infer that our method results are in agreement with the experts. The advantage of our method is that it provides a metric and rationale for the decision arrived at using AHP and TOPSIS unlike experts who rely more on experience and intuition.

7 Conclusion and Future Work

In this paper, we have made an attempt to integrate GORE method with AHP and TOPSIS in arriving at a decision. The method gives consistent results which depend on the quality of the goals identified and their priority. The same process might yield a different outcome if the priorities on the soft goals change. We plan to come up with suitable metrics for comparison with other approaches. A tool to support the entire process is also planned for the future.

References

1. Doerr J, Kerkow D, Koenig T, Olsson T, Suzuki T (1995) Non-functional requirements in industry: three case studies adopting an experience-based NFR method. In: Proceedings of the 2005 13th IEEE international conference on requirements engineering
2. Mylopoulos J, Chung L, Nixon B (1992) Representing and using nonfunctional requirements: a process-oriented approach. IEEE Trans Software Eng 18(6):483–497
3. van Lamsweerde A (2001) Goal-oriented requirements engineering: a guided tour. In: Proceedings of the 5th IEEE international symposium on requirements engineering. IEEE computer society, Washington
4. Ruhe G (2003) Software engineering decision support and empirical investigations—a proposed marriage. Workshop Empirical Stud Softw Eng (WSESE), 2: 25–34
5. Omasreiter H (2007) Balanced decision making in software engineering: general thoughts and a concrete example from industry. In: First international workshop on the economics of software and computation (ESC), 2007

6. Easterbrook S, Singer J, Storey M-A, Damian D (2008) Selecting empirical methods for software engineering research. In: Guide to advanced empirical software engineering, Section III, 285–311
7. Hannay JE, Sjøberg DIK, Dyba T (2007) A systematic review of theory use in software engineering experiments. IEEE Trans Softw Eng 33(2):87–107
8. Price J, Cybulski J (2006) The importance of IS stakeholder perspectives and perceptions to requirements negotiation. In: AWR. Adelaide, Australia
9. Ivanović A, America P (2010) Information needed for architecture decision making. In: Proceedings of the ICSE workshop on product line approaches in software engineering
10. Lakshminarayanan V, Liu WQ, Chen CL, Easterbrook S, Perry DE (2006) Software architects in practice: handling requirements. In: Proceedings of the conference of the center for advanced studies on collaborative research (CASCON)
11. Hasan MS, Mahmood AA, Alam MJ, Hasan SN, Rahman F (2010) An evaluation of software requirement prioritization techniques. Int J Comp Sci Inf Security (IJCSIS), 8(9)
12. Saaty TL (2008) Decision making with the analytic hierarchy process. Int J Services Sci, 1(1),83–98
13. Chen SJ, Hwang CL (1992) Fuzzy multiple attribute decision making: methods and applications. Springer, Berlin
14. Hwang CL, Yoon K (1981) Multiple attribute decision making. in: lecture notes in economics and mathematical systems, 186. Springer, Berlin
15. Vinay S, Aithal S, Sudhakara A (2012) A quantitative approach using goal-oriented requirements engineering methodology and analytic hierarchy process in selecting the best alternative. In: Proceedings of international conference on advances in computing, advances in intelligent systems and computing, 174, pp 441–454
16. Vinay S, Aithal S, Sudhakara A (2012) Integrating TOPSIS and AHP into GORE decision support system. Int J Comp Appl, 56(17): 46–53
17. van Lamsweerde A (2004) Goal-oriented requirements engineering: a roundtrip from research to practice. In: Proceedings of RE '04
18. van Lamsweerde A (2009) Reasoning about alternative requirements options, in conceptual modeling: foundations and applications. In: Borgida A, Chaudhri V, Giorgini P, Yu E (eds) Essays in Honor of John Mylopoulos, Springer LNCS 5600, 380–397
19. Castro J, Kolp M, Mylopoulos J (2002) Towards requirements-driven information systems engineering: the Tropos project. Inf Syst, 27(6): 365–389
20. Kaiya H et al. (2004) Identifying stakeholders and their preferences about NFR by comparing use case diagrams of several existing systems. 12th IEEE international requirements engineering conference (RE'04)
21. Vinay S, Aithal S, Sudhakar G (2011) A goal-oriented requirements engineering method for analysing conflicts, ICCANA, Nitte, Jan 2011
22. Akao Y (1994) Development history of quality function deployment. The customer driven approach to quality planning and deployment. Minato, Tokyo 107 Japan: Asian Productivity Organization, pp 339–351
23. Herzwurm G et al. (2003) QFD for customer focused requirements engineering. In: 11th IEEE international requirements engineering conference, 2003
24. Hierholzer A, Herzwurm G, Schlang H (2003) Applying QFD for software process improvement at SAP AG, Walldorf, Germany. In: Proceedings of the third workshop on software quality, ACM, 2003
25. de Felice F, Petrillo A (2010) A multiple choice decision analysis: an integrated QFD: AHP model for the assessment of customer needs. Int J Eng Sci Technol 2(9):25–38
26. Kaiya H, Nagano Horai H, Saeki M (2002) AGORA: attributed goal-oriented requirements analysis method. In: International conference on requirements engineering, 2002 proceedings

Transformation of Artistic Form Text to Linear Form Text for OCR Systems Using Radon Transform

Vishwanath C. Kagawade, C. S. Vijayashree and T. Vasudev

Abstract The existing optical character readers (OCRs) are capable of reading linear-form text and have limitations to read artistic- and nonlinear-form text. This paper presents a technique to transform printed English artistic-form text to linear-form text in order to make an OCR to read the text. The technique starts with artistic-form text as input and transforms the same to linear form. First, the characters in artistic text are segmented and extracted using connected component analysis technique. Due to the intrinsic nature of artistic text, the extracted characters exhibit skew. In the next stage, such implicit skew in extracted characters is detected using Radon transform and corrected. Further, skew corrected characters are concatenated to put in linear form. Experimental results of the proposed method show an average 85 % of readability efficiency by OCR.

Keywords Artistic text · Linear text · OCR · Connected component analysis · Canny edge detector · Skew detection · Radon transform

1 Introduction

A significant area in the field of digital image processing is document image analysis. Document analysis is very important in applications such as document identification/recognition, language identification, automatic reading from

V. C. Kagawade (✉) · C. S. Vijayashree · T. Vasudev
P.E.T Research Centre, PES College of Engineering, Mandya, Karnataka 571401, India
e-mail: Vishwanath.1312@gmail.com

C. S. Vijayashree
e-mail: cs.vijayashree@yahoo.com

T. Vasudev
e-mail: banglivasu@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_76,

document. Many researchers are working on different problems on document images starting from image acquisition to image understanding [1, 2]. Document image analysis (DIA) and processing activities can be divided into pre-processing, segmentation, script identification, page layout analysis (PLA) and classification, character recognition, etc. [3], and these have led to many vibrant research problems [4]. The results of the research on the above problems are gradually converging towards the generic solutions to major issues in DIA.

In spite of considerable research work in the area of DIA, a major issue that is not sufficiently addressed is reading or extracting the contents of the text that appear in artistic form in a document [1]. Many documents, such as certificates, marks cards, sign boards, logos, have artistic text. In addition, many official seals/stampings used for document authentication purposes are also artistic in nature. The contents of such artistic text definitely have some valuable information that has to be processed. Most of the graduation certificates issued by the universities contain the name of the university in artistic form. If such documents have to be processed by an optical character reader (OCR), then proper pre-processing is required to make that text readable by OCR as the available OCRs are capable of reading only linear-form text. Few such artistic texts in documents are text appearing in triangular form, arc form, circular form, wave form, shadow form, telescopic form, etc. Figure 1 shows few such examples of artistic-form text. The contents of such text normally conveys the identity information such as company's name, type of document, brand name, event name, type of document, , which is the main source for classification of the document. Therefore, text conversion from artistic-form text to linear-form text should be done before processing the document with OCR. Hence, it is necessary to transform artistic-form text into linear-form text and make it suitable for reading by an OCR. This motivated us to make an attempt to transform artistic-form text to linear-form text.

In this research work, we propose a methodology to transform artistic-form text to linear-form text suitable for OCR processing. The model uses connected component analysis technique for character extraction from document image having artistic-form text. The extracted characters have implicit skew [1] due to the intrinsic nature of artistic text, and the OCRs fail considerably to read such skewed characters. The Radon transform is used for detection of skew in characters, and the skew is corrected. Finally, all the skew corrected characters are concatenated to present the text in linear form. The input image is assumed to be free from noise and contains only artistic text.

Fig. 1 Samples of artistic-form texts

The rest of the paper is organized as follows: The sequence of stages in the proposed work along with character extraction using connected component analysis is discussed in Sect. 2. The theory of Radon transform along with detection and correction of skew in characters is explained in Sect. 3. Experimental results are illustrated in Sect. 4. Conclusion on the work is given in Sect. 5.

2 Stages in the proposed Work

The block diagram shown in Fig. 2 indicates the sequence of different processing stages performed in the proposed system.

The printed artistic text document image obtained at 300-dpi resolution is taken as input to the system. Character segmentation from artistic text is made through a series of image processing operations. To begin with, the image is normalized to 340 × 340 pixels and image is binarized to black and white, in order to make invariant to image size and colour. Then, the characters are extracted from the image through connected component analysis. Labelling of characters is made using 8-connectivity logic [5]. All the objects containing pixel values less than the threshold value are treated as noise and are to be removed. Using the labels, the connected components (characters) are extracted. The extracted characters are then resized to 45 × 45 pixels.

The steps for Text Extraction from document image are given below in the form of an algorithm:

Input: Scanned image document containing artistic text
Output: A set of character images extracted from input image

1. Read the image
2. Convert input image to grey scale and grey scale to black and white image
3. Scan the image document from left to right to label characters
4. Label and count connected components that form the candidate region
5. Remove all objects containing less than threshold pixel value
6. Crop the character using candidate region
7. Repeat step 6 until all characters are extracted from the input document image.

Once all the labelled components are extracted, it is noticed that some additional blocks of residues are produced due to the limitation in cropping procedure. Such residues are addressed as noises in this work and are to be removed. It is evident from the experimental results that the connected component regions of

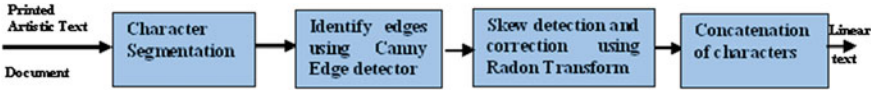


Fig. 2 Block diagram for skew detection and correction using Radon transform

Fig. 3 Sample input image for connected component analysis technique

Fig. 4 Extracted characters using connected component

Fig. 5 Extracted characters after noise is removed

such noises have less than 20 pixels (width-to-height ratio), that is, connected component region Area < connected components regions MaxArea/20. Considering 20 pixels as the threshold, the regions that are less than 20 pixels are removed from the resultant image.

The result of the implemented methodology is shown through an artistic-form text shown in Fig. 3.

Characters are extracted using connected component analysis technique, and it has been observed that once the character has been extracted, additional noise may appear along with characters as shown in Fig. 4. Using the defined threshold value, such noises are removed and the output after removing noises is shown in Fig. 5.

It is noticed that the extracted characters are not free from skew. Since the OCRs fail to read such skewed characters, the skew in characters is to be corrected. The skew in characters is detected using Radon transform and is explained in next section.

3 Radon Transform

The Radon transform is a well-known fundamental tool used for wide range of imaging applications such as radar imaging, geophysical imaging, nondestructive testing and medical imaging [6, 7]. The Radon transform has been represented in many different forms and notations [8]. The Radon transform for a 2-dimensional application is:

$$R(\theta, \rho) = \iint_D f(x, y)\, \delta(\rho - x\, \cos\theta - y\, \sin\theta)\, \mathrm{d}x\mathrm{d}y$$

where, D denotes the entire image plane, $\delta()$ is the Dirac delta function, and $\rho = x\, \cos\theta + y\, \sin\theta$.

The Radon transform is very suitable for character skew detection. The reason is that there is a large variation in the angle of alignment of character in the

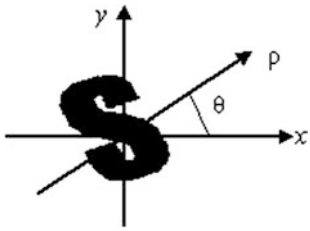


Fig. 6 Two parameters (ρ, θ) determining the position of the line in (x, y) co-ordinate system

Fig. 7 Skew corrected characters

artistic-form text. By applying the Radon transform to character, a mapping is established between the domain determined by the co-ordinate system (x, y) and the Radon domain (ρ, θ), that is, the distance and angle of alignment (ρ, θ) are determined as shown in Fig. 6. The maximum Radon transform value over all angles is found. Value of maximum angle is the skew angle. Implementation of Radon Transform is shown in Fig. 7.

In the last stage, the character is rotated using skew angle for skew correction to the base line and all the skew corrected characters are unified to form the linear-form text.

The output of skew detection and correction on the characters extracted in previous section is shown in Fig. 7.

The complete steps for transformation of artistic-form text to linear-form text are given below in the form of algorithm:

Input: Printed document image containing artistic-form text
Output: Image containing linear-form text

1. Extract the characters from the document image using connected component technique
2. Perform region-based heuristic filtering to remove noise in extracted character
3. Identify the edges using Canny edge detector
4. Identify the distance and angle of alignment using Radon transform
5. Find the maximum Radon transform value over all angles
6. Rotate the image with maximum angle
7. Repeat step 2 to step 6 for all the extracted characters
8. Merge all the output characters of step 7 to get resultant linear-form text.

The proposed method is efficient and simple, and it is suitable for most of font's sizes and shapes of English texts. However, Radon transform peaks may not be clearly marked due to low pixel density distribution of the domain, that is, in low-density areas. In the proposed method, the problem appears in the case of characters with nonuniform pixel density distribution. In such cases, character skew angle may not be estimated properly.

4 Experimental Results

The proposed methods have been implemented in the MATLAB R2009a. The different documents from newspapers, journals, text books, magazines, advertisement articles, pamphlets, and the artistic text document created from paint/word are considered. The experiments have been conducted on more than 150 image samples. Figure 8 shows some sample input texts and corresponding results. Within each figure, first row indicates the input artistic-form text, second row indicates extracted characters using connected component, and third row indicates linear-form text.

Experimental results illustrated from Fig. 8 indicate that the proposed approach considerably transforms artistic-form text to linear-form text better suitable for OCRs. The experiments are conducted using MATLAB R2009a as programming tool. The transformation model on characters increases the readability of OCR. Analysis of readability by an OCR before transformation, after character extraction and after skew correction is performed with respect to English text using the OCR "Readiris Pro 9" (http://www.irislink.com). Few instances from the worst case to the best case are considered, and readability by OCR at three stages is shown in Table 1.

Table 1 indicates that when an artistic-form text is given as input to OCR, it recognizes the text as picture rather than text and retains it in the picture form and the readability of text by OCR is obviously 0 %. This demands that most of OCRs do require linear text for reading. The extracted characters using connected component analysis technique introduces readability for OCR to some extent as show in Table 2, but still many input samples were recognized as pictures and broken characters because of nonlinear form text (skew). Table 3 shows that OCR

Fig. 8 Artistic-form text to linear-form text

Table 1 OCR's readability of artistic-form text

Input artistic form text	OCR Recognition	Readability
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture
		0.0 % Recognized as picture

readability considerably increases as character skew is corrected through Radon transform. Table 4 indicates average readability of OCR is 0–55 % in case of extracted characters and also shows that the Radon transform-based skew detection and correction in extracted characters introduces readability for OCR up to 85 %. The experimental results are illustrated in Fig. 9.

Table 4 indicates that average readability of OCR of artistic form text is 0 %, indicating that OCRs are not capable of reading text other than linear form. After extracting the characters and placing the characters linearly without skew correction improve the readability. Once the skew is corrected in the extracted characters, OCR performance enhances readability at an average of 85 %.

Table 2 OCR's readability after character extraction use CC (prior to skew correction)

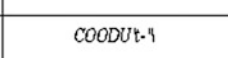
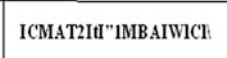
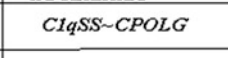
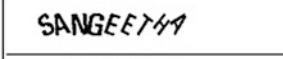
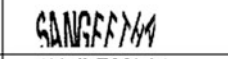
Input text	OCR Recognition	Readability
		0.0 %
		0.0 %
		57.14 % (4/7)
		0.0 %
		46.67 %(7/15)
		0.0 %
		72.7% (8/11)
		0.0 %
		0.0 %
		0.0 %

Table 3 OCR's readability after character skew correction

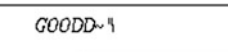
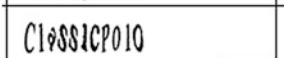
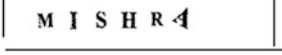
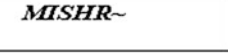
Input text	OCR Recognition	Readability
	SUCCESS	7 / 7 → 100 %
	CRII-FTF41R	6/ 9 → 66.6 %
		5 / 7 → 71.4 %
	AAI pcare6rto	9 / 11 → 81.8 %
	KMI\T10"1"1MBJIIMCI\	11 / 15 → 73.3 %
	SITESSAIE	8 / 9 → 88.9 %
		10 /11 → 90.9 %
	C>A NG EETHA	8 / 9 → 88.9 %
	AIVNI\IERSI\R'(	8 / 11 → 72.7 %
		5 / 6 → 83.3 %

Table 4 Readability analysis

Input text	Readability range (%)	Average readability (%)
Artistic-form text	0.00	0.00
Extracted character text with skew	00–57	5.70
After character skew correction	67–100	85

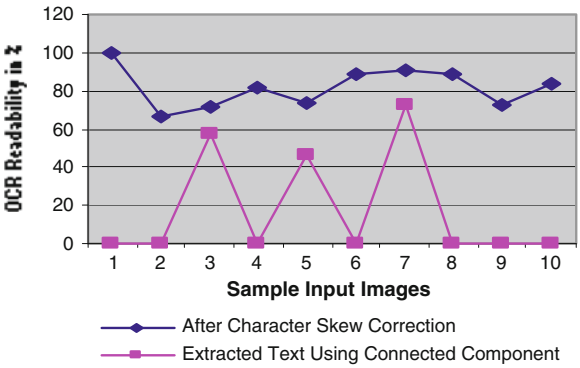


Fig. 9 OCR readability analysis for different image samples

5 Conclusion

In this paper, a new technique is proposed to transform artistic-form text to linear-form text of machine printed English text document. The paper describes the image processing stages used such as binarization, normalization, noise reduction, character segmentation and finding out the unknown skew angle of the character. The implementation of the proposed methodology exhibits an average readability efficiency of 85 % by OCR. The proposed technique is not suitable when document text appears in circle, circle with multiple lines of linear text, semi-arc with linear text, triangular form, and text with complex background. There is scope in this approach to investigate a new transformation model and minimize character tilt at the transformation level.

References

1. Vasudev T, Hemanthakumar GH, Nagabhushan P (2007) Transformation of arc-form-text linear-form-text- suitable for OCR. Sci Direct, Pattern Recog Lett 28(2007):2343–2351
2. Pal U, Mitra M, Choudari BB (2001) Multi-skew detection of Indian script documents. In: Proceedings international conference on document analysis and recognition (ICDAR 2001)
3. Vasudev T, Hemanthakumar GH, Nagabhushan P (2005) Segmentation of characters in an arc-form text. In : Proceedings 7th international conference on cognitive systems (ICCS 2005), CD version, India

4. O'Gorman L, Kasturi R (1998) Executive briefing: document image analysis. IEEE Computer Society Press
5. Kapoor R, Bagai D, Kamal TS (2004) A new algorithm for skew detection and correction. Sci Direct, Pattern Recogn Lett 25(2004):1215–1229
6. Nagabhushan P, Anagadi SA et al. (2007) Geometric model and projection based algorithms for Tilt correction and extraction of ascenders/descenders for cursive word recognition, IEE-ICSCN, pp 488–491
7. Velu CM, Vivekandan P (2010) Automated letter sorting for Indian postal address recognition system based on PIN codes. J Internet Inf Syst 1(1):6–15
8. Jain AK (2004) Fundamentals of digital image processing, Eastern Economy Edition, PHI Publication

Boolean Arithmetic Polynomials and Equivalence Checking

Rajkumar Vedam

Abstract Boolean arithmetic polynomials are used for the combinational equivalence checking problem in this tutorial paper. Boolean arithmetic polynomial multiplication methods are presented. Several methods for equivalence checking are discussed, including nonconventional methods.

Keywords Boolean functions · Nonlinear equations · Combinational equivalence checking · Boolean polynomials

1 Introduction

Boolean circuits are increasingly large and complex and require sophisticated methods to verify their functional correctness. Equivalence checking is an important part in the design of ASICs, permitting the rapid checking of combinational functional correctness, while the Boolean functions undergo design optimizations. The equivalence checking problem is to check if the worked-upon version of a Boolean circuit is functionally identical with the reference version of the same circuit.

Equivalence checking tools consider the logic function between the boundaries of sequential elements, also called logic cones. The outputs of sequential elements are the inputs to the logic cones. The inputs to the next stage of sequential elements are the outputs of the logic cones. A typical equivalence checking problem is constructed by identifying such logic cones in the reference and implementation circuits, and constructing a logic circuit wherein the matched inputs of the two logic cones are tied together, and the matched outputs of the two logic cones are

R. Vedam (✉)
Invensys, 10900 Equity Dr, Houston, TX 77041, USA
e-mail: rajvedam@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_77,

input to XOR gates, called compare-points. Let $y_R = F_R(X)$ be the output of the reference circuit F_R and $y_I = F_I(X)$ be the output of the implementation circuit, F_I, both given the same inputs, X. The combinational equivalence checking problem seeks answering the question, $y_R \oplus y_I = 0$ for all values of Boolean X?

A number of mathematical methods have been proposed for the equivalence checking problem. These include the use of binary decision diagrams [1], ATPG methods [2], and Boolean satisfiability methods [3]. These methods have found great success in industry to solve a wide variety of equivalence checking problems.

In practice, verifying a large circuit with many logic cones requires orchestration of several solvers such as BDD-based, SAT, justification, and ATPG. When solvers solve some compare-points and determine functional internal equivalence points, such information is communicated to other solvers which can bootstrap from such accumulated knowledge of the circuit. With appropriate heuristics, solver orchestration can perform well on classes of problems. Current trends in orchestration are more invasive, where BDD, SAT, ATPG, and justification solvers are deployed to verify potential internal equivalences [7].

Difficulties in equivalence checking arise when the reference circuit is close to RTL, and the implementation is gate-level that has undergone several design transformations. With low similarity between the circuits, and compounded by deep logic cones, the equivalence checking problem becomes difficult even with clever solver orchestrations. Additionally, when Boolean circuits contain complex arithmetic elements that cannot be easily identified (and extracted) in the reference and implementation circuits, it leads to exceedingly difficult equivalence checking problems. Traditional methods fail to solve complex data path problems that are embedded in the logic circuit, in a time- and memory-expedient manner. Thus, it continues to be of interest to investigate methods for Boolean equivalence checking.

In this paper, we provide a tutorial exposition of the equivalence checking problem. Making use of Boolean arithmetic polynomials [4], the underlying analytical problem in equivalence checking is explained with the examples.

2 Boolean Arithmetic Polynomials

Consider the n-dimensional Boolean space, $\mathbf{B^n}$. Let $\mathbf{x}$ be a vector of Boolean variables of size n, $\mathbf{x} \in \mathbf{B^n}$. The mapping of logical operation to the arithmetic domain is carried out via the following elementary operations [5, 6], resulting in multivariate Boolean arithmetic polynomials (BAP).

$$\mathrm{NOT}(x) = \bar{x} = 1 - x \tag{1}$$

$$\mathrm{AND}\,(x_1, x_2) = x_1 \wedge x_2 = x_1 x_2 \tag{2}$$

$$\mathrm{OR}\,(x_1, x_2) = x_1 \vee x_2 = x_1 + x_2 - x_1 x_2 \tag{3}$$

$$\text{Idempotence: } x \wedge x = x_2 = x \tag{4}$$

$$\text{Orthogonality: } x \wedge x = x(1-x) = 0. \tag{5}$$

Note that, the '+' and '−' operators are arithmetic operators.

Let $\mathbf{Y_S}$ be the set of Boolean polynomials, where Boolean polynomial $y(\mathbf{x}) \in \mathbf{Y_S}$, such that $y: \mathbf{B^n} \rightarrow \mathbf{B}$, $\mathbf{x} \in \mathbf{B^n}$. Let $y_1(\mathbf{x})$, $y_2(\mathbf{x}) \in \mathbf{Y_S}$ be two BAPs with the same Boolean vector argument, $\mathbf{x}$. Boolean operations using the two Boolean polynomials will result in a Boolean polynomial, that is,

$$y = y_1 \,(\text{Bool Operation})\, y_2 \in Y_S. \tag{6}$$

Sample Boolean operations are OR, AND, XOR, etc.

3 Construction of a BAP

In order to construct a BAP, we should determine the structure of the polynomial and then the corresponding coefficients of the terms. If there are N primary input lines, then there are up to 2^N terms in the BAP, in the general case.

Determination of the coefficients of the terms in the polynomial has computational complexity of the order of 2^N, where N = number of primary input lines. Consider the standard model [5]

$$\begin{aligned} y(x_1, x_2, \ldots, x_n) = a_0 &+ a_1x_1 + a_2x_2 + a_{1,2}x_1x_2 + a_3x_3 + a_{1,3}x_1x_3 + \cdots \\ &+ a_{1,2,\ldots,n}x_1x_2\ldots x_n \end{aligned} \tag{7}$$

We generate the coefficients by the solution of the following equations:

$$\begin{aligned} a_0 &= y_A(0,0,\ldots,0) \\ a_0 + a_1 &= y_A(1,0,\ldots,0) \\ a_0 + a_2 &= y_A(0,1,0,\ldots,0) \\ a_0 + a_1 + a_2 + a_{1,2} &= y_A(1,1,0,\ldots,0) \\ &\ldots \\ a_0 + a_1 + \cdots + a_{1,2,\ldots,n} &= y_A(1,1,1,\ldots,1) \end{aligned} \tag{8}$$

Consider the example of a 2-input XOR gate. With two inputs, the model has four terms.

$$y(x_1, x_2) = a_0 + a_1x_1 + a_2x_2 + a_{1,2}x_1x_2 \tag{9}$$

Define the vectors,

$$\phi = [\,1 \quad x_1 \quad x_2 \quad x_1x_2\,]$$
$$\theta = [\,a_0 \quad a_1 \quad a_2 \quad a_{1,2}\,]$$

We have four parameters to be estimated. We thus require at least four measurements of y and x to estimate the four parameters. Given the truth table (note 2^N data points)

X_1	X_2	Y
0	0	0
1	0	1
0	1	1
1	1	0

We have $Y = [0\ 1\ 1\ 0]^{\mathrm{T}}$ and concatenating four ϕ vectors using the four truth table entries, that is,

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

We solve $\Psi\ \theta = Y$ and get $\theta = [0\ 1\ 1\ -2]^{\mathrm{T}}$, which yields the XOR model in BAP form as $y = x_1 + x_2 - 2x_1x_2$

The procedure above leads to the following lemma.

Lemma 1: *The standard model (7) of the BAP is a canonical structure.*

Proof We have seen that there are up to $1 + \Sigma\ ^N C_k$ terms in a Boolean arithmetic polynomial, where the sum is over $k = 1$ to N, and N is the number of inputs. This evaluates to 2^N terms. Estimating 2^N coefficients requires 2^N measurements. An N-input truth table has 2^N entries for inputs and outputs. Thus, we get a linear system of the form

$$\Psi\ \theta = Y \tag{10}$$

The matrix Ψ described as the arithmetic transform or Reed-Muller transform on GF(2) [11] is easily proven to have rank 2^N. Thus, there exists a unique solution to the linear system. Thus, the logic represented in a truth table can be represented in a BAP, with unique coefficients, and is canonical.

Lemma 1 allows us to use BAP as a tool for equivalence checking. If two versions of a Boolean circuit have the same BAP (i.e., same terms and same coefficients), then they encode exactly the same logic.

The inverse of the matrix Ψ can be defined as follows, for $N = 2^n$, where $n = 0,1,2,\ldots$, is the number of bits [8]:

$$A_N = \begin{bmatrix} A_{\frac{N}{2}} & 0 \\ -A_{\frac{N}{2}} & A_{\frac{N}{2}} \end{bmatrix}, \quad A_1 = 1 \tag{11}$$

4 Arithmetic Operations with BAP

The equivalence checking problem using BAP is observed to incorporate arithmetic operations such as addition, subtraction, and multiplication of polynomials. Key to efficient polynomial arithmetic is the efficient representation of the BAP and the efficient multiplication of BAPs.

An *N*-bit unsigned word can represent *N* Boolean variables in a pre-determined sequence. Stringing together of *N*-bit words of an appropriate length can be used to represent an arbitrary number of Boolean variables.

Consider the example of representing the term $x_1\ x_2\ x_9\ x_{20}\ x_{33}\ x_{63}$, using word length, $N = 32$. Without signature translation (presented later), this requires two 32-bit unsigned words. The words are arranged in a pre-determined order, say, the LSB on the right and the MSB on the left. Then, the first 32-bit unsigned word has the 1st, 2nd, 9th, and 20th bit positions set, and all others zero. The second 32-bit unsigned word has the 1st and 32nd bits set. Taken together, they represent the desired term in symbolic form.

Given that several polynomials may have the same terms, rather than duplicating the symbolic terms, the terms can be uniquely kept in a polynomial-terms hash table and shared as needed by the different polynomials. The number of references made to a term can be kept count of. When new polynomials are created, the references to the re-used terms can be incremented, and when the polynomials are destroyed, the corresponding references can be decremented. When the reference count for a term goes to zero, the term can be released, thus allowing efficient representation of the symbolic terms.

Addition and subtraction of polynomials using hash table lookup take a linear time in the size of the polynomials. Note that, addition or subtraction operations arise in the evaluation of Boolean functions. Let us consider the worst case of the length of the polynomial in n-inputs, that is, 2^n. Addition or subtraction would involve a worst case of 2^{n+1} operations. Polynomial multiplication is expensive and, in the worst case, takes 2^{2n} operations. However, some optimizations can improve the efficiency.

5 Efficient Polynomial Multiplication

It is well known that univariate polynomial multiplication can be performed efficiently using FFTs. In the case of the multivariate BAP, we do not have such polynomial-time methods, and unless $P = NP$, it is unlikely that polynomial-time methods can be found to perform BAP multiplication. There exists scope to use heuristics to order the inputs and order of gate evaluation that might speed the computation.

We consider the paradigm of polynomial evaluation followed by dot product, followed by interpolation, for multivariate polynomial multiplication. In order to

perform the multiplication of a group of polynomials (such as those at a multi-input AND gate), we first consider two methods to evaluate a polynomial.

Given an n-input BAP, we can evaluate it for an instance of the n-inputs. Consider the representation of the n-inputs in bit form, ordered identical to the terms of the BAP. Let this bit input be U. Consider each term in a BAP, say B_i for $i = 1$ to N, the number of terms in the BAP. When at a bit-level AND we have B_i & $U = B_i$, then we can add the corresponding coefficient of B_i to a cumulative sum. The cumulative sum for a valid BAP will evaluate to zero or one. For an n-input BAP, we have 2^n input combinations to evaluate the BAP completely and yields a 2^n sized output vector, termed the forward transform, or the output vector in the basis of the inputs.

The second method of evaluating a polynomial involves the matrix equation in (10). By constructing an appropriately sized matrix Ψ, and using the ordered coefficients of the polynomials in θ, we can evaluate the output vector Y, or the forward transform. Note that, this computation is also of the order of 2^n, where n is the number of inputs.

When multiplying a group of polynomials, it is expedient to evaluate all polynomials in the group for the same input. When all polynomials in the product-group have been evaluated for the same input, the product term for that input is easily computed as the product of each BAP's output for that input. Note that, on incidence of a zero in evaluating a polynomial, the other polynomials in the product-group need not be evaluated for that input because their product for that input will be zero. This optimization speeds up processing. The analogous optimization for the matrix method, if a zero should be obtained on the multiplication of that row and θ, is to omit that row of Ψ for future computations.

The dot product of the resultant transformed BAPs yields the transformed product. The 2^n product vector can then be converted to a BAP by solving the linear equation (8) for θ, or using the structured inverse matrix (11). This interpolation operation is termed the inverse transformation.

We introduce *signature translation* to denote an operation wherein a term signature that may be noncontiguous in the original inputs can be aliased to a contiguous signature using fictitious inputs. As an example, a signature such as x_1 x_{12} x_{39} can be signature-translated to x_1 x_2 x_3, where the aliasing is $\{x_1 = x_1, x_{12} = x_2, x_{39} = x_3\}$. Such signature aliasing permits efficient BAP transformation and multiplication.

Signature translation also permits retaining the transformed product as an alternate representation of a BAP. In such a case, we also retain information of the aliases, so that the result can be inverse-transformed when required.

Let $Y_1(x)$ and $Y_2(x)$ be the transformed outputs of the BAPs $y_1(x)$ and $y_2(x)$, respectively. While y_1: $\mathbf{B^n}$ $\mathbf{->}$ $\mathbf{B}$, Y_1 is a Boolean vector of size 2^n. Just as Boolean operations on $y_1(x)$ and $y_2(x)$ yield new Boolean polynomials, Boolean operations on Y_1 and Y_2 yield new transformed Boolean outputs. This allows us to use either the BAP or the transformed output as alternate representations.

Shannon expansion can also be used to reduce the complexity of BAP multiplication and also for efficient multithreading. The Shannon expansion for a BAP is shown below, expanding on x_p.

$$p(x) = x_p p_1(x)|_{x_p=1} + (1 - x_p) p_2(x)|_{x_p=0} \tag{12}$$

Let $p_a(x)$ and $p_b(x)$ be two polynomials to be multiplied. If both are represented in Shannon form, expanded on the same input, say x_p, then their product is expressed as $x_p\, p_{a1}(x)\, p_{b1}(x) + (1 - x_p)\, p_{a2}(x)\, p_{b2}(x)$, where the sub-product terms have one bit less than the parent polynomials, and thus simplifies large polynomial multiplications. Note also that, the two sub-products can be computed in parallel.

The Horner form can also be used to reduce the bit complexity of polynomial multiplication. Factoring the pivot variable of interest, a BAP can be represented in Horner form as shown below.

$$p(x) = x_p p_1(x) + p_2(x) \tag{13}$$

Here, both p_1 and p_2 are independent of x_p. In a manner analogous to the product with Shannon expansion, $p_a(x)$ and $p_b(x)$ can be represented in Horner form using the same pivot variable, x_p. The product of p_a and p_b is then $x_p\ (p_{a1}\ p_{b1} + p_{a2}\ p_{b1} + p_{a1}\ p_{b2}) + p_{a2}\ p_{b2}$. The four sub-terms can be computed in parallel.

Efficient polynomial multiplication is thus seen to have several components—transformed multiplication, signature translation, and use of Shannon or Horner forms. While these methods can help in fast multiplication of polynomials, one still needs to consider methods for expedient verification of complex circuits.

6 Equivalence Checking Methods using BAP

6.1 Naïve Construction of Formulas

Let the logic circuits whose equivalence is to be checked be mitered (i.e., pairs of matched outputs tied to XOR gates) and represented by M gates, N-inputs, and one output. We first define N unique polynomials for the N-inputs. Subsequently, we construct polynomials at the output of each logic gate, by composing the logic of that gate using the polynomials at its input nets. We construct M such polynomials for the M gates. Figure 1 shows a sample circuit, and Table 1, the polynomials setup for each net. If we have a logic cone that passes the equivalence test, then the output should be a constant zero, that is, for all Boolean inputs, there is no difference between the reference and implementation circuits. In such a case, the polynomial constructed at the output net will be identically zero. This method works well for small circuits. The drawback of the method is that the size of the intermediate polynomials can grow quite large and may exceed time and memory

resources. It is thus necessary to devise mechanisms that can help deal with the time and space complexity.

6.2 Use of Transformed Output

To help manage complexity, we can use the alternate representation of the BAP, using the forward transform. In this case too, efficient methods can be devised for quick computation. Additionally, the transformed output vector Y, being Boolean, can be represented as a bit signature, storing W of them in an unsigned W-bit word, thus reducing storage complexity by a factor of W. Table 2 illustrates the method for the example in Fig. 1, assuming full-size (3-bit) transforms at all nets, for convenience. Note that, in general, with signature translation, we can have shorter bit sequences. The drawback of this method is that it is expensive when the bit count in the product exceeds a threshold.

6.3 Use of Lifting Polynomials/Cutpoints

We introduce the concept of a "lifting polynomial", that is, a new input that is an alias for a BAP. We use lifting polynomials in two cases. An internal equivalence point is one whose BAP is an exact copy of a BAP computed for some other net. In

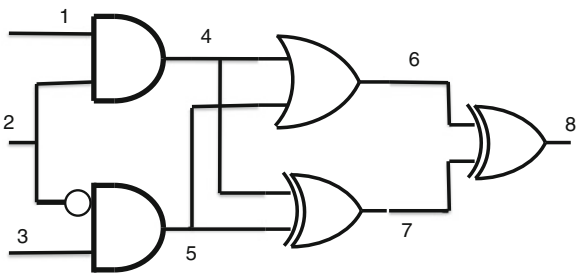


Fig. 1 Example circuit for equivalence checking

Table 1 BAP for example circuit

Net	Polynomial
1	X_1
2	X_2
3	X_3
4	X_1X_2
5	$X_3 - X_2X_3$
6	$X_3 + X_1X_2 - X_2X_3$
7	$X_3 + X_1X_2 - X_2X_3$
8	0

Table 2 Alternate BAP for example circuit

Net	Bits	Poly. transform	Bit signature
1	X_1	[0 0 0 0 1 1 1 1]	15
2	X_2	[0 0 1 1 0 0 1 1]	51
3	X_3	[0 1 0 1 0 1 0 1]	85
4	X_1X_2	[0 0 0 0 0 0 1 1]	3
5	X_2X_3	[0 1 0 0 0 1 0 0]	68
6	$X_1X_2X_3$	[0 1 0 0 0 1 1 1]	71
7	$X_1X_2X_3$	[0 1 0 0 0 1 1 1]	71
8	$X_1X_2X_3$	[0 0 0 0 0 0 0 0]	0

such a case, the two nets are equivalent. We introduce a new input—a lifting polynomial—at this equivalent point in lieu of the equivalent BAP.

When the number of bits in a potential product exceeds some threshold deemed to require excessive time and space requirements, we can introduce new inputs—lifted polynomials, in lieu of the BAP that they are aliasing. This artifice hides the complexity while marching from inputs to outputs. However, we cannot determine equivalence easily at the output, which may falsely indicate failing circuits. In such a case, it is necessary to unwind the lifted polynomials to the original inputs before equivalence determination can be made reliably. Table 3 shows the use of lifted polynomials x_4 and x_5 when bit complexity exceeds 2, for the example circuit.

6.4 BAP Equivalence Checking Using Justification

Constructing the traditional Justification arguments, we set the output to 1 and seek inputs that justify the assignment. If no consistent assignment can be found, then we have a succeeding circuit, or if one can be found, we have a failing pattern.

In the example circuit, we set $x_8 = x_4x_5 = 1$. This implies that $x_4 = 1$ and $x_5 = 1$. For x_4, from Table 3, this implies that $x_1 = 1$ and $x_2 = 1$. For x_5, this

Table 3 Lifted BAP for example circuit

Net	Polynomial
1	X_1
2	X_2
3	X_3
4	$X_4 = X_1X_2$
5	$X_5 = X_3 - X_2X_3$
6	$X_4 + X_5 - X_4X_5$
7	$X_4 + X_5 - 2X_4X_5$
8	X_4X_5
8—Unwinding	$(X_1X_2)\ (X_3 - X_2X_3) = 0$

implies that $x_2 = 0$ and $x_3 = 1$. We cannot simultaneously have x_2 at 0 and 1. Therefore, we conclude that x_8 is not equal to 1 and, hence, is 0, and succeeding.

6.5 Using Lifting and Transformed Outputs

If we have a transform representation of the circuit with lifted polynomials, we can check equivalence by looking at the justification and/or evaluation of a number of smaller terms. Consider the example circuit and Table 3. A 2-bit transformation of x_8 yields [0 1 1 0], as expected for a 2-input XOR gate. This implies that we have to check for the patterns $x_4 = 1$ and $x_5 = 0$, $x_4 = 0$ and $x_5 = 1$, to see if either of them evaluate to 1. One can justify the two patterns to determine that neither is possible. Or we could evaluate the products $x_4\ (1 - x_5)$ and $(1 - x_4)x_5$. The evaluation of the products leads to $x_1x_2 = 1$ and $x_3 - x_2x_3 = 1$. Clearly, the two sub-products cannot be satisfied with Boolean values for x_1, x_2, and x_3. Hence, neither product is feasible, and the circuit passes the equivalency test.

7 Concluding Remarks

BAP exhibits a size disadvantage when compared with BDDs. For example, an m-input OR gate has an optimized length of O(m) using BDDs, whereas the standard form BAP (7) would require $O(2^m)$. However, the factorized form of the BAP for an m-input OR gate would require length of O(m) [5]. The size disadvantage can show up significantly when constructing formulas for difficult circuits. Some reasons to study the benefits of this technology vis-à-vis the more efficient BDDs are (a) analytical processing is easier; (b) this is a complementary technology to BDDs, and may be able to assist in cases where BDDs fail; (c) we can consider some unconventional methods for equivalence checking.

When the Boolean derivative [9] of the output with respect to every input variable is zero, it indicates a constant output. A single random simulation determines if the constant is zero or one and concludes equivalence. The method gives a convenient mechanism to contain the complexity in unwinding lifted polynomial. However, it may show false negatives because depending on the derivative path, we may look at only a subset of functionality. These false negatives can be quickly simulated for verification of equivalence. Multivariate polynomial factorization [10] can reduce the size of the formulas. If across a circuit-cutline, all polynomials have a common factorization (say GCD), then we can reduce all formulas on that cutline by dividing them by the common factorization, thus reducing the size of the forward formulas. When the maximum of the miter output with respect to the inputs is 0, then we have equivalence, but have the issue of proving a global optimum for this Boolean optimization problem.

Being a technology complementary to those offered by BDD-based solvers, BAP can occupy a place of importance in the orchestration of solvers for equivalence checking. Oftentimes, it is observed in solver orchestration that where one solver experiences complexity difficulties, a complementary solver may be able to perform better. While BDDs exhibit variable function sizes based on variable ordering, the size of the canonical BAP is invariant with respect to ordering. BAP is also of pedagogic interest because the nature of formula management as well as the complexity issues experienced has a close correspondence with BDD-based methods.

References

1. Bryant RE (1986) Graph-based algorithms for Boolean function manipulation. IEEE TC C-35(8):677–691
2. Brand D (1993) Verification of large synthesized designs. In: Proceedings of international conference on CAD, pp 534–537
3. Marques-Silva J, Sakallah KA (1996) GRASP: a new search algorithm for satisfiability. In: IEEE/ACM international conference on CAD, pp 220–227
4. Zhegalkin II (1927) On the technique of calculating propositions in symbolic logic. Matematicheskii Sbornik 43:9–28
5. Schneeweiss WG (1998) On the polynomial form of Boolean functions: derivations and applications. IEEE Trans Comput 47(2):217–221
6. Falkowski BJ (1999) A note on the polynomial form of Boolean functions and related topics. IEEE Trans Comput 48(8):860–864
7. Mishchenko A, Chatterjee S, Brayton R, Een N (2006) Improvements to combinational equivalence checking. In: Proceedings of ICCAD 2006, pp 836–847
8. Falkowski BJ (1999) Relationship between arithmetic and Haar wavelet transforms in the form of layered Kronecker matrices. Elect Let 35(10):799–800
9. Akers S (1959) On a theory of Boolean functions. SIAM J 7:487–498
10. Kaltofen E (1985) Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization. SIAM J Comput 14:469–489
11. Zilic Z, Radecka K (2009) Enabling practical uses of arithmetic transform: a comprehensive analysis. Proceedings of 9th international workshop on Reed-Muller transform and its applications, RM'09

Generation of High-Frequency Pulses for Transient Studies

Eranna, B. V. Sumangala and G. R. Nagabhushana

Abstract This paper presents a transient voltage distribution in a small transformer winding model developed using the lumped RLC—parameters for the high-frequency lightning impulse voltages of very steep time to front and the voltage distribution due to switching surge over voltages which are damped asymmetrical sinusoidal in wave shape for a frequency of 65.8 kHz. The impulse waveforms of different front times and damped sinusoidal waveform (DSW) (of frequency 65.8 kHz) are generated experimentally using the custom-built pulse generators. The sources are applied between the line end and the grounded end of the winding model. The voltage distribution at various sections with reference to the grounded end of the winding model is studied. The waveforms for which the voltage distribution studied are simulated using the OrCAD PSpice software and applied to the winding model to validate the results and are agreeable with acceptable approximations.

Keywords High-frequency pulses · Impulse wave form · Damped sinusoidal waveform · Transients

1 Introduction

The power transformer constitutes a vital link in the integrated power systems. The crucial nature is due to the fact that power flow is invariably through these transformers. With very large power transformers, it is essential to optimize the

Eranna (✉) · B. V. Sumangala
Department of EEE, Dr. Ambedkar Institute of Technology, Bangalore, India
e-mail: eranna64@gmail.com

G. R. Nagabhushana
High Voltage Engineering, Indian Institute of Science, Bangalore, India

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_78,

insulation structure with a view to economize on cost and also reduce the size and weight of the transformer. If any transformer in a system fails, the power flow through that completely stops, thus causing a serious disturbance. The seriousness increases enormously with the system voltage since the power ratings go up as the square of the voltage. In addition, failure of a transformer (of extra high-voltage class) invariably causes losses of several hundreds of crores or even more. Further, failure could be violent and cause serious injuries to personnel and also damage nearby equipment. Further, almost always, failure of power apparatus is due to failure of the electrical insulation system. Therefore, it is of great importance to study the causes of failure of power transformers and prevent them to the best possible extent. One of the possible reasons of failures has been thought to be due to switching surges of relatively fast fronts. Switching over voltages of longer times to front and tail do not pose a serious problem to transformer. On the other hand, lightning over voltages having very sharp rise times, as low as 1 μs, pose considerable hazard to the transformer insulation. Some of the parameters of study are the magnitude of the wave, and effect of frequency on oscillatory switching surge voltage distribution in power transformers under surges of various wave shapes including a large range of time to front.

For fast front voltages, the inductances of the transformer may be treated, to a first degree of approximation, as open circuits. Thus, the voltage distribution is governed by the capacitance network—capacitances to ground as well as across the sections, which is somewhat similar to a ladder network. This distribution involves severe voltage gradients at or just inside the winding at the line end stressing that insulation severely. Because of this, surge distribution in transformer windings has attracted a great deal of attention. On considering the effect of inductances and resistances, the voltage distribution tends to become better. The exact voltage distribution behaviour is extremely difficult to analyse but because of the great importance, studies have been carried out using many approaches with simplifying assumptions. High-voltage transformer windings for core type of transformers can be made to have a more uniform initial voltage distribution than a simple continuous disc winding by appropriate interleaving [1] of turns. The purpose of interleaving winding is to secure a better capacitive coupling through the winding.

2 High-Frequency Pulses

2.1 Generation-1: Lightning Impulse

A capacitor C1 previously charged to a particular DC voltage is suddenly discharged into a wave-shaping network (R1, R2 and C2) by closing the switch S. The discharge voltage $v(t)$ across C2 gives the desired exponential wave shape. The wave front and wave tail times can be independently controlled by changing

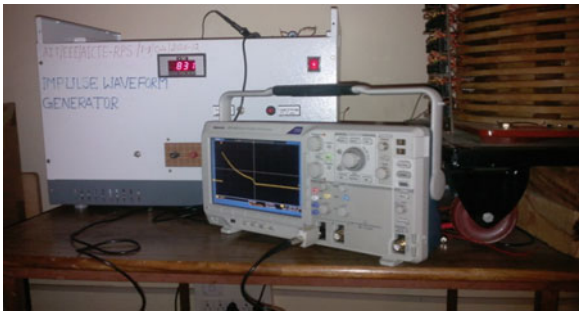


Fig. 1 Photograph of impulse generator with DSO

either R1 or R2 to obtain the desired wave shape. The photograph of the impulse generator along with DSO is shown in Fig. 1.

The schematic diagram of the impulse waveform generator and the experimental and simulated waveform of 0.46/48.3 μs is shown in Figs. 2, 3, and 4.

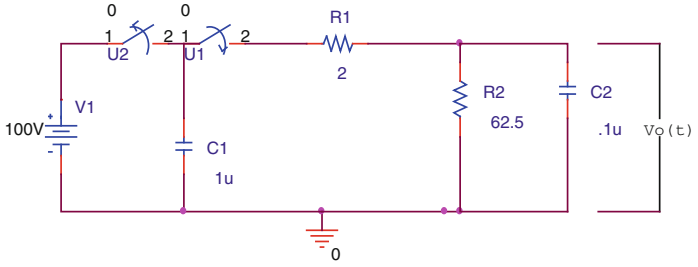


Fig. 2 Schematic diagram of the lightning impulse waveform generator

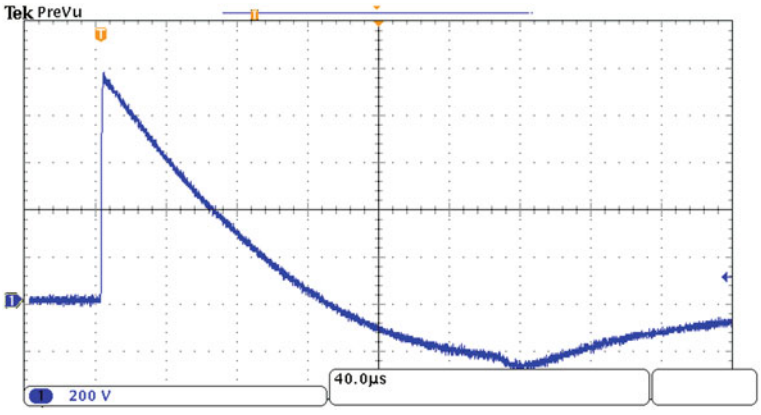


Fig. 3 Experimental impulse waveform ($t1/t2 = 0.46/48.3$) μs

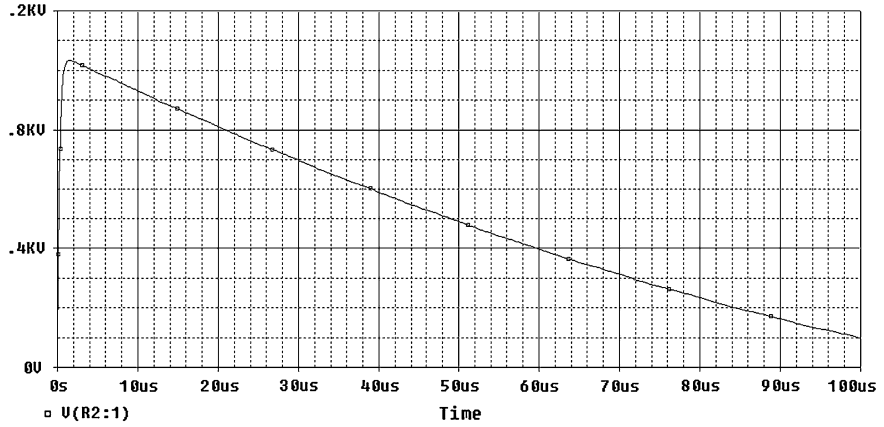


Fig. 4 Simulated impulse waveform (0.46/48.3) μs

3 Damped Sinusoidal Waveform

A pulse generator was developed for the generation of high-frequency damped sinusoidal pulses. This generator functions on the principle of discharging a charged capacitor through an appropriate inductance. The circuit parameters R, L and C are so chosen to obtain the desired magnitude and frequency of damped sinusoidal waveform (DSW). This pulse representing the VFTO is used to study the response of the transformer winding model to very fast transients. The photograph of the DSW generator along with DSO is shown in Fig. 5.

The schematic diagram of the DSW generator and experimental and simulated DSW of frequency 65.8 kHz is shown in Figs. 6, 7, and 8.

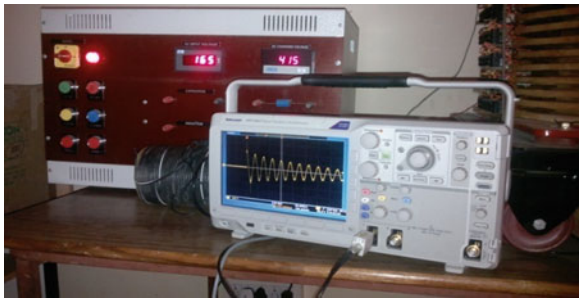


Fig. 5 Photograph of damped sinusoidal waveform generator with DSO

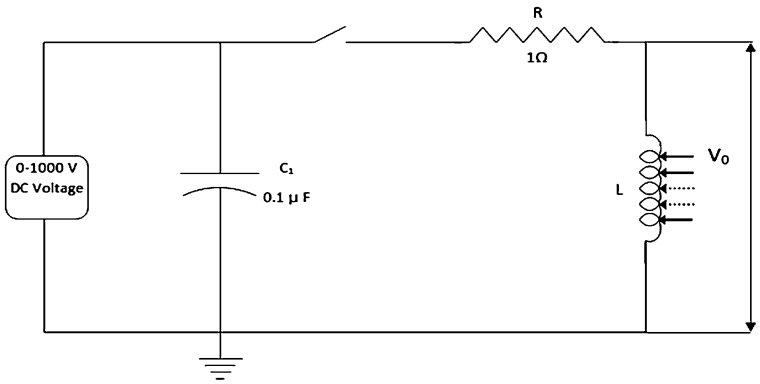


Fig. 6 Schematic diagram of the damped sinusoidal pulse generator

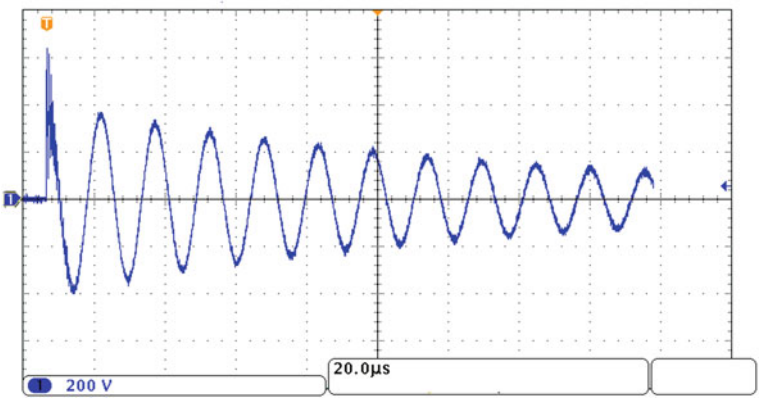


Fig. 7 Damped sinusoidal waveform (65.8 kHz)

4 Transient Studies on a Coil

A small winding model (custom built) with interleaving has been developed for surge voltage distribution measurements. The winding model has twenty discs wound on a coil former having length of 400 mm and a diameter of 500 mm made out of glass–epoxy material. The first ten discs are interleaved, having eight turns per disc with a rectangular conductor of double paper covered having size 12 × 3 mm. The remaining ten discs are of continuous type placed below interleaved winding having the same geometry. The lead outs (tappings) are brought out ensuring no interturn short using PTFE or PVC covered wires soldered at specified turns. All the discs are interconnected so as to leave one end lead to be connected to ground and the other end to the HV terminal, and adequate numbers

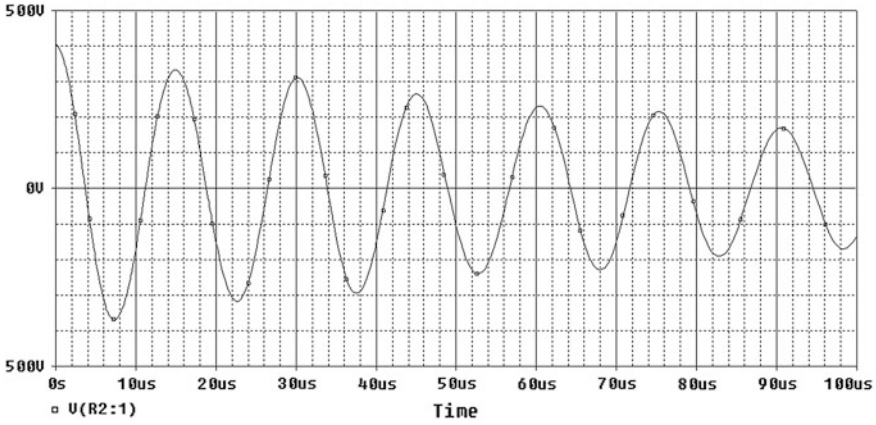


Fig. 8 Simulated damped sinusoidal waveform (f = 65.8 kHz)

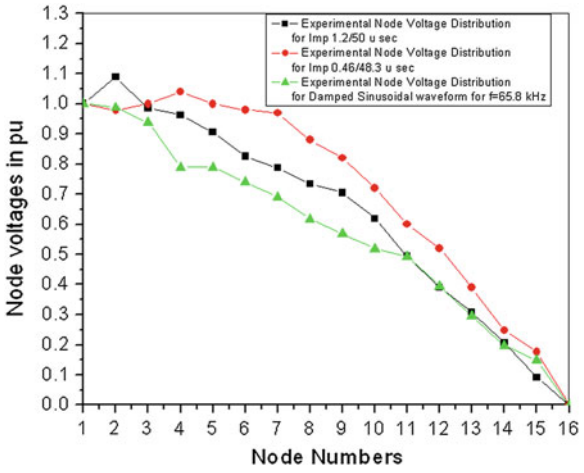


Fig. 9 Experimental node voltage distribution for impulses of 0.46/48.3 and 1.2/50 μs and damped sinusoidal waveform (f = 65.8 kHz)

of press board strips are provided between the discs to make the model more robust and rugged. The above model is supported on a metal trolley provided with wheels and a jack insulated with respect to the winding. The sources generated using the custom-built pulse generators are applied to the winding model, and voltage distribution is studied for the 1.2/50 μs, 0.46/48.3 μs and for the DSW of frequency 65.8 kHz. The results are shown in the Fig. 9.

5 Simulation

The equivalent circuit topology is usually defined by the ladder-type network and must describe the surge behaviour of the windings for wide range of variation in the input pulse wave shape and frequency. The equivalent circuit is a series of π circuits with mutual magnetic coupling. Each of these circuits is a section of the winding, which can be a disc, or a layer or a set of any of these elements according to the details required for the analysis. The elements of a Section [2] are series resistance, self-inductance, mutual inductance between the section and the other sections, series capacitance, and the capacitance to ground. The correct and reliable calculation of the inductive and capacitive parameters of the equivalent circuit is very important to obtain an acceptable model. The self-inductance and mutual inductance between the section and the various other sections or discs are calculated using the well-known equations [3].

In the OrCAD PSpice software used for simulation, there is no provision for incorporation of mutual inductances directly. Therefore, the coupling coefficients are calculated by making use of the mutual inductance values given by

$$K_{i-j} = \frac{M}{\sqrt{L_r L_s}}$$

K_{i-j} is the coupling coefficient between the discs '*i*' and '*j*' that is L_i and L_j.

It is to be noted here that the coupling between the sections less than 15 % is neglected. Also, as the core behaves like an air core around 50 kHz and higher, reduces the effective coupling between the turns. Therefore, coupling between far off discs can be neglected. The DC resistance is measured using the digital R-L-C metre.

In practice, the inter-section capacitances were calculated (as the analysis was carried out by approximating the complete winding with a very few number of sections) by assuming the voltage to be evenly distributed among the turns within a section. But, this is not so in practice as the various points (turns) within a section are not at the same potential. Therefore, the evaluation of various capacitances in a transformer winding is very important, but a complicated task. In this section, capacitance values between the sections and capacitances from section to ground are calculated by treating the adjacent coils as parallel plate capacitor by using the well-known formula, that is,

$$C = \frac{\varepsilon A}{d} = \frac{\varepsilon_0 \varepsilon_r A}{d}$$

where $\varepsilon_0 = 8.854 \times 10^{-12}$F/m—absolute permittivity of the space. ε_r = relative permittivity of the dielectric medium–combination of air and paper is taken as 2.3. And d is the distance between the two coils. The electrical equivalent network of the transformer winding model with source at its terminal is shown in Fig. 10.

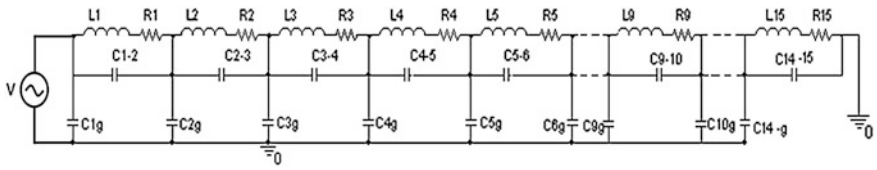


Fig. 10 Electrical equivalent network of the transformer winding model with source

6 Simulation Results and Comparison with Experimental Results

The well-known OrCAD PSpice software has been used for the simulation of the transformer using the above data of resistances, inductances, and capacitances. OrCAD PSpice is a simulation programme that models the behaviour of a circuit containing analogue elements(R, L and C) as well as sources and devices. The OrCAD capture provides a professional design entry environment with many advanced capabilities. The schematic created in OrCAD capture generates a circuit file set, which in turn is read by PSpice for computation. The results can be obtained in tabular or graphical format. PSpice introduces the concept of simulation profiles. Each simulation profile refers to one schematic in a design and includes one analysis type (AC, DC or transient) with options like sensitivity, temperature, parametric, etc.

The electrical equivalent network of the winding model shown in Fig. 9 is used for digital simulation. The line end of the winding (Node-1) is connected to the source, and the other end of the winding (Node-16) is grounded. Care is taken to see that all the resistances, capacitances, self-inductances, and mutual inductances between the various sections in terms of the coupling coefficients are incorporated in the network model. The different types of sources which were used during the experiments are generated in the simulation to be very close to the pulses from the experimental pulse generators as shown in Figs. 3 and 6. The simulation and the experimental results for DSW and lightning impulses of 0.46/48.3 and 1.2/50 μs are shown in Figs. 11, 12, and 13 respectively.

7 Results and Discussions

1. A partly interleaved winding model was fabricated (custom built) for the purpose of transient studies.
2. Experiments were conducted for different surges of different frequencies—imitating lightning, switching and VFTO surges.
3. An electrical network model for the fabricated winding model was developed using R, L and C at section level (8 to 16 turns in each section).

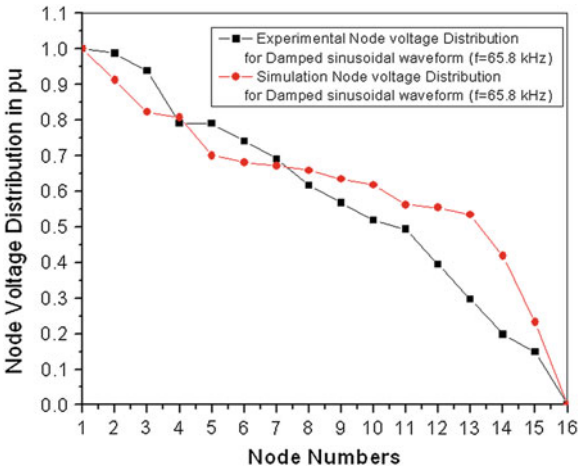


Fig. 11 Comparison of experimental and simulation node voltage distribution for damped sinusoidal waveform of frequency 65.8 kHz

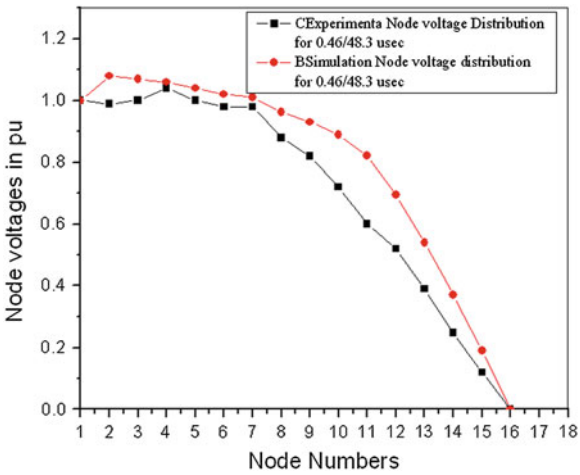


Fig. 12 Comparison of experimental and simulation node voltage distribution for impulse 0.46/48.3 μs

4. Simulations were carried out using OrCAD PSpice software for different sources as obtained in the experiments.
5. The voltage distribution is nonlinear along the length of the winding, and the degree of nonlinearity increases with increase in frequency.
6. It is seen that the ground end turns are stressed more compared to line end for both impulse and DSWs.
7. The results obtained from the simulation are in agreement with acceptable approximations.

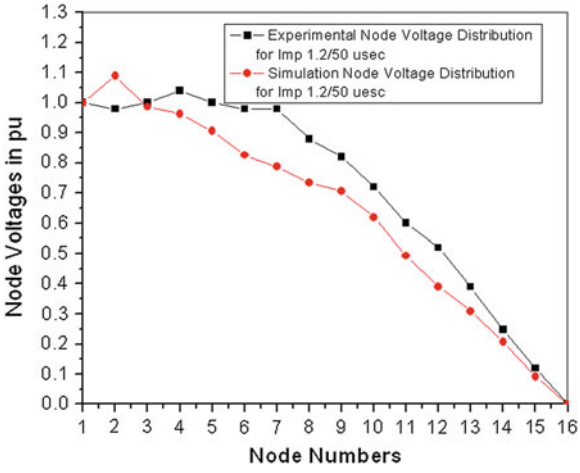


Fig. 13 Comparison of experimental and simulation node voltage distribution for impulse 1.2/50 μs

Acknowledgments We are grateful to the AICTE, New Delhi, for providing us the financial assisstance to carry out the project under AICTE-RPS.

We are thankful to the Principal and Management of Dr. Ambedkar Institute of Technology for their kind support in carrying out this work.

References

1. Van Nuys R (1978) Interleaved high-voltage transformer windings. IEEE Transactions on power apparatus and systems, Pas-97, No. 5, Sept 1978
2. Malewsky R, Francheck MA, McWhirter JH (1994) Experimental validation of a computer model simulating an impulse voltage distribution in HV transformer windings. IEEE Trans Power Delivery 9(4):1789–1798
3. Grover FW (1946) Inductance calculations. Dover Publications, Inc., New York

Toffoli Cascade Synthesis of an Optimized Two-Bit Comparator

H. R. Bhagyalakshmi and M. K. Venkatesha

Abstract Reversible logic gates are in high demand in the world of low-power digital circuits and quantum computers since the power dissipation problems which occur in the computational operations of a computer can be avoided by using reversible structures instead of the conventional irreversible logic structures. One important task is to compare the inputs using a computer and decide what to do next. In a quantum computer, this must be built using reversible logic circuits. In this paper, we are presenting a two-bit comparator using new reversible logic gate called BVMF gate which is a multifunction gate. The optimization of the designed circuit is achieved by keeping the number of gates and number of garbage outputs to a minimum value.

Keywords Comparator · Low-power digital circuits · Power dissipation · Reversible logic

1 Introduction

R Landuaer, in 1960, stated in his principle that the loss of one bit of information dissipates KTln2 joules of energy where k is the Boltzmann's constant and T is the absolute temperature at which the operation is performed [1]. Later Bennett, in 1973, showed that in order to avoid KTln2 joules of energy dissipation in a circuit, it must be built from reversible circuits [2]. Frank [3] showed that in order for the computer to attain very high levels of performance at fixed levels of power consumption and to keep the energy efficiency of computation close to 100 %, *the*

H. R. Bhagyalakshmi (✉)
BMS College of Engineering, Bangalore, Karnataka 560019, India
e-mail: hrbbmsce@gmail.com

M. K. Venkatesha
RNS Institute of Technology, Bangalore, Karnataka 560098, India
e-mail: mkvenkatehsa@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_79,

computational blocks must be designed such that the level of energy dissipation of individual logic operations becomes very much less than 40 KT.

A reversible logic gate is an n-input, n-output logic device with one-to-one mapping. This helps to uniquely determine the outputs from the inputs but also the inputs can be uniquely recovered from the outputs. Extra inputs are added to make the number of inputs equal to the number of outputs in a reversible logic gate [4, 5]. Fan-out in reversible gates must be derived from the gate, maintaining one or more inputs constant.

A reversible circuit should be designed using minimum number of reversible gates. One key requirement to achieve optimization is that the designed circuit must produce minimum number of garbage outputs. From the survey of existing research works on various reversible logic circuits, we observe that the important parameters of interest are minimum garbage outputs, minimum delay, and minimum quantum cost. To achieve this, many researchers find an optimized way of connecting discrete reversible logic gates to build the circuit. In the present work, one-bit and two-bit comparator circuits are constructed using the Toffoli module of BVMF gate. For this, we have used RevKit [6], an open source toolkit to construct the module. We are also using a fan-out gate called BVF gate for fan-out purposes.

This paper is organized as follows: Sect. 1 gives the necessary introduction on some of the basic reversible logic gates related to the present work. In Sect. 2, the new reversible logic gates and their quantum implementation used in the present work are described. In Sect. 3, the design of one-bit and two-bit comparators using reversible logic gates along with their quantum implementation is discussed. In Sect. 4, comparison of other existing designs is given. Section 5 presents conclusions and scope for further research.

2 Introduction to Reversible Logic Gates

In this section, some basic reversible logic gates are discussed.

2.1 Feynman Gate

Figure 1 shows a 2×2 Feynman gate [7]. The output is given by $P = A$, $Q = A \oplus B$. Its quantum cost is 1. It is also called as controlled-NOT gate or CNOT gate.

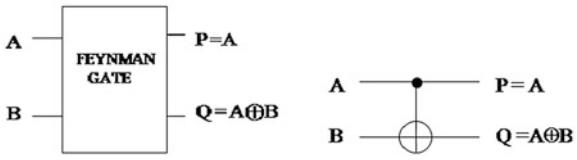


Fig. 1 Feynman gate and its graphical representation

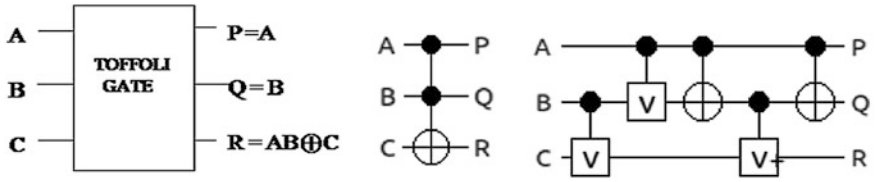


Fig. 2 Toffoli gate, symbol, and its quantum representation

2.2 Toffoli Gate

Figure 2 shows a 3 × 3 Toffoli gate, its symbol, and its quantum representation [4]. The output is defined by $P = A$, $Q = B$, $R = AB \oplus C$. The inputs A and B are known as control inputs, and C is known as the target. This gate passes the inputs A and B directly to the corresponding outputs without any change and toggles the target C if and only if both inputs A and B are 1.

2.3 Fredkin Gate

Figure 3 shows a 3 × 3 Fredkin gate [5] and its symbolic representation. The output is defined by $P = A$, $Q = \bar{A}B \oplus AC$, and $R = \bar{A}C \oplus AB$. The quantum cost of a Fredkin gate is 5. It is also called as a controlled SWAP gate as the two bits are swapped if and only if the control input A is 1.

Figure 4 shows the quantum representation and its Toffoli gate implementation and its symbolic representation [8, 14].

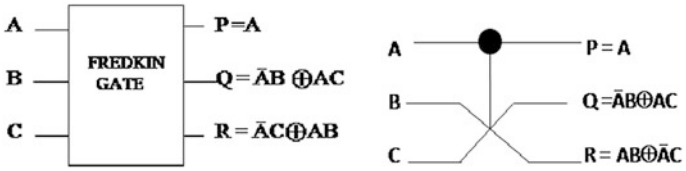


Fig. 3 Fredkin gate and its symbolic representation

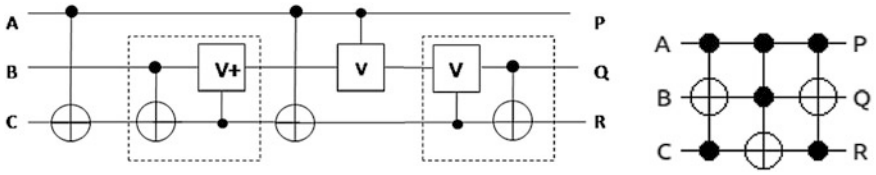


Fig. 4 Quantum and Toffoli gate representation of Fredkin gate

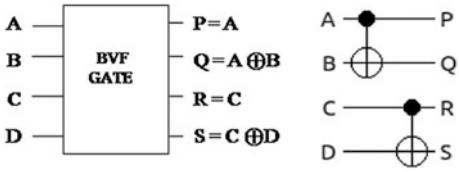


Fig. 5 BVF gate and its Toffoli representation

3 Reversible Logic Gates Used

Two reversible logic gates [9] that are used along with primitive gates for the construction of comparator circuits are presented. The details of each of these gates are given below.

3.1 BVF Gate

The BVF gate is designed to have controlled fan-out of inputs. It is a 4 × 4 reversible gate. Figure 5 shows BVF gate and its Toffoli gate representation. The quantum cost of BVF gate is 2.

3.2 BVMF Gate

The BVMF gate is a 4 × 4 reversible logic gate with output vectors as shown in the Fig. 6a. The BVMF gate, its Toffoli gate representation, and quantum gate implementation are shown in Fig. 6b.

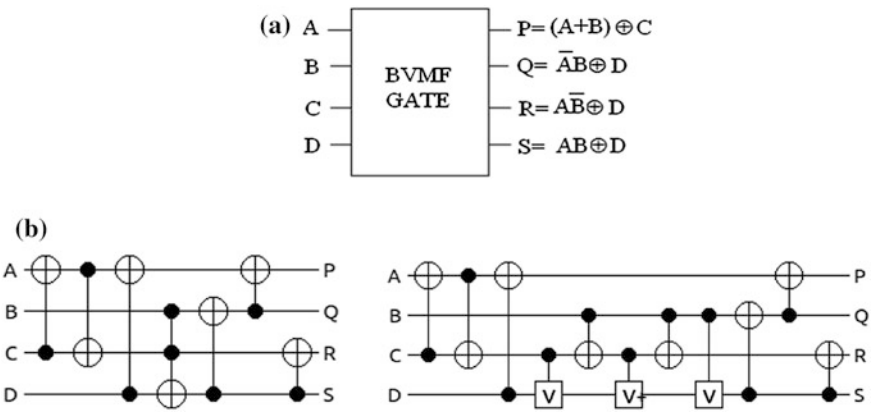


Fig. 6 **a** BVMF gate. **b** Toffoli gate and quantum gate implementation of BVMF gate

Table 1 BVMF as a one-bit comparator

Inputs				Outputs			
A	B	C	D	$(A = B)0$	$A < B$	$A > B$	$(A = B)1$
0	0	1	0	1	0	0	0
0	1	1	0	0	1	0	0
1	0	1	0	0	0	1	0
1	1	1	0	0	0	0	1

4 Comparator Design

Synthesis of one-bit and two-bit comparators is explained in the following sections.

4.1 One-Bit Comparator

A single BVMF gate can be used as a one-bit comparator with $C = 1$ and $D = 0$. When two one-bit numbers $A = A0$ and $B = B0$ are given to inputs A and B, respectively, and with $C = 1$ and $D = 0$, the outputs of BVMF gate produce the comparator outputs as shown in the Table 1. It is observed that when two one-bit numbers are equal to 0, it indicated an output $P = 1$. Similarly, if two one-bit numbers are 1, it indicated an output $S = 1$. BVMF gate $C = 1$ and $D = 0$ compares input A with input B. Also equal inputs are clearly indicated at the outputs P and S. Table 1 shows the comparator table.

Implementation of one-bit comparator circuit using BVMF gate and the simulation results using Dinotrace waveform are shown in Fig. 7.

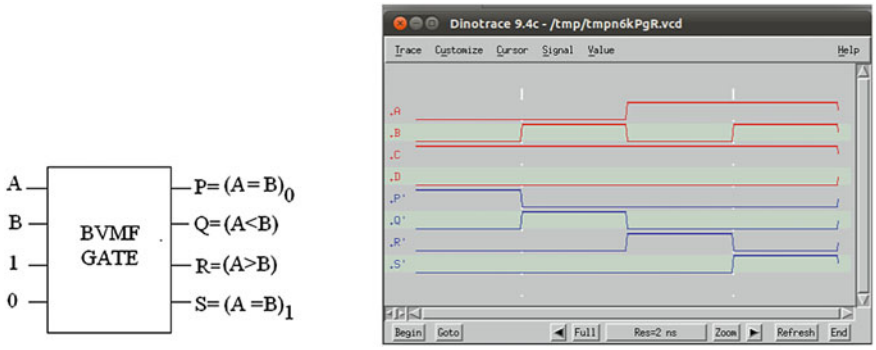


Fig. 7 BVMF gate as a comparator and the simulation waveform

4.2 Two-Bit Comparator

The BVMF gate is used to design a two-bit comparator with the following features:

- It compares two numbers of two-bit lengths.
- The comparison result shows the relative magnitude.
- It also produces the relative magnitude of individual bits, which is not done by a conventional comparator.

The algorithm of comparison of two n-bit numbers is given below.

1. Compare the most significant bits of the two numbers. If there is an inequality, output the result and stop comparison.
2. If the two bits are equal, then proceed to compare the next lower significant bits. If the lower bits are unequal, output the result.
3. If the lower bits are equal, compare the next lower bits. Continue to compare the next lower bits till the inequality is reached.

The design of two-bit comparator uses BVMF gates, BVF gates, Fredkin gates, and Peres gates. The two numbers $A1A0$ and $B1B0$ are compared.

4.3 Synthesis of Two-Bit Comparator

There are many reversible logic gates available for the synthesis. The most popular reversible gates are the Toffoli gate and the Fredkin gate. But in the present work, exact synthesis using Toffoli cascades is adopted. A Toffoli gate is a universal gate, and any function can be realized using it. Toffoli gates are used to build systems that will perform any desired Boolean function computation in a reversible manner. Due to probable technological restrictions, the synthesis of reversible logic is done with no feedback and no fan-out [10]. So the cascade structure is the only model satisfying those conditions. Thus, we consider cascades of Toffoli gates. Toffoli gate circuits have been demonstrated as a promising alternative to achieve minimal results for functions with up to six variables.

From the existing literature on exact synthesis [11–13] of reversible logic, it is observed that

- Smaller circuits than the currently best known realizations.
- Computation of larger circuits using the smaller as building blocks.
- Circuit cost can be calculated by knowing the cost of each multiple controlled Toffoli gate.

The block diagram of a two-bit comparator using reversible logic gates obtained from RevKit tool implementation is shown in the Fig. 8. Toffoli

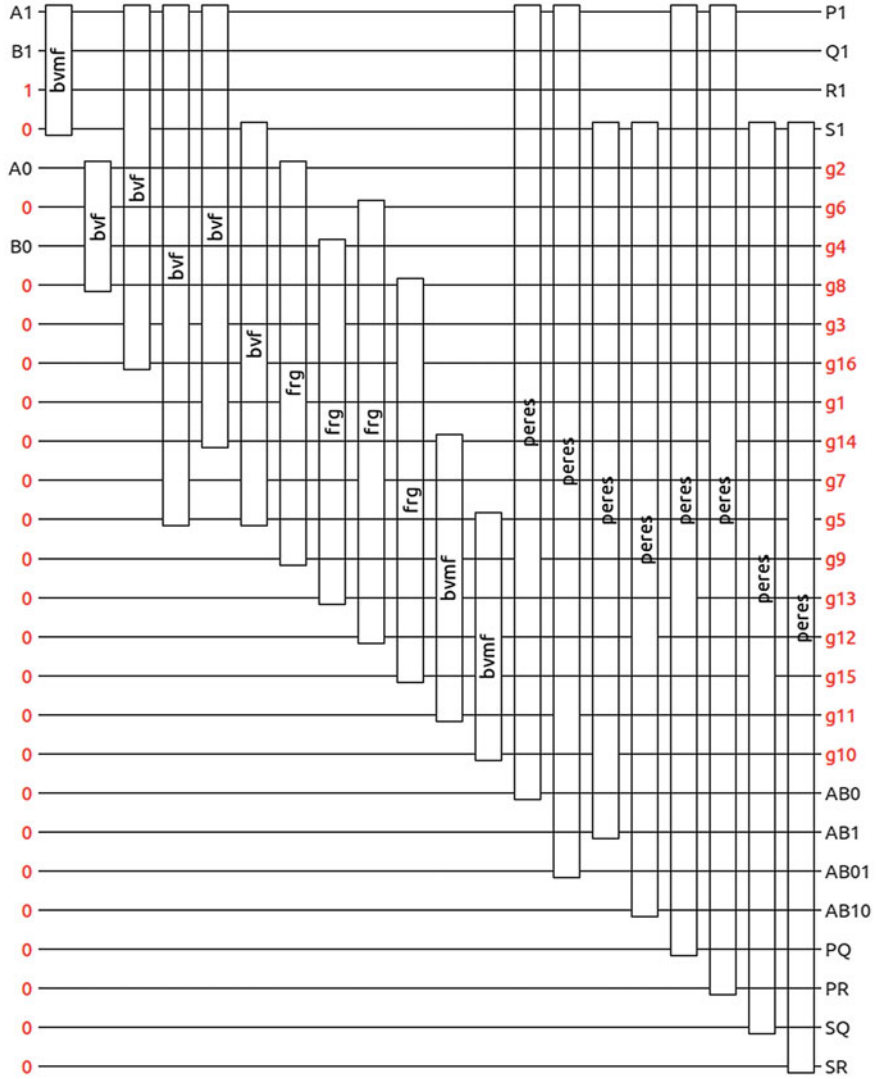


Fig. 8 Block diagram of two-bit comparator

implementation of two-bit comparator is shown in the Fig. 9. The simulation results of two-bit comparator circuit using Dinotrace waveform viewer are shown in Fig. 10. The resulting parameters of two-bit comparator using BVMF gates are as shown in Table 2.

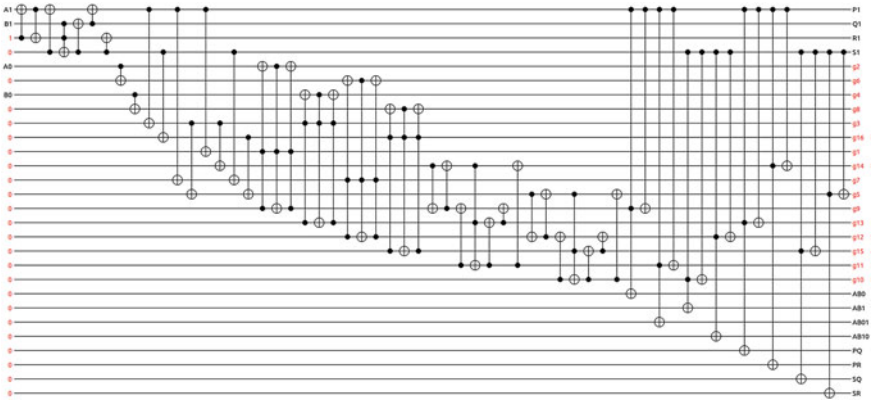


Fig. 9 Toffoli cascade structure of two-bit comparator

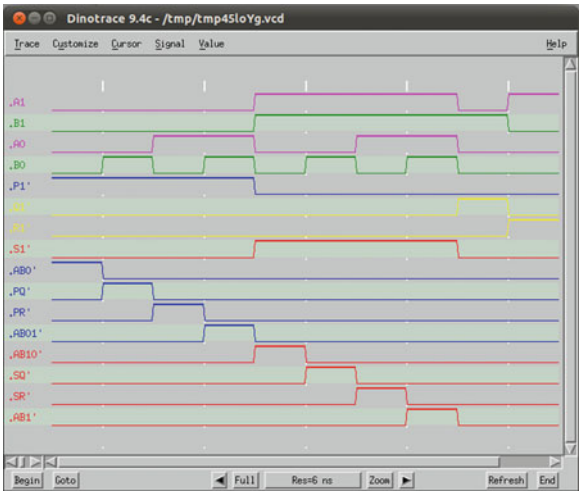


Fig. 10 Simulation results of two-bit comparator

Table 2 Parameters of two-bit comparator circuit

Parameters	Value
Gate count	59
Line count	28
Quantum cost	151
Total number of gates	20
Number of garbage outputs	16
Number of constant inputs	24
Transistor count	656

5 Conclusions and Future Scope

The focus of this paper is the application of a reversible logic gate to realize comparator using a multifunction reversible logic gate. In the present work, a multifunction reversible logic gate is designed keeping in view the optimization factors of the reversible circuits and is synthesized using exact synthesis. Toffoli gate circuits have been demonstrated as a promising alternative to achieve minimal reversible circuits than that achieved using heuristic synthesis approach, namely the transformation-based method. The quantum costs directly depend on the used Toffoli gates and its control inputs. In the present circuit, there are 36 single-control input Toffoli gates and 23 two-control input Toffoli gates. Cost of a single-control input Toffoli gate (Feynman gate) is 1 and that of a two-control input Toffoli gate is 5. The total circuit cost is equal to $36 + 115 = 151$. Other synthesis methods and cost reduction techniques are under investigation as a future work.

Acknowledgment The authors wish to thank the ECE Department of BMS College of Engineering, Bangalore, Karnataka, India, for supporting this work.

References

1. Landuaer R (1961) Irreversibility and heat generation in the computing process. IBM J Res Dev 3:183–191
2. Bennett CH (1973) Logical reversibility of computation. IBM J Res Dev 17(6):525–532
3. Frank MP (2005) Approaching the physical limits of computing, 35th international symposium on multiple-valued logic, vol 3, pp 168–185
4. Toffoli T (1980) Reversible computing, tech memo MIT lab for computer science
5. Fredkin E, Toffoli T (1982) Conservative logic, vol 21, pp 219–253
6. Soeken M, Frehse S, Wille R, Drechsler R (2011) RevKit: an open source toolkit for the design of reversible circuits, lecture notes in computer science, vol 7165, pp 64–76, in reversible computation. RevKit is available at www.revkit.org
7. Feynman R (1985) Quantum mechanical computers. Optic News 11:11–20
8. Wille R, Grobe D, Dueck GW, Drechsler R (2008) RevLib: an online resource for reversible functions and reversible circuits. Int Symp Multiple-Valued Logic (ISMVL) 220–225
9. Bhagyalakshmi HR, Venkatesha MK (2011) Design of a multifunction BVMF reversible logic gate and its applications. Int J Comput Appl 32(3):36–41
10. Nielsen M, Chuang I (2000) Quantum computation and quantum information. Cambridge University Press, Cambridge
11. Wille R, Drechsler R (2010) Towards a design flow for reversible logic. Springer, Heidelberg
12. Fazel K, Thorten M, Rice JE (2007) ESOP-based Toffoli gate cascade generation. In: proceedings of the IEEE Paci_c rim conference on communications, computers and signal processing (PACRIM), Victoria, BC, 22–24 Aug 2007, pp 206–209
13. Nayeem NM, Rice JE (2011) Improved ESOP-based synthesis of reversible logic, RM work shop, Tuusula, 25–26 May 2011, vol 32(3), pp 57–62
14. Smolin JA, Di Vincenzo DP (1996) Five two-bit quantum gates are sufficient to implement the quantum Fredkin gate. Phys Rev 53:2855–2856

Design and Development of Biosignal Conditioning, and Processing Electronic System by Employing Analog Devices Chip (AD-549) and LabVIEW

A. Usha, B. Ramachandra and M. S. Dharmaprakash

Abstract This research work is focused on the design of bio-sensor, bio-signal conditioning, and processing electronic system by employing analog devices (AD-549) chip and LabVIEW. A simple, low-cost data conditioning and processing system for very low power (of the order of pico-amperes) signal is designed and effectively implemented in the laboratory. This is a new innovative approach which is developed using novel signal conditioning (SC) and processing technique and can be applicable for biological applications. Design and development of electronics circuitry using AD549 for effective signal conditioning has been implemented for possible applications in future technology. A new attempt for biosignal processing using LabVIEW software and hardware module appears to be the better solution especially for biomedical applications. The real-time signal from any sensor or transducer has been accessed (simulated), conditioned, processed, studied, and monitored for values in the range of pico-amperes and nano-volts by using MULTISIM software and PXI-4071 digital multimeter (DMM card). In this work, the signal acquired from the sensor has been conditioned using analog devices chip, AD549 IC, and control signals for this system for further processing are generated using LabVIEW (Laboratory Virtual Instrument Engineering Bench Work) program. The proposed sensor and novel signal conditioning and processing circuit have several advantages such as high sensitivity, greater resolution, better accuracy, precision, and wide application in various biomedical fields.

A. Usha (✉)
Department of Electrical and Electronics Engineering, IEEE Member,
BMSCE, Bangalore, India
e-mail: ushaajoshi@yahoo.com

B. Ramachandra
Department of Electrical and Electronics Engineering, PESCE, Mandya, India
e-mail: bramachandra1@yahoo.co.in

M. S. Dharmaprakash
Department of Chemistry, BMSCE, Bangalore, India
e-mail: msdharmaprakash@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_80,

Keywords Biosensor · Signal conditioning and processing technique · AD-549 chip · LabVIEW

1 Introduction

Owing to the remarkable advance in Integrated Chip (IC) technologies, there are many electronic chips (released by Analog Devices) with built-in advanced circuits along with their development boards. This offers a starting point for researchers to customize their systems. Further, advantages of this approach are simplicity and multichannel capacity. In this investigation, a Labview-based single or multichannel functioning system is proposed. Using the data conditioning plug-in board and LabVIEW software, novel electronic system is proposed to display, analyze and store the data in digital form [1]. This helps to increase the efficiency of process and monitor the signal. The parameters composed amplitude, shape, phase, and wavelength of the responses are displayed and analyzed by the LabVIEW program. Furthermore, the LabVIEW can be configured for appropriate commands or instructions to generate timing signals for controlling the conditioned signal output of the sensor or transducer [2]. Signals may be digital or analog depending on the transducer used. Signal conditioning may be necessary if the signal from the transducer is not suitable for further processing. The signal often needs to be amplified, filtered, or demodulated. Various other examples of signal conditioning circuitry include bridge configuration, providing current, or voltage excitation to the sensor, isolation, and linearization. For transmission purposes, single-ended analog signals, which are more susceptible to noise can be converted into discrete or digital form. Once digitized, the signal can be encoded to achieve efficient transmission and reduce errors [3].

Ideal biosensor for use to detect the presence of glucose or leukocyte counts in the blood sample/analyte has been effectively designed and this unit can be easily incorporated onto a suitable SC and processing electronic system [4]. The sensor has excellent sensitivity with atmospheric temperature and pressure which is combined with a quick response time and consumes only a microliter of the blood sample.

Signal conditioning is an effective technique by which an analog signal is manipulated in such a way that it satisfies the requirements of the next stage for further processing. In applications, it is usually required to have a sensing stage (which consists of sensor), a signal conditioning circuit (operational amplifier, AD-549) is generally employed to carry out the amplification of the signal in the circuit and a processing unit (normally carried out by an ADC, microprocessor, or microcontroller). Signal inputs accepted by various sources include DC voltage and current, AC voltage and current, frequency, and electric potential. Sensor inputs can be accelerometer, thermocouple, thermistor, resistance thermometer, strain gauge, LPG sensor, carbon monoxide/CO, or biosensor [5]. Specialized data signals include encoder, counter or tachometer, timer or clock, relay or switch, and

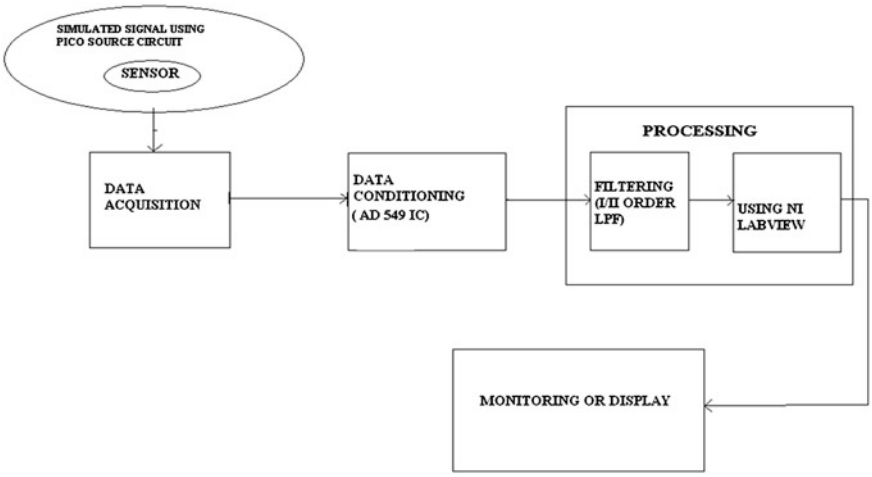


Fig. 1 Block diagram of the system

some particular inputs. Outputs expected for SC equipment can be voltage, current, frequency, or resistance and other specialized signal outputs (Fig. 1).

LabVIEW is a graphical programming language that uses icons instead of lines of text to create applications. In contrast to text-based programming languages, where instructions determine program execution, LabVIEW uses dataflow programming, where the flow of data determines execution. In LabVIEW, we build a user interface with a set of tools and objects. The user interface is known as the front panel. We then add code using graphical representations of functions to control the front panel objects. The block diagram contains this code. In some ways, the block diagram resembles a flowchart. We can purchase several add-on software and hardware toolsets for developing specialized applications.

LabVIEW programs are called virtual instruments or VIs, because their appearance and operation imitate physical instruments, such as oscilloscopes and multimeters. Every VI uses functions that manipulate input from the user interface or other sources and display that information or move it to other files or other computers.

2 Smart Sensor Design

Design of the smart sensor has been optimized using conducting polymers and nano-composites. The conducting type polymers are mechanically and electrically compatible with such conductors as copper and such semiconductors as silicon. Typical conductivities of the latest conducting polymers are in the range of 10^{-4} and 107 s/meter, as compared with 1.1×107 s/m for copper and 10^{-2} to 1 s/m for lightly doped silicon both at room temperature and pressure. The electrical

conductivity of any conducting polymer can be varied over a wide range, depending on the amount and reactivity of the dopant or oxidant used.

Using sol-gel or drop-cast technique, thin film of the sensor is synthesized, which constitutes mainly the novel conducting polymers such as polyaniline (PANI), thin film with nano-materials that is, gold nano-particles coated on the surface of the sensor. With maximum surface area, that is, more number of active sites which can assist the enzyme or antibody to enhance the immobilization of enzyme while interacting with the specific analyte or the sample. The analyte employed here is the blood sample, in which the Glucose or Leukocyte counts are initially traced and identified by observing and analyzing various performance characteristics of the biosensor at room temperature and pressure by using the two-probe technique. Further, synthesis and characterization of the biosensor is carried out at IISC NANO-Electronics Center under INUP Program (Indian NANO-Electronics User Program) funded by MCIT and Department of Science and Technology (DST), Government of India.

3 Signal Conditioning Electronics

3.1 Pico-amps Current Source & Signal Conditioning Circuit Implementation

The circuit effectively designed, constructed, and tested to obtain pico-current in the range of 0–1,000 pA is as shown in the below Fig. 2. The hardware circuit connection was found better to implement on the Teflon board and the same has been implemented in this research work. The signal conditioning circuit shown in the below Figs. 3 and 4 reveals that the generated pico-current source is interfaced with an amplifier circuit of the analog device chip, that is, AD549 IC. This AD549 IC chip in turn converts the obtained pico-current values into millivolts range, which is suitable for further processing techniques. It is more convenient to

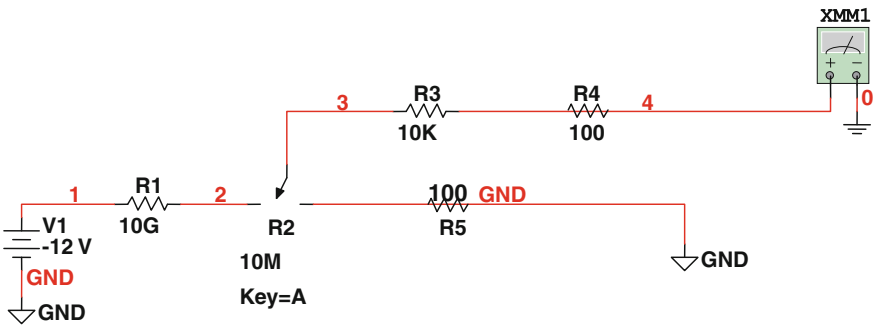


Fig. 2 Practical pico-current source circuit

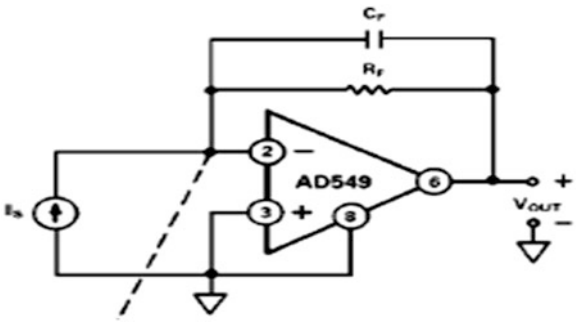


Fig. 3 Preamplifier circuit for conditioning signal

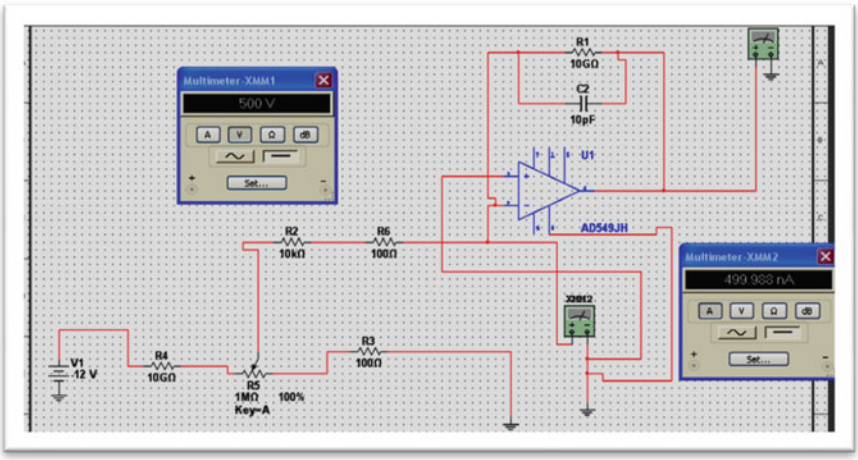


Fig. 4 Signal conditioning circuit using AD549

process the signals which are at higher power values rather than the lower power values which is in the order or nano- or pico-current/voltage levels. Hence, the conditioning circuit designed and constructed using AD549 IC plays a vital role in converting, conditioning, and testing of very low current value signals acquired from any sensor or transducer.

Precautions to be taken during the hardware set up:

The tolerance values of the components should be very low.
The tolerance value of the used 10-G resistor is 1 %.
The hardware setup implemented on the Teflon board should be shielded from external fields.
Care should be taken during the soldering of the AD549 IC chip.
The connection between the components should be within the shorter distances as much as possible.
The whole metal enclosed hardware setup should be wrapped by aluminum sheet to prevent external interferences.

The connecting wires employed should have negligible or minimum lead resistances.
For better implementation of the hardware circuit of AD549 IC, islets are recommended while soldering the IC with the peripheral components on the Teflon board.

A simple pico-current source is constructed by −12-V output from an ideal voltage source and a voltage divider to adjust the voltage, V. The voltage divider consists of a 10-GΩ resistor in series with a 10-MΩ potentiometer. The 100-Ω resistor with a 10-K resistor in series is connected to the moveable contact (wiper) of the potentiometer. The other end of the potentiometer is connected to ground through a 100-Ω resistor. The output is variable between 0 to 1,000 pA. In this work, signal conditioning circuit is designed and implemented which constitutes mainly current to voltage converter and a basic amplifier using analog devices AD549 IC chip. Usually, the output of the sensor is a weak current signal which will be in the range of pico- or nano-amperes. This signal in turn is converted to voltage with required amplification and reducing all offsets. The current to voltage conversion is achieved by using AD-549 chip. There are many parameters and factors to be followed while selecting a specific opamp for the required application.

The chip AD-549 is an operational amplifier (OPAMP) with very low input bias current, input offset voltage and input offset voltage drift are laser trimmed for precision performance. The ultra-low input current of the opamp is achieved with top gate JFET technology, a process development exclusive to analog devices. This technology allows fabrication of extremely low input current JFETs compatible with a standard junction–isolated bipolar process. The 1015-Ω common-mode impedance, which results from the bootstrapped input stage, ensures that the input current is essentially independent of the common-mode voltage.
The AD549 is suited for applications requiring very low input current and low input offset voltage. It excels as a preamp for a wide variety of current output transducers, such as photodiodes, photomultiplier tubes, or oxygen sensors. The AD549 can also be used as a precision integrator or low-droop sample-and-hold circuit applications. The AD549 chip is compatible with standard FET and electrometer opamps, allowing designers to upgrade the performance of present systems at a minimum additional expenditure. The AD549 is available in a TO-99 hermetic package. The AD549 case is connected to Pin 8, thus, in turn the metal case can be independently interfaced to a point at the same potential as the input terminals, minimizing stray leakage to the case. The AD549 is available in four performance grades. The J, K, and L versions are rated over the commercial temperature range from 0 to +70 °C. The S grade is specified over the military temperature range from −55 to +125 °C.

In this research work, AD549 JH version (subversion of AD549 J) is selected and implemented efficiently to achieve signal conditioning of pico-current values, which is designed and implemented using discrete hardware components in the laboratory.

3.2 Implementation of Bio-signal Processing Electronic System

3.2.1 Filtering Using Hardware Device

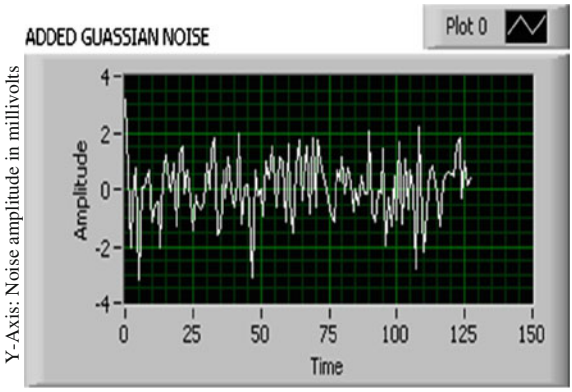
Filtering is one of the most commonly used signal processing techniques and, in this work, is effectively implemented by employing LabVIEW Signal Express software and PXI-4071 DMM hardware card. Signal processing systems can filter unwanted signals or noise components from the accessed and measured data. A noise filter on low-rate or slowly changing signals such as temperature to eliminate higher frequency signals that can reduce signal accuracy has been employed. A common use of a filter is to eliminate the noise from 50- or 60-Hz AC power line. A low-pass filter of 4 Hz removes the 50- or 60-Hz AC noise from signals sampled at low rates. Further, a low-pass filter eliminates all signal frequency components above the designed cutoff frequency. Many signal processing modules have low-pass filters that have software-selectable cutoff frequencies which ranges from 10 Hz to 25 kHz.

3.2.2 Windowing Technique

Spectral leakage is a phenomenon whereby the measured spectral energy appears to leak from one frequency into other frequencies. It occurs when a sampled waveform does not contain an integral number of cycles over the time period during which it was sampled. The technique used to reduce spectral leakage is to multiply the time-domain waveform by a window function. Discrete Fourier transform (DFT) and fast Fourier transform (FFT) are mathematical techniques that resolve a given signal into the sum of sines and cosines, which is the basis for spectrum analysis. Using the DFT/FFT sampling, a non-integer number of cycles, such as 7.5 cycles, return a spectrum in which it appears as if the energy at one frequency leaks into all the other frequencies because the FFT assumes that the data are the single period of a periodically repeating waveform. The artificial discontinuities appear as very high frequencies that were not present in the original signal, because these frequencies are higher than the Nyquist frequency and they appear aliased between 0 and $fs/2$ range.

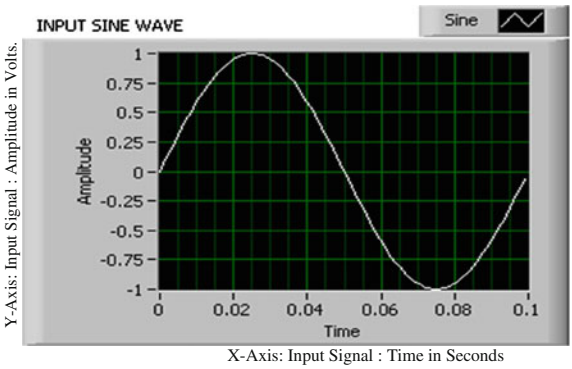
The type of window to use depends on the type of signal that is acquired and also based on the applications. Choosing the correct window requires proper knowledge of the analyzing signals.

In this research work, the conditioned signal further processed by designing the second-order low-pass filter (FIR Butterworth filter) to filter out all high-frequency noise components and interferences. Finite impulse response second-order Butterworth filter has been optimized for the real-time sampled and conditioned signals as the signals are generated/acquired, conditioned, and amplified using the specific electronic circuits (Fig. 5).



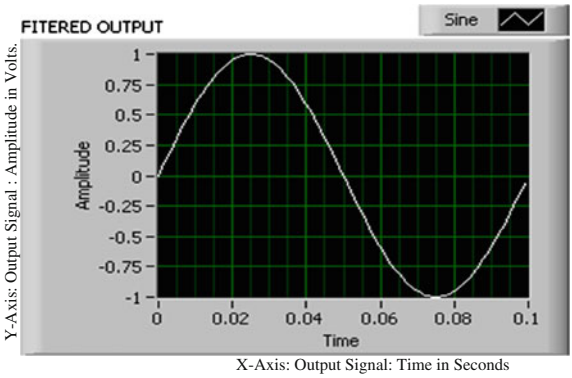
Y-Axis: Noise amplitude in millivolts

X-Axis: Time in milliseconds



Y-Axis: Input Signal : Amplitude in Volts.

X-Axis: Input Signal : Time in Seconds



Y-Axis: Output Signal : Amplitude in Volts.

X-Axis: Output Signal: Time in Seconds

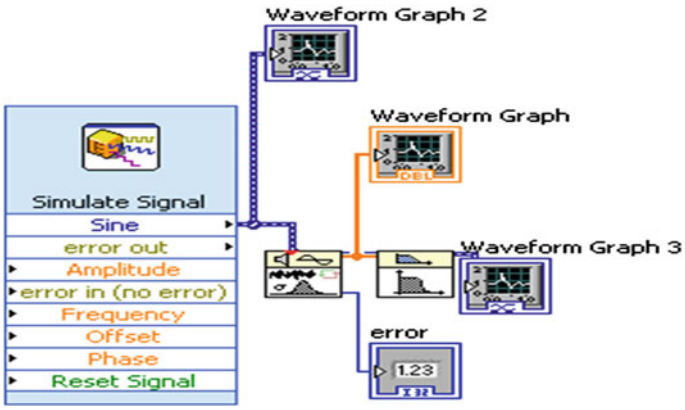


Fig. 5 Block diagram of Butterworth filter using LabVIEW

4 Experimental Results and Discussion

The designed pico-current source and SC circuit has been successfully simulated using (discrete components and AD549 chip) MULTISIM software in the laboratory. Later, the practical pico-current source circuitry which produces the pico-current in the range of 0–1,000 pA is effectively constructed and the results of practical discrete pico-source current circuit in the laboratory are as listed in the below Table 2. Further, practical results of the conditioned output voltage values in the laboratory are as listed in the below Table 1 and the conditioning circuit results at the output stage ranges from 0 to 1,000 mV.

The processing of the AC/DC signal, that is, filtering of the conditioned output signal of the AD549 IC, has been efficiently implemented using LabVIEW Signal Express which is interfaced with the PXI-4071 DMM (digital multimeter)

Table 1 Signal-conditioned output voltage values

Sensor output ΔI (pico-amps)	Practical signal-conditioned output (millivolts)
0	0
100	121
200	245
300	312
400	419
500	538
600	629
700	739
800	845
900	927
1000	1037

Table 2 Pico-source output voltage values

Potentiometer values (M Ω) (%)	Practical pico-current source output (pico-amps)
100	10
90	50
80	120
70	180
60	260
50	320
40	390
30	550
20	750
10	950

hardware card. A second-order low-pass FIR Butterworth filter is designed and constructed in the block diagram window of the LabVIEW with the simulated signals and also the added Gaussian noise which is filtered and reconstructed again the original simulated or acquired or conditioned AC/DC signal (Table 2).

5 Conclusion

In this research work, a simple, low-cost data conditioning and processing electronic system for very low power (of the order of pico-amperes) signal is designed and constructed. The real-time signal from any sensor or transducer has been acquired or simulated, constructed, studied, and monitored for values in the range of pico-amperes and nano-volts by using MULTISIM software and PXI-4071 digital multimeter (DMM card). Further, the simulated design is implemented on the Teflon board to obtain the actual pico-current values. The generated pico-source current values are conditioned using analog devices IC, AD549 chip to obtain the output in millivolts range, which is more suitable for processing techniques. The filtering of the conditioned signal has been done using LabVIEW and LabVIEW Signal Express software, which is a graphical user interface and the processing techniques such as analog to digital conversion and digital to analog conversions are achieved. The design of the various filter types are implemented using the LabVIEW block diagram, where the filters are constructed and the corresponding waveforms are obtained in the front panel window. It also shows the variations in the resulted filtered output with the corresponding change in the sampling frequency, cutoff frequencies, and the order of the filter. The hardware setup has been efficiently implemented and processed to obtain the required and specified results using PXI-4071 DMM hardware card. Synthesis, characterization, and testing of the biosensor are carried out at IISc NANO-Electronics Center and the results of which are compatible with this signal conditioning and processing electronic system.

Acknowledgment The authors sincerely thank all IISC-INUP professors for giving an opportunity to execute the research work, sensor synthesis, characterization, and testing at IISC NANO-Electronics Center under INUP Program (Indian NANO-Electronics User Program) funded by Ministry of Communication and Information Technology (MCIT) and Department of Science and Technology (DST), Government of India. They acknowledge the valuable contributions of Dr. Pranesh, Dr. Ravishankar Deekshit, and also acknowledge Nancy Subhashini for her tremendous work in this field which further succeeded in obtaining the fruitful results.

References

1. Cheol Jeong (2007) Development of bio signal measurement system for vehicles, department of biomedical engineering, Yonsei University, international conference on convergence information technology, Korea
2. Upasani MDE (2010) Wireless recorder for bio-medical signals. Int J Comput Appl 1(18):0975–8887, SGREFS, COEM, Ahmednagar
3. Asaoka T (2008) Development of the device to detect human's bio-signals by easy sensing, 30th annual international IEEE EMBS conference, student member, IEEE, Yoshiaki KANAEDA, student member, IEEE, Kazushige MAGATANI, member, IEEE, Vancouver, British Columbia, Canada
4. Shojaei-Baghini M (2009) A lower power and compact analog CMOS processing chip for portable ECG recorders, Department of electrical engineering, Indian institute of technology, Mumbai, India
5. Baglio S (2003) Bio–geochemically inspired capacitive sensors for heavy metals pollution monitoring. IEEE Trans Instrum Meas IEEE 52(5):1474–1481

Design of an Optimal Yaw Damper for 747 Jet Aircraft Model

Seema Singh and T. V. Rama Murthy

Abstract The theory of optimal control is concerned with operating a linear dynamic system at a minimum cost. It also provides the best possible performance with respect to given measure of performance index. This paper presents the design of an optimal control law for yaw damper of an aircraft model. This modern controller based on linear quadratic regulator (LQR) technique is stable, robust, and optimized with respect to energy. It is compared and analyzed with classical controller available in the literature. The classical controller uses washout filter which is well known in aircraft control design. The comparative assessment is based on time response specification performance for a yaw control system of an aircraft. The LQR-based controller has effectively taken care of Dutch roll mode which is an area of concern in lateral dynamics of aircraft. The performances of yaw control systems are investigated and analyzed based on common criteria of impulse response in order to identify which control strategy delivers better performance with respect to the desired bank angle and yaw rate. Simulation is carried out for 747 jet aircraft model using MATLAB and Simulink. According to simulation results, LQR controller delivers better performance than classical controller.

Keywords Aircraft · Flight control · LQR controller · Yaw damper · Dutch roll · Lateral dynamics

S. Singh (✉)
Department of Electronics and Communication, BMS Institute of Technology, Bangalore, India
e-mail: seema.singh1980@gmail.com

T. V. Rama Murthy
Department of Electronics and Communication, Reva Institute of Technology and Management, Bangalore, India
e-mail: drtvramamurthy@revainstitution.org

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_81,

Nomenclature

x	State-space matrix of aircraft model, [4 × 1]
u	Input matrix, [2 × 1]; rudder, aileron (radian)
β, ψ, θ, ϕ	Aircraft states; sideslip angle (radian), yaw rate (rad/sec), roll rate (rad/sec), bank angle (radian)
y	Output matrix, [2 × 1]; ψ, ϕ
A, B, C, D	State-space specifications of aircraft model
J	Performance index/quadratic cost function
Q, R	Weighting matrices of J

1 Introduction

Aircraft design relies heavily on automatic control system which monitors and controls many aircraft subsystems. The development of automatic control system has played an important role in the growth of civil and military aviation. Automatic control system in aircraft aids the flight crew in navigation, flight management, and augmenting the stability characteristic of the airplane.

Dutch roll mode is primarily excited by rudder, while aileron deflection produces roll rate and roll angle, with some Dutch roll oscillation superimposed on the aircraft motion. However, considerable Dutch roll does show up in the yaw rate and sideslip response of the aircraft [1]. This causes the airplane to roll as well as yaw. Dutch roll can sometimes happen in bad weather or for other reasons. If a gust pushes the nose of the aircraft to the side, it will also roll the aircraft. As the aircraft corrects the yaw, the plane will roll the other way. This will keep on going until the pilot corrects it or the yaw damper sends inputs through the autopilot to the rudder to correct it. Dutch roll is a type of aircraft motion, consisting of an out-of-phase combination of "tail-wagging" and rocking from side to side. This yaw–roll coupling is one of the basic flight dynamic modes (others include phugoid, short period, and spiral divergence). This motion is normally well damped in most light aircraft, though some aircraft with well-damped Dutch roll modes can experience degradation in damping, as airspeed decreases and altitude increases. Dutch roll stability can be artificially increased by the installation of a yaw damper. A yaw damper is a device used on many aircraft to damp the rolling and yawing oscillations due to Dutch roll mode. The use of the yaw damper helps to provide a better ride for passengers. For some aircraft, yaw damper ensures that the aircraft stability remains within the certification values.

In classical control design, the usual method of damping the Dutch roll is to detect the yaw rate with a rate gyro and use this signal to deflect the rudder. A first-order washout filter is used for Dutch roll damper [1]. A classical control design which gives satisfactory impulse response for both ψ (yaw rate) and ϕ (bank angle) is discussed in [2].

This paper presents the design of optimal yaw damper which uses linear quadratic regulator (LQR) technique of modern control theory. LQR uses state-space approach of the plant for controller design. The 747 jet aircraft model used in this paper is completely controllable and observable which is required for LQR design. The notion of optimality is closely tied to MIMO control system design. For MIMO plants, only LQR satisfies both the performance index and stability [3].

The LQR generates a static gain matrix K, and hence, the order of the closed-loop system is as same as that of the plant.

LQR produces a robust controller as it achieves infinite gain margin and also guarantees phase margin $\gamma \geq 60$ degrees.

This is in good agreement with the practical guidelines for control system design [4]. The performance index of LQR is

$$J = \int_0^\infty \left(x^T Q x + u^T R u\right) \mathrm{d}t. \tag{1}$$

When $R \gg C^T Q_1 C$, where Q_1 is a constant, it gives expensive control as the cost function is dominated by the control effort u, and so the controller minimizes the control action itself. The expensive control solution puts stable closed-loop poles at the mirror images of the unstable plant poles. When $R \ll C^T Q_1 C$, it gives cheap control as the cost function is dominated by the output errors y, and there is no penalty for using large u [4].

Optimal control solution discussed here uses an automated design procedure where optimization is done for energy consumption. LQR technique is well used by designers for aircraft control systems. Wahid and Rahmat [5] found LQR controller to perform better than fuzzy controller for pitch control system for aircraft. Zhichao et al. [6] used LQR technique for control design of 3-DOF helicopter model. Nasir et al. [7] carried out performance comparison between LQR and PID controllers for an inverted pendulum system and found better results with LQR controller. Dinesh kumar [8] used LQR controller for design of Navion aircraft pitch control system. Kulczycki [9] carried out controller design for autonomous airships using sequential loop closure (SLC) and LQR methods. Zheng et al. [10] used LQR for optimal control of an aircraft of aileron-less folding wing. Zheng et al. [11] used LQR state-feedback control method to design the rate damping and stabilizing loop in order to achieve desirable dynamic characteristics for lateral control for an aircraft of folding wing. Usta et al. [12] and Vishnu et al. [13] compared LQR controller and fuzzy logic controller for aircraft roll control system and aircraft yaw control system, respectively. Both found LQR to be better than fuzzy logic controller. The authors [12, 13] discuss the optimization of only one state and compare the results with that of fuzzy controller; there is no comparison of their results with that of classical controller. The significant contribution of this work lies in the discussion of LQR design for MIMO system (lateral dynamics) and comparison of the performance with classical control design [2]. The results reported here show improved performance with that discussed in [2].

2 Flight Control System

Two types of dynamical equations are present for an aircraft. The lateral dynamic equations of motion represent the dynamics of aircraft with respect to lateral axis. Longitudinal dynamic equations of motion represent the dynamics of aircraft with respect to longitudinal axis. Lateral dynamics includes yaw, roll, and sideslip motions of aircraft. Pitching motion comes under longitudinal dynamics. The control surfaces of aircraft are shown in Fig. 1.

2.1 747 Jet Aircraft Model

The lateral dynamics of 747 jet aircraft model has four states, two inputs, and two outputs. The simplified trim model of the aircraft during cruise flight at MACH = 0.8 and $H = 40{,}000$ ft is given in state-space form in Eqs. 2 and 3.

$$\dot{x} = Ax + Bu \tag{2}$$

$$y = Cx + Du \tag{3}$$

$A = [4 \times 4]$, $B = [4 \times 2]$, $C = [2 \times 4]$, $D = [2 \times 2]$ matrices.

$$A = \begin{bmatrix} -0.0558 & -0.9968 & 0.0802 & 0.0415 \\ 0.598 & -0.115 & -0.0318 & 0 \\ -3.05 & 0.388 & -0.465 & 0 \\ 0 & 0.0805 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0.0729 & 0 \\ -4.75 & 0.00775 \\ 0.153 & 0.143 \\ 0 & 0 \end{bmatrix}$$
$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{4}$$

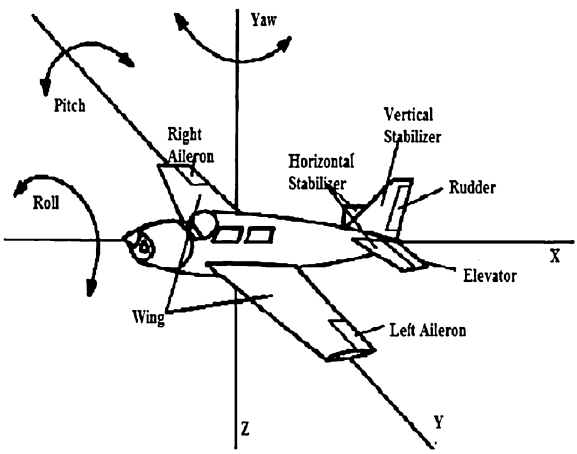


Fig. 1 Yaw, roll, and pitch motions of aircraft [13]

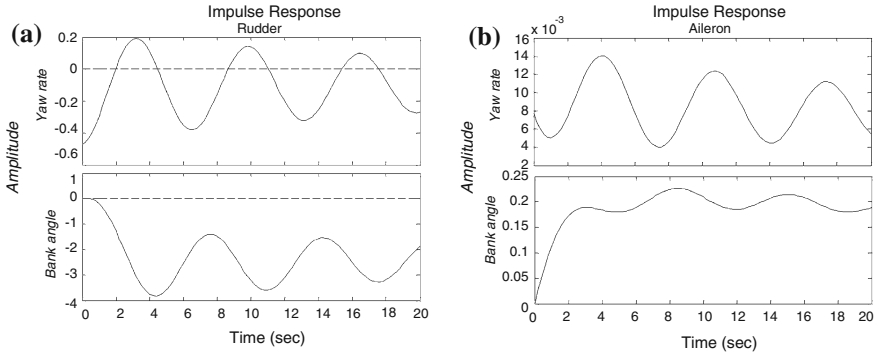


Fig. 2 Impulse response to both outputs [yaw rate and bank angle]: **a** from input 1: rudder and **b** from input 2: aileron. Dutch roll can be seen in second graph between aileron and bank angle

The poles of this model are located at -0.0329 ± 0.9467i, -0.5627, and -0.0073. This model has one pair of complex conjugate lightly damped poles. They correspond to "Dutch roll mode".

The open-loop model is marginally stable and its characteristics are shown in Fig. 2. Oscillations are seen in all the four plots.

2.2 Linear Quadratic Regulator Controller

Using state-space methods, it is relatively simple to work with an MIMO system. The system can be stabilized using full-state feedback. For the derivation of the LQR, we assume the plant to be written in state-space form $\dot{x} = Ax + Bu$, and all the n states of x are available for the controller. The feedback gain is a matrix K, implemented as

$$u = -K(x - \chi) \tag{5}$$

where χ represents the vector of desired states and serves as the external input to the closed-loop system. The system dynamics is then written as follows:

$$\dot{x} = (A - BK)x + BK\chi \tag{6}$$

The $\|A$-matrix$\|$ of the closed-loop system is $(A - BK)$, and the $\|B$-matrix$\|$ of the closed-loop system is BK. The closed-loop system has exactly as many outputs as inputs, n. The column dimension of B equals the number of channels available in u and must match the row dimension of K. Pole placement is the process of placing the poles of $(A - BK)$ in stable, suitably damped locations in the complex plane.

3 Linear Quadratic Regulator Controller Design

The idea is to transfer the designer's iteration on pole locations as used in full-state feedback (pole placement method) to iterations on Q and R matrices (LQR method) which contributes to the cost function, J. This method determines the feedback gain matrix that minimizes J in order to achieve some compromise between the use of control effort, the magnitude, and the speed of response that will guarantee a stable system.

Algorithm For a given system in state-variable form, $\dot{x} = Ax + Bu$, a set of full-state feedback gains are calculated that minimize the quadratic cost function as given in Eq. 1 and rewritten here for further discussion.

$$J = \int_0^\infty \left(x^T Q x + u^T R u\right) \mathrm{d}t$$

The weighting matrices, Q and R, are symmetric and positive semi-definite and positive definite, respectively. Second term on RHS of the above equation accounts for the expenditure of energy on the control efforts. The matrices Q and R determine the relative importance of the error and the expenditure of this energy. The resulting control law is

$$u(t) = -K\,x(t) \text{ where, } K = R^{-1}B^T P \tag{7}$$

and P is a symmetric positive-definite matrix satisfying the matrix Riccati equation as given in Eq. 8.

$$PA + A^T P + Q - PBR^{-1}B^T P = 0 \tag{8}$$

As will be seen later, the linear control law given by the above equation is the optimal control law. Therefore, if the unknown elements of the matrix K are determined so as to minimize performance index, then K is optimal for any initial state x (0). Figure 3 shows the optimal configuration.

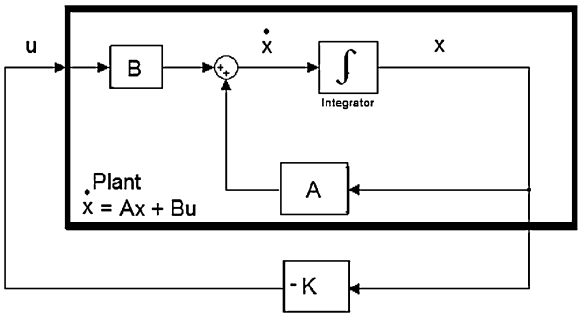


Fig. 3 Quadratic optimal regulator system

3.1 Weight Matrix Selection

The LQR design makes use of the given weight matrices Q and R such that the performances of the closed-loop system can satisfy the desired requirement. The selection of Q and R is weakly connected to the performance specifications, and a certain amount of trial and error is required with an interactive computer simulation before a satisfactory design is arrived at.

3.2 Yaw Damper Design

As the matrices Q and R are symmetric, there are 10 distinct elements in Q and 3 distinct elements in R which need to be selected. Further, the matrices Q and R should satisfy the positive definiteness. One practical method is to set Q and R to be diagonal matrix such that only six out of ten elements need to be decided. The value of the elements in Q and R is related to its contribution to the cost function J. R matrix is chosen to be identity matrix to keep input value unaffected for controller design. As system states are affected with diagonal values of Q matrix, it is judicially chosen to achieve the required stability in the specified time. In designing the LQR controller, MATLAB function '*lqr*' is used to determine the value of K as given below.

$$[K, S, e] = \text{lqr}(A, B, Q, R)$$

It gives state-feedback gain K, solution S of the associated Riccati equation, and the closed-loop eigenvalues (e), $e = \text{eig}\ (A - B * K)$. Input parameter values are A, B, $R = I$, and $Q = C^T C$, where C^T is the matrix transpose of C from state equation.

The controller can be tuned by changing the non-zero elements in Q matrix which is done in m-file.

$$R = I; \quad Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

The value of $m = 30$ was found to give the optimum impulse response. The corresponding value of matrix K is

$$K = \begin{bmatrix} 5.0488 & -7.3098 & -1.0013 & -0.6897 \\ -1.1464 & 0.5569 & 0.4776 & 0.3420 \end{bmatrix} \tag{10}$$

The complete simulation diagram based on LQR technique is shown in the Fig. 4, which was simulated in MATLAB/Simulink. A, B, C matrices and K matrix are given in Eqs. 4 and 10, respectively. Since output consists of only two states,

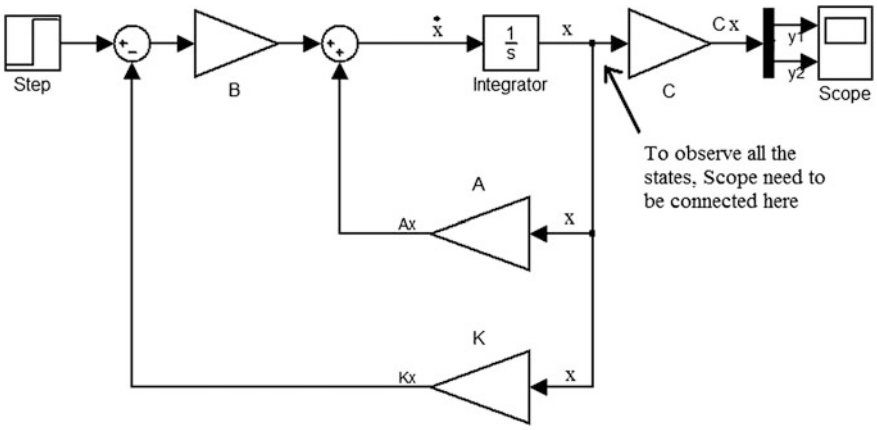


Fig. 4 Simulink model for aircraft state-space model and LQR controller

that is, yaw rate and bank angle, the scope in Fig. 4 shows two outputs. However, to see all the four states, separate scopes are needed before 'C' block in the simulation diagram. Results discussed in the next section show plots of two outputs (ψ, ϕ) for both the inputs (rudder and aileron).

4 Discussion of Results

Closed-loop responses of both LQR and classical control are quite stable compared to open-loop responses as seen in Fig. 2. Figure 5 shows comparison of results from classical controller (dash dot line) using washout filter with $K = 2.34$ and LQR controller (regular line). Result shows improvement as a whole. For all the

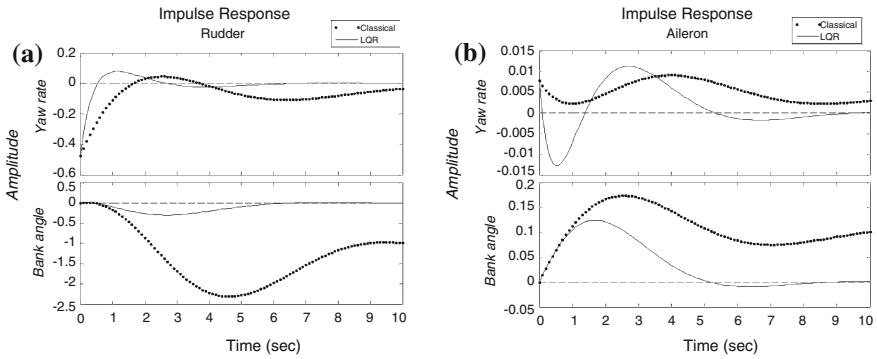


Fig. 5 Impulse response to both outputs (yaw rate and bank angle): **a** from input 1: rudder, and **b** from input 2: aileron

Table 1 Performance comparison of two methods

Performance Characteristics	(o/p,i/p)	LQR	Classical control (Washout filter)	(Output, Input) pair details
Settling Time (sec)	S(1,1)	2.2192	8.4318	S(1,1) From input 1(rudder) to output 1(yaw rate) S(2,1) From input 1(rudder) to output 2(bank angle) S(1,2) From input 2(aileron) to output 1(yaw rate) S(2,2) From input 2(aileron) to output 2(bank angle)
	S(2,1)	8.1664	16.2798	
	S(1,2)	8.4692	15.2409	
	S(2,2)	7.3193	14.1263	
Error (variance)	S(1,1)	0.0021	0.0037	
	S(2,1)	0.0061	0.1815	
	S(1,2)	0.00001	0.0000025	
	S(2,2)	0.001	0.001	
Maximum peak value (Overshoot)	S(1,1)	-0.475	-0.475	
	S(2,1)	-0.2954	-2.3031	
	S(1,2)	-0.0124	0.01	
	S(2,2)	0.1245	0.1735	

cases, LQR controller achieved stability within about 2.2–8 s which is considerably faster than the classical controller (8.4–16 s).

Significant improvement in LQR controller is seen in the second plot of Fig. 5b, which shows response from aileron to bank angle. This accounts for efficient solution for Dutch roll mode. This approach has shown improvement for all the cases except for response from aileron to yaw rate. In majority of cases, the peak overshoot is less in LQR controller. Overall performance comparison can be seen in Table 1 which shows significant improvement with LQR controller for various performances listed.

Acknowledgment The authors thank the authorities of BMS Institute of Technology, Reva Institute of Technology and Management, and the Director, R & D Cell, JNTU, Hyderabad, for their encouragements.

References

1. Blakelock JH (1990) Automatic control of aircraft and missiles, 2nd ed, A wiley interscience publication
2. Case study (2011) Design of Yaw damper for 747 Jet aircraft. Mathworks Documentation
3. Scokaert POM, Rawlings JB (1998) Constrained linear quadratic regulation. IEEE Trans Autom Control 43(8):1163–1169
4. Triantafyllou MS, Hover FS (2004) Maneuvering and control of marine vehicle. Massachusetts Institute of Technology, Lecture Notes (13.49), 2004
5. Wahid N, Rahmat MF (2010) Pitch control system using LQR and fuzzy controller. In: Proceedings of IEEE symposium on industrial electronics and applications, 3–5 Oct 2010, pp 389–394
6. Liu Z, Choukri El Haj Z, Shi H (2010) Control strategy design based on fuzzy logic and LQR for 3-DOF helicopter model. Intelligent control and information processing (ICICIP), pp 262–266

7. Nasir ANK, Ahmad MA, Rahmat MF (2008) Performance comparison between LQR and PID controller for an inverted pendulum system. In: Proceedings of international conference on power control and optimization, Chiang May, Thailand, 18–20 July
8. Dinesh Kumar M (2012) Design of modern and classical controllers for navion aircraft pitch control system. In: Proceedings of international conference on computing and control engineering (ICCCE 2012)
9. Kulczycki EA, Joshi SS, Hess RA (2006) Towards controller design for autonomous airships using SLC and LQR methods. AIAA guidance, navigation and control conference and exhibit, pp 67–78
10. Wang ZJ, Guo S, Li W (2008) Modeling, simulation and optimal control for an aircraft of aileron-less folding wing. WSEAS Trans Syst Control, 3(10):869–878
11. Wang ZJ, Li W, Guo S (2008) Lateral control for an aircraft of folding wing. In: Proceedings of 8th WSEAS international conference on simulation, modelling and optimization (SMO '08), pp 99–104
12. Usta MA, Akyazi O, Akpinar AS (2011) Aircraft roll control system using LQR and fuzzy logic controller. In: Proceedings of IEEE international symposium on innovations in intelligent systems and applications (INISTA), pp 223–227
13. Nair VG, Dileep MV, George VI (2012) Aircraft Yaw control system using LQR and fuzzy logic controller. Int J Comp Appl, 45(9):25–30

Design, Implementation and Verification of CAN Node Interface to Peripheral Ports in FPGA

R. Harish and L. Swarna Jyothi

Abstract A controller area network (CAN) node interface with serial and parallel peripheral ports is suggested for debugging and configuration of CAN network. The suggested interface reduces the effort and engineering cost in debugging the N/W. Serial interface includes RS232 and USB. Parallel interface is IEEE1284 parallel mode of operation. Peripheral port aids in connecting a host for debugging and monitoring the health of a network. CAN IP, serial and parallel interface wrapper for peripheral interface are implemented in Verilog. Prototyping of the system is done on ALTERA's CYCLONE device. The complete reusable verification environment is built with components from 'e' language and Verilog.

Keywords Design · Verification · Prototyping FPGA · ASIC · Verilog · Specman-e · System verilog · CAN · USB · RS232 · Serial interface · Parallel interface

1 Introduction

Increase in the use of microcontrollers in the automotive industry leads to more control circuitry for controlling different subsystems of a vehicle. To coordinate and connect between the components of control circuitry, a network mechanism is needed. Communication protocols like controller area network (CAN) have been evolved to connect microcontrollers in an automotive network. As different

R. Harish (✉)
VTU-Research Resource Centre, Visvesvaraya Technological University, Belgaum, India
e-mail: tharishr@gmail.com

L. Swarna Jyothi
Yellamma Dasappa Institute of Technology, Bangalore, India
e-mail: swarnajyothi57@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_82,

networks need transfer of data, a gateway system that allows separate protocols to engage with each other can be used.

CAN originally developed for the European car industry is a fast, secure and cost-effective data bus for multi-master and real-time applications. In addition to automotive applications, it is suitable as a general data bus for industrial control functions [1]. CAN is a serial, asynchronous, multi-master, highly secured communication bus protocol for distributed real-time networks connecting electronic control modules, sensors, actuators in automotive and several other industrial applications [2].

Peripheral ports can be interfaced to CAN node. Peripheral node provides an interface to connect a host or personal computer to CAN node. Configuring and debugging of the complete network is done with the help of these peripheral ports. It can also be used to monitor the health of the network.

2 Automotive Field Bus System

Automotive field bus system consists of a series of nodes connected by communication paths. The nodes can be connected in bus, star, ring or hybrid topologies for message transmission in automotive field bus systems [3]. Conventional fieldbus addressing schemes use both source and destination addresses.

CAN avoids using addresses as it requires each node to transmit its messages to all other nodes in the network. Receiving nodes filter incoming messages, accepting only messages which have message types that have been registered. This approach increases network flexibility for adding or deleting nodes [4].

3 Problem Definition

Protocol is able to meet the designer and manufacturers requirements such as equipment costs, bandwidth requirements and determinism. If a protocol is unable to meet the specified requirements and [5] an alternative protocol allows these specified requirements to be met, then migration can take place between these two protocols.

Protocol converter makes all components compatible within one network so that they can communicate harmoniously. A protocol converter has to provide message translation, message routing, message monitoring, bus error management and network management all in real time with minimal transmission delay and minimum loss of messages which gives flexibility to the system.

CAN to USB converter is needed to debug the CAN network from an USB interface and is implemented on cyclone FPGA device. The DUT is designed using Verilog. The protocol convertor is needed to debug the CAN network from a PC using USB interface.

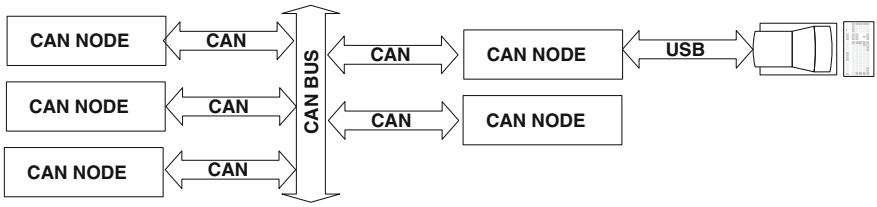


Fig. 1 Network architecture

4 Can Debugging Architecture

Network architecture for connecting the CAN nodes with peripheral ports is shown in Fig. 1. A host or a personal computer is used to configure and debug the CAN nodes and the network. Host configures the nodes with message ID parameters which aids in filtering of the message. Each node has a host interface as well as parallel or serial (RS232 and USB) interface. The network can be debugged from any node in the network, which has peripheral interface port support. The network debugging is done by monitoring the responses from the other node.

5 Design and Verification of CAN USB Converter Node

Details of CAN to USB protocol converter demo set-up are shown in Fig. 2. Different blocks of demo setup are

- CAN device
- CAN transceiver (PHY) device
- Serial/parallel interface to USB
- USB transceiver (PHY) device

6 CAN to USB Conversion

Details of data conversion process from CAN to USB are shown in Fig. 3.

Once the data is received by the convertor, it extracts the message ID, length and data and maps it on to serial or USB format for transmission from the CAN controller to the host machine.

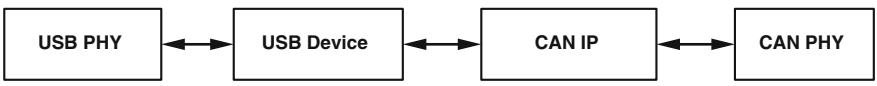


Fig. 2 Block diagram of USB to CAN converter in a CAN node

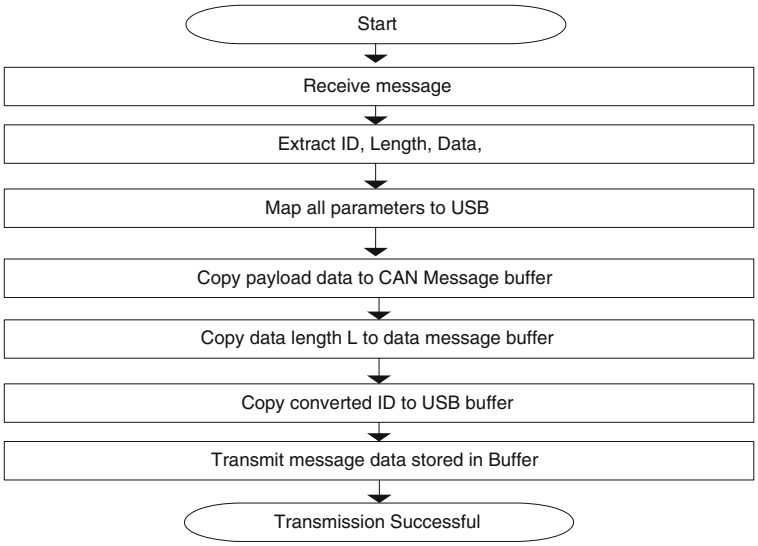


Fig. 3 Flow chart of CAN to USB protocol converter process

7 CAN to Serial/Parallel/USB Convertor

Figure 4 shows the block diagram of CAN to serial/parallel/USB convertor. Each node consists of CAN IP interfaced to CAN transceiver (PHY) device on one side and host controller interface on the other side. It supports serial, parallel port interface or USB interface.

CAN to serial/parallel/USB convertor was prototyped on a customized board that has serial, Parallel and USB interfaces.

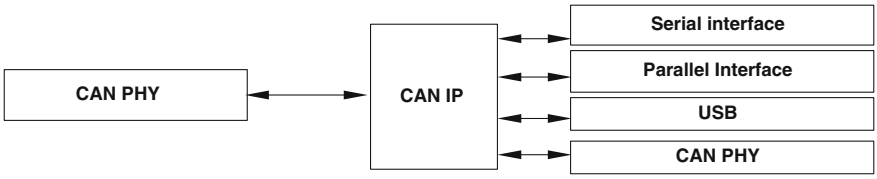


Fig. 4 Block diagram of CAN to serial/parallel/USB convertor

8 Data Transmission and Reception

8.1 Serial Interface

Figure 5 shows the serial interface to CAN controller. It is used to provide the interface between CAN controller and USB or UART.

The data transfer is through the serial interface wrapper. Wrapper interfaces RS232 and USB link data to CAN controller's Rx and Tx lines. Tx_d and Rx_d are connected to CAN bus through CAN PHY.

8.2 Parallel Interface

Figure 6 shows the parallel interface between CAN and CAN Transceiver chip. Parallel interface provides handshake signals between CAN controller and parallel bus structure.

9 CAN IP Design and Verification

CAN IP consist of transmitter and receiver interfaced to RAM (8×256). Processor interface block is interfaced to RAM and CAN transceiver. Verification is carried out using reusable verification components. The module-level components have been reused in top-level verification environment.

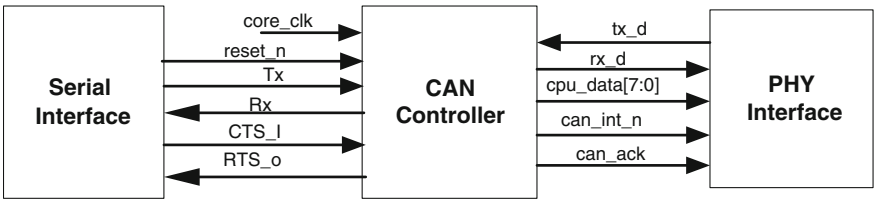


Fig. 5 Serial CAN interface

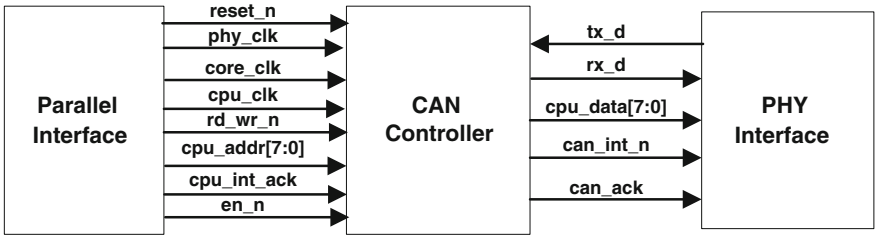


Fig. 6 CAN parallel interface

9.1 Transmitter Block

Transmitter block (TXB) interacts with RAM, physical layer and receiver block. TXB is responsible for the functions such as construction of frame by adding SOF, RTR, reserved bits, CRC, acknowledgement bit and end of frame bits. Initiation of the transmission process after recognizing bus idles serialization of frames, insertion of stuff bits, CRC sequence generation, switching to receive mode when the arbitration is lost, error detection, construction, transmission of overload frame and error frame.

9.2 Receiver Block

Receiver block (RXB) is responsible for functions such as reception of serial bit stream from the physical layer, bit synchronization, de-serialization of the frame structure, deletion of stuff bits, error detection, acceptance filtering, transmission of acknowledgement and recognition of overload frames.

9.3 Verification Environment

CAN verification environment consists of CAN reference models, host processor model, CAN bus protocol monitor, packet checker, bit injector and test case generators to completely simulate a true CAN bus. Complete environment is automated. Verification is targeted for 100 % code coverage, 90 % functional coverage and 48 h of 'no error' 'non-stop' running of random generator.

CAN IP core application interface is wishbone bus interface. CAN core supports transmission and reception of standard CAN protocol messages, automatic stuffing or de-stuffing, automatic CRC calculation, CRC error checking and automatic retransmission. Test cases consist of test cases for directed, random tests and host interface test cases (Table 1).

Table 1 Target device details

FPGA vendor/FAB	Device/Technology	LE count/Cell count	Frequency (MHz)
Altera	Cyclone	1,901(LE)	75
Xilinx	Virtex-2	1,641(CLB)	81
TSMC	0.13	12,600	525

10 Conclusions

Use of this node with serial ports (RS232 and USB) or parallel port will aid in configuration, faultfinding and isolation of faulty node with the help of a host in an easier and cost-effective method compared to time-consuming method. Complete solution is prototyped on FPGA. CAN IP and wrappers for serial and parallel ports interfaces are in Verilog.

Design verification environment is designed using reusable components for module- and top-level verification. The verification environment demonstrates reusability, use of assertions and randomization of test scenarios and test cases, to meet the functional coverage. FPGA prototyping demonstrates the working of the CAN node with peripheral interfaces on silicon.

References

1. Standard-11898 (1993) Road vehicles-interchange of digital information controller area network (CAN) for high-speed communication. International Standards Organisation (ISO)
2. CAN Specification (2003) CAN specification 2.0. Retrieved from http://www.can-cia.org/can/, 18 Oct 2009
3. Navet N, Song Y, Simonot Lion F, Wilwert C (2005) Trends in automotive communication systems. Proc IEEE 93(6):1204–1223
4. Johansson K, Torngren M, Nielsen L (2005) Handbook of networked and embedded control systems. In: Hristu Varsakelis D, Levine WS (eds). Birkhauser, Boston, MA
5. Harish R, Dr. Swarna Jyothi L (2012) Design and verification of FlexRay to CAN protocol converter node, CUBE 2012, Sept 2012, pp 192–197

Mining Information from Model Build Without Information Disclosure

M. Sumana and K. S. Hareesh

Abstract This paper confers an approach of categorizing multiple datasets that are distributed at various sites. The variation in this method of classification is that all the parties jointly build a decision tree model revealing only sufficient information and hiding superfluous data. We have used secure protocols such as secure sum and secure union using commutative encryption technique while model construction. The process discussed in our paper builds efficient, binary classification trees which can in turn predict a test data given by any of the parties involved in the classification.

Keywords Privacy-preserving data mining · Decision tree classification · Secure protocols

1 Introduction

Privacy-preserving data mining (PPDM) algorithms aim in extraction of relevant knowledge from huge volumes of data, while protecting sensitive information at the same time. The existing PPDM techniques can be classified according data distribution (centralized or distributed), the modification applied to the data (encryption, perturbation, generalization, and so on), the data mining algorithm (classification, association mining, or clustering) which the privacy preservation technique is designed for, the data type (single data items or complex data

M. Sumana (✉)
Information Science and Engineering Department, M S Ramaiah Institute of Technology, Bangalore, India
e-mail: sumana.a.a@gmail.com

K. S. Hareesh
Computer Science and Engineering, Manipal Institute of Technology, Manipal, India
e-mail: Harish_dvg@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_83,

correlations) that needs to be protected from disclosure and the approach adopted for preserving privacy (heuristic- or cryptography-based approaches) as in [1]. Heuristic-based algorithms proposed aim at hiding sensitive raw data by applying perturbation techniques based on probability distributions and is mainly conceived for centralized data sets. Current trends in several heuristic-based approaches for hiding both raw and aggregated are k-anonymization, adding noises, data swapping, generalization, and sampling.

Several data mining processes include privacy preservation in order to avoid disclosure of sensitive information while discovering knowledge. Data mining algorithms use various adaptation techniques to construct models or patterns from private data. It is essential to evaluate the quality of the data resulting from the alteration applied by each algorithm, as well as the performance of the algorithms. A broad set of decisive factors have been identified to assess the existing algorithms and determine which algorithm meets specific requirements.

As mentioned in [2] due to the varying typical features of the privacy-preserving algorithms, there is no one single approach that outperforms the others. An algorithm may perform better than the other with respect to hiding sensitive information, privacy level, and data quality as seen in [1] and [5]. Users are provided with a comprehensive set of privacy-preserving–related metrics which will enable them to select the most appropriate privacy-preserving technique for the data at hand; with respect to some specific parameters, they are interested in optimizing.

The PPDM algorithms are constructed based on some of the following privacy definitions. As seen in [1], first, privacy can be defined as the right of an individual which means that the person involved can indicate what data and how much of data can be revealed. Secord, privacy as a controlled access to information, and lastly, privacy as limited to the person and all features related to the person, which means there is no compromise to individual's privacy.

2 Criteria for Evaluation

(a) *Level of privacy* Here, the level up to which sensitive information that has been hidden can be evaluated. The aspects under which metrics for privacy level needs to be used are data privacy and results privacy. Data privacy is quantified by degree of uncertainty. Higher the degree of uncertainty betters the data privacy. Results privacy should include metrics to indicate the degree of inferring the private/sensitive data from the results obtained after data mining.

(b) *Failure in hiding sensitive information* The *hiding failure* (HF) parameter is estimated by the percentage of sensitive information that is still discovered, after the data have been hidden. Most of the developed privacy-preserving algorithms are designed with the goal of obtaining zero hiding failure. Oliveira and Zaiane define the HF as the percentage of restrictive patterns that are

discovered from the sanitized database. It is measured as follows: HF = #RP(D′) / #RP(D), where #*RP*(*D*) and #*RP*(*D*′) denote the number of restrictive patterns discovered from the original data base *D* and the sanitized database *D*,′ respectively, as has been seen in [2]. Under an ideal situation, the HF should be 0.

(c) *Data Quality* Data metrics should be included to evaluate the state of the individual items contained in the database after the enforcement of a privacy-preserving technique such as perturbation or anonymizing data in a database. The quality of the data mining results evaluates the alteration in the information that is extracted from the perturbed or anonymized data on the basis of intended data use. Quality of the result extracted includes various parameters such as accuracy, completeness, and consistency.

(d) *Complexity* This evaluation criterion comprises metrics to measure the *efficiency* and *scalability* of a PPDM algorithm. Efficiency measures the space and time requirements of the algorithm to indicate its performance. Evaluation of the time requirements is the average number of operations performed in order to reduce the frequency of appearance of sensitive information. The communication cost incurred during the exchange of information among the collaborating sites should also be evaluated. Scalability describes the efficiency trends of the algorithm when size of the data sets enhances.

(e) *Resistance* The level of sanitization for different data mining techniques differs. When a sanitization algorithm is developed on a dataset for a particular data mining technique the endurance of this algorithm so that no other data mining technique can disclose, the sensitive information has to be measured. Such a parameter is called as endurance traversal as in [7].

3 Basic Algorithm

3.1 Decision Tree Algorithm

1. Check for terminating conditions
   (a) If all attributes other than the class label attribute have been split.
   (b) If all tuples belong to the same class.
2. If terminating conditions not satisfied, For each attribute *a*
   (a) Find the Gini index of each attribute with its splitting criteria on *a*
3. Let *asplit* be the attribute with the lowest Gini value
4. Create a binary decision *node* that splits on *asplit* splitting criteria.
5. Recurse on the instances/tuples obtained by splitting on *asplit* and add those nodes as children of *node.*

This approach of decision tree construction builds smaller trees faster that are memory efficient.

4 Commonly Used Cryptographical Approaches and Their Realizations

The commonly used cryptographic private approaches used in PPDM are secure multiparty protocols, mentioned in [3, 5]. Most of these protocols are used based on the data mining task to be performed on the data sets. These protocols are either homomorphic or commutative in nature.

Let $Epk(.)$ denotes the encryption function with public key pk and $Dpr(.)$ denotes the decryption function with private key pr. A secure public key cryptosystem is called homomorphic if it satisfies the following requirements:

1. Given the encryption of $m1$ and $m2$, $Epk(m1)$ and $Epk(m2)$, there exists an efficient algorithm to compute the public key encryption of $m1 + m2$, denoted $Epk(m1 + m2) := Epk(m1) + h\ Epk(m2)$.
2. Given a constant k and the encryption of $m1$, $Epk(m1)$, there exists an efficient algorithm to compute the public key encryption of $k \cdot m1$, denoted $Epk(k \cdot m1) := k \times h\ Epk(m1)$.

An encryption algorithm is commutative if the order of encryption does not matter. Thus, for any two encryption keys $E1$ and $E2$, and any message m, $E1\ (E2\ (m)) = E2\ (E1\ (m))$.

4.1 Secure Sum

Secure sum securely calculates the sum of values from individual sites [6]. Assume that each site i has some value vi and all sites want to securely compute $v = \sum_{i=1^m} v_i$, where v is known to be in the range $[0 \ldots n]$. Homomorphic encryption could be used to calculate secure sum.

4.2 Secure Dot Product

Securely computing the dot product of two vectors [4] is another important subprotocol required in many privacy-preserving data mining tasks. Alice has a n-dimensional vector $X = (x1, \ldots, xn)$ while Bob has a n-dimensional vector $Y = (y1, \ldots, yn)$. At the end of the protocol, Alice should get $ra = X \cdot Y + rb$ where rb is a random number chosen from uniform distribution that is known only

to Bob, and $X \cdot Y = \sum_{i=1}^{n} x_i \cdot y_i$. The key idea behind the protocol is to use a homomorphic encryption system. However, this protocol works well only for 2-party situation.

4.3 Secure Union and Secure Intersection

Secure union [6, 8] and secure intersection compute the secure union and intersection of vectors. These protocols can be used for multiple parties. They can be either homomorphic or commutative in nature.

5 Our Implementations

The following protocols have been implemented based on their role in the below-mentioned PPDM algorithm. The complexities of each one of them have also been discussed.

5.1 Secure Sum Computation

One of the significant conditions in privacy-preserving decision tree classification is to compute class to which the classified instances belong to, when all the attributes in distributed datasets have been split. This program was realized in Java. This protocol is initiated by Party 1 who obtains counts (indicated as count[i]) of the number of its instances belonging to each of the class i.

It then does the following steps:

1. Party 1 uses a generate random function to obtain a random number (*rand*). It then adds the count[i] it has to this random number and sends it to the neighboring party.
2. For $i = 2, \ldots, k$

Each party i adds its number of instances [i.e., its count[i] to the corresponding number obtained from its party $(i - 1)$ and forwards it to its party $(i + 1)$].

3. Party 1 receives the added result from party k (last party). Then, largest count[i] obtained in Party 1 is the majority class.

Based on the above for the secure sum protocol with k parties, the computation and communication costs are both $O(k \log(|T|))$. $|T|$ indicates the maximum size of the messages passed between the parties and the outputs of the parties.

Secure sum is also used in our approach to obtain the best attribute for split and condition for split. Secure sum maintains aggregate of the individual counts and presents only the aggregate counts. Though this approach guarantees confidentiality of individual count, it still exposes the accurate sum to the miner, which compromises the individual record privacy. In our research, we would further improve secure sum approach.

5.2 Secure Union Protocol

Here, we discuss the working of the secure union protocol used to compute the union of the data available on multiple sites. In our approach, this method is used to check the terminating condition whether all the instances belong to the same class. Pohlig Hellman protocol has been used to perform commutative encryption and to generate the encryption keys required for secure union. The working of the secure union protocol for the following approach is as follows.

1. Party 1 encrypts its class values for the instances and sends this set to its neighbor.
2. Each of the other parties in turn encrypts its class values for the instances it has and the encrypted class values it has received from the other parties and passes it to their neighbors.
3. In this manner, each of the class values should be encrypted n times is n is the number of parties involved in model building.

Note: Union of all the values placed in the set. This is done by sorting all the elements and eliminates duplication.

4. Finally, if the set has only 1 value in the union set, then Party 1 concludes that all instances after split belong to the same class. Else it continues splitting and model building.

To use Pohlig Hellman for commutative approach, the modulus value n is common to all the sites. All sites based on this value compute the encryption key.

Security of the data is maintained by using this approach. Data are also hidden while mining. However, the complexity depends on the type of commutative encryption, and the amount of data to be encrypted. Communication cost depends on the number of items encrypted and to be forwarded to the other sites and the key size used for encrypting data.

Set of parties wishes to jointly compute some conclusions based on their inputs. But these parties do not want to disclose their entire information while computing. Our paper discusses an approach of classifying data available at multiple sites. Significant requirements to be followed for this method are privacy and correctness. The privacy requirement states that not anything should be learned beyond what is extremely necessary. The correctness requirement states that each party

should be able to interpret the output generated and use them to derive appropriate conclusions. Hence, all parties collaborate to induce a decision tree model from the data available with them. This decision tree can be further used by the sites to indicate a new decision.

5.3 Decision Tree Construction on Horizontal Partitioned Datasets

Concepts Involved

Assumptions: Data are preprocessed
Require: R, The set of attributes.
Require: C, The class attribute.
Require: S, the horizontally distributed data set which means all the parties know the number of attributes. All parties hold sufficient information about all the attributes and include the class label attribute.

Party 1

1. if all attributes in R have been split, then
   (a) Active secure sum protocol send the count of the number of each of the class values to its nearest neighbor in encrypted (same encryption key) format.
   (b) Based on the final encrypted result obtained from the last site (class value that has the highest result), return a leaf with class value c_i.
2. else if // check whether all the tuples have the same class value
   (a) Activate secure union protocol using commutative encryption
   That is, Party 1 encrypts its values and the encrypted values that it receives from last and sends it to the neighboring parties. After all the data are encrypted the number of times as the number of parties, it computes union.(e.g., $E1(E2(C_1)) = E2(E1(C_1))$ where C_1 is present in both sites 1 and 2 and $E1$ and $E2$ are the encryption methods in Party 1 and Party 2 then Union($E1(E2(C_1))$, $E2(E1(C_1))$) is $E1(E2(C_1))$)
   Note: will not encrypt a value that is 0.
   (b) If only single value in the set, then return a leaf with class value c_i.
3. else
   (a) // to determine the best value for split.
   Compute Gini index and inform other parties also to compute the Gini index of all the attributes not yet split. Again activate the secure sum protocol to find the attribute and the splitting criteria having the least Gini index. Indicate this information to all the parties.

(b) Perform a binary split on S, $S(a_1)$ and $S(a_2)$ such that a_1 and a_2 are the 2 possible splitting condition values of A.
(c) Repeat the operation with splits until condition 1 and 2 are satisfied.

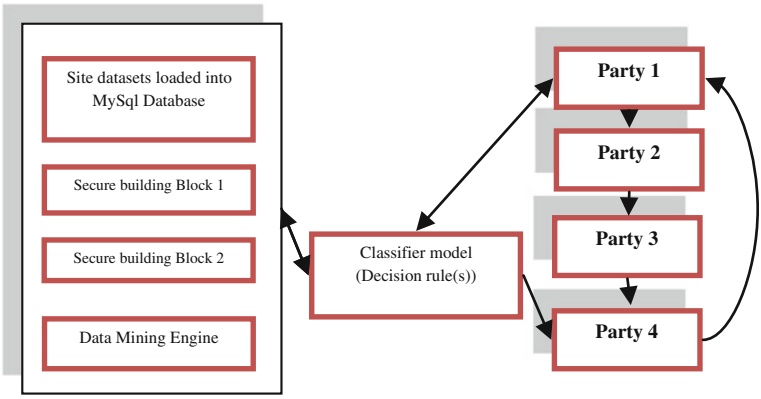
4. end if

Remaining Parties

1. If a value is obtained from the neighboring site and secure sum activated then perform addition of the value and pass to the next neighbor.
2. If secure union or secure intersection included then compute encryption on the values present in the dataset + values obtained from the other sites.
3. Compute Gini index when indicated by Party 1.
4. Use the final decision tree to predict the given data set.

The entire process of decision tree computation was done in the following manner

1. Initially, the number of parties who wanted to perform the joint computation where indicated. It is assumed that the data set used was preprocessed.
2. PPDM was performed on these multiple datasets using secure computations to build a decision tree.
3. The decision tree was interpreted by the parties in the form of decision rules.
4. The parties finally extract the knowledge from the output they receive.



6 Conclusions

This paper converses a framework for evaluating PPDM algorithms. Such framework allows one to assess the different features of a privacy-preserving algorithm according to a variety of evaluation criteria. Parameters like level of privacy, data quality, and HF have been defined and the evaluations of such

parameters over a set of privacy-preserving algorithms have been presented. Communication and security analysis of the algorithms such as secure union, sum, and dot product have been performed. Further, these above-mentioned approaches have been implemented keeping in mind the usage of these methods in privacy preserving. Our implementations work well for semi-honest models.

In our future work, we would like to evaluate the algorithms for malicious models. Security and quality issues also need to be considered while using an efficient encryption method such as Pohlig Hellman. Further, we would build privacy-preserving techniques using the secure multiparty protocols realized.

References

1. Bertino E, Lin D, Jiang W (2008) A survey of quantification of privacy preserving data mining algorithms. Scientific Commons
2. Aggarwal CC, Yu Philip S (2007) Privacy preserving data mining: models and algorithms. Springer, Berlin
3. Lindell Y, Pinkas B (2008) Secure multiparty computation for privacy-preserving data mining. Israel Science Foundation (grant number 860/06)
4. Hussein M, El-Sisi A, Ismail N (2008) Fast cryptographic privacy preserving association rules mining on distributed homogenous data base. Springer, Berlin, pp 607–616
5. Clifton C, Kantarcioglu M, Vaidya J (2004) Tools for privacy preserving distributed data mining. SIGKDD Explor 4(2)
6. Shrikant Vaidya J (2004) Privacy preserving data mining over vertically partitioned data. A thesis submitted to Purdue University for the Degree of Doctor of Philosophy, Aug 2004
7. Kenthapadi K (2006) Models and algorithms for data privacy. A dissertation submitted to the Department of Computer Science for the Degree of Doctor of Philosophy, Sept 2006
8. Du W, Zhan Z (2002) Building decision tree classifier on private data. IEEE international conference on data mining workshop on privacy, security, and data mining, vol 14, pp 1–8, Maebashi City, Japan, 9 Dec 2002. Australian Computer Society

Quantitative Analysis of the Openness of the Major Temporal Arcade in Retinal Fundus Images of Retinopathy of Prematurity

Faraz Oloumi, Rangaraj M. Rangayyan and Anna L. Ells

Abstract Monitoring the openness of the major temporal arcade (MTA) and how it changes in the presence of various forms of retinopathy, such as retinopathy of prematurity (ROP), could assist in timely detection and effective treatment. We present image processing techniques, including Gabor filters and a form of the generalized Hough transform, for the detection and modeling of the MTA via a graphical user interface (GUI). An arcade angle was also obtained via the GUI. An area under the receiver operating characteristic curve of $A_z = 0.75$ was obtained using the results of single- and dual-parabolic modeling in the discrimination of Stage 0 ROP from Stage 3 ROP; $A_z = 0.71$ was obtained in screening for ROP (Stage 0 versus Stages 1, 2, and 3). The arcade angle provided similar results in terms of A_z values.

Keywords Blood vessels · Fundus images · Gabor filters · Hough transform · Major temporal arcade · Parabolic modeling · Retina · Retinopathy of prematurity · Temporal arcade · Arcade angle

F. Oloumi · R. M. Rangayyan (✉)
Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada
e-mail: ranga@ucalgary.ca

A. L. Ells
Division of Ophthalmology, Department of Surgery, Alberta Children's Hospital, Calgary, AB T3B 6A8, Canada

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_84,

1 Introduction

1.1 Retinopathy of Prematurity

Various forms of retinopathy, such as proliferative diabetic retinopathy and retinopathy of prematurity (ROP), can alter the structure of the retinal vasculature in different ways [1–3]. Premature infants with a birth weight of less than 1,500 g or a gestational age of less than 28–31 weeks are required to be screened for possible development of ROP [4–9]. The staging of ROP is based on the abnormal vascular response at the junction of the vascularized and the avascular parts of the retina and consists of five stages [3]. Other posterior signs that are indicative of the presence of ROP are the straightening of the major temporal arcade (MTA), as well as a decrease in the angle of insertion of the MTA [3, 10–12]. Figure 1 shows retinal fundus images of premature infants with and without ROP. The cases in which no ROP is present (Fig. 1a) are also referred to as Stage 0 ROP.

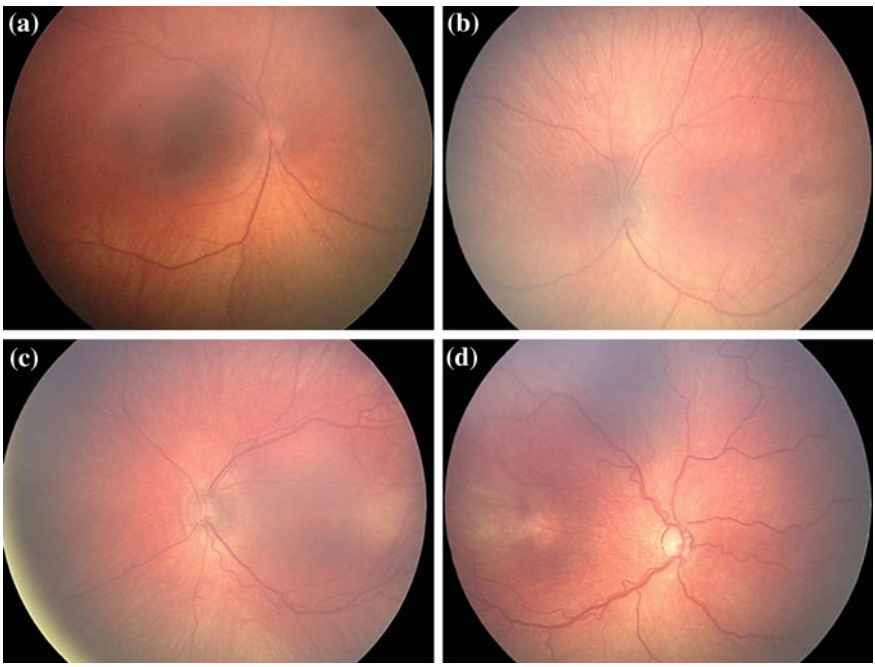


Fig. 1 Retinal fundus images of preterm infants showing various stages of ROP. **a** A normal case showing no signs of ROP; normal cases are also referred to as Stage 0 ROP. **b** An example of a case with Stage 1 ROP. Cases of **c** Stage 2 and **d** Stage 3 ROP; both cases show signs of plus disease, that is, increase in arteriole tortuosity and venular dilation

1.2 Changes in the Vessel Width and Tortuosity in the Presence of Plus Disease

In addition to the changes that describe the five stages of ROP, the presence of plus disease can also be indicative of the severity of active ROP. Plus disease is diagnosed by the presence of a certain amount of increase in venular dilation and arteriole tortuosity in at least two quadrants of the eye [3]. The presence of sufficient dilation and tortuosity of the posterior vessels for the diagnosis of plus disease is determined by visual qualitative comparison to a gold standard retinal fundus photograph [3, 13, 14]. The presence of venular dilation and arteriole tortuosity that is not sufficient for the diagnosis of plus disease is indicative of preplus disease [3]. A severe form of ROP, called aggressive posterior ROP, also shows increase in the dilation and tortuosity of the blood vessels in all four quadrants at early stages of its development [3]. Figures 1c and d show cases of Stages 2 and 3 ROP, respectively, that show arteriole tortuosity and venular dilation that are sufficient for the diagnosis of plus disease.

A direct relationship appears to exist between increasing venule thickness and increasing severity of ROP; however, this has not been shown to have statistical significance across all studies [15]. Furthermore, the detection of small changes in venule thickness requires high-resolution imaging as these changes are at, or below, the spatial resolution of typical retinal imaging systems [14]. Arteriole tortuosity has shown higher correlation with the presence of plus disease, but this factor has also not been consistent across all trials [16]. Moreover, the detection of arterioles presents an image processing challenge; they have a lower contrast as compared to the venules, and their thickness can be below the resolution limit of even high-resolution images of premature infants [15]. These factors and limitations indicate a great need for quantification of other diagnostic manifestations of ROP, such as the angle of insertion and the straightening of the MTA.

1.3 Changes in the Architecture of the MTA Due to ROP

Changes in the overall architecture of the MTA in terms of a decrease in its angle of insertion as well as the straightening of the superior and the inferior temporal arcades (STA and ITA) are used as indicators of the compromised structural integrity of the macular region [11, 13, 17, 18]. A change in the architecture of the MTA can be seen as a sequela of ROP [3, 10].

1.4 Measurement of Vessel Thickness and Tortuosity

Using a binary vessel map, Heneghan et al. [16] computed the vessel width by extending a line segment, in different directions, from a pixel that belongs to the vessel on both sides in opposite directions until nonvessel pixels were encountered. The width of the vessel was taken as the smallest distance over all angles. Heneghan et al. also obtained the tortuosity of the vessels by first dividing a vessel into smaller linear segments and then dividing the total derived length of a vessel by the length of its chord.

Wilson et al. [14] measured the vessel width using two different methods. As one measure, the standard deviation of the Gaussian profile that was used to detect the blood vessels was used. As a second measure, the correlated measure of total isotropic contrast at the vessel centerline along the entire length of the vessel was used. Wilson et al. [14] also measured the tortuosity of blood vessel by dividing a vessel segment into smaller segments using the bisector of each segment and the length of their chords until a minimum chord length was reached. The tortuosity of the vessel was computed by considering the change in the chord length after division of each segment into two parts.

Fiorin and Ruggeri [19] proposed semiautomated methods for measurement of vessel thickness and tortuosity in narrow-field retinal fundus images of preterm infants using a Web-based graphical user interface (GUI). The centerline of the vessel to be analyzed was selected manually, and the edges of the vessel were extracted using a Canny filter. The vessel width and tortuosity were then computed using the centerline and by associating pairs of points on opposite edges so that a line through the two points would be perpendicular to the vessel centerline.

By finding the branching and the endpoints of a vessel skeleton map, Gelman et al. [20] divided a vessel into smaller segments. The tortuosity was measured as the sum of the length of all such segments divided by the length of the chord of the entire vessel. The width of the vessel was measured as the total area of the vessel divided by its length.

Poletti et al. [21] used various forms of vessel-level tortuosity such as angle-based tortuosity, caliber-weighted angle-based tortuosity, and twisted-based tortuosity to derive eight different measures of tortuosity. A linear weighted combination of the eight different measures of vessel-level tortuosity was then obtained to represent an image-level tortuosity measure for each image.

1.5 Detection of the MTA for Localization of Other Anatomical Features

By using the supremum of opening operator on an enhanced grayscale image of blood vessels, Welfer et al. [22] extracted the STA and the ITA to locate the optic nerve head (ONH). A set of 24 linear structuring elements of length 80 pixels was

used to extract the MTA. The resulting image was binarized, skeletonized, and pruned to obtain a binary image that represented the STA and the ITA.

Ying and Liu [23] obtained a vascular topology map using an energy function. A circle fitting method was applied to the extracted pixels to model the MTA as a circle, which was then used to localize the macula.

Using steerable filters and color thresholding, Kochner et al. [24] extracted edge points on the main blood vessels. An ellipse was then fitted to the extracted points using the generalized Hough transform (GHT). The end of the long axis of the ellipse was taken as an estimate of the location of the ONH.

With prior knowledge of the location of the ONH, Tobin et al. [25] estimated the retinal raphe angle by using the least squares method and a parabolic model applied to the statistical distribution of points on the vascular skeleton. Even though they estimated the parameter a, of the parabolic model, it was only used to draw a parabola on the image.

Fleming et al. [26] proposed a method to extract the MTA by vessel enhancement and semielliptical curve fitting using the GHT.

Li and Chutatape [27] used an active shape model along with a point distribution model to extract the main course of the retinal vessels. A combination of the Hough transform [28] and linear least squares fitting was used to estimate a parabolic model.

By assuming that the main vessels originate from the ONH and extend in paths that can be geometrically modeled as parabolas, Foracchia et al. [29] defined a directional model for the major blood vessels. With the model and data indicating the center points, direction, and caliber of the vessels, by using a residual sum of squares method, the parameters of the model were identified.

The MTA has not been previously modeled for parameterization of its openness. As shown in our recent studies [2, 30], the parabolic shape of the MTA, or the dual-parabolic nature of the STA and ITA, allows for effective single- and dual-parabolic modeling using a form of the GHT. In such models, changes in the architecture of the MTA, STA, and ITA are expected to be reflected as changes in the openness parameters of the parabolic models. The present work is concerned with quantitative analysis of the openness of the MTA in retinal images of cases with and without ROP using single- and dual-parabolic models as well as a form of the arcade angle measure, obtained using a GUI [2, 31].

2 Materials and Methods

2.1 Database of Fundus Images of the Retina

The proposed methods were tested with images from the Telemedicine for ROP In Calgary (TROPIC) database [32] (see Fig. 1). The images of the TROPIC database are wide-field (130°) fundus images of the retina, captured using the RetCam 130

camera, and have a size of 640 × 480 pixels. The spatial resolution of the images is estimated to be 30 μm/pixel [33]. The TROPIC database does not include any cases of Stages 4 and 5 ROP. Eighteen images for each of Stages 0, 1, 2, and 3 were selected at random from the database. The selected images are from 29 different individuals. At most, two images of the same patient were included for the same stage of ROP: one image from each eye. Images of the same eye from the same patient were included only if the ROP stages were different at the time of imaging. A total of 18 images were also diagnosed with plus disease, which were categorized as Stages 2 and 3 of ROP.

2.2 Procedure for Measurement of the Arcade Angle

The present work employs the method of Wong et al. [12] for the measurement of the arcade angle via a GUI [2, 31]. Using the module for the measurement of the arcade angle prompts the user to mark the center of the ONH, after which a circle with a radius that is specified by the user is automatically drawn on the image [12, 31]. The user is then prompted to mark the point of intersection of the circle with the inferior venule; the same is repeated for the superior venule. The arcade angle is measured as the angle between the three marked points with the center of the ONH being the vertex of the angle (see Oloumi et al. [2, 31] for details of the procedure for the measurement of arcade angles). In the present work, circles of radii $r = 120$ and 60 pixels were used to measure the arcade angle.

2.3 Detection and Modeling of the MTA

Figure 2 provides a flowchart representation of the image processing algorithms developed for the detection and modeling of the MTA by single- and dual-parabolic modeling. In the present work, Gabor filters are used for the detection of the MTA (see Rangayyan et al. [34] for details of the Gabor filters and the detection of blood vessels). A type of the GHT for the detection of parabolic forms is used with the results of Gabor filtering for single- and dual-parabolic modeling of the MTA, STA, and ITA (see Oloumi et al. [30] for details of the modeling procedures). The proposed methods are implemented via a GUI, which facilitates user input for selection of the parameters of the Gabor filters and the binarization steps used for detection and modeling of the MTA, as well as measurement of the arcade angle (see Oloumi et al. [2, 31] for details of the GUI).

The green component (G-component) from the original color image is used to perform the proposed procedures. Thresholding the G-component at 0.075 of its maximum value provides a mask image indicating the effective area of the image. The G-component is extended beyond the effective area to avoid the detection of its edges by the Gabor filters. In the present work, 30 Gabor filters over the range

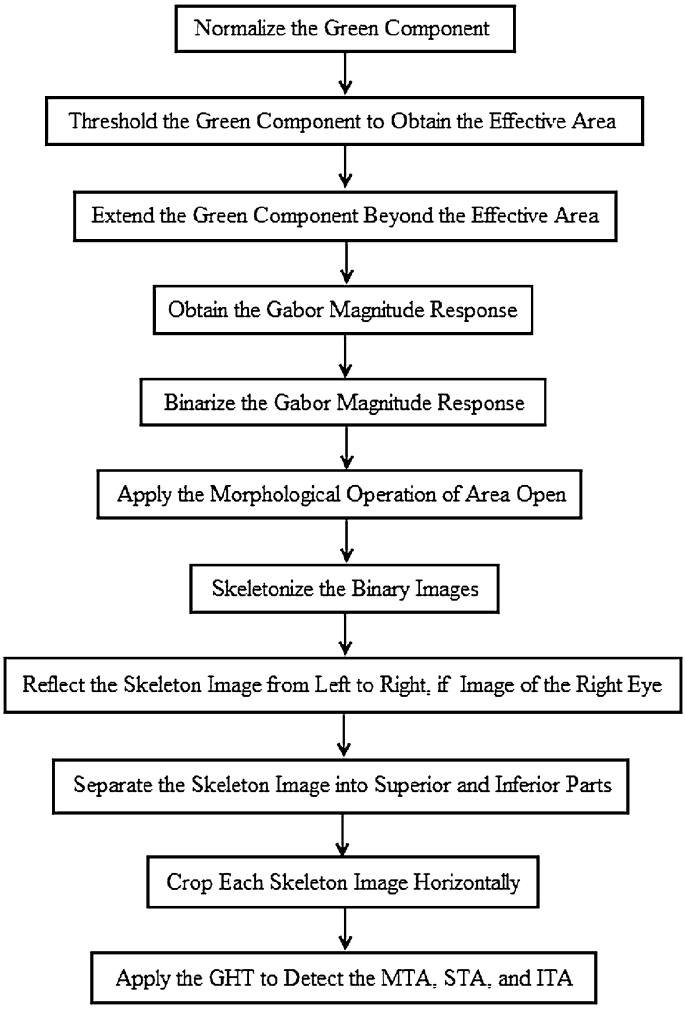


Fig. 2 Flowchart representation of the image processing algorithms used for the detection and modeling of the MTA

$[-90°, 90°]$ are used to detect the MTA. The set of values used for the thickness parameter of the Gabor filters is $\tau = \{10, 12, 14\}$ pixels. The value for the elongation parameter of the Gabor filters is fixed at $l = 2$; see Rangayyan et al. [34] for details.

The Gabor magnitude response image can be thresholded using a sliding threshold via the GUI. The user can specify the maximum number of connected pixels to be removed to eliminate small vessel segments that may remain after the thresholding step by using the morphological operation of area open [35]. For each

image, a suitable threshold was selected via the GUI to obtain a binary image containing only the MTA.

The modeling procedure requires a binary image to perform the GHT procedure. The user is required to indicate whether the current image is an image of the left eye or the right eye. The user is then prompted to mark the approximate location of the ONH in a separate window. The range of the openness parameter, a, used in the parabolic modeling procedure, was set to $10 \leq a \leq 400$ for the selected images of the TROPIC database. Given an average width of the ONH (ONHW) of about 1.05 mm in preterm infants [33], and the spatial resolution of the RetCam images, the vessel skeleton map used to derive the GHT is horizontally restricted from 0.25 × ONHW nasal to 2 × ONHW temporal to the ONH center (size of each Hough space slice = 80 × 480 pixels).

The p-values, indicating the statistical significance of the differences between values of the parameter a, as well as the arcade angle for the Stage 0 cases as compared to the values of the same for the cases of Stages 1, 2, and 3, were obtained via the t-test. To assess the diagnostic performance of the parameters derived, receiver operating characteristic (ROC) [36] analysis was performed using ROCKIT [37].

3 Results

Figure 3 shows the results of single- and dual-parabolic modeling as well as the measurement of the arcade angle using circles of radii $r = 60$ and 120 pixels for an image from the TROPIC database. It can be observed that both the single- and the dual-parabolic models are providing close fits to the arcade, but the dual-parabolic model provides a closer fit to the entire length of the arcade up to the macular region. The arcade angle using the circle of radius $r = 120$ pixels is providing a measure closer to the macular region, whereas the arcade angle obtained using the circle of radius $r = 60$ pixels is providing a measure closer to the posterior pole.

Table 1 shows the values of the mean ± standard deviation (STD) obtained using the parameters of the single- and dual-parabolic models, as well as the arcade angles using a radius of $r = 120$ pixels for the cases of Stages 0, 1, 2, and 3 from the TROPIC database. It can be observed that the values for all of the parameters decrease as the severity or stage of ROP increases. There is a large STD in the dual-parabolic model parameters for Stages 0 and 1 because of a few outliers, where the two arcades (STA and ITA) are asymmetric and one yields a high openness parameter and the other a much lower one.

Table 2 shows the values of the area under the ROC curve (A_z) and their respective standard errors (SE) for the discrimination of Stage 0 ROP from Stages 1, 2, 3, and some of their combinations, using the openness parameters obtained from the single-and dual-parabolic models as well as the arcade angles obtained using a circle of radius $r = 120$ pixels. It can be observed that the best classification results are obtained for the discrimination of Stage 0 from Stage 3 and

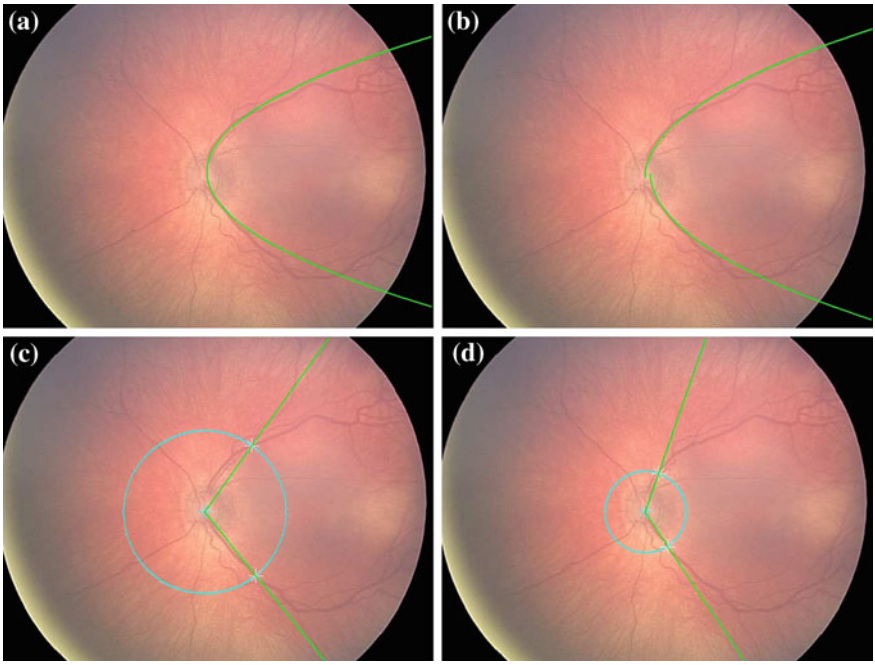


Fig. 3 Image 2304 of the TROPIC database. **a** Single-parabolic model with $a_{\mathrm{MTA}} = 30$. The model is missing parts of the ITA toward the periphery. **b** Dual-parabolic model with $a_{\mathrm{STA}} = 36$ and $a_{\mathrm{ITA}} = 36$; both models are providing close fits. Arcade angle measures using circles of radii **c** $r = 120$ pixels with $\theta = 106.2°$ and **d** $r = 60$ pixels with $\theta = 127.0°$. The arcade angle measure obtained using the circle of radius $r = 60$ pixels provides a measure closer to the posterior pole as compared to the arcade angle obtained using the circle of radius $r = 120$ pixels

Table 1 Values of the mean ±STD for the |a| parameters of the single- and dual-parabolic models as well as the arcade angle (in degrees) obtained using a radius of $r = 120$ pixels for 18 images for each Stage 0, 1, 2, and 3

Parameter	Stage 0	Stage 1	Stage 2	Stage 3
	Mean ± STD	Mean ± STD	Mean ± STD	Mean ± STD
Arcade angle	119.98 ± 15.79	116.92 ± 18.35	107.58 ± 12.18	106.20 ± 15.63
$\lvert a_{\mathrm{MTA}} \rvert$	54.28 ± 33.53	39.06 ± 14.75	35.72 ± 20.27	36.00 ± 39.74
$\lvert a_{\mathrm{STA}} \rvert$	56.11 ± 47.11	56.00 ± 81.62	37.17 ± 17.70	34.22 ± 24.34
$\lvert a_{\mathrm{ITA}} \rvert$	69.83 ± 67.32	70.89 ± 83.13	40.50 ± 21.91	44.83 ± 59.01

combination of Stages 2 and 3. The discrimination of cases of Stage 0 ROP from Stage 1 yielded low A_z values, as expected.

Table 3 shows the p-values related to the statistical significance of the differences between the various stages of ROP for the parameters of the single- and dual-parabolic models as well as the arcade angles obtained with a circle of radius $r = 120$ pixels. The differences between the values obtained for the arcade angle

Table 2 Values of the area under the ROC curve (A_z) and their standard error (SE) in the discrimination of Stage 0 ROP (no ROP) against Stages 1, 2, 3, and their combinations using the $|a|$ parameters of the single- and dual-parabolic models as well as the arcade angle (in degrees) obtained using a radius of $r = 120$ pixels

ROP Stage	Arcade angle, A_z (SE)	$\lvert a_{\mathrm{MTA}}\rvert$, A_z (SE)	$\lvert a_{\mathrm{STA}}\rvert$, A_z (SE)	$\lvert a_{\mathrm{ITA}}\rvert$, A_z (SE)
0 versus 1	0.55 (0.095)	0.67 (0.088)	0.57 (0.095)	0.50 (0.095)
0 versus 2	0.73 (0.083)	0.71 (0.085)	0.62 (0.093)	0.67 (0.088)
0 versus 3	0.74 (0.081)	0.75 (0.083)	0.69 (0.087)	0.75 (0.081)
0 versus 2 + 3	0.74 (0.069)	0.73 (0.068)	0.66 (0.079)	0.71 (0.072)
0 versus 1 + 2 + 3	0.64 (0.068)	0.71 (0.067)	0.64 (0.078)	0.65 (0.072)

There are 18 images per stage of ROP. Note that the absolute values of the parameters of the parabolic models are used to perform the calculations

Table 3 The p-values indicating the statistical significance of the differences between various stages of ROP and their combinations for the $|a|$ parameters of the single- and dual-parabolic models as well as the arcade angle (in degrees) obtained using a radius of $r = 120$ pixels

ROP Stage	Arcade angle, p-value	$\lvert a_{\mathrm{MTA}}\rvert$, p-value	$\lvert a_{\mathrm{STA}}\rvert$, p-value	$\lvert a_{\mathrm{ITA}}\rvert$, p-value
0 versus 1	0.5947	0.0869	0.9960	0.9669
0 versus 2	0.0125*	0.0525	0.1195	0.0878
0 versus 3	0.0127*	0.1451	0.0889	0.2443
0 versus 2 + 3	0.0029**	0.0508	0.0314*	0.0801
0 versus 1 + 2 + 3	0.0281*	0.0284*	0.3143	0.2985
1 versus 2	0.0811	0.5764	0.3455	0.1429
1 versus 3	0.0679	0.7616	0.2856	0.2859
2 versus 3	0.7691	0.9791	0.6807	0.7720

There are 18 images per stage of ROP

Note that the absolute values of the parameters of the parabolic models are used to perform the calculations. (*$p < 0.05$ and **$p < 0.01$)

measure for cases with no ROP (Stage 0) as compared to cases of Stages 2 and 3 (Stage 2 + 3) combined show high statistical significance. The differences between the parameter a of the STA ($|a_{\mathrm{STA}}|$) models for the same stages (Stage 0 versus Stage 2 + 3) are statistically significant. The results of arcade angle measurement using a circle of radius $r = 60$ pixels are similar to the measures obtained using a circle of radius $r = 120$ pixels, but slightly worse in terms of both A_z and p-values; hence, they are not included in the present paper.

4 Discussion

This is the first study to quantify the effects of various stages of ROP on the openness of the MTA, STA, and ITA, using semiautomated methods to perform single- and dual-parabolic modeling. Previous studies that have quantified the

openness of the MTA have used manual and semiautomated methods to measure an angle based on three different points [11, 12]. In the present work, the statistical and diagnostic performance (see Tables 1, 2, and 3) of the parameters of the single- and dual-parabolic models is similar to that provided by the arcade angle measures obtained based on the method of Wong et al. [12] via the GUI. All of the studied measures show a similar trend: As the severity of ROP increases, the openness of the MTA decreases. Although the A_z values obtained for the arcade angle measures and the parameters of the parabolic models are not high, the results are encouraging. The Gaussian assumption made in ROCKIT and the *t*-test may be inappropriate for the set of parameters derived in the present work; other classifiers and measures of performance may lead to better results.

Ells and MacKeen [38] illustrated that the changes that occur in the MTA in the presence of ROP are dynamic as they alter the posterior architecture of the MTA. Based on this point, we believe that the arcade angle measures proposed by Wilson et al. [11] and Wong et al. [12] may not accurately reflect such changes that occur over the entire posterior architecture of the MTA, as they only define the openness of the MTA based on three points and the methods are similar to fitting a triangle to the MTA. Furthermore, the value of the arcade angle is sensitive to the exact location of the center of the ONH provided by the user; a difference of 10 pixels in the location of the marked center of the ONH can result in a difference of 10° in the measured arcade angle. The parabolic modeling procedure is dependent only on an approximate location of the ONH instead of the specific location of the center of the ONH.

The size of the radius of the circle used to obtain the arcade angle measures needs to be fixed based on a physiologically meaningful parameter, such as the average ONHW or the distance from the center of the ONH to the fovea. In the present work, this value is set based on the work of Wong et al. [12], in which the value has been specified with no explanation. This value should be set based on how far from the ONH the arcade angle should be measured by considering a physiological limit.

Even though the parameters of the single-parabolic models have provided slightly better results as compared to the parameters of the dual-parabolic models, the results of single-parabolic modeling may not be as accurate as those of dual-parabolic modeling. It was observed that in cases of asymmetric arcades, the single-parabolic modeling procedure captures the more dominant semi-parabola and may not incorporate both the STA and the ITA. The dual-parabolic modeling procedure may be more suitable for parameterization of the openness of the architecture of the MTA in fundus images of the retina.

The large STD in the parameters of the dual-parabolic models, caused by a few outliers, is mainly due to the fact that if the retinal raphe angle (the line going through the center of the ONH and the fovea) is large with respect to the horizontal axis of the image, it could lead to a much larger openness parameter for one of the dual-parabolic models than the other. The retinal raphe angle may be corrected either manually or by using methods to detect the fovea [30]. However, as shown by Chiang et al. [39], retinal fundus images of preterm infants typically lack a clear

depiction of the fovea. The two parameters of the dual-parabolic models may be combined using a neural network or a nonlinear classifier to incorporate the information from both the STA and the ITA models into the classification methods.

The approximate shape of the overall architecture of the MTA may appear to be parabolic or semielliptical; however, upon close inspection, it becomes clear that first, the STA and the ITA are asymmetric, and second, more accurate modeling of each arcade may be possible by applying higher-order models. A high-order curve fitting method may provide more accurate results in terms of modeling and parameterization of the MTA.

Quantification and combination of measures of tortuosity and thickness of the blood vessels with the quantification of the openness of the MTA may yield better results in terms of diagnosis of ROP as well as discrimination between its various stages; this work is currently in progress.

Acknowledgments This work was supported by the Natural Sciences and Engineering Research council of Canada. We thank April Ingram for help with the TROPIC images.

References

1. Danis RP, Davis MD (2008) Proliferative diabetic retinopathy. In: Duh EJ (ed) Diabetic retinopathy, Contemporary diabetes. Humana Press, Totowa, pp 29–65
2. Oloumi F, Rangayyan RM, Ells AL (2012) Computer-aided diagnosis of proliferative diabetic retinopathy. In: Engineering in Medicine and Biology Society (EMBS), 34th annual international conference of the IEEE, pp 1438–1441, San Diego
3. International committee for the classification of retinopathy of prematurity (2005) The international classification of retinopathy of prematurity revisited. Archives of Ophthalmology, pp 123:991–999
4. American Academy of Pediatrics, American Association for Pediatric Ophthalmology and Strabismus, American Academy of Ophthalmology (2001) Screening examination of premature infants for retinopathy of prematurity. Pediatrics, pp 108:809–811
5. Association of Perinatal Medicine (1996) Retinopathy of prematurity: guidelines for screening and treatment. Early Human Dev, pp 46(3):239–258
6. KIDROP: Karnataka Internet assisted Diagnosis of Retinopathy of Prematurity. http:/www.narayananethralaya.org/kidrop.html
7. Vinekar A (2011) IT-enabled innovation to prevent infant blindness in rural India: The KIDROP experience. J Indian Bus Res 3(2):98–102
8. Murakami Y, Silva RA, Jain A, Lad EM, Gandhi J, Moshfegh DM (2010) Stanford University Network for Diagnosis of Retinopathy Of Prematurity (SUNDROP): 24 month experience with telemedicine screening. Acta Ophthalmol 88(3):317–322
9. Stanford University Network for Diagnosis of Retinopathy Of Prematurity (SUNDROP): four-years with telemedicine screening. http:/med.stanford.edu/profiles/frdAction-Servlet.stanford.edu/profiles/frdAction-Servlet?choiceId=showPublication&pubid=15688823&fid=17803&
10. Cryotherapy for Retinopathy of Prematurity Cooperative Group (2001) Multicenter trial of cryotherapy for retinopathy of prematurity: ophthalmological outcomes at 10 years. Archives of Ophthalmology, 119:1110–1118
11. Wilson C, Theodorou M, Cocker KD, Fielder AR (2006) The temporal retinal vessel angle and infants born preterm. Br J Ophthalmol 90:702–704

12. Wong K, Ng J, Ells AL, Fielder AR, Wilson CM (2011) The temporal and nasal retinal arteriolar and venular angles in preterm infants. Br J Ophthalmol 95(12):1723–1727
13. Watzke RC, Robertson JE, Palmer EA, Wallace PR, Evans MS, Delaney Soldevilla JE (1990) Photographic grading in the retinopathy of prematurity cryotherapy trial. Archives of Ophthalmology, 108(7):950–955
14. Wilson CM, Cocker KD, Moseley MJ, Paterson C, Clay ST, Schulenburg WE, Mills MD, Ells AL, Parker KH, Quinn GE, Fielder AR, Ng J (2008) Computerized analysis of retinal vessel width and tortuosity in premature infants. Invest Ophthalmol Vis Sci 49(1):3577–3585
15. Swanson C, Cocker KD, Parker KH, Moseley MJ, Fielder AR (2003) Semiautomated computer analysis of vessel growth in preterm infants without and with ROP. Br J Ophthalmol 87(12):1474–1477
16. Heneghan C, Flynn J, O'Keefe M, Cahill M (2002) Characterization of changes in blood vessels width and tortuosity in retinopathy of prematurity using image analysis. Med Image Anal 6(1):407–429
17. Jelinek HF, Cree MJ (2010) Introduction. In: Jelinek HF, Cree MJ (eds) Automated image detection of retinal pathology. CRC Press, Boca Raton, pp 1–26
18. Fledelius HC, Goldschmidt E (2010) Optic disc appearance and retinal temporal vessel arcade geometry in high myopia, as based on follow-up data over 38 years. Acta Ophthalmol 88(5):514–520
19. Fiorin D, Ruggeri A (2011) Computerized analysis of narrow-field ROP images for the assessment of vessel caliber and tortuosity. In: Engineering in Medicine and Biology Society, 33rd annual international conference of the IEEE, pp 2622–2625, Boston
20. Gelman R, Martinez-Perez ME, Vanderveen DK, Moskowitz A, Fulton AB (2005) Diagnosis of plus disease in retinopathy of prematurity using retinal image multiscale analysis. Invest Ophthalmol Visual Sci 46(12):4734–4738
21. Poletti E, Grisan E, Ruggeri A (2012) Image-level tortuosity estimation in wide-field retinal images from infants with retinopathy of prematurity. In: Engineering in Medicine and Biology Society, 34th annual international conference of the IEEE, pp 4958–4961, San Diego
22. Welfer D, Scharcanski J, Kitamura CM, Dal Pizzol MM, Ludwig LWB, Marinho DR (2010) Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach. Comput Biol Med 40(2):124–137
23. Ying H, Liu JC (2010) Automated localization of macula-fovea area on retina images using blood vessel network topology. In: Acoustics speech and signal processing, IEEE international conference on, pp 650–653
24. Kochner B, Schuhmann D, Michaelis M, Mann G, Englmeier KH (1998) Course tracking and contour extraction of retinal vessels from color fundus photographs: most efficient use of steerable filters for model based image analysis. In: SPIE medical imaging, 3338:755–761, San Diego
25. Tobin KW, Chaum E, Govindasamy VP, Karnowski TP (2007) Detection of anatomic structures in human retinal imagery. IEEE Trans Med Imaging 26(12):1729–1739
26. Fleming AD, Goatman KA, Philip S, Olson JA, Sharp PF (2007) Automatic detection of retinal anatomy to assist diabetic retinopathy screening. Phys Med Biol 52:331–345
27. Li H, Chutatape O (2004) Automated feature extraction in color retinal images by a model based approach. IEEE Trans Biomed Eng 51(2):246–254
28. Hough PVC (1962) Method and means for recognizing complex patterns. US Patent 3, 069, 654
29. Foracchia M, Grisan E, Ruggeri A (2004) Detection of optic disc in retinal images by means of a geometrical model of vessel structure. IEEE Trans Med Imaging 23(10):1189–1195
30. Oloumi F, Rangayyan RM, Ells AL (2012) Parabolic modeling of the major temporal arcade in retinal fundus images. IEEE Trans Inst Meas 61(7):1825–1838
31. Oloumi F, Rangayyan RM, Ells AL (2012) A graphical user interface for measurement of temporal arcade angles in fundus images of the retina. In: Canadian conference on electrical and computer engineering (CCECE), Proceedings of IEEE Canada 25th Annual, pages 4 on CD– ROM, Montreal, Canada

32. Hildebrand PL, Ells AL, Ingram AD (2009) The impact of telemedicine integration on resource use in the evaluation ROP, Analysis of the telemedicine for ROP in Calgary (TROPIC) database. Invest Ophthalmol Visual Sci, 50:E–Abstract 3151
33. De Silva DJ, Cocker KD, Lau G, Clay ST, Fielder AR, Moseley MJ (2006) Optic disk size and optic disk-to-fovea distance in preterm and full-term infants. Invest Ophthalmol Vis Sci 47(11):4683–4686
34. Rangayyan RM, Ayres FJ, Oloumi Faraz, Oloumi Foad, Eshghzadeh-Zanjani P (2008) Detection of blood vessels in the retina with multiscale Gabor filters. J Electron Imaging, 17(023018):1–7
35. Acton ST (2000) A pyramidal algorithm for area morphology. In Proceedings of IEEE international conference on image processing, 10–13, Vancouver, Canada
36. Metz CE (1978) Basic principles of ROC analysis. Seminars in Nuclear Medicine, 8(4):283–298
37. ROCKIT. Metz ROC Software. metz -roc.uchicago.edu/MetzROC/software
38. Ells AL, MacKeen LD (2008) Dynamic documentation of the evolution of retinopathy of prematurity in video format. J Am Assoc Pediatric Ophthalmol Strabismus 12(4):349–351
39. Chiang MF, Thyparampil PJ, Rabinowitz D (2010) Inter expert agreement in the identification of macular location in infants at risk for retinopathy of prematurity. Arch Ophthalmol 128(9):1153–1159

DCT-SVM-Based Technique for Off-line Signature Verification

B. H. Shekar and R. K. Bharathi

Abstract In this paper, a new method for off-line signature verification is proposed based on discrete cosine transform (DCT). The proposed approach has two stages, namely feature extraction and representation of signature using DCT followed by classification through support vector machine (SVM). The training signature samples are subjected to preprocessing to obtain binarized image, and DCT is employed on the binarized image. The upper-left corner block of size $m\ X\ n$ is chosen as a representative feature vector for each trained signature sample. These small feature vector blocks are fed as an input to the SVM for training purpose. The SVM is used as a verification tool and trained with different number of training samples including genuine, skilled, and random forgeries. The proposed approach produces excellent results on the standard signature databases, namely CEDAR, GPDS-160, and MUKOS—a Kannada signature database. In order to demonstrate the superiority of the proposed approach, comparative analysis is provided with many of the standard approaches.

Keywords Discrete cosine transform · Support vector machine · Signature verification

B. H. Shekar (✉)
Department of Computer Science, Mangalore University, Mangalore, Karnataka, India
e-mail: bhshekar@gmail.com

R. K. Bharathi
Department of Master of Computer Applications, S.J. College of Engineering, Mysore, Karnataka 570001, India
e-mail: rkbharathi@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_85,

1 Introduction

Biometric recognition in general forms a strong link between a person and his/her identity as the biometric traits cannot be easily shared, lost, or duplicated. Handwritten signature in particular happens to be the most natural behavioral biometric which generally established mode of providing the personal identity for authentication. The signatures have been established as one of the most acceptable means for authenticating a person's identity by many of the administrative and financial institutions [9]. Though human signature is a behavioral, natural, and intuitive, there exist few limitations in using this as the mode for personal identification and verification, such as there are some inconsistencies to a person's signature due to physical health, due to course of age, psychological or mental status, and so on [4]. The major threat/challenge in signature verification is the forgery of the signature, an illegal modification or reproduction of a signature. A signature is also considered forged if it is claimed that it was made by someone who did not make it [18]. Here, forgeries can be classified into three basic types: random forgery, when the forger has the access neither to the genuine signature nor to any information about the author's name, still reproduces a random signature; simple forgery, where the forger has no access to the sample signature but knows the name of the signer; and lastly, skilled forgery, when forger reproduces the signature, having access to the sample genuine signature. The signature is classified as online signature and off-line signature based on the procedure of capturing them. Online signature acquired at the time of its registration provides the intrinsic dynamic details, viz. velocity, acceleration, direction of pen movement, pressure applied and forces, whereas in off-line signature, the static image of the signature is captured once the signing process is complete. In spite of the fact that off-line signature lacks the dynamic details, it withholds many of the global and local features, viz. signature image area, height, width, zonal information, characteristic points such as end points, cross points, cusps, loops, presence and absence of zonal information. Devising an efficient and accurate off-line signature verification algorithm is a challenging task, as signatures are sensitive to geometric transformations, interpersonal signature in a corpus, complex background, skilled forgery, scalability of the system, noise introduced while capturing the image, difference in pen width, ink pattern, and so on. In this context, we proposed an efficient and simple to implement DCT-SVM technique for verification of the off-line signatures. The details of the proposed technique are brought out in the following subsections. The paper is organized as follows. Review of the related works is brought down in Sect. 2. In Sect. 3, we present the proposed technique. In Sect. 4, the experimental setup and a brief note on the dataset used in the experimentation along with the result analysis are presented. Comparative analysis of the proposed work with the work in literature is brought down in Sect. 5, followed by conclusion in Sect. 6.

2 Review of Related Work

Features play a vital role in the pattern recognition problems; hence, feature selection and extraction contribute significantly in the overall performance of any recognition algorithms. Off-line signatures are the scanned images of the signatures and possess three types of features, namely global, statistical, and geometrical features. The global features describe the signature image as a whole, statistical feature is extracted from the distribution of pixels of a signature image. The geometrical and topological features describe the characteristic geometry and topology of a signature, thereby preserving both global and local properties [1].

Some of the well-accepted off-line signature verification methods based on varying features and feature selection/extraction techniques have been introduced by many researchers. Huang and Yan [8] proposed a combination of static and pseudo-dynamic structural features for off-line signature verification. Shekar and Bharathi [19] concentrated on reducing the dimension of feature vectors, preserving the discriminative features obtained through principal component analysis on shape-based signature, and extended model is given based on kernel PCA [20]. Rekik et al. [16] has worked on the features consisting of different kinds of geometrical, statistical, and structural features. For comparison, they have used two baseline systems, global and local, both based on a larger number of features encoding the orientations of the stroke using mathematical morphology, and experimented on BioSecure DS2 and GPDS off-line signature databases. Kumar et al. [13] proposes a novel set of features based on surroundedness property of a signature image. The surroundedness property describes the shape of a signature in terms of spatial distribution of black pixels around a candidate pixel, which provides the measure of texture through the correlation among signature pixels in the neighborhood of that candidate pixel. They also developed the feature reduction techniques and experimented on GPDS-300 and CEDAR datasets. Kalera et al. [10] extracts the features based on quasi-multi-resolution technique using gradient, structural, and concavity (GSC) which was earlier used for identifying the individuality of handwriting, and experimented on CEDAR dataset. Nguyen and Bluemnstein [14, 15] concentrate on the global features derived from the total energy a writer uses to create their signature following the projections (both horizontal and vertical), which focuses on the proportion of the distance between the key strokes. They also described a grid-based feature extraction technique that utilizes directional information extracted from the signature contour and applied 2D Gaussian filter on the feature matrices. Chen and Srihari [2] tries to look at the online flavor in off-line signature by extracting the contours of the signature and combining the upper and lower contours neglecting the smaller middle portion to define a pseudo-writing path. To match two signatures, dynamic time warping (DTW), a nonlinear normalization is applied to segment them into curves, followed by extraction of features using Zernike moments (shape descriptor) and experimental results were given for CEDAR dataset. Ruiz-Del-Solar et al. [17] claims a new approach of using local interest

points in the signature image followed by computing local descriptors in the neighborhood of these points and later compared using local and global matching procedures. Verification is achieved through Bayes classifiers on GPDS signature corpus. Vargas et al. [21] focuses on pseudo-dynamic characteristics, that is, features presenting information of high-pressure points (HPP) of a handwritten signature contained in a grayscale image are analyzed. They used KNN and PNN standard model to analyze the robustness of simple forgeries of their own GPDS-160 signature corpus. Kumar et al. [12] comes with the objective in two-fold: one to propose a feature set based on signature morphology and the second is to use a feature analysis technique to get a more compact set of features that in turn makes the verification system faster through SVC on CEDAR signature database. Vargas et al. [22] work at the global image level and measure the gray-level variation in the image using statistical texture features, and the co-occurrence matrix and local binary pattern are analyzed and used as features. Experimental analysis is given using both random and skilled forgery for training and testing considering SVM classifier on MCYT and GPDS datasets. Ferrar et al. [5] present a set of geometric features based on the description of the signature envelope and the interior stroke distribution in polar and Cartesian coordinates and tested with different classifiers such as HMM, SVM, and Euclidean distance classifier on GPDS-160 dataset. Thus, we have seen different class of algorithm that aims to achieve better classification accuracy. In this context, we devised a transform-based approach that exhibit compact representation and SVM is used for accurate classification. The details of the proposed technique are brought out below.

3 Proposed Model

The proposed DCT-SVM technique of off-line signature verification involves three major phases: preprocessing, DCT-based feature extraction, and SVM-based verification. In the preprocessing stage, the signature images are binarized using Otsu's method to remove the complex background which might have occurred due to scanning, ink distribution, and so on. Once we binarize the signature image, the noise is eliminated using a simple morphological filter that results in a clear noise-free signature image. This preprocessed image is transformed to frequency domain using discrete cosine transform (DCT) (see Fig. 1).

Fig. 1 **a** Initial image. **b** Binarized and noise removed. **c** DCT-transformed image. **d** Block feature of DCT coefficients and their selection in zigzag pattern. **e** Marked *top-left* block of the image (feature block)

Discrete Cosine Transform: The DCT is a popular technique in image processing and video compression which transforms the input signal present in the spatial domain into a frequency domain. We proposed to use DCT-II in our work introduced by Wang [23]. The forward 2D DCT of M X N block image **f** is defined as

$$C(u,v) = \propto (u).\propto (v)\sum_{X=0}^{M-1}\sum_{Y=0}^{N-1} f(x,y)X\cos\left[\frac{\pi(2x+1)u}{2M}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (1)$$

and the inverse transform is defined as

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)C(u,v)X\cos\left[\frac{\pi(2x+1)u}{2M}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (2)$$

where

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, & u = 0 \\ \sqrt{\frac{2}{M}}, & u = 1,2,\ldots M-1 \end{cases} \quad (3)$$

and x and y are spatial coordinates in the image block, and u and v are coordinates in the transformed image. It is well-known fact that the discriminative features are available in the top-left portion of the image, say of size M_N, that is, the most energy being compacted to the low-frequency coefficients. The size of the subset is chosen such that it can sufficiently represent the input signature image and help in verification later. The size of the subset may be quite small compared to the whole vector with all the coefficients in it. It was observed that the DCT coefficients exhibit the expected behavior in which a relative large amount of information about the original signature image is stored/represented in fairly small number of coefficients [6]. The DCT coefficients, which are located at the upper-left corner, hold most of the image energy and have considered for further processing. For instance, the top $10 \times 10 = 100$ DCT coefficients are enough to represent the signature image and hence considered as a feature vector in our approach [11].

4 Experimental Setup and Result Analysis

The performance is evaluated through false acceptance rate (FAR) and false rejection rate (FRR). The FAR is the frequency that an unauthorized person signature (forge) is accepted as authorized (genuine), whereas the FRR is defined as the frequency that an authorized person (genuine) is rejected their access considering their signature to be forged. Hence, FAR is the percentage ratio of the total number of accepted forgeries to the total number of tested forgeries, whereas FRR is defined by the percentage ratio of total number of genuine rejected to the total number of tested genuine.

For the purpose of classification, we proposed to use support vector machine (SVM) classifier [7]. Unlike other techniques, where radial basis function (RBF) kernel [5] is used for training the system, we proposed to use linear kernel to reduce computational complexity. It is understood from the literature that the linear kernel performs well in terms of classification accuracy also. Training the SVM with suitable number of genuine and forgery (considered to be two-class problem) is a challenging task. We have conducted an exhaustive experimentation to exhibit the performance of the proposed algorithm with varying number of training and testing signature combinations. The training and testing samples include genuine signatures with random forgery and skilled forgery. Experimentation was conducted by extracting varying number of DCT coefficients from the DCT-transformed image, starting from dimension 8, that is, 64 coefficients, then to 10 with 100 coefficients and moved up to 14. Out of the experimental results analyzed, we decided to have 10×10 DCT coefficients as the feature vector for our rest of the experimentations. Extensive experimentation is conducted on publicly available databases, namely GPDS-160, CEDAR including one regional language database called MUKOS (Mangalore University Kannada Off-line Signature). Various combinations of samples are considered for training through SVM and hence testing to exhibit the efficacy of the proposed algorithm.

Initially, we started training the SVM with first 5 genuine and first 5 skilled forge samples from each signer of the dataset and tested with the remaining samples of genuine and skilled forge of corresponding signer. Later, we extended to train with first 10 genuine and first 10 skilled forge of each signer in the dataset and test with remaining samples. Similarly, we also extended to train with first 15 genuine and first 15 skilled forge samples and test with remaining samples of each signer of the dataset. The performance accuracy with different training sets on all databases considered is revealed in Table 1. Then, the experimentation was continued with training the SVM by 70 % of the genuine sample along with 100 randomly chosen forge (random forgery is the genuine sample of the other signer) from every database and tested with the remaining 30 % of genuine and other 50 randomly chosen forge. To justify the random selection, the above experimentations are repeated 5 times.

Table 2 gives the best performances of the DCT-SVM technique on all the databases considered and experimented with both random and skilled forgery. In the proposed technique, 0 % of false acceptance is observed when trained and tested with random forgery on CEDAR dataset. This was observed when we

Table 1 Performance accuracy (FAR and FRR) of the proposed DCT-SVM technique

	CEDAR		GPDS-160		MUKOS-Set-1	
No. of training samples	FAR	FRR	FAR	FRR	FAR	FRR
First 5 genuine + 5 skilled forgeries	8.32	10.64	11.77	13.56	11.44	13.43
First 10 genuine + 10 skilled forgeries	7.35	8.75	11.24	10.33	8.24	11.76
First 15 genuine + 15 skilled forgeries	6.44	7.28	11.64	7.63	4.18	10.49

Note Testing samples are the remaining signatures of the respective datasets

Table 2 Best performances of the DCT-SVM technique

Dataset	Experimentation with genuine and random forgery			Experimentation with genuine and skilled forgery		
	Accuracy	FRR	FAR	Accuracy	FRR	FAR
CEDAR	95.76	4.24	0.0	92.72	7.28	6.44
GPDS-160	96.29	3.70	0.046	92.47	7.63	11.64
MUKOS-Set-1	94.62	5.46	0.025	89.61	10.49	4.18
MUKOS-Set-2	95.59	4.41	0.063	91.75	8.22	6.54

trained the system with 12 genuine and 24 random forgeries of each signer, keeping the FRR to 4.24 % with an accuracy of 95.76 %. MUKOS-Set-1 has the accuracy of 89.61 %, with FAR of 4.18 %. Low FAR also projects the good performance of the model as it avoids the forge to be accepted as genuine.

4.1 Experimentation on MUKOS Database

We have also worked on one regional language off-line signature database called MUKOS. The MUKOS database consists of two sets: Set-1 consisting of 1,350 signatures and Set-2 with 760 signatures. In Set-1, we collected 30 genuine signatures and 15 skilled forgeries from 30 distinct signers. Each genuine signature was collected using black ink on A4 size white paper featuring 14 boxes on each paper. Once the genuine signatures were collected by all thirty signers, the forgeries were produced imitating a genuine signature from the static image of the genuine after a time gap where they were allowed to practice the forgery of other signers (other than genuine signers). In Set-2, we collected 20 genuine and 20 skilled forgery samples from 38(19 + 19) different individuals. Skilled forgery was obtained by arbitrarily selecting the people who in practice used English for their genuine signature. These signatures were acquired with a standard scanner with 75 dpi resolution in an 8-bit grayscale image. MUKOS performance is exhibited in Table 3, with convincing accuracy of 89.61 % for Set-1 and 91.75 % on Set-2.

Table 3 Experimental results for MUKOS dataset

Proposed by	Feature	Classifier	Accuracy	FAR	FRR
Shekar and Bharathi [19]	Shape-based eigen-signature	Euclidean distance measure	93.00	11.07	6.40
Proposed algorithm Set-1	DCT coefficients	SVM	89.61	10.49	4.18
Set-2		SVM	91.75	8.22	6.54

5 Comparative Analysis

Apart from the above experimentations, we have also conducted the experiments as proposed in the literature, which we have considered for result analysis. For instance, [14] in each test, 12 genuine signatures and 400 random forgeries were employed for training, and remaining 59 genuine signatures were chosen from the remaining 59 writers along with the remaining 12 genuine samples for testing from the GPDS-160 corpus. Kumar et al., [12] have selected 5 individuals randomly and kept them aside as test data and the remaining 50 individuals comprise the training data with a combination of 274 genuine–genuine pairs and 576 genuine-forged pairs of each individual from CEDAR corpus.

5.1 Experimentation on CEDAR Database

In Table 4, we present the results on CEDAR database. It shall be observed that the proposed DCT-SVM-based approach exhibit better performance when compared to other approaches. Although the classification accuracy is similar to that of [3], it should be noted that the number of feature used are only 100. High-performance rate is shown by the proposed technique when tested on CEDAR dataset (both with skilled and random forge samples). To boost the performance, 0 % false acceptance has been recorded when trained and tested with random forgery, indicating 100 % rejection of forge samples.

5.2 Experimentation on GPDS-160 Database

GPDS-160, a subcorpus of GPDS-300, is used for experimentation for the analysis of the performance of DCT-SVM technique. Extensive experimentation was conducted with this dataset as the literature has revealed a variety of combinations of number of training and testing samples. However, the proposed DCT-SVM technique out performed considerably high with the accuracy of 92.47 % when

Table 4 Experimental results obtained for CEDAR dataset

Model proposed by	Feature type	No. of features	Classifier	Accuracy	FAR	FRR
Kalera et al. [10]	Word shape (GSC)	1,024	PDF	78.50	19.50	22.45
Chen and Srihari [2]	Zernike moments	640	DTW	83.60	16.30	16.60
Kumar et al. [12]	Signature morphology	44	SVM	88.41	1,159	11.59
Chen and Srihari [3]	Graph matching	>1,024 + 8	DTW	92.10	8.20	7.70
Kumar et al. [13]	Surroundedness	24 out of 44	SVM	91.67	8.23	8.33
Proposed algorithm	DCT coefficients	100	SVM	92.72	7.28	6.44

Table 5 Experimental results obtained for GPDS-300 corpus

Model proposed by	Feature type	No. of features	Classifier type	Accuracy	FAR	FRR
Ferrar et al. [5]	Geometric features	42 * 2 + 22 * 2 sequence	SVM	86.65	13.12	15.41
			HMM		12.60	14.10
Vargas et al. [21]	High-pressure points (HPP)	Two vectors * no. of HPP	KNN & PNN	87.67	14.66	10.01
Vargas et at. [22]	GLCM + LBP	>(4 + 4 * radius)	SVM	87.28	6.17	22.49
Solar et al. [17]	Local interest points + local descriptors	12 * no. of descriptors	Bayesian	84.70	14.20	16.40
Kumar et al. [13]	Surroundedness	29 selected	SVM	86.21	13.76	13.76
Nguven and Blumenstein [15]	MDF, energy, maxima and ratio	Every black pixel	SVM	82.25	17.25	17.25
Nguven & Blumenstein [14]	Grid based + 2D Gaussian filter	(12 * 6 * 4) * no. of grids formed	SVM	86.32	13.68	14.18
Proposed model	DCT coefficients	100	SVM	92.47	7.63	11.64

first 15 genuine sample and 15 skilled forgeries of every signer is used to train the SVM, and remaining 9 genuine and 15 skilled forgery samples of every signer is used for testing. The detail result analysis is shown in Table 5.

6 Conclusion

We have proposed an efficient, robust, and less computational cost technique for off-line signature verification. The prominent DCT features of a preprocessed signature are fed to the layers of SVM for training and hence recognized the genuine and forged signatures by testing the samples. Exhaustive experimentations were conducted to exhibit the performance of the proposed technique with varying training and testing configurations. We also demonstrated the performance on various datasets with state-of-the-art models and databases. The success of the proposed technique performance is revealed through FAR and FRR.

References

1. Arya S, Inamdar VS (2010) A preliminary study on various off-line hand written signature verification approaches. Int J Comput Appl 1(9):55–60
2. Chen S, Srihari S (2005) Use of exterior contours and shape features in off-line signature verification. In: ICDAR, pp 1280–1284
3. Chen S, Srihari S (2006) A new off-line signature verification method based on graph matching. In International conference on pattern recognition (ICPR06), vol 2, pp 869–872
4. Fairhurst MC (1997) Signature verification revisited: Promoting practical exploitation of biometric technology. Electron Commun Eng J 9:273–280
5. Ferrer M, Alonso J, Travieso C (2005) Off-line geometric parameters for automatic signature verification using fixed-point arithmetic. IEEE Trans Pattern Anal Mach Intell 27(6):993–997
6. Hafed ZM, Levine MD (2001) Face recognition using the discrete cosine transform. Int J Comput Vis 43(3):167–188
7. Hsu C-W, Chang C-C, Lin C-J (2003) A practical guide to support vector classification
8. Huang K, Yan H (2002) Off-line signature verification using structural feature correspondence. Pattern Recogn 35:2467–2477
9. Impedovo D, Pirlo G (2008) Automatic signature verification: the state of the art. IEEE Trans Syst Man Cybern C (Appl Rev) 38(5):609–635
10. Kalera MK, Srihari S, Xu A (2003) Off-line signature verification and identification using distance statistics. Int J Pattern Recogn Artif Intell 228–232
11. Khayam S (2003) The discrete cosine transforms (DCT): theory and application. Michigan State University
12. Kumar R, Kundu L, Chanda B, Sharma JD (2010) A writer-independent offline signature verification system based on signature morphology. In: Proceedings of the first international conference on intelligent interactive technologies and multimedia, IITM'10, pp 261–265, New York, NY, USA. ACM
13. Kumar R, Sharma JD, Chanda B (2012) Writer-independent off-line signature verification using surroundedness feature. Pattern Recogn Lett 33(3):301–308

14. Nguyen V, Blumenstein M (2011) An application of the 2D Gaussian filters for enhancing feature extraction in offline signature verification. In ICDAR'11, pp 339–343
15. Nguyen V, Blumenstein M, Leedham G (2009) Global features for the offline signature verification problem. In: Proceedings of the 2009 10th international conference on document analysis and recognition, pp 1300–1304, Washington, DC, USA. IEEE Computer Society
16. Rekik Y, Houmani N, Yacoubi MAE, Garcia-Salicetti S, Dorizzi B (2011) A comparison of feature extraction approaches for offline signature verification. In IEEE international conference on multimedia computing and systems, pp 1–6
17. Ruiz-Del-Solar J, Devia C, Loncomilla P, Concha F (2008) Offline signature verification using local interest points and descriptors. In: Proceedings of the 13th Iberoamerican congress on pattern recognition: progress in pattern recognition, image analysis and applications, CIARP'08. Springer, Berlin, pp 22–29
18. Saikia H, Sarma KC (2012) Approaches and issues in offline signature verification system. Int J Comput Appl 42(16):45–52
19. Shekar BH, Bharathi RK (2011) Eigen-signature: a robust and an efficient offline signature verification algorithm. In International conference on recent trends in information technology (ICRTIT), June 2011, pp 134–138
20. Shekar BH, Bharathi RK, Sharmilakumari M (2011) Kernel eigen-signature: an offline signature verification technique based on kernel principal component analysis. In (Emerging Applications of Computer Vision), EACV-2011 Bilateral Russian-Indian Scientific Workshop, Nov 2011
21. Vargas FJ, Ferrer MA, Travieso CM, Alonso JB (2008) Off-line signature verification based on high pressure polar distribution. In: Proceedings of the 11th international conference on frontiers in handwriting recognition, ICFHR 2008, Aug 2008, pp 373–378
22. Vargas J, Ferrer M, Travieso C, Alonso J (2011) Off-line signature verification based on grey level information using texture features. Pattern Recogn 44(2):375–385
23. Wang Z (1984) Fast algorithms for the discrete *w* transform and for the discrete Fourier transform. IEEE Trans Acoust Speech Signal Process 32(4):803–816

Headset Prototype Design for Industrial Noise Reduction Using DSP

K. V. Mahendra Prashanth and V. Sridhar

Abstract In many industries, depending on the noise characteristics, when the noise level exceeds permissible limits, hearing protectors are being used. Passive-type hearing protectors are effective for noise of frequencies in the range ≥500 Hz. However, the majority of the noise-generating machineries produce noise below 500 Hz. From the studies, it is observed that the non-auditory physiological systems of workers, who have exposed to these low-frequency noises, are being severely affected. This in turn leads to reduce their effectiveness at the work place. This paper describes the design and implementation of a headset prototype, based on filtered-x least mean square (FXLMS) algorithm using DSP to reduce noise levels. The industrial noise spectrum power reduction of around 23–30 dB in the frequency range of 0–500 Hz has been accomplished.

Keywords Industrial noise · Active noise control (ANC) · FXLMS algorithm

1 Introduction

In most of the industries, dominant noise frequencies from the machinery are of low- and mid-frequency characteristics [1]. It is evident that interference of noise is a serious health issue that is spreading across its effects more than the auditory effects. In industries, both passive and active noise headsets are being used. The passive-type noise protectors' performances well at mid- and high frequencies (≥500 Hz). But they are unsuitable for low frequencies. Since most of the noise-

K. V. M. Prashanth (✉)
Department of Electronics and Communication Engineering, S. J. B. Institute of Technology, Bangalore, India
e-mail: kvmprashanth@yahoo.com

V. Sridhar
P. E. S. College of Engineering, Mandya, India
e-mail: venusridhar@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_86,

generating machinery produces noise below 500 Hz, these noise protectors will not be effective in noise reduction. Noise cancellation in time-varying environments, adaptive filters is suitable [2–4]. Digital signal processing technology is being used in a wide range of applications, such as in communications, controls, and signal processing [5]. Real-time processing is becoming more applicable to a wide range of problems [6]. Extensive research has been done in the field of headrest systems [7, 8]. Successful applications have also been achieved in propeller aircraft cabin, helicopter cabin noise control, and in the automobile industry, controlling interior noise in cars [9, 10].

Though the concept of active noise control (ANC) has been known from the last four decades, its applications in hearing protector have become prominent recently with the support of digital systems. Active noise control suits very much for low-frequency quieting treatment. ANC headphones have achieved widespread commercial success. Headphone is one of the most convenient structures to apply ANC. Headphones with ANC are available as commercial products [11–14], and some of them have quite satisfactory ANC performance in many practical situations like cockpits, airport ground personnel, where noise reduction is definitely relaxing for human beings.

This paper explores implementation of adaptive ANC algorithm applied on a prototype headphone system, and performance is evaluated. Filtered-x least mean square (FXLMS) algorithm [15] was implemented using DSP.

2 Experimental Setup

Figure 1 illustrates the block diagram of the experimental setup of a headset prototype for noise reduction. The setup was designed using DSP. The personal computer and DSP starter kit (TMS320C6713) were inter-connected using USB. The audio daughter card (DSK_AUDIO4) was mounted on DSK, using the slot provided. The audio daughter card was programmed to set amplifier gains, select the input source, and control various codec capabilities. The ANC system uses the reference microphone to pick up the input noise, processes this input with an adaptive filter to generate an anti-noise to cancel the primary noise acoustically, and uses an error microphone to measure the error and to update the adaptive filter coefficients (see Figs. 1, 2). Code Composer Studio (CCS) was used as the integrated development environment (IDE) for coding, debugging, and configuration.

3 Filtered-xLMS Algorithm

Due to its robustness and relative ease of implementation, FXLMS algorithm [16] is used in most of the ANC applications. The conventional least mean square (LMS) algorithm has to be modified to ensure convergence. A solution is to place

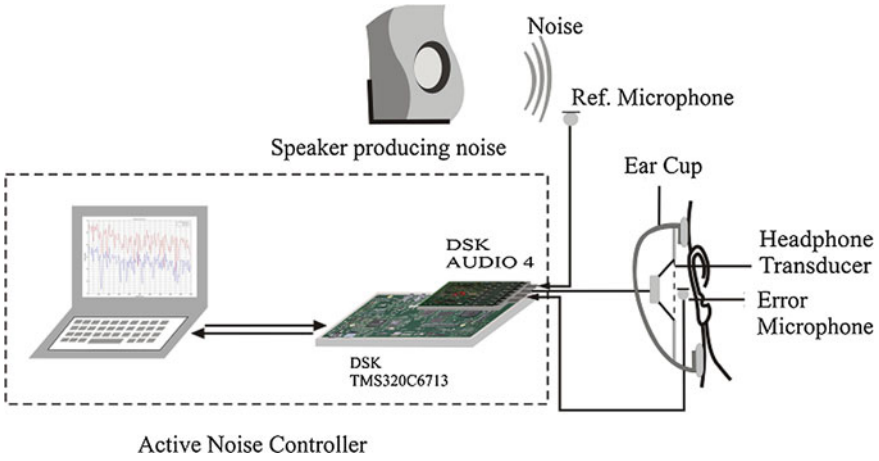


Fig. 1 Block diagram of the experimental setup

Fig. 2 A snapshot of the experimental setup

an identical filter of $S(z)$ to the reference signal path to the weight update of the LMS algorithm. As the input is being filtered, this algorithm is called FXLMS algorithm. The placement of the secondary path transfer function following the digital filter $w(z)$ controlled by the LMS algorithm is shown in Fig. 3.

The FXLMS configuration is as shown in Fig. 3, and relevant expressions [16] are as follows:

The residual error is

$$e(n) = d\,(n) - \; s(n) * \; \left[w^T(n)\, x(n)\right] \tag{1}$$

where $s(n)$ is the impulse response of the secondary path $S(z)$ at time n.

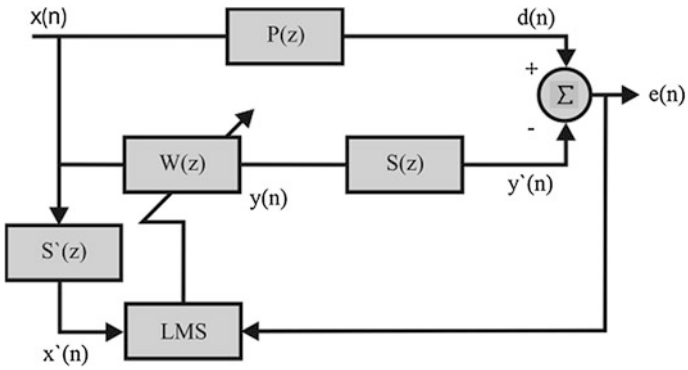


Fig. 3 Block diagram of ANC system using the FXLMS algorithm [15]

$$\mathbf{w}(n) = [w_0(n)\, w_1(n) \ldots w_{L-1}(n)]^T$$

is the coefficient vector of w(z) at time n,

$$\mathrm{x}(n) = [x(n)x(n-1) \ldots x(n-L+1)]^T$$

is the signal vector at time *n,* and *L* is the order of the filter $w(z)$.

The adaptive filter minimizes the instantaneous squared error,

$$\xi\hat{}(n) = e^2(n)$$

Using LMS algorithm, the coefficient is updated in the negative gradient direction with step size μ, thus,

$$w(n+1) = w(n) - \mu/2\,[\nabla \xi\hat{}(n)] \tag{2}$$

where $\xi(n)$ is an instantaneous estimate of the MSE gradient at time n and can be expressed as

$$\nabla \xi = \nabla^2 e(n) = 2[\nabla e\,(n)]e\,(n)$$

From Eq. (1),

$$\nabla e(n) = -\,s(n) * x(n) = -x\hat{}(n)$$

where

$$x\hat{}(n) = [x\hat{}(n)x\hat{}(n-1) \ldots x\hat{}(n-L+1)]^T \textit{ and } \mathbf{x}\hat{}(n) = s(n) * \mathbf{x}(n)$$

Therefore, the gradient estimate becomes

$$\nabla \xi\hat{}(n) = -2\mathrm{x}\hat{}(n)e(n) \tag{3}$$

Substituting Eq. (3) in (2), the adaptive filter coefficient update for the FXLMS algorithm is

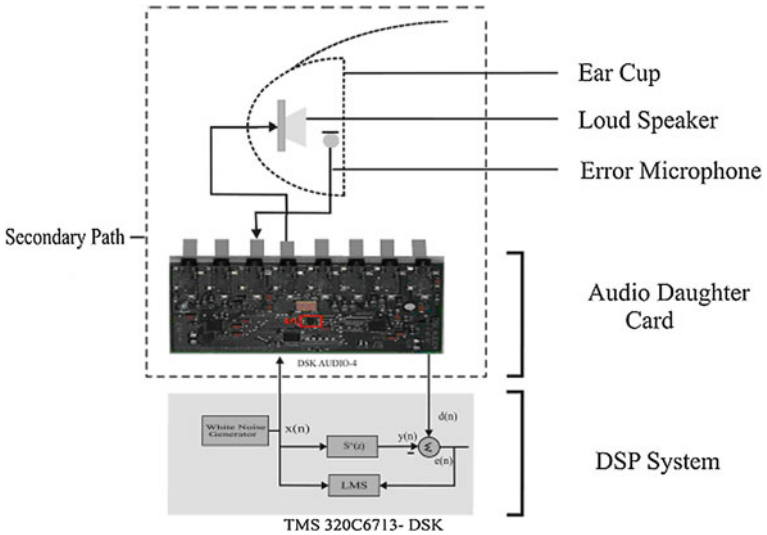


Fig. 4 Experimental setup for off-line secondary path modeling

$$w(n+1) = w(n) + \mu \mathbf{x}'(n)e(n)$$

In the FXLMS algorithm, the filtered reference signal has to be generated. The stability of the algorithm depends on the accuracy of the estimated filter $S'(z)$), modeling the true secondary path [16]. The FXLMS algorithm requires knowledge of the secondary path transfer function $S(z)$. Assuming that the characteristics of $S(z)$ are time invariant but unknown, $S(z)$ can be estimated in the absence of primary noise. This is called off-line modeling, and this estimate $S'(z)$) can be used for the ANC operation(see Fig. 4) [17–20]. Signal processing was carried out on DSP with an audio daughter board (DSK_AUDIO4). The LMS algorithm was implemented in a fixed point C using a varying number of filter taps to identify the system. The generated white noise was used as the training signal. The experimental setup used to estimate the secondary path transfer function is shown in Fig. 4.

4 Results and Discussion

The performance of the designed headset prototype system has been evaluated considering recorded industrial noise samples. In the tests, 128-tab filters have been used as the secondary path model and noise cancellation filters. In the figure (see Figs. 5, 6), solid red-lined plots depict the actual noise levels (ANC off) of the recorded samples and solid blue-lined plots show the reduced noise levels inside the ear cup of the headset prototype (ANC on). From the performance test results,

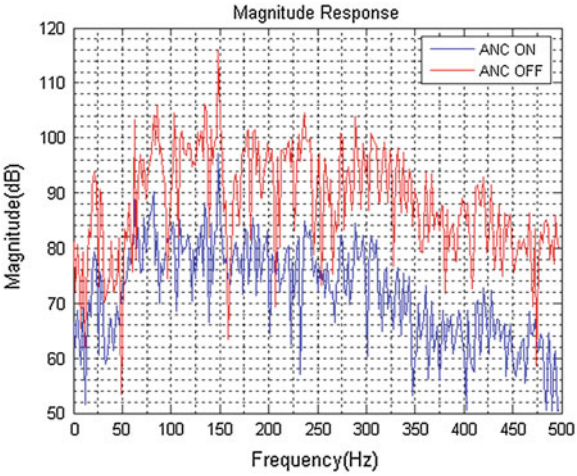


Fig. 5 Performance test of headset prototype system (industrial noise sample 1)

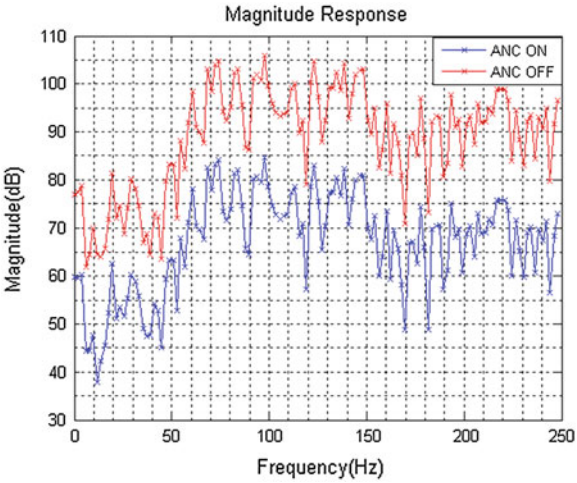


Fig. 6 Performance test of headset prototype system (industrial noise sample 2)

it is observed that the attenuation achieved by the designed headset prototype system varied at different frequencies see (Figs. 5, 6). The reason for this phenomenon is that every signal path has a different gain contribution at different frequencies of the input signal.

5 Conclusion

The paper described the design and development of an ANC-based headset prototype using DSP. These results show a significant noise reduction. The approach implemented has yielded a good overall response. The headset prototype system that has been developed works effectively in reducing industrial noise samples, which is useful for workers in a noisy environment. The performance evaluation results show that the prototype system can effectively attenuate industrial noises with a maximal power reduction in dominant noise spectrum by 23–30 dB in the low-frequency range up to 500 Hz.

References

1. Mahendra Prashanth KV, Sridhar V (2008) Relationship between noise frequency components and physical, physiological and psychological effects among industrial workers. Noise Health 10(23):90–98
2. Haykin S (2001) Adaptive filter theory, 4th edn. Pearson Education (Singapore) Pte Ltd, Delhi, pp 231–311
3. Widrow B, Stearns SD (2004) Adaptive signal processing. Pearson Education (Singapore) Pte. Ltd., Delhi
4. Omoto et al (2002) Behaviour of adaptive algorithms in active noise control systems with moving noise sources. Acoust Sci Tech 23(2):84–89
5. Chassing R (2005) "Digital signal processing and applications" with the C6713 and C6416 DSK. Wiley, New York
6. Kuo SM, Gan GS (2005) Digital signal processors architectures, implementations, and applications. Prentice Hall, Englewood Cliffs
7. Pawelczyk M (2004) Adaptive noise controls algorithms for active headrest system. Control Eng Pract 12:1101–1112
8. Rafaely B, Elliott SJ, Garcia-Bonito J (1999) Broadband performance of an active headrest. J Acoust Soc Am 106(2):787–793
9. Garcia-Bonito J, Elliott SJ, Boucher CC (1997) Generation of zones of quiet using a virtual microphone arrangement. J Acoust Soc Am 101:3498–3516
10. Kuo SM, Morgan DR (1996) Active noise control systems: algorithms and DSP implementations. Wiley, New York
11. Hansen CH (2005) Current and future industrial applications of active noise control. Noise Control Eng J 53(5), Sept–Oct 2005
12. Bose Corporation Bose, Worlds Leader in audioElectronics. Available at: www.bose.com
13. Sennheiser Company Sennheiser worldwide-evolution Kopf, rerMikrofonAkustik. Available at: www.sennheiser.com
14. Bateman A, Paterson-Stephens I (2002) The DSP handbook algorithms, applications and design techniques. Prentice hall, Englewood Cliffs
15. Kuo SM, Morgan DR (1999) Active noise control: a tutorial review. Proc IEEE 87(6):943–973
16. Kuo SM, Morgan DR (1999) Active noise control: a tutorial review. Proc IEEE 87(6), June 1999
17. Thom J, Peters C, McIntyre E, Winters M (2005) Active noise control communication headsets for the entertainment industry, pp 2–24

18. Solo V, DeKalb I (1995) Adaptive signal processing algorithms. Prentice hall, Englewood cliffs
19. Morgan DR (1980) Analysis of multiple correlation cancellation loop with a filter in the auxiliary path. IEEE Trans ASSP ASSP-28(4):454–467
20. Widrow B (1975) Adaptive noise canceling: principles and applications. Proc IEEE 63(12):1692–1716

Optimal Location of Series FACTS Device Using PSO Technique to Reduce the Losses and to Enhance Power Transfer Capability in a Power System

Usha Surendra and S. S. Parthasarathy

Abstract Electric power systems, around the globe, are changing in terms of structure, operation, management, and ownership due to technical, financial, and ideological reasons. Recent trend involves augmentation of power systems in terms of geographical area, assets additions, and penetration of new technologies in generation, transmission, and distribution sectors. In this regard, flexible alternative current transmission system (FACTS) devices play a key role in enhancing controllability and increasing power transfer capability of the network. Thyristor-controlled series compensator (TCSC) is an emerging FACTS device designated to achieve this objective. The conventional methods in solving optimization problems in power systems suffer from several limitations due to necessities of derivative existence, providing suboptimal solutions, etc. Computational intelligence plays an important role in determining the optimal solutions for multiobjective functions. A combinatorial analysis problem in power systems can be solved by modern heuristic methods in finding optimal solution. Thus, in this paper, particle swarm optimization (PSO), a wing of evolutionary computation (EC) encapsulated with heuristic approach, has proposed for finding optimal location of TCSC. An IEEE standard 5-bus and 14-bus systems have been considered to test the credibility of the proposed algorithm. The simulation result proved the efficiency of the proposed approach by optimal placement of the FACTS device to minimize the losses and to improve the power transfer in a power system network.

U. Surendra (✉)
EEE Department, NIE Institute of Technology, Mysore, Karnataka, India
e-mail: us_1963@yahoo.com

S. S. Parthasarathy
EEE Department, PES College of Engineering, Mandya, Karnataka, India
e-mail: vsarathypartha@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_87,

Keywords Computational intelligence · Particle swarm optimization (PSO) · Flexible alternative current transmission system (FACTS) · Thyristor-controlled series compensator (TCSC) · Loss reduction · Power transfer capability

1 Introduction

Electric power systems have forced to operate to almost their full capacities due to the environmental and/or economic constraints to build new generating plants and transmission lines. The amount of electric power that can transmit between two locations through a transmission network is limited by security and stability constraints. Power flow in the lines and transformers should not be allowed to increase to a level where a random event could cause the network collapse because of angular instability, voltage instability, or cascaded outages [1]. Hence, economic operation of power system along with the assurance of refined quality of power supply to consumers is a challenging task. Due to the introduction of deregulation in electricity market, installation of flexible alternative current transmission system (FACTS) devices has become inevitable [3]. Because of the economic considerations, installation of FACTS controllers in all the buses or the lines in a system is not feasible. There are several sensitivity-based methods for finding optimal locations of FACTS devices in power systems described [4]. However, it is required to find the optimal location of FACTS devices by heuristic method to overcome both economical and technical barriers in accomplishing the objective. The use of thyristor-controlled series compensator (TCSC) which is a FACTS device gives a number of benefits for the user of the grid, all contributing to increase the power transmission capability of new as well as existing transmission lines. These benefits include improvement in system stability, voltage regulation, reactive power balance, load sharing between parallel lines, and reduction in transmission losses [2]. Optimal location of TCSC is a task assigned to particle swarm optimization (PSO) where PSO is an approach to find optimal solutions for search problems through application of the principles of swarm intelligence technique [4]. The sensitivity analysis has been carried out for the location of the different FACTS devices for the solution of optimal power flow [5]. System load ability can be increased using evolutionary strategies that use different types of FACTS controllers have been discussed [6]. The application of PSO method for location of FACTS devices to achieve maximum system load ability with minimum cost of installation of FACTS devices has been evaluated [7].

This paper has focused on loss reduction and to increase the power transfer capability of the transmission line by optimal location of TCSC using PSO technique in a power system. An overview of modeling of TCSC and brief description of PSO are presented in Sects. 2 and 3.1, respectively. Optimal location of TCSC using PSO has explained in Sect. 4.1.2, which elaborates the traits of an IEEE 5-bus and 14-bus systems under consideration. It also reveals the results

obtained on applying the evolutionary strategy to solve the optimization problem. Finally, the conclusion of the paper is described in Sect. 5.

2 Thyristor-Controlled Series Compensator Modeling

The IEEE TCSC is a capacitive reactance compensator, which consists of three main components: capacitor bank C, bypass inductor L, and bidirectional thyristors SCR1 and SCR2 [8]. Series capacitive compensation has used to increase line power transfer as well as to enhance system stability. Figure 1 shows the main circuit of a TCSC.

The firing angles of the thyristors are controlled to adjust the TCSC reactance according to the system control algorithm, normally in response to some system parameter variations. According to the variation in the firing angle, this process can modeled as a fast switch between corresponding reactance offered to the power system. Assuming that the total current passing through the TCSC is sinusoidal, the equivalent reactance at the fundamental frequency can represented as a variable reactance X_{TCSC}.

The TCSC can control to work in either the capacitive or the inductive zones avoiding steady-state resonance [8]. There exists a steady-state relationship between the firing angle α and the reactance X_{TCSC}, as described by the following equation [8]:

$$X_{\mathrm{TCSC}}(\alpha) = \frac{X_C X_l(\alpha)}{X_l(\alpha) - X_C} \tag{1}$$

where,

$$X_l(\alpha) = X_L \frac{\pi}{\pi - 2\alpha - \sin\,\alpha} \tag{2}$$

where α is the firing angle, X_L is the reactance of inductor, and X_l is the effective reactance of inductor at firing angle [8].

A model of transmission line with a TCSC connected between bus i and bus j is shown in Fig. 2. During the steady state, the TCSC has considered as a static

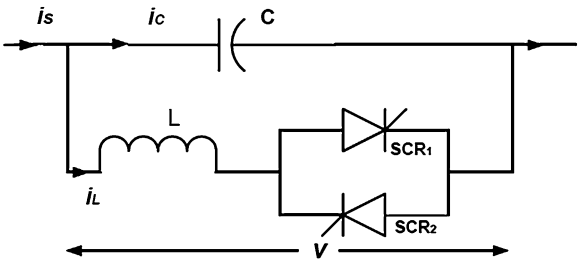


Fig. 1 Configuration of a TCSC

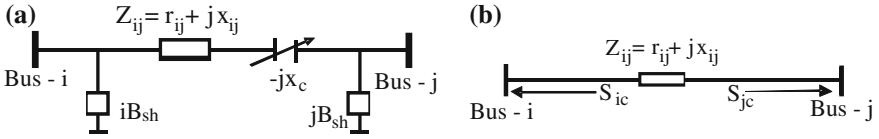


Fig. 2 **a** TCSC model. **b** Injection model of TCSC

reactance $-jx_c$. The controllable reactance x_c is directly used as a control variable in the power flow equations.

The real power injections at bus $i(P_{ic})$ and bus $j(p_{jc})$ are given by the following equations [2]:

$$P_{ic} = V^2{}_i\,\Delta G_{ij} - V_i Vj\left[\Delta G_{ij}\cos\delta_{ij} + \Delta B_{ij}\sin\delta_{ij}\right] \tag{3}$$

$$P_{jc} = V^2{}_j\,\Delta G_{ij} - V_i Vj\left[\Delta G_{ij}\cos\delta_{ij} - \Delta B_{ij}\sin\delta_{ij}\right] \tag{4}$$

Similarly, the reactive power injections at bus $i(Q_{ic})$ and bus $j(Q_{jc})$ can be expressed as follows:

$$Q_{ic} = -V^2{}_i\,\Delta B_{ij} - V_i Vj\left[\Delta G_{ij}\sin\delta_{ij} - \Delta B_{ij}\cos\delta_{ij}\right] \tag{5}$$

$$Q_{jc} = -V^2{}_j\,\Delta B_{ij} + V_i Vj\left[\Delta G_{ij}\sin\delta_{ij} + \Delta B_{ij}\cos\delta_{ij}\right] \tag{6}$$

where

$$\Delta G_{ij} = \frac{x_c r_{ij}(x_c - 2x_{ij})}{\left(r_{ij}^2 + x_{ij}^2\right)\left(r_{ij}^2 + \left(x_{ij} - x_c\right)^2\right)} \tag{7}$$

$$\Delta B_{ij} = \frac{-x_c\left(r_{ij}^2 - x_{ij}^2 + x_c x_{ij}\right)}{\left(r_{ij}^2 + x_{ij}^2\right)\left(r_{ij}^2 + \left(x_{ij} - x_c\right)^2\right)} \tag{8}$$

where ΔG_{ij} and ΔB_{ij} are the changes in conductance and susceptance of the line i and line j, respectively.

This model of TCSC is used to properly modify the parameters of transmission lines with TCSC for optimal location.

3 Proposed Method

3.1 Particle Swarm Optimization

The inherent rule adhered by the members of birds and fishes in the swarm enables them to move and synchronize, without colliding, resulting in an amazing

choreography, which was the basic idea of PSO technique. PSO is similar to evolutionary computation (EC) techniques in which a population of potential solutions to the problem under consideration has used to probe the search space. The major difference between the EC techniques and swarm intelligent (SI) techniques is that EC technique uses genetic operators, whereas SI techniques use the physical movements of the individuals in the swarm. PSO has developed through simulation of bird flocking in two-dimensional space. The position of each agent is represented in *X–Y* plane with position (*Sx*, *Sy*), *Vx* (velocity along *X*-axis), and *Vy* (velocity along *Y*-axis). Modification of the agent position has realized by the position and velocity information. Bird flocking optimizes a certain objective function. Each agent knows its best value so far, called 'Pbest', which contains the information on position and velocities. This information is the analogy of personal experience of each agent. Moreover, each agent knows the best value so far in the group, 'Gbest' among all 'Pbest'. This information is the analogy of knowledge, how the other neighboring agents have performed. Each agent tries to modify its position by considering current positions (*Sx*, *Sy*), current velocities (*Vx*, *Vy*), the individual intelligence (Pbest), and the group intelligence (Gbest).

The following equations are utilized, in computing the position and velocities, in the *X–Y* plane:

$$V_i^{k+1} = W \times v_i^{k} + C1 \times \text{rand}_1 \times \left(P_{\text{besti}} - s_i^{k}\right) + C_2 \times \text{rand}_2 \times \left(G_{\text{best}} - s_i^{k}\right) \tag{9}$$

$$s_i^{k+1} = s_i^{k} + v_i^{k+1} \tag{10}$$

where

v_i^{k+1}	Velocity of *i*th individual at $(k + 1)$th iteration;
v_i^{k}	Velocity of *i*th individual at *k*th iteration;
W	Inertial weight;
C_1, C_2	Positive constants both equal to 2;
rand_1	Random number selected between 0 and 1;
rand_2	Random number selected between 0 and 1;
P_{besti}	Best position of the *i*th individual;
G_{best}	Best position among the individuals (group best);
s_i^{k}	Position of *i*th individual at *k*th iteration;

The velocity of each agent has modified according to (9), and the position has modified according to (10). The inertia weight '*W*' is modified using (11), to enable quick convergence.

$$W = W_{\max} - \frac{(W_{\max} - W_{\min})}{\text{iter}_{\max}} \times \text{iter} \tag{11}$$

where

$W_{\max}$	Initial value of inertia weight;
$W_{\min}$	Final value of inertia weight;
Iter	Current iteration number;
$\text{Iter}_{\max}$	Maximum iteration number

3.2 Problem Formulation Equations

Equations that considered for optimization to reduce the losses are given below.

$$P_{gi} - P_{di} = \sum_{j=1}^{N_B} \left|V_i\right|\left|V_j\right|\left|Y_{ij}\right| \cos\left(\delta_i - \delta_j - \theta_{ij}\right) \tag{12}$$

$$Q_{gi} - Q_{di} = \sum_{j=1}^{N_B} \left|V_i\right|\left|V_j\right|\left|Y_{ij}\right| \sin\left(\delta_i - \delta_j - \theta_{ij}\right) \tag{13}$$

$$P_{gi,\min} \le P_{gi} \le P_{gi,\max}$$

$$Q_{gi,\min} \le Q_{gi} \le Q_{gi,\max} \tag{14}$$

$$P_{di,\min} \le P_{di} \le P_{di,\max}$$

$$Q_{di,\min} \le Q_{di} \le Q_{di,\max} \tag{15}$$

$$V_{i,\min} \le V_i \le V_{i,\max} \tag{16}$$

$$-0.8\,X_L \le X_{\text{TCSC}} \le 0.2\,X_L \text{ p.u} \tag{17}$$

$$\text{Ploss}_{ij} = \sum_{K=1}^{N} G_{ij}\left(V_{1ij}^2 + V_{2ij}^2 - 2\left(V_{1ij}V_{2ij}\cos(\delta_1 - \delta_2)\right)\right) \tag{18}$$

$$\min \sum_{i=1}^{N_G} C_{Gi}(P_{Gi}) - \sum_{i=1}^{N_D} B_{Di}(P_{Di}) \tag{19}$$

P_{gi}, Q_{gi} are the real and reactive power generation at bus i.
P_{di}, Q_{di} are the real and reactive power demands at bus i.
V_i, δ_i are voltage and angles at bus i.
$P_{gi,\min}, P_{gi,\max}$ real power minimum and maximum generation limits at bus i.

$Q_{gi,\min}, Q_{gi,\max}$ reactive power minimum and maximum generation limits at bus i.
$P_{di,\min}, P_{di,\max}$ real power minimum and maximum demand limits at bus i.
$Q_{di,\min}, Q_{di,\max}$ reactive power minimum and maximum demand limits at bus i.
In the objective function, $C_{Gi}(P_{Gi})$ is cost function for generating real power P_{Gi} at bus i, and $B_{Di}(P_{Di})$ is the demand function.
N number of lines
V voltage at the bus

3.3 Proposed Algorithm for Location of TCSC

The algorithm steps for the proposed optimal placement of TCSC device with PSO are as follows:

Step 1: Initialize the line and bus parameters, particle size, and maximum number of iterations.
Step 2: The initial population of individuals is created satisfying the FACTS device constraints given by (17), and it has verified that only one device is placed in each line individually.
Step 3: Using initial parameters, run load flow to find losses using Newton–Raphson method.
Step 4: Populate the dependent variable, that is, voltage between its limits with the number of populations mentioned.
Step 5: Calculate the power loss using dependent variables.
Step 6: The minimum power loss and its position are taken as the P_{best}(position) and fP_{best}(power loss)
Step 7: Update each population with the new population and run load flow with modified admittance matrix, which reflects the reactive power injected variation.
Step 8: Iteration starts from this point.
Step 9: Process from Step 4 and Step 5 repeated until the stopping criterion, which is the total number of iteration, reached.
Step 10: The P_{best} value becomes the best for the 1st iteration.
Step 11: From the 2nd iteration, the new P_{best} compared with the Gbest, that is, update and interchanges for the lower power loss.
Step 12: Check whether the final best individual obtained satisfies the above equations, which means that the load voltage magnitude deviations and real power losses are minimum.
Step 13: After the final iteration, the G_{best} value and its position are taken to find the placement of TCSC.
Step 14: Stop the process.

4 Simulation Result

4.1 Case Study of IEEE 5-bus System

One line diagram of the above system has shown in Fig. 3. The system consists of 5 buses, 7 branches, and generators connected to buses 1, 2, and 3. The range of TCSC is taken as −0.8 to +0.2 % from line reactance, and the power flow has carried out before and after placing the TCSC to determine their benefits.

4.1.1 Line Flow for 5-bus System

The load flow analysis by Newton–Raphson method has been carried out using MATLAB, and the results are tabulated as shown in Table 1.

4.1.2 Optimal Location of TCSC Found by the PSO Method

Proposed PSO methodology has applied to the IEEE 5-bus system. In this paper, reactance of TCSC has considered as variable parameter. From the simulation results, we can infer that TCSC has been optimally located in one of the seven branches where minimum loss occurs rather than locating all the branches, which in turn reduces the cost. The location of TCSC and the corresponding reactance with PSO method have tabulated in Table 2.

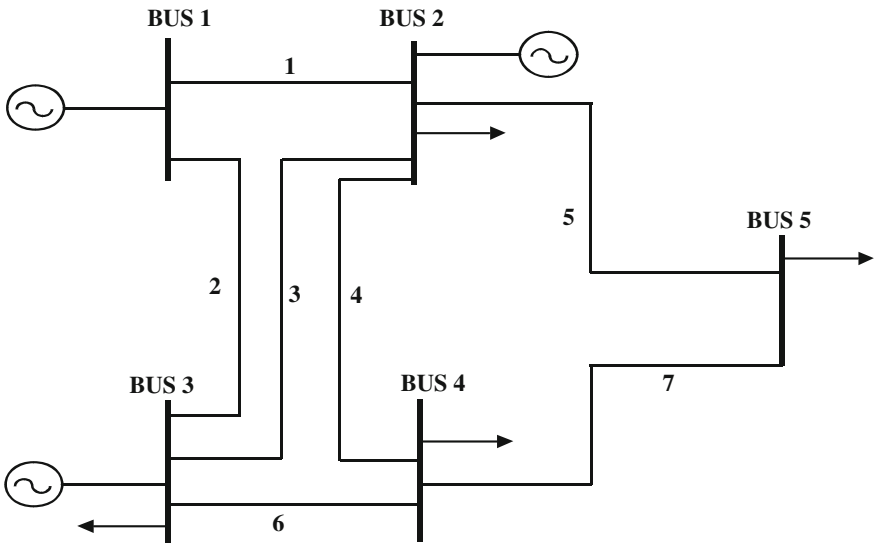


Fig. 3 Single *line diagram* of 5-bus system

Table 1 Line flow result for IEEE 5-bus system

Branch	Line	X_{old}	Line losses		Power in MW	Power in MVAr
			MW	MVAr		
1	1–2	0.06	0.648	−4.702	59.891	4.050
2	1–3	0.24	0.408	−4.239	23.153	3.217
3	2–3	0.18	0.080	−4.066	10.918	2.961
4	**2–4**	**0.18**	**0.231**	**−3.565**	**18.221**	**7.254**
5	2–5	0.12	1.298	0.784	50.136	30.437
6	3–4	0.03	0.237	−1.389	43.573	23.660
7	4–5	0.24	0.154	−4.582	11.333	5.853

Table 2 Result for TCSC allocation with PSO

Line	X_{old}	X_{TCSC}	$X_{eff} = X_{old} - X_{TCSC}$	Line losses		Power in MW	Power in MVAr
				MW	MVAr		
1–2	0.06	0.000	0.06	0.749	−40,400	64.583	2.662
1–3	0.24	0.000	0.24	0.288	−4.598	18.732	4.489
2–3	0.18	0.000	0.18	0.038	−4.191	3.656	5.307
2–4	**0.18**	**-0.1337**	**0.04623**	**0.718**	**−3.714**	**34.531**	**8.491**
2–5	0.12	0.000	0.12	1.154	0.352	45.725	30.932
3–4	0.03	0.000	0.03	0.142	−1.675	31.970	21.013
4–5	0.24	0.000	0.24	0.234	−4.355	15.763	4.931

It has noticed that the insertion of the TCSC found by PSO method into the system has resulted in:

Total P_{loss} reduces, real power flow increases, reactive flow changes, and optimal location of TCSC is in 4th branch.
Graph of line loss verses branch shown in Fig. 4 and graph of power flow verses branch shown in Fig. 5.

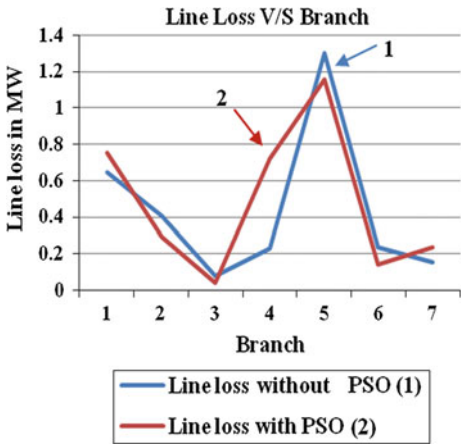


Fig. 4 Line loss verses branch

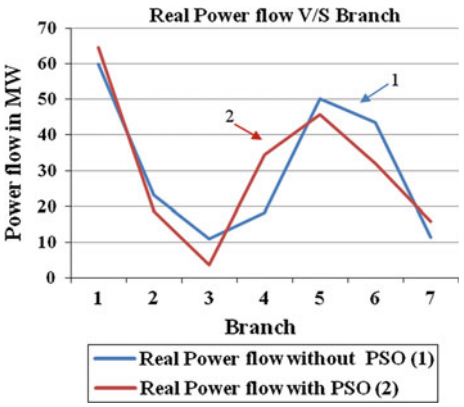


Fig. 5 Power flow verses branch

4.2 Case Study of IEEE 14-bus System

The data of the IEEE 14-bus system are given in Fig. 6. The system consists of 14 buses, 20 branches, and generators connected to buses 1 and 2. The reactive power sources connected to buses 3, 6, and 8.

Data: $S_{base} = 100$ MVA, $V_{max} = 1.06$ p.u, $V_{min} = 0.94$ p.u, P_{max} gen. at bus $1 = 250$ MW, P_{max} gen .at bus $2 = 50$ MW, Q_{max} gen. at bus $1 = 10$ MVAr, Q_{max} gen. at bus $2 = 50$ MVAr, Q_{max} gen. at bus $3 = 40$ MVAr, Q_{max} gen. at bus $6 = 24$ MVAr, Q_{max} gen. at bus $8 = 24$ MVAr, Q_{min} gen. at bus $1 = 0$ MVAr, Q_{min} gen. at bus $2 = -40$ MVAr, Q_{min} gen. at bus $3 = 0$ MVAr, Q_{min} gen. at bus $6 = -6$ MVAr, Q_{min} gen. at bus $8 = -6$ MVAr.

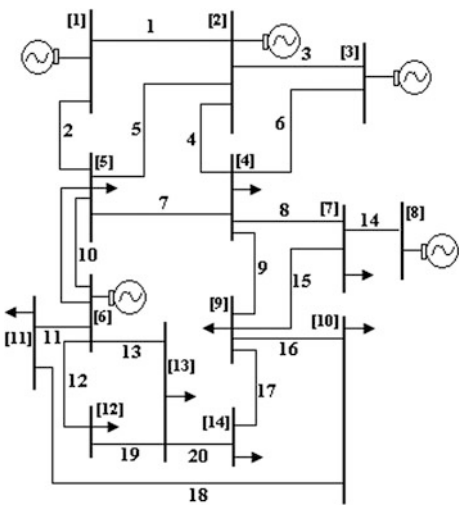


Fig. 6 Single *line diagram* for 14-bus system

Table 3 Line flow result for IEEE 14-bus system

Branch	Line	X_{old}	Line losses		Power in MW	Power in MVAr
			MW	MVAr		
1	1–2	0.05917	5.017	3.563	157.556	−66.776
2	1–5	0.22304	2.726	0.569	74.441	−10.116
3	2–3	0.19797	2.370	0.463	75.052	0.971
4	2–4	0.17632	1.798	−1.824	57.188	10.520
5	2–5	0.17388	0.815	−5.100	39.192	−0.044
6	3–4	0.17103	0.526	−1.304	−21.553	17.636
7	4–5	0.04211	0.472	1.490	−60.014	12.656
8	4–7	0.20192	0	1.632	28.596	5.257
9	4–9	0.55618	0	2.215	16.935	13.402
10	5–6	0.25202	0	4.336	42.720	3.880
11	6–11	0.1989	0.050	0.104	6.543	3.554
12	6–12	0.25581	0.076	0.158	7.721	2.469
13	6–13	0.13027	0.220	0.433	17.339	7.215
14	7–8	0.17615	0	0.186	0.006	−10.613
15	7–9	0.11001	0	1.058	28.592	14.373
16	9–10	0.08450	0.017	0.045	6.095	4.186
17	**9–14**	**0.27038**	**0.137**	**0.291**	**9.932**	**3.567**
18	10–11	0.19207	0.009	0.022	−2.968	−1.636
19	12–13	0.19988	0.006	0.006	1.498	0.830
20	13–14	0.34802	0.049	0.099	5.090	1.751

4.2.1 Line Flow for 14-bus System

The line losses and power generation as per Newton–Raphson method have given in Table 3.

4.2.2 Optimal Location of TCSC Found by the PSO Method

Proposed PSO methodology has been applied to IEEE 14-bus systems. The location of TCSC and the corresponding reactance with PSO method have been tabulated in Table 4.

PSO method into the system has resulted in:

Total P_{loss} reduces, real power flow increases, change in reactive flow, and optimal location of TCSC is in 17th branch.

Graph of line loss verses branch shown in Fig. 7 and graph of power flow verses branch shown in Fig. 8.

Table 4 Result for optimal location of TCSC found by PSO

Line	X_{old}	X_{TCSC} added	$X_{eff} = X_{old} - X_{TCSC}$	Line losses		Power in MW	Power in MVAr
				MW	MVAr		
1–2	0.0592	0	0.0592	5.015	3.558	157.530	−66.770
1–5	0.2230	0	0.2230	2.727	0.558	74.411	−10.622
2–3	0.1980	0	0.1980	2.369	0.459	75.040	0.962
2–4	0.1763	0	0.1763	1.798	−1.827	57.302	10.097
2–5	0.1739	0	0.1739	0.807	−5.132	39.772	−0.674
3–4	0.1710	0	0.1710	0.515	−1.335	−21.508	17.214
4–5	0.0421	0	0.0421	0.487	1.536	−61.151	12.010
4–7	0.2019	0	0.2019	0	1.721	29.427	5.184
4–9	0.5562	0	0.5562	0	2.272	17.434	13.231
5–6	0.2520	0	0.2520	0	4.044	41.418	2.148
6–11	0.1989	0	0.1989	0.062	0.130	7.597	3.535
6–12	0.2558	0	0.2558	0.069	0.143	7.227	2.733
6–13	0.1303	0	0.1303	0.183	0.361	15.398	7.729
7–8	0.1762	0	0.1762	0.0	0.167	−0.001	−10.074
7–9	0.7100	0	0.7100	0.0	1.082	29.420	13.548
9–10	0.0845	0	0.0845	0.013	0.022	4.989	4.254
9–14	**0.2704**	**−0.2163**	**0.0540**	**0.195**	**0.083**	**12.345**	**−2.568**
10–11	0.1921	0	0.1921	0.015	0.035	−4.019	−1.564
12–13	0.2209	0	0.2209	0.004	0.004	1.048	1.002
13–14	0.3480	0	0.3480	0.024	0.048	2.765	2.560

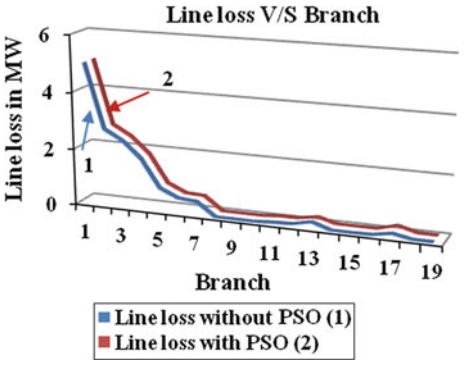


Fig. 7 Loss verses branch

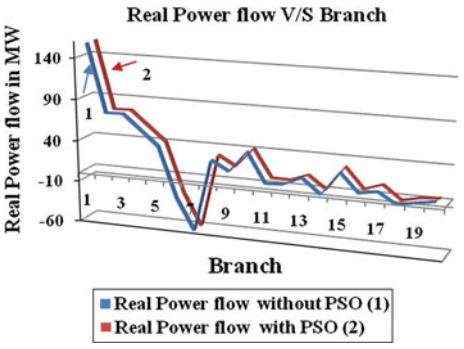


Fig. 8 Power flow verses branch

4.3 Summary of the Results

	5-bus system		14-bus system	
Parameter	Without PSO	With PSO	Without PSO	With PSO
Real power flow in MW	18.221 (Line 2–4)	34.531	9.932 (Line 9–14)	12.345
Total loss	3.055 MW	3.0 MW	14.288 MW	14.283 MW
Degree of compensations		5 %		5 %

5 Conclusion

It can be observed from the results that using PSO method, TCSC has been optimally placed in a weak line of the system and resulted in loss reduction in the lines. The total loss effectively reduces and resulted in a loss reduction, and it increases the power transfer capability of the line. Hence, this method can be extended to any practical systems with more number of buses. Further, the same method can also be effectively applied to shunt FACTS devices to enhance the voltage stability.

References

1. Ramasubramanianl P, Uma Prasana G, Sumathi K (2012) Optimal location of FACTS devices by evolutionary programming based OPF in deregulated power systems. Br J Math Comput Sci 2(1):21–30
2. Rajalakshmi L et al (2011) Congestion management in deregulated power system by locating series FACTS devices. IJCA 13(8)
3. Kodsi SKM, Cañizares CA (2003) Modelling and simulation of IEEE 14 bus system with FACTS controllers. IEEE benchmark technical report
4. Venayagamoorthy GK, Harleyl RG (2007) Swarm intelligence for transmission system control. 1-4244-1298- 6/07/IEEE 2007
5. Chandrasekarn K et al (2005–2009) A new method to incorporate FACTS devices in optimal power flow using particle swarm optimization. J Theor Appl Inf Technol 67–74
6. Santiago-Luna M, Cedeno-Maldonado JR (2006) Optimal placement of facts controllers in power systems via evolution strategies. PES transmission and distribution conference and exposition Latin America, Venezuela, IEEE
7. Saravanan M et al (2005) Application of PSO technique for optimal location of facts devices considering system loadability and cost of installation
8. Samimi A, Naderi P (2012) A new method for optimal placement of TCSC based on sensitivity analysis for congestion management. Smart Grid Renew Energy 3:10–16

Effective Integration of Wind-Distributed Generation to Power Grid with STATCOM

Surekha Manoj and P. S. Puttaswamy

Abstract Worldwide fast depletion of conventional energy resources necessitates the implementation of renewable energy sources for generation to satisfy the growing demand. Since last decade, technological innovations and a changing economic and regulatory environment have resulted considerable revival of interest in connecting wind generation to the grid. Utilities are seeking to understand possible impacts on system operations when a large amount of wind power is introduced into the electric power system. Producers of renewable energy must condition the power produced in order to interconnect with the power grid and not interface with the grid's overall performance. In these aspects, flexible AC transmission systems (FACTS) technology plays a vital role in enhancing the power system performance and improving the power quality of the system. This paper concentrates on power quality issues when wind power integrates with grid and the solution with the usage of STATCOM. An attempt is made with IEEE 16 Bus, 3 feeder test system, and modeled for simulation study using MATLAB/SIMULINK simulation. Scopes obtained from the simulation results are proven for the improvement of voltage profile which in turn improves the overall power quality issues.

Keywords FACTS · Wind energy · Power quality · Grid integration

S. Manoj (✉)
Department of Electrical and Electronics, Vidya Vikas Institute of Engineering and Technology, Mysore, India
e-mail: surekhamanoj12@yahoo.co.in; surekhamanoj.vviet@gmail.com

P. S. Puttaswamy
Department of Electrical and Electronics, P.E.S. College of Engineering, Mandya, India
e-mail: psputtaswamy_ee@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_88,

1 Introduction

India's rapidly growing economy and population leads to relentlessly increasing electricity demand. As a result, the country's installed power generation capacity has increased from just 1.4 GW in 1947 to over 170 GW in 2010. According to the ministry of new and renewable energy(MNRE), today, the share of renewable based capacity is 10.9 % (excluding large hydro) of the total installed capacity of 170 GW in the country. This includes 13,065.78 MW of wind, 2939 MW of small hydro power, 1,562 MW of cogeneration, 997 MW of biomass, 73.46 MW of waste to power, and 17.8 MW of solar PV for grid-connected renewable at the end of 2010 [1].

Present scenario is not only to satisfy demand but also to maintain higher power quality along with the rising concern about problems related to our environment, such as global warming. These are new challenges to the power grid, and it is expected to perform better and be "Greener". The drawbacks of centralized generation such as, long gestation period, high transmission and distribution losses, poor efficiency, high carbon footprint and peak demand management through load shedding caused the rapid development of distributed generation (DG) technology and is gradually reshaping the conventional power systems in number of countries [2].

DG applications in the vicinity of the load had shown greater operational and power quality advantages, in addition to transmission loss reduction. DGs are very appropriate for particular site and specific applications as they require short period of construction and need low investment. It is defined on the basis of size of the plant, which may vary from few KW to MW (10–50 MW). DG options can be classified based on the fuel source as renewable or nonrenewable [3]. DG has brought greater attention from the power community, when it is associated with renewable energy sources, as a sustainable alternative energy.

In addition to the environmental benefits, renewable energy conversion systems penetration increases the utility's reserve capacity by adding converted power into electricity. They can provide power to remote areas, relieving the generating and distribution utilities from expanding their resources. They can also be integrated with existing transmission or distribution networks [4]. Grid-connected wind capacity is undergoing through fastest rate of growth compared to any form of renewable electricity generation, achieving global annual growth rates of 20–30 % [5]. Figure 1 depicts the global wind power installed capacity.

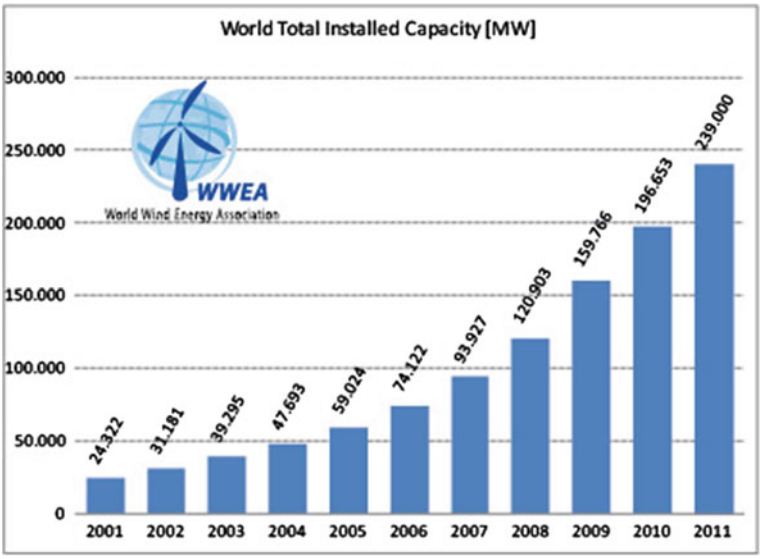


Fig. 1 Global wind power installed capacity

2 Grid Integration with Wind Energy

2.1 Challenges: Integrating Renewable Technologies with the Traditional Power Grid May Face Many Challenges

- When DG is fed into the power grid, it produces changes on the power flow pattern and may bring different challenges including voltage quality problems, limiting the maximum penetration capacity of DG units on the network and their location [2, 6, 7]. The installations must meet requirements of the grid connection codes such as fault ride through capability, reactive power range, voltage control, and frequency range and control [8, 9].
- Majority of the wind power–based DG technologies employ induction generators instead of synchronous generators, for the technical advantages of induction machines like, reduced size, increased robustness, lower cost, increased electromechanical damping, and its capability to be synchronized directly to the grid [5]. The main disadvantage is that the grid has to supply not only the load and line reactive power but also the generator. Its reactive power consumption depends on active power production. It is therefore very important to meet the variable VAR requirements within the wind farm locality and to relieve the hosting utility from supplying extra VARs [4].
- Following the fault conditions, the induction generator draws more reactive power from the grid, and the voltage recovery may become impossible, and

consequently, the wind farm may experiences voltage collapse at its terminals [5, 9, 10].

- During a grid fault, the wind turbines are rapidly disconnected from the power network and reconnected when normal operation has been resumed. This is possible, as long as wind power penetration remains low. However, the penetration of wind power is increasing rapidly, and also, demand is starting to influence overall power system behavior [10, 11].
- Wind being a geographically and climatically uncontrollable resource, the nature of distributed wind induction generators, the stability, and power quality issues of integrating large wind farm in grid may become pronounced [12].

2.2 Solution

To minimize reactive power exchange between wind power plant and distribution network, dynamic compensation of reactive power can be employed which would help in preventing the voltage collapse at the terminals of wind farms and lead to improving the stability of the wind farm. There are several ways offered to fulfill the requirements of reactive power compensation and to overcome the drawbacks mentioned above.

2.2.1 Cables

The cable itself represents a VAR source that should be used to supply reactive power into the grid. With slightly under compensated cables, it is possible to provide a considerable contribution to VAR generation [9]. However, to increase the transmission capacity and to reduce the losses, long AC cables always need compensation by shunt reactors on both sides. Besides, the switching capability of circuit breakers is also limited.

2.2.2 Transformers Equipped with on-Load Tap Changers (TC)

By using tap changers, the voltage on the wind farm side can be controlled in a definite range. It is also possible to reduce the reactive power demand by operating the system at higher voltage levels [13]. But TC cannot solve the power quality issues, such as power fluctuations, voltage fluctuations, and harmonics, satisfactorily because these devices are not fast enough [14].

2.2.3 Mechanical Switched Capacitor Banks (MSC)

Switched shunt capacitor banks may also be an option to generate capacitive power, but it presupposes a proper medium or low voltage level where the

capacitors can be connected. Although mechanically switched capacitors can play a significant part in an overall VAR compensation system, it does not have the response or the repeatability of operation that are generally needed for the dynamic compensation of power system. Precise and consistent control of the MSC closure is not possible.

Moreover, the frequent switching of MSC and TC causes resonance and transient overvoltage, which add additional stress on wind turbine gearbox and shaft, make themselves and turbines wear out quickly and hence increase the maintenance and replacement cost.

2.2.4 Flexible AC Transmission Systems (FACTS)

It is a power electronics revolution, which not only offered the advantage of high speed and reliability of switching but enhanced the value of electric energy. FACTS technology opens up new opportunities for controlling power and enhancing the usable capacity of present as well as new and upgraded lines and overtook MSC and TC [14]. The ability of FACTS controllers to control the interrelated parameters that govern the operation of transmission systems including series impedance, shunt impedance, current, voltage, phase angle and the damping of oscillations opens a wide door for power system. When smooth reactive power control is needed, FACTS may be the right option.

Several studies have found that FACTS technology not only provides solutions for efficiently increasing transmission system capacity but also increases available transfer capability, relieve congestion, improve reliability and enhances operation and control [9]. Different FACTS can be used to generate/absorb the reactive power required. They are classified as, Shunt connected controllers, Series connected controllers, Combined shunt and series connected controllers. Static Synchronous Compensator (STATCOM) and Static VAR Compensator (SVC) are shunt controllers and widely used for voltage control, VAR compensation and for damping oscillation. Static Synchronous Series Compensator (SSSC) is a series controller and used for current control, damping oscillations fault current limiting. Unified Power Flow Controller (UPFC) is combination of shunt and series controllers, which is used for active and reactive power control, voltage control, VAR compensation etc. [14].

The installation of the DFACTS in the distribution network allows better and higher penetration of DGs [2]. Simulation results show that the FACTS controllers prevents large deviations of bus voltage magnitude induced by reactive power drawn from grid during an external three-phase fault and under wind speed changes [5]. Usage of fixed type VAR across each wind turbine generator would not be practical as the VAR level changes with the power delivered and hence suggested a FACT system to regulate [4]. The placement of both STATCOM and SSSC at suitable locations proved that dynamic voltage control of the wind farm to ride through the grid disturbances is achieved [10]. The effectiveness of STATCOM in facilitating the integration of a large windmill into a weak power system

has been studied and a STATCOM is proposed for dynamic voltage control and steady state of the system [12].

The best improvement can also be achieved by connecting static shunt capacitors to the system during fault conditions [15]. SVC and TCSC applied to damp sub synchronous resonance caused by the induction generator connected to the grid through series compensated line [16]. The inclusion of SVC in fixed speed wind turbine improves voltage stability and prevents the disconnection [17]. Several researches show that SVC and STATCOM can improve power system performance when wind farm connects the grid [18]. But STATCOM implementation is better than SVC for overall improvements of power system.

The capabilities of FACTS made ENERCON power electronics technology to implement FACTS controller in the ENERCON wind energy converters itself [19]. ABB has experience of installing over 600 FACTS application in over 50 countries and seen an improvement in capability and flexibility of transmission corridors at least by 10–15 % [6]. Many researchers have proposed about FACTS implementation in windmills integrated with grid in order to overcome the associated problems discussed above.

3 Case Study and Simulation Results

A modified IEEE 16 node and three feeder test system is selected for the case study as shown in Fig. 2.

Here grid is supplying feeder 1, 2, and 3. One of the feeders is replaced by wind DG whose total generation is 9 MW to see the effect of renewable integration with the grid. This test system is modeled using MATLAB/SIMULINK as shown in Fig. 3 and analyzed the system with wind integration at feeder three point from simulation studies.

For this position, simulation is carried out with STATCOM and without STATCOM. This is carried out by changing the position of switches in the simulation model. The obtained results from the scopes are shown in Figs. 4 and 5.

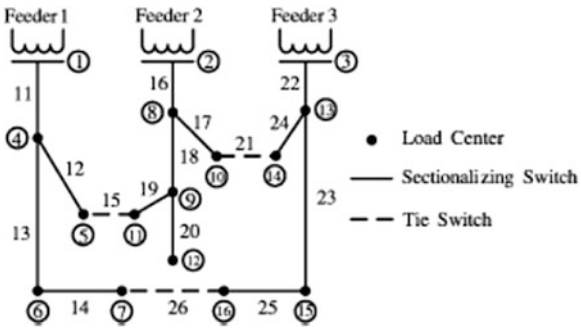


Fig. 2 IEEE 16 node three feeder distribution system

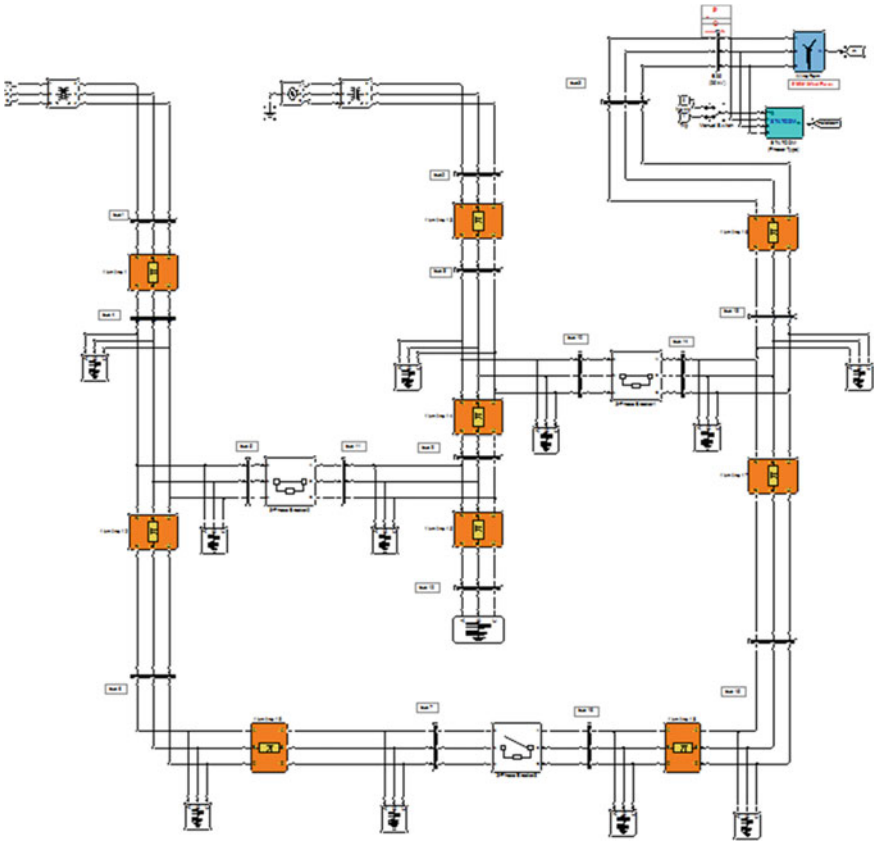


Fig. 3 Simulink model of wind farm integrated with grid

It is very well seen from the scope that a considerable improvement in voltage profile and considerable improvement in power output when STATCOM is connected at the point of common coupling than when there was no STATCOM connected at the point of common coupling.

Further, a double line to ground fault is created at one of the wind turbine generator as disturbance in the system to observe how the total system behaves with STATCOM and without STATCOM. The results obtained are shown in Figs. 6 and 7. All the performance parameters like voltage, real power, and reactive power shows considerable improvement with STATCOM when compared with no STATCOM at the point of common coupling.

The response of real power P, reactive power Q, and voltage behavior in scope is observed for 20 s and fault is created at 15 s of the study considered. From the scope, it is very clear that after the fault active power goes to zero indicating that Wind Energy DG has been isolated from the system and voltage improvement with STATCOM is observed.

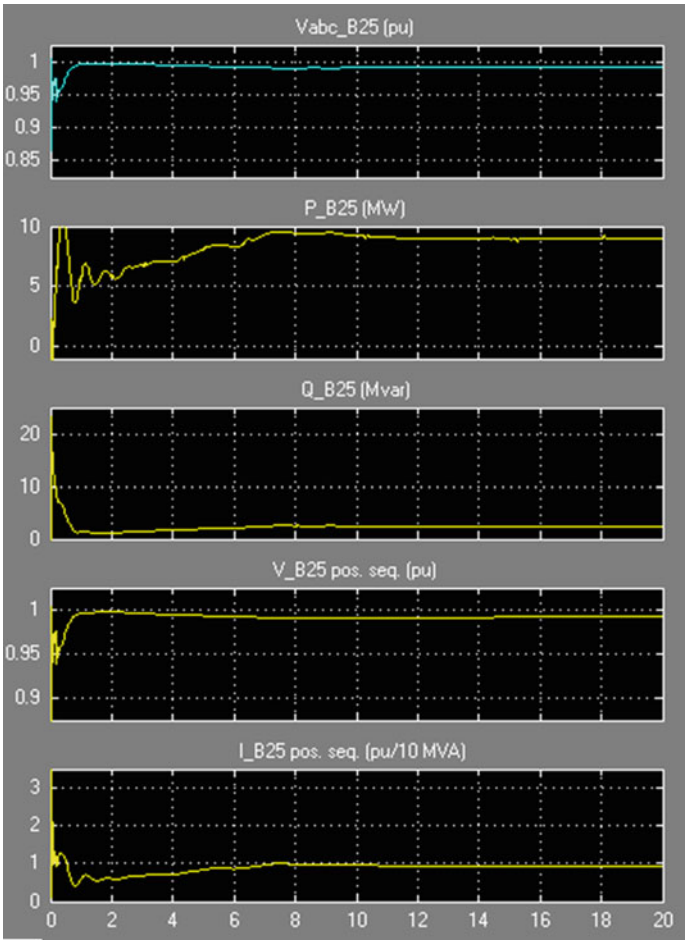


Fig. 4 Steady state analysis with STATCOM

The pitch of the blade is varied in the given range in order to maintain the power constant for the varying wind with its cut-in and cut-out speed ranges to handle different loads, here, the maximum pitch angle is given as 45 degrees, and maximum rate of change of pitch is given as 2 deg/sec. Maximum power of 1 pu can be generated at a wind speed of 9 m/s, and it can be observed from Fig. 8.

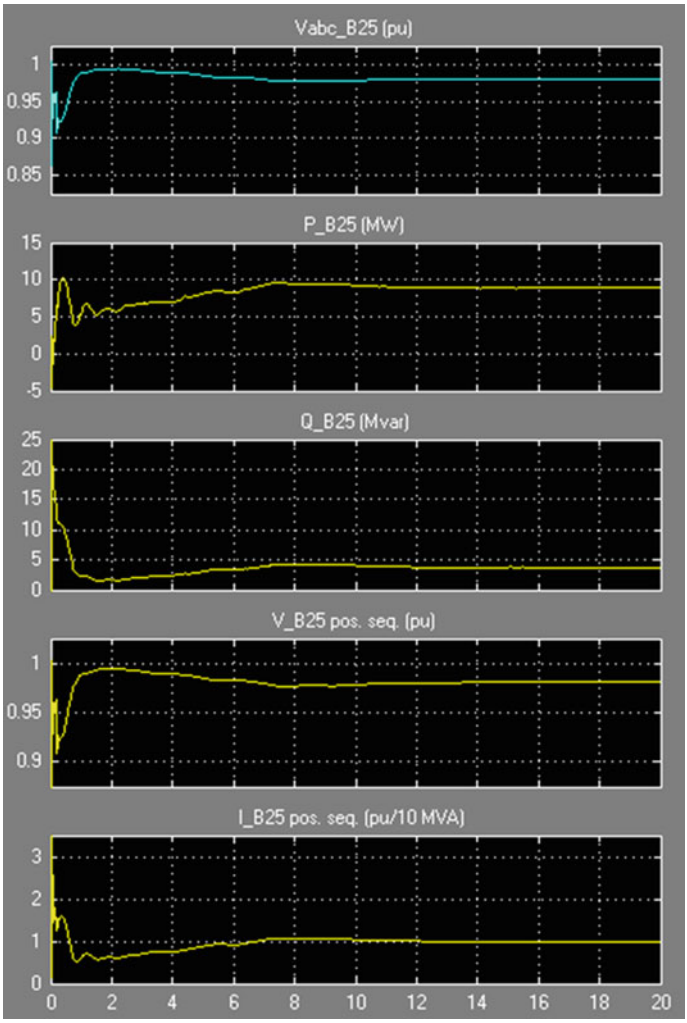


Fig. 5 Steady state analysis without STATCOM

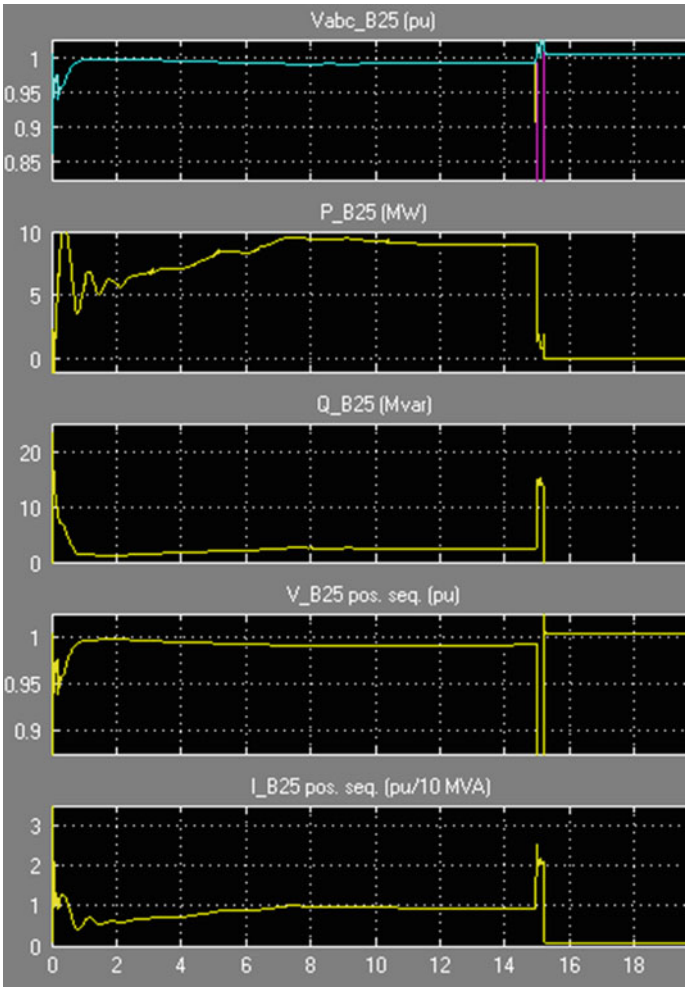


Fig. 6 Transient state analysis with STATCOM

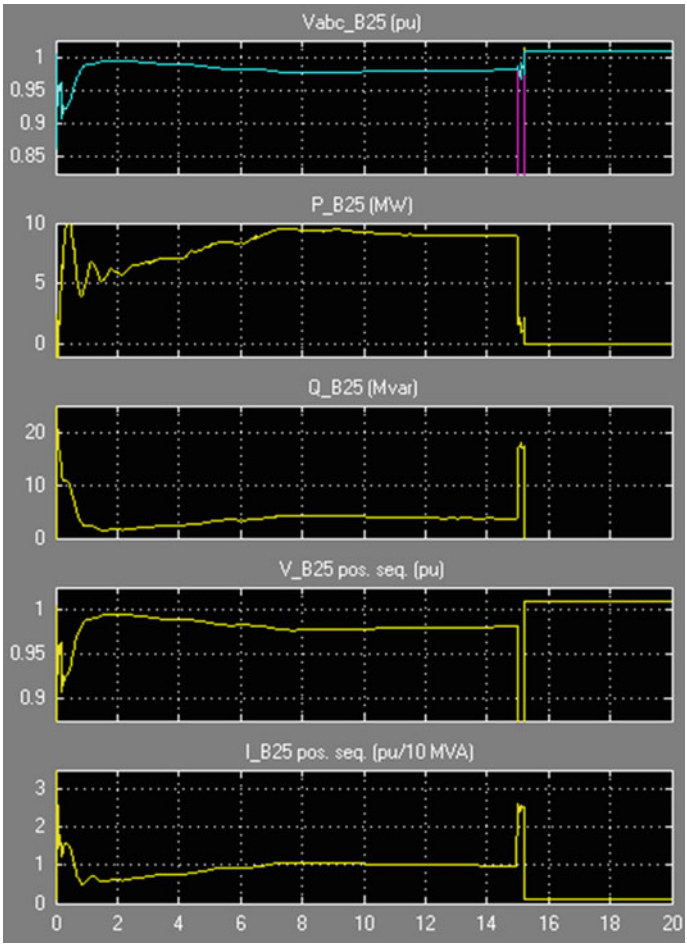


Fig. 7 Transient state analysis without STATCOM

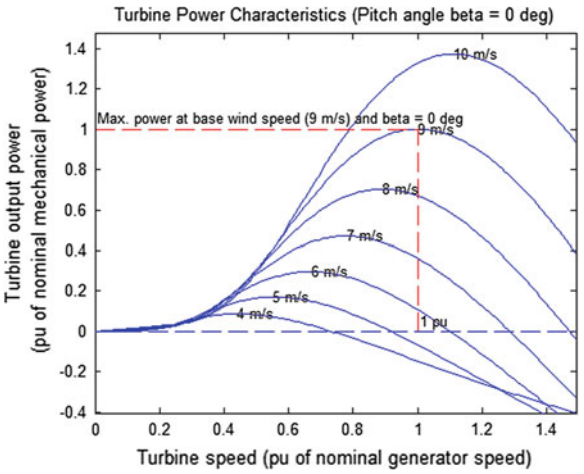


Fig. 8 Turbine power characteristics

4 Conclusion

The attempt made to verify the performance of power system with STATCOM and without STATCOM found satisfactory as it is seen very clearly from the scopes, the improvement in the performance curves at the point of common coupling. Simulation studies have been carried out to verify the system performance under different scenarios such as with fault, without fault, with STATCOM and without STATCOM for the IEEE 16 bus test system considered, and obtained results were satisfactory. FACTS controller implementation plays a vital role in the improvement of overall performance of the power system. As the wind integration causes the system voltage deterioration, at the point of coupling, a controller helps the system to uplift to maintain grid codes. With this test study, further work is planned to apply the same on real system case study. This work made the confident report to apply on real system.

References

1. Indian wind energy outlook (2011) Global wind energy council
2. Ren H, Yu X, Watts D (2009) Application of DFACTS for the improvement of penetration capacity of distributed generation. DFACTS 2009 (web.ing.puc.cl/∼power/paperspdf/Dfacts2009.pdf)
3. Banerjee R (2006) Comparison of options for distributed generation in India. Energ Policy 34(1):101–111
4. Zobaa AF, Nigim KA, Bansal RC (2006) Technology of VAr compensation for induction generator applications in wind energy conversion systems. J Electrical Syst 2(3):172–184
5. Panda S, Padhy NP (2007) Power electronics based FACTS controller for stability improvement of a wind energy embedded distribution system. Int J Electron, Circuits Syst 1(1):30–37

6. Jones P (2007) Harder and smarter. IET Power Eng, p 34–37 (www.theiet.org/power)
7. Douglas J, Orme M (2006) Time to connect. IET Power Eng, p 31–33 (www.theiet.org/power)
8. Cartwright P, Anaya-Lara O, Wu X, Xu L, Jenkins N (2004) Grid compliant offshore wind power connections provided by FACTS and HVDC solutions. Proceedings EWEA conference, pp 1–8
9. Wilch M, Pappala VS, Singh SN, Erlich I (20047) Reactive power generation by DFIG based wind farms with AC grid connection. Proc Power Tech Conf, p 626–632
10. Qiao W, Harley RG, Venayagamoorthy GK (2006) Effects of FACTS devices on a power system which includes a large wind farm. Proc Power Syst Conf Exposition, p 2070–2076
11. Wessels C, Fuchs FW (2009) High voltage ride through with FACTS for DFIG based wind turbines. EPE
12. Han C, Huang AQ, Litzenberger W, Anderson L, Edris AA (2006) STATCOM impact study on the integration of a large wind farm into a weak loop power system. Proc Power Syst Conf Exposition, pp 1266–1272
13. Keane A, Ochoa LF, Vital E, Dent CJ, Gareth P (2010) Harrison- enhanced utilization of voltage control resources with distributed generation. IEEE transaction on power systems, paper accepted for inclusion in future issue of the journal
14. Hingorani NG, Gyugyi L (2000) Understanding FACTS: concepts and technology of flexible AC transmission system. IEEE Press
15. Salman SK, Teo ALJ (2005) Dynamic behaviour of integrated multiple wind farms during fault conditions on the hosted distribution network. Power Tech, IEEE Russia, pp 1–4
16. Varma RK, Auddy S, Semsedini Y (2008) Mitigation of subsynchronous resonance in a series-compensated wind farm using FACTS controllers. IEEE Trans Power Delivery 23(3):164
17. Salehi V, Afsharnia S, Kahrobaee S (2006) Improvement of voltage stability in wind farm connection to distribution network using FACTS devices. IEEE industrial electronics, IECON 2006. 32nd annual conference publication year, IEEE Conferences, pp 4242–4247
18. Wachtel S, Hartge S (2007) Technical and economical benefits of wind energy converters with FACTS capabilities for power systems and the grid integration of wind power-EWEC, Milan, ENERCON
19. Young GO (1964) Synthetic structure of industrial plastics. In: Peters J (ed) Plastics, vol 3, 2nd edn. McGraw-Hill, New York, pp 15–64
20. Chen WK (1993) Linear networks and systems. Belmont, Wadsworth, pp 123–135
21. Poor H (1985) An introduction to signal detection and estimation. Springer, New York, Ch 4
22. Smith B (1982) An approach to graphs of linear forms, unpublished
23. Miller EH A note on reflector arrays (Periodical style—Accepted for publication), IEEE Trans Antennas Propag
24. Wang J Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication), IEEE J Quantum Electron, submitted for publication

Novel Shannon-Based Low-Power Full-Adder Architecture for Neural Network Applications

G. A. Lalithamma and P. S. Puttaswamy

Abstract This paper proposes novel low-power full-adder cell to be used as Shannon adder in the neural network applications. By using the Shannon's theorem, the gate count is reduced, thereby the total chip area gets minimized. Hence, the power also gets reduced to a considerable amount. The designs are implemented using synthesis tool which results in significant reduction in area and power for the modified Shannon-based full-adder cell when compared with multiplexing control input technique (MCIT)-based full-adder cell.

Keywords Full adder · Datapath · Shannon's technique · Power · Gate count · Area · Propagation delay

1 Introduction

In state of the art, consumer electronics power optimization is an important challenge in very large scale integration (VLSI) circuits. The battery-operated portable devices continue to grow. The traditional approaches for designing these systems vary according to the need of low-power design. The power minimization is one of primary design constraint for current day VLSI systems [1]. Datapath/arithmetic unit forms the heart of most of the state of the art systems [2]. Hence, optimizing the datapath unit is one of the hot topics of research in the present day VLSI domain [3]. The artificial neurons or nodes form the neural networks. The artificial neural networks are formed by connecting artificial neurons (similar to

G. A. Lalithamma (✉)
Department of Electrical & Electronics Engineering, SJBIT, Bangalore, India
e-mail: lalithasjb@gmail.com

P. S. Puttaswamy
Department of Electrical & Electronics Engineering, PESCE, Mandya, India
e-mail: psputtaswamy_ee@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_89,

the properties of biological neurons). An artificial neural network involves a network of simple processing elements. Complex systems can be built by connecting the neurons in order to achieve the required functionality. The layered technique is used to connect neurons. The output of the first layer neurons forms the input to the next layer neurons. The primary inputs of the system will drive the first neuron layer. These neurons are used as buffer (to hold the input values). The next layer is a hidden layer (it is invisible). The neurons themselves consists of a set of inputs from the first (input) layer, an implementation of a mathematical function of those inputs, then an output that carries the result of the function to the next layer. This next layer can potentially be another layer whose neurons are now functions of the outputs of the previous layer. After the final hidden layer, a last layer, called the output layer, actually supplies the predicted output values to the user. The self-organizing neural networks are capable of designing themselves. The initialization of a network to perform a specific task is called training. During training, the hidden neurons are where the training actually occurs. Every input combination must be tested to see its effects on the output of the system. A comparison between the input and output results in a weight or coefficient that is associated with that input value.

2 Shannon Theorem

The Shannon Theorem can be a function of many variables, $f(\mathrm{b0}, \mathrm{b1}, \mathrm{b2}, \mathrm{y}, \mathrm{bi}, \mathrm{y}, \mathrm{bn})$ It is written as the sum of two terms, say one with a particular variable *ai*, set to 0, and one with it set to 1.

$$\begin{aligned} \mathrm{f(b0,\ b1,\ b2, \ldots,\ bi, \ldots y, bn)} \ = \ & \mathrm{bi'f\ (b0, b1, b2, \ldots, 0, \ldots y, bn)} \\ & + \ \mathrm{bi\ f\ (b0, b1, b2, \ldots, 1, \ldots y, bn) \ldots} \end{aligned} \tag{1}$$

3 Existing Architecture

The two existing architecture are

a. Multiplexing control input technique (MCIT).
b. Existing shannon-based full-adder cell.

3.1 Multiplexing Control Input Technique

Based up on the Karnaugh map, the MCIT technique is developed. The truth-table for full adder is derived from the Boolean expressions for the sum and carries

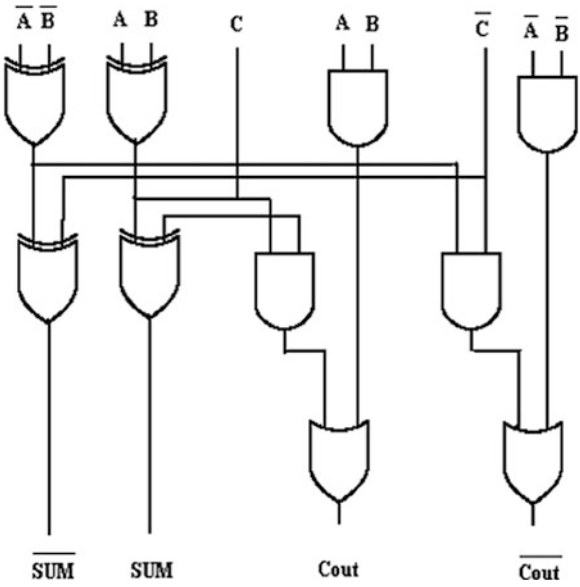


Fig. 1 MCIT for a full adder circuit using logic gates

signals. Figure 1 shows the carry and sum path. The Boolean expressions are as follows:

$$C = AB + BC + CA\ldots \tag{2}$$

$$S = ABC + A'B'C + AB'C' + A'BC'\ldots \tag{3}$$

3.2 Existing Shannon-Based Full-Adder Cell

The existing full adder was designed based on combining the MCIT technique for sum, and Shannon operation for carry [4] is as shown in Fig. 2.

4 Proposed Architecture

Novel low-power architecture for XOR gate is proposed. The XOR gate is built using combination of OR-AND gate and NAND gate. In the proposed architecture, since we are sharing the AND gate, area will be reduced. The inverters at input are eliminated which will result in reduced power and improved performance. By exploiting the available resources in FPGA, we can achieve better optimizations by implementing carry path logic in one 6-input look-up table (LUT) (since it requires six variables). Refer to Fig. 3.

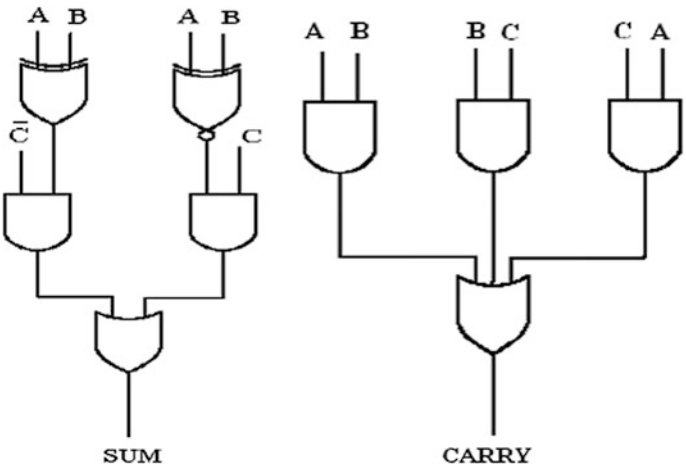


Fig. 2 Shannon-based full adder using logic gates

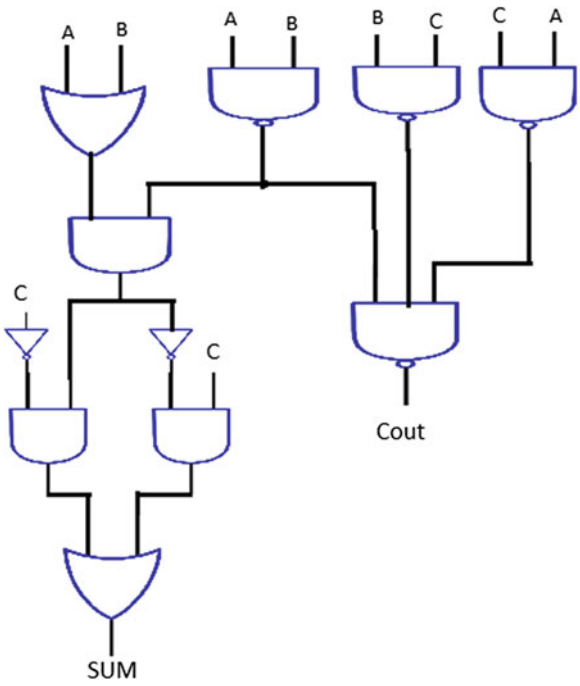


Fig. 3 Proposed Shannon-based full adder using logic gates

5 Implementation

The existing and proposed architectures are implemented using both AISC & FPGA methodologies.

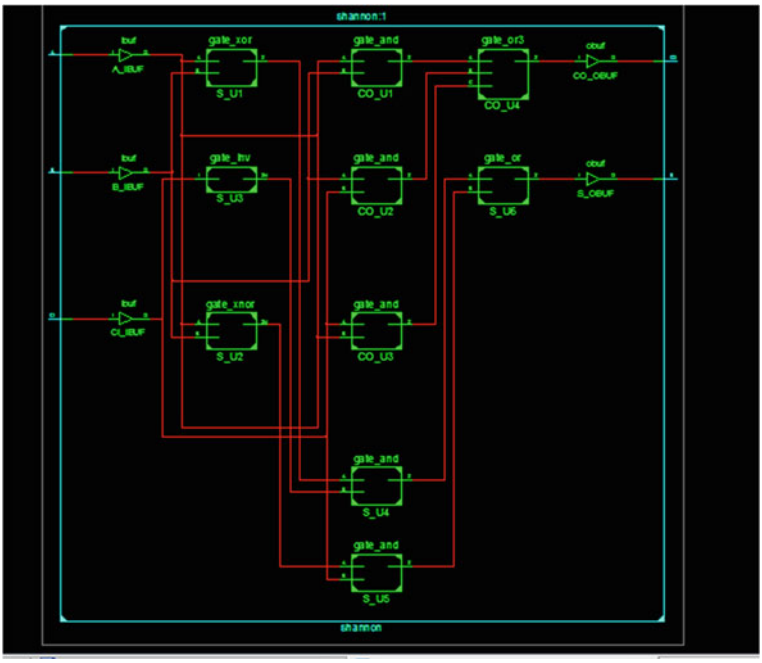


Fig. 4 Shannon adder basic

5.1 Fpga Results

Both the architectures are modeled using verilog, functionally verified using model-sim, and synthesized using ISE. The designs were mapped to Virtex2 FPGA as shown in Figs. 4 and 5.

Simulation results of sequential 4-bit adder using the proposed full-adder cell is as shown in Fig. 6.

6 Benchmarking Methodology

6.1 Field Programmable Gate Array

A generic top-down FPGA design methodology is shown in Fig. 7. It is used as baseline for the hardware realization of algorithms. The first step in FPGA design methodology is to capture the algorithm to be implemented on FPGA using hardware description languages (HDLs) or schematic depending on the complexity of the design. After capturing the design using HDLs or schematic, one must validate the logical correctness of the design against the required specification.

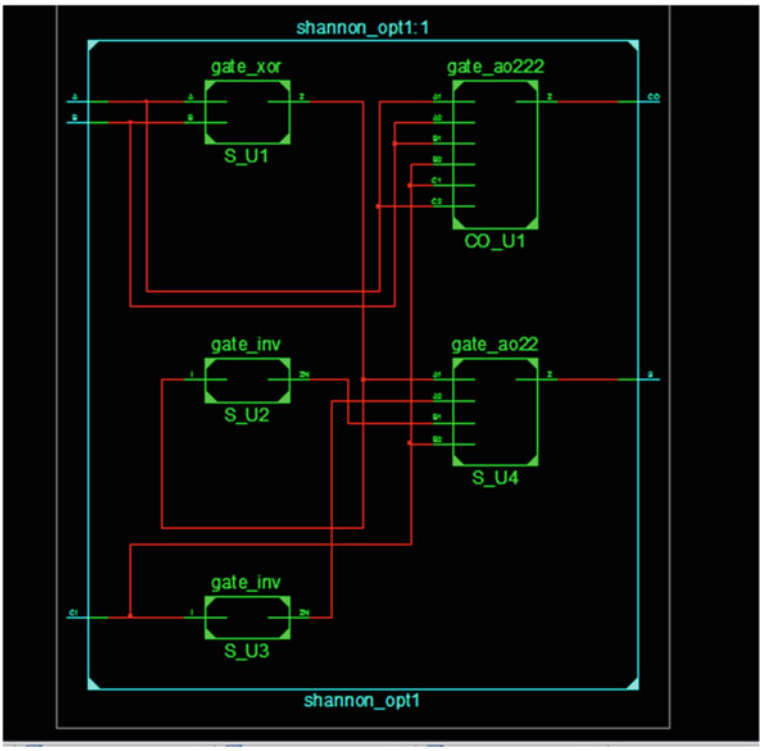


Fig. 5 Shannon adder propose

This is performed using functional or behavioral simulation. Designers execute this design phase soon after completing the coding phase and logic synthesis phase. Logic synthesis is the process of converting HDL or schematic-based design into a netlist of actual gates/blocks specified in FPGA devices. After logic synthesis, the gates are mapped to specified technology library. During this phase, the tool converts a netlist of technology independent logic gates into one consisting of logic cells and input/output blocks (IOBs) in the target FPGA architectures. The process of selecting the optimal position for each gate of the netlist is called placement.

The primary goal of an FPGA placement is to locate functional blocks such that the interconnect delay required to route the signals between them is minimal. Since interconnect delay is dominant in FPGA, a good placement is extremely important. It has the direct impact on the routability and the performance of a design on FPGA. A poor placement will lead to lower maximum operating speed and increased power consumption. Typically, FPGA placement algorithms are classified as routability driven and timing driven. The main objective of routability-driven algorithms is to create a placement that minimizes the total interconnect required. In addition to optimizing for routability, timing-driven algorithms use

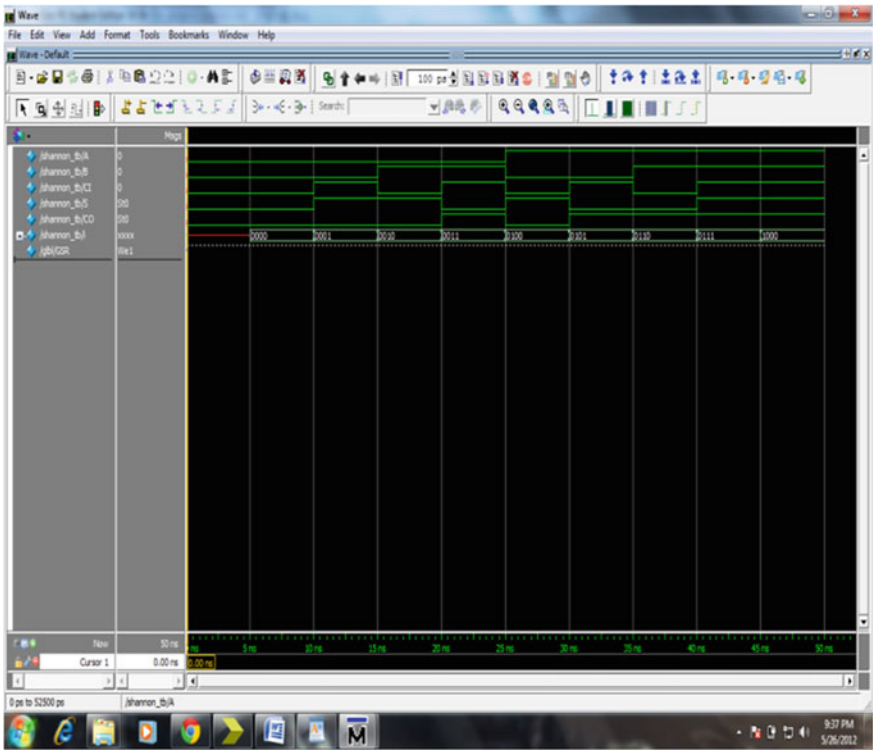


Fig. 6 Simulation results of sequential 4-bit adder using the proposed full-adder cell

timing analysis to identify critical paths and/or connections and optimize the delay of those connections.

The process of connecting the placed gates is called routing. It is the last phase in the design methodology prior to generating the bitstream to program the FPGA. It is a tedious process, since it must use only the prefabricated-routing resources such as wire segments, programmable switches, and multiplexers. Always, it is a challenging task to achieve 100 % routability in FPGAs. Timing simulation is performed after placement and routing to validate the logical correctness of the design taking into account the delays of the FPGA device. Power estimation of the design is done; the power analysis such as XPower and PowerPlay tools are used in XilinxISE and Altera Quartus II tools, respectively. The bitstream generation is the final step in the FPGA design methodology. The input to this phase is the mapped, placed, and routed design. It generates the necessary bitstream to program the logic and interconnects to implement the intended logic design and layout on the target device.

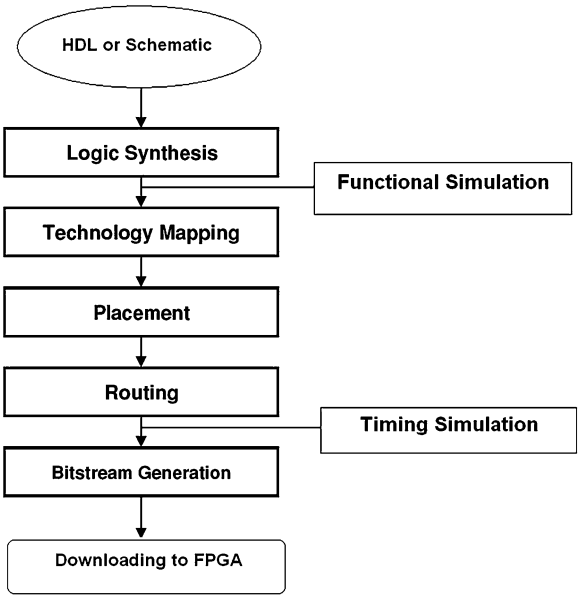


Fig. 7 FPGA design methodology

6.2 Application-Specific Integrated Circuit

Referring to the Fig. 8, the HDL design capture phase implements the "topdown" methodology of design. It takes a designer from abstract concept or algorithm down to hardware in manageable and verifiable steps. This involves developing a design specification that will be used to create a high-level behavioral abstraction with high-level programming languages such as C/C++.

Additionally, this model may also be created using the verilog hardware description language (HDL). The behavioral model should be simulated in order to verify that the desired functionality is captured completely and correctly. The behavioral abstraction is then used as a reference to create and refine a synthesizable register transfer-level (RTL) abstraction that captures the desired functionality required by the design specification.

The HDL design capture phase is completed with "pre-synthesis" simulations to verify that the RTL abstraction fully provides the desired functionality. The functional verification of the design that occurs at this point must be as complete and thorough as possible. This requires that the designer fully understands both the design specification and the RTL implementation. The test vectors employed during simulation should provide the coverage necessary to ensure the design will meet specifications. The HDL design synthesis phase involves using a synthesis tool to translate the abstract RTL design description to register elements and combinational logic. Optimize the combinational logic by minimizing, flattening and factoring the resulting boolean equations. Translate the optimized logic-level

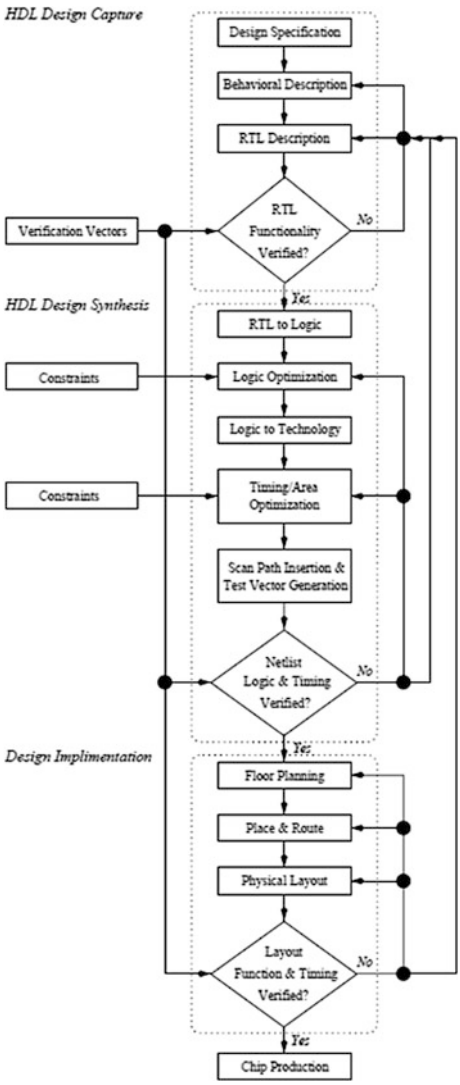


Fig. 8 ASIC design methodology

description to a gate-level description using cells from the specified technology library. Optimize the gate-level description using cell substitution to meet the specified area and timing constraints. Produce a gate-level netlist of the optimized circuit with accurate cell timing information. HDL design synthesis finishes with "post-synthesis" simulations to verify that the gate-level circuit fully provides the desired functionality and meets the appropriate timing requirements.

Table 1 FPGA benchmarking result

Module	Shannon adder basic	Shannon adder proposed	Percentage gain
Number of LUTs	10	6	40
Timing	8.137 ns	7.838 ns	3.7

Table 2 ASIC benchmarking results

Module	Existing architecture	Proposed architecture	Percentage gain
Full adder	$A = 11.6$	$A = 7.62$	$A = 52$
	$T = 314$ ns	$T = 223$ ns	$T = 40.8$
	$DP = 507.304$	$DP = 473.255$	$DP = 7.2$
	$CLP = 0.125$	$CLP = 0.098$	$CLP = 27.6$

Note
A = Area in square microns
T = Delay in ns
DP = Dynamic power in uW
CLP = Cell leakage power uW

7 Results & Discussion

Since area and power are directly proportional, area reduction results in power reduction (Leakage & Dynamic). Hence, we can conclude that the proposed architecture is power optimized.

As depicted in Table 1, the proposed architecture is well-suited for area-optimized applications. The performance of the proposed architecture is also better. Since we are eliminating a NAND gate, we are also reducing the number of interconnects which in turn reduces the adder delay. The Table 2 shows the benchmarking results of the existing and proposed architectures after implementing using ASIC design methodology. From Table 2, it is clear that the proposed architecture outperforms the existing architecture in all the design aspects (area, timing, and power). It is interesting to note that both dynamic and leakage power are reduced. Since it is an architectural innovation, below are the low-power advantages which are no area or performance penalty, minimum verification effort, since it is correct by design. It is pervasive, and it is independent of adder width.

8 Conclusion

The proposed modified Shannon-based full-adder cell has been simulated, and results are compared with MCIT-based full-adder cell and the existing Shannon-based full-adder cell in terms of power, transistor count, and area. This proposed adder cell is having improvement in all these aspects. The proposed low-power concept is proven in both FPGA & ASIC design methodologies. In the future

work, the proposed adder will be used to implement neural network. Since adder forms the basic datapath component in the neural network, the proposed low-power concept will result in the power optimization during both training of the network and normal functioning of the neural network. The plan is to develop a servo motor speed control logic based upon artificial neural network.

References

1. Najm F (1994) A survey of power estimation techniques in VLSI circuits. IEEE Trans VLSI Syst 2:446–455
2. Lee JD, Yoony YJ, Leez KH, Park B-G (2001) Application of dynamic pass-transistor logic to an 8-bit multiplier. J Kor Phys Soc 38(3):220–223
3. Khatibzadeh A, Raahemifar K (2005) A novel design of a 6 GHz 8×8 bit pipelined multiplier, IEEE Proceedings of the 9th international database engineering and application symposium (IDEAS'25'05), pp 1–5
4. Nehru K, Shanmugam A, Deepa S Priyadarshini R (2010) A shannon based low power adder cell for neural network training. IACSIT Int J Eng Technol, 2(3)

Novel Algorithm for Mobile Location Identification

Pranav Ravikumar, Jobin John, Megha Daga and Abdul Razak

Abstract The scope of this chapter is to develop an algorithm to attain current position of a mobile user whether stationary or moving. It will efficiently disclose the actual coordinates of a mobile phone bearer. A new hybrid technique has been proposed in this chapter, where the localization methods have been modified in order to improve accuracy in location estimation. This chapter can be significantly manifested in a series of mobile tracking services.

Keywords Base station (BS) · Angle of arrival (AOA) · Time of arrival (TOA) · Time difference of arrival (TDOA) · Taylor series algorithm (TSA) · Hybrid lines of position algorithm (HLOP)

1 Introduction

Mobile communication is a means of transmission of data packets and voice either via computer or a mobile device without any physical or fixed connection. Mobile node localization has been always a critical task in wireless communication network. With the advancement in technology and various applications, the localization has become more challenging. Many localization techniques have been proposed in the past few years. These localization techniques attempt to locate a

P. Ravikumar (✉) · J. John · M. Daga · A. Razak
Electrical and Electronics Engineering Department, BITS Pilani Dubai,
Dubai International Academic City, Dubai, UAE
e-mail: rkpranav.cooldude@gmail.com

J. John
e-mail: jobinjohn245@gmail.com

A. Razak
e-mail: abdulrazak@bits-dubai.ac.ae

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_90,

mobile node by estimating and computing various parameters of radio signals travelling between the mobile station and the base station.

2 Related Work

The paper on "Survey on wireless geolocation techniques" commences with the basic geolocation system and its components. Followed up is the fundamental principles employed in the traditional localization algorithm. All possible localization algorithms have been discussed briefly. Positioning methods have been isolated into network-based and MS-based localization techniques [1]. The method that estimates the mobile location using TDOA estimates acquired by using delay-locked loop method in CDMA cellular network was proposed in [2]. The performance was thereby tested in various radio propagation environments. The new approach to mobile localization known as lee localization algorithm (LLA) was presented in the chapter [3] to mitigate measurements errors in TOA technique with low complexity and high accuracy.

3 The Proposed Hybrid TOA/AOA Technique

We know that a AOA measurement can be very well represented by a line extended from MS to a BS and TOA measurement can be presented by a circle with BS as its centre. The mathematical equations of the three TOA circles and three AOA lines used for MS location estimation can be expressed as

$$\begin{aligned}
&\text{Circle}\quad 1: \quad (x - X1)^2 + (y - Y1)^2 = R_1^2 \qquad (X1,\ Y1) = (0, 0)\\
&\text{Circle}\quad 2: \quad (x - X2)^2 + (y - Y2)^2 = R_2^2 \qquad (X2,\ Y2) = (X2, 0)\\
&\text{Circle}\quad 3: \quad (x - X3)^2 + (y - Y3)^2 = R_3^2 \qquad (X3,\ Y3) = (X3, Y3)
\end{aligned}$$

$$\begin{aligned}
\text{Line1} &: x.\tan\Theta_1 - y = 0\\
\text{Line2} &: x.\tan\Theta_2 - y = \tan\Theta_2.\ X2\\
\text{Line3} &: x.\tan\Theta_3 - y = X3.\tan\Theta_3 - Y3
\end{aligned}$$

The proposed hybrid technique uses three TOA circles and three AOA lines to find out all the feasible intersections in order to locate a MS. Since the NLOS errors are always positive, the measured TOA will always be greater than the original value if measurements due to noise are neglected. The true location of MS should be within the region covered by the intersection of all these three circles. The intersections that lie within this region of overlap are termed as feasible points of intersections. Therefore, all the feasible points of intersections must satisfy given inequalities at the same time:

$$(x - X1)^2 + (y - Y1)^2 \leq R_1^2 \quad (X1, Y1) = (0, 0)$$
$$(x - X2)^2 + (y - Y2)^2 \leq R_2^2 \quad (X2, Y2) = (X2, 0)$$
$$(x - X3)^2 + (y - Y3)^2 \leq R_3^2 \quad (X3, Y3) = (X3, Y3)$$

However, all these feasible points are not useful to give information about MS location. Thereby in order to improve the accuracy in location estimation putting less efforts, various schemes of localization that were proposed in [2] have been modified.

3.1 Distance-weighted Method

The following steps govern the location estimation using this method:

Step 1: Any error in AOA measurement can be detected if the scatterer point S lies outside the region of intersections of three TOA circles. It can be seen that the true AOA measurement for jth and kth circles is given by ΘAjk.

$$\theta_{Ajk} \in \left\{\min\left(\theta_{Ajk}\right), \max\left(\theta_{Ajk}\right)\right\}, j, k = 1, 2 \ldots N (j \neq k)$$

Step 2: Determine all the feasible points of intersections of these three TOA circles and three AOA lines.

Step 3: $\overline{\mathrm{X}} = \overline{\mathrm{X}} = \frac{1}{N}\sum_{k=1}^{N} xk$, $\overline{\mathrm{Y}} = \frac{1}{N}\sum_{k=1}^{N} yk$

Step 4: Compute the distance *dK* between each feasible point (*xk,yk*) and the average values $\left(\overline{X}, \overline{Y}\right)$

$$\mathrm{d_k} = \sqrt{(xk - \overline{x})^2 + (yk - W)^2}, 1 \leq k \leq N$$

Step 5: Place the weights for *k*th feasible points of intersection to $(d_k^2)^{-1}$. Then, the Mobile station location is given by

$$X_{\mathrm{MS}} = \left[\sum \left(d_k^2\right)^{-1} \cdot xk\right] / \sum \left(d_k^2\right)^{-1},$$
$$Y_{\mathrm{MS}} = \left[\sum \left(d_k^2\right)^{-1} \cdot yk\right] / \sum \left(d_k^2\right)^{-1}$$

3.2 Sort Averaging Method

Step: 1–4 are same as that of the distance-weighted scheme.
Step 5: Now rank the computed distances d_k in the increasing order and also arrange the points of intersections in the same order.
Step 6: The MS location (x_L, y_L) is computed by the mean value of first L feasible intersections.

$$xL = \frac{1}{L}\sum_{k=1}^{L} xk$$

$$yL = \frac{1}{L}\sum_{k=1}^{L} yk\left(L = \frac{N}{2} \leq \mathrm{N}\right)$$

4 Results

Simulations were performed using all the methods of localization, and performance of the proposed methods was demonstrated. The coordinates for BS1 are (0, 0), BS2: (15, 0), and BS3: (7.5, 12.99). The location of MS was chosen randomly within the area of overlap of the TOA circles as shown in Fig. 6.1. Some different approaches were used to test various parameters affecting the performance of locating a MS.

Reading was recorded to estimate location of a moving person. Since the person is moving from one location to anther, it will form different TOA circles around the BS. The moment it comes stationary at some point its approximate location can be identified. Plot of RMSE versus Upper bound on NLOS error is plotted to check the performance for various localization schemes (Fig. 1).

Thus, the Upper bound on NLOS range error η is given by

max η_1 = min{AB, EF}. Similarly, Upper bounds η_2 and η_3 are given as
max η_2 = min{AB, CD} max η_3 = min{CD, EF}

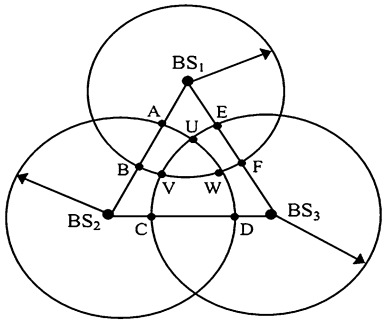


Fig. 1 Geometry of TOA-based location showing range *circles* and region of overlap

Table 1 Readings were taken for different TOA circles formed by mobile user in motion

BS1: R = 10 AOA = 30° BS2: R = 10 AOA = 155° BS3: R = 10 AOA = 84.9998°			
(7.5, 6.61)	(5.52, 3.18)		(9.48, 3.19)
Distance-weighted	X = 7.5029	Y = 4.3282	RMS = 5.2204
Threshold	X = 8.49	Y = 4.9	RMS = 6.5292
Sort averaging	X = 8.49	Y = 4.9	RMS = 6.5292
HLOP	X = 7.5	Y = 4.3299	RMS = 5.2204
TSA	X = 8.52	Y = 4.8651	RMS = 6.5507

$$rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((x_i - \hat{x}_i)^2 \quad (y_i - \hat{y}_i)^2 \right)}$$

Another simulation was performed on performance for number of AOA lines on location estimation. The proposed algorithm uses three AOA information. Performance when no AOA information and when one or two is available at the BSs is also estimated.

5 Observation Tables and Graphs

Readings were taken for different TOA circles formed by mobile user in motion and one of the observations for $\eta = 6.00033$ are shown in Table 1. Also readings for different number of AOA information were taken and one of the observations is shown in Table 2. The respective graphs for the different observations are also shown below. The first graph plots RMS location error versus Upper bound on NLOS error. The second graph plots RMS location error versus number of AOA line information (Fig. 2).

Table 2 Readings for different numbers of AOA information were taken and one of the observations

BS1: R = 10 AOA = 40° BS2: R = 13 AOA = 160° BS3: R = 11 AOA = 71.5913°				
(3.262, 2.839)	(5.2, 8.5417)	(9.178, 2.119)	(9.7499, 2.2225)	(4.026, 2.553)
(4.454, 3.838)	(7.61, 6.487)	(2.784, 4.446)	(5.871, 8.094)	(2.4099, 3.2386)
Distance-weighted				
	X = 4.6476	Y = 4.0827	RMS = 12.1821	
Threshold				
	X = 4.5757	Y = 4.3907	RMS = 12.1794	
Sort averaging				
	X = 4.4272	Y = 4.0326	RMS = 12.6245	
HLOP				
	X = 5.1856	Y = 4.8543	RMS = 11.6506	
TSA				
	X = 4.4612	Y = 4.9626	RMS = 12.6668	

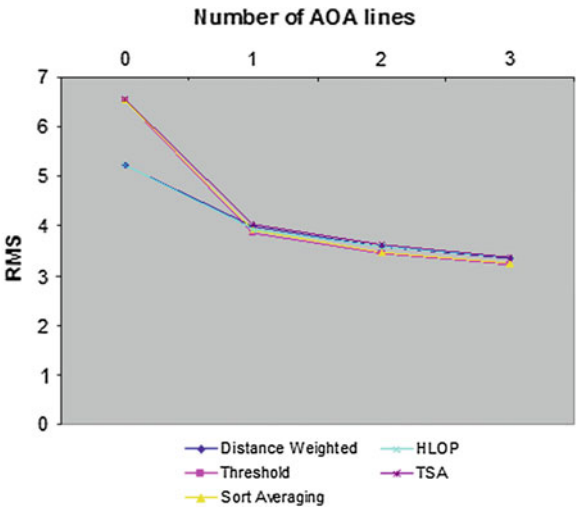


Fig. 2 Number of AOA lines vs RMS location error

6 Conclusion

In this report, we presented a modified hybrid TOA/AOA scheme to estimate MS location by hear ability from three BSs. The scope of the work consisted of learning and understanding TOA and AOA techniques of localization and then developing a hybrid scheme for the purpose of localization. The proposed scheme uses all possible intersection points of three TOA circles of BSs and three AOA line information for locating a mobile station without needing to know any knowledge about the NLOS error. The proposed algorithm is based on weighted summation of intersections points of three TOA circles and three AOA lines. The technique takes into account the non-line-of-sight errors in measurement and ways to mitigate these errors. The scheme aims to find the location of mobile node with better accuracy and less complexity.

References

1. Roxin A, Gaber J, Wack M, Nait-Sidi-Moh A (2007) Survey of wireless geolocation techniques. IEEE Globecom Workshops, pp 26–30
2. Caffery JJ, Stuber GL (1995) Radio location in urban CDMA microcells. In: Proceedings of IEEE international symposium personal, indoor, mobile radio communications, 2, pp 858–862
3. Li R, Fang Z (2010) LLA: A new high precision mobile node localization algorithm based on TOA. J Commun, 5(8)

Handwritten Kannada Numerals Recognition Using Discrete Cosine Transform Coefficients

S. K. Shreedharamurthy and H. R. Sudarshana Reddy

Abstract The process of pattern recognition poses quiet a lot of challenges especially in recognizing handwritten scripts of different languages in India, in spite of several advancement in technologies pertaining to optical character recognition (OCR). Handwriting continues to persist as means of documenting information for day-to-day life especially in rural areas. There exists a need to develop handwritten character recognition system for its applications in post offices, bank check processing, handwritten document processing, etc. In this paper, a handwritten Kannada numerals recognition system is proposed which makes the use of discrete cosine transform coefficients as feature vectors and neural network as the classifier. Higher degree of accuracy in results has been obtained with the implementation of this approach on a comprehensive database.

Keywords: Pattern recognition · OCR · Discrete cosine transformation · Kannada numerals

1 Introduction

Character recognition can solve more complex problems and help ease the drudgery involved in maintaining obscure image files. Basically, converting scanned images into text document can enable manipulation through word processing applications.

S. K. Shreedharamurthy (✉)
Electronics and Communication Department, University BDT College of Engineering, Davangere 577004, Karnataka, India
e-mail: sks_murthy@yahoo.com

H. R. Sudarshana Reddy
Electrical and Electronics Department, University BDT College of Engineering, Davangere 577004, Karnataka, India
e-mail: hrsreddy@hotmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_91,

Optical character recognition (OCR) has gained a momentum since the need for digitizing or converting scanned images of machine printed or handwritten text (numerals, letters, and symbols) into a format recognized by computers (such as ASCII). Handwriting recognition is the task of transforming a language represented in its own spatial form of graphical marks into a symbolic representation. Handwriting recognition inherited a number of technologies from OCR. The main difference between handwritten and typewritten characters is in the variations that come with handwriting. Traditionally, the field of handwriting recognition is divided into off-line and on-line recognition [1]. In off-line recognition, only the image of the handwriting is available for the computer, while in the on-line, case temporal information such as pen-tip coordinates as a function of time is also available. Typical data acquisition devices for off-line and on-line recognition are scanners and digitizing tablets, respectively. Due to the lack of temporal information, off-line handwriting recognition is considered more difficult than on-line. Furthermore, it is also clear that the off-line case is the one that corresponds to the conventional reading task performed by humans [2].

The need for OCR arises in the context of digitizing Kannada documents from the ancient and old era to the latest, which helps in sharing the data through the Internet [3]. Kannada, the native language of Karnataka (southern state) in India, has several million speakers across the world and has obtained cultural status from central government very recently. The penetration of information technology (IT) becomes harder in a country such as India where the majority people read and write in their native language especially in rural areas. Therefore, enabling interaction with computers in the native language and in a natural way such as handwriting is absolute necessary.

In the literature, many papers have been published with research detailing new techniques for the classification of handwritten characters. Some researchers have obtained very promising results for isolated/segmented characters using neural network-based techniques [4]. However, the results for the segmentation and recognition of handwritten characters have not been very good, and still, there is a need for improvement so that they can be used in real-world applications. In this paper, we propose a complete off-line OCR system for handwritten Kannada numerals. The scanned document image is preprocessed to ensure that the characters are in a suitable form. Finally, neural network is used for the recognition of numerals using discrete cosine transform features. A survey of feature extraction methods for character recognition is reported in [5]. Considerable amount of work has been carried out in numeral recognition through regional decomposition, histogram methods, Hough transformations, principal component analysis, support vector machines (SVMs), nearest neighbor, neural computing based approaches, and fuzzy theory-based approaches. An extensive survey of recognition performance for large handwritten database through many kinds of features and classifiers is reported in [6].

1.1 Kannada Numerals Set

The Kannada language is one of the four major south Indian languages. It is spoken by about 50 million people in the Indian states. The Kannada alphabet consists of 16 vowels and 36 consonants. It also includes 10 different symbols representing the ten numerals of the decimal number system and is given in Fig. 1.

Kannada is one of the important Dravidian languages like Malayalam, Telugu, and Tamil, which is the mother tongue of the Kannada people spoken in Karnataka. The language has gained official status in India and conferred as a classical language by the government of India very recently. The language is also the administrative language of the Indian state of Karnataka.

2 Numerals Recognition System

A typical numerals recognition system consists of three stage processes as shown in Fig. 2. The first stage is pre-processing, second stage is feature extraction, and the third stage is classification.

2.1 Pre-processing

Pre-processing involves normalizing the raw data given to the computer so that the further processing is easier. The initial stage of the process, the preprocessing steps, includes general signal processing algorithms and also more application-specific algorithms such as thinning, rotating, etc. This is a step where certain normalizations are done.

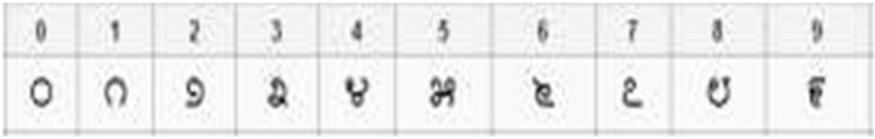


Fig. 1 Kannada numerals

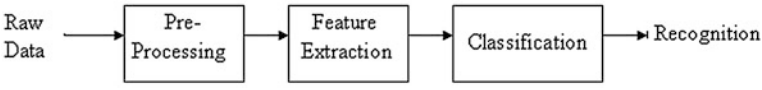


Fig. 2 Typical pattern recognition system

2.2 Feature Extraction

When the input data to an algorithm are too large to be processed and it is suspected to be redundant, then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called *features extraction*. If the features extracted are carefully chosen, it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full-size input. The set of features that are used makes up a feature vector, which represents each member of the population. Then, character recognition system classifies each member of the population on the basis of information contained in the feature vector.

Selection of a feature extraction method is probably the single most important factor in achieving high recognition performance in character recognition systems. The features that are extracted are fed to the classifiers for further processing. Discrete cosine transform feature vectors are used to classify the Kannada numerals in our system. Developed by Ahmed, the DCT is a close relative to the Discrete Fourier transform (DFT) which transforms a signal or image from the spatial domain to the frequency domain. The DCT is a technique for converting a signal into elementary frequency components. Each DCT uses N orthogonal real basis vectors whose components are cosines. The DCT approach has an excellent energy compaction property and requires only real operations in transformation process. Its most important feature is its superior energy compacting property. On applying DCT, a frequency spectrum (or the DCT coefficients) B_{pq} of an $M \times N$ image represented by A_{mn} for m = 0, 1, …, $M-1$, n = 0, 1, …, $N-1$ can be defined as

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{matrix}$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2/M}, & 1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2/N}, & 1 \le q \le N-1 \end{cases}$$

For most images, much of the signal energy lies at low frequencies; these appear in the upper left corner of the DCT. The lower right values represent higher frequencies and are often small enough to be neglected with little visible distortion.

Proposed method:

In the proposed method, the captured image (hand written numeral) has to be binarized so that the numeral images have pixel values 0 and 1. Each numeral image represents a numeral (binary 1) that is unconstrained, isolated, and clearly discriminated from the background (binary 0). This can be done by converting the captured image to BMP format, and then, preprocessing of the same is done.

Fig. 3 BMP image values

0 0 0 0 **1 1 1 1** 0 0 0 0

0 0 0 **1** 0 0 0 0 **1** 0 0 0

0 0 **1** 0 0 0 0 0 0 **1** 0 0

0 0 **1** 0 0 0 0 0 0 **1** 0 0

0 0 **1** 0 0 0 0 0 0 0 **1** 0

0 0 **1** 0 0 0 0 0 0 0 **1** 0

0 0 **1** 0 0 0 0 0 0 0 **1** 0

The BMP format of the image of Kannada numeral '1' contains only zeroes and ones and is as shown in Fig. 3.

The features from the bmp image can be obtained by reading the spatial coordinate (x, y) values of the pixels having the value '1'only.

The algorithm for generating the discrete cosine transform features is given below:

Algorithm

1. Read the pattern in bmp format.
2. Start from top left scan image line by line.
3. Note down the spatial coordinate values (x, y) of 'ON' cells in the first line.
4. Obtain remaining spatial coordinates repeating above step for all 'ON' cells in the image.
5. Apply 1D DCT for x-vector to obtain DCT coefficients. Select first few coefficients.
6. Apply 1D DCT for y-vector to obtain DCT coefficients. Select first few coefficients.
7. Selected coefficients form the feature vector.

2.3 Classification

Classification is a step in numerals recognition which accepts extracted features from the feature extraction step and identifies the pattern written. A large number of classifiers are available: parametric and nonparametric statistical classifiers, neural networks, SVMs, hybrid classifiers, etc. Artificial neural network (ANN) has been used as a classifier in our system.

2.3.1 Artificial Neural Network

Artificial neural network systems have great ability to learn by experience and generalize the inputs to produce reasonable outputs for inputs that were not encountered during learning (training).

The Multi-Layer Perceptron

Multi-layer perceptrons are one of many different types of the existing neural networks. They comprise a number of neurons connected together to form a network. The strength or a weight of the links between the neurons is where the functionality of the network resides. Its basic structure is shown in Fig. 4.

The idea behind neural networks stems from studies of the structure and function of the human brain. Neural networks are useful to model the behaviors of real-world phenomena. Being able to model the behaviors of certain phenomena, a neural network is able subsequently to classify the different aspects of those behaviors, recognize what is going on at the moment, diagnose whether this is correct or faulty, predict what it will do next, and if necessary respond to what it will do next.

Feed-Forward Back-Propagation Network

Feed-forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. Back-propagation learning rule is used to train multiple-layer networks. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities. The structure of feed-forward neural network is shown in Fig. 5.

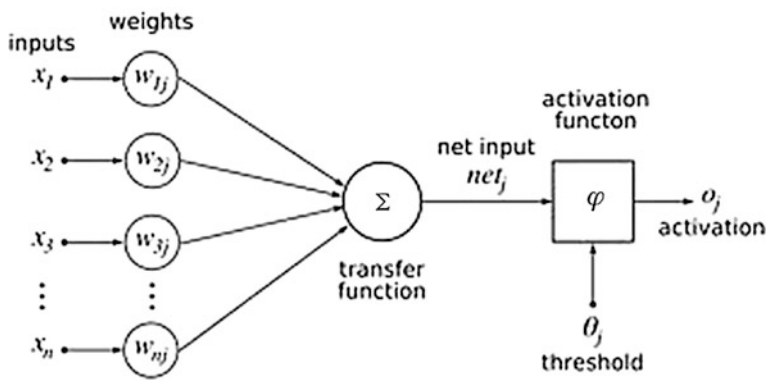


Fig. 4 Artificial neuron model

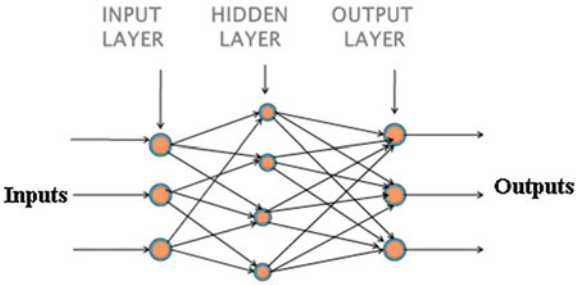


Fig. 5 Feed-forward neural network

Forward Propagation

Forward propagation is the process whereby each of all of the neurons calculates its output value, based on inputs provided by the output values of the neurons that feed it. The input neuron distributes the signal along multiple paths to hidden layer neurons.

A weight is associated with a hidden neuron. Each node of input layer is connected to every node of hidden layer. Likewise, each node of hidden layer is connected to every node of output layer again by some weights as shown in Fig. 5. Also, the data flow from left to right. Hence, the network is called feed-forward network. The output of a neuron is a function of its net input. This function can be trigonometric, hyperbolic, or sigmoid function.

Error Back-Propagation Learning

Back-propagation learning algorithm is popularly used to train a feed-forward neural network. The error back-propagation consists of two passes through the different layers of the network: a forward pass and a backward pass.

Forward pass is same as the forward propagation. Back-propagation is an iterative process that starts with the last layer and moves backwards through the layers until the first layer is reached.

For each set of inputs, a set of target values is provided. During learning, the difference between the set of output and target values is found to get the error value. After the feed-forward process, this error is back-propagated and weights between the layers are adjusted starting from output layer to hidden layer and then from hidden layers to input layer to minimize the error. This process is repeated till it reaches the required minimum error value.

2.4 Training

The process of training is preparing the ANN to recognize the desired set of characters. For character to be recognized, a set of similar characters with different size and little variation in their shape are written and used for training. Standard back-propagation is a gradient descent algorithm, in which the network weights are moved along the negative of the gradient of the performance function. The term back-propagation refers to the manner in which the gradient is computed for nonlinear multi-layer networks. There are a number of variations on the basic algorithms those are based on other standard optimization techniques, such as conjugate gradient and Newton methods.

With standard steepest descent, the learning rate is held constant throughout training. The performance of the algorithm is very sensitive to the proper setting of the learning rate. If the learning rate is set too high, the algorithm may oscillate and become unstable. If the learning rate is too small, the algorithm will take too long to converge. It is not practical to determine the optimal setting for the learning rate before training, and, in fact, the optimal learning rate changes during the training process, as the algorithm moves across the performance surface.

The gradient descent algorithm for training the multi-layer perceptron was found slow especially when getting close to a minimum (since the gradient is disappearing). One of the reasons is that it uses a fixed-size step. In order to take into account the changing curvature of the error surface, many optimization algorithms use steps that vary with each iteration. In order to solve this problem, an adaptive learning rate can be applied to attempt keeping the learning step size as large as possible while keeping learning stable. The learning rate is made responsive to the complexity of the local error surface. In this approach, new weights and biases are calculated using the current learning rate at each epoch. New outputs and errors are then calculated. As with momentum, if the new error exceeds the old error by more than a predefined ratio, for example, 1.04, the new weights and biases are discarded. In addition, the learning rate is decreased. Otherwise, the new weights are kept. If the new error is less than the old error, the learning rate is increased. This procedure increases the learning rate.

In our system, a feed-forward multi-layer perceptron with a single hidden layer and trained by gradient descent with momentum and a learning rate back-propagation method was applied to the digit classification problem.

2.5 Recognition

Once the ANN is trained to recognize a set of numerals, it is ready to use for recognizing digits in the numerals recognition system. During recognition phase, ANN has the capacity to generalize and identify the numerals written with little variations when compared to the numerals used for training.

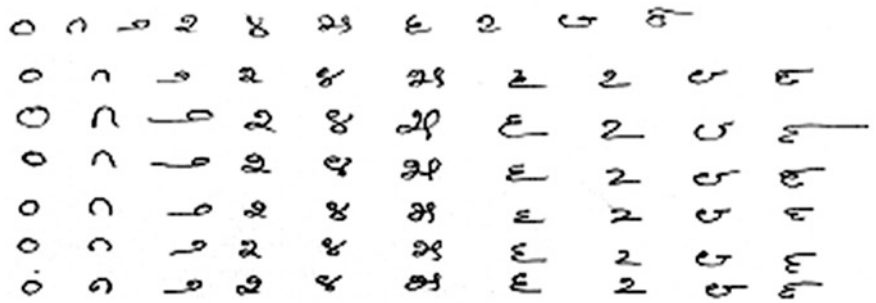


Fig. 6 A sample patterns of Kannada handwritten numerals 0–9

3 Results and Conclusion

The proposed system describes an approach which uses discrete cosine transform features and ANN as classifier to recognize handwritten Kannada numerals. We have used 100 samples of each numeral from the created database, sample patterns of which are shown in Fig. 6. Out of which 80 patterns used for training phase and 20 samples for testing phase. We achieved around 90 % of recognition rate.

References

1. Plamondon R, Srihari SN (2000) On-line and off-line handwriting recognition: a comprehensive survey. IEEE Trans Pattern Anal Mach Intell 22(1):63–79
2. Bahlmann C (2006) Directional features in online handwriting recognition. Pattern Recogn Lett 39:115–125
3. Kannan RJ, Prabhakar R (2008) Off-line cursive handwritten Tamil character recognition, WSEAS transactions on signal processing, vol 4(6). ISSN: 1790-5052
4. Prema KV, Subbareddy NV (2002) Two-tier architecture for unconstrained handwritten character recognition. Sadhana 27(5):585–594
5. Due Trier O, Jain AK, Taxt T (1995) Feature extraction methods for character recognition. A surve, July 1995
6. Pal U, Chaudari BB (2004) Indian script character recognition: a survey. Pattern Recogn 37:1887–1899
7. Chen GY, Bui TD, Krzyzak A (2003) Contour based numeral recognition using mutiwavelets and neural networks. Pattern Recogn Lett 36:1597–1604
8. Goltsev A, Rachkovskij D (2005) Combination of the assembly neural network with a perceptron for recognition of handwritten digits arranged in numeral strings. Pattern Recogn Lett 38:315–322
9. Liu CL, Nakashima K, Sako H, Fujisawa H (2004) Handwritten digit recognition: investigation of normalization and feature extraction techniques. vol. 37, Elsevier Publication, Pattern Recognition, pp 265–279
10. Zhang B, Fu M, Yan H (2001) A nonlinear neural network model of mixture of local principal component analysis: application to handwritten digits recognition. Pattern Recogn Lett 34:203–214

11. Rajput GG, Horakeri R, Chandrakant S (2010) Printed and handwritten mixed Kannada numerals recognition using SVM. IJCSE 02(05):1622–1626
12. Arica N, Yarman-Vural FT (2000) One dimensional representation of two dimensional information for HMM based handwriting recognition. Pattern Recogn Lett 21:583–592
13. Kim H-C, Kim D, Yang Bang S (2002) A numeral character recognition using PCA mixture model. Pattern Recogn Lett 23:103–111

Development of Randomized Hybrid Cryptosystem Using Public and Private Keys

A. C. Nuthan, M. S. Naveen Kumar, Shivanand S. Gornale and M. Basavanna

Abstract Before the modern era, cryptography was concerned solely with message confidentiality (i.e., encryption), conversion of messages from a comprehensible form into an incomprehensible one and back again at the other end, rendering it unreadable by interceptors without secret knowledge (namely the key needed for decryption of that message). This paper presents a highly secured randomized encryption/decryption technique. Implementation presented in this paper chooses a randomly switching model which combines the symmetric cryptosystem Advanced Encryption Standard (AES) and asymmetric cryptosystem elliptic curve cryptography (ECC) to work together as a single unit. Symmetric algorithms provide high confidentiality using a single key, incorporating low power and memory resources. Asymmetric algorithms provide strong security strength using a pair of keys enabling rapid implementations. This model being implemented on FPGA provides excellent area results, and a higher level of security is achieved.

A. C. Nuthan (✉) · M. S. Naveen Kumar
G. Madegowda Institute of Technology, Mandya, Karnataka, India
e-mail: nuthancnayak@gmail.com

M. S. Naveen Kumar
e-mail: navi.katte@gmail.com

S. S. Gornale · M. Basavanna
Department of Computer Science, Government College (Autonomous), Mandya, Karnataka, India
e-mail: shivanand1971@gmail.com

M. Basavanna
e-mail: basavanna_m@yahoo.com

S. S. Gornale · M. Basavanna
Department of Computer Science, Government College (Autonomous), Mandya, Karnataka, India
e-mail: shivanand1971@gmail.com

M. Basavanna
e-mail: basavanna_m@yahoo.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_92,

Keywords Hybrid cryptosystem · Advanced Encryption Standard (AES) · Elliptic curve cryptography (ECC) · Random number generator (RNG)

1 Introduction

Cryptography is the art and science of protecting information from undesirable individuals by converting it into a form non-recognizable by its attackers while stored and transmitted [1, 2]. Data cryptography mainly is the scrambling of the content of data, such as text, image, audio, video, and so forth to make the data unreadable during transmission or storage called encryption. The reverse of data encryption is data decryption.

In recent decades, the field has expanded beyond confidentiality concerns to include techniques for message integrity checking, sender or receiver identity authentication, digital signatures, interactive proofs and secure computation, among others.

The crypto systems that are in use are static, that is, only one of the standard algorithms is used at a time. If same encryption methods are used throughout the day and for days together continuously, the crypto-analysers may crack the code or find out the encryption method and may intrude or spy or hack the confidential data. This necessitates a robust encryption technique. This paper highlights a novel approach of highly secured cryptic algorithm in FPGA which makes nearly impossible for data theft. Many researches have been made with the multiple algorithms. Instead of dedicating the hardware, that is, FPGA for single algorithm, the paper [3] presents a hardware implementation of three standard cryptography algorithms on a universal architecture. But in this design, multiple algorithms are implemented on a universal architecture and there is switching between the algorithms at random times.

Though multiple crypto algorithms can be implemented, for simplicity, only two algorithms are deployed:

- AES Advance Encryption Standard
- ECC Elliptic Curve Cryptography

Implementation of AES and ECC are dealt in [4] , [5] and [6]. The Implementation of the combination of AES and ECC is done in [7] , [8] and [9].

2 Advanced Encryption Algorithm

AES Advanced Encryption Standard implementation on FPGA with encryption and decryption [10] and AES embedded hardware implementation in [11] National Institute of Standards and Technology (NIST) in 1997 issued a call for proposals for a new Advanced Encryption Standard (AES) with requirements [1] and

evaluation criteria [12]. NIST selected Rijndael as the proposed AES algorithm and published a final standard Advanced Encryption Standard in November of 2001 [13]. The algorithm processes data block of size 128 bits using a cipher key of lengths 128, 192, and 256 bits. There are four basic transformations:

1. Substitute bytes: Byte-by-byte substitution of the block using S-box (table)
2. Shift rows: Circular shift or a simple permutation operation is performed.
3. Mix columns: Each column of the block is multiplied with a constant matrix, and the result obtained is substituted back to the block which makes use of arithmetic over GF (Galois field).
4. Add round key: Simple bitwise XOR operation between the current block and expanded key.

These transformations are applied to the data block (i.e., state). The structural resemblance for both encryption and decryption makes hardware implementations easier [14].

3 Elliptic Curve Cryptography

Elliptic curve cryptography, an approach to Elliptic Curve Cryptography, was independently suggested by Neil Koblitz and Victor S. Miller in the year 1985. It is based on the algebraic structure of elliptic curves defined over finite fields. The prime advantage of elliptic curve cryptography is that the key length can be much smaller.

An elliptic curve is a plane curve which is isomorphic to a curve defined by a cubic equation of the form:

$$\mathbf{y^2 = x^3 + ax + b}$$

If $\mathbf{y^2 = x^3 + ax + b}$ contains no reprised factors or if $\mathbf{4a^3 + 27b^2 \neq 0}$, then the elliptic curve can be used to form an Abelian group, with the point at infinity as the identity element. There are four basic operations performed on points in elliptic curve: point addition, point subtraction, point doubling, and point multiplication [1].

One-to-one point mapping is used to map each point on elliptic curve to message [15]. Since ECC encrypts 8 bits at a time, in order to cope with 128 bit AES, here ECC is considered for 128 bits, that is, even ECC can encrypt 128 bits at a time. Authenticated ECC implementation is given in [16]. Analytical study and implementation of ECC are given in [17] with their complexity and practical issues.

4 Random Number Generator

For simplicity, instead of 4-bit leap-forward Random Number Generator [18], a 2-bit leap-forward LFSR is considered. Therefore, the equation becomes

$$q_{0\_next} = q_1$$

$$q_{1\_next} = q_0$$

Hence, if the seed is '01', then the output sequence is 10 01 10 01.

5 Proposed System, Execution, and Results

The communication between desktop and FPGA is done using HyperTerminal application supported by Windows OS and UART cable. The universal asynchronous receiver and transmitter (UART) transmits 11 bits in a frame out of which 8 bits are data bits. The first bit is the start bit, and the last bit of the frame is the stop bit.

Eight characters (64 bits) given after every reset or start-up acts like password.

- Two copies of password which are concatenated (16 characters and 128 bits) are key for AES encoder/decoder and a private key [1] for sender in ECC encryption.
- First two bits of password act like seed for random number generator.
- Each bit is select line between AES and ECC, that is, if bit is high, then AES is active or else ECC is active.

Consider

- Password:
  - NNNNNNNN.
  - 01001110010011100100111001001110010011100100111001001110.

Therefore

- AES key and ECC private key:
  - NNNNNNNNNNNNNNNN.
  - 01001110010011100100111001001110010011100100111001001110010011100100111001001110010011100100111001001110.
- Seed to RNG: 01 (First 2 bits of Password-NNNNNNNN).

For simplicity, let us consider first character of the password, that is, N (01001110), then the switching between the AES and ECC is done as shown in Table 1. Similarly, this is continued with the next bit of the password infinitely. Taking '01' as seed, the Random Number Generator gives '10'(2) as the output which is fed to down counter. Since the first bit of password is '0,' the ECC line is selected. The 8 bit data appear at the demultiplexer, and the 8 bits of data are serially shifted to a 128 bit (16 characters) SIPO shift register. And counter is decremented once, that is, count value 2 is decremented to 1. When the SIPO

Table 1 Sequence of switching between Advanced Encryption Standard and Elliptic Curve Cryptography

N	Switching between ECC and AES	Rounds	
0	Starts with ECC	2 times ECC	2 times ECC then switching to AES
1	ECC→AES	1 time AES	1 time AES then switching to ECC
0	AES→ECC	2 times ECC	3 times ECC then switching to AES
0	Remains in ECC	1 time ECC	
1	ECC→AES	2 times AES	5 times AES then switching to ECC
1	Remains in AES	1 time AES	
1	Remains in AES	2 times AES	
0	AES→ECC	1 time ECC	1 time ECC

Fig. 1 Proposed module of encoder/decoder

contains 128 bits of data to be encoded (or decoded), the 128 bits of data are loaded to the AES encoder (or decoder) as shown in Fig. 1.

The encoded (or decoded) data are shifted out serially to the UART through multiplexer. Since the count value is non-zero, again the ECC line is selected and the above steps are repeated for next 16 characters. But this time when the counter is decremented, then the count value becomes '0', and hence

- It signals the RNG to give the next random number which is '01'(1)
- Checks the next bit, that is, second bit of password and since its '1' AES line is selected

AES encoder(or decoder) takes 10 clock pulses to complete the 10 rounds of encoding(or decoding) operations, and at the 11th clock pulse, the DONE signal goes high and AES encoder loads encrypted data into 128 bit PISO shift register. The encoded (or decoded) data are shifted out serially to the UART through multiplexer. But this time when the counter is decremented, then the count value becomes '0', and hence

- It signals the RNG to give the next random number which is '10'(2)
- Checks the next bit, that is, third bit of password and since its '0' ECC line is selected

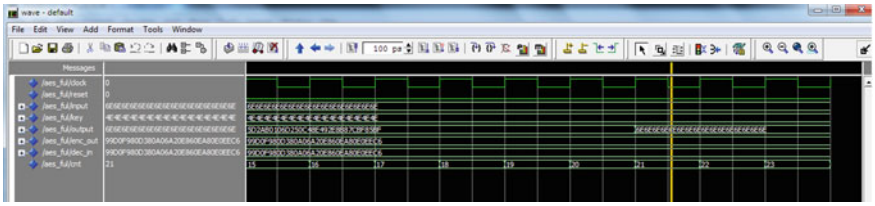


Fig. 2 Simulation results of AES encoder and decoder

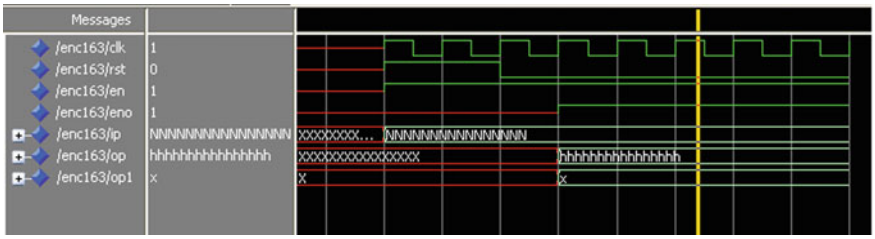


Fig. 3 Simulation results of ECC encoder and decoder

Similarly, this is continued with the next bit of the password infinitely.

The simulation done on Modelsim of AES and ECC encoder and decoder clubbed into one is shown in the Figs 2 and 3.

When 16 bytes of data in case of AES or ECC are fed into Xilinx simulator, we can observe that all the 16 bytes or 8 bytes of encrypted or decrypted data for AES or ECC, respectively, but when the data are dumped on FPGA kit, because of some design constraints, that is, for same clock pulse, from encoder data are loaded to 128-bit register and also contents of 4-bit TEMP register get loaded into 128-bit register (into LSB). As a result, we are losing one byte of data in case of both AES and ECC, respectively.

Encoder and decoder have slight inconsistency, and the encoded characters correlate with the input data.

6 Conclusions

- The concept of universal coding with random switching can be practically implemented.
- Implementation of universal coding on the FPGA is successful, and multiple data encryption and decryption standards can be successfully integrated and can be switched randomly among them. So instead of using single algorithm, the usage of multiple algorithms helps to increase the security.

- There is a flexibility to remove or add any other cryptographic standards. These changes can be on the fly.
- The cipher key (password) can be changed with respect to the user or the company requirements.
- Since AES and ECC are used, the universal coding inherits the statistical and other advantages of AES and ECC.
- Switching randomly between the algorithms also enhances the security.
- Since the password acts as AES key and seed for RNG and controls switching between AES and ECC, universal coding is completely randomized depending upon the password, that is, the intruder will not know which algorithm is active at what time and for how long.

In future, this prototype can also be imported to ASIC.

References

1. William S (2005) Cryptography and network security principles and practices, 4th edn. Printice Hall, NY
2. Forouzan BA (2007) Cryptography and network security. Tata Mc-Graw Hill, New Delhi (Special Indian Edition)
3. Eslami Y, Sheikholeslami A, Gulak PG, Masui S, Mukaida K (2006) An area-efficient universal cryptography processor for smart cards. IEEE Trans Very Large Scale Integr (VLSI) Sys 14:43–56
4. Daemen J, Rijndael RV (2001) The Advanced Encryption Standard. Dr. Dobb's J
5. Batina L, Mentens N, Sakiyama K, Preneel B, Verbauwhede I (2006) A low-cost elliptic curve cryptography for wireless sensor networks. Springer, Berlin
6. Abomhara M, Zakaria O, Khalifa OO, Zaidan AA, Zaidan BB (2010) Enhancing selective encryption for H.264/AVC using advance encryption standard. Int J Comput Electr Eng (IJCEE) 2(2):224–229, ISSN: 1793-8198
7. Li X, Chen J, Qin D, Wan W (2010) Research and realization based on hybrid encryption algorithm of improved AES and ECC. Audio Language and Image Processing (ICALIP), 2010 International Conference on IEEE, pp 396–400 (978-1-4244-5858-5/2010)
8. Ganesh AR, Manikandan PN, Sethu SP, Sundararajan T, Pargunarajan K (2011) An improved AES-ECC Hybrid encryption scheme for secure communication in co-operative diversity based wireless sensor networks. Recent Trends in Information Technology (ICRTIT), 2011 International Conference on IEEE, pp 1209–1214 (978-1-4577-0590-8,2011)
9. Wang J, Zeng X, Chen J (2006) A VLSI implementation of ECC combined with AES. IEEE 8th International Conference on Solid-State and Integrated Circuit Technology, pp 1899–1904 (1-4244-0161-5/2006)
10. Federal Information Processing Standards Publication (1999) FIPS PUB 46-3, Reaffirmed 25 Oct 1999
11. Zaidan AA, Zaidan BB, Majeed A (2009) High securing cover-file of hidden data using statistical technique and AES encryption algorithm. World Acad Sci Eng 54:468–479
12. Jyrwa B, Paily R (2009) An area-throughput efficient FPGA implementation of block cipher AES algorithm. IEEE Comput Soc 328–332
13. Advanced Encryption Standard (AES) (2001) FIPS 197, 26 Nov 2001
14. Deshpande AM, Deshpande MS, Kayatanavar DN (2009) FPGA implementation of AES Encryption and Decryption. International conference on "control, automation, communication and energy conservation" Proc.CACEC'09,2009, paper 1, pp 1–6

15. Konheim A (1981) Cryptography: a primer. Wiley, NY
16. Chang Q, Zhang Y, Qin L (2010) A node authentication protocol based on ECC in WSN. Computer design and applications (ICCDA), 2010 international conference on IEEE, vol 2, pp V2606–V2609 (978-1-4244-7164-5.2010)
17. Shah PG, Huang X, Sharma D (2010) Analytical study of implementation issues of elliptical curve cryptography for wireless sensor networks Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on IEEE, pp 589–592, (978-0-7695-4019-1/10.2010)
18. Bardell PH, McAnney WH, Savir J (1987) Build-in test for VLSI: pseudo-random techniques. Wiley, NY

Author Biographies

Nuthan A.C completed BE in Electronics and Communication Engineering and M. Tech in Digital Electronics and Communication System. Presently he is working as an Assistant Professor in Electronics Communication Engineering, Department of ECE, GMIT, Bharathi Nagar, Mandya, Karnataka, India. He is pursuing his Ph.D. from Jain University, Bangalore.

Naveen Kumar M S completed BE in Electronics and Communication Engineering and M. Tech in Digital Electronics and Communication. Presently he is working as an Assistant Professor in Electronics Communication Engineering, Department of ECE, GMIT, Bharathi Nagar, Mandya, Karnataka, India.

Dr. Shivanand S Gornale completed M. Sc in Computer Science, M. Phil in Computer Science, and Ph.D. in Computer Science from University of Pune, Pune, and has been recognized as guide for Ph.D. in Computer Science and Engineering from Jain University, Bangalore, and he has published 55 research papers in various national and international journals and conferences. He is an editorial member for International Journal on Computer Science and IT, International Journal of Bioinformatics and Soft Computing (IJBSC), and International Journal of Computer Science and Application and also working as a reviewer for various international and national journals and conferences. He has worked as National/International Conferences and Workshops Technical Committee Member like ICCMS-2009, ICFME-2009, ICIFE 2010, and ICKEM 2011. He is a fellow of IETE, New Delhi; Life Member of CSI, Member of Indian Association for Research in Computer Science (IARCS); Tata Institute of Fundamental Research (TIFR), Mumbai; Member of International Association of Computer Science and Information Technology (IACS&IT), Singapore; Member of International Association for Engineers, Hong Kong; Member of Computer Science Teachers Association, USA; and Graduate Member of IEEE, Life Member of Indian Science Congress Association, Kolkata, India. Presently he is working as a Head, Postgraduate Department of Computer Science, Government College Mandya (Autonomous), Karnataka. His areas of research interest are biometrics, image processing and pattern recognition and information communication technology (ICT).

Basavanna M completed M. Sc in Computer Science, M. Phil in Computer Science, and pursuing Ph.D. in Computer Science from VELS University, Chennai, Tamil Nadu. Presently he is working as an Assistant Professor, Department of Computer Science, Government College Mandya (Autonomous), Karnataka. His areas of research interest are image processing and pattern recognition.

Dynamic Coverage by Homogeneous Mobile Sensor Nodes Using Tri-Hexagonal Scheme

R. J. D' Souza and Ganala Santoshi

Abstract Mobile sensor nodes (MSNs) are energy constrained, recharge, and replacement of battery is not possible after deployment. They spend more energy during their traversal and can be conserved by minimizing the traversal. High degree of coverage is preferred in many applications. Energy conservation is one of the most demanded researches. MSNs with variable sensing range are best suitable in search-related operations. Degree of coverage can be increased whenever the situation demands.

Keywords Mobile sensor nodes · Traversal algorithms · Coverage and connectivity

1 Introduction

Mobile sensor nodes (MSNs) inherit all resource constraint properties from wireless sensor nodes. MSNs are equipped with locomotive which can move around after been deployed. A collection of such low-powered electro-mechanical de-vices with limited on-board processing and wireless communication capabilities is called a mobile sensor network. Placement of static sensor nodes might not be possible in many un-habitation situations. MSNs are the best suitable in such hazardous and disastrous situations.

MSNs are energy constrained. They have limited battery power. Battery replacement or recharge is not possible once they are deployed. Energy depletion of MSN is largely proportional to total distance it travels. In order to conserve the energy of an MSN to extend network lifetime, their traveling distance must be

R. J. D' Souza · G. Santoshi (✉)
Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka Surathkal, Mangalore, India
e-mail: ganala.santoshi@gmail.com

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0_93,

minimized. When a group of MSNs are deployed for a specific task, all MSNs should travel approximately an equal distance. The MSN which travel longer distance die early compared to the MSN which travel shorter distance.

A point is said to be k-covered, if it lies within the sensing range of k distinct sensors. The term k is refereed as degree of coverage, and it is considered as one of the basic requirements in search-related operations. When a group of MSNs are deployed to scan a region of interest (ROI), they have to coordinate each other to perform search- and coverage-related operations along with regular operations like identification of location, data collection and further traversal patterns.

2 Related Work

MSNs that are equipped with locomotive and GPS facilities can move around. Energy conservation is an important issue in MSNs and it can be done by minimizing the total distance to be traveled. Along with main objective of minimizing the distance to be traveled, a high degree of coverage should also be maintained. In many applications, high degree of coverage is preferred, because a point which is covered by k sensors can tolerate up to $(k - 1)$ failed sensors.

In this section, we discuss the literature related to mobile traversal algorithms (MTA) and variable sensing ranges separately. To best of our knowledge, no research has been done in MTA where MSNs have variable sensing range. So, the related work is reviewed in two deferent aspects.

2.1 Mobile Traversal Patterns

Khan et al. [1] divided rectangular ROI into equilateral triangles. Area under each equilateral triangle is 1-covered by the three MSNs which are placed at the vertices of the equilateral triangle. To cover the horizontally neighboring triangle, the sensor that is not adjacent to the neighboring triangle is moved to the third vertex of the adjacent triangle, so that it is covered. When the sensors reach one end, they move either up or down to cover the next row. Here again, the node that is not adjacent to the next triangle is moved to cover the area under the adjacent triangle.

D' Souza et al. [2] divided rectangular ROI into regular hexagons in tessellation fashion. An equilateral triangle is formed by connecting the center points of two horizontally adjacent regular hexagons and another hexagon which is vertically sharing sides with them. Three MSNs are placed at the vertices of equilateral triangle.

To 2-cover a rectangular ROI, Purohit et al. [3] divided it into regular hexagons. Each regular hexagon is further divided into two isosceles triangles and a rectangle. Area under a rectangle is 2-covered by the four MSNs, which are placed at the vertices of rectangle. Three out of four MSNs are used to 2-cover the area

occupied by the isosceles triangle. This method is not free from coverage holes. In addition to coverage holes, the algorithm does not balance the total distance traveled among the deployed nodes.

To 2-cover a rectangular ROI, D' Souza and Ganala [4] divided it into regular hexagons. Three homogeneous MSNs are placed at the vertices of equilateral triangle. Another mobile sensor node, with large sensing range, is placed at the center of the equilateral triangle.

2.2 Variable Sensing Ranges

In order to prolong the network lifetime by minimizing the energy consumption of sensor nodes, Zhou et al. [5] used variable radii sensor nodes. These sensor nodes can adjust both the sensing and communication ranges. With experimental results, they showed that sensors with variable sensing range and transmission range are more efficient in prolonging the network lifetime with desirable degree of coverage and connectivity.

To provide reliable surveillance on physical or environmental conditions, Wang and Medidi [6] designed dynamic coverage algorithms. With simulation results, they showed that complete coverage is guaranteed with sensors having variable sensing ranges. Along with coverage, the other performance improvements are observed in coverage ratio, energy consumption, node failure rate and network lifetime.

According to Zhou [7], maximal lifetime with constrained energy (MLCE) problem is NP-complete. A large number of sensors are randomly deployed to cover the large number of targets which are randomly deployed over an area of interest. Energy consumption for sensing varies with the sensing range of the sensor. So, to maximize the total operational time of the network, sensors with continuously adjustable sensing ranges are used.

To provide energy efficient barrier coverage, Wang et al. [8] deployed sensors with adjustable sensing ranges. From their simulation results, it was observed that sensors with adjustable sensing range are reasonable in power consumption and can greatly improve network lifetime.

In this paper, problem is formulated in Sect. 3. ROI partition is discussed in Sect. 4. In Sect. 5, Dynamic Coverage Patterns are detailed. MTA is explained in Sect. 6. Results are analyzed in Sect. 7. We conclude the paper in Sect. 8.

3 Problem Formulation

All the above MTAs are coverage degree specific. The degree of coverage remains the same throughout the traversal. In order to conserve the energy of MSNs, the total distance to be traveled must be minimized, and sensing range of the sensor

must be varied as per the requirement. So, we propose an MTA in which the sensing range of the sensor can be varied. With the variable sensing ranges, the degree of coverage can be increased whenever it is required.

4 ROI Partition

A rectangular ROI of length L and width W is considered, where $L \gg W$. It is divided into regular hexagons in tessellation fashion. Side of the regular hexagon is limited by the normal sensing range r_{s1} of the MSN. The ROI is covered using $N_{\max}$ rows and each row having $M_{\max}$ tri-hexagons. Rows are numbered from 1 to $N_{\max,}$ and columns are numbered from 1 to $M_{\max}$. The values of $N_{\max}$ and $M_{\max}$ can be determined from Eqs. (1) and (2).

This partition is similar to the partition detailed in [4], as shown in Fig. 1.

$$L = \frac{\sqrt{3}}{2}(3M_{\max} - 1)r_{s1} \tag{1}$$

$$W = \left[3(N_{\max} - 1)\frac{5}{2}\right]r_{s1} \tag{2}$$

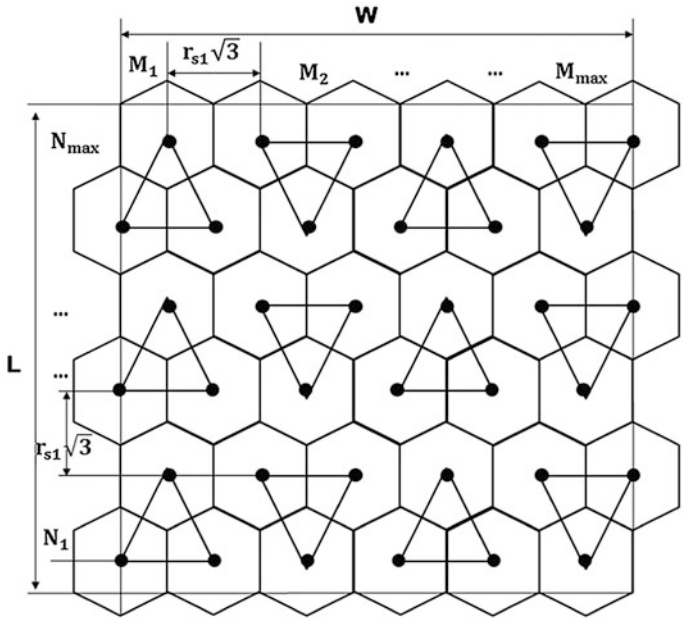


Fig. 1 ROI partition

The distance between the center points of adjacent regular hexagons is $r_{s1}\sqrt{3}$. Connecting the center points of two horizontally adjacent regular hexagons and another hexagon which is vertically sharing sides with them forms an equilateral triangle of side length $r_{s1}\sqrt{3}$. The structure of three regular hexagons which are connected by an equilateral triangle is refereed as tri-hexagon.

D'Souza and Ganala [4] defined the sensing point as a location, where the MSNs perform the predetermined operations. We use the same definition in this work.

5 Dynamical Coverage Patterns

MSNs are equipped with variable sensing range capabilities. Normal sensing range of the MSNs is $r_{s1,}$ and the extended sensing range is r_{s2}, where $r_{s2} \geq 2 \times r_{S1}$. The normal sensing range of the MSN is shown as dark circle, and extended sensing range is shown as dotted circle in Fig. 2.

Three MSNs with normal sensing range are placed at the vertices of the equilateral triangle, as shown in Fig. 3. In the event of detection of an object in any one of the MSNs sensing range, such information will be passed on to the remaining MSNs. One of the MSNs will increase its sensing range to $2r_{s1}$ to provide additional degree of coverage.

In Fig. 4, two MSNs, namely B and C, are with the normal sensing range. The third MSN, namely A, extends its sensing range and moves to the center point the tri-hexagon. With this, the covered area under the MSN B and C is 2-covered, and the area under MSN A is 1-covered. Whenever a demand arises, one MSN is selected among the three MSN on alternate basis. This procedure is repeated for all the sensing points of the ROI.

The distance between the vertex and center point of the equilateral triangle is r_{s1}. The MSN which extends its sensing range will move to the center point of equilateral triangle. On successful completion of predetermined operation, the

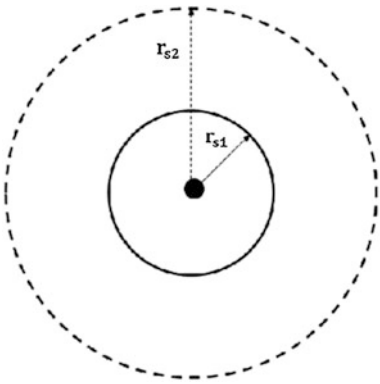


Fig. 2 MSNs with normal sensing range

MSN goes back to its original position and reduces the sensing range to r_{s1}. This makes MSN to have to an additional traversal of distance $2r_{s1}$. This procedure is repeated whenever the situation demands. This working procedure is detailed in Algorithm 1.

Algorithm 1 Dynamic Coverage Patterns

```
1: Select the MSN on alternative basis
2: Extend the Sensing Range to r_S2
3: Move MSN to the center point of the equilateral triangle
4: if Operation == Successful then
5:     Move MSN from the center point of the equilateral triangle to back to the original
       position
6:     Reduce Sensing Range to r_S1
7: end if
```

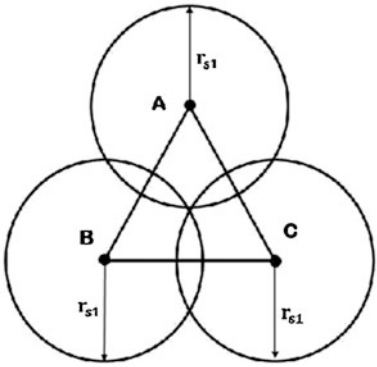


Fig. 3 MSNs at the vertices of equilateral triangle

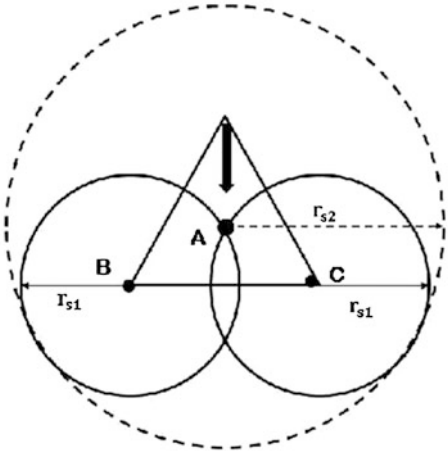


Fig. 4 MSN "A" with extended sensing range

6 Mobile Traversal Algorithm

Initially, three MSNs, namely A, B and C, are placed at the vertices of tri-hexagon with index (M_1 and N_1). On successful completion of predetermined operations at this sensing point, these three MSNs are moved to next tri-hexagon which is vertically adjacent to the current tri-hexagon.

This procedure is repeated for all intermediate sensing points of the current column. On successful completion of predetermined operations at the last sensing point of the current column (M_1 and N_{max}), three MSNs are moved horizontally to occupy next tri-hexagon in next column, that is,(M_2 and N_{max}). Operations at all the intermediate sensing points of current column are repeated as was done in previous column. On reaching last sensing point of current column, that is, (M_2, N_1), all the three MSNs are moved to horizontally adjacent tri-hexagon. This procedure is repeated till the whole ROI is covered. An algorithm for MTA is detailed in Algorithm 2, and mobility pattern is shown in Fig. 5.

7 Results and Analysis

The distance between the center points of two vertically adjacent tri-hexagons is $3r_{s1}$. Total number of tri-hexagons in a column is N_{max}. Distance traveled by three

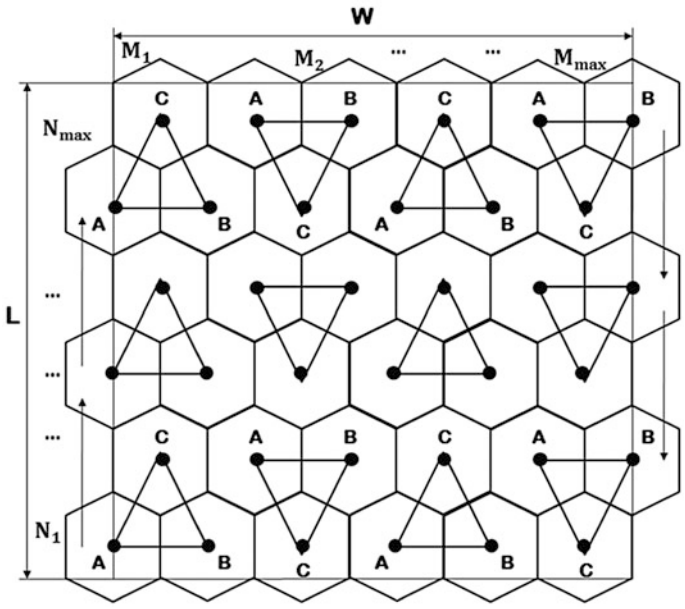


Fig. 5 Mobility pattern for MSNs

MSNs to cover one column is $(N_{\max} - 1) \times 3r_{s1} \times 3$. ROI contains $M_{\max}$ number of columns. So, the distance traveled by the MSNs to cover all the columns is $[(N_{\max} - 1) \times 3r_{s1} \times 3] \times M_{\max}$. The distance between the center points of two horizontally adjacent tri-hexagons is $3r_{s1}$. Distance traveled by the MSNs in horizontal direction is $(M_{\max} - 1) \times 3r_{s1} \times 3$. Total distance TDT_N with normal sensing range, by the three MSNs to cover the entire ROI, is

Algorithm 2 Mobile Traversal Algorithm

```
1: Set N=1 and M=1 [Place three MSNs at tri-hexagon with index (M,N)]
2: while M ≤ M_max do
3:     if N == 1 then
4:         while N ≤ N_max do
5:             if Object Found == True then
6:                 Call Dynamic Coverage Pattern
7:             end if
8:             N =←N + 1 [Move the MSNs to next tri-hexagon ]
9:         end while
10:        M ←M + 1 [Move the MSNs to next column]
11:    end if
12:    if N == N_max then
13:        while N_max ≥ 1 do
14:            if Object Found == True then
15:                Call Dynamic Coverage Pattern
16:            end if
17:            N =←N − 1 [Move the MSNs to next tri-hexagon ]
18:        end while
19:        M ←M + 1 [Move the MSNs to next column]
20:    end if
21: end while
```

$$\text{TDT}_N = 3 \times 3\, r_{S1} \times (N_{\max} \times M_{\max} \times -1) \tag{3}$$

To provide high degree of coverage, one MSN has to make an extra traversal of distance $2r_{S1}$. Extra distance traveled by the three MSNs for providing high degree of coverage throughout the ROI is $N_{\max} \times M_{\max} \times 2\, r_{S1}$. Total distance traveled including extra distance TDT_E by the three MSNs is

$$\text{TDT}_E = r_{S1} \times (11 \times N_{\max} \times M_{\max} - 9) \tag{4}$$

For simulation purpose, rectangular ROI of measure 4,500 × 2,000 units is considered. Normal sensing range of MSN is varied from 45 to 65 units in step 5 units. Total distance traveled by the MSNs is calculated with both normal and extended sensing range. Results are tabulate in Table 1 and compared with the results of [1, 3].

Mathematically, the lower and upper bounds can be given as,

$$\begin{aligned} \mathrm{T\ DT}_N \leq \mathrm{T\ DT} \leq\ \mathrm{T\ DT}_E \\ 9 \times \mathrm{C} \leq \mathrm{TDT} \leq (9+2) \times \mathrm{C} \\ \mathrm{D} \leq \mathrm{TDT} \leq \mathrm{D} + (2 \times r_{\mathrm{s1}} \times M_{\max} \times N_{\max}) \end{aligned} \tag{5}$$

where $C = (r_{\mathrm{s1}} \times M_{\max} \times N_{\max})$.

ROI consists of $M_{\max} \times N_{\max}$ number of tri-hexagons. Simulation is run for two deferent events. In first event, the whole ROI is considered to have no objects to be sensed. Hence, the MTA is run with normal sensing range r_{s1}, and the ROI is 1-covered. The values of TDT_N are tabulated in Table 1 which are considered as the lower bounds of total distance traveled with respect to the concerned sensing range.

The value of TDT_N is 220,050 units when the value of r_{s1} is 50 units. At this value of r_{S1}, the total distance traveled by the MSNs is reduced by 70.07 % compared with value 735,372 of Khan et al. [1].

In second event, the whole ROI is considered to have very large number of objects to be sensed. So, the MTA is run with normal sensing range $\mathrm{r_{s1}}$ and extended sensing range r_{s2}. This makes MSNs to travel an extra distance of $2 \times r_{\mathrm{s1}} \times M_{\max} \times N_{\max}$. The entire ROI is 2-covered. The values of TDT_E which are tabulated in Table 1 are the maximum values of TDT_E which are taken as the upper bounds of total distance traveled with respect to the concerned normal sensing range.

The value of TDT_E is 269,050 units when the value of r_{S1} is 50 units. At this value of r_{S1}, the total distance traveled by the MSNs is reduced by 56.91 % compared with value 624,499 of Purohit et al. [3].

Table 1 Total distance traveled by MSNs

$\mathrm{r_{s1}}$	Triangulation based coverage [1]	$\mathrm{T\ DT_N}$ LB-present study	% Reduction in $\mathrm{T\ DT_N}$	Hexagonal Coverge [2]	$\mathrm{T\ DT_E}$ -UB-present study	% Reduction in $\mathrm{T\ DT_E}$
45	810678	236520	70.82	690678	289170	58.13
50	735372	220050	70.07	624499	269050	56.91
55	656266	205425	68.89	569479	251185	55.89
60	610687	193860	68.25	536880	237060	55.84
65	567533	173160	64.48	465345	211770	54.49

LB lower bound, *UB* upper bound

From the experimental results, it can be observed that the total distance traveled by the MSNs is reduced by 68 % in 1-coverage and 55 % in 2-coverage. Moreover, the coverage is dynamic, and coverage can be configured dynamically.

8 Conclusion

Rectangular ROI divided into tri-hexagons in tessellation fashion. MSNs with variable sensing range are deployed to scan the ROI. MSNs are traversed along the vertices of tri-hexagons. Whenever an object is sensed within sensing range of a MSN, one MSN among the other two MSNs will extend their sensing range to provide high degree of coverage. This degree of coverage can be increased whenever the situation demands. This approach is compared with similar approaches and shows reduction in more than 55 % in total distance traveled by the MSNs with 1 and 2 coverage.

References

1. Khan A, Qiao C, Tripathi SK (2007) Mobile traversal schemes based on triangulation coverage. Mob Netw Appl 12(5):422–437
2. D'Souza RJ, Ganala S, Jose J (2012) Optimal 1-coverage by homogenous mobile sensor nodes using tri-hexagonal scheme. In: IEEE international conference on computer communication and informatics (ICCCI), Jan 2012, vol 3, pp 554–558
3. Purohit GN, Verma S, Sharma M (2010) Hexagonal coverage by mobile sensor nodes. Int J Comput Netw Sec 2(4):41–44
4. D'Souza RJ, Ganala S (2012) Optimal 2-coverage by heterogeneous mobile sensor nodes using triangular scheme. Elsevier Procedia Technol 4(0):187–195. 2nd international conference on computer, communication, control and information technology(C3IT-2012) on 25–26 Feb 2012
5. Zhou Z, Das SR, Gupta H (2005) Fault tolerant connected sensor cover with variable sensing and transmission ranges. In: IEEE SECON, IEEE Press, pp 594–604
6. Wang J, Medidi S (2007) Energy efficient coverage with variable sensing Radii in wireless sensor networks. In: Proceedings of the third IEEE international conference on wireless and mobile computing, networking and communications. WIMOB'07, Washington, DC, USA. IEEE Computer Society, p 61
7. Zhou F (2011) Energy-efficient coverage using sensors with continuously adjustable sensing ranges. In: 2011 seventh international conference on natural computation, pp 109–113
8. Wang C, Wang B, Xu H, Liu W (2012) Energy-efficient Barrier Coverage in WSNs with adjustable sensing ranges. In: IEEE 75th vehicular technology conference (VTC Spring), 2012, May 2012, pp 1–5

About the Editors

Dr. V. Sridhar obtained his B.E. from University of Mysore in Electronics and Communication Engineering, M.E. (Electronics and Telecommunication Engineering) from Jadavpur University, Calcutta and Ph.D from Indian Institute of Technology, New Delhi. He carried out his Post Doctoral Research at Dept. of Electrical and Electronics, University Tenaga Nasional, Malaysia. Dr. Sridhar has been working as a Principal of PES College of Engineering for last 4 years and has been associated with this Institute since 1980, from beginning of his career. His fields of interest include VLSI Design, Digital Signal Processing, Bio-medical Engineering & Instrumentation and Cognitive studies, Computer and Mobile communications. He was also the Registrar (Evaluation) at Visweswaraya Technological University, Belgaum. He has presented and published many papers in National and International peer-reviewed journals and conferences. He has also been a guiding force for M.Sc and Ph.D students. He is instrumental in establishing a CAD Center for VLSI Design as Center of Excellence at the Institution. He is a life member of IEEE (M), IETE (F), ISTE (M), and BMSI (M) and has participated in various national and international conferences.

Dr. H. S. Sheshadri completed his B.E. in Electronics and Communication Engineering from University of Mysore, M.E. in Applied Electronics from Bharathiar University, Coimbatore and Ph.D from Anna University, Chennai. Dr. Sheshadri is currently working as a Professor with Department of Electronics and Communication Engineering in PES College of Engineering, Mandya and has been serving this institution since 1982. His areas of interest are Medical Image Processing and Embedded Systems. He has published many papers in National and International Journals and Conference Proceedings. He is also guiding research candidates for Ph.D programme. He has conducted several short term courses and conferences. He is a life member of IEI, IETE, ISTE, SSI.

Dr. M. C. Padma obtained her B.E. from University of Mysore in Computer Science and Engineering, M.Sc Tech by Research from University of Mysore and Ph.D from VTU, Belgaum. She has been associated with PES College of Engineering since 1990 and is presently working as Professor and Head of Department of Computer Science and Engineering. Her areas of interest are Image Processing, Pattern Recognition, Natural Language Processing, Machine

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0, © Springer India 2014

Translation, Artificial Intelligence and Expert Systems and Document Image Analysis and Recognition. She has published many papers in National and International journals. She is also guiding research candidates for Ph.D programme. She is a life member of ISTE, AKECTA, IVPR, and CSI.

Author Index

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0, © Springer India 2014

Subject index

V. Sridhar et al. (eds.), *Emerging Research in Electronics, Computer Science and Technology*, Lecture Notes in Electrical Engineering 248,
DOI: 10.1007/978-81-322-1157-0, © Springer India 2014