Summer 6-17-2012

Proceeding of International Conference on Advances in Electronics & Communication and Computer Science Engineering AECCSE-2012

Prof. (Dr.) Srikanta Patnaik

Editorial

If we make a review of the 21st century generation, their planning and activities, their interest and involvement driven by some electronic device coupled with an advanced technical functioning. There is a spectacular focus on Information science and Communication technology as it drives the present socio-technical system of the global society. ICT appears as an effective tool for empowering all the civic and anti civic systems of the world. ICT is emerging as an investment area in the millennium development goals of various organizations like UNO, WTO, IBRD and other international apex bodies. It has become an integrated discipline in all the academic spheres as it increases global adaptability

Let me highlight some of the recent developments in Electronics discipline. The new integrated devices did not find a ready market. Users were concerned because the individual transistors, resistors, and other electronic circuit components could not be tested individually to ensure their reliability. Also, early integrated circuits were expensive, and they impinged on the turf that traditionally belonged to the circuit designers at the customer's company. Again, Bob Noyce made a seminal contribution. He offered to sell the complete circuits for less than the customer could purchase individual components to build them. (It was also significantly less than it was costing us to build them!) This step opened the market and helped develop the manufacturing volumes necessary to reduce manufacturing costs to competitive levels. To this day the cost reductions resulting from economies of scale and newer high-density technology are passed on to the user—often before they are actually realized by the circuit manufacturer. As a result, we all know that the high-performance electronic gadget of today will be replaced with one of higher performance and lower cost tomorrow.

The integrated circuit completely changed the economics of electronics. Initially we looked forward to the time when an individual transistor might sell for a dollar. Today that dollar can buy tens of millions of transistors as part of a complex circuit. This cost reduction has made the technology ubiquitous—nearly any application that processes information today can be done most economically electronically. No other technology that I can identify has undergone such a dramatic decrease in cost, let alone the improved performance that comes from making things smaller and smaller. The technology has advanced so fast that I am amazed we can design and manufacture the products in common use today. It is a classic case of lifting ourselves up by our bootstraps—only with today's increasingly powerful computers can we design tomorrow's chips.

The mushrooming growth of the IT industry in the 21st century determines the pace of research and innovation across the globe. In a similar fashion Computer Science has acquired a path breaking trend by making a swift in a number of cross functional disciplines like Bio-Science, Health Science, Performance Engineering, Applied Behavioral Science, and Intelligence. It seems like the quest of Homo Sapience Community to integrate this world with a vision of Exchange of Knowledge and Culture is coming at the end. Apparently the quotation “Shrunken Earth, Shrinking Humanity” holds true as the connectivity and the flux of information remains on a simple command over an internet protocol address. Still there remains a substantial relativity in both the disciplines which underscores further extension of existing literature to augment the socio-economic relevancy of these two fields of study. The IT tycoon Microsoft addressing at the annual Worldwide Partner Conference in Los Angeles introduced Cloud ERP (Enterprise Resource Planning,) and updated CRM (Customer Relationship Management) software which emphasizes the ongoing research on capacity building of the Internal Business Process. It is worth mentioning here that Hewlett-Packard has been with flying colors with 4G touch pad removing comfort ability barriers with 2G and 3G. If we progress, the discussion will never limit because advancement is seamlessly flowing at the most efficient and state-of-the art universities and research labs like Laboratory for Advanced Systems Research, University of California. Unquestionably apex bodies like UNO, WTO and IBRD include these two disciplines in their millennium development agenda, realizing the aftermath of the various application projects like VSAT, POLNET, EDUSAT and many more. 'IT' has magnified the influence of knowledge management and congruently responding to social and industrial revolution.

It's my pleasure to welcome all the participants, delegates and organizer to this international conference. In the process of organizing this conference ASTAR family members have shown their commitment and dedication. I sincerely thank all the authors for their invaluable contribution to this conference. I am indebted towards the reviewers and Board of Editors for their generous gifts of time, energy and effort.

Editor-in-Chief

Prof. (Dr.) Srikanta Patnaik
Chairman IIMT
Intersceince Campus, Bhubaneswar
Email: srikantapatnaik@hotmail.com

An Improved Face Recognition using Dimensionality Reduction Technique

Shaik Rahamtula & K. Veera Swamy
QIS College of Engineering and Technology, Ongole, India.

Abstract - Face recognition plays an important role in many applications such as voter databases, biometric security and many other important activities. The objective of this paper is to improve the face recognition rate. In this work PCA is considered for face recognition. Dimensionality reduction technique performs satisfactorily when test image to be recognized is captured under conditions similar to those of the training images of Yale face database. In this method, overall mean for the entire database is calculated. Mean value is subtracted for all the images. Covariance matrix is computed for all images. Eigenvectors are calculated for each image. Important Eigen values are used to retrieve the relevant image. Minkowski distance is considered to retrieve the images from the database. The experimental results on the several images indicate that the proposed algorithm is superior to an existing algorithm. As compare to the existing method proposed method performs better results.

Keywords*-component: Principle Component Analysis (PCA); Peak Signal to Noise Ratio (PSNR); Minkowski distance; Linear Discriminant Analysis (LDA); Mean Square Error (MSE).*

I. INTRODUCTION

Facial recognition is useful in the areas like Driver's licenses, Pass ports, Access control, Airport Security, Defense, and Army etc. Face recognition is an important part of today's emerging biometrics market. Significant literature is available in face recognition. So many types of commercial systems are available currently and many research organizations are working on the development of more reliable, accurate, and efficient systems.

During last two decades many approaches have been proposed to accomplish the task of face recognition. Later Kohonen demonstrated that a simple neural net could perform face recognition for aligned and normalized face images [1]. One of the Common methods to recognize faces are to use the Eigen face method is popular. Humans can recognize face even when the matching image is distorted, such as a person wearing glasses [7].

Face recognition can be accomplished in three steps. They are Face Detection, Feature Extraction and Face Recognition. The face Detection is to detect the face from the normalized image. The feature extraction is to extract the feature from the detected face. The face recognition is to recognition the face compared with face data base.

It is difficult to directly deal with raw data whenever the amount of data is hug. Dimension reduction is the best method to solve the above problem. This method extracts the structured data and removes the redundant data. Whenever the training images are increased then the matrix size will increase. This problem is called as "curse of dimensionality". Which is solved by dimensionality reduction techniques[8].

Two types of dimensionality reduction techniques; they are linear dimensionality reduction and non-linear dimensionality reduction. The linear dimensionality reduction techniques are Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Locality Preserving Projections (LPP) and so on. The non-linear dimensionality reduction techniques are Isometric Mapping (ISOMAP), Locally Linear Embedding (LLE) and so on.

The aim of this paper is to give the improved face recognition using linear dimensionality reduction technique (PCA). Dimension reduction is presented in section II. Face recognition and PCA are presented in section III. Proposed algorithm is presented in section IV. Experimental results are discussed in section V. conclusions are presented in section VI.

II. DIMENSIONALITY REDUCTION

The basic flow of dimension reduction in face recognition is as shown in the figure1

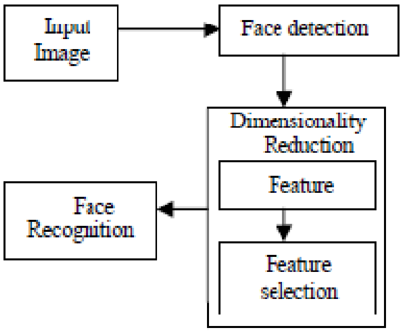


Figure 1: Basic flow of dimensionality reduction technique

Face recognition commonly includes three parts: feature extraction [11], feature reduction and classification. Face extraction is to find the most representative description of the faces, making them can be most easily distinguished from others. Face reduction is to not only decompose and compress the original features but also not destroy the most important information. Classification is to choose the available measure method, which is used to classify the feature data.

III. FACE RECOGNITION

Face recognition has received substantial attention from researches in biometrics, pattern recognition field and computer vision communities. Face recognition can be applied in Security measure at Air ports, Passport verification, Criminals list verification in police department, Visa processing , Verification of Electoral identification and Card Security measure at ATM's**.**

Figure 2 illustrates the face recognition system with different parts.

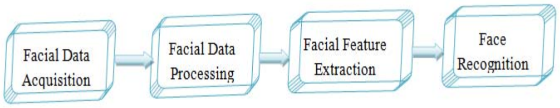


Figure 2 Face Recognition System

Face extraction is to find the most representative description of the faces, making them can be most easily distinguished from others. Face reduction is to not only decompose and compress the original features but also not destroy the most important information. Classification is to choose the available measure method, which is used to classify the feature data. Because the face image is often with a high dimension, it is difficult to use the original data directly, so it is crucial to choose the effectively distinguished features for extraction and reduction. In all kinds of the algorithms of face recognition, PCA [1-6] method is just the effective feature extraction method based on face global feature. PCA is considered as one of the most successful linearity analysis algorithm. It can reduce the dimension effectively and hold the primary information at the same time. But it is only considered the differences between individuals in the traditional PCA method, not taking the differences between classes into account, Therefore the differences of the images in the same person are also increasing when the differences of all images are increasing. It is a disillusionary defect of the PCA algorithm, revolutionized the way of functioning of the society from the earlier satellite transmission, radio and television broadcasting to the current ubiquitous mobile telephones.

A) Standard Face Databases

While there are many databases in use currently, the choice of an appropriate database to be used should be made based on the task given [2]. Some face data sets that are commonly used: Color FERET Database, Yale Face Database, PIE Database, FIA video Database, CBCL Face recognition Database, Expression Image Database, Indian Face Database, Face Recognition Data, and University of Essex, UK.

B) Different Approaches for Face Recognition

There are several methods for different types of task i.e. geometric approach, elastic matching, neuron nets and vide based face recognition.

C) principal component analysis (pca)

The Principal Component Analysis (PCA) is one of the most successful techniques that have been used in image recognition and compression. PCA is a statistical method under the broad title of factor analysis. The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically [10]. This is the case when there is a strong correlation between observed variables. The jobs which PCA can do are prediction, redundancy removal, feature extraction, data compression, etc.

Because PCA is a classical technique which can do something in the linear domain, applications having linear models are suitable, such as signal processing, image processing, system and control theory, communications, etc.

PCA is a technique that effectively and efficiently represents pictures of faces into its Eigen face components. It reduces data dimensionality by performing a covariance analysis between factors. When applied on conditions, PCA will explore correlations between samples or conditions [4]. If we consider an image as a point in a very high dimensional space, these principal components are essentially the eigenvectors of the covariance matrix of this set of face images, which Turk and pent land termed the Eigen face. Each individual face can then be represented exactly by a linear combination of Eigen faces, or approximately, by a subset of "best" eigenfaces [9] - those that account for the most variance within the face database characterized by its eigenvalues, as depicted in Figure 3.

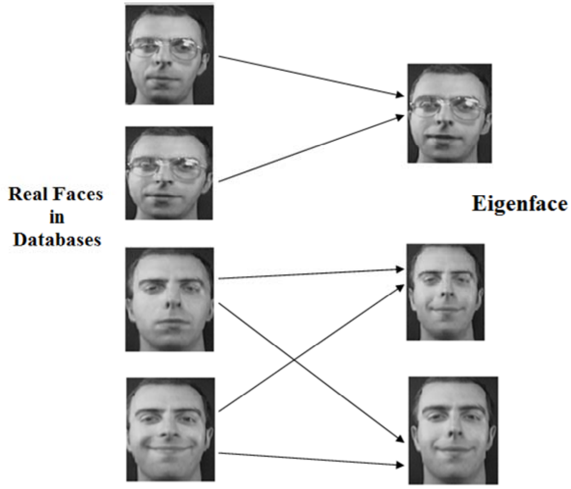


Figure 3 Faces are Linear Combinations of Eigen faces

A) Peak Signal to Noise Ratio:

The peak signal-to-noise ratio (PSNR) is needed to be calculated against the no of subtracted eigenvectors. The PSNR computes the peak signal-to-noise ratio, in decibels, between two images. The higher the PSNR value, the better the quality of the compressed or reconstructed image [5]. To compute the PSNR, the mean-squared error (MSE) [6] is calculated using the following equation:

$$MSE = \frac{\sum_{M,N}[I1(m,n) - I2\ (m,n)]^2}{M\ X\ N}$$

Here *I* is the main image and h is the test image. The dimension of the input image is (M N).

Then the PSNR value is calculated using the following equation:

$$PSNR = 10log_{10}\left(\frac{R^2}{MSE}\right)$$

Here R is the maximum fluctuation in the input image data type.

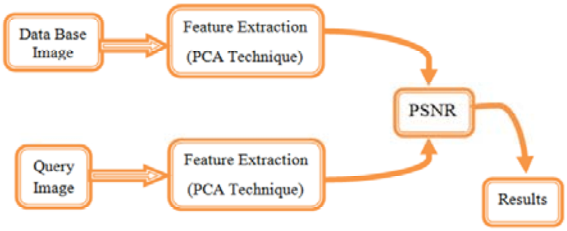


Figure 4 Face Recognition Using PSNR Architecture

Figure 4 illustrates the face recognition using PSNR - peak signal to noise ratio. Here find out the face recognition image due to PSNR method. Database images are extracted using PCA technique and same processor is applied to the query image. The results of two images are compared with PSNR. Finally we get the result in the result section.

B) Minkowski Distance:

Minkowski distance is one of the distance measure techniques. Which can be used to measure the distance between two points concentrated on Euclidean [6] space, which can be considered as a generalization of both Euclidean and Manhattan distance for getting more recognition efficiency.

Let us consider the order of minkowski distance is ‘p’.

The minkowski distance between two points P and Q are define as

$$d(P,Q) = \ L_p(P,Q) = (\sum_{i=1}^{n} |x_i - y_i|^p\ )^{1/p}$$

Where P = (x_1, x_2, …., x_n) and Q = (y_1, y_2, …., y_n)

The value of p is 1 or 2 for minkowski distance. In the limiting case of p reaching infinity we obtain the chebyshev distance.

Figure 5 illustrates the face recognition using Minkowski distance. Here find out the face recognition image due to minkowski method. Database images are extracted using PCA technique and same processor is applied to the query image. The results of two images are compared with minkowski distance. Finally we get the result in the result section.

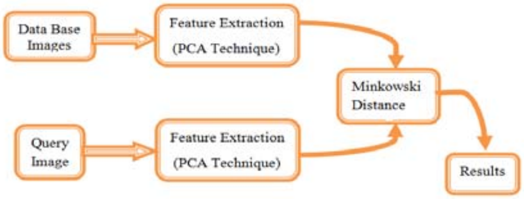

Data Base Images
Feature Extraction (PCA Technique)
Minkowski Distance
Query Image
Feature Extraction (PCA Technique)
Results

Figure 5 Face Recognition Using Minkowski Architecture

Minkowski distance is often used when variables are measured on ration scales with absolute zero value. Variables with a wider range can overpower the result.

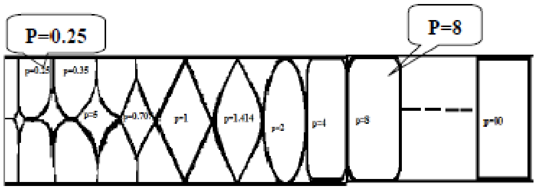


Figure 6 Minkowski Distance

IV. PROPOSED ALGORITHM

Proposed algorithm consists of the following steps.

1. Consider Yale face data base.
2. Resize the images into (N×N) pixel size to reduce the memory requirements of the overall application.
3. Reshape the images into (1×N^2).
4. Each of the training images is stored in a vector of size p.

$$x^i = [\,x_1^i, x_2^i, \dots \dots \dots \dots \dots x_p^i\,]^T$$

Where i = 1, 2, - - - - - - - - - - , n.

Where n is the number of training images.

The training images are then combined into a data matrix of size p×n.

X = [$x^1, x^2, \dots \dots \dots \dots \dots \dots \dots, x^n$]

5. Overall mean of the database is calculated.

$$\mu = [\,\mu_1, \mu_2, \dots \dots \dots \dots \dots \dots, \mu_p]^T$$

Where $\mu_j = \frac{1}{n} \sum_{i=1}^{n} x_j^i$

Where j = 1, 2……………., p.

6. Mean centered images are generated by subtracting the mean image from each of the training images.

Let $\emptyset_i$ be a mean centered image, then

$$\emptyset_i = x^i - \mu$$

Once the training images are centered, they are combined into a centered data matrix of size p × n.

$$\emptyset = [\,\emptyset_1, \emptyset_2, \dots \dots \dots \dots \dots \dots \dots, \emptyset_n\,]$$

7. Compute the covariance matrix (C). It is the product of data matrix and its transpose.

$$C = \emptyset\, \emptyset^T$$

8. Calculate the Eigen values of the covariance matrix in descending order.

$$C : \lambda_1 > \lambda_2 > ------- > \lambda_p$$

9. Compute the Eigen vectors corresponding to Eigen values of covariance matrix C.

$$v_1, v_2, \dots \dots \dots \dots \dots \dots, v_p$$

10. Finally project the training images. Each of the centered training images $\emptyset_i$ is projected onto the eigenspace by selected Eigen vectors (V).

$$L = V^T \emptyset_i$$

11. Repeat the same procedure for query image.
12. By using different distance measure techniques the relevant images from the database are retrieved by considering shortest distance.

V. EXPERIMENTAL RESULTS

Recognition performance in terms of average recognition rate and recognition time of the proposed face recognition system is tested by conducting experiments on Yale data base. The Yale database can be downloadable from the AT & T Laboratories. It contains ten different images of each of 40 distinct subjects. Some images contain facial expressions like open/close eyes, smiling /not smiling, glasses/ no glasses. In this paper first step is resize the each image of database. Find the size of the image. Find the mean value and subtract the mean from original image. Compute the covariance. Collect the eigenvectors and eigenvalues of the covariance matrix. Take largest eigenvectors corresponding to eigenvalues. Based on this concept we mentioned the feature as shown in below figure.

A) Feature selection

The use of co-variance method is to highlight the original local feature depends up on the Eigenvector with respect to eigenvalue. In this paper selected four features for each database image using dimensionality reduction technique. Apply the same process to query image.

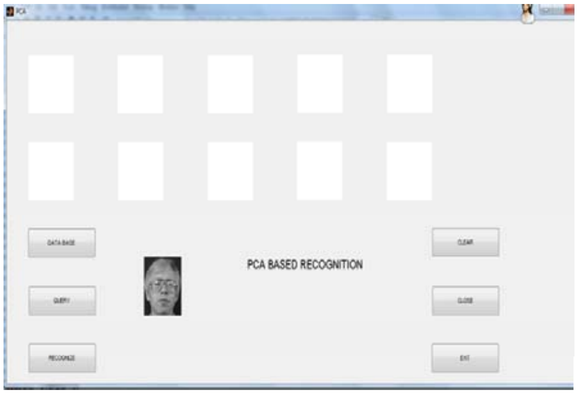


Figure 7 Sample image from face database.

Finally this database is collecting the global features based on the local feature of the technique such as minkowski distance of the database images.

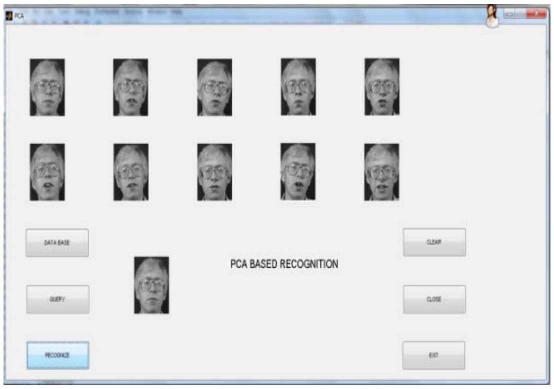


Figure 8 Recognized images

A) Average Recognized Rate

The average recognized rate for the query is measured by counting the number of images from the same category which are found in the top ‘N’ matches.

Table 1: The Results of Recognition Rate

	Number of top matches				
Methods	1	3	5	7	10
PSNR	100	90.5	81	68	55
Minkowski Distance	100	95	88	84	74

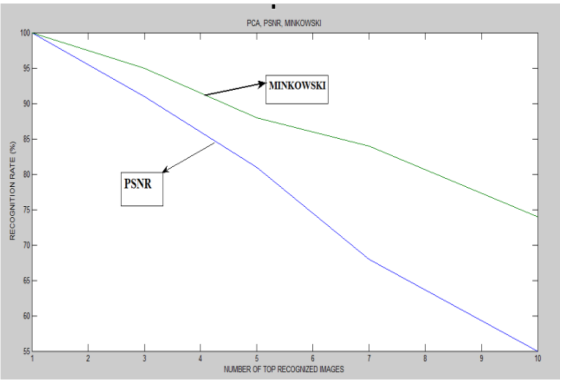


Figure 9 Comparative recognition rates

Face recognition using PSNR and Minkowski distance measures is presented. Time taken for generating the features for the entire database is 53.98 seconds for approach1. Similarly database time for approach2 is 52.84 seconds. Recognition time for PSNR and Minkowski are 0.46, 0.44seconds respectively for approach1. Recognized time for PSNR and Minkowski are 0.41, 0.39 seconds respectively for approach2.

CONCLUSION

An improved face recognition using dimensionality reduction technique is presented. Face recognition using dimensionality reduction technique is relatively simple and robust enough. Comparision between peak signal noise ration (PSNR) and minkowski distanse has been done. The recognized rate of PSNR is 100% recognition for single image and 55% for top ten images. Minkowski distance is 100% recognition for single image and 74% for top ten images.

REFERRENCE

[1] Tahia Fahrin Karim, Molla ShahadatHossain Lipu, Face Recognition using PCA- Based Method. IEEE-2010.

[2] H. Moon and P. J. Phillips, “Computational and performance aspects of PCA-based face-recognition algorithms”, Perception, 2001, Volume 30, pages 303-321.

[3] Maxim A. Grodin, "Elastic graph matching (EGM) ", On internal representation in face recognition systems, Pattern recognition, Volume 33, Issue 7, July 2000, Page 1161-1177.

[4] Cooley, W. W. and Lohnes, "PCA correlation", P. R. Multivariate Data Analysis John Wiley & Sons, Inc., New York, 1971).

[5] Thomson, N., Boulgouris, N. V., & Strintzis, M.G. (2006, January) peak signal-to-noise ratio (PSNR).

Optimized Transmission of JPEG2000 Streams Over Wireless Channels. IEEE Transactions on Image Processing, 15 (I).

[6] George Casella & E.L. Lehmann, Mean-squared error (MSE), "Theory of Point Estimation". Springer, (1999).

[7] Rama Challappa Charles L. Wilson, and Saad Sirohey, "Human and machine recognition of faces: A Survey", proceedings of the IEEE, 1995.

[8] S.K.Sandhu, Sumith Budhiraja, "Combination of Non-linear Dimensionality Reduction Techniques for Face Recognition System", published in IJERA.

[9] Turk, M. and A. Pentland, eigenfaces for Recognition. *Journal of Cognitive neuroscience*, 3(1), 1991, 71-86.

[10] Y.V.Lata,Chandra Kiran Bharadwaj Tungathurthi,H.RamMohanRao,Dr.AGovardhan,Dr.L.P.Reddy,Facial Recognition using eigenfaces by PCA, *International journal of Recent Trends in Engineering*,Vol.1,No.1,may2009

[11] Xudong Jiang, B. Mandal and Alex Kot, Eigenfeature Regularization and Extraction in Face Recognition, *IEEE Trans. On PAMI 30(3), 2008.*

◈◈◈

The Wireless Communication Bus Systems Using Zigbee Technology

B.Venkateswrlu[1], **Ch.Amarnatha Sarma**[2]

[1]Department of Electronics & Communication Engineering, QIS College of Engineering & Technology, JNTUK
[2] Assistant professor, ECE Department, QIS College of Engineering & Technology, JNTUK

Abstract - This paper describes a research on The wireless communication Bus Systems using Zigbee as a communication medium. The Wireless Communication Bus System is a demand responsive transit (DRT) but it is more efficient and convenient in a sense that it entertains passenger's demands and gives bus locations in real time. The real time synchronization of The Wireless Communication Bus System makes it information rich and unique as compared to other DRTs. The Wireless Communication Bus Systems is a system that can replace the Traditional Bus Systems with its flexibility and efficiency. This paper discusses the use of wireless technologies in The Wireless Communication Bus Systems and how to make it more reliable using short range wireless technology Zigbee.

Keywords— *ITS, Zigbee, Mobile communication, DRT, RFID.*

I. INTRODUCTION

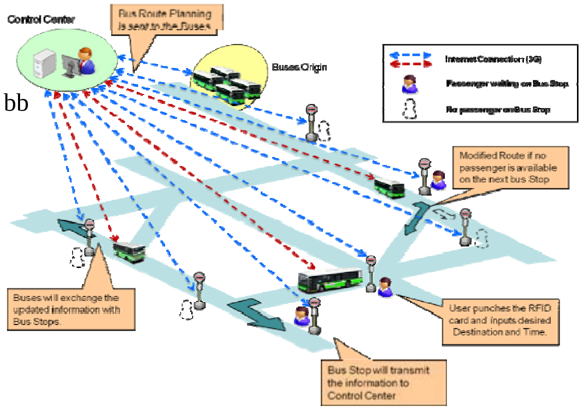
The main objective of this paper is to do a research on the use of short range wireless technology called "Zigbee" in Demand Responsive Transit (DRT), making it much more efficient, reliable and *less expensive*. This research is not the only way to develop this kind of a system and by no means suggested the best solution but it can definitely be one of the better alternatives we have till date and can be used in the areas where there are no 3G, WiMax or other long range wireless technologies available. This research will also help us understand the potential of Zigbee. Till now Zigbee is being used as in-house or in-vehicle technology but this research brings an idea of using Zigbee as communication tool for Inter-Vehicle and Vehicle to Infrastructure. Using Zigbee to communicate between Bus and the Bus Stop will also reduce the total cost of the system as Zigbee devices are far cheaper than WiFi, 3G and WiMax devices. Due to the fact that Zigbee is low power as compared to other short range wireless technologies like WiFi, this system can be deployed in mountainous areas where power is a major concern.

The algorithm of Flexible Bus System is devised in a way that this system replaces the scheduled bus lines systems and buses can dynamically change their routes according to passenger's demands. Passengers are informed about the real time location of the buses which makes it easy for the passengers to decide whether to ride a particular bus or not making this system passenger friendly.

II. SYSTEM DESCRIPTION

Most of the researches today on DRT are carried out in a way that communication medium is long range Wireless technologies like 3G (cellular data) as shown in the Fig. 1.

bb

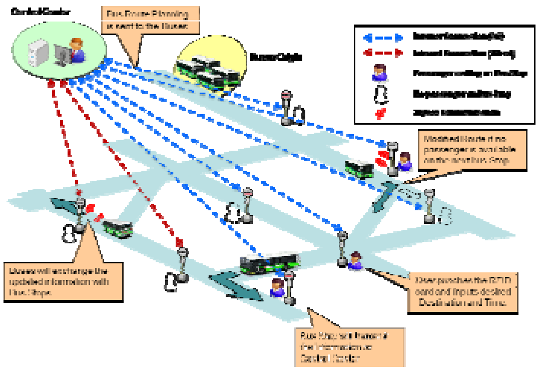

Control Center
Bus Route Planning is sent to the Buses
Buses Origin
Internet Connection (3G)
Passenger waiting on Bus Stop
No passenger on Bus Stop
Modified Route if no passenger is available on the next bus Stop
Buses will exchange the updated information with Bus Stops.
User punches the RFID card and inputs desired Destination and Time.
Bus Stop will transmit the information to Control Center

Everything in the System is connected to the Control Centre and all the information is shared with the Control Centre. Communication between Bus and Control Centre is through 3G; similarly communication between Control Centre and Bus Stop is through 3G. Installing the 3G modules on Bus Stop and Bus which are more than 1 in most of the cases will greatly increase the total installation cost of the system. Following are some drawbacks to this approach.

1. Installation cost is very high.

2. Maintenance cost is very high.
3. High power consumption.
4. Cannot be applied in areas where there are no long range wireless (3G, WiMax) signals available (rural or mountainous areas).

To overcome these problems we propose a model which is less costly than the system discussed above. We call our proposed system “The Flexible Bus System”. This paper will only include the wireless communication part of this system. Control Centre is connected to Bus Stops through internet (Wireless or Wired). Fig. 2 shows our proposed system


Control Center
Bus Route Planning is sent to the Buses
Buses Origin
Modified Route if no passenger is available on the next bus Stop
Buses will exchange the updated information with Bus Stops
User punches the RFID card and inputs desired Destination and Time
Bus Stop will transmit the information to Central Center

As shown in Fig.2 Control Centre and Bus Stops are connected to each other through the internet (wired or wireless). Bus Stop and Bus are connected through short range Wireless Technology “Zigbee”. An interesting feature here is that Bus and Control Centre are not connected to each other directly, they communicate to each other through Bus Stop, means the information transfer from Control Centre is first transferred to Bus Stop and then from Bus Stop to Bus vice versa. This way there will be no need for 3G module installation in Buses. This will greatly reduce the total cost of the system.

The Bus Stops in this system are very smart and we call them “Intelligent Bus Stops”. These Bus Stops are equipped with different devices like RFID card reader; Touch Screen etc. The passengers carry RFID card which contains all the information about the passenger. Whenever a passenger goes to the Bus Stop, he/she will punch his/her RFID card to those RFID card readers installed at Bus Stop and gets recognized by the system. After getting recognized by the system, the passenger can enter the destination and will get the response from the system e.g. on which Bus to ride and how much is the wait time, leaving the decision of riding the Bus to the passenger.

III. PROPOSED ALGORITHM

The algorithm for this system is devised in such a way that passengers have to wait less on the bus stops and buses drive to the bus stops where passengers are waiting instead of driving to the Bus Stops where there are no passengers.

Fig. 3 shows the algorithm flow for The Flexible Bus Systems. Passenger after reaching the Bus Stop punches the RFID Card and all the information (Passenger ID, Destination etc) is transferred to Control Centre which then sends the info on which Bus to ride. Similarly info for all the Buses are transferred to Control Centre through Bus Stops and then Route info for the Buses are transferred to Buses through the Bus Stops. Navigation is installed in the bus which guides the Buses about the routes.

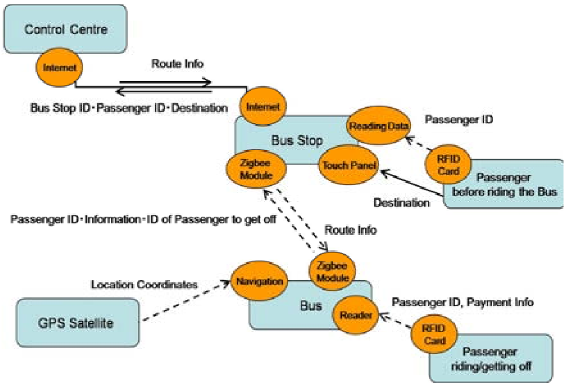


Fig.3 Algorithm flow for The Flexible Bus Systems

Fig. 4 shows that a Traditional Bus System has a fixed route, it drives from 1~25 and then back from 25~1. The route will remain the same even if there are no passengers on Bus Stop3 and Bus Stop4. However in The Flexible Bus System the buses can change the routes dynamically depending upon the demand of the passenger. As shown in Fig. 4 every Bus Stop is connected to the Bus Stop next to it. For example Bus Stop1 is connected to Bus Stop2 and Bus Stop10. For example if BusA is on Bus Stop2 and a passenger is waiting on Bus Stop9, then instead of going to the Bus Stop9 by driving to all the Bus Stops from 2~9, the Bus will directly drive to Bus Stop9 to pick up the passenger. This way the wait time of the passenger waiting on Bus Stop9 will be less as compared to the case in which Bus has to drive from 2~9. The Flexible Bus System will not only decrease the wait time of the passengers but it will also decrease the drive time of the Buses which will greatly reduce the total cost of the System.

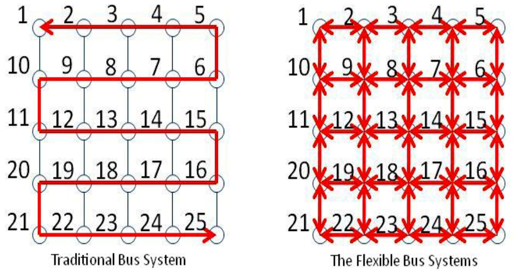


Fig.4 Traditional Bus System and the Flexible Bus System Grid Model

A smart algorithm is required to efficiently perform the tasks discussed in Fig. 4. We are working on an algorithm that can perform these jobs by keeping the wait time of the passengers on the Bus Stops and Ride time of the passengers on the Bus at a low side. Different simulation are been carried out for this purpose using our Simulator called "Konno Simulator".

Fig.5 shows the comparison of Wait time of passengers on the Bus Stop in Traditional Bus Systems and The Flexible Bus Systems.

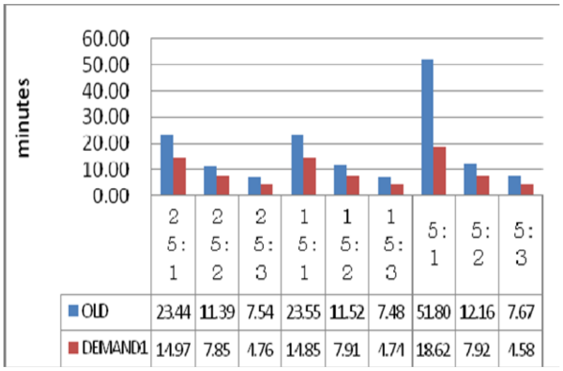


Fig.5 Average Wait Time for Passengers on Bus Stop

In Fig. 5 OLD is "Traditional Bus Systems" and DEMAND1 is "The Flexible Bus Systems". Ratios (25:1, 15:2, 5:3 etc) shown in Fig.5 are No. of Passenger: No. of Buses.

Similarly the ride time of the passengers in the Bus is also Very important our simulation results show that using our Algorithm ride time of passengers in the buses is also less than that of ride time of the passengers in the bus inTraditional Bus Systems.

IV. ZIGBEE COMMUNICATION

Zigbee Communication between Bus and Bus Stops are the major concern of this paper. The reason Zigbee is given priority over the other short range wireless technologies like WiFi is that in this research only few bytes of data is to be transferred using short range wireless technologies, no heavy data like Audio or Video is transferred so Zigbee seems to be a good alternative to WiFi which offers heavy data transfer. With heavy data comes more power consumption which can be a concern in rural areas where power is not easily available. As described earlier Buses and Bus Stops are going to communicate with each other through Zigbee. Since Control Centre is aware of all the ongoing and carrying all the information about all the buses and bus stops, the information is first transferred to corresponding bus stop about the upcoming bus. Bus Stops and Control Centre are always connected to each other through the internet (Wired or Wireless) so information between Bus Stops and Control Centre can be easily shared.

Major concern is the transfer of information between the Bus Stops and Buses and that is to be done through Zigbee. This is the core of this research and so the information transfer between the Buses and the Bus Stops are carefully thought of. Fig. 6 shows the sequence of information transfer between Buses and the Bus Stops through Zigbee

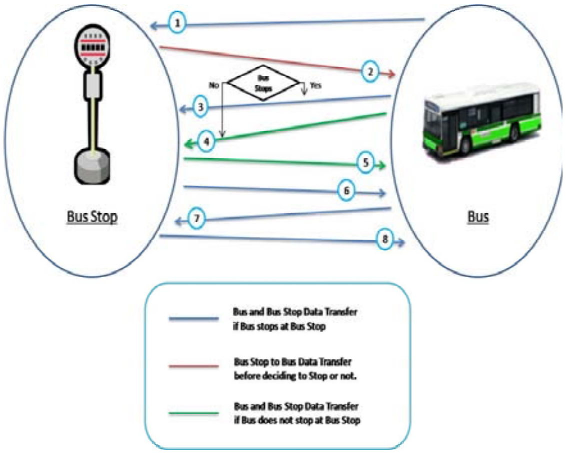

1
2
No
Bus Stops
Yes
3
4
5
6
7
8
Bus Stop
Bus
Bus and Bus Stop Data Transfer if Bus stops at Bus Stop
Bus Stop to Bus Data Transfer before deciding to Stop or not.
Bus and Bus Stop Data Transfer if Bus does not stop at Bus Stop

Fig. 6 Data Transfer Sequence between Buses and Bus Stops

As soon as Bus enters the communication range of the Bus Stop, information exchange starts. The numbers 1~8 in Fig. 6 are described below.

1. Bus sends the Bus ID to Bus Stop.
2. Bus Stop authenticates the request and requests the Bus to Stop or Don't Stop at current Bus Stop.
3. If Bus is requested to stop at current Bus Stop,

Bus sends the stopping signal to Bus Stop.

4. If Bus is requested to pass the Bus Stop without stopping, it sends the Non-Stop passing signal to Bus Stop and requests Next Bus Stop ID
5. Bus Stop sends the Next Bus Stop ID to the Bus.
6. In case of Bus stopped at current Bus Stop, after the Passengers rode the Bus, Bus Stop requests for the No. of Passengers in the Bus.
7. Bus sends the No. of passengers in the Bus to the Bus Stop and requests for Next Bus Stop ID.
8. Bus Stop sends the Next Bus Stop ID to the Bus.

The decision of stopping at Bus Stop is made by Control Centre and is to be made on following conditions

- No. of passengers to get off on this Bus Stop
- No. of passengers to ride on this Bus
- Seats available in this Bus
- Destination of the passengers

Each time information/request is sent by the Bus to the Bus Stop, the information is sent to the Control Centre by the Bus Stop. The Control Centre checks for the validity of the data and sends it back to the Bus Stop from where the data is sent to the Bus. Fig. 7 is a flow chart depiction of Fig. 6 showing an easy to understand flow of information transfer between the Bus and the Bus stop.

All the information received from Bus (No. of passengers in the Bus) which is now carried by the Bus Stop will be transferred to the Control Centre through the internet after the Bus leaves the Bus Stop.

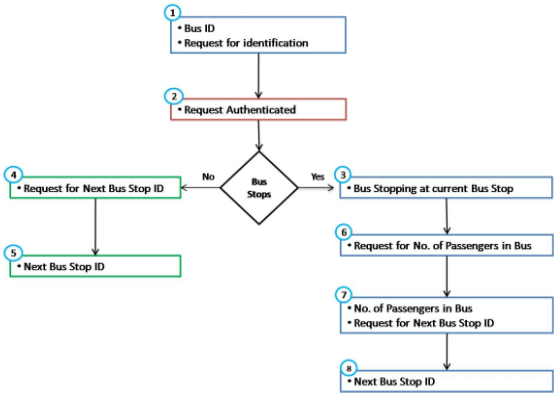

1
• Bus ID
• Request for identification
2
• Request Authenticated
Bus Stops
No
Yes
4
• Request for Next Bus Stop ID
5
• Next Bus Stop ID
3
• Bus Stopping at current Bus Stop
6
• Request for No. of Passengers in Bus
7
• No. of Passengers in Bus
• Request for Next Bus Stop ID
8
• Next Bus Stop ID

Fig.7 Data Transfer Flow Chart between Buses and Bus Stops

V. CONDUCTED EXPERIMENTS

Different experiments are performed to check whether Zigbee is suited for this kind of System (Intelligent Bus Systems) or not. Below are the details of experiments. All the experiments are conducted using Max Stream Xbee Pro.

Below are the conditions in which experiments are conducted.

- Maximum distance between Xbee Bus and Xbee Bus stop is 100m with maximum height of 2m.
- Maximum No. of Bytes sent is 7 Bytes.
- Experiments are conducted using the Data Transfer Sequence shown in Fig. 6 and Fig. 7.

Three types of experiments are conducted to check the eligibility of Zigbee for The Flexible Bus Systems.

1. Zigbee Communication on straight road with clear Line of Sight.

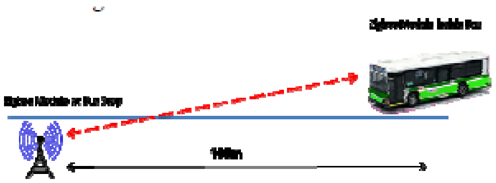


Fig.8 Zigbee Communicate with Clear Line of Sight

2. Zigbee Communication on curve with No Line of Sight (Trees as Hurdles)

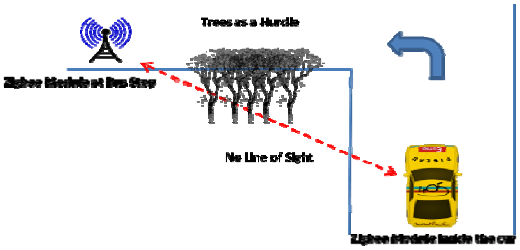


Fig.9 Zigbee Communicate with No Clear Line of Sight (trees as hurdle)

3. Zigbee Communication on curve with No Line of Sight (Buildings as Hurdles.

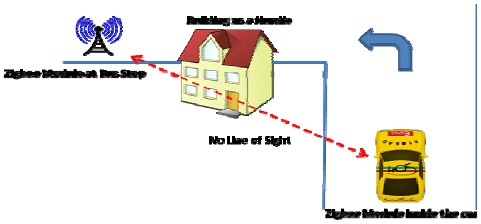


Fig.10 Zigbee Communicate with No Clear Line of Sight (Buildings as hurdle)

Table I shows the data collected for Experiment 1 when there is clear line of sight and the road is straight.

TABLE I

ZIGBEE COMMUNICATION ON STRAIGHT ROAD (LOS)

Distance (m)	No. of Times Sent	No. of Times received	Sent every	Error Percentage
100m	6	6	1 sec	0.00%
100m	6	6	1 sec	0.00%
100m	6	6	1 sec	0.00%

As shown in Table I, there are no errors when Zigbee Data transfer is checked on the straight road with clear line of sight which shows that Zigbee can be used on straight roads with clear line of sight without any doubts.

Table II shows the data collected for Experiment 2 when there is no clear line of sight and there is a curve/turn on road which has some rough trees around.

TABLE II

ZIGBEE COMMUNICATION ON STRAIGHT ROAD (NO LOS WITH TREES)

Distance (m)	No. of Times Sent	No. of Times received	Sent every	Error Percentage
100m	6	6	1 sec	0.00 %
100m	6	5	1 sec	16.67%
100m	6	5	1 sec	16.67%
100m	6	3	1 sec	50.00%
100m	6	5	1 sec	16.67%
100m	6	4	1 sec	33.33%
Average	**6**	**4.67**	**1 sec**	**22.22%**

Table II shows the data which is still quite encouraging although there are a few errors, but the average shows that every 6 times data is sent 4.67 times it is received. As described earlier if the data is received even only once the communication can be done easily and bus can get to know if it has to stop on the upcoming Bus Stop or not but as data is received more than 1 times, in this case 4.67 times which is far more times it should be received, so there shouldn't be any problem in data communication.

Table III shows the data collected for Experiment 3 when there is no clear line of sight and there is a curve/turn on road which has some buildings around.

TABLE III

ZIGBEE COMMUNICATION ON STRAIGHT ROAD (NO LOS WITH BUILDINGS

Distance (m)	No. of Times Sent	No. of Time	Sent every	Error Percentage
100m	6	3	1 sec	50.00 %
100m	6	2	1 sec	66.67%
100m	6	1	1 sec	83.33%
Average	**6**	**2**	**1 sec**	**66.67%**
80m	6	4	1 sec	33.33%
80m	6	2	1 sec	66.67%
80m	6	3	1 sec	50.00%
80m	6	5	1 sec	16.67%
80m	6	4	1 sec	33.33%
Average	**6**	**3.6**	**1 sec**	**40.00%**
50m	6	4	1 sec	33.33%
50m	6	5	1 sec	16.67%
50m	6	6	1 sec	0.00%
50m	6	3	1 sec	50.00%
50m	6	6	1 sec	0.00%
Average	**6**	**4.8**	**1 sec**	**20.00%**

The data shown in Table III is not very encouraging as error percentage is at high side. Table shows that at distance of 100m error rate is 66.67% which means that data is received only 2 times out of 6 which is at a risky side. Similarly at a distance of 80m on average data is received 3.6 times out of 6 which is better than that at 100m but still below par. At a distance of 50m the data is encouraging getting received 4.8 times out of 6. If we concentrate on data for 50m, it seems as if data communication can be done easily as data is received quite frequently but the problem with 50m is that if Bus is sent a signal by the Bus Stop to stop at this Bus Stop when Bus is just 50m away from the Bus Stop, it will be difficult for the bus to stop with a speed of 60km/h which is the maximum speed limit here in Japan.

A. Experiments outcome

The conducted experiments showed the following results

Different experiments showed that the main factor affecting Zigbee communication is the distance between the Zigbee modules (Xbee Bus and Xbee Bus Stop in this case). The more the distance the less efficient is the Zigbee communication.

- The bigger and denser the hurdles, more difficult it is to communicate between the Zigbee Modules.
- The straighter the roads, the better it is to communicate between Zigbee Modules.

- A suitable height of about 2m can increase the efficiency of Zigbee Data Transfer.
- Speed of the car has no significant effect on Zigbee Data Transfer

B. *Solutions to the problems*

Below are some proposed solutions to the problems occurred during the experiments.

- The main problem is less frequency of receiving of Data when there are hurdles like Buildings on road with curves or turns. The proposed solution to this problem is to avoid making the Bus Stops where there are buildings on the roads with curve or turns. Another solution can be to use Relays with Zigbee to increase the range of Zigbee Modules but this will increase the cost of the System.

VI. SUMMARY

The Flexible Bus Systems is an efficient and a smart Demand Responsive Transit (DRT). It is flexible in a sense that it can change dynamically according to the demand of the passenger. The system can fulfil the demands of the passengers in a way that they have to wait less on the Bus Stops and even if they miss the Bus they can be entertained by the next bus without waiting for very long. The use of Zigbee for communication between the Buses and the Bus Stops greatly reduce the total cost of the system. Everything is connected to the Control Centre which is the brain of the system. Control Centre and Bus Stops are connected through the internet and Buses and Control Centre are connected to each other through Bus Stops. All the characters (Buses, Bus Stops and Passengers) are updated with latest information all the times which makes The Flexible Bus Systems more information rich and reliable.

REFERENCES

[1] Shahin Farahani, Zigbee Wireless Networks and Transceivers

[2] Lee, E., Ryu, K., Paik, I.: A Concept for Ubiquitous Transportation Systems and Related Development Methodology. In: International IEEE Conference on Intelligent Transportation Systems, pp.37-42 (2008)

[3] Yuwei Li, Jessica Wang, Justin Chen, Michael Cassidy, Design of a Demand-Responsive Transit System (California PATH Working Paper UCB-ITS-PWP-2007-4)

[4] ZHANG Feizhou, CAO Xuejun, YANG Dongkai, Intelligent Scheduling of Public Traffic Vehicles Based on a Hybrid Genetic Algorithm (TSINGHUA SCIENCE AND TECHNOLOGY, ISSN 1007-0214 09/25 pp625-631, Volume 13, Number 5, October 2008)

[5] Jin Xu, Zhe Huang, An Intelligent Model for Urban Demand- responsive Transport System Control (Journal of Software, Vol. 4, No. 7, September 2009)

[6] R´emy Chevrier, Philippe Canalda, Pascal Chatonnay and Didier Josselin, Comparison of three Algorithms for solving the Convergent Demand Responsive Transportation Problem (Proceedings of the IEEE ITSC 2006)

[7] Zigbee Standards Organization, Zigbee Specification, ZigBee Document 053474r17, January 17, 2008.

[8] Tomohiro Utsumi, Shin Konno, Yusuke Kanno, Ken-ichi Yukimatsu, Mahito Kobayashi, Masashi Hashimoto, A Feasibility study on the Buffer-Less Routing Networks using Deflection Routing Control (Proceedings of IEICE Vol. J92-B No. 11 pp. 1741-1749 2009).

◈◈◈

FPGA Implementation of a DA based 1D DCT Processor

Sirisha.surampalli, P. Kalyan Chakravarti & # Govinda Rao Locharla

Dept. of ECE, Sri Sivani College Of Engineering Chilakapalem, India
# GMR Institute of Technology, Rajam, India

Abstract - In this paper a DA (Distributed Arithmetic) algorithm based architecture for DCT calculation is proposed and results are compared with the MATLAB results. Distributed Arithmetic (DA) is an important technique to implement digital signal processing DSP functions in FPGA. It is a powerful technique for reducing the size of a parallel hardware multiply-accumulate that is well suited to FPGA designs. XILINX's Spartan 3E FPGA is targeted for this implementation. XILINX ISE Foundation (10.1i) software is used.

Keywords*- DA, FPGA, DCT, XILINX 10.1i*

I. INTRODUCTION

Video and Audio data streams require a huge bandwidth to be transferred in an uncompressed form. Image compression is very important for image storage and image transmission. The goal of image compression is to reduce the amount of data required to represent a digital image by removing redundant data. Over the past three decades, many image and video coding methods have been developed for the compression task. JPEG (Joint Photographic Experts Group), H.261/H.263, MPEG1/2/4 (Motion Pictures Expert Group) and H.264. These coding standards cover almost all applications we can enumerate from our daily life, such as digital cameras, video phones and videoconferencing, storage-based applications (e.g., VCD and DVD), Internet media, digital TV broadcasting, and HDTV.

II. REVIEW OF DA ALGORITHM

Distributed Arithmetic (DA) is a different approach for implementing DSP circuits. The basic idea is to replace all multiplications and additions by a look up table, shifter-and accumulator. In DSP applications like convolution i.e., h(n)*x(n), if the coefficients are known beforehand then partial products of the h (n)*x (n) can be calculated for all possible x(n)'s. These partial products are stored in a look up table (LUT) and then accessed by the input samples each bits of which are used as the address for ROM. This process is graphically narrated in Fig.1. Mathematical representation of the convolution process is given in eq (1) & eq(2):

$$y(n)=\sum_{k=0}^{n} h(k) * x(n-k) \qquad \ldots (1)$$

$$\begin{aligned}
y(n)=\{\ & h(0)\bullet x_{n,0}+h(1)\bullet x_{n-1,0} \\
&+\ \ldots +h(n)\bullet x_{n-k,0}\ \} \qquad + \\
&\{h(0)\bullet x_{n,1}+h(1)\bullet x_{n-1,1} \\
&+\ \ldots\ +h(n)\bullet x_{n-k,1}\}\bullet 2^{-1}+ \\
&\{h(0)\bullet x_{n,2}+h(1)\bullet x_{n-1,2}+\ldots \\
&+h(n)\bullet x_{n-k,2}\}\bullet 2^{-2} \qquad + \\
&. \\
&. \qquad + \\
&. \\
&\{h(0)\bullet x_{n,7}+h(1)\bullet x_{n-1,7}+\ldots \\
&+h(n)\bullet x_{n-k,7}\}\bullet 2^{-7}
\end{aligned} \qquad \ldots(2)$$

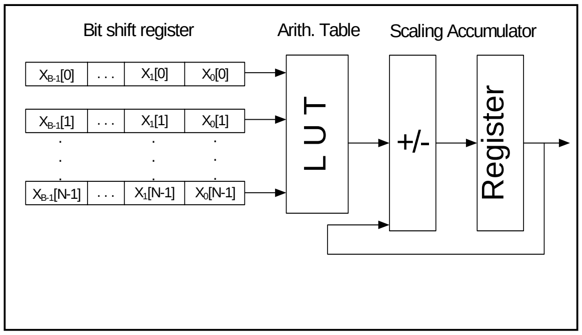

Bit shift register
Arith. Table
Scaling Accumulator
X_{B-1}[0] . . . X1[0] X0[0]
XB-1[1] . . . X1[1] X0[1]
XB-1[N-1] . . . X1[N-1] X0[N-1]
L U T
+/-
Register

Fig.1 Implementation of sum of products

The distributed arithmetic requires 2^N word LUT which is pre-programmed to the target device. These partial products are accessed using input vector x_b= ($x_b(0)$, $x_b(1)$,…..$x_b(N-1)$) are weighed by the power of 2 and accumulated. The inner product of y is computed after N lookup

III. DA BASED DA ARCHITECTURE

Implementation computation is reduced by decomposing (1) in two 8x1 1-D DCT given by,

$$F(u)=1/2*c(u)*\sum_{i=0}^{7} X(i)*\cos(\frac{(2i+1)u\pi}{16}) \qquad ...(3)$$

where 0 ≤ u≤ 7 and 0 ≤ v ≤ 7 and c(u),c(v) = 1/ 2 for u,v=0,c(u),c(v) =1 otherwise.

Eq.3 can be expanded as below:

$F(0) =[X(0) + X(1) + X(2) + X(3) + X(4) + X(5) + X(6) + X(7)]T$

$F(1) =[X(0)- X(7)]A+[X(1)- X(6)]B+[X(2)-X(5)]C+[X(3)-X(4)]D$

$F(2) =[X(0)- X(3)- X(4)+ X(7)]R+[X(1)-X(2)-X(5)+ X(6)]S$

$F(3)=[X(0)-X(7)]B+[X(1)-X(6)](-D)+[X(2)-X(5)](-A)+[X(3)-X(4)](-C)$

$F(4) =[X(0) - X(1) - X(2) + X(3) + X(4) - X(5) - X(6) + X(7)]T$

$F(5) =[X(0)- X(7)]C+[X(1)- X(6)](-A)+[X(2)- X(5)]D+[X(3)- X(4)]B$

$F(6) =[X(0)- X(3)- X(4)+ X(7)]S+[X(1)- X(2)- X(5)+ X(6)](-R)$

$F(7) =[X(0)- X(7)]D+[X(1)-X(6)](-C)+[X(2)-X(5)]B+[X(3)-X(4)](-A)$

where,

R = ½* COS(π/8); S = ½* COS(3π/8); T = ½* COS(π/4);

A = ½* COS(π/16); B = ½* COS(3π/16); C = ½* COS(5π/4); D = ½* COS(7π/4);

By using the sum values inside the brackets as address lines and storing the possible pre computed values in the look up tables. Pin layout of the DA based DCT processor is illustrated in Fig.4. Comparison of FPGA simulation results and MATLAB results are given in table.1

Table.I : Results Comparision

X(n)	89 115 140 166 192 217 243 64
$DCT(X(n))_{MATLAB}$	433.4565 -64.8799 -94.5063 67.8618 -72.1249 51.3948 -39.6870 18.7338
$DCT(X(n))_{FPGA}$	433 -64.8 -94.499 67.81 -72.12 51.39 -39.680 18.733

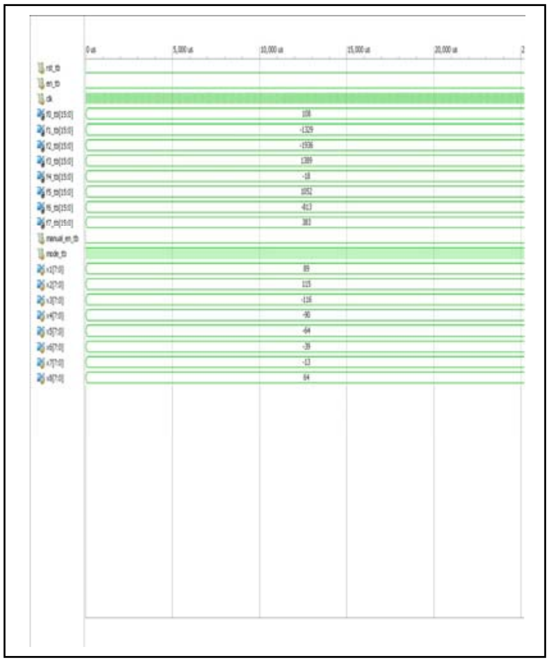
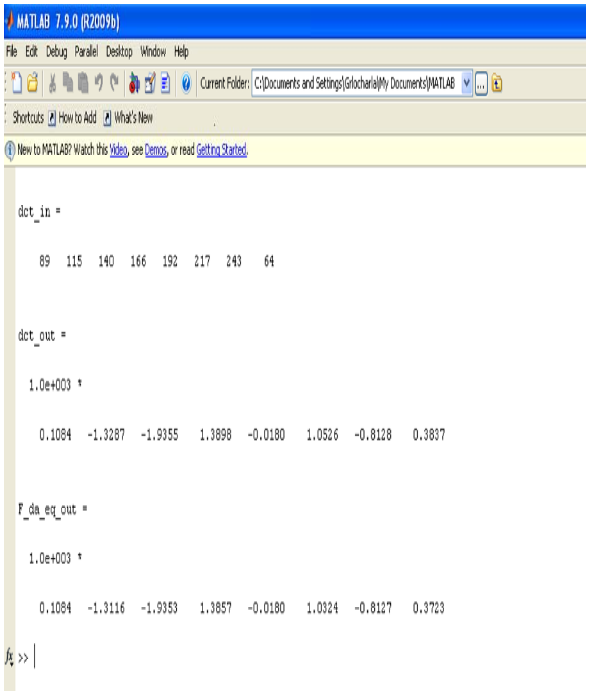


Fig.2. FPGA simulation results (scaled values)


MATLAB 7.9.0 (R2009b)
File Edit Debug Parallel Desktop Window Help
Current Folder: C:\Documents and Settings\Grlocharla\My Documents\MATLAB
Shortcuts How to Add What's New
New to MATLAB? Watch this Video, see Demos, or read Getting Started.

```
dct_in =

   89   115   140   166   192   217   243    64


dct_out =

  1.0e+003 *

   0.1084  -1.3287  -1.9355   1.3898  -0.0180   1.0526  -0.8128   0.3837


F_da_eq_out =

  1.0e+003 *

   0.1084  -1.3116  -1.9353   1.3857  -0.0180   1.0324  -0.8127   0.3723

fx >>
```

Fig.3. Comparision of the FPGA simulation with MATLAB results (scaled values)

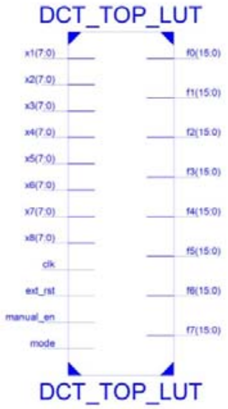


Fig.4 Pin layout of the DA based DCT processor

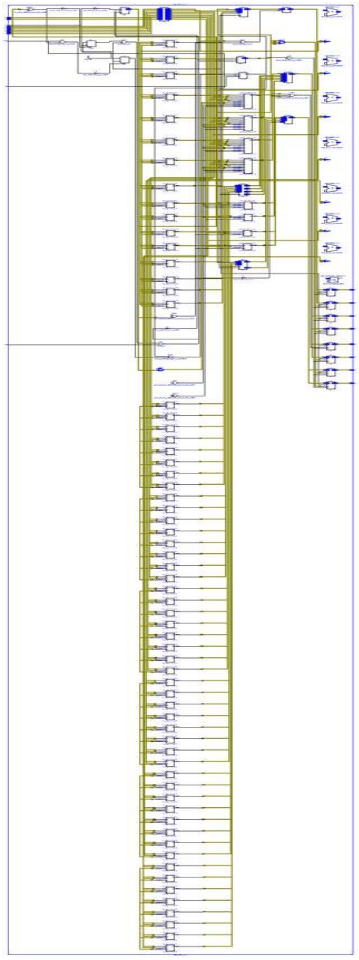


Fig.5 RTL Schematic of the DA based DCT processor

TABLE. II DEVICE UTILIZATION SUMMERY

ARCITECTURE FOR SIN^{-1}X/COS^{-1}X EVOLUTION		
S.NO.	Logic Utilization	UTILIZED HARDWARE
1.	Number of Slices	600
2.	Number of Slice Flip Flops	328
3.	Number of 4 input LUTs	1000
4.	Number of bonded IOBs	196

IV. CONCLUSION

An area efficient architecture for the computation of 1-D DCT using DA is proposed and implemented in both Xilinx FPGA and MATLAB, results are compared. With proposed 1-D DCT architecture, 2-D DCT can be implemented using row column decomposition technique. and area and power results are tabulated.

REFERENCES

[1] R. C. Gonzalez, R. E. Woods, “Digital Image Processing,” 2nd.Ed.,Prentice Hall, 2002.

[2] F.H.P. Fitzek, M. Reisslein, “MPEG-4 and H.263 Video Traces for Network Performance Evaluation ,” *IEEE Network*, vol.15, no.6, pp.40-54, Nov/Dec 2001.

[3] Luciano Volcan Agostini, Ivan Saraiva Silva and Sergio Bampi, “Multiplierless and fully pipelined JPEG compression soft IP targeting FPGAs,” *Microprocessors and Microsystems*, vol. 31(8), 3 pp.487-497, Dec. 2007.

[4] S. A. White, “Applications of distributed arithmetic to digital signal processing: a tutorial review,” *IEEE ASSP Magazine*, vol.6, no.3, pp.4-19, Jul.1989.

[5] M.-T. Sun, T.-C. Chen, A.M. Gottlieb, ‘‘VLSI Implementation of a 16x16Discrete Cosine Transform,” *IEEE Transactions on Circuits and Systems,* vol.36, no. 4, pp. 610 – 617, Apr. 1989.

[6] A. Shams, A. Chidanandan, W. Pan, and M. Bayoumi, “NEDA: A low power high throughput DCT architecture,” *IEEE Transactions on Signal Processing,* vol.54(3), Mar. 2006.

[7] Peng Chungan, Cao Xixin, Yu Dunshan, Zhang Xing, “A 250MHz optimized distributed architecture of 2D 8x8 DCT,” 7th International Conference on ASIC, pp. 189 – 192, Oct. 2007.

[8] M. Kovac, N. Ranganathan, “JAGUAR: A Fully Pipelined VLSI Architecture for JPEG Image Compression Standard,” *Proceedings of the IEEE*, vol.83, no.2, pp. 247-258,Feb.1995.

[9] Yuan-Ho Chen, Tsin-Yuan Chang, Chung-Yi Li, “High Throughput DABased DCT With High Accuracy Error-Compensated Adder Tree,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. PP, issue 99, pp. 1-5, Jan 2010.

[10] A. Kassem, M. Hamad, E. Haidamous, “Image Compression on FPGA using DCT," *International Conference on Advances in Computational Tools for Engineering Applications, 2009*, *ACTEA '09*, pp.320-323, 15-17 July 2009.

◈◈◈

Comparison of Various FIR Lowpass Filter Design Techniques With PSO Algorithm

Vijay.P[1] & M. Venkata Sudhakar[2]

SRI MITTAPALLI COLLEGE OF ENGINEERING, Guntur, Andhrapradesh

Abstract - This paper presents an optimization technique for the design of optimal digital FIR low pass filter. The design of digital FIR Filters possible by solving a system of linear equations .In this paper we are using PSO (particle Swarm Optimization).PSO method is used to determine the frequency response of Digital FIR filters, consequently the optimal filter coefficients are obtained with fast convergence speed. PSO technique is purely random algorithm. PSO technique provides optimal filter coefficients such that error function is minimized when compared with the traditional FIR filter design techniques such as windowing techniques. The convergence speed of PSO is faster than the traditional FIR filter design techniques. Particle Swarm Optimization technique which has been applied in many areas such as function optimization, fuzzy system method and other areas.

I. INTRODUCTION

Particle Swarm Optimization is a population based stochastic optimization technique developed by Dr.Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling. In PSO the potential solutions, called particles, fly through the problem space by following the current optimum particles. Each particle keeps track of its Coordinates in the problem space which is associated with best solutions (fitness) it has achieved so far. This value is called pbest. When a particle takes all the population as its topological neighbors. The best value is called towards its gbest location.

II. PROBLEM FORMULATION AND MATHEMATICAL BACKGROUND OF PSO

The frequency response of a linear-phase FIR filter is given by ref(5)

$$H(e^{jw})=\sum_{i=0}^{N} h(n)e^{-jwn} \quad \text{ref(5)} \qquad (1)$$

Where h(n) is the real valued impulse response of filter, N+1 is the length of filter and ω is frequency according to the length being even and odd and the symmetry being an even and odd four types of FIR filters described. The linear phase is possible if the impulse response h(n) is either symmetric(h(n) =h(N-n)) or is anti symmetry (h(n)=-h(N-n)) for $0 \leq n \leq N$ In general the frequency response for type1 FIR filter can be expressed in the form

$$H(e^{j\omega}) = e^{-jn\omega/2}\, \tilde{H} \quad \text{ref(5)} \qquad (2)$$

Where amplitude response $\tilde{H}(\omega)$, also called the zero response, is given by

$$\tilde{H}(\omega) = h(N/2) + \sum_{n=1}^{N/2} h(N/2-n)\cos(\omega n) \qquad (3)$$

the amplitude response for the type 1 linear phase FIR filter (using the notation N=2M) is expressed as

$$\tilde{H} = \sum_{k=0}^{M} a(k)\cos(\omega k) \quad \text{ref(5) (4)}$$

Where a(0)=h(M) and a(k) = 2h(M-k), 1<=k<M

The amplitude response for the type 2 linear phase FIR filter is given by

$$A(\omega) = \sum_{i=1}^{k} \{W(\omega_i)[\sum_{k=0}^{M} a(k)\cos(\omega_i k) - D(\omega_i)]\}\wedge 2 \qquad (5)$$
$$\partial \in / \partial a(k) = 0$$

$$A(\omega) = \sum_{k=1}^{2M+1/2} b(k)\cos[\omega(k-1/2)$$

Where b(k)=2h[2m+1/2-K], 1<K<2M+1/2.

The amplitude response for the case of type 3 linear phase FIR filter is given as

$$A(\omega) = \sum_{k=1}^{M} c(k)\sin(\omega k)$$

Where c (k) = 2h (M-k) 1<=k<=M

The amplitude response for the case of type 4 linear phase FIR filter is given as

$$\overset{\square}{H}(\omega) = \sum_{k=1}^{2M+1/2} d(k)\sin[\omega(k-1/2)$$

Where d(k)=2h[2M+1/2-k]   1<=k<=2M+1/2

The design of a linear phase FIR filter with least mean square error criterion, we find the filter coefficients a(k) such that error is minimized. Corresponding to the coefficients the filter coefficients are obtained as shown by the equations.

The Least mean square design function for this design is given as

$$\text{Є}=\sum_{i=1}^{k} W(\omega_i)[A(\omega_i - D(\omega_i)] \text{ ref(5)}$$

for type 1 FIR filter the amplitude response A(w) is a function of a(k) to arrive at the minimum value of ϵ, we set

$$\partial \in / \partial a(k) = 0 \quad 0<=k<=M$$

Which results in a set of (M+1) linear equations that can be solved for a(k)

For the type 1 the expression for mean square error is ε expressed as

$$\sum_{i=1}^{k}\{W(\omega_i)[\sum_{k=0}^{M} a(k)\cos(\omega_i k) - D(\omega_i)]\}\ \hat{}\ 2 \quad (6)$$

A similar formation can be derived for the other three types of linear phase FIR filters. This Design approach can be used to design a linear phase FIR filter with arbitrarily shaped desired response. Where D(w)is the frequency response.

III. PROPOSED PSO ALGORITHM

Particle Swarm Optimization (PSO) algorithm is a population based optimization algorithm. Its population is called a swarm and e

Step 1:

Error function is to be minimized is expressed in equation

Step 2:

Initial population (swarm) is generated where each particle in the swarm is a solution vector containing M=5 elements, then initial population can be expressed as

$$U_i^0 = [U_{i1}^0, U_{i2}^0, U_{i3}^0, U_{i4}^0, U_{i5}^0]$$

Particle U_{ij}^0 of particle U_j^0 is generated from uniform distribution [0 , $U_{ij,\max}^0$].

Step 3:

Initial velocities of each particle are written as follows

$$V_i^0 = [V_{i1}^0, V_{i2}^0, V_{i3}^0, V_{i4}^0, V_{i5}^0] \text{ i=1, 2 ,.....5}$$

Each elements V_{ij}^0 of V_j^0 is selected a random digits for example, between

[0, 0.1 $x_{i,\max}$]

Step 4:

Set iteration count K=1.

Step 5:

Calculate Error value by using eq (6). $\in = \min[\in_1, \in_{2,} \in_{3,} \in_{4,} \in_5]$

And corresponding particle is the gbest is the particle which leads to Є.

Step 6: Update velocities of each particle using following relation

Where $Pbest_i^k$ is the best previous position of the i th particles, $gbest^{k-1}$ is particle which leads to ε

$$V_i^k = W*V_i^{k-1} + c1*r1(pbest_i^{k-1} - U_i^{k-1})$$

$$+c2*r2(gbest_i^{k-1} - U_i^{k-1})$$

the position among all partilcles.r1 and r2 are random digits between [0,1].c1 and c2are acceleration constants and W an inertia weight typically selected in the range 0.1 to 2.ref(1)

Step 7:

Update position of each individual particle as

Where i=1, 2....M.

Step 8:

Update $Pbest_i^k$ and $Gbest_k$

$Pbest_i^k = U_i^k$

if C(U_i^k)<C($Pbest_i^k$)

$= Pbest_i^k$

if C (U_i^k)>=C($Pbest_i^k$)

Out of all particles which give min error gives the Gbest.

Step 9:

Repeat 5 to 9 for maximum no of times to get the least possible error.

IV. SIMULATION RESULTS

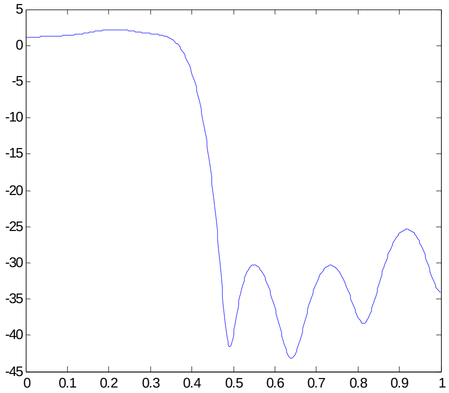


Fig 1. Frequency response of first order low pass FIR digital filter (frequency vs gain(db))

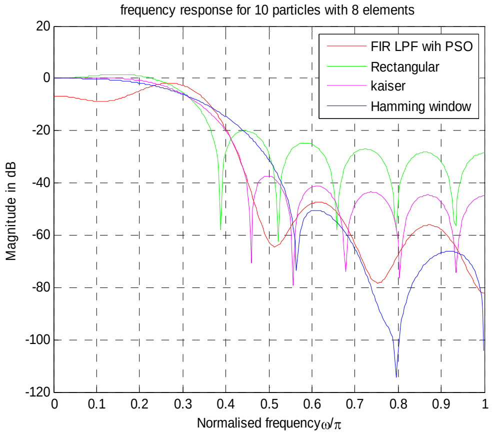


Fig 2. Comparison of Fir low-pass filter response using PSO with Rectangular, Kaiser and Hamming windowing techniques.

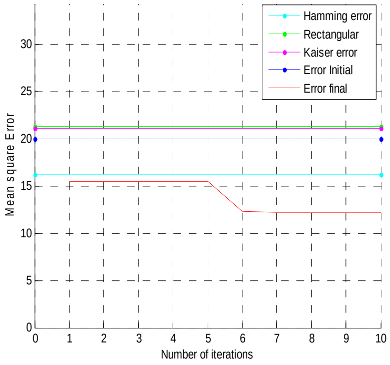


Fig 3. Error plot

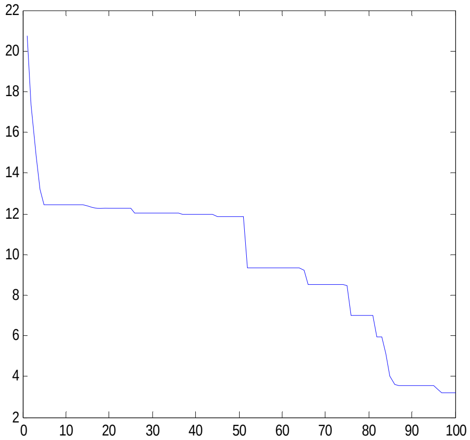


Fig 4. No of Iterations (N=100) VS MSE

V. CONCLUSIONS

The information mechanism in PSO is significantly different. So the whole population moves like a one group towards an optimal area. In PSO, only gbest gives out the information to others. PSO converges to the best solution quickly and gives min error. The frequency response of PSO is better than rectangular, Kaiser and hamming windowing techniques, as the no iterations increases error is reduced. How ever PSO does not have any genetic operators like crossover and mutation. So PSO is better than traditional windowing techniques such as rectangular, Kaiser and hamming windowing techniques.

REFERENCES

[1] Youlian Zhu, Chng hung, Hybrid optimization design of FIR filters, IEEE.int conference neural networks &signal processing zhenjing, June 8~10, 2008.

[2] L.S.Titare and L.D. arya 'A particle swarm optimization for improvement of voltage stability reactive power reserve management' a journal of the institute of engineers (India) Sep 2006

[3] www.swarmintelligenc.org/tutorials.php

[4] Genetic search approach .A new learning algorithm for adaptive IIR filters, IEEE. Int by Leung S.H, Chung.C.Y. Jan 1996.

[5] S.K.Mitra, Digital Signal Processing: A Computer-Based Approach. New York, New York: McGraw-Hill, third ed., 2006.

[6] Adaptive Filter theory by Simon Haykins (fourth edition).

[7] Genetic search approach .A new learning algorithm for adaptive IIR filters, IEEE.int Chung.C.Y., Jan 1996.

[8] N.J.Fliege, Multirate Digital Signal Processing. New York, New York: John Wiley & Sons, 1994.

◈◈◈

The New Adaptive Active Constellation Extension Algorithm For PAR Minimization In OFDM Systems

K. B. Pavan Kumar & R. S. Rao

Sree Vidyanikethan Engineering College

***Abstract* -** In this paper, the Peak-to-Average Ratio reduction in OFDM systems is implemented by the Adaptive Active Constellation Extension (ACE) technique which is more simple and attractive for practical downlink implementation purpose. However, in normal constellation method we cannot achieve the minimum PAR if the target clipping level is much below than the initial optimum value. To get the better of this problem, we proposed Active Constellation Algorithm with adaptive clipping control mechanism to get minimum PAR. Simulation results exhibits that the proposed algorithm reaches the minimum PAR for most severely low clipping signals to get minimum PAR.

Keywords: *Peak-to-Average ratio, OFDM, adaptive, active constellation extension*

I. INTRODUCTION

OFDM is a well known method for transmitting high data rate signals in the frequency selective channels. In OFDM system, a wide frequency selective channel is sub divided in several narrow band frequency nonselective channels and the equalization becomes much simpler [1]. However, one of the major drawbacks of multitone transmission systems, such as OFDM systems has higher PAR compared to that of single carrier transmission system.

To solve this problem, many algorithms have been proposed. In some of the algorithms, modifications are applied at the transmitter to reduce the PAR. In some of the algorithms like Partial Transmit Sequence and selective mapping (SLM) the receiver requires the Side Information (SI) to receive data without any performance degradation. In some other methods the receiver can receive the data without SI; example is clipping and filtering, tone reservation [2] and ACE [3]. In the ACE method, the constellation points are moved such that the PAR is reduced, but the minimum distance between the constellation points remains the same. Thus the BER at the receiver does not increase, but a slight increase in the total average power. To find the proper movement constellation points, an iterative Projection onto Convex Set (POCS) method has been proposed [3] for OFDM systems.

Among various peak-to-average (PAR) reduction methods, the active constellation extension (ACE) technique is more attractive for down-link purpose. The reason is that ACE allow the reduction of high-peak signals by extending some of the modulation constellation points towards the outside of the constellation without any loss in data rate. This favour, however, comes at the cost of slight power penalty. For practical implementation, low complexity ACE algorithms based on clipping were proposed in [3-4]. The elementary principle of clipping based ACE (CB-ACE) involves the switching between the time domain and switching domain [5]. Clipping in the time domain, filtering and employing the ACE constraint in the frequency domain, both require iterative process to control the subsequent regrowth of the peak power. This CB-ACE algorithm provides a suboptimal solution for the given clipping ratio, since the clipping ratio is predetermined at the initial stages. Even this method has the low clipping ratio problem in that it cannot achieve the minimum PAR when the clipping level is set below an unknown optimum level at the initial stages, because many factors, such as the initial PAR and signal constellation, have an impact on the optimal target clipping level determination [3]. To the best of knowledge, a practical CB-ACE algorithm cannot predetermine the optimal target clipping level.

In this method, to solve the low clipping problem, we introduce a new method of ACE for PAR reduction. The approach combines a clipping-based algorithm with an adaptive control, which allows us to find the optimal clipping level. The rest of the paper consist OFDM system with CB-ACE, proposed ACE algorithm. Finally the last section includes the simulation results for M-QAM.

II. PAR PROBLEM WITH CB-ACE

An OFDM, the input bit stream is interleaved and encoded by a channel coder. These coded bits are mapped into complex symbols using QPSK or QAM modulation. The signal consists of the sum of N independent signals modulated in the frequency domain onto sub channels of equal bandwidth. As a continuous-time equivalent signal, the oversampled OFDM signal is expressed as

$$\mathbf{x_n} = \frac{1}{\sqrt{\mathbf{JN}}} \sum_{\mathbf{k=0}}^{\mathbf{JN-1}} \mathbf{X_k} \mathbf{e}^{\mathbf{j2\pi\frac{k}{JN}n}} \qquad (1)$$

where n=0,1,2,…….,JN-1, N is the number of subcarriers; X_k are the complex data symbols at kth subcarrier; J is the oversampling factor where J≥4, which is large enough to accurately approximate the peaks [6]. In matrix notation , (1) can be expressed as **x=Q*X**. where Q* is inverse discrete Fourier transform matrix of size JN × JN, ()* indicate the Hermitian conjugate, the complex time-domain signal vector $x=[x_0,x_1,\ldots,x_{JN-1}]^T$, and the complex symbol vector $X=[X_0,X_1,\ldots,X_{N/2-1},0_{1\times(J-1)N},X_{N/2},X_{N/2+1},\ldots,X_{N-1}]^T$. Here the guard interval is not considered because it does not have any impact on PAR, which is defined as x_n

$$\mathbf{PAR(x)} \triangleq \frac{\mathbf{\max_{0\le n\le JN-1}|x_n|^2}}{\mathbf{E[|x_n|^2]}} \qquad \mathbf{(}2\mathbf{)}$$

Note that (2) does not include the power of the anti-peak signal added by the PAR reduction. Let $\mathcal{L}$ be the index set of all data tones;$\mathcal{L} = \{\forall k\ s.t\ 0 \le k \le N-1\} = \{\mathcal{L}a \cup \mathcal{L}_a^c\}$, where $\mathcal{L}_a$ is the index set of active sub channels for the reducing PAR. The PAR problem in ACE is formulated as [7]

$$\min_{\mathbf{c}} \|\mathbf{x} + \mathbf{Q}^*\mathbf{C}\|_\infty^2$$

Subjected to: X_k+C_k be flexible for k∈ $\mathcal{L}_a$,

C_k=0, for k∉ $\mathcal{L}_a$ (3)

Where C is the extension vector whose components are non zero only if k∈ $\mathcal{L}_a$. However, this optimal solution for this ACE formulation for PAR reduction is not appropriate for practical implementation due to high computational complexity. Thus, the CB-ACE algorithm is introduced [3-4].

The basic idea of the CB-ACE algorithm is to generate the anti-peak signal for PAR reduction by projecting the clipping in-band noise into the feasible extension area while removing the out-of-band distortion with filtering. Thus, the CB-ACE formulation is considered as a repeated-clipping-and-filtering (RCF) process with ACE constraint as follows;

$$\mathbf{x}^{(\boldsymbol{i}+1)} = \mathbf{x}^{(\boldsymbol{i})} + \boldsymbol{\mu}\tilde{\mathbf{c}}^{(\boldsymbol{i})} \qquad (4)$$

Where μ is a positive real step size that determine the convergence speed, i is the iteration index, the initial signal is $x^{(0)}$, and $\tilde{c}^{(i)}$ is the anti-peak at the ith iteration as follows: $\tilde{c}^{(i)} = k^{(i)}c^{(i)}$, where $k^{(i)}$ is the transfer matrix of size JN × JN and $k^{(i)} = \widehat{Q}^{*(i)}\,\widehat{Q}^{(i)}$, and $\widehat{Q}^{(i)}$ is determined by the ACE constraint that $X_k^{(i)} + C_k^{(i)}$ is a feasible for k∈ $\mathcal{L}_a$. Here, $c^{(i)}$ is the peak signal above the predetermined clipping level A and $c^{(i)} = [c_0^{(i)}, c_1^{(i)}, \ldots, c_{JN-1}^{(i)}]^T$, where $c_n^{(i)}$ is the clipping sample, which can be obtained as follows:

$$\mathbf{c_n^{(\boldsymbol{i})}} = \begin{cases} \left(\left|\mathbf{x_n^{(\boldsymbol{i})}}\right| - \boldsymbol{A}\right)\boldsymbol{e^{j\theta_n}}, & \boldsymbol{if}\ \left|\boldsymbol{x_n^{(i)}}\right| > \boldsymbol{A} \\ \mathbf{0}, & \boldsymbol{if}\ \left|\boldsymbol{x_n^{(i)}}\right| > \boldsymbol{A} \end{cases}, \quad (5)$$

Where $\theta_n = \arg(-x_n^{(i)})$. This clipping level A is related to the clipping ratio $\varkappa$ as $\varkappa = \frac{A^2}{E\{|x_n|^2\}}$. In general, we expect more PAR reduction gain with a lower target clipping level. The existing CB-ACE algorithm cannot achieve the minimum PAR for low target clipping ratios, because the reduced power by low clipping reduces the PAR reduction gain. The original constellation move towards the with the decreasing clipping ratio in [8], which places the clipping signal constellation outside the feasible extension area. The number of $\mathcal{L}_a$, corresponding to the number of reserved tones into tone reservation (TR),as in [9]. Decrease in the clipping ratio degrades the PAR reduction capacity in ACE.

III. SUGGESTED ACE ALGORITHM

The main objective of the proposed algorithm is to control both clipping level and convergence factor at each iteration and to iteratively minimize the peak power signal greater than the target clipping level. The basic cost function is defined as

$$\boldsymbol{\xi}(\boldsymbol{I}^{(i)}) \triangleq \mathbf{min}_{\boldsymbol{\mu},\boldsymbol{A}} \left\|\boldsymbol{x}^{(i)} + \boldsymbol{\mu}\tilde{\boldsymbol{c}}^{(i)} - \boldsymbol{A}\boldsymbol{e}^{\boldsymbol{j\Phi}^{(i)}}\right\|_{\mathbf{2}}^{\mathbf{2}} \qquad (6)$$

Where $\Phi^{(i)}$ is the phase vector of $x^{(i)} + \mu\tilde{c}^{(i)}$ at the ith iteration and $I^{(i)}$ represents the set of time indices at the ith iteration, $I^{(i)} = \{\forall n\ s.t\ n \in [0, JN-1]\}$.

The summary the proposed algorithm is,

Step 1: Initialize the parameters

(a) Selecting the target clipping level A

(b) Setting the maximum number of iterations L.

Step 2: Set$i = 0$, $x^{(0)} = x$ and $A^{(0)} = A$.

Step 3: Computing the clipping signal in (5); if there is no clipping signal, transmit signal, $x^{(i)}$.

Step 4: Transfer the clipping signal into anti-peak signal subjected to ACE constraint;

(a) Convert $\mathrm{c}^{(i)}$ into$\mathrm{C}^{(i)}$.

(b) Removing the out-of-band of $\mathrm{C}^{(i)}$ by projecting $\mathrm{C}^{(i)}$ onto the feasible region in ACE.

(c) By taking the IDFT obtain $\tilde{c}^{(i)}$.

Step 5: Update $\mathrm{x}^{(i)}$ in (4) and minimizing (6).

(a) Computing the optimal step size μ,

$$\boldsymbol{\mu} = \frac{\mathfrak{R}[\langle \mathbf{c}^{(i)}, \tilde{\boldsymbol{c}}^{(i)} \rangle]}{\langle \tilde{\boldsymbol{c}}^{(i)}, \tilde{\boldsymbol{c}}^{(i)} \rangle} \qquad (7)$$

Where $\mathfrak{R}$ defines the real part and $\langle , \rangle$ is the complex inner-product.

(b) Adjust the clipping level A

(c)

$$\boldsymbol{A}^{(i+1)} = \boldsymbol{A}^{(i)} + \boldsymbol{\nu}\boldsymbol{\nabla}_{\boldsymbol{A}} \qquad (8)$$

Where the gradient with respect to A is

$$\boldsymbol{\nabla}_{\boldsymbol{A}} = \frac{\sum_{n \in I_1^{(i)} \cup I_3^{(i)}} \left| c_n^{(i)+1} \right|}{N_p} \qquad (9)$$

and ν is the step size with 0≤ ν ≤ 1 and N_p is the number of peak signals larger than A.

Step 6: Increase the iteration counter $\boldsymbol{i} = \boldsymbol{i} + \mathbf{1}$. If $i < L$, go to Step 3 and repeat ; otherwise, transmit signal, $\mathrm{x}^{(i)}$.

Compared to the existing CB-ACE with complexity of order $O\ (JN\log JN)$, the complexity of the proposed algorithm slightly increases whenever the adaptive control is calculated in (8).however this complexity is negligible compared to that of order $O\ (JN\log JN)$.

IV. SIMULATION RESULTS

In this section, we illustrate the performance of our proposed algorithm using computer simulations. In the simulations, we use an OFDM system with 2048 sub carriers and M-QAM constellation on each subcarrier. To approximate the continuous-time peak signal of an OFDM signal, the oversampling rate factor J = 8 is used in (1).

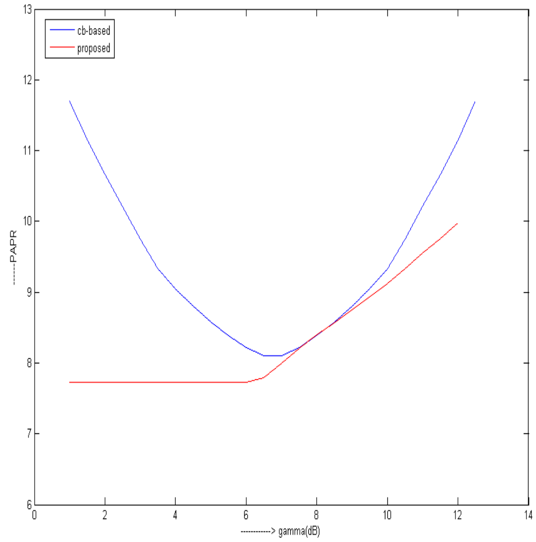


Figure1. The achievable PAR of CB-ACE and the proposed algorithm for an OFDM signal with a 12dB PAR, for different target clipping levels.

Fig. 1 compares the achievable PAR of CB-ACE with the optimal adaptive scaling with that of our proposed algorithm for an OFDM signal with an initial 11.7 dB PAR and 16-QAM modulation, for different target clipping ratios$\varkappa$ from 0dB to 12dB. In the case when CB-ACE is applied, we find the minimum achievable PAR, 7.72Db, is obtained with a target clipping ratio of 6dB, which shows that CB-ACE depends on the target clipping ratio, as we mentioned in the previous section. The PAR reduction gain becomes smaller with a decreasing target clipping ratio from the optimal value of 6dB. Thus, we must carefully select the target clipping ratio for CB-ACE. On the other hand, we observe that our proposed algorithm can achieve the lower minimum PAR even when the initial target clipping ratio is set below the CB-ACE optimal value of

6.4dB, It is obvious that our proposed algorithm solves the low target clipping ratio problem associated with the CB-ACE, as shown in Fig.1.

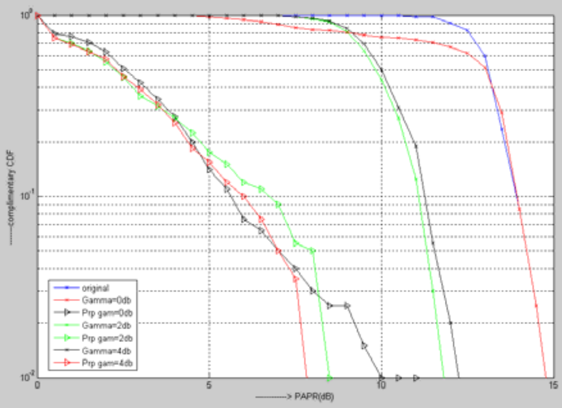


Figure2. PAR CCDF comparison of the CB-ACE and proposed method for different initial target clipping ratios: ϰ=0dB, 2dB, and 4dB, L=10.

Fig,2 considers two algorithms: CB-ACE and our proposed method for three different initial target clipping ratios, ϰ=0dB, 2dB, and 4db, in terms of their complementary cumulative density function (CCDF). The solid line curve at the right is plotted for the original OFDM signal. The marked lines correspond to the PAR reduced signals of CB-ACE and our proposed method after 10-iterations, which we have confirmed is sufficient for convergence. For a 10^{-3} CCDF, CB-ACE with initial target clipping ratios of ϰ=0dB, 2dB, and 4dB can be achieve a 0.14dB,0.89dB, and 2.95dB PAR reduction from the original PAR of 11.7dB,respectively. In other words, when the target clipping ratio is set low, the achievable gain in PAR reduction decreases, which is opposite to out general expectation, but is consistent with the trend shown in Fig.1.On other hand, our proposed algorithm shows about a 4dB reduction gain in PAR at10^{-3} CCDF for all three of the initial low target clipping ratios.

V. CONCLUSION

In this paper, we proposed a new CB-ACE algorithm for PAR reduction using adaptive clipping control. We observed that the existing CB-ACE depends on initial target clipping ratio. The lower the initial target clipping ratio is from the optimal clipping value, the smaller the PAR reduction gain. However, our proposed algorithm provided the minimum PAR even when the initial target clipping ratio was set below the unknown optimum clipping point.

REFERENCES

[1]. Wu, y. and Zou, W.Y. "Orthogonal Frequency Division Multiplexing: a multi-carrier modulation scheme," *IEEE Trans. Consumer Electronics*, 41(3), pp. 392-399 , 1995.

[2]. Krongold, B.S. and Jones, D.L. "An active-set approach for OFDM PAR reduction via tone reservation," *IEEE Trans, Signal Processing*, pp. 495-509, Feb. 2004.

[3]. Krongold, B.S. and Jones, D.L. "PAR reduction in OFDM via active constellation extension," *IEEE trans, Broadcasting*, pp. 258-268 , Sept. 2003.

[4]. L. Wang and C. Tellambura, "An adaptive-scaling algorithm for OFDM PAR reduction using active constellation extension," in *Proc. IEEE Veh. Technology Conf.*, pp. 1-5, Sep. 2006.

[5]. E. Van der Ouderaa, J. Schoukens, and J. Renneboog, "Peak factor minimization using a time-frequency domain swapping algorithm," *IEEE Trans. Instrum. Meas.*, vol. 37, no. 1, pp. 145-147, Mar. 1988.

[6]. C. Tellambura, "Computation of the continuous-time PAR of an OFDM signal with BPSK subcarriers," *IEEE Commun. Lett.*, vol.5, no. 5, pp.185-187, May 2001.

[7]. Y. Kou, W.-S. Lu, and A. Antoniou, "New peak-to-average power-ratio reduction algorithm for multicarrier communication," *IEEE Trans. Circuits and Syst.*, vol. 51 no. 9, pp. 1790-1800, Sep. 2004.

[8]. M. Friese, "On the degradation of OFDM-signal due to peak-clipping in optimally predistorted power amplifiers," in Proc. *IEEE Globecom*, pp. 939-944, Nov. 1998.

[9]. J. Tellado, *Multicarrier Modulation with Low PAR: Applications to DSL and Wireless.* Boston: Kluwer Academic Publishers, 2000.

◈◈◈

A FPGA-Based AES and LDPC code for Wireless Security System

Vinay K G & Usha Rani K R

Dept. of Electronics and Communication, R.V. College of Engineering, Bangalore, India

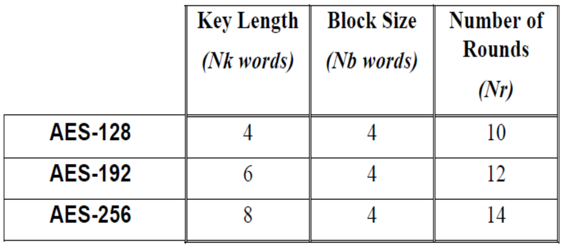
***Abstract* -** Wireless networks are forcing organizations to completely rethink how they secure their networks and devices to prevent attacks and misuse that expose critical assets and confidential data. By their very nature, wireless networks are difficult to roll out, secure and manage, even for the most savvy network administrators. Unlike traditional wired networks in which communications travel along a shielded copper wire pair or optical cable, wireless radio frequency (RF) signals literally traverse through open air. As a result, RF signals are completely exposed to anybody within range and subject to fluctuating environmental factors that can degrade performance and make management complex. As a result, it's very important that enterprises define effective wireless security policies that guard against unauthorized access to important resources and also take care of Environmental interference and physical defects in the communication medium which can cause random bit errors during data transmission. In this paper, a VHDL (Very High Speed Integrated Circuit hardware description language) code is developed for AES128 algorithm (Encryption and Decryption) used for secured data transmission along with LDPC code(Encoder and Decoder) for error detection and correction and is verified using ModelSim 6.3f Simulator for Wireless Security Systems and is implemented on Spartan 3 FPGA.

Key Words: *AES, LDPC Code, FPGA.*

I. INTRODUCTION

Wireless networks present a number of security challenges not only to the authentication but also to data privacy and integrity. Extremely strong authentication and encryption are requirements for any secure wireless network. In today's digital world, encryption is emerging as a disintegrable part of all communication networks and information processing systems, for protecting both stored and transmit data. Encryption is the transformation of plain data (known as plaintext) into unintelligible data (known as cipher text) through an algorithm referred to as cipher. There are numerous conventional encryption algorithms that are now commonly used in computation, but the U.S. government has adopted the Advanced Encryption Standard (AES) to be used by Federal departments and agencies for protecting sensitive information.

The National Institute of Standards and Technology (NIST) has published the specifications of this encryption standard in the Federal Information Processing Standards (FIPS) Publication 197. AES is a block cipher. This means that the number of bytes that it encrypts is fixed. If the bytes being encrypted are larger than the specified block then AES is executed concurrently. . All the AES algorithm operations are performed on a two dimensional 4x4 array of bytes which is called the State [1]. The algorithm is composed of three main parts: Encryption, Decryption and Key Expansion. Encryption and Decryption are composed of specific number of rounds (Table 1). The number of rounds to be performed during the execution of the algorithm is dependent on the key length [2].



	Key Length *(Nk words)*	Block Size *(Nb words)*	Number of Rounds *(Nr)*
AES-128	4	4	10
AES-192	6	4	12
AES-256	8	4	14

Table 1

Environmental interference and physical defects in the communication medium can cause random bit errors during data transmission. Error coding is a method of detecting and correcting the errors. Error coding is required to ensure information is transferred intact from its source to its destination. Different error coding schemes are chosen depending on the types of errors expected, the communication medium's expected error rate, and whether or not data retransmission is possible. Low-density parity-check (LDPC) codes are forward error-correction codes, first proposed in the 1962 PhD thesis of Gallager at MIT. Unlike many other classes of codes LDPC codes are equipped with very fast (probabilistic) encoding and decoding algorithms. They can recover the original code word even in the presence of large amounts of noise. This makes LDPC codes not only attractive from a theoretical point of view, but also perfect for practical applications [3].

II. DESCRIPTION OF AES ALGORITHM

The AES cipher either operates on individual bytes of the State or an entire row/column. The algorithm consists of four different transformations. These are: Sub Byte Transformation, Shift Row Transformation, Mix Column Transformation and Add Round key Transformation.

During Sub Byte Transformation each value of the state is replaced with the corresponding value from the pre-defined look up table called S-Box which is derived from a multiplicative inverse of a finite field.The ShiftRows transformation cyclically shifts the last three rows of the state by different offsets. For Mix Column, the state is arranged into a 4 row table. The multiplication is performed one column at a time (4 bytes). Each value in the column is eventually multiplied against every value of the predefined matrix (16 total multiplications). The results of these multiplications are XORed together to produce only 4 result bytes for the next state. The multiplication is performed one matrix row at a time against each value of a state column. During the AddRoundKey transformation, the key values are bitwise XORed with the State. Each key is the same size as the state [4].

The AES algorithm requires four words of key for each encryption round. The key expansion routine executes a maximum of four consecutive functions which involves shift row transformation, box substitution, function which returns a four byte value based on the predefined matrix and a constant which changes every round. The decryption algorithm involves operations: InvShift-Rows, InvSubByte, and InvMixColumns. These operations are the inverse of the ShiftRows, SubByte and MixColumns operations, respectively [5].

In general the Complete Encryption and Decryption process is depicted in Figure 1 below.

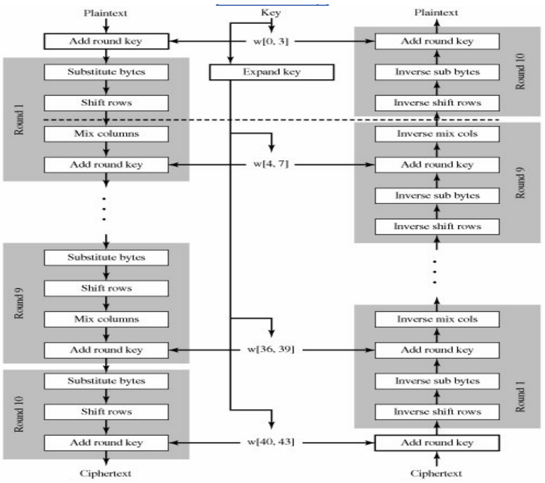


Figure 1: AES Encryption and Decryption Process

LDPC CODE

R. Gallager defined an (N, j, k) LDPC codes as a block code of length N having a small fixed number (j) of ones in each column of the parity check H matrix, and a small fixed number (k) of ones in each rows of H as shown in figure 2. Encoding is done by using equations derived from the H matrix to generate the parity check bits [6]. Decoding is accomplished using "soft-inputs" with these equations to generate new estimates of the sent values. This process is repeated in an iterative manner resulting in a very powerful decoder. There are three distinct types of LDPC codes: Simple codes, Gallager codes and Irregular Repeat Accumulate (IRA) codes.

$$
\begin{array}{cccccccccccccccccccc}
\hline
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
\hline
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
\hline
\end{array}
$$

Figure 2: Example of a low-density code matrix:N=20,j=3,k=4.

A. *LDPC ENCODER AND DECODER MODULE*

A codeword can be formed from a message, s, by the following formula: x = $G^{T}\cdot$ s. For code words of length n, encoding k information bits requires a Generator matrix, G, of size k by n. The generator matrix is of the form:

$$G^{T} = \begin{bmatrix} I_k \\ P^{T} \end{bmatrix}$$

Where I_k denotes a k × k Identity matrix and P^T, a parity sub-matrix of size n-k by k. Encoder in this paper takes 3 binary data bits at a time and converts it into 7 bit codeword. Four parity bits are added to the original data; the number off parity bits added depends on the total number of bit errors the decoder can detect at the receiver side. Different type of decoding algorithms are present depending on the type of messages passed or on the type of operation performed at the nodes such as Bit-flipping decoding, Sum-product decoding, Message-passing on the binary erasure channel etc. This paper uses the bit flipping algorithm for LDPC decoder. In bit flipping algorithm, the decoder computes each parity check, using only hard-sliced binary input signals with simple XOR operations. It then schedules a bit to be flipped if the number of failed parity checks exceeds a fixed flipping threshold value b. The flipped bits are then used in the next iteration of the decoding process. The decoding algorithm stops when either all of the parity checks are satisfied or a pre-defined maximum iteration limit is reached [7]

Compared with other LDPC decoding algorithms, Bit flip algorithms use only one bit information and the parity check decision is based on a simple XOR operation. Because of its simplicity, a Bit Flip decoder could save large amount of power and silicon area.

IV. DESIGN AND OPTIMIZATION

A. *AES ENCRYPTION AND DECRYPTION OPTIMIZATION*

Here all the basic AES Transformations along with the key Expansion Schedule are defined as “Functions” and are called as and when required. In order to optimize, the SubByte and Shift-Row transformations of the AES Algorithm are combined together and defined as one “function for efficient device area and memory utilization. The Initial round is performed by calling a Function which performs the XOR operation of state and Key. For the remaining rounds two “Blocks” have been defined. Block 1 performs the first 9 rounds and block 2 performs the final round as shown in figure 3.

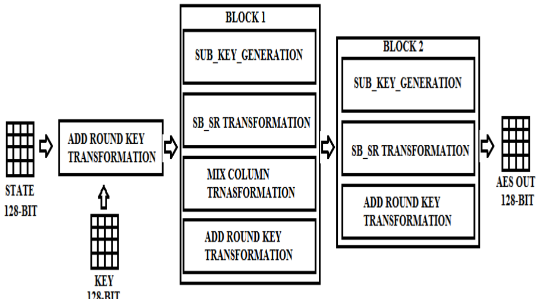


Figure 3: AES Encryption Design

Similarly at the Decryption the The Inverse SubByte and Shift-Row transformation are combined to form one transformation. Two blocks are defined Block 1 which performs Inv_Sb_Sr, InvMixColumn and Add Round key Transformation and Block 2 which performs Inv_Sb_Sr and Add Round Key Transformation as shown in figure 4.

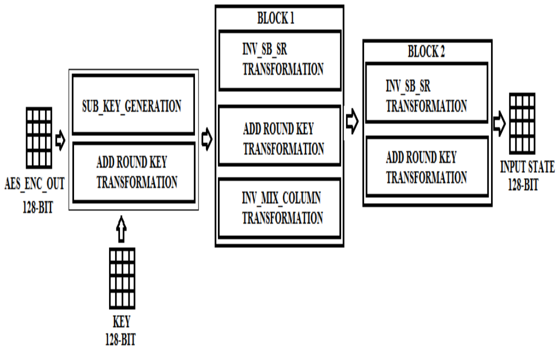


Figure 4: AES Decryption Design

B. *TRANSMITTER DESIGN AND IMPLEMENTATION*

The transmitter mainly consists of two blocks: 128-AES Encryption and 3-Bit LDPC Encoder. Since the output of AES block is 128-bit which is not compatible with the input of Encoder which is 3-bit, a shifter is used to connect these two blocks as shown in figure 5. At each clock cycle the shifter takes 3 bit data from the encryptor output and gives it as input to encoder. The Encoder produces 7 bit output data at each clock cycle. A total of 43 clock cycles are required for encoding 128-bit encrypted data. The AES Encryption along with LDPC encoder which forms the transmitter is designed in VHDL and synthesized using Xilinx ISE 12.2

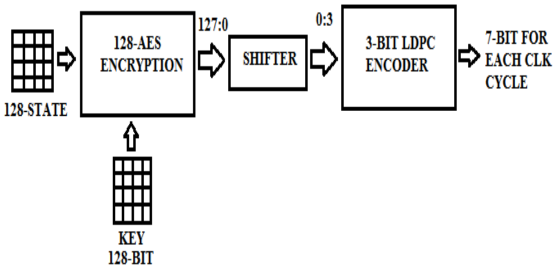


Figure 5: Transmitter

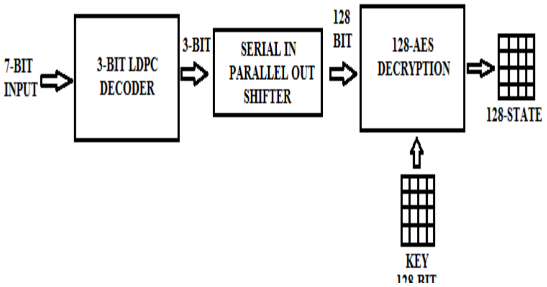


Figure 6: Receiver

The synthesis & mapping results of Transmitter are summarized in Table 2.

Target FPGA Device	Spartan 3 xc3s400-5tq144
Maximum Frequency	92.259MHz
Number of Slices	3041 out of 3584 (84%)
Number of Slice Flip Flops	560 out of 7168 (7%)
Number of 4 input LUTs	5716 out of 7168 (79%)
Number of GCLKs	1 out of 8 (12%)
256x8-bit ROM	20
Total Memory Usage	266040 kilobytes

C. *RECEIVER DESIGN*

As shown in figure 6 receiver consists of three blocks: 3-bit LDPC Decoder, serial-in-parallel-out shifter and 128-AES Decryptor. The LDPC decoder receives 7 bit input for each clock cycle. The decoder is capable of correcting up to two bit errors from the received data. The original 3-bit data is decoded from the 7-bit input after error correction, if any errors have occurred during transmission. The decoder takes 43 clock cycles to decode 128-bits of the transmitted data. Serial-in-parallel-out shifter receives 3 bit input from the decoder and stores it in a register, when the shifter receives all the 128 bits from the decoder, the output of shifter which is 128 bit is given as input to the 128-AES Decryptor. After 10 clock cycles the original 128-bit data is produced.

V. SIMULATION RESULTS

A behavioural description of designed AES Encryption, Transmitter, Decryption and Receiver was carried out using VHDL hardware description language. In order to verify the correctness of logic functions of 128-bit mode AES encryption and decryption in the transmitter and receiver, simulation of the system with MODELSIM SE 6.3f simulator was carried out. All the data used are testing data provided by FIPS.

A. *ENCRYPTION PROCESS*

Plain Text: 00112233445566778899aabbccddeeff

Key: 000102030405060708090a0b0c0d0e0f

Output/Cipher Text:

69c4e0d86a7b0430d8cdb78070b4c55a

Figure 7 represents the waveforms generated by the 128- bit complete encryption Process. The inputs are reset, clock, enable, 128-bit input and output is 128-bit encrypted data.

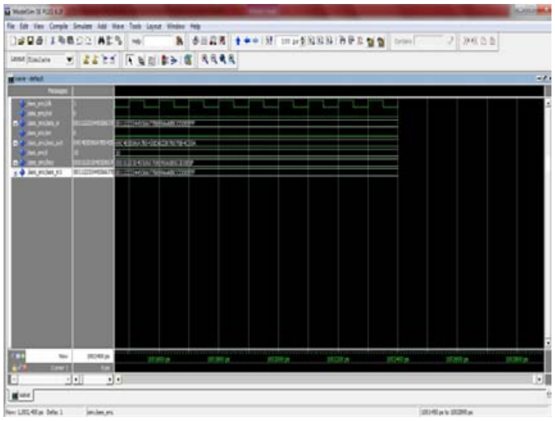


Figure 7: Simulation Waveform of Encryption Process

B. *DECRYPTION PROCESS*

Input /Cipher Text: 69c4e0d86a7b0430d8cdb78070b4c55a

Key: 000102030405060708090a0b0c0d0e0f

Output/Plain Text:

00112233445566778899aabbccddeeff

Figure 8 represents the waveforms generated by the 128- bit complete decryption Process. The inputs are reset, clock, enable, 128-bit encrypted input and output is 128-bit decrypted data.

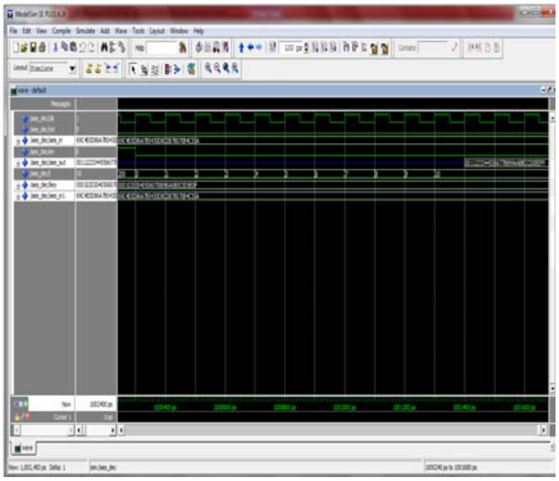


Figure 8: Simulation Waveform of Decryption Process

C. *TRANSMITTER*

Figure 9 represents the waveforms generated by the 128- bit Encryptor and LDPC encoder which forms transmitter. The inputs are reset, clock, load, 128-bit input and output is 7-bit data (encrypted and encoded).

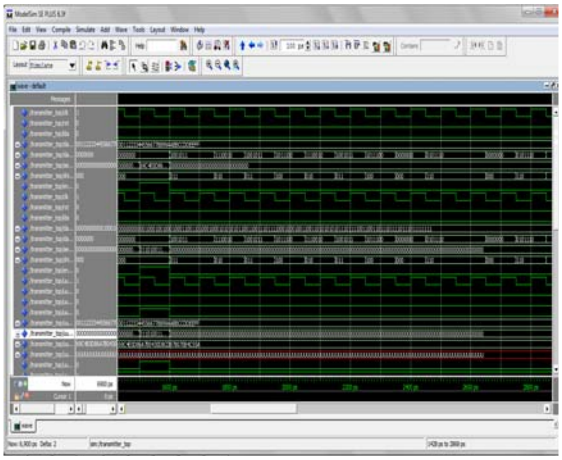
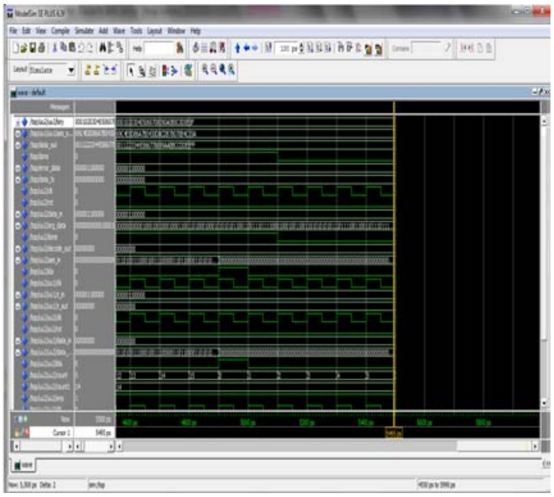


Figure 9: Simulation Waveform of Transmitter

D. *RECEIVER*

Figure 10 represents the waveforms generated by the 128- bit Decryptor and LDPC decoder which forms receiver.

The inputs are reset, clock, load, 7-bit input and output is 128-bit data (decoded and decrypted).



The analysis of simulation results of the above 128-bit input data, under the control of the same key, plaintext is Encrypted and encoded by the system and transmitted, and the resulting cipher text is again Decoded and decrypted by the system at the receiver, the final resulting data is consistent with the input plaintext. Thus it can be proved that the transmitter and receiver module is normal.

IV. CONCLUSION

An efficient FPGA implementation of 128 bit AES Encryption and LDPC encoder has been presented in this paper. Optimized and Synthesizable VHDL code is developed for the implementation of both 128 bit data encryption and decryption process using Xilinx 12.2 & description is verified using ModelSim 6.3f simulator. In order to minimize the hardware consumption all the transformations of algorithm are simulated using an iterative design approach. For receiver the simulation results are verified and the system uses the LDPC code for error correction.

REFERENCES

[1] FIPS 197, “Advanced Encryption Standard (AES)”, November 26,2001.

[2] Atul M. Borkar, R.V. Kshirsagar, M. V. Vyawahare, ”FPGA Implementation of AES

Algorithm", Third International Conference on Electronics Computer Technology (ICECT), 2011.

[3] Hao Zhong, Tong Zhang, "Block-LDPC: A Practical LDPC Coding System Design Approach" IEEE Transactions on Circuits and Systems-I, VOL. 52, NO. 4, April 2005.

[4] Dur-e-Shahwar Kundi, Saleha Zaka, Qurat-Ul-Ain, Arshad Aziz, "A Compact AES Encryption Core on Xilinx FPGA", Second International Conference on Computer, Control and Communication, 2009.

[5] Luanlan, "The AES Encryption And Decryption Realization Based On FPGA", Seventh International Conference on Computational Intelligence and Security, 2011.

[6] R.G. Gallager, "Low-Density Parity-Check Codes", IRE Transactions on Information Theory, 1962.

[7] Xin Sheng Zhou, Bruce F. Cockburn, Stephen Bates, "Improved Iterative Bit Flipping Decoding Algorithms for LDPC Convolutional Codes", IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 2007.

[8] Jing-Jie Wu, Rui Huang "A FPGA-Based Wireless Security System", Third International Conference on Multimedia Information Networking and Security, 2011.

[9] Webber J, Nishimurs T, Ohgane T, Ogawa Y, "A Study on Adaptive Thresholds for Reduced Complexity Bit-Flip Decoding" fourteenth International Conference on Advanced Communication Technology (ICACT), 2012

◈◈◈

Implementation of Robust Architecture for Error detection and Data Recovery in Motion Estimation on FPGA

V.V.S.V.S. Ramachandram & Finney Daniel. N

Department of ECE, Pragati Engineering College, Surampalem, Kakinada, A.P, India.

Abstract - Video compression is necessary in a wide range of applications to reduce the total data amount required for transmitting or storing video data. Among the coding systems, Motion Estimation is of priority concern in exploiting the temporal redundancy between successive frames, yet also the most time consuming aspect of coding. This paper presents an error detection and data recovery (EDDR) design, based on the residue-and quotient (RQ) code that is embed into ME for video coding testing applications. Based on the Concurrent Error Detection (CED) concept, this work develops a robust EDDR architecture based on the RQ code to detect errors and recovery data in PEs of a ME and, in doing so, further guarantee the excellent reliability for video coding applications. We synthesized this design using Xilinx tool.

***Keywords**: Concurrent Error Detection, Motion Estimation, Residue quotient code generation.*

I. INTRODUCTION

At its most basic level, compression is performed when an input video stream is analyzed and information that is indiscernible to the viewer is discarded. Each event is then assigned a code – commonly occurring events are assigned few bits and rare events will have more bits. These steps are commonly called signal analysis, quantization and variable length encoding respectively. There are four methods for compression; discrete cosine transforms (DCT), vector quantization (VQ), fractal compression, and discrete wavelet transform (DWT). Discrete cosine transform is a lossy compression algorithm that samples an image at regular intervals, analyzes the frequency components present in the sample, and discards those frequencies which do not affect the image as the human eye perceives it. DCT is the basis of standards such as JPEG, MPEG, H.261, and H.263.

Advance Video Coding is widely regarded as the next generation video compression standard. So, video compression is necessary to reduce the total data amount required for transmitting or storing video data. Among the coding systems, a ME is of priority concern in exploiting the temporal redundancy between successive frames, yet also the most time consuming aspect of coding. Additionally, while performing up to 60%–90% of the computations encountered in the entire coding system, a ME is widely regarded as the most computationally intensive of a video coding system. To explore the feasibility of a BIST to detect errors and recover data of a ME is of worthwhile interest. Additionally, the reliability issue of numerous PE's in a ME can be improved by enhancing the capabilities of concurrent error detection (CED). CED can also test the circuit at full operating speed without interrupting a system. Based on the CED concept, this work develops a robust EDDR architecture based on the RQ code, to detect errors and recovery data in PEs of a ME. This guarantees the excellent reliability for video coding testing applications.

II. SYSTEM MODEL

The proposed EDDR architecture is as shown in Fig.1. Which comprises two major circuit designs, i.e. data recovery circuit (DRC),to detect errors and recover the corresponding data in a specific PE and error detection circuit (EDC).

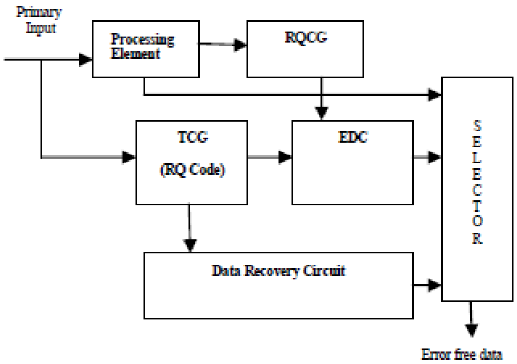


Fig.1. Conceptual View of the Proposed EDDR architecture

The test code generator (TCG) utilizes the concepts of RQ code to generate the corresponding test codes for errordetection and data recovery. In other words, the test codes from TCG and the primary output from CUT are delivered to EDC to determine whether the CUT has errors. DRC is in charge of recovering data from TCG. Additionally, a selector is enabled to export error-free data or data-recovery results. Importantly, an array-based computing structure, such as ME, discrete cosinetransform (DCT), iterative logic array (ILA), and finite impulse filter (FIR), is feasible for the proposed EDDR scheme to detect errors and recover the corresponding data.

III. SYSTEM DESIGN

A. PROCESSING ELEMENT

A ME consists of many PE's incorporated in a 1-D or 2-D array for video encoding applications. It consists of PEs with a size of 4X4.Fig.2. shows the internal structure of PE. A PE generally consists of two ADDs (i.e. an 8-bit ADD and a 12-bit ADD) and an accumulator (ACC). Next, the 8-bit ADD (a pixel has 8-bit data) is used to estimate the addition of the current pixel (Cur_pixel) and reference pixel (Ref_pixel). Additionally, a 12-b ADD and an ACC are required to accumulate the results from the 8-b ADD in order to determine the sum of absolute difference (SAD) value for video encoding applications.

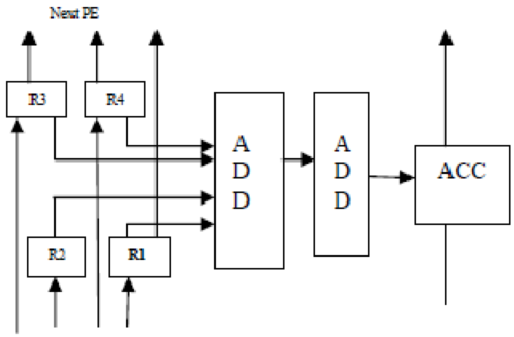


Fig.2. Internal Structure of Processing Element

Notably, some registers and latches exist in ME to complete the data shift and storage applications. The PEs is essential building blocks and are connected regularly to construct a ME. Generally, PE's issurrounded by sets of ADDs and accumulators that determine how data flows through them. Additionally, the visual quality and peak signal-to-noise ratio (PSNR) at a given bit rate are influenced if an error occurred in ME process. A testable design is thus increasingly important to ensure the reliability of numerous PEs in a ME.PEs can thus be considered the class of circuits called ILAs, whose testing assignment can be easily achieved by using the fault model, cell fault model (CFM). Using CFM has received considerable interest due to accelerated growth in the use of high-level synthesis, as well as the parallel increase in complexity and density of integration circuits.

B. RQ CODE GENERATION

To detect circuit errors in design applications coding approaches such as parity code, Berger code, and residue code have been considered. Residue code is generally separable arithmetic codes by estimating a residue for data and appending it to data. Error detection logic for operations is typically derived by a separate residue code, making the detection logic is simple and easily implemented. However, only a bit error can be detected using residue code. Therefore a quotient code which is derived from the residue code, to help the residue code in detecting errors and recovering the data. If the input data is represented as X bits, the binary data is expressed as

$$X = \{b_{n-1}b_{n-2} \dots\dots\dots\dots\dots\dots b_2b_1b_0\} = \sum_{i=0}^{n-1} b_i\,2^i \quad (1)$$

To accelerate the circuit design of RQCG, the binary data in (1) can generally be divided into two parts:

$$X = \sum_{i=0}^{n-1} b_i\,2^i$$

$$= \left(\sum_{i=0}^{k-1} b_i 2^i\right) + \left(\sum_{i=k}^{n-1} b_i 2^{i-k}\right) 2^k$$

$$= Y_0 + Y_1 2^k \qquad (2)$$

Significantly, the value of *k* is equal to *n*/2. The modulus value i.e. *m*= $2^k - 1$. The residue code is given by

$$\mathrm{R} = |\mathrm{X}|\mathrm{m} \qquad (3)$$

$$\mathrm{Q} = \left\lfloor \frac{\mathrm{X}}{\mathrm{m}} \right\rfloor \qquad (4)$$

Based on the equations, the corresponding circuit design of RQCG is easily realized. The basic architecture of RQ code is in Fig.3.

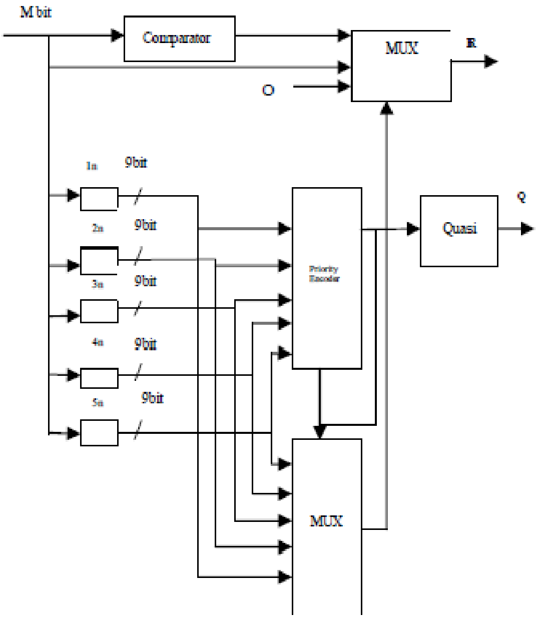
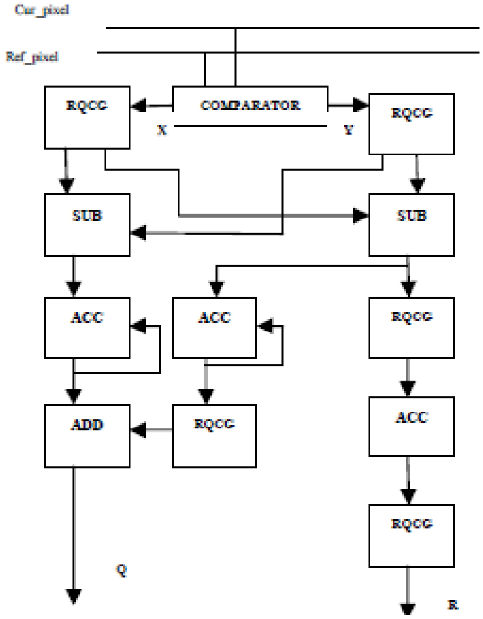


Fig.3. Residue Quotient Code Generator

The main elements of RQCG are the priority encoder and the multiplexer. These elements generate the quotient value by passing through a Quasi. The residue value is obtained directly by the comparator.

C. *TEST CODE GENERATION*

The important component of the TCG design is the RQCG circuit.TCG design is based on the ability of the RQCG circuit to generate corresponding test codes in order to detect errors and recover data. The specific PE estimates the absolute difference between the Cur_pixel of the search area and the Ref_pixel of the current macro block. The absolute difference is obtained based on SAD (Sum of Absolute Difference). The TCG consists of 5 RQCG block and comparator, accumulators and subtractors.



D. *SAD TREE*

The SAD of a macro block with the size *NXN*can be compute as follows:

$$SAD = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left|X_{ij} - Y_{ij}\right|$$

$$= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left|(q_{xij}.\,m + r_{xij}) - (q_{yij}.\,m + r_{yij})\right| \qquad (5)$$

The 2-D intra-level architecture called the Propagate Partial SAD is better suited for the absolute difference evaluation. The architecture of SAD is shown in Fig.5. The SAD is composed of PE arrays with a 1-D adder tree in the vertical direction. Current pixels are stored in each PE, and two sets of continuous reference pixels in a row are broadcasted to PE arrays at the same time.

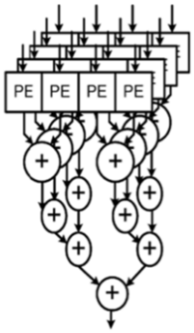


Fig.5. Propagate Partial SAD

In each PEarray with a 1-D adder tree, distortions are computed and summed by a 1-D adder tree to generate one-row SAD. The row SADs are accumulated and propagated with propagation registers in the vertical direction. The reference data of searching candidates in the even and odd columns are inputted by Ref_pixel 0 and Ref_pixel 1, respectively. After initial cycles, the SAD of the first searching candidate in zero columns is generated, and the SADs of the other searching candidates are sequentially generated in the following cycles. When computing the last searching candidates in each column, the reference data of searching candidates in the next columns begin to be inputted through another reference input. In Propagate Partial SAD, by broadcasting reference pixel rows and propagating partial-row SADs in the vertical direction, it provides the advantages of fewer reference pixel registers and a shorter critical path. Since Rt(Qt) is equal to RPEi (QPEi) EDC is enabled and a signal “0” is generated to describe a situation in which the specific PEi is error-free. Conversely, if SA1 and SA0 errors occur in bits 1 and 12 of a specific PE i.e. the SAD value of PEi= 2124. Distortion is the difference between the current pixel and the reference pixel, and SAD is the total distortion of this searching candidate. The row column) SAD is the summation of distortions in a row (column). After all searching candidates are examined; the searching candidate that has the smallest SAD is selected as the motion vector of the current macro block.

E. CONCURRENT ERROR DETECTION (CED)

Thevisual quality and peak signal-to-noise ratio (PSNR) at a given bit rate are influenced if an error occurred in ME process. While the extended BIST schemes generally focus on memory circuit, testing-related issues of video coding. Thus, exploring the feasibility of an embedded testing approach to detect errors and recover data of a ME is of worthwhile interest. So, the reliability issue of numerous PE’s in a ME can be improved by enhancing the capabilities of concurrent error detection (CED). The CED approach can detect errors through conflicting and undesired results generated from operations on the same operands. This scheme can also test the circuit at full operating speed without interrupting a system. Hence, based on this concept, a robust EDDR architecture can be achieved based on the RQ code to detect errors and recovery data in PEs of a ME.This further guarantees the excellent reliability for video coding testing applications. However, only a bit error can be detected based on the residue code. Additionally, an error can’t be recovered effectively by using the residue codes. Therefore, a quotient code which is derived from the residue code to assist the residue code in detecting multiple errors and recovering errors. The corresponding circuit architecture of the RQCG is easily realized and the RQ code can be generated with a low complexity and little hardware cost.

Concurrent test methods enable integrated circuits to verify the correctness of their results during normal operation. Quality assessment of concurrent test methods relies on several parameters, including the model of detectable faults or errors, the worst-case detection latency, and the incurred area overhead. Several low-cost, non-intrusive, concurrent fault detection (CFD) methods have been proposed for stuck-at faults in combinational circuits. In the R-CBIST method, the requirement for input combinations is relaxed at the cost of a small RAM. Alternatively, latency is reduced.

IV. ERROR DETECTION AND DATA RECOVERYPROCESS.

To detect errors and recover the corresponding data in a specific ME our EDDR architecture, which comprises two major circuit designs, i.e. error detection circuit (EDC) and data recovery circuit (DRC) is sufficient. The test code generator utilizes the RQ code to generate the corresponding test codes for error detection and data recovery. The test codes from TCG and the primary output are delivered to EDC to determine whether the input has errors. DRC is in charge of recovering data from TCG. A selector is enabled to export error-free data or data-recovery results. Inorder todetermine whether errors have occurred,the outputs between TCG and RQCG is compared. The errors in a specific PE can be detected when the values of $R_{PE} \neq R_T$ and similarly the quotient values. 0/1 signal is generated by error detecting circuit to indicate the error.

The lost data can be recovered by DRC. This plays an important role in recovering RQ code from TCG. The mathematical model for data recovery is

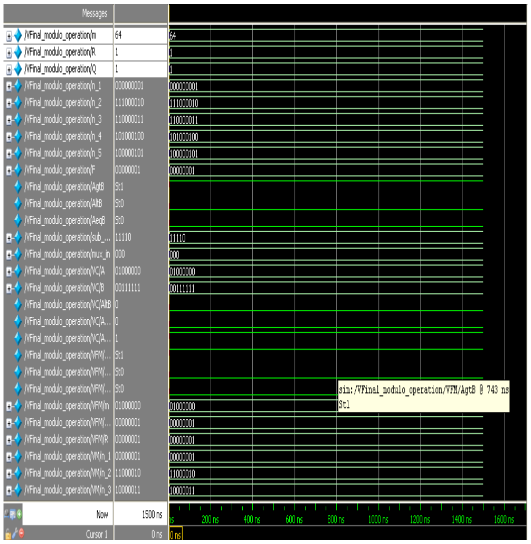
$$\mathrm{SAD} = m \times \mathrm{Q}_T + R_T$$

$$= (2^{\mathrm{i}} - 1) \times \mathrm{Q}_{\mathrm{T}} + \mathrm{R}_{\mathrm{T}}$$

The Barrel shift and Corrector circuits are necessary to achieve the functions $2^i \times Q_T$ and$-Q_T + R_T$ respectively for data recovery.

V. RESULTS

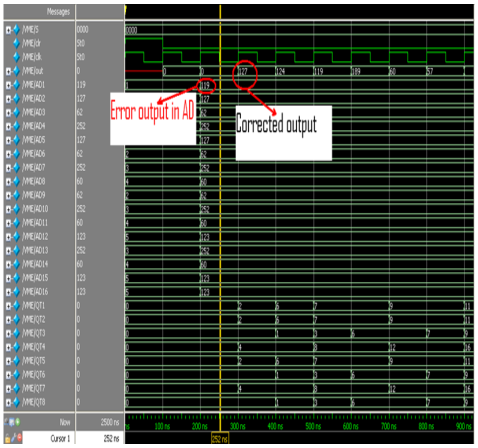
Extensive verification of the circuit design is performed using the Verilog and then synthesized by the Xilinx9.2 to demonstrate the feasibility of the proposed EDDR architecture design for ME testing applications.



Messages	
/VFinal_modulo_operation/m	64
/VFinal_modulo_operation/R	1
/VFinal_modulo_operation/Q	1
/VFinal_modulo_operation/n_1	000000001
/VFinal_modulo_operation/n_2	111000010
/VFinal_modulo_operation/n_3	110000011
/VFinal_modulo_operation/n_4	101000100
/VFinal_modulo_operation/n_5	100000101
/VFinal_modulo_operation/F	00000001
/VFinal_modulo_operation/AgtB	St1
/VFinal_modulo_operation/AltB	St0
/VFinal_modulo_operation/AeqB	St0
/VFinal_modulo_operation/sub_...	11110
/VFinal_modulo_operation/mux_in	000
/VFinal_modulo_operation/VC/A	01000000
/VFinal_modulo_operation/VC/B	00111111
/VFinal_modulo_operation/VC/AltB	0
/VFinal_modulo_operation/VC/A...	0
/VFinal_modulo_operation/VC/A...	1
/VFinal_modulo_operation/VFM/...	St1
/VFinal_modulo_operation/VFM/...	St0
/VFinal_modulo_operation/VFM/...	St0
/VFinal_modulo_operation/VFM/m	01000000
/VFinal_modulo_operation/VFM/...	00000001
/VFinal_modulo_operation/VFM/R	00000001
/VFinal_modulo_operation/VM/n_1	00000001
/VFinal_modulo_operation/VM/n_2	11000010
/VFinal_modulo_operation/VM/n_3	10000011
Now	1500 ns
Cursor 1	0 ns

sim:/VFinal_modulo_operation/VFM/AgtB @ 743 ns
St1

Fig. 6.1 Simulation results of RQCG



Messages	
/VME/S	0000
/VME/clr	St0
/VME/clk	St0
/VME/out	0
/VME/AD1	119
/VME/AD2	127
/VME/AD3	62
/VME/AD4	252
/VME/AD5	127
/VME/AD6	62
/VME/AD7	252
/VME/AD8	60
/VME/AD9	62
/VME/AD10	252
/VME/AD11	60
/VME/AD12	123
/VME/AD13	252
/VME/AD14	60
/VME/AD15	123
/VME/AD16	123
/VME/QT1	0
/VME/QT2	0
/VME/QT3	0
/VME/QT4	0
/VME/QT5	0
/VME/QT6	0
/VME/QT7	0
/VME/QT8	0
Now	2500 ns
Cursor 1	252 ns

Fig. 6.2 Simulation results of EDDR

CONCLUSION

We presented an efficient RQCG architecture for a error detection and data recovery in MEfor video coding testing applications.Simulation results shows that the proposed design detects the errors and recovers the data. An error in processing elements key components of a ME can be detectedand recovered effectively by using the proposed EDDR design.

REFERENCES

[1] C. Y. Chen, S. Y. Chien, Y. W. Huang, T. C. Chen, T. C. Wang, and L. G. Chen, “Analysis and architecture design of variable block-size motion estimation for H.264/AVC,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 53, no. 3, pp. 578–593, Mar. 2006.

[2] T. H. Wu, Y. L. Tsai, and S. J. Chang, “An efficient design-for-testabilityscheme for motion estimation in H.264/AVC,” in Proc. Int. Symp.VLSI Design, Autom. Test, Apr. 2007, pp. 1–4.

[3] M. Y. Dong, S. H. Yang, and S. K. Lu, “Design-for-testability techniquesfor motion estimation computing arrays,” in Proc. Int. Conf. Commun., Circuits Syst., May 2008, pp. 1188–1191.

[4] Y. S. Huang, C. K. Chen, and C. L. Hsu, “Efficient built-in self-test for video coding cores: A case study on motion estimation computing array,” in Proc. IEEE Asia Pacific Conf. Circuit Syst., Dec. 2008, pp. 1751–1754

[5] W. Y Liu, J. Y. Huang, J. H. Hong, and S. K. Lu, “Testable design and BIST techniques for systolic motion estimators in the transform domain,” in Proc. IEEE Int. Conf. Circuits Syst., Apr. 2009, pp. 1–4.

[6] J. M. Portal, H. Aziza, and D. Nee, “EEPROM memory: Threshold S. Bayat-Sarmadi and M. A. Hasan, “On concurrent detection of error sin polynomial basis multiplication,” IEEE Trans. Vary Large Scale Integr.(VLSI) Systs., vol. 15, no. 4, pp. 413–426, Apr. 2007.

VHDL Implementation of a Novel Low Power BIST using bit Swapping LFSR and Scan Chain Re-ordering

Kusumanchi Avinash Kumar, P. Krishna Rao & #T. Ravindra

Dept. of E.C.E., SSCE, Chilakapalem Jn., Srikakulam, India.
# ASGRTAW, Kavulavada, Viziangaram, India.

Abstract - A low-transition linear feedback shift register (LFSR) that is based on some new observations about the output sequence of a conventional LFSR is discussed in this paper. Bit-swapping LFSR (BS-LFSR) is the proposed design which is composed of an LFSR and a 2 × 1 multiplexer. As compared to the conventional LFSR, the number of transitions occurs at the scan-chain input during scan chain shift operation reduced by 50% to generate test patterns for scan-based built-in self-tests. Hence, it reduces the overall switching activity in the circuit under test during test applications. A scan-chain-ordering algorithm is also combined with the BS-LFSR that orders the cells in a way that reduces the average and peak power (scan and capture) in the test cycle or while scanning out a response to a signature analyzer. These techniques have a substantial effect on average- and peak-power reductions with negligible effect on fault coverage or test application time. Experimental results on ISCAS'89 benchmark circuits show up to 65% and 55% reductions in average and peak power, respectively.

Keywords*- Built-in self-test (BIST), linear feedback shift register (LFSR), low-power test, pseudorandom pattern generator, scan-chain ordering.*

I. INTRODUCTION

In recent years, the design for low power has become one of the greatest challenges in high-performance very large scale integration (VLSI) design. As a consequence, many techniques have been introduced to minimize the power consumption of new VLSI systems. However, most of these methods focus on the power consumption during normal mode operation, while test mode operation has not normally been a predominant concern. However, it has been found that the power consumed during test mode operation is often much higher than during normal mode operation [1]. This is because most of the consumed power results from the switching activity in the nodes of the circuit under test (CUT), which is much higher during test mode than during normal mode operation [1]-[3].Several techniques that have been developed to reduce the peak and average power dissipated during scan-based tests can be found in [4] and [5]. A direct technique to reduce power consumption is by running the test at a slower frequency than that in normal mode. This technique of reducing power consumption, while easy to implement, significantly increases the test application time [6]. Furthermore, it fails in reducing peak-power consumption since it is independent of clock frequency. Another category of techniques used to reduce the power consumption in scan-based built-in self-tests (BISTs) is by using scan-chain-ordering techniques [7]-[13].

These techniques aim to reduce the average-power consumption when scanning in test vectors and scanning out captured responses. Although these algorithms aim to reduce average-power consumption, they can reduce the peak power that may occur in the CUT during the scanning cycles, but not the capture power that may result during the test cycle (i.e., between launch and capture). The design of low-transition test-pattern generators (TPGs) is one of the most common and efficient techniques for low-power tests [14]-[20]. These algorithms modify the test vectors generated by the LFSR to get test vectors with a low number of transitions. The main drawback of these algorithms is that they aim only to reduce the average-power consumption while loading a new test vector, and they ignore the power consumption that results while scanning out the captured response or during the test cycle.

Furthermore, some of these techniques may result in lower fault coverage and higher test- application time. Other techniques to reduce average-power consumption during scan-based tests include scan segmentation into multiple scan chains [6], [21], test-scheduling techniques [22], [23], static-compaction techniques [24], and multiple scan chains with many scan enable inputs to activate one scan chain at a time [25]. The latter technique also reduces the peak power in the CUT. On the other hand, in addition to the techniques mentioned earlier, there are some new approaches that aim to reduce peak-power consumption during tests, particularly the capture power in the test cycle.

One of the common techniques for this purpose is to modify patterns using an X-filling technique to assign values to the don’t care bits of a deterministic set of test vectors in such a way as to reduce the peak power in the test vectors that have a peak-power violation [26]-[29].This paper presents a new TPG, called the bit-swapping linear feedback shift register (BS-LFSR), that is based on a simple bit-swapping technique applied to the output sequence of a conventional LFSR and designed using a conventional LFSR and a 2 × 1 multiplexer. The proposed BS-LFSR reduces the average and instantaneous weighted switching activity (WSA) during test operation by reducing the number of transitions in the scan input of the CUT. The BS-LFSR is combined with a scan-chain-ordering algorithm that reduces the switching activity in both the test cycle (capture power) and the scanning cycles (scanning power).

II. PROPOSED APPROACH TO DESIGN THE BS - LFSR

The proposed BS-LFSR for test-per-scan BISTs is based upon some new observations concerning the number of transitions produced at the output of an LFSR. Definition: Two cells in an n-bit LFSR are considered to be adjacent if the output of one cell feeds the input of the second directly (i.e., without an intervening XOR gate).

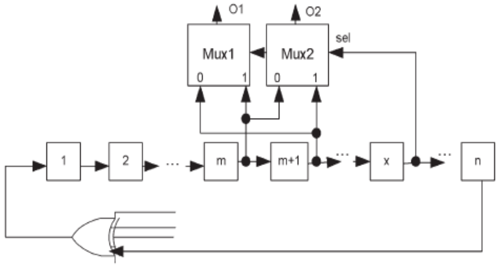


Fig. 1. Swapping arrangement for an LFSR.

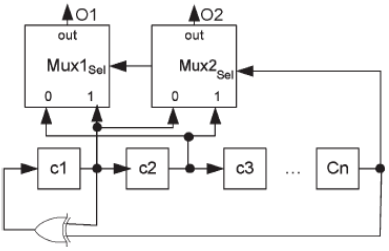


Fig. 2. External LFSR that implements the prime polynomial $x^n + x + 1$ and the proposed swapping arrangement.

Lemma 1: Each cell in a maximal-length n-stage LFSR (internal or external) will produce a number of transitions equal to 2^{n-1} after going through a sequence of 2^n clock cycles.

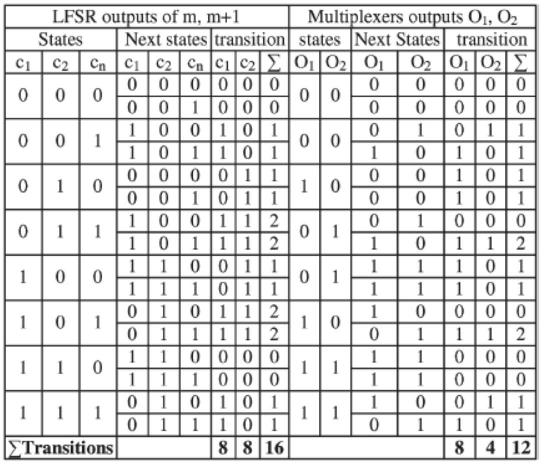
Proof: The sequence of 1s and 0s that is followed by one bit position of a maximal-length LFSR is commonly referred to as an m-sequence. Each bit within the LFSR will follow the same m-sequence with a one-time-step delay. The m-sequence generated by an LFSR of length n has a periodicity of $2^n - 1$. It is a well-known standard property of an m-sequence of length n that the total number of runs of consecutive occurrences of the same binary digit is 2^{n-1} [3], [30]. The beginning of each run is marked by a transition between 0 and 1; therefore, the total number of transitions for each stage of the LFSR 2^{n-1}. This lemma can be proved by using the toggle property of the XOR gates used in the feedback of the LFSR [32].

Lemma 2: Consider a maximal-length n-stage internal or external LFSR (n > 2). We choose one of the cells and swap its value with its adjacent cell if the current value of a third cell in the LFSR is 0 (or 1) and leave the cells un-swapped if the third cell has a value of 1 (or 0). Fig. 1 shows this arrangement for an external LFSR (the same is valid for an internal LFSR). In this arrangement, the output of the two cells will have its transition count reduced by $T_{saved} = 2^{(n-2)}$ transitions. Since the two cells originally produce $2 \times 2^{n-1}$ transitions, then the resulting percentage saving is $T_{saved\%} = 25\%$ [32]. In Lemma 2, the total percentage of transition savings after swap-ping is 25% [31]. In the case where cell x is not directly linked to cell m or cell m + 1 through an XOR gate, each of the cells has the same share of savings (i.e., 25%).

Lemmas 3-10 show the special cases where the cell that drives the selection line is linked to one of the swapped cells through an XOR gate. In these configurations, a single cell can save 50% transitions that were originally produced by an LFSR cell. Lemma 3 and its proof are given; other lemmas can be proved in the same way.

Lemma 3: For an external n-bit maximal-length LFSR that implements the prime polynomial $x^n + x + 1$ as shown in Fig. 2, if the first two cells (c_1 and c_2) have been chosen for swapping and cell n as a selection line, then o_2 (the output of MUX2) will produce a total transition savings of 2^{n-2} compared to the number of transitions produced by each LFSR cell, while o_1 has no savings (i.e., the savings in transitions is concentrated in one multiplexer output, which means that o_2 will save 50% of the original transitions produced by each LFSR cell).

TABLE I
POSSIBLE AND SUBSEQUENT STATES FOR CELLS c_1, c_2, AND c_n
(SEE FIG. 2)

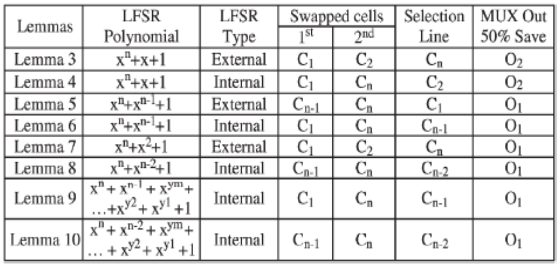


LFSR outputs of m, m+1									Multiplexers outputs O$_1$, O$_2$						
States			Next states			transition			states		Next States		transition		
c$_1$	c$_2$	c$_n$	c$_1$	c$_2$	c$_n$	c$_1$	c$_2$	∑	O$_1$	O$_2$	O$_1$	O$_2$	O$_1$	O$_2$	∑
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			0	0	1	0	0	0			0	0	0	0	0
0	0	1	1	0	0	1	0	1	0	0	0	1	0	1	1
			1	0	1	1	0	1			1	0	1	0	1
0	1	0	0	0	0	0	1	1	1	0	0	0	1	0	1
			0	0	1	0	1	1			0	0	1	0	1
0	1	1	1	0	0	1	1	2	0	1	0	1	0	0	0
			1	0	1	1	1	2			1	0	1	1	2
1	0	0	1	1	0	0	1	1	0	1	1	1	1	0	1
			1	1	1	0	1	1			1	1	1	0	1
1	0	1	0	1	0	1	1	2	1	0	1	0	0	0	0
			0	1	1	1	1	2			0	1	1	1	2
1	1	0	1	1	0	0	0	0	1	1	1	1	0	0	0
			1	1	1	0	0	0			1	1	0	0	0
1	1	1	0	1	0	1	0	1	1	1	1	0	0	1	1
			0	1	1	1	0	1			0	1	1	0	1
∑Transitions						**8**	**8**	**16**					**8**	**4**	**12**

Proof: There are eight possible combinations for the initial state of the cells c$_1$, c$_2$, and c$_n$. If we then consider all possible values of the following state, we have two possible combinations (not eight, because the value of c$_2$ in the next state is determined by the value of c$_1$ in the present state; also, the value of c$_1$ in the next state is determined by "c$_1$xor c$_n$" in the present state). Table I shows all possible and subsequent states.

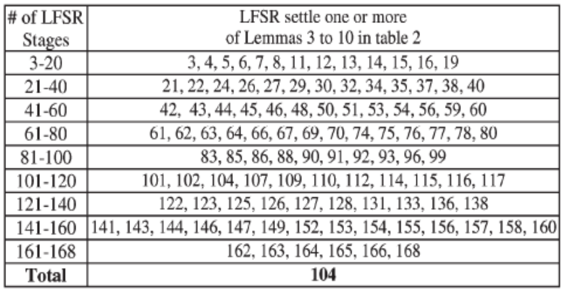
It is important to note that the overall savings of 25% is not equally distributed between the outputs of the multiplexers as in Lemma 2. This is because the value of c$_1$ in the present state will affect the value of c$_2$ and its own value in the next state ($c_{2(Next)} = c_1$ and $c_{1(Next)}$ = "c$_1$ xor c$_n$"). To see the effect of each cell in transition savings, Table I shows that o$_1$ will save one transition when moving from state (0,0,1) to (1,0,0), from (0,1,1) to (1,0,0), from (1,0,1) to (0,1,0), or from (1,1,1) to (0,1,0). In the same time, o$_1$ will increase one transition when moving from (0,1,0) to (0,0,0), from (0,1,0) to (0,0,1), from (1,0,0) to (1,1,0), or from (1,0,0) to (1,1,1). Since o$_1$ increases the transitions in four possible scenarios and save transitions in other four scenarios, then it has a neutral overall effect because all the scenarios have the same probabilities. For o$_2$, one transition is saved when moving from (0,1,0) to (0,0,0), from (0,1,0) to (0,0,1), from (0,1,1) to (1,0,0), from (1,0,0) to(1,1,0), from (1,0,0) to (1,1,1), or from (1,0,1) to (0,1,0). At the same time, one additional transition is incurred when moving from state (0,0,1) to (1,0,0) or from (1,1,1) to (0,1,0). This gives o$_2$ an overall saving of one transition in four possible scenarios where the initial states has a probability of 1/8 and the final states of probability 1/2; hence, P$_{save}$ is given by P_{save} = 1/8 × 1/2 + 1/8 × 1/2 + 1/8 × 1/2 + 1/8 × 1/2 = 1/4. (1).

TABLE II
SPECIAL CASES WHERE ONE CELL SAVES 50% OF THE TRANSITIONS



Lemmas	LFSR Polynomial	LFSR Type	Swapped cells		Selection Line	MUX Out 50% Save
			1st	2nd		
Lemma 3	x^n+x+1	External	C$_1$	C$_2$	C$_n$	O$_2$
Lemma 4	x^n+x+1	Internal	C$_1$	C$_n$	C$_2$	O$_2$
Lemma 5	$x^n+x^{n-1}+1$	External	C$_{n-1}$	C$_n$	C$_1$	O$_1$
Lemma 6	$x^n+x^{n-1}+1$	Internal	C$_1$	C$_n$	C$_{n-1}$	O$_1$
Lemma 7	x^n+x^2+1	External	C$_1$	C$_2$	C$_n$	O$_1$
Lemma 8	$x^n+x^{n-2}+1$	Internal	C$_{n-1}$	C$_n$	C$_{n-2}$	O$_1$
Lemma 9	$x^n+x^{n-1}+x^{ym}+ \ldots +x^{y2}+x^{y1}+1$	Internal	C$_1$	C$_n$	C$_{n-1}$	O$_1$
Lemma 10	$x^n+x^{n-2}+x^{ym}+ \ldots +x^{y2}+x^{y1}+1$	Internal	C$_{n-1}$	C$_n$	C$_{n-2}$	O$_1$

TABLE III
LFSRS THAT SATISFY ONE OR MORE OF LEMMAS 3–10



# of LFSR Stages	LFSR settle one or more of Lemmas 3 to 10 in table 2
3-20	3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 16, 19
21-40	21, 22, 24, 26, 27, 29, 30, 32, 34, 35, 37, 38, 40
41-60	42, 43, 44, 45, 46, 48, 50, 51, 53, 54, 56, 59, 60
61-80	61, 62, 63, 64, 66, 67, 69, 70, 74, 75, 76, 77, 78, 80
81-100	83, 85, 86, 88, 90, 91, 92, 93, 96, 99
101-120	101, 102, 104, 107, 109, 110, 112, 114, 115, 116, 117
121-140	122, 123, 125, 126, 127, 128, 131, 133, 136, 138
141-160	141, 143, 144, 146, 147, 149, 152, 153, 154, 155, 156, 157, 158, 160
161-168	162, 163, 164, 165, 166, 168
Total	**104**

If the LFSR is allowed to move through a complete cycle of 2^nstates, then Lemma 1 shows that the number of transitions expected to occur in the cell under consideration is 2^{n-1}. Using the swapping approach, in 1/4 of the cases, a saving of one transition will occur, giving a total saving of 1/4 × $2^n = 2^{n-2}$. Dividing one figure by the other, we see that the total number of transitions saved at o$_2$ is 50%.In the special configurations shown in Table II (i.e. Lemmas 3-10), if the cell that saves 50% of the transitions is connected to feed the scan-chain input, then it saves 50% of the transitions inside the scan-chain cells, which directly reduces the average power and also the peak power that may result while scanning in a new test vector. Table III shows that there are 104 LFSRs (internal and external) whose sizes lie in the range of 3-168 stages that can be configured to satisfy one or more of the special cases in Table II to concentrate the transition savings in one multiplexer output.

III. IMPORTANT PROPERTIES OF THE BS-LFSR

There are some important features of the proposed BS-LFSR that make it equivalent to a conventional LFSR. The most important properties of the BS-LFSR are the following.

1) The proposed BS-LFSR generates the same number of 1s and 0s at the output of multiplexers after swapping of two adjacent cells; hence, the probabilities of

having a 0 or 1 at a certain cell of the scan chain before applying the test vectors are equal. Hence, the proposed design retains an important feature of any random TPG. Furthermore, the output of the multiplexer depends on three different cells of the LFSR, each of which contains a pseudorandom value. Hence, the expected value at the output can also be considered to be a pseudorandom value.

2) If the BS-LFSR is used to generate test patterns for either test-per-clock BIST or for the primary inputs of a scan-based sequential circuit (assuming that they are directly accessible) as shown in Fig. 3, then consider the case that c_1 will be swapped with c_2 and c_3 with c_4, . . . , c_{n-2} with c_{n-1} according to the value of c_n which is connected to the selection line of the multiplexers (see Fig. 3). In this case, we have the same exhaustive set of test vectors as would be generated by the conventional LFSR, but their order will be different and the overall transitions in the primary inputs of the CUT will be reduced by 25% [32].

IV. CELL REORDERING ALGORITHM

Although the proposed BS-LFSR can achieve good results in reducing the consumption of average power during test and also in minimizing the peak power that may result while scanning a new test vector, it cannot reduce the overall peak power because there are some components that occur while scanning out the captured response or while applying a test vector and capturing a response in the test cycle. To solve these problems, first, the proposed BS-LFSR has been combined with a cell-ordering algorithm presented in [11] that reduces the number of transitions in the scan chain while scanning out the captured response.

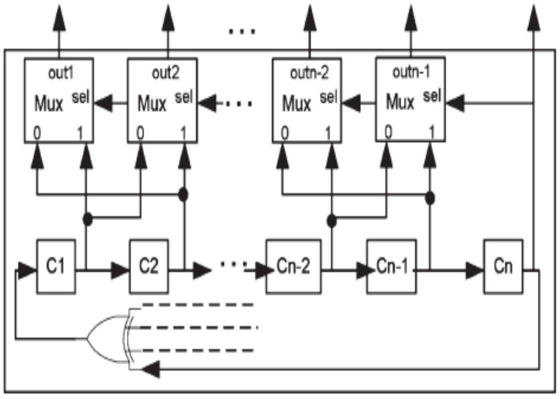


Fig. 3. BS-LFSR can be used to generate exhaustive patterns for test-per-clock BIST

This will reduce the overall average power and also the peak power that may arise while scanning out a captured response. The problem of the capture power (peak power in the test cycle) will be solved by using a novel algorithm that will reorder some cells in the scan chain in such a way that minimizes the Hamming distance between the applied test vector and the captured response in the test cycle, hence reducing the test cycle peak power (capture power).In this scan-chain-ordering algorithm, some cells of the ordered scan chain using the algorithm in [11] will be reordered again in order to reduce the peak power which may result during the test cycle. This phase mainly depends on an important property of the BS-LFSR. This property states that, if two cells are connected with each other, then the probability that they have the same value at any clock cycle is 0.75. (In a conventional LFSR where the transition probability is 0.5, two adjacent cells will have the same value in 50% of the clocks and different values in 50% of the clocks; for a BS-LFSR that reduces the number of transition of an LFSR by 50%, the transition probability is0.25, and hence, two adjacent cells will have the same value in 75% of the clock cycles.) Thus, for two connected cells (cells j and k), if we apply a sufficient number of test vectors to the CUT, then the values of cells j and k are similar in 75% of the applied vectors. Hence, assume that we have cell x which is a function of cells y and z.

If the value that cell x will have in the captured response is the same as its value in the applied test vector (i.e., no transition will happen for this cell in the test cycle) in the majority of cases where cells y and z have the same value, then we connect cells y and z together on the scan chain, since they will have the same value in 75% of the cases. This reduces the possibility that cell x will undergo a transition in the test cycle. The steps in this algorithm are as follows.

1) Simulate the CUT for the test patterns generated by the BS-LFSR.

2) Identify the group of vectors and responses that violate the peak power.

3) In these vectors, identify the cells that mostly change their values in the test cycle and cause the peak-power violation.

4) For each cell found in step 3), identify the cells that play the key role in the value of this cell in the test cycle.

5) If it is found that, when two cells have a similar value in the applied test vector, the concerned cell will most probably have no transition in the test cycle, then connect these cells together. If it is found that, when two cells have a different value, the cell under consideration

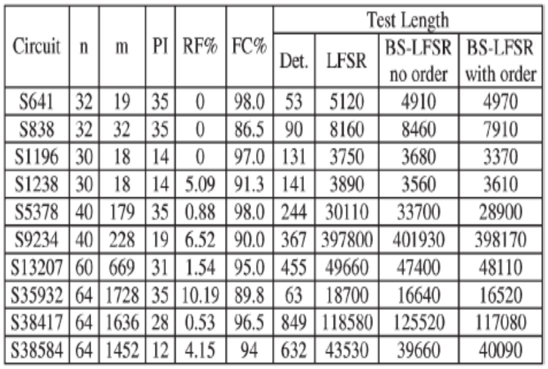
will most probably have no transitions in the test cycle, then connect these cells together through an inverter. It is important to note that this phase of ordering is done when necessary only, as stated in step 2 of the algorithm description that the group of test vectors that violates the peak power should be identified first. Hence, if no vector violates the peak power, then this phase will not be done. In the worst case, this phase is performed in few subsets of the cells. This is because, if this phase of ordering is done in all cells of the scan chain, then it will destroy the effect of algorithm found in [11] and will substantially increase the computation time.

V. EXPERIMENTAL RESULTS

A group of experiments was performed on full-scan ISCAS’89 benchmark circuits. In the first set of experiments, the BS-LFSR is evaluated regarding the length of the test sequence needed to achieve a certain fault coverage with and without the scan-chain-ordering algorithm. Table IV shows the results for a set of ten benchmark circuits. The columns labeled n, m, and PI refer to the sizes of the LFSR, the number of flip-flops in the scan chain, and the number of primary inputs of the CUT, respectively. The column labeled RF indicates the percentage of redundant faults in the CUT, and fault coverage (FC) indicates the target fault coverage where redundant faults are included. The last four columns show the test length needed by a deterministic test (i.e., the optimal test vector set is stored in a ROM), a conventional LFSR, a BS-LFSR with no scan-chain ordering, and the BS-LFSR with scan-chain ordering, respectively.

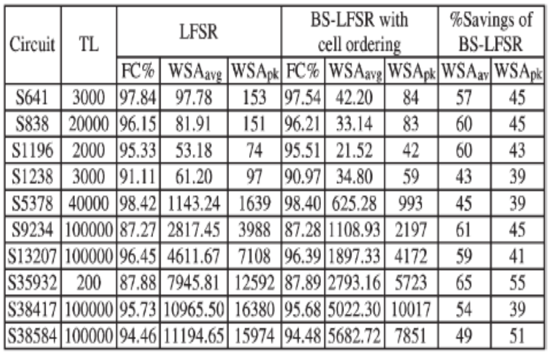
The results in Table IV show that the BS-LFSR needs a shorter test length than a conventional LFSR for many circuits even without using the scan-chain-ordering algorithm. It also shows that using the scan-chain-ordering algorithm with BS-LFSR will shorten the required test length. The second set of experiments is used to evaluate the BS-LFSR together with the proposed scan-chain-ordering algorithm in reducing average and peak power. For each benchmark circuit, the same numbers of conventional LFSR and BS-LFSR patterns are applied to the full scan configuration. Table V shows the obtained results for the same benchmark circuits as in Table IV. The column labeled test length (TL) refers to the number of test vectors applied to the CUT. The next three columns show the FC, average WSA per clock cycle(WSA_{avg}), and the maximum WSA in a clock cycle (WSA_{peak}) for patterns applied using the conventional LFSR.

TABLE IV
TEST LENGTH NEEDED TO GET TARGET FAULT COVERAGE FOR LFSR AND BS-LFSR



Circuit	n	m	PI	RF%	FC%	Test Length			
						Det.	LFSR	BS-LFSR no order	BS-LFSR with order
S641	32	19	35	0	98.0	53	5120	4910	4970
S838	32	32	35	0	86.5	90	8160	8460	7910
S1196	30	18	14	0	97.0	131	3750	3680	3370
S1238	30	18	14	5.09	91.3	141	3890	3560	3610
S5378	40	179	35	0.88	98.0	244	30110	33700	28900
S9234	40	228	19	6.52	90.0	367	397800	401930	398170
S13207	60	669	31	1.54	95.0	455	49660	47400	48110
S35932	64	1728	35	10.19	89.8	63	18700	16640	16520
S38417	64	1636	28	0.53	96.5	849	118580	125520	117080
S38584	64	1452	12	4.15	94	632	43530	39660	40090

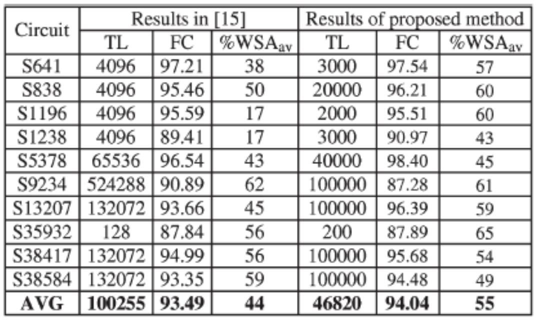
TABLE V
EXPERIMENTAL RESULTS OF AVERAGE- AND PEAK-POWER REDUCTION OBTAINED BY USING THE PROPOSED TECHNIQUES



Circuit	TL	LFSR			BS-LFSR with cell ordering			%Savings of BS-LFSR	
		FC%	WSA_{avg}	WSA_{pk}	FC%	WSA_{avg}	WSA_{pk}	WSA_{av}	WSA_{pk}
S641	3000	97.84	97.78	153	97.54	42.20	84	57	45
S838	20000	96.15	81.91	151	96.21	33.14	83	60	45
S1196	2000	95.33	53.18	74	95.51	21.52	42	60	43
S1238	3000	91.11	61.20	97	90.97	34.80	59	43	39
S5378	40000	98.42	1143.24	1639	98.40	625.28	993	45	39
S9234	100000	87.27	2817.45	3988	87.28	1108.93	2197	61	45
S13207	100000	96.45	4611.67	7108	96.39	1897.33	4172	59	41
S35932	200	87.88	7945.81	12592	87.89	2793.16	5723	65	55
S38417	100000	95.73	10965.50	16380	95.68	5022.30	10017	54	39
S38584	100000	94.46	11194.65	15974	94.48	5682.72	7851	49	51

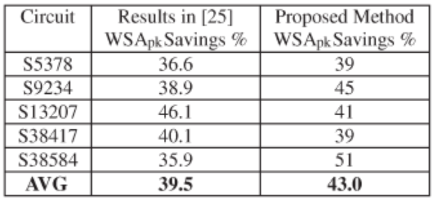
The next three columns show FC, WSA_{avg} , and WSA_{peak} for the BS-LFSR with ordered scan chain. Finally, the last two columns show the savings in average and peak power by using the BS-LFSR with the scan-chain-ordering algorithm. In order to provide a comparison with the techniques published previously by other authors, Table VI compares the results obtained by the proposed technique with those obtained in [15]. Table VI com-pares the TL, FC, and average-power reduction (WSA_{avg}). It is clear that the proposed method is much better for most of the circuits, not only in average-power reduction but also in the test length needed to obtain good fault coverage. Finally, Table VII compares the results obtained by the proposed technique for peak-power reduction with those obtained in [25]. It is clear from the table that the proposed method has better results for most of the benchmark circuits.

TABLE VI
COMPARISON WITH RESULTS OBTAINED IN [15]



Circuit	Results in [15]			Results of proposed method		
	TL	FC	%WSA$_{av}$	TL	FC	%WSA$_{av}$
S641	4096	97.21	38	3000	97.54	57
S838	4096	95.46	50	20000	96.21	60
S1196	4096	95.59	17	2000	95.51	60
S1238	4096	89.41	17	3000	90.97	43
S5378	65536	96.54	43	40000	98.40	45
S9234	524288	90.89	62	100000	87.28	61
S13207	132072	93.66	45	100000	96.39	59
S35932	128	87.84	56	200	87.89	65
S38417	132072	94.99	56	100000	95.68	54
S38584	132072	93.35	59	100000	94.48	49
AVG	**100255**	**93.49**	**44**	**46820**	**94.04**	**55**

TABLE VII
COMPARISON OF PEAK-POWER REDUCTIONS WITH RESULTS IN [25]



Circuit	Results in [25] WSA$_{pk}$ Savings %	Proposed Method WSA$_{pk}$ Savings %
S5378	36.6	39
S9234	38.9	45
S13207	46.1	41
S38417	40.1	39
S38584	35.9	51
AVG	**39.5**	**43.0**

VI. CONCLUSION

The proposed low-transition TPG is based on some observations about transition counts at the output sequence of LFSRs. Along with this TPG, a novel algorithm for scan-chain ordering has been presented. The proposed TPG is used to generate test vectors for test-per-scan BISTs in order to reduce the switching activity while scanning test vectors into the scan chain. The average and peak power are substantially reduced if the proposed scan-chain-ordering algorithm is used along with the BS-LFSR. The effects like fault coverage, test-application time, and hardware area are negligible in the proposed design. This proposed design can achieve better results for most tested benchmark circuits while this proposed design is compared with the previously published methods.

REFERENCES

[1] Y. Zorian, "A distributed BIST control scheme for complex VLSIdevices," in Proc. 11th IEEE VTS, Apr. 1993, pp. 4-9.

[2] A. Hertwig and H. J. Wunderlich, "Low power serial built-in self-test," in Proc. IEEE Eur. Test Workshop, May 1998, pp. 49-53.

[3] P. H. Bardell, W. H. McAnney, and J. Savir, Built-in Test for VLSI:Pseudorandom Techniques. New York: Wiley, 1997.

[4] P. Girard, "Survey of low-power testing of VLSI circuits," IEEE Des. TestComput., vol. 19, no. 3, pp. 80-90, May/Jun. 2002.

[5] K. M. Butler, J. Saxena, T. Fryars, G. Hetherington, A. Jain, and J. Lewis,"Minimizing power consumption in scan testing: Pattern generation And DFT techniques," in Proc. Int. Test Conf., 2004, pp. 355-364.
[6]J. Saxena, K. Butler, and L. Whetsel, "An analysis of power reductiontechniques in scan testing," in Proc.Int. Test Conf. 2001670-677.

[7] V. Dabhholkar, S. Chakravarty, I. Pomeranz, and S. M. Reddy, "Techniques for minimizing power dissipation in scan and combinational circuits during test applications," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 17, no. 12, pp. 1325-1333, Dec. 1998.

[8] Y. Bonhomme, P. Girard, L. Guiller, C. Landrault, S. Pravossoudovitch,and V. Virazel, "Design of routing-constrained low power scan chains," in Proc. Des. Autom. Test Eur. Conf. Exhib., Feb. 2004, pp. 62-67.

[9] W. Tseng, "Scan chain ordering technique for switching activity reductionduring scan test," Proc. Inst. Elect. Eng.—Comput. Digit. Tech., vol. 152, no. 5, pp. 609-617, Sep. 2005.

[10] C. Giri, B. Kumar, and S. Chattopadhyay, "Scan flip-flop ordering withdelay and power minimization during testing," in Proc. Annu. IEEEINDICON, Dec. 2005, pp. 467-471.

[11] Y. Bonhomme, P. Girard, C. Laundrault, and S. Pravossoudovitch, "Powerdriven chaining of flip-flops in scan architectures," in Proc. Int. Test Conf.,Oct. 2002, pp. 796-803.

[12] M. Bellos, D. Bakalis, and D. Nikolos, "Scan cell ordering for lowpower BIST," in Proc. IEEE Comput. Soc. Annu. Symp. VLSI, Feb. 2004,pp. 281-284.

[13] K. V.A. Reddy and S. Chattopadahyay, "An efficient algorithm to reducetest power consumption by scan cell and scan vector reordering," in Proc. IEEE 1st India Annu. Conf. INDICON, Dec. 2004, pp. 373-376.

[14] S. Wang, "A BIST TPG for low power dissipation and high fault cover-age," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 15, no. 7, pp. 777-789, Jul. 2007.

[15] S. Wang and S. Gupta, "LT-RTPG: A new test-per-scan BIST TPG for lowswitching activity," IEEE Trans. Comput.-Aided Design Integr. CircuitsSyst., vol. 25, no. 8, pp. 1565-1574, Aug. 2006.

[16] S. Wang and S. K. Gupta, "DS-LFSR: A BIST

TPG for low switchingactivity," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.,vol. 21, no. 7, pp. 842-851, Jul. 2002.

[17] H. Ronghui, L. Xiaowei, and G. Yunzhan, "A low power BIST TPGdesign," in Proc. 5th Int. Conf. ASIC, Oct. 2003, vol. 2, pp. 1136-1139.

[18] L. Jie, Y. Jun, L. Rui, and W. Chao, "A new BIST structure for low powertesting," in Proc. 5th Int. Conf. ASIC, Oct. 2003, vol. 2, pp. 1183-1185.

[19] M. Tehranipoor, M. Nourani, and N. Ahmed, "Low transition LFSRfor BIST-based applications," in Proc. 14th ATS, Dec. 2005, pp. 138-143.

[20] I. Pomeranz and S. M. Reddy, "Scan-BIST based on transition proba-bilities for circuits with single and multiple scan chains," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 25, no. 3, pp. 591-596, Mar. 2006.

[21] N. Nicolici and B. Al-Hashimi, "Multiple scan chains for power min-imization during test application in sequential circuits," IEEE Trans.Comput., vol. 51, no. 6, pp. 721-734, Jun. 2002.

[22] V. Iyengar and K. Chakrabarty, "Precedence-based, preemptive, andpower-constrained test scheduling for system-on-a-chip," in Proc. IEEEVLSI Test Symp., 2001, pp. 368-374.

[23] R. Chou, K. Saluja, and V. Agrawal, "Power constraint scheduling oftests," in Proc. IEEE Int. Conf. VLSI Des., 1994, pp.271-274.

[24] R. Sankaralingam, R. Oruganti, and N. Touba, "Static compaction tech-niques to control scan vector power dissipation," in Proc. IEEE VLSI Test Symp., 2000, pp. 35-42.

[25] S. Wang and W. Wei, "A technique to reduce peak current and averagepower dissipation in scan designs by limited capture," in Proc. Asia SouthPacific Des. Autom. Conf., Jan. 2007, pp. 810-816.

[26] N. Badereddine, P. Girard, S. Pravossoudovitch, C. Landrault,A. Virazel, and H. Wunderlich, "Minimizing peak power consumptionduring scan testing: Test pattern modification with X filling heuristics," in Proc. Des. Test Integr. Syst. Nanoscale Technol., 2006, pp. 359-364.

[27] R. Sankaralingam and N. Touba, "Controlling peak power during scantesting," in Proc. 20th IEEE VLSI Test Symp., 2002, pp. 153-159.

[28] S. Remersaro, X. Lin, S. M. Reddy, I. Pomeranz, and J. Rajski, "Low shift and capture power scan tests," in Proc. Int. Conf. VLSI Des., 2007, pp. 793-798.

[29] X. Wen, Y. Yamashita, S. Kajihara, L. Wang, K. Saluja, and K. Kinoshita, "On low-capture-power test generation for scantesting," in Proc. 23rdIEEE VLSI Test Symp., 2005, pp. 265-270.

[30] R. David, Random Testing of Digital Circuits, Theory and Applications. New York: Marcel Dekker, 1998.

[31] A. Abu-Issa and S. Quigley, "LT-PRPG: Power minimization tech-nique for test-per-scan BIST," in Proc. IEEE Int. Conf. DTIS, Mar. 2008,pp. 1-5.

[32] A. Abu-Issa and S. Quigley, "Bit-swapping LFSR for low-power BIST," Electron. Lett., vol. 44, no. 6, pp. 401-402, Mar. 2008.

◈◈◈

Power Optimization For A Datapath of A Genral Purpose Processor

Srinivasanaidu Nalla & Kalpana Tandle

Department of Electronics & Communication Engineering, Sri Sivani College of Engineering
Chilakapalem, Srikakulam, Andhra Pradesh, India

Abstract - Data path is the core of the processor; it is where all computations are performed. The other blocks in the processor are supporting units. At present, most of the popular processor hardware synthesis tools give higher priority to delay. So the processor synthesis tools tend to generate data path architecture for faster implementation. With increasing importance of power reduction on a processor, it is becoming necessary to evaluate different data path architectures from the point of view of both delay and power. This work is aimed at characterizing various architectures of common operators for power, delay and area and selecting a particular low power architecture where delay is not critical. Multiplier Architecture that consists of both Array and Tree Multipliers, Mixed style Multiplier was found to be a suitable Architecture for better Performance. Performance Evaluation was performed using Xilinx Power Analyzer and it was observed that the mixed style Multiplier gives an optimized power delay product.

Keywords*- data path, delay, power optimized multipler, bypassing technique, mixed style architecture.*

I. INTRODUCTION

Data path is the core of a processor; it is where all computations are performed. The other blocks in the processor are support units that either store the results produced by the data path or help to determine what will happen in the next cycle. A typical data path consists of an interconnection of basic combinational functions such as logic or arithmetic operators. Intermediate results are stored in registers.

Different strategies exist for the implementation of data paths-structure custom cells versus automated standard cells, or fixed hard-wired versus flexible field-programmable fabric. The choice of the implementation plat form is mostly influenced by the tradeoff between different design metrics such as area, speed, energy, design time and reusability. The control module determines what actions happen in the processor at any given point in time. A controller can be viewed as a finite state machine (FSM). The memory module serves as the centralized data storage area. The interconnect network joins the different processor modules to one another, while the input/output circuitry connects to the outside world. Unfortunately, the wires composing the interconnect network are less than ideal and present a capacitive, resistive and inductive load to a driving circuitry.

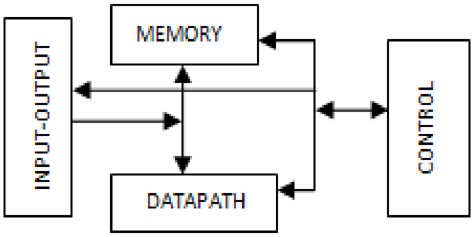


Fig. 1: A general purpose processor

The project is aimed at developing a power aware utility and ASIC flow of different operators such as adders and multipliers. The two main parts for the implementation are (i) Analyzing the power for different architectures and bring out the power time area relationship. (ii) Developing the utility that is having the minimum power for the desired frequency of operation. For analyzing the power for different architectures arithmetic operators and their word lengths are important. The selection of a low power operator architecture is done on the basis of a look-up table like consisting of name of architecture along with its power, area and delay for different bit widths.

Multiplication is an important arithmetic operation. Multipliers are much larger than adders and also more power consuming. Multiplication is actually a process of addition of multiple partial products. These partial products are formed by operating the multiplicand by each bit/bits of the multiplier. There are different multiplier architectures formed depending on the method of generation of partial product and its addition to finally give the complete product. Three important criteria to be considered in the design of multipliers are the chip area, speed of computation and power dissipation. Most advanced digital systems incorporate a parallel multiplication unit to carry out high-speed mathematical operations. The microprocessor requires multipliers in its ALU. Today high-speed parallel multipliers with much larger areas and higher complexity are used extensively, in RISC and graphic accelerators. Some examples of the parallel multiplier are the array multipliers Braun and Baugh-Wooley multipliers as well as the tree multipliers like Wallace tree multipliers, which has the advantage of a logarithmic circuit delay, and the array form, like the Carry-Save array, where the delay is linear. The advantage of the array multiplier is its regular structure, which leads to a dense layout, ideal for fabrication. Although tree multipliers are generally faster. The major drawback of these multipliers is the relatively larger chip area consumption. It presents high speed performance, but it is expensive in terms of silicon area and power consumption. This is because for parallel multipliers both operands are input to the multiplier in a parallel manner. As a result, the circuitry occupies a much larger area.

II. NORMAL AND BYPASS MULTIPLIERS

Low power issues have become an important factor in modern VLSI design, for both environmental reasons as well as to preserve battery life. However, different implementation technologies present different power optimization opportunities. In technologies above 0.1µm, dynamic power is the dominant contribution to the total power dissipated while in smaller technologies; leakage power is gaining more importance.

Dynamic power dissipation is the result of charging the load capacitances in a circuit. It is given by equation $P_d = C_L V_{DD}^2 E(sw) f_{clk}$, Where C_L is the output capacitance, V_{DD} The supply voltage, E(sw) (called switching activity) is the average number of transitions and f_{clk} The clock frequency.

Leakage power dissipation is divided in two major parts, the sub-threshold leakage and the gate-oxide leakage. The sub-threshold leakage is caused by short channel effects and low threshold voltage (Vth), while the gate-oxide leakage is exponentially increasing with decreasing oxide thickness. As in each new technology the supply voltage decreases to improve performance and dynamic power dissipation, this requires the threshold voltage being scaled down also. Unfortunately, sub-threshold leakage currents increase exponentially with decreasing threshold voltage.

The functionality of the carry save array multiplier is as follows .X= $(x_{n-1}, \ldots\ldots, x_1, x_0)$ and Y= $(y_{n-1}, \ldots\ldots y_1, y_0)$These inputs are fed to the array of FA* cells shown in figure 3(a), Each FA* cell performs $x_i * y_j$ using an AND gate and then adds the resultant with the incoming carry bits, to produce an output sum and output carry.

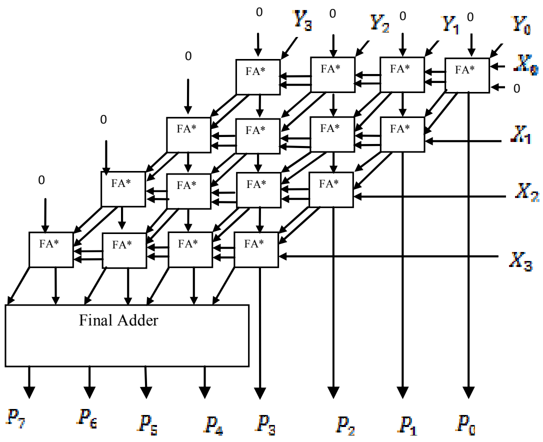
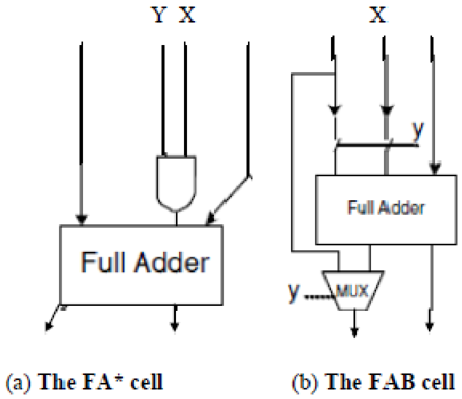


Fig. 2: A 4X4 carry save array multiplier



(a) The FA* cell

(b) The FAB cell

Fig. 3: Multiplier cells

This full adder cell is repetitively used in the multipliers operation. Multiplication is nothing but the repetitive addition of partial products. In general the array multipliers give the advantage of dynamic power savings. Due to the large no. of almost identical critical paths, increasing the performance of the structure through transistor sizing yields marginal benefits. A more efficient realization can be obtained by noticing that the multiplication result does not change when the output carry bits are bypassed diagonally downwards instead of only to the right. We include an extra adder called vector merging adder to generate a final result. The resulting multiplier is nothing but carry save multiplier, because carry bits are not immediately added, but rather are saved for the next stage. In the final stage carries and sums are merged in a fast carry propagate adder stage. While this structure has a slightly increased area cost, it has an advantage of that its worst case critical path and are uniquely defined.

All the FA* cells are placed appropriately as shown in the fig. 2 to perform multiplication. The final adder in the figure is used to merge the sums and carries from the last row of the every array, since in every row the carry bits are not immediately added but rather propagated to the row. There is drawback of the normal multiplier that is uses FA* cell that when y_j is zero the corresponding cells are functioning unnecessarily. So to block the operation a new cell structure that blocks unwanted diagonal. It uses a new architecture, bypass multiplier.

To minimize the power dissipation in digital multipliers, starting from dynamic power and then to switching activity. Minimization of switching activity in a digital circuit can be performed by isolating units producing non valuable partial products, in order to save power. The estimation and minimization of switching activity in combinational circuits by considering both and transitions 1 → 0 and 0 → 1. To minimize the switching activity, eliminating input inverters or increasing the number of inputs to a gate.

There have been lot of techniques to reduce the switching activity; the one of the most important technique is that clock gating. To save power, clock gating support adds more logic to a circuit to prune the clock tree, thus disabling portions of the circuitry so that its flip flops do not change state. Their switching power consumption goes to zero and only leakage currents are incurred. Here a new technique in the place of clock gating is used i.e. use of transmission gates.

FAB Cell

The main drawback of the normal carry save multiplier is that when y_j is 0 then the corresponding cells are functioning unnecessarily. If $x_i * y_j$ and carry input is zero then the sum output is nothing but sum input. Consequently, the sum output of the above cell can bypass the unimportant diagonal with the use of transmission gates. So, FA* cell can be replaced by FAB cell means full adder cell with bypassing technique as shown in fig. 3(b)

The operation of the FAB cell is that X and Y are multiplier and multiplicand bits, Sin and C_{in} are the sum input and carry input respectively. In figure 3(b) Y is used as control signal. Here the transmission gates are implemented using CMOS gates. In the FAB cell, if Y=0, the X and S_{in} inputs are blocked rather than that Sin input is send to the multiplexer and the same way if Y=0 the S_{in} input is send to the S_{out} terminal. Suppose if Y=1 then the sum output of the cell is propagated from the full adder. The carry input does not generate any new value since the diagonal carry input is always 0. If Y=0, to fix the erroneous carry generated from previous blocks, an AND gate is used before the final adder to make the final diagonal carry output to 0. This is the modified bypass multiplier that can be used in our mixed style multiplier architecture. Because of this bypassing technique the delay can be reduced with the area overhead

Bypass Multiplier

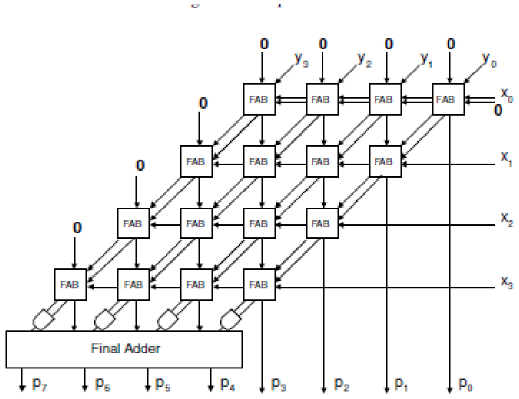


Fig. 4: The proposed 4x4 carry save array multiplier with bypass

All the FAB cells are placed appropriately as shown in the fig. 4 to perform multiplication. The final adder in the fig. 4 is used to merge the sums and carries from the last row of the every array, since in every row the carry bits are not immediately added but rather propagated to the row.

III. WALLACE TREE AND MIXED STYLE MULTIPLIER

Wallace tree multiplier

The partial sum adders can also be rearranged in a tree like fashion, reducing both the critical path and the number of adder cells needed. The number of full adders needed for this operation can be reduced by observing that only column 3 in the array has to add four bits. All other columns are somewhat less complex. To arrive at the minimal implementation, we iteratively cover the tree with FA's and HA's, starting from its densest part. The tree multiplier as shown in fig. 5 realizes substantial hardware savings for larger multipliers. The propagation delay is reduced as well. It can be shown that the propagation delay through the tree is equal to O ($\log_{\frac{3}{2}}(N)$). While substantially faster than the carry save structure for larger multiplier word lengths, The Wallace tree has the disadvantage of being very irregular, which complicates the task of coming up with dense layout. This irregularity visible even in the four bit implementation.

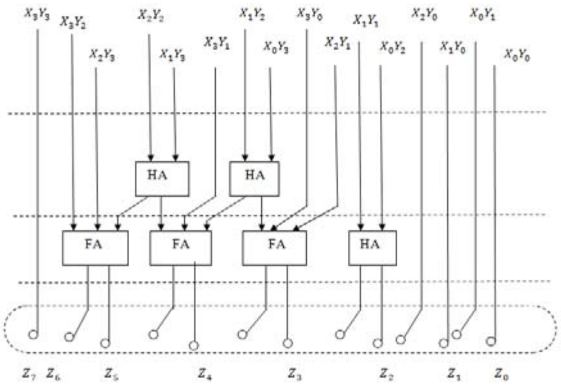


Fig. 5: Wallace tree for four bit multiplier

Mixed style multiplier

In general, array multipliers offers dynamic power saving but it fails to give a reduced area and fast speed advantages because of their regular interconnection pattern. So, tree multipliers are introduced to achieve reduced area and fast speed, to achieve both delay and power advantages it is better to use a different architecture that is having two parts one is tree based part and other is bypass architecture. The idea for this combination is given in figure 6. Two 8 bit values can be multiplied by splitting each one of them in two 4 bit parts. If the first 8 bit value is (A, B) and the second (C, D), four 8 bit partial products are generated, X = A × D, Y = B × D, Z = A × C and W = B × C. These four partial products shifted and added together generate the final 64 bit multiplication result.

The key point behind operand splitting is to use different multiplier architectures for each partial multiplication. In the example of figure 6, X = A × D and Z =A × C are performed in Wallace multipliers while the others in a bypass architecture. So, from X = A × D and Z =A × C performance is gained while from Y = B × D and W = B × C power is gained. If half of one or both operands usually contain more 0s than 1s, this specific half should be passed through the bypass multiplier for greater power improvements. For example, the architecture of figure 6 gives better results when B contains on average more 0s than 1s.

The modified bypass technique offers minimum dynamic power compared to normal carry save array multiplier. The tree based part of the mixed architecture has enough timing slack to be implemented using high threshold voltage components.

IV. FPGA IMPLEMENTATION

Overview of FPGA

FPGAs comprise an array of uncommitted circuit elements, called *logic blocks*, and interconnect resources. FPGA configuration is performed through programming by the end user. FPGAs have been responsible for a major shift in the way digital circuits are designed. FPGA generally consists of configurable logic blocks (CLBs), Input/output (I/O) blocks and programmable interconnections (Routing channels). FPGA based design has the advantages of re-programmability & on-board programming. It is highly suitable for high speed signal processing applications.

As one of the largest growing segments of the semiconductor industry, the FPGA market-place is volatile. There are two basic categories of FPGAs on the market today, SRAM-based FPGAs and Anti-fuse-based FPGAs. Xilinx and Altera are the leading manufacturers in terms of number of users, with the major competitor being AT&T. For Anti-fuse based products, Actel, Quick logic and Xilinx offer competing products.

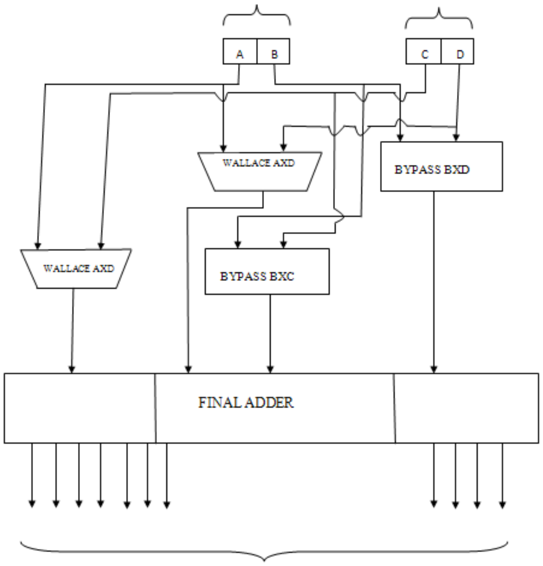


Fig. 6: 16-bit multiplication split in parts

Xilinx introduced the first FPGA family, called the XC2000 series, in about 1985 and now offers three more generations: XC3000, XC4000, and XC5000. Although the XC3000 devices are still widely used, we will focus on more popular XC4000 family. The Xilinx 4000 family devices range in capacity from about 2000 to more than 15,000 equivalent gates. The XC4000 features a logic block (called a *Configurable Logic Block* (CLB) by Xilinx) that is based on look-up tables (LUTs). A LUT is a small one bit wide memory array, where the address lines for the memory are inputs of the logic block and the one bit output from the memory is the LUT output.

Implementation results

All the algorithms discussed in multiplication are synthesized to Spartan 3E FPGA. Power dissipating were estimated using Xlinx "X Power analyzer". From the table it is clear that the mixed style architecture gives an optimized Power Delay Product.

Multiplier Topology	Bits	Switching Power Dissipation (mW)	Critical Path Delay (ns)	Area (in terms of LUTs)
Carry Save Array Multiplier	8x8	1.87	23.76	131
Mixed Style Multiplier	8x8	3.31	19.862	135
Wallace	4x4	2.67	13.088	29
Bypass	4x4	2.65	11.6	25

TABLE 1: BASIC ARCHITECTURES - POWER DISSIPATION, AREA AND TIME

CONCLUSION

In this paper, evaluation of different data path architectures from the point of view of both delay and power is done. First, the application of a bypass technique in the architecture of the Carry-Save array multiplier was evaluated. This bypass technique gives us significant gains in power consumption, making it ideal for implementation of power efficient DSP VLSI systems. However, performance is affected by the limitations of the Carry-Save architecture where we have to sacrifice logarithmic delay. By combining this architecture with the architecture of the traditional Wallace tree multiplier, bypassing only columns (or rows) with high probability to be 0, a mixed architecture can be constructed that offers significant power (dynamic and leakage) and timing savings.

REFERENCES

[1] Dimitris Bekiaris, George Economakos and Kiamal Pekmestzi, "Mixed Style Multiplier Architecture for Low Dynamic and Leakage Power Dissipation," in *International Symposium on Circuitsand Systems*. IEEE, 2010, pp. 258–261.

[2] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective - Third Edition*. Addison-Wesley, 2004

[3] J. Rabaey and M. Pedram, *Low Power Design Methodolies*. Kluwer Academic Publishers, 1996.

[4] G. Economakos and K. Anagnostopoulos, "Bit level architectural exploration technique for the design of low power multipliers," in *International Symposium on Circuits and Systems*. IEEE, 2006

[5] C. C. Wang and G. N. Sung, "Low-power multiplier design using a bypassing technique," *Journal of Signal Processing Systems*, 2008

[6] M. Karlsson, "A generalized carry-save adder array for digital signal processing," in *4th Nordic Signal Processing Symposium*. IEEE, 2000, pp. 287–290.

[7] P. Meier, R. A. Rutenbar, and L. R. Carley, "Exploring multiplier architecture and layout for low power," in *Custom Integrated Circuits Conference*. IEEE, 1996, pp. 513–516.

◈◈◈

Introduction to Virtual Prototype Modeling and General Design Considerations for its Efficiency

Varun S [a], Chethana G [b] , Surendra Mahala [c]

[a, b] Dept. of Electronics and Communication, R.V. College of Engineering, Bangalore, India
[c]IMC-WLS-RD-CD-VP, Intel Mobile Communications, India Pvt. Ltd, Bangalore, India

Abstract - The silicon world is not the same as it was a few decades earlier where a SoC included a simple processor, with minimal peripherals. It is very fast growing with increasing complexity where a fully functional SoC include complex processors, digital signal processors, multi-layer buses, multiple memories and other blocks that might have been separate ASICs in the past. So when we see the traditional approach of 'hardware-then-software' design flow which is a serial process is very time consuming and also costly. Development cost and time to market are the two critical issues which need to be addressed in the recent times, in the development cycle of any silicon based product. But the traditional design approach fails to do the same in both the cases. To achieve these challenges, system designers and embedded software developers came up with an idea of virtual system prototype. This paper briefly highlights the need for VP based flow vs. traditional approach. Requirements for a VP to be developed and how efficiently you can do considering the various aspects in the development flow.

Key Words*: Virtual Prototype, TLM.*

I. INTRODUCTION

The traditional approach to system development is to design the hardware (HW), make a physical prototype, write the code, and then integrate the hardware and software (SW). Because many projects target fast-moving markets, this approach is now too slow and too risky. A three-month delay on a product with a total lifetime of 30 months can reduce profits by 50% [1].

What many software teams want is to start developing the software as soon as possible in the project lifecycle. That means finding an alternative to the traditional 'hardware-then-software' design flow and getting started before the hardware is ready is shown in figure 1.

Increasingly, semiconductor companies are looking for ways to differentiate themselves beyond the feature sets of their chips.

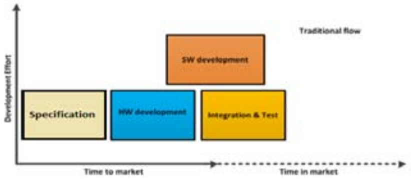


Figure 1: Traditional HW then SW embedded project design flow

By adopting virtual prototyping they can offer better support to their systems and customers. This approach can lead to a 9 to 12-month market advantage, which for a growing number of development teams is impossible to ignore [1]. And many find that the secondary benefits of virtual prototyping are making their teams so much more productive and efficient – for example, during debugging.

Virtual prototypes can be made available just a few weeks into the project schedule, which allows the software team to begin porting operating systems and developing device drivers without having to wait for the hardware team to write a single line of RTL code. Figure 2 shows how a "parallelized" flow leads to a significant schedule improvement.

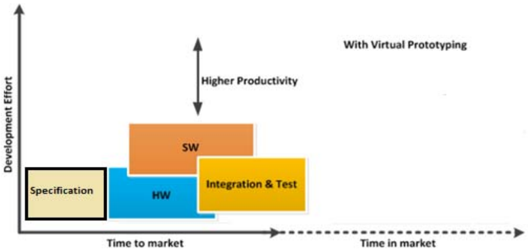


Figure 2: Parallelized HW and SW design flow using virtual prototypes

II. VP ABSTRACTIONS

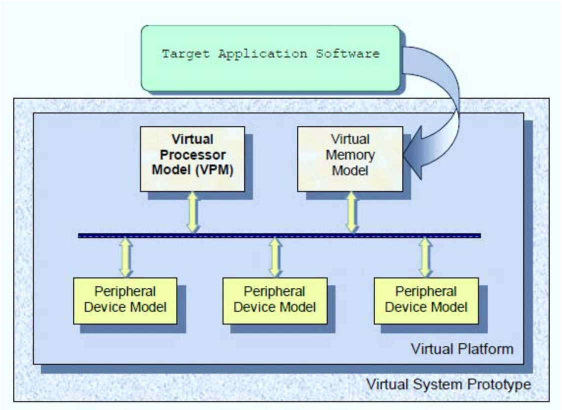


Figure 3: Virtual System Prototype

WHY ABSTRACTION?

- Simulations are event driven, where every event simulates stable state of hardware –Clock Edge, Interrupts, Signal transitions, State transitions.
- Every event involves computation on Native Host.
- RTL simulations cover all hardware states.
- For fast VP simulations, numbers of simulated events are reduced.
- Higher the abstraction, lesser the number of events, faster the simulation.

III. PRODUCT DEVELOPMENT WITH VIRTUAL PROTOTYPE

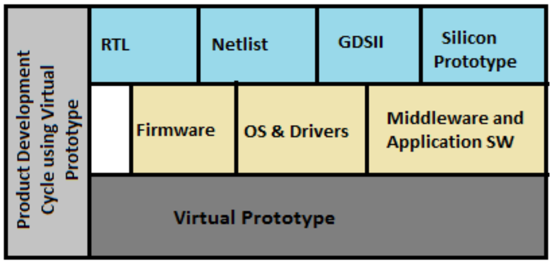

Product Development Cycle using Virtual Prototype
RTL
Netlist
GDSII
Silicon Prototype
Firmware
OS & Drivers
Middleware and Application SW
Virtual Prototype

Figure 4: Typical Product Development Cycle with Virtual Prototype

A virtual prototype is a fully functional software representation of a hardware design that encompasses virtually any combination of hardware. It provides an unambiguous executable specification of the system design, which developers can use to develop, integrate and test the software.

A virtual prototype runs on a general-purpose PC or workstation and is detailed enough to execute unmodified production code, including drivers, the operating system, middleware and applications at speeds close to or even faster than real-time. SW developers can refine the virtual prototype as the project progresses, for example, extending the virtual prototype based on the software design task. To support an initial operating system port will require the instruction set simulators (ISS) and a few key peripherals such as Universal Asynchronous Receiver/Transmitter (UART) and timer. The development team can add more complex peripherals to support the needs of the schedule.

Virtual prototypes should also integrate with the hardware verification and system validation flows. These flows use additional tools such as Field-programmable gate array (FPGA) prototypes, RTL simulation, physical system simulation, test bench equipment and even virtual IO that enable the virtual prototype to connect to physical hardware, such as USB or Ethernet.

IV. SYSTEMC & TLM 2.0 FOR VIRTUAL PROTOTYPING

A. *WHY SYSTEMC?*

SystemC is a C++ based modeling platform supporting design abstractions at the register-transfer, behavioral, and system levels. Consisting of a class library and a simulation kernel, the language is an attempt at standardization of a C/C++ design methodology, and is supported by the Open SystemC Initiative (OSCI), a consortium of a wide range of system houses, semiconductor companies, Intellectual property (IP) providers, embedded software developers, and design automation tool vendors. The advantages of SystemC include the establishment of a common design environment consisting of C++ libraries, models and tools, thereby setting up a foundation for hardware-software co-design; the ability to exchange IP easily and efficiently; and the ability to reuse test benches across different levels of modeling abstraction. All these features of SystemC make it an attractive language for design specification, verification, and synthesis at different levels of abstraction.

B. *WHY TRANSACTION LEVEL MODELING?*

Transaction-Level Model (TLM) is motivated by a number of practical reasons:

- Providing an early platform for software development.
- System level design exploration and verification.
- The need to use system level models in block level verification [6].

The TLM modeling level of the SystemC language emphasizes the transactions in a complex system, considered at a very high level of abstraction. COMPLEX hardware/software systems include increasingly sophisticated components; consequently, the demand for efficient verification solutions is growing [4]. In that context, TLM is perceptibly being adopted and has become the basis of advanced verification [5]. TLM (Transaction Level Modeling) is a modeling level of the SystemC language. Implemented as an extension of C++, SystemC provides a library of classes for modeling complex systems [6] and covers various levels of accuracy [7]: structural, timing, functional, etc. TLM provides communication models for complex data types ("transactions") between IPs or circuits and allows us to describe characteristics of the design at a much more abstract level than RTL. It greatly improves simulation performance and is therefore becoming a key factor for gaining time to market.

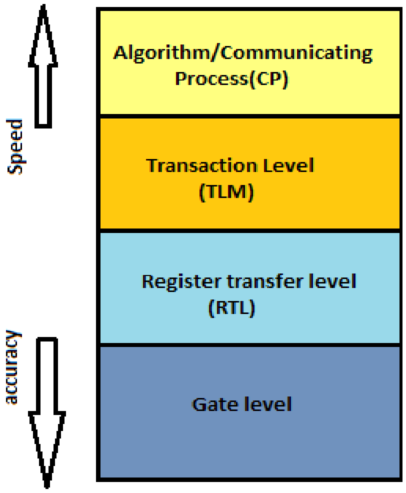

Speed
accuracy
Algorithm/Communicating Process(CP)
Transaction Level (TLM)
Register transfer level (RTL)
Gate level

Figure 5: TLM Abstraction Level Overview

Compared with the RTL-level prototypes, the TLM-level prototypes have the following advantages:

- Reduce details and simplify modeling.
- Migrate functional implementations between software and hardware efficiently.
- Hold higher simulation speed.
- Provide appropriate simulation accuracy for analysis.
- Transactions across interfaces are modeled as function calls (instead of individual signals).
- Abstract protocol and payload representation.
- Event based modeling with concurrent processes.
- Clocks are abstracted (>1000x faster than RTL).
- Possibility to model complete systems (VP).

C. TLM MODELING AND SIMULATION SCOPE

- Model consists of generated stub shell + functional kernel
- It is embedded into a test bench that is just another SystemC module.
- Model and test bench can be simulated as a compiled standalone binary without any additional tool.
- Field of application:

A. Component Development

B. Component Verification

- Component can later be integrated into larger system context

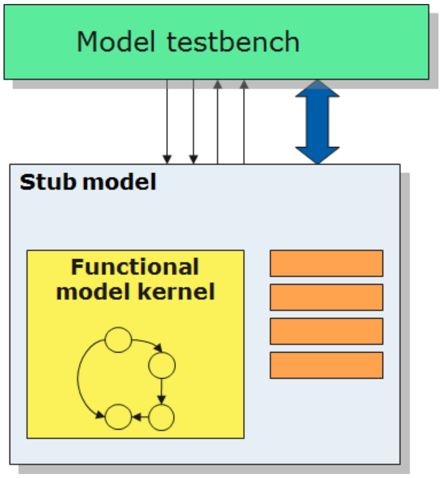


Figure 6: TLM Modeling Overview

D. TLM MODEL USAGE

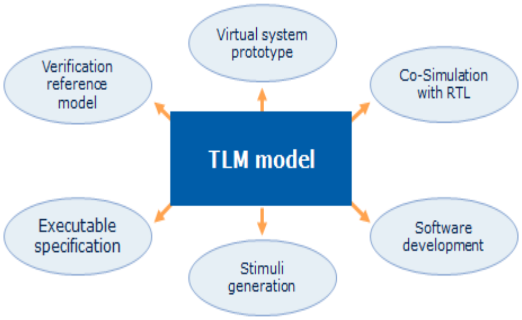


Figure 7: TLM Modeling Usage Options

V. DESIGN CONSIDERATIONS FOR EFFICIENCY

As we know that Development cycle of any SoC involve various stages like

- Concept/Architecture Definition
- Software/Firmware Development
- Development of Drivers
- Silicon Validation
- Development of Boot Code
- System Verification and so on

A. VP TARGET APPLICATION/USERS

i. Concept/Architecture definition team should maintain that VP is functional and timing accurate. They have to make sure that design meets the application requirements. It may be in terms of bus bandwidth, memory requirements, processor MIPs and so on.

ii. Software/Firmware development team: to start development of the Software/Firmware much ahead of Target board availability. This increase the efficiency of SW/FW debugging.

iii. Drivers/FW development team: approximately timed model should be sufficient maintaining functional accuracy and configurability. However the simulation speed of VP should be high (around 30-60 real time factor should be targeted).

iv. Application s/w team: loosely timed model and in some cases untimed is also sufficient while still keeping the functional accuracy.

v. Timing accuracy can be achieved by maintaining cycle accuracy at signal level or at boundary level, however maintaining timing accuracy at signal level will prove to be very costly in terms of simulation speed.

vi. It would be lot of effort and time consuming to develop different types of models for different needs, so it is essential to design and develop VP in such a way that it suits all the above target users.

B. CLOCKING

- By defining clock as sc_in <bool> clock; and if Methods/Threads are made sensitive to clock edges, in such cases simulation speed reduces.
- So proposed idea is to define clock as sc_in <sc_time> clock; where Input clock can be obtained from sc_time and can be directly read from the port which increases the simulation speed of the SystemC Models [9].

C. REUSABILITY OF COMPONENTS

Generally when we see any IP, there are several components which are common like clock controller, Interrupt Controller, Reset handler logic and many more. So by creating a library of all these components and making use across different Models can reduce the development effort to some extent [9].

VI. CONCLUSION

There is mounting pressure on embedded software teams. Increasingly, product success rests on their shoulders, and with software development time dominating the project, the software team cannot afford to wait for the hardware in order to get started.

Virtual prototypes offer software teams a way of creating software somewhat independently of physical hardware (silicon), which can result in a 9 to 12 months market advantage. Software teams can start developing code even before the hardware team has produced any RTL. Development teams can use this productivity advantage to get to market faster or allow for more testing time, including corner testing. As well as the time-to-market advantage, software teams are finding the other benefits of using virtual prototypes difficult to ignore.

Virtual prototypes offer excellent control and visibility for debugging – far better than the access available when debugging the hardware itself. Virtual models also enable globally distributed teams to work together – far more easily than when they have to share physical hardware.

ACKNOWLEDGEMENT

I would like to thank Musalimadugu Balasubrahmanyam (IMC-WLS-RD-CD-VP), Surendra Mahala (IMC-WLS-RD-CD-VP) for their technical guidance and giving an opportunity to carry out the work.

REFERENCES

[1] Marc Serughetti– Synopsis Inc. “Virtual Prototyping for Software Developers”, issue2, 2011.

[2] Preeti Ranjan Panda- Synopsys Inc. “SystemC - A modeling platform supporting multiple design abstract ions” *ISSS’Ol*, October 1-3, 2001, Montreal, Quebec, Canada.

[3] David C. Black, Jack Donovan, Bill Bunton, Anna Keist "SystemC From the ground up" 2nd edition, Springer publication.

[4] Liang Liang, Bo Zhou, Xue-Gong Zhou, and Cheng-Lian Peng "System Prototyping Based on SystemC TLM" Proc. IEEE Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06).

[5] R. Dubey, "Elements of Verification," SOC Central: white paper, Mar. 2005.

[6] M. Glasser, "Applying Transaction-Level Models for Design and Test benches," SOC Central, June 2006.

[7] IEEE Standard 1666-2005, IEEE Standard SystemC Language Reference Manual.

[8] T. Grotker, S. Liao, G. Martin, and S. Swan, "System Design with SystemC", Kluwer Academic, 2002.

[9] Praveen Kumar Kondugari, Aravinda Thimmapuram - Intel Mobile Communications Inc., "Design and Implementation Techniques for improving simulation speed of SystemC models" Indian SystemC User Group Conference, 9th Apr 2012,unpublished. http://www.iscug.in

◈◈◈

Design and Implementation of the ATM Operations using IRIS & Universal Subscriber Identification Modules

J. Lakshmi Narayana , A.kabirdas & N.V.G.Prasad
Electronics and Communication Engineering, Sasi institute of technology & Engineering, Tadepalligudem, India

Abstract - Automated teller machines (ATMs) are well known devices typically used by individuals to carry out a variety of personal and business financial transactions and/or banking functions. ATMs have become very popular with the general public for their availability and general user friendliness. ATMs are now found in many locations having a regular or high volume of consumer traffic. For example, ATMs are typically found in restaurants, supermarkets, Convenience stores, malls, schools, gas stations, hotels, work locations, banking centers, airports, entertainment establishments, transportation facilities and a myriad of other locations. In this proposed system we have created the new generation ATM machine which can be operator without the ATM card. By using this system ATM machine can be operator by using our SIM in the mobile phone. When we insert our SIM in the reader unit of the ATM machine it transfers the mobile to the server. In server we can collect the related information of the mobile number (i.e.) the users account details, their photo etc. the camera presented near the ATM machine will capture the users iris and compare it with the user iris in the server using MATLAB. Only when the iris matches it asks the pin number and further processing starts. Otherwise the process is terminated. So by using this system need of ATM card is completely eliminated we can operate the ATM machine by using our SIM itself. By using this system malfunctions can be avoided. Our transaction will be much secured. One more application can also be added in this system for helping the blind people. In the existing system all the transactions are done through keyboard only. It may be difficult for blind people so we can also add voice enunciator to indicate each and very process to the blind people. It that enables a visually and/or hearing impaired individual to conveniently and easily carry out financial transactions or banking functions. And if unauthenticated means it will send one mms to the corresponding person, after receive the message it will ask whether it want to continue, or not, if he will reply yes means only it will to the further process, if he will reply no means it give the alarm in the ATM section. Experts assert that the iris is the most data rich part of the body. Iris recognition devices can use 260 degrees of freedom making it significantly more accurate than fingerprint recognition. The iris is the colored part of the eye at the front of the lobe. Each iris is unique. A person's right and left iris are not the same. Iris recognition may not be effective in poorly lit or highly reflective lighting environments or when the user is wearing designer contact lenses or mirrored sunglasses.

I. INTRODUCTION

ATMs are typically available to consumers on a continuous basis such that consumers have the ability to carryout their ATM financial transactions and/or banking functions at any time of the day and on any day of the week.

Banks can use biometrics to authenticate employees, contractors, and customers for both physical and computer systems access. Biometrics applies statistical theory to biological characteristics for the purpose of determining similarities and differences. Biometric authentication involves a person claiming to be someone and comparing his or her current biometric data with a sample previously provided. Some financial institutions have already deployed automated biometric solutions and others are considering cost-effective solutions to improve authentication techniques. Additionally, in response to the September 11, 2001 terrorist attacks, many industries are looking more closely at improving authentication systems that may include biometrics solutions.

Banks have used biometric solutions for a long time. Historically, banks have relied on two types of non-automated biometric recognition: facial and handwriting. Facial recognition techniques require customers to provide photo identification for comparison. Handwriting recognition compares current customer signatures (e.g., on checks or receipts) to previously provided signatures (e.g., on signature cards or the back of credit card). Banks approve activities based upon a positive match or deny an activity (or requests additional information) if a positive match is not made.

As with any authentication method, successfully designing, implementing, and maintaining an automated biometrics authentication solution depends on more than technology. It also depends on appropriate policies, procedures, and controls. An effective authentication method should have customer acceptance, reliable performance, scalability to accommodate growth, and interoperability with existing systems and future plans.

Existing ATMs are convenient and easy to use for most consumers. Existing ATMs typically provide instructions on an ATM display screen that are read by a user to provide for interactive operation of the ATM. Having read the display screen instructions, a user is able to use and operate the ATM via data and information entered on a keypad. However the drawback in the existing system is that the user should carry their ATM card without fail. But in many cases we forget it. So only we designed a system which helps us to use the ATM machine without the ATM card.

RELIAB'ITY MODELING

Reliability is a probabilistic concept that has been applied by engineers in product development. Frankel defined reliability as "the probability that a system or component will perform its intended function for a specified period of time, under required conditions."[5] In order to determine the reliability of a system, the following steps should be performed.

1. Assign a probability of success to each component
2. Determine attributes' values of each component.
3. Ascertain the degree of interdependencies among Components.
4. Employ a mathematically combinatorial model for calculating overall system reliability by using the probabilities of components' successes and relationships among components.

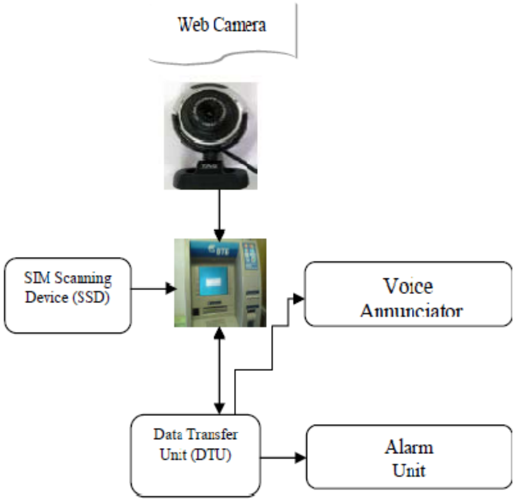


Fig : Design Implementation

OVERVIEW OF THE SYSTEM CONNECTION-LESS AUTHENTICATION SYSTEM:

A onetime password (OTP) is generated without connecting the client to the server. The mobile phone will act as a token and use certain factors unique to it among other factors to generate a one-time password locally. The server will have all the required factors including the ones unique to each mobile phone in order to generate the same password at the server side and compare it to the password submitted by the client. The client may submit the password online or through a device such as an ATM machine. A program will be installed on the client's mobile phone to generate the OTP.

SMS-BASED AUTHENTICATION SYSTEM: In case the first method fails to work, the password is rejected, or the client and server are out of sync, the mobile phone can request the one time password directly from the server without the need to generate the OTP locally on the mobile phone. In order for the server to verify the identity of the user, the mobile phone sends to the server, via an SMS message, information unique to the user. The server checks the SMS content and if correct, returns a randomly generated OTP to the mobile phone. The user will then have a given amount of time to use the OTP before it expires. Note that this method will require both the client and server to pay for the telecommunication charges of sending the SMS message**.**

A. STATISTICS:

The entire subject of statistics is based around the idea that you have this big set of data, and you want to analyse that set in terms of the relationships between the individual points in that data set. I am going to look at a few of the measures you can do on a set of data, and what they tell you about the data itself.

B. STANDARD DEVIATION:

To understand standard deviation, we need a data set. Statisticians are usually concerned with taking a *sample* of a *population*. To use election polls as an example, the population is all the people in the country, whereas a sample is a subset of the population that the statisticians measure. The great thing about statistics is that by only measuring (in this case by doing a phone survey or similar) a sample of the population, you can work out what is most likely to be the measurement if you used the entire population. In this statistics section, I am going to assume that our data sets are samples 2 of some bigger population. There is a reference later in this section pointing to more information about samples and populations. I could simply use the symbol to refer to this entire set of numbers. If I want to refer to an individual number in this data set, I will use subscripts on the symbol to indicate a specific number There are a number of things that we can calculate about a data set. For example, we can calculate the mean of the sample. I assume that the reader understands what the mean of a sample is, and will only give the formula: All this formula says is "Add up all the numbers and then divide by how many there are". Unfortunately, the mean doesn't tell us a lot about the data except for a sort of middle point. For example, these two data sets have exactly the same mean (10), but are obviously quite different: So what is different about these two sets? It is the *spread* of the data that is different. The Standard Deviation (SD) of a data set is a measure of how spread out the data is. How do we calculate it? The English definition of the SD is: "The average distance

C) Choosing components and forming a feature vector

In general, once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives you the components in order of significance. Now, if you like, you can decide to *ignore* the components of lesser significance. You do lose some information, but if the eigenvalues are small, you don't lose much. If you leave out some components, the final data set will have less dimensions than the original. To be precise, if you originally have _ dimensions in your data, and so you calculate _ eigenvectors and eigenvalues, and then you choose only the first { eigenvectors, then the final data set has only { dimensions.

What needs to be done now is you need to form a *feature vector*, which is just a fancy name for a matrix of vectors. This is constructed by taking the eigenvectors that you want to keep from the list of eigenvectors, and forming a matrix with these eigenvectors in the columns

Control Attributes

The Exploratory Study

To explore the validity of using a reliability model for the ATM internal control system, twelve users from an EDP auditing class (11 graduate students and the professor) participated. Each ser evaluated the hard copy case and then answered an associated questionnaire provided by the computerized decision aid. This decision aid converts the user assigned scale values (front-end conversion), computes the reliabilities of both control areas and the entire system based on model structures**,** and converts these reliabilities to Corresponding scale values (back-end conversion). For determining the effectiveness of the reliability model, the computed model values, the original user assigned values, and their differences are analyzed

E) IMAGE REGNITION:

Image recognition is composed of two parts: classification and validation. The classification can be done somewhat easily by statistics of dimensions and pattern features of each type of image. On the other hand, validation is very difficult because we cannot obtain counterfeits that might appear in future, while we can collect plenty of genuine images.[3] Moreover, statistics for a two-class (genuine and counterfeit banknotes) problem has less power because counterfeits could not actually be collected. Our approach is therefore to carefully select observation points at which a physical feature has a small deviation amongst genuine banknotes and looks difficult to imitate

F) WIRELESS COMMUNICATION:

A GSM modem is a wireless modem that works with a GSM wireless network. A wireless modem behaves like a dial-up modem. The main difference between them is that a dial-up modem sends and receives data through a fixed telephone line while a wireless modem sends and receives data through radio waves.A GSM modem can be an external device or a PC Card / PCMCIA Card. Typically, an external GSM modem is connected to a computer through a serial cable or a USB cable. A GSM modem in the form of a PC Card / PCMCIA Card is designed for use with a laptop computer.. As mentioned in earlier sections of this SMS tutorial, computers use AT commands to control modems. Both GSM modems and dial-up modems support a common set of standard AT commands. You can use a GSM modem just like a dial-up modem.

F) IRIS RECOGNITATION:

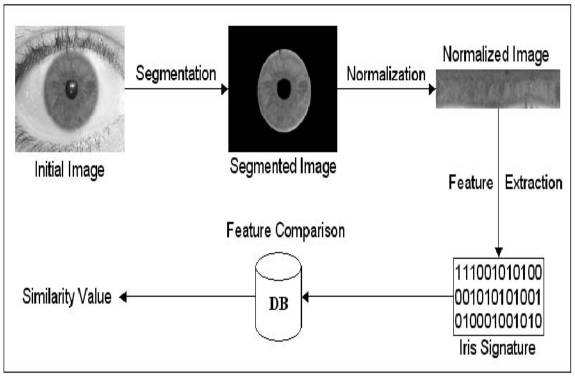


Fig: IRIS RECOGNITATION

USER ACCEPTANCE:

Because biometrics involve such personal information, privacy and security concerns may always be present even if user acceptance rates are high[2]. Early survey results show bank customers like automated biometric solutions. For example, an independent study found most customers had a favorable response to using iris recognition automatic teller machines (ATM).[1] The three fastest growing biometric solutions, in terms of sales, are fingerprints, voice, and dynamic signature recognition. Some experts project growth in the biometric-systems market from $400 million in 2000 to $1.4 billion by the year 2004.[2] [1] User acceptance of biometric solutions may be limited if users believe that the biometric readers will physically harm them or their personal characteristics can be recreated.

OPERATING PRINCIPLE:

An iris-recognition algorithm first has to identify the approximately concentric circular outer boundaries of the iris and the pupil in a photo of an eye. The set of pixels covering only the iris is then transformed into a bit pattern that preserves the information that is essential for a statistically meaningful comparison between two iris images. The mathematical methods used resemble those of modern lossy compression algorithms for photographic images. In the case of Daugman's algorithms, a Gabor wavelet transform is used in order to extract the spatial frequency range that contains a good best signal-to-noise ratio considering the focus quality of available cameras. The result are a set of complex numbers that carry local amplitude and phase information for the iris image. In Daugman's algorithms, all amplitude information is discarded, and the resulting 2048 bits that represent an iris consist only of the complex sign bits of the Gabor-domain representation of the iris image. Discarding the amplitude information ensures that the template remains largely unaffected by changes in illumination and virtually negligibly by iris color, which contributes significantly to the long-term stability of the biometric template. To authenticate via identification (one-to many template matching) or verification (one-to one template matching) a template created by imaging the iris, is compared to a stored value template in a database. If the Hamming distance is below the decision threshold, a positive identification has effectively been made.

A practical problem of iris recognition is that the iris is usually partially covered by eye lids and eyelashes. In order to reduce the false-reject risk in such cases, additional algorithms are needed to identify the locations of eye lids and eyelashes, and exclude the bits in the resulting code from the comparison operation.

A graphic, entitled "The Challenge of Rapid Iris Capture", illustrates the movement of the eyelids and eyelashes, as well as the contemporaneous movements of the eyeball within the orbit.

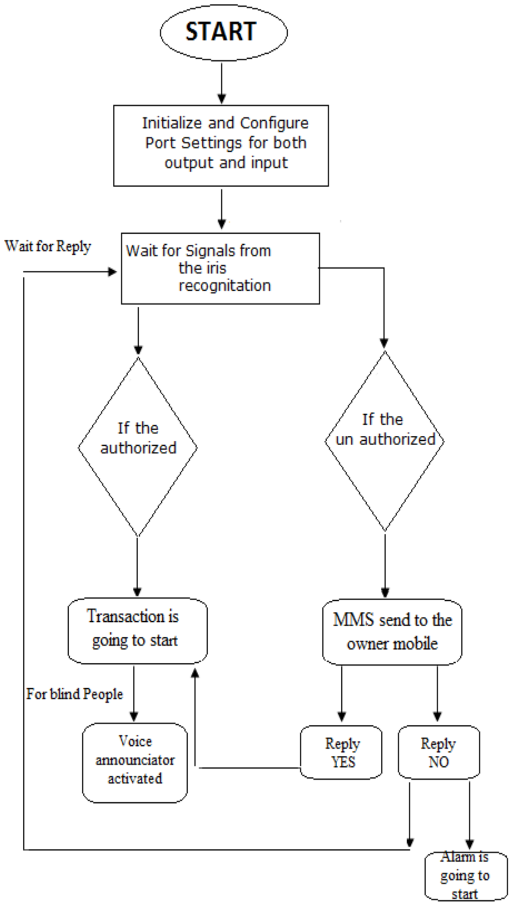

START
Initialize and Configure Port Settings for both output and input
Wait for Reply
Wait for Signals from the iris recognitation
If the authorized
If the un authorized
Transaction is going to start
MMS send to the owner mobile
For blind People
Voice announciator activated
Reply YES
Reply NO
Alarm is going to start

Fig: Flow Chart

USE OF THE MODEL:

Given the increasing complexity and the changing technologies of the ATM environment, this reliability model may be used to support the decision made by practitioners, to facilitate the learning process of students compare it **to** the model result **to** support or reevaluate his/her decision making. For designing a new ATM system, this tools *can* also be employed to simulate different ATM control systems to explore the best trade-off among system security requirements, budget provided, and technologies available. Based on this model, a computerized decision aid (e.g., a decision support system or an expert system) may be developed for facilitating the user's decision making. The systems designer or the auditor may use this computerized aid for computing overall system reliability and two area reliabilities, performing related analyses (e.g., what if, goal seeking, or sensitivity analysis), and retrieving explanation of reliability modeling, model structures, or ATM terms. The effectiveness of different computerized decision aids may also be examined in future research. In order to validate the potential usefulness, the evaluation of this model aiding human judgments is necessary. A small sample of experienced users can be asked to evaluate the ATM controls of a case without and then with this decision aid. A verbal protocol may be used to collect information used and considerations performed in the decision making process. An empirical study using more subjects can be conducted **to** test the significance of the model. Consensus and consistency measures can also be used to compare the effectiveness of this model versus human judgments. Consensus means the extent of agreement/disagreement

ADVANTAGES:

An iris scan is similar to taking a photograph and can be performed from about 10 cm to a few meters away. There is no need for the person to be identified to touch any equipment that has recently been touched by a stranger, thereby eliminating an objection that has been raised in some cultures against finger-print scanners, where a finger has to touch a surface, or retinal scanning, where the eye can be brought very close to a lens (like looking into a microscope lens).

Some argue that a focused digital photograph with an iris diameter of about 200 pixels contains much more long-term stable information than a fingerprint. The originally commercially deployed iris recognition algorithm, John Daugman's IrisCode, has an unprecedented false match rate (better than 10−11). While there are some medical and surgical procedures that can affect the colour and overall shape of the iris, the fine texture remains remarkably stable over many decades. Some iris identifications have succeeded over a period of about 30 year

CONCLUSION:

All the functions of the atm, the authors are now concentrating on developing the intention recognition, mobile based processing and alert module.

This paper presents a novel architecture that can be used as a means of interaction between mobile phone, ATM machine and a Banking application for the purpose of withdrawing cash. The proposed design; the secure M-cash withdrawal allows the use of mobile phones as a tool of interaction and provide flexibility through a robust identity management architecture. The first part of the architecture is the process of being implemented and all the process involved has been analyzed and justified where possible

REFERENCES:

[1] P. J. Phillips, A. Martin C. L. Wilson and M. Przybocki,“An Introduction to Evaluating Biometric Systems,” *IEEE Computer*, Vol.33, No.2, Feb. 2000, pp. 56-63.

[2] S. Pankanti, R. M. Bolle and A. Jain, “Biometrics: The Future of Identification,” *IEEE Computer*, Vol.33, No.2, Feb. 2000, pp. 46-49.

[3] M. Ejiri, T. Uno, M. Mese and S. Ikeda, “A Process for Detecting Defects in Complicated Patterns,” *Computer Graphics and Image Processing*, Vol.2, No.3-4, 1973, pp.326-339.

[4] S. Kashioka, M. Ejiri and Y. Sakamoto, “A Transistor Wire-Bonding System Utilizing Multiple Local Pattern Matching Techniques,” *IEEE Trans. on SMC*, Vol.SMC-6, No.8, Oct. 1977, pp. 562-570.

[5] H. Yoda, Y. Ohuchi, Y. Taniguchi and M. Ejiri, “An Automatic Wafer Inspection System Using Pipelined Image Processing Techniques,” *IEEE Trans. on PAMI*, Vol. PAMI-10, No.10, 1988, pp.4-16.

◈◈◈

Underground Wireless Communication Channel Modeling And Simulation Using 0.3GMSK Modulation Considering Penetration And Scattering Loss

M.N.Jayaram & C.R.Venugopal,

Department of E&C , SJCE , Mysore570006 , Karnataka , India

Abstract - Wireless communication inside mines and tunnels is very different from that in terrestrial environment because of the strong attenuation of signals. Here, we are developing an empirical model for the underground wireless communication channel based on experimental data which help in predicting the average received signal strength at a given distance from transmitter. The model aims at adding correction factors to the available outdoor and indoor propagation models such as Okumara-hata model, cost231 model, ITU indoor propagation models etc. Modeling is done by choosing the most appropriate model among the available ones and performing regression methods to the model based on experimental data. Correction factors are then added based on two parameters which we are considering namely- Penetration and Scattering loss for 0.3GMSK.

I. INTRODUCTION

Uunderground mines which are characterized by their tough working conditions and hazardous environments require fool-proof mine-wide communication systems for smooth functioning of mine workings and ensuring better safety. Proper and reliable communication systems not only save the machine breakdown time but also help in immediate passing of messages from the vicinity of underground working area to the surface for day-to-day normal mining operations as well as for speedy rescue operations in case of disaster. Therefore, a reliable and effective communication system is an essential requisite for safe working, and maintaining requisite production and productivity of underground mines. Most of the existing systems generally available in underground mines are based on line (wired) communication principle, hence these are unable to withstand in the disaster conditions and difficult to deploy in inaccessible places. Therefore, wireless communication is an indispensable, reliable, and convenient system and essential in case of day-to-day normal duty or disaster situations. The wireless communication systems used on surface cannot be applied straightaway in underground mines due to high attenuation of radio waves in underground strata, besides presence of inflammable gases and hazardous environment. Non symmetric mine topology, uneven mine structure and complex geological structures put further hindrance on the way of communication. Wireless communication in underground mine is a very complex technique .We have used simple modulations like BFSK ,BPSK &0.3GMSK .

II. AVILABLE PRACTCAL DATA , CALCULATION AND RESULTS

1. *Terrain Profile*

Underground mines consist of slopes, rocks, edges etc. This peculiar type of confined environment is characterized by very rough surfaces and a frequent absence of a line-of-sight between transmitting and receiving antennas. The resulting propagation characteristics differ from those frequently encountered in more typical indoor environments such as offices and corridors. This type of terrain profile limits the signal strength which makes it difficult for wireless communication. An approximate model of the underground channel taking such a terrain into consideration is done. Wireless communication in an environment which consists of soil or rock is more challenging than through air.

The approximate terrain profile inside mine is shown in figure 1.1

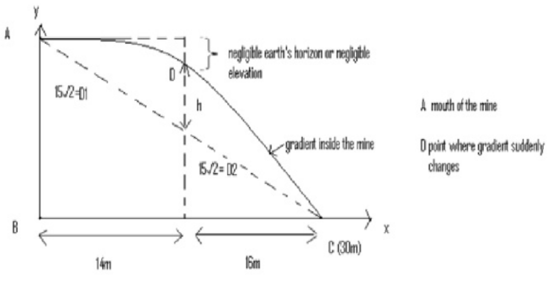

y
A
negligible earth's horizon or negligible elevation
D
15√2=D1
h
A mouth of the mine
gradient inside the mine
D point where gradient suddenly changes
15√2= D2
B
x
C (30m)
14m
16m

Figure 1.1: Approximate terrain elevation inside mine

Power measurements are done at distances of 1m (point A), 14m (point D) and 30m (point C).These points are selected because these are critical points.

2. *Available Practical Data*

Measurements carried out in a gold mine using a high performance transceiver . probable losses that occurs at points A,D&C are listed in table 1.1 .

Table 1.1: The prominent losses at points A,D & C in the mine

Points	Losses considered during propagation
A	Penetration loss+ machinery loss+ power line interference loss
D	Penetration loss+ multi path loss+ power line interference loss.
C	Penetration loss+ multi path loss+ power line interference loss+ bending loss.

The measured values of the total loss(transmitted signal strength – measured signal strength in dB) at A,D and C point considering three modulation formats BPSK, BFSK and 0.3-GMSK are given in Table 1.2:

Table 1.2: The measured values of losses at points A, D & C

Points	BPSK	BFSK	0.3GMSK
A	119.7185	119.7185	119.7185
D	137.4461	134.7918	132.5873
C	150.2984	147.9036	146.1967

3. *Modeling.*

Three basic models which are used in terrestrial comunication are considered .

3.1 *Selection of model*

We consider 3 propagation models and then use one of them as the base for modeling.

- **<u>Cost231 model</u>**

Mathematical equation is given by equation(1.1)

$$L=46.3+33.9\log f-13.82\ \log(h_B)-\ a(h_R)+(44.9-6.55\log h_B) \quad \ldots\ldots\ldots\ (1,1)$$

Where $a(h_R)=(1.1\log f-0.7)(h_R)-(1.56\log f-0.8)$, L=Median path loss in dB, f=Frequency of transmission in MHz, h_B =Base station antenna height in meter, d=Link distance in Km, h_R =Mobile station antenna height in meter and $a(h_R)$=Mobile station antenna correction factor.

In the GSM 900 MHz, receiving antenna height h_R =1m ,transmitting antenna height h_B = 50m and considering distance d in m, the losses at points A,D and C are given in Table:1.3

Table 1.3: Losses obtained from cost 231 model at A,D & C

Points	Loss(dB)
A(d=1m)	124.877
D(d=14m)	163.58
C(d=30m)	174.562

- <u>ITU model for indoor propagation</u>

Mathematical equation is given by

$$L = 20\log f + N\log d + Pf(n) - 28 \text{ --------------------------------- } (1.2)$$

Where L= Total path loss in dB, f= Frequency of transmission in MHz, d= Distance in meter, N= Distance power loss coefficient, n= Number of floors between the transmitter and receiver, Pf(n)=Floor loss penetration factor. In GSM 900MHzband , floor attenuation factor Pf(n)=24, no of floors N=33 and considering distance d in m, the losses at points A,D and C are given in Table:1.4

Table 1.4: Losses obtained from ITU model at A,D & C

Points	Loss(dB)
A(d=1m)	55.462
D(d=14m)	93.285
C(d=30m)	104.2076

- Log distance path loss model

Mathematical equation is given by

L= 20 log (4*π*d/ λ) + 10* γ *log(d/d $_0$) + X $_g$ ------------(1.3) Where λ=Wavelength, d=Distance in meter, X $_g$ =Fading factor, γ=Path loss exponent, d $_0$ =Reference distance. For Wave length in GSM 900MHz band , reference distance d $_0$ =1m, fading factor X $_g$ =0,path loss exponent γ =2, the losses at points A,D and C are given in Table:1.5

Table 1.5: Losses obtained from Log distance path loss model at A,D & C

Points	Loss(dB)
A(d=1m)	55.462
D(d=14m)	93.285
C(d=30m)	104.2076

The best model for further calculations is found by comparing the measured values of loss with the losses obtained from model equations.

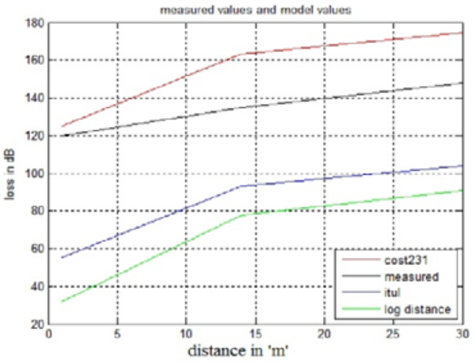


Figure 1.2 Measured & model path loss plot

3.2 Deviation of measured loss from model losses.

The deviation between the two losses are given in Table1.6

Table 1.6: The mean error from the differences of losses calculated from the models and measured losses

Models	Mean error
Cost231 model	20.202
ITU model	49.82
Log distance path loss model	67.253

Since Cost231 model has the lowest mean error. Hence, it is considered for all further calculations.

3.3 Calculation of correction factor to the selected model.

In the first step, correction factors w.r.t distance are added to the model using differences between model calculated and measured values.

Table 1.7: Differences between losses obtained from cost231 model and measured losses for 0.3GMSK modulation.

Points	Loss(dB)
A(d=1m)	5.159
D(d=14m)	30.993
C(d=30m)	28.57

First a plot of the differences vs. log d is drawn as in Fig 1.3 below .

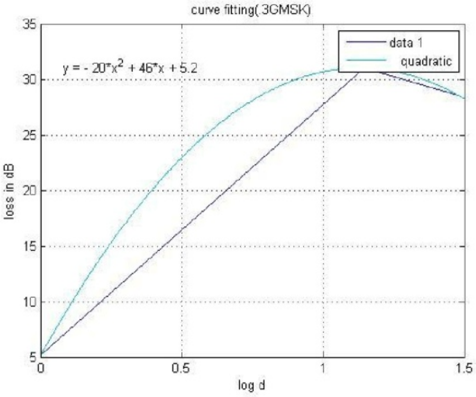


Fig 1. 3 loss vs logd curve

Polynomial regression is then performed by using the basic fitting tool in mat lab to obtain correction factor.

Thus, we obtain a correction factor

C1=5.2+41(log $_{10}$ d)-18(log $_{10}$ d) (1.4)

Incorporating in eq(1.1) ,

we have

L=46.3+33.9logf-13.82log(h $_B$)-a(h $_R$)+(44.9-6.55logh $_B$)

-C1 -- (1.5)

Figure 1.4 shows the corrected cost231 model w.r.t distance

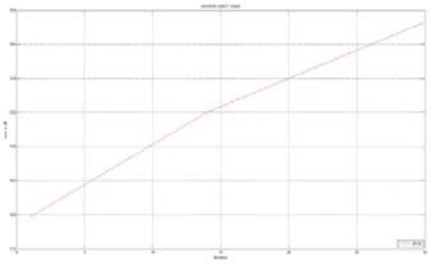


Figure 1.4: Plot of corrected cost231 model vs. distance

III. ANALYSIS OF PARAMETERS

Further corrections in the model are obtained by considering two parameters: penetration & scattering losses. Multi path loss is assumed to be 40 dB , using Fresnel's model bending loss 34.63 dB & power line loss is 57.9 dB (measured using magnetic dipole) .

1. Penetration loss.

There was a 35 cm thick concrete brick wall at the entrance which results in penetration loss.

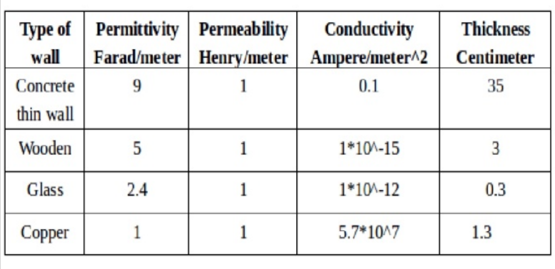
For lossy dielectric medium

$\gamma^2 = (\alpha + j\beta)^2 = (\sigma + j\omega\epsilon)(j\omega\mu)$………...(1.6)

So Attenuation constant $\alpha \approx \sigma/2\sqrt{(\mu/\epsilon)}$

Dielectric constants at 900MHz GSM band are listed in table1.8.For concrete this loss is 16 dB .

Table 1.8: Value of constants.



Type of wall	Permittivity Farad/meter	Permeability Henry/meter	Conductivity Ampere/meter^2	Thickness Centimeter
Concrete thin wall	9	1	0.1	35
Wooden	5	1	1*10\^-15	3
Glass	2.4	1	1*10\^-12	0.3
Copper	1	1	5.7*10^7	1.3

Now at point C the losses considered are:

Penetration loss+ multipath loss+ bending loss+ power line interference loss, We take penetration loss to be variable , let it be given by some term k3 ,hence the total loss becomes:

L=K3 +40+34.63+57.9

The correction factors are derived using the basic fitting tools in MATLAB. Thus, we obtain a second correction factor $C2=10.58\log_{10}f-173$

Adding it to the already corrected cost231 model we get the final cost231 model as,

$L=46.3+33.9\log f-13.82 \log(h_B)- a(h_R)+(44.9-6.55\log h_B) \log d - C1-C2$------(1.7)

Scattering loss:

At point C total loss is

L=Penetration loss+ multipath loss+ power line interference loss.

Considering multipath loss (MPC) as variable we have L= 16+40+57.9+MPC

Using polynomial regression method we use MATLAB for curve fitting and find the third correction factor as

C3=-0.01*d^2+0.45*d-40

Adding it to already corrected Cost 231 model we get the model after scattering correction as:

$L=46.3+33.9\log f-13.82 \log(h_B)- a(h_R)+(44.9-6.55\log h_B) \log d - C1-C2-C3.$

The final plot of corrected model is as in Fig 1.4

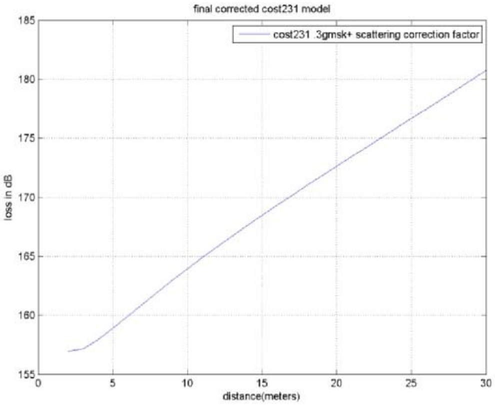


Figure 1.4: Shows the Plot of final corrected cost231 model.

III. CONCLUSION

1. Point C (at a depth of 30m) is selected because it is the critical point in the terrain .
2. 0.3 GMSK modulation is used .
3. High power of is used to overcome penetration loss at entrance .
4. Beyond 30m as more than 75% of the received signal was attenuated measurements was not possible.
5. We have used super position while modeling i.e for penetration loss we have assumed scattering loss as constant 7 vice versa .

6. Model can be updated by considering antenna height loss (due different planes of TX &RX) , bending loss, low frequency interference loss as variables .
7. Parameters like loss due to fan , blowers , lighting , ground reflection can be considered .
8. This model may not work in other mines due to change in geometry.

REFERENCES

1. Theodore.S. Rappaport, Wireless communications, Principle and practice, Eastern Economy Edition, 2^{nd} ed, 2002.
2. William C.Y.Lee, Mobile communications engineering, Theory and applications
3. Gilles, Y. Delisle. School of information technology & engg, University of Ottawa, Canada Underground mine wireless propagation modeling
4. Signal propagation techniques in underground mines. http://www.ece.gatech.edu/research/labs/bwn/surveys/underground09.pdf
5. Helchel. IEEE transaction on antenna & propagation, 'Comparison of 900 MHz band and 800 MHz band indoor propagation 'vol 54, No 12, Dec 2006
6. Shi Lihua, Xu Qiwei, Chan Bin Gao Cheng Measurement of the frequency dependent Dielectric constant of concrete materials by TDR & wavelet modeling method Asia pacific Conference on environmental electromagnetic. CEEM 2003, Nov 4 – 7.
7. Hindawi Publishing Corporation , International Journal of Antennas and Propagation , Volume 2008, Article ID 806326, 10 pages,doi:10.1155/2008/806326
8. Raj Jain, "Channel models a tutorial, Feb 1,2007"
9. About Modulation Format. http://en.wikipedia.org/ wiki/Modulation
10. C. Nerguizian, C. L. Despins, S. Aff`es, and M. Djadel, "Radiochannel,characterization of an underground mine at 2.4GHz,",IEEE Transactions on Wireless Communications, vol. 4, no. 5,pp. 2441–2453, 2005.
11. Zhi, S. and I. F. Akyildiz, \Channel modeling of wireless net-works in tunnels," IEEE Global Telecommunications Conference,(GLOBECOM'08), New Orleans, LA, Nov. 30{Dec. 4, 2008).
12. Empirical formula for propagation loss in land mobile radio service, IEEE transactions on Vehicular Technology, VOL. VT-29, NO.3, AUGUST 1980
13. About Radio Propagation Models. http://en.wikipedia.org/wiki/Radio_ propagation model
14. About Path Loss. http://en.wikipedia.org/wiki/Path_loss
15. Bertoni, H. L., Radio Propagation for Modern Wireless Systems Prentice Hall, Englewood Cli®s, NJ, 2000.
16. M. Boutin, A.Benzakour, C.Despins and S. Affes. (2008). Radio wave characterization and modeling in underground mine tunnels, IEEE. Transaction. on Antennas and Propagation, 2:540-549.
17. M.Lienard and P.Degauque. (2000). Natural Wave Propagation in Mine Environments, IEEE Transaction on Antennas and propagation, 48(9): 1326-1339.
18. W.E.Pakala and V.L.Chartier Radio noise measurements on overhead power lines from 2.4to 800kv,
19. M.Ndoh and G.Y.Delisle. (2004). Underground mines wireless propagation modeling, 60th IEEE Vehicular Technology Conference, Los Angeles, CA, 3584-3588.
20. Kjeldsen, E. and M. Hopkins, \An experimental look at RF prop-agation in narrow tunnels," Proc. IEEE Military CommunicationsConf. (MILCOM'06), Washington D.C., Oct. 23{25, 2006
21. Li Wenfeng, Lv Yingli, Li Baiping. (2008). Experiment on characteristic of radio propagation in mine. Journal of Xi an University of Science and Technology, 2:327-330 (in Chinese

◈◈◈

Implementation and Comparison of Wavelet Transform and Fourier Transform in Wi-max OFDM System

P. Mahesh Kumar & K. Sarada

ECE Department , Narayana Engineering College. Nellore. INDIA

Abstract - WIMAX wireless communication is based on ofdm technology which enables going towards 4G Based on IEEE 802.16d-2004. But the reliability of ofdm is limited with the problems of time varying nature of the channel. This can reduced by adding cyclic prefix or guard interval between each block of data symbols. This guard interval minimizes the spectral efficiency in ofdm system. Recently, it was proved that discrete wavelet transform in ofdm system will reduce the isi Inter carrier interference, which is produced by loss of ortho-gonality between carriers and also provides more spectrum efficiency. It is found that proposed wavelet design achieves much lower bit error rates, Increases signal to noise power ratio (SNR), and can be used as an alternative to the Conventional OFDM WIMAX. The proposed OFDM system was modeled tested, and its Performance was found under different International Telecommunications Union (ITU Channel models.

Keywords*: WIMAX, SFF SDR, OFDM, RS, Coding, AWGN, DWT, FFT.*

I. INTRODUCTION

The Wi-max is the standard harmonic for wide variety of Broadband wireless access (BWA) technologies. In the present paper we compare the performance with the schemes of the Fourier and wavelet transforms. The simulated packet error rate (PER) and the throughput results are shown for Wi-max technique. The expected throughput for each technique is computed as a function of base station terminal separation distance.

The Fourier transform Orthogonal Frequency Division Multiplexing (OFDM) has the complex based exponential function and it is replaced by the wavelets to reduce the level interference. It is founded that the Haar based wavelet transform is used for reducing the inter symbol interference and inter carrier interference which will occur due to the loss of orthogonality between the carriers [1, 2, 3]. The simulation results are shown that BER performance of the OFDM system with different orthogonal bases like Fourier based OFDM and Wavelet based OFDM. The simulation results are found a great deal of channel dependence with the performance of Fourier and wavelet filters

The main aim of using the wavelet based OFDM WIMAX is the superior spectral properties of the Wavelet filters over the Fourier filters [4]. Further performance gains are made by looking at the alternative orthogonal basis function and found a better transform than the Fourier transform. In this paper further performance using the DWT OFDM in Wi-max technology is executed and compared with the FFT OFDM in the different international Union (ITU) channel modes.

This paper analysis the Discrete Wavelet Transform Orthogonal Frequency Division Multiplexing (DWT-OFDM) based on the WIMAX. It is organized as follows: Section 2 describes the proposed system model for DWT OFDM based on WIMAX. Section 3 presents the SFF SDR development platform and section 4 presents the simulation results and section 5 describes the conclusion.

1. Proposed system model for DWT OFDM based on WIMAX

From the figure shown below the system is divided in to 3 main sections. They are

- Transmitter
- Channel
- Receiver

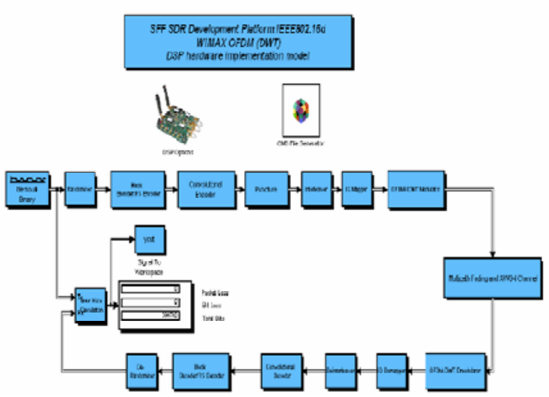


Fig: Block Diagram of Proposed IEEE802.16d WIMAX based DWT-OFDM

The data is generated from random source of information to the receiver section through channel. The generated data consists of series of ones and zeros. The transmission is done on block wise with forward error correction (FEC) and the size of data that is generated is depended on the size of the block. The generated data is converted in to lower rate sequence through series and parallel conversion and then randomized. The randomized data is encoded and it consists of concatenation of outer reed Solomon (RS) code. The implemented RS encoder is derived from the systematic RS code using the field generator and the inner convolution code. That is the data is first pass in block format through the RS encoder and then pass through the convolution encoder[5]. Finally interleaving is done by a two stage permutation.

- To avoid the mapping of the adjacent coded bits on the adjacent sub carriers

- It insures the adjacent coded bits that are mapped alternately on to more or less significant bits of the constellation, so it avoids the low reliable bits.

The pilot sub carrier's frame will be inserted and sent according to the information frame. These pilot sub carriers are used for estimating the channel which is used to compensate the channel effects on the signal. Symbol mapping is the method used for mapping the bits to form the symbol. The 16 QAM modulation scheme is used with half of code with gray coding in the constellation map. In some case the symbol is normalized so the average power is unity irrespective of modulation technique used. In the modulation the data is converted to the corresponding value of M-ary constellation which is complex word i.e. it has real component and imaginary component.

The band width $B = \left(\frac{1}{Ts}\right)$is divided in to equally spaced N sub carriers at frequencies $(k\Delta f)$,

$where\ k = 0,1,2..,N-1\ \ with\ \Delta f = \frac{B}{N}$ and Ts is the sampling interval.

In the transmitter section the information bits are grouped and mapped in to the complex symbols which have real and imaginary parts. In this project constellation with QAM is the modulation technique used to map the bits to symbols. To modulate the spread data symbols to orthogonal carriers N point inverse Wavelet Transform (IDWT) is used with conventional OFDM. In the IDWT zeros are inserted in some bins in order to make the transmitted spectrum compact and to reduce the adjacent carrier's interference. The zeros that are added to the sub carriers will limit the bandwidth of the system, and the system without zeros pad has a spectrum which is spread in frequency. The total number of bits in OFDM symbol is equal to $\log_2 M * Nc$ [where Nc is the sub set of total sub carriers (Nt)]. Orthogonality between the carriers is destroyed when the transmitted signal is passed to the dispersive channel. When the orthogonality between carriers is destroyed then the inverse transformation at the receiver cannot recover the data that is transmitted perfectly. The interference is occurred in the symbol when the energy from one sub channel is transferred or leaked to the other sub channel. But the above problem is solved by introducing the cyclic prefix (CP). The cyclic prefix consists of the v samples from the original K samples that to be transmitted, and it is prefixed to the transmitted symbol. The length v is calculated as the channel's impulse response and it is used to minimize the Inter symbol interference (ISI). If the impulse response channel length is less than or equal to the v, then cyclic prefix is sufficient to completely eliminate Inter Symbol Interference and Inter Carrier Interfernce (ICI).

The efficiency of the transceiver is reduced by a factor of $\frac{k}{k+v}$. It is desirable to calculate v as small or to take K as a large as possible. If the number of the sub channels is sufficiently large, then the channel power spectral density can be assumed virtually flat with in the sub channel. The size of the sub channel is required to approximate the optimum performance of the channel transfer function. The optimum performance depends on how the channel transfer function varies with the frequency [6]. After 16 QAM the data is converted to serial to parallel and applied the Inverse Discrete wavelet Transform (IDWT) after that the data is converted from parallel to serial and these data is fed to the channel. WIMAX model of the receiver performs the same operation as the transmitter but is completely reverse to the transmitter i.e. the data is converted to serial to parallel and Discrete Wavelet Transform (DWT) is applied to the output of the channel. It also includes the operations for compensation and synchronization for the destructive channel. After that the transformed output is modulated and finally produces the received signal.

2. SFF SDR Development platforms

The SFF SDR development platform consists of mainly 3 distinct hardware modules.

1. Digital processing module
2. Data conversion module
3. RF module

The above three modules offer the flexible development capabilities.

Fig: SFF SDR Development Platform

The Digital processing module is used for Virtex-4 FGPA and DM 6446 SoC which are used for necessary performance for implementing the custom IP and the acceleration function with varying from one protocol to another protocol which supports the same hard ware.

The Data conversion module is equipped with the dual channel analog to digital and digital to analog converter The RF module covers different types of frequency ranges in both transmission and reception by allowing it to support a wide range of applications [7].

3.1 System performance analysis and optimization to achieve

The two companies Math works and Texas Instruments both together are used for Mat-Lab or Simulink development of a DSP development tool. The Object modules are designed to their own needs and the programming system is implemented through the RTW (Real time Work Shop) and S-function with the TLC (Target Language compiler). The function of the system design is directly converted to the DSP programming language when it is completed.

3.2 System integration and implementation of Workflow

The overall WIMAX PHY physical layer system is opened in the simulink interface and Matlab is used to communicate the internal function of RTW and TLC. Here we need to build a finished system in to a module according with the code of each block. The over all work flow is shown in the below figure. The figure describes the system that is build based on the simulink IEEE 802.16d Wireless MAN-OFDM PHY standard modules. The first step is the configuration by simulink of the parameters interface and development for connecting node configuration. The information is set to leave the bulk form of a fixed number of patterns and the Real Time Workshop development module is used to and replaced by the C language.

In the mean while the Target Language Complier file, SDR development module and the simulink system development are scheduled for DSP link module by an external module. The configuration of IEEE 802.16d wireless OFDM is achieved with the DSP option Block, simulink is used to develop interfaces connecting node and development platform. The use of compiler option and DSP option block are to optimize the system.

In the SFF SDR development platform of the DSP configuration is classified in to 3 types of memories.

- ➢ LIDRAM (8KB)
- ➢ L2RAM (64KB)
- ➢ SDRAM (8MB)

From the above 3 memories LIDRAM and L2RAM are used for internal memory and SDRAM is used for external memory.

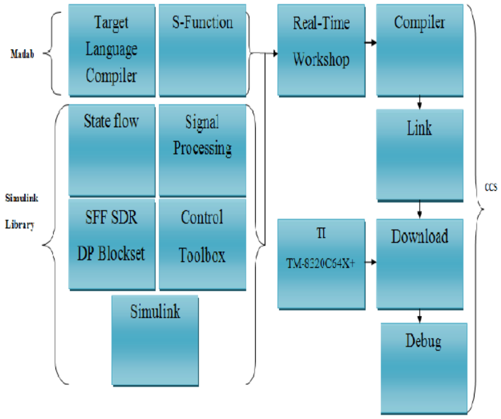


Fig: System Work Flow diagram

Due to the retention of internal memory the speeds become quicker, and if the information is placed in the internal memory system the speeds and executive would enhance the performance.

2.3 Target Language Compiler (TLC)

TLC is a Mat-lab Program and it uses as the syntax. Developers that are using the Real Time Workshop (RTW) tool can use the TLC for creating the self designed syntax language code by adding to executive after the RTW is generated. The usage of S function in the input and output of the set is to design its own system for c programming and creates the simulink objects in the box to use; anyway the RTW is used for producing the C language program. It will not check by performing the actions or debugging code in to the editor. The design of the TLC and the features are shown below

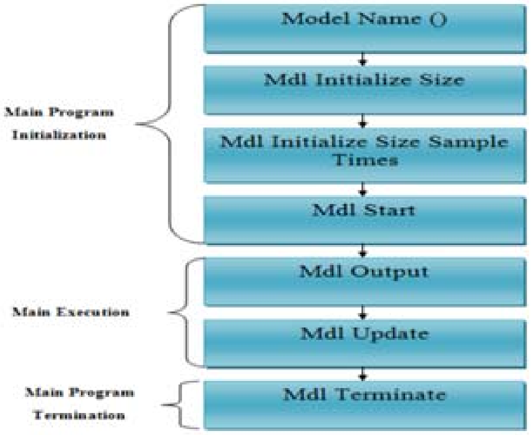


Fig: Target Language Compiler (TLC) structure

3. SIMULATION RESULTS

The simulation results are proposed on DWT OFDM with WIMAX and comparing with the FFT OFDM system and the BER performance of the OFDM system is considered with the AWGN channel.

4.1 Performance with AWGN channel

The results are for the proposed DWT OFDM and it gives the BER performance of the DWT OFDM in AWGN channel. The figure clearly shows the DWT OFDM is better than the FFT OFDM. So we can conclude that the orthogonal base of the Wavelets is more significant than the orthogonal base of the Fourier transform.

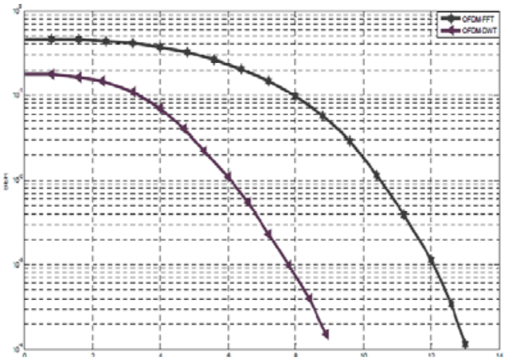


Fig: BER performance of OFDM with DWT and FFT for AWGN channel

4.2 Pedestrian channel A

In Pedestrian situation we have considering two different situations: Moving and Stationary Person. The results are shown in below figure.

In the stationary case we can seen that BER is10^{-3} and the SNR for DWT-OFDM is about 13.2 dB and for FFT-OFDM the SNR is about 17.5 dB.

In the moving case is seen that BER=10^{-3} and the SNR for DWT-OFDM is about 16.2 dB and for FFT-OFDM the SNR is about 21.8 dB

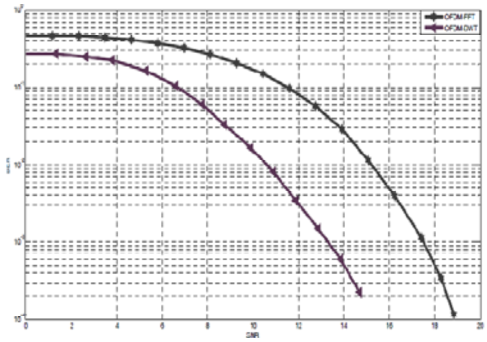


Fig: BER performance of DWT -OFDM in AWGN & Multipath for stationary Pedestrian Channel

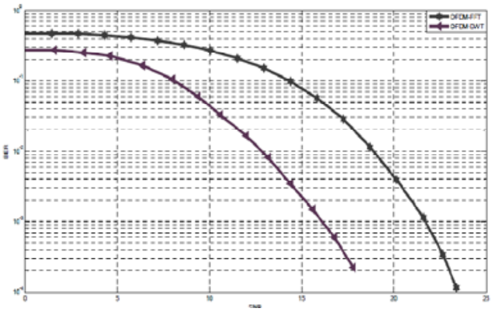


Fig: BER performance of DWT OFDM in AWGN & Multipath for moving Pedestrian Channel

CONCLUSION

The DSP of the SFF SDR Development Platform are completely integrated to the model based design flow, which integrates MATLAB, Simulink, and Real-Time Workshop from The Math-Works. The SFF SCA Development Platform optional package allows SCA waveform development and implementation. The key contribution of this paper was the implementation of the IEEE 802.16d PHY layer based the DWT-OFDM structure was proposed simulate and tested. Simulations provided proved that proposed design achieves much lower bit error rates and better performance than FFT-OFDM assuming reasonable choice of the bases function and method of computations. Proposed DWT-OFDM systems is robust for multi-path channels and does not require cyclically prefixed guard interval, which means that it obtains higher spectral efficiency than conventional OFDM and it can be used at high transmission rates. From obtained results it can be concluded, that SNR can be successfully increased sing proposed wavelet designed method and using a desired wavelet bases function. Therefore this structure can be considered as an alternative to the conventional OFDM.

REFERENCES

1. M. J. Manglani and A. E. Bell, "Wavelet Modulation Performance in Gaussian and Rayleigh Fading Channels," Proc. of MILCOM 2001, McLean, Virginia, Oct. 2001.
2. Research of DFT-OFDM and DWT-OFDM on Different Transmission Scenarios. Proceedings of the 2nd International Conference on Information Technology for application (ICITA).
3. S. R. Baig, F.U.R., and M. J. Mughal, Performance Comparison of DFT, Discrete Wavelet Packet and Wavelet Transforms in an OFDM Transceiver for Multipath Fading Channel. 9th IEEE International Multitopic Conference,, 2005

Improvement of Satellite Image Resolution Appearance Using DWT

D. Leela Rani & V.N.S. Raghavendra Kumar

Department of ECE, Sree Vidyanikethan Engg College, Tirupathi, India

Abstract - In many research fields satellite images are mostly used. The resolution is one of the major issues of these type of images. In this paper, we propose a new satellite image resolution enhancement technique based on the interpolation of the high-frequency subbands obtained by discrete wavelet transform (DWT) and the input image. DWT technique is used by the proposed image resolution enhancement technique to decompose the input image into different subbands. Then, the high-frequency subband images and the input low-resolution image have been interpolated, then by combining all these images a new resolution-enhanced image is generated by using inverse DWT technique. To achieve a sharper image an intermediate stage for estimating high frequency subbands has been proposed. The proposed technique has been tested on satellite benchmark images. The quantitative (peak signal-to-noise ratio and root mean square error) and visual results show the superiority of the proposed technique over the conventional and state-of-art image resolution enhancement techniques.

Keywords:Discrete wavelets transform, Interpolation, Satellite image resolution enhancement, Wavelet zero padding.

I. INTRODUCTION

Resolution of an image has been always an important issue in many image and video processing applications. These include the applications such as video resolution enhancement [1], feature extraction [2] and satellite image resolution enhancement [3].

Images are usually processed in order to obtain more enhanced resolution. In image processing interpolation is a method, which increases the pixels in a digital image. Interpolation technique has been very widely used in many processing applications of an image, such as facial reconstruction [4], multiple description coding [5], and image resolution enhancement [6]-[8].The interpolation based image resolution enhancement has been used for a long time and many techniques are developed in order to increase the image quality. There are three interpolation techniques namely nearest neighbor, bilinear, and bi-cubic. We use the bi-cubic interpolation as it is the sophisticated than the other two techniques as it produces smoother edges.

Wavelets are also playing a significant role in many image processing applications. The 2-D wavelet decomposition of an image is performed by applying 1-D discrete wavelet transform along the rows of the image and then along the columns, where the results are decomposed.By this operation four subband images are obtained namely low-low (LL), low-high (LH), high-low (HL) and high-high (HH). The frequency components of those subbands cover the entire frequency spectrum of the original image. The another wavelet transform which is used in several image processing applications is stationary wavelet transform. It is same as the discrete wavelet transform and will not use down sampling. The size of the subbands obtained will be same as the input image.

The image resolution enhancement using wavelets is a very new subject and new algorithms have been proposed [9]-[15]. Carey estimation algorithm have been attempted to estimate the unknown details of the wavelet coefficients in order to improve the sharpness of the image which is obtained or reconstructed. This estimation was carried out by understanding the wavelet transform evolution among the subbands, which are of same type. Over smoothed edges are obtained in common image interpolation techniques, obtained by image continuity. A wavelet based interpolation method to place no continuity constraints is introduced. This algorithm estimates the edges regularity by measuring the decay of the wavelet transform coefficients across scales and preserves the regularity by extrapolating a new subband to be used in image resynthesis. This algorithm produces veryvisible sharp edges than the other traditional techniques [9].

The edge detection algorithm is used to identify the edges in the lower frequency subbands, which are used to prepare a model for estimating edges in higher frequency subbands. The coefficients with significant values are estimated for the evolution of the wavelet coefficients. The edge guided non-linear interpolation technique comprises directional filtering and data fusion techniques. To interpolate the pixel two observation sets are defined in the orthogonal directions. Each set produces an estimate of the pixel value. These estimates are modeled as different noise measurements of the pixel which is missing. This is fused by linear minimum mean square error estimator. To improve into a more robust estimate, by using the statistics of the two other estimates. Many researches are implemented in order toestimate the coefficients [16].

In this paper, we propose a resolution enhancement technique which generates a sharper high resolution image. The technique uses DWT to obtain different subbands by using a low resolution image. Then the bicubic interpolation has been applied to all the obtained high frequency subbands. To obtain a sharper image we introduce an intermediate stage for estimating the high frequency subbands by utilizing the LL subband. This can be used by subtracting the LL subband image and the input image. This image is called difference image.The input image also interpolated in parallel.To combine all these images, to get a final resolution enhanced image inverse discrete wavelet transform has been used. The proposed technique has been compared with standard interpolation techniques, wavelet zero padding (WZP), where zeros are replaced in the place of unknown coefficients of the high frequency subbands. The state-of-art techniques, such as WZP and cycle spinning (CS) [17], stationary wavelet transform (SWT) [16] and complex wavelet transform (CWT) image resolution enhancement technique, which is previously used [3]. The proposed enhancement technique uses, Daubechies wavelet as the mother wavelet function and bicubic interpolation as the interpolation technique.

II. WAVELET-BASED IMAGE RESOLUTION ENHANCEMENT

In the last few years, researchers have been actively exploring many mechanisms to enhance the resolution of the image. One of thetechniques proposed to enhance the image resolution is CS based enhancement. The two more are the CWT image resolution enhancement and SWT image enhancement techniques.

2.1 CS Based Image Resolution Enhancement

This method adopts the CS methodology in the wavelet domain [17]. The algorithm consists of two main steps as follows:

1) An initial approximation to the unknown high resolution image is generated by using wavelet domain zero padding (WZP).

2) The cycle-spinning methodology has been adopted to operate the following tasks:

a) A number of low resolution images are generated from the obtained and estimated high resolution image by spatial shifting, wavelet transforming and discarding the high frequency subbands.

b) The WZP processing is applied to all those low resolution images yielding N high resolution images.

c) These intermediate high resolution images are realigned and averaged to give the final reconstructed high resolution image.

2.2 CWT Based Image Resolution Enhancement

In this technique, dual-tree CWT (DT-CWT) is used to decompose an input image into different subband images. DT-CWT is used to decompose an low resolution input image into different subband images. The high frequency subband images and the input image are then interpolated. By combining all these images, we generate a new reconstructed high resolution image by using inverse DT-CWT. The image resolution enhancement is achieved by using directional selectivity, which is provided by CWT. Here the subbands of high frequency will contribute the sharpness of the image in six different directions. These are called as edges. Here the enlargement factor in the resolution enhancement is α .

2.3 SWT Based Image Resolution Enhancement

In this technique, the three high frequency subbands (LH, HL, HH) contains the high frequency components of the input image. The bi-cubic interpolation with enlargement factor of 2, is applied to the high frequency subbands images. The downsampling in the each DWT subband images causes the loss of information in the particular subbands. The SWT is employed to minimize the loss.

The high frequency subbands obtained by applying the SWT and DWT techniques to the input image can be added each other as they are of same size. For further enhancement of resolution the subbands are interpolated. It is also known that in the wavelet domain, the low resolution image is obtained by lowpass filtering of the high resolution image [16]. The low frequency subband is the low resolution of the original image. Instead of using LL subband image which contain less information than the original input image, we are using the input image directly by applying interpolation. By using the input image rather than LL subband image, increases the quality of the reconstructed image.

Interpolating the input image by $\alpha/2$, and the high frequency subbands by 2 and α in the intermediate and in the final interpolation stages. By applying the IDWT to all the interpolated images, we achieve the final reconstructed image.

III. DISCRETE WAVELET TRANSFORM (DWT)

The discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations obeying some defined rules. In other words, this transform decomposes the signal into mutually orthogonal set of wavelets, which is the main difference from continuous wavelet transform, or its implementation for the discrete time series sometimes called discrete-time continuous wavelet transform (DT-CWT) [3]. *The wavelet can be constructed from a scaling function which describes its scaling properties. The restriction that the scaling functions must be orthogonal to its discrete translations implies some mathematical conditions on them which are mentioned everywhere e. g. the dilation equation.*

$$\emptyset(x) = \sum_{k=-\infty}^{\infty} a_k \emptyset(S_x - k)$$

Where S is a scaling factor. Moreover, the area between the function must be normalized and scaling function must be orthogonal to its integer translates e. g.

$$\int_{-\infty}^{\infty} \boldsymbol{\emptyset(x)\emptyset(x+1)dx = \delta 0, l}$$

After introducing some more conditions (as the restrictions above does not produce unique solution) we can obtain results of all this equations, e. g. finite set of coefficients a_k which define the scaling function and also the wavelet. The wavelet is obtained from the scaling function as

$$\boldsymbol{\Psi(x) = \sum_{-\infty}^{\infty}(-1)^k a_{N-1-k}\Psi(2x - k)}$$

Where N is an even integer. The set of wavelets than forms an orthonormal basis which we use to decompose signal. Usually only few of the coefficients a_k are nonzero which simplifies the calculations.

IV. ENHANCING RESOLUTION OF AN IMAGE USING DWT

As it was mentioned before,resolution is an important feature in satellite imaging. It makes the enhancing of resolution in the images. They are of very vital importance because increase of resolution directly affects the performance of the system and these images cannot be used as input images. The main loss of an image after being resolution enhanced by applying interpolation is on its high-frequency components, which is due to the smoothing caused by interpolation. Hence, in order to increase the quality of the enhanced image, preserving the edges is essential. DWT [8] has been employed in order to preserve the high-frequency components of the image. DWT separates the image into different sub band images, namely, LL, LH, HL, and HH. High-frequency sub bands contain the high frequency component of the image. The interpolation can be applied to these four sub band images. In the wavelet domain, the low-resolution image is obtained by low-pass filtering of the high-resolution image.The lowresolutionimage (LL subband), without quantization (i.e., withdouble-precision pixel values) is used as the input for theproposed resolution enhancement process. In other words, lowfrequency subband images are the low resolution of the originalimage. Therefore, instead of using low-frequency subband images,which contains less information than the original inputimage, we are using this input image through the interpolationprocess. Hence, the input low-resolution image is interpolatedwith the half of the interpolation factor, $\alpha/2$ are used to interpolatethe high-frequency subbands. The redundancy and shift variance of the DWT mean that DWT coefficients are interpolable inherently.In this correspondence, one level DWT (with Daubechies 9/7 aswavelet function) is used to decompose an input image into differentsubband images.

In order topreserve more edge information, i.e., obtaining a sharper enhancedimage, we have proposed an intermediate stage in highfrequencysubband interpolation process. As the low-resolution input satellite image and the interpolatedLL image with factor 2 are highly correlated. The differencebetween the LL subband image and the low-resolution inputimage are in their high-frequency components. Hence, thisdifference image can be used in the intermediate process tocorrect the estimated high-frequency components.

This estimationis performed by interpolating the high-frequency subbandsby factor 2 and then including the difference image (whichis high-frequency components on low-resolution input image)into the estimated high-frequency images, followed by anotherinterpolation with factor α/2 in order to reach the requiredsize for IDWT process. The intermediate process of addingthe difference image, containing high-frequency components,generates significantly sharper and clearer final image. Thissharpness is boosted by the fact that, the interpolation ofisolated high-frequency components in HH, HL, and LH willpreserve more high-frequency components than interpolating the low-resolution image directly.

The quantitative comparisons of the proposed method with all the other methods can show the perfect superiority of the proposed method. Peak-signal-to-noise-ratio (PSNR) and Root mean square error (RMSE) can be used in order to obtainsome quantitative results for the comparison purpose.

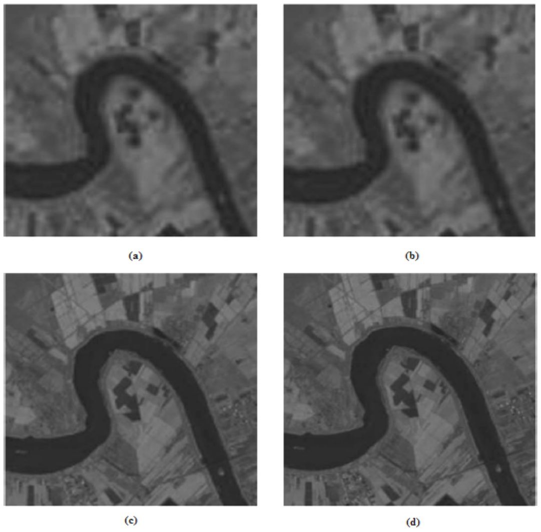


Fig1: (a) Input Image, (b) CS Based Image,(c) CWT BasedEnhanced Image, (d) DWT Based Enhancement Technique

PSNR can be achieved by using the formula:

$$PSNR = 10\, log10\left(\frac{R^2}{MSR}\right)$$

Where R is the maximum fluctuation in the input image. MSE is representing the MSE between the input image I_{in} and the original image I_{org} which can be achieved by the formula:

$$MSE = \frac{\sum_{i,j}\left(I_{in}(i,j) - I_{org}(i,j)\right)^2}{M \times N}$$

Where M and N are the size of the images. The RMSE, the square root of MSE can be calculated by:

$$RMSE = \sqrt{\frac{\sum_{i,j}\left(I_{in}(i,j) - I_{org}(i,j)\right)^2}{M \times N}}$$

V. CONCLUSION:

Here we proposed a new resolution enhancement technique by applying DWT. In order to show the superiority of the proposedmethod over the conventional and state-of-art techniques from visual point of view in Fig1. It is clear that the resultant image, enhanced by using theproposed technique, is sharper than the other techniques.The technique can be tested on satellite and benchmark images, where the PSNR, RMSE and visual results show the superiority of the proposed technique when compared to the previous techniques. The PSNR improvement of the proposed technique can be up to 7.19dB compared with the standard bi-cubic interpolation technique.

REFERENCES:

[1]. L. Yi-bo, X. Hong, and Z. Sen-yue, “The wrinkle generation methodfor facial reconstruction based on extraction of partition wrinkle linefeatures andfractal interpolation,” in *Proc. 4th Int. Conf. Image Graph.*,Aug. 22–24, 2007, pp. 933–937.

[2]. Y. Rener, J. Wei, and C. Ken, “Downsample-based multiple descriptioncoding and post-processing of decoding,” in *Proc. 27th ChineseControl Conf.*, Jul. 16–18, 2008, pp. 253–256.

[3]. H. Demirel, G. Anbarjafari, and S. Izadpanahi, “Improved motionbasedlocalized super resolution technique using discrete wavelettransform for low resolution video enhancement,” in *Proc. 17th Eur. Signal Process. Conf.*, Glasgow, Scotland, Aug. 2009, pp.1097–1101.

[4]. Y. Piao, I. Shin, and H. W. Park, “Image resolution enhancement usinginter-subband correlation in wavelet domain,” in *Proc. Int. Conf. ImageProcess.*, 2007, vol. 1, pp. I-445–448.

[5]. H. Demirel and G. Anbarjafari, “Satellite image resolution enhancementusing complex wavelet transform,” *IEEE Geoscience and RemoteSensing Letter*, vol. 7, no. 1, pp. 123–126, Jan. 2010.

[6]. C. B. Atkins, C. A. Bouman, and J. P. Allebach, “Optimal imagescaling using pixel classification,” in *Proc. Int. Conf. Image Process.*,Oct. 7–10, 2001, vol. 3, pp. 864–867.

[7]. W. K. Carey, D. B. Chuang, and S. S. Hemami, “Regularity-preservingimage interpolation,” *IEEE Trans. Image Process.*, vol. 8, no. 9, pp.1295–1297, Sep. 1999.

[8]. S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York:Academic, 1999.

[9]. J. E. Fowler, “The redundant discrete wavelet transform and additivenoise,”Mississippi State ERC, Mississippi State University, Tech. Rep.MSSU-COE-ERC-04-04, Mar. 2004.

[10]. X. Li and M. T. Orchard, “New edge-directed interpolation,” *IEEETrans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[11]. K. Kinebuchi, D. D. Muresan, and R. G. Baraniuk, “Waveletbasedstatistical signal processing using hidden Markov models,”in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2001, vol. 3, pp. 7

[12]. S. Zhao, H. Han, and S. Peng, “Wavelet domain HMT-based imagesuper resolution,” in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2003,vol. 2, pp. 933–936.

[13]. A. Temizel and T. Vlachos, “Wavelet domain image resolution enhancementusing cycle-spinning,” *Electron. Lett.*, vol. 41, no. 3, pp.119–121, Feb. 3, 2005.

[14]. A. Temizel and T. Vlachos, “Image resolution upscaling in the waveletdomain using directional cycle spinning,” *J. Electron. Imag.*, vol. 14,no. 4, 2005.

[15]. G. Anbarjafari and H. Demirel, “Image super resolution based oninterpolation of wavelet domain high frequency subbands and thespatial domain input image,” *ETRI J.*, vol. 32, no. 3, pp. 390–394,Jun. 2010.

[16]. A. Temizel, “Image resolution enhancement using wavelet domainhidden Markov tree and coefficient sign estimation,” in *Proc. Int.Conf. Image Process.*, 2007, vol. 5, pp. V-381–384.

[17]. A. Temizel and T. Vlachos, “Wavelet domain image resolution enhancementusing cycle-spinning,” *Electron. Lett.*, vol. 41, no. 3, pp. 119–121,Feb. 3, 2005.

[18]. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. EnglewoodCliffs, NJ: Prentice-Hall, 2007.

◈◈◈

An Efficient Technique of Data Hiding For STEGO Using Double Key

P.V. Mahesh & **R. S. Rao**

Sree Vidyanikethan Engineering College

Abstract - *Steganography* gained significance in the past few years due to the increasing need for providing secrecy in an open environment like the internet. With almost anyone can observe the communicated data all around, steganography attempts to hide the very existence of the message and make communication undetectable. In this paper we propose a modern technique with Integer Wavelet transform (IWT) and double key to accomplish high hiding capability, high security and good visual quality. Here cover image is transformed in to wavelet transform co-efficients and the coefficients are selected randomly by using Key-2 for embedding the data. Key-1 is used to calculate the number of bits to be embedded in the randomly selected coefficients. Finally the Optimum Pixel Adjustment Process(OPAP) is applied to the stego image to reduce the data embedding error.

Keywords- *Steganography, Integer wavelet transform (IWT), Optimum Pixel Adjustment Process(OPAP).*

I. INTRODUCTION

Steganography is the art and science of hiding secret data in plain sight without being noticed within an innocent cover data so that it can be securely transmitted over a network. The word steganography is originally composed of two Greek words *steganos* and *graphia,* which means "covered writing". Modern steganography is generally understood to deal with electronic media rather than physical objects. Image steganography, of all has gained much impetus and reputation in the recent past [1-18]. It comes under the general assumption that if the feature is visible, the point of attack is evident. Thus the goal here is to always conceal the coherence of the embedded data. The basic model of secret key steganography consists of cover, secret data, stego image and key. Any digital file such as image , video, audio, etc can be used as cover. Cover is also known as cover-object or cover image, is the plain digital image with no secret data deposited in it. After the embedment it is called the stego image or stego object [1, 2, 3, 10]. In image steganography [1] the critical data is camouflaged in a cover image with immense dexterity.

The most popular hiding techniques are spatial domain based steganographic techniques and transform domain based steganographic techniques. A useful, practical steganographic method should be robust and should retain the hidden data even after many pixel values have been modified. The other type of hiding method is the transform domain techniques which appeared to overcome the robustness and imperceptibility problems found in the LSB substitution techniques. There are many transforms that can be used in data hiding, the most widely used transforms are; the discrete cosine transform (DCT) which is used in the common image compression format JPEG and MPEG, the discrete wavelet transform (DWT) and the discrete Fourier transform (DFT).One approach to this problem is to transform the image and embed the data in the transformed pixels[4-6,8,9]. We say that the original image exists in the spatial domain and the transformed image in the transformed domain. The data is then embedded in the transformed pixels and the image is transformed back to the spatial domain. The idea is that the image may now be exposed to various operations that will change the pixels, but when this modified image is transformed again, the hidden data will still be embedded in the transformed pixels. The disadvantage of the DCT based steganographic technique [9, 13], is the hiding capacity. Wavelet transform based stego technique provides high capacity as much as possible. In [4] the secret message is embedded into the high frequency and low frequency coefficients of the wavelet transform to high hiding capacity, but it provides less PSNR at high hiding rate. In this paper we propose a new modified version of the methodology in [4], which can embed a larger amount of data in integer wavelet transform (IWT) domain with high PSNR.

II. RELATED WORKS

A. *Integer Wavelet Transforms*

Generally wavelet domain allows us to hide data in regions that the human visual system (HVS) is less sensitive to, such as the high resolution detail bands (HL, LH and HH), Hiding data in these regions allow us to increase the robustness while maintaining good visual quality. Integer wavelet transform maps an integer data set into another integer data set. In discrete wavelet transform, the used wavelet filters have floating point coefficients so that when we hide data in their coefficients any truncations of the floating point values of the pixels that should be integers may cause the loss of the hidden information which may lead to the failure of the data hiding system.

To avoid problems of floating point precision of the wavelet filters when the input data is integer as in digital images, the output data will no longer be integer which doesn't allow perfect reconstruction of the input image [12] and in this case there will be no loss of information through forward and inverse transform [11].

The Haar Wavelet Transform is the simplest of all wavelet transform. The four bands obtained are LL, LH, HL, and HH which is shown in Fig 1. The LL band is called as approximation band, which consists of low frequency wavelet coefficients, and contains significant part of the spatial domain image. The other bands are called as detail bands which consist of high frequency coefficients and contain the edge details of the spatial domain image. Integer wavelet transform can be obtained through lifting scheme. Lifting scheme is a technique to convert DWT coefficients to Integer coefficients without losing information. [5, 8].

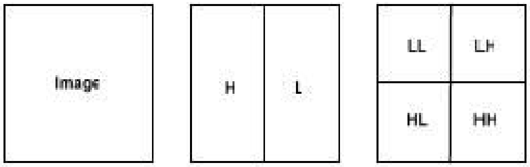


Figure 1 Image and its transform domain bands

B. *Forward Lifting scheme in IWT.*

Step1: Column wise processing to get H and L

H = (Co-Ce) (1)

L = (Ce- $\left| H/2 \right|$) (2)

Where Co and Ce is the odd column and even column wise pixel values.

Step 2: Row wise processing to get LL,LH,HL and HH,

Separate odd and even rows of H and L,

Namely, Hodd – odd row of H

Lodd- odd row of L

Heven- even row of H

Leven- even row of L

LH = Lodd - Leven (3)

LL = Leven - $\left| LH / 2 \right|$ (4)

HL = Hodd – Heven (5)

HH = Heven - $\left| HL / 2 \right|$ (6)

C. *Reverse Lifting scheme in IWT*

Inverse Integer wavelet transform is formed by Reverse lifting scheme. Procedure is similar to the forward lifting scheme.

D. *LSB Embedding*

Simple LSB embedding [2]is detailed in this section. Consider a 8-bit gray scale image matrix consisting m × n pixels and a secret message consisting of k bits. The first bit of message is embedded into the LSB of the first pixel and the second bit of message is embedded into the second pixel and so on. The resultant stego-image which holds the secret message is also a 8-bit gray scale image and difference between the cover image and the stego-image is not visually perceptible.

This can be further extended, and any number of LSB's can be modified in a pixel. The quality of the image, however degrades with the increase in number of LSB's. Usually up to 4 LSB's can be modified without significant degradation in the message. Mathematically, the pixel value 'P' of the chosen pixel for storing the k-bit message Mk is modified to form the stego-pixel 'Ps' as follows:

$Ps=P-mod(P,2^k)+M_k$ (7)

The embedded message bits can be recovered by

$Mk=mod(Ps,2^k)$ (8)

One method to recover the quality of the LSB substitution is Optimal Pixel adjustment Process (OPAP)[2].

E. *Optimal Pixel adjustment Process*

The projected Optimal Pixel adjustment Procedure (OPAP) reduce the error caused by the LSB substitution method. In OPAP method the pixel value is adjusted after the secret data is concealed. It is done to improve the quality of the stego image without disturbing the data hidden.

F. *Adjustment Process*

- Let ‘n’ LSBs be substitute in each pixel.
- Let d= decimal value of the pixel after the substitution.
- d1 = decimal value of last n bits of the pixel.
- d2 = decimal value of n bits concealed in that pixel.

If(d1~d2)<=(2^n)/2

then no adjustment is made in that pixel.

Else

If(d1<d2)

d = d – 2^n .

If(d1>d2)

d = d + 2^n .

This d is converted to binary and written back to pixel.

III. PROPOSED METHODOLOGY

Fig. 2 shows the proposed system is a high capacity steganography system. Preprocessing includes R, G and B plane separation and Histogram modification. Then Integer wavelet transform is applied to the cover image to get wavelet coefficients. Wavelet coefficients are randomly selected by using key-2 for embedding the secret data. Key -2 is 8x8 binary matrix in which ‘1’ represents data embedded in the corresponding wavelet coefficients and ‘0’ represents no data present in the wavelet coefficients.

Key-1(K1) is a decimal number varying from 1 to 4 and it will decide the number of bits to be embedded in the cover object.

.

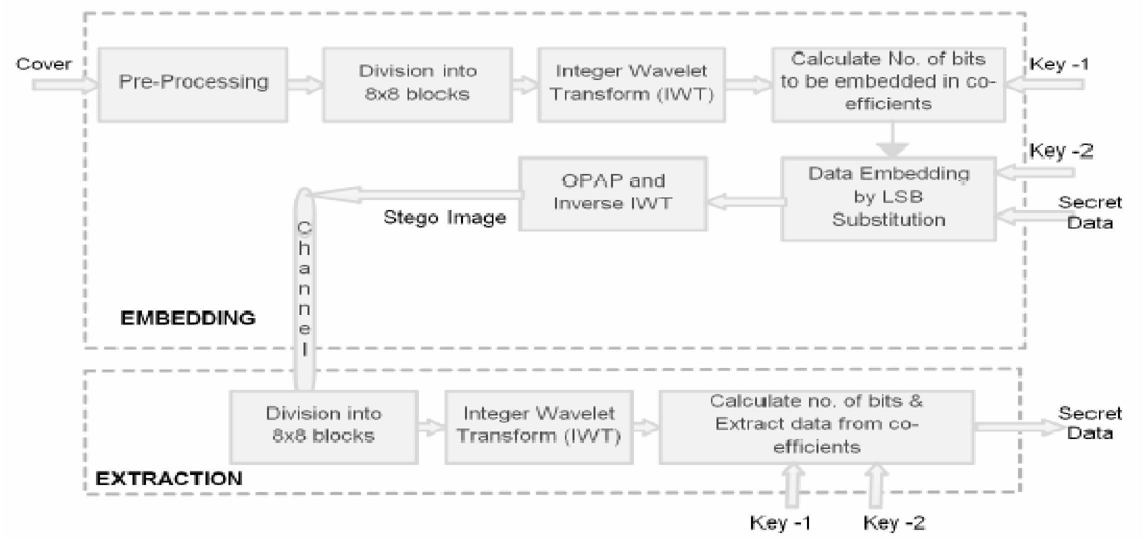

Cover
Pre-Processing
Division into 8x8 blocks
Integer Wavelet Transform (IWT)
Calculate No. of bits to be embedded in co-efficients
Key -1
Key -2
Data Embedding by LSB Substitution
Secret Data
OPAP and Inverse IWT
Stego Image
Channel
EMBEDDING
Division into 8x8 blocks
Integer Wavelet Transform (IWT)
Calculate no. of bits & Extract data from co-efficients
Secret Data
EXTRACTION
Key -1
Key -2

Figure 2 The Block Diagram of the Proposed Methodology

.*A*. *Algorithm (For Embedding of Data):*

Step 1: Read the cover image as a 2D file with size of 256×256 pixels.

Step 2: R, G and B planes are separated.

Step 3: Consider a secret data as text file. Here each character will take 8 bits.

Step 4: Histogram adjustment[16] is done in all planes, Because, the secret data is to be embedded in all the planes, while embedding integer wavelet coefficients produce stego-image pixel values greater than 255 or lesser than 0. So all the pixel values will be ranged from 15 to 240.

Step 5: Each plane is divided into 8×8 blocks.

Step 6: Apply Haar Integer wavelet transform to 8 × 8 blocks of all the planes, This process results in LL1, LH1, HL1 and HH1 sub bands.

Step 7: Using Key-1(K1) calculate the Bit length(BL) for corresponding wavelet co-efficients (WC), Here we used modified version of Bit length calculation used in [4]. Using the following equation, we get the high capacity steganography.

$$BL = \begin{cases} K1+3 & if \quad WC \geq 2^{K1+2} \\ K1+1 & if \quad WC < 2^{K1+2} \end{cases} \qquad (9)$$

Step 8: Using key-2 select the position and coefficients for embedding the 'BL' length data using LSB substitution[2]. Here data is embedded only in LH1,HL1and HH1 subbands. Data is not embedded in LL1 because they are highly sensitive and also to maintain good visual quality after embedding data. An example of key-2 is shown below.

$$Key-2- \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix} \qquad (10)$$

Step 9: Applying Optimal Pixel adjustment Procedure (OPAP) reduces the error caused by the LSB substitution method.

Step 10: Take inverse wavelet transform to each 8×8 block and combine R,G&B plane to produce stego image.

B. Algorithm (For Extracting of Data):

Step 1: Read the Stego image as a 2D file with size of 256 × 256 pixels.

Step 2: R, G and B planes are separated.

Step 3: Each plane is divided into 8×8 blocks.

Step 4: Apply Haar Integer wavelet transform to 8×8 blocks of all the planes, This process results LL1,LH1,HL1 and HH1 subbands.

Step 5: Using Key-1 calculate the Bit length(BL) for corresponding wavelet co-efficients(WC), using the 'BL' equation used in Embedding procedure.

Step 6: Using key-2 select the position and coefficients for extracting the 'BL' length data.

Step 7: Combine all the bits and divide it in to 8 bits to get the text message.

IV. ERROR METRICS

A performance estimate in the stego image is calculated by means of two parameters namely, Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). The MSE is calculated by using the equation,

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (X_{i,j} - Y_{i,j})2 \qquad \textbf{(11)}$$

.

key - 1(K1) = 1								
Various key -2	Cover Image	Total No. of bits embedded	Channel - I Red		Channel - II Green		Channel - III Blue	
			MSE	PSNR(dB)	MSE	PSNR(dB)	MSE	PSNR(dB)
Key - A	Flower	59168	1.31219	53.9016	1.86553	50.8456	1.05667	55.7828
	Elephant	80512	3.34099	45.7841	3.10357	46.4244	2.80991	47.2878
Key - B	Flower	78696	1.32099	53.8436	1.86851	50.8317	1.06836	55.6872
	Elephant	92968	3.31004	45.8649	3.10463	46.4214	2.79836	47.3235
Key - C	Flower	104944	1.32786	53.7986	1.86316	50.8566	1.07731	55.6148
	Elephant	196480	3.37844	45.6873	3.14449	46.3106	2.82475	2.82475
Key - D	Flower	110320	1.29164	54.0388	1.84346	50.949	1.04401	55.8875
	Elephant	252784	3.36729	45.716	3.14812	46.3006	2.81612	47.2686
Key - E	Flower	240840	1.33505	53.7517	1.88163	50.7709	1.0914	55.5019
	Elephant	335560	3.40763	45.6126	3.18555	46.1979	2.8945	47.0301

.where M and N denote the total number of pixels in the horizontal and the vertical dimensions of the image $X_{i,\ j}$ represents the pixels in the original image and $Y_{i,j}$ represents the pixels of the stego-image.

The Peak Signal to Noise Ratio (PSNR) is expressed as

$$PSNR = 10\log_{10}\left(\frac{I_{max}^{2}}{MSE}\right) db \qquad (1$$

V. RESULT AND DISCUSSION

In this present implementation, Flower and Elephant 256 × 256 × 3 color digital images have been taken as cover images, as shown in Figure 3&4- a,b, c & d, tested with key-2(key-E) and various key-1s. The effectiveness of the stego process proposed has been studied by calculating MSE and PSNR for the two digital images in RGB planesandtabulated.

First analysis is used to select the Key-2 for random selection of coefficients for embedding data (in this analysis Key-1 has been set as K1=1) and the results are tabulated in Table-I for various Key-2 using the proposed method. From table –I we will understand that Key-E provides high capacity and Key – A provides low capacity.

In the second analysis, Key-E will be taken with various 'K1' values and the results are tabulated in Table-II. Combining Key-E with K1=4 will yield high hiding capacity with high PSNR

.

Table – II MSE, PSNR for fixed Key-2(Key-E) with varying Key-1 in all the three planes

Key-1	Cover Image	Total No. of bits embedded	Channel - I Red		Channel - II Green		Channel - III Blue	
			MSE	PSNR(dB)	MSE	PSNR(dB)	MSE	PSNR(dB)
K1=1	Flower	240840	1.33505	53.7517	1.88163	50.7709	1.0914	55.5019
K1=2	Flower	299256	1.35302	53.6355	1.88619	50.7499	1.0997	55.4357
K1=3	Flower	307280	1.34458	53.6899	1.87415	50.8055	1.0992	55.4397
K1=4	Flower	348192	1.35817	53.6025	1.87792	50.7881	1.0911	55.504
K1=1	Elephant	335560	3.40763	45.6126	3.18555	46.1979	2.8945	47.0301
K1=2	Elephant	348192	3.39305	45.6498	3.16413	46.2565	2.8627	47.126
K1=3	Elephant	370288	3.38845	45.6616	3.16842	46.2447	2.8782	47.0791
K1=4	Elephant	435304	3.38991	45.6578	3.18234	3.18234	2.8663	47.1151

.

Figure 3. K1=1,2,3 & 4 respecively

Figure 4. K1=1,2,3 & 4 respecively

VI. CONCLUSION

Data hiding with steganography has two primary objectives firstly that steganography should grant the maximum possible payload, and the second, embedded data must be undetectable to the observer. It should be stressed on the fact that steganography is not meant to be robust. It was found that the proposed method gives high payload (capacity) in the cover image with very little error. . This is of course on the expense of reducing PSNR and increasing the MSE. By modifying the equation (9) to get high capacity for the various applications using wavelet transform, Key-1 and Key-2 provides high protection. The drawback of the proposed method is the computational overhead. This can be reduced by high speed computers.

REFERENCE :

[1] Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc kevitt" Digital image steganography: Survey and analysis of current methods Signal Processing", 90 (2010),727–752.

[2] C.K. Chan, L.M. Chen, "Hiding data images by simple LSB substitution", Pattern recognition, 37 (3) (2004), 469–474.

[3] R.O. EI Safy, H. H. Zayed, A. EI Dessouki, " An Adaptive Steganographic Technique Based on Integer Wavelet Transform", International conference on Networking and media convergence ICNM-(2009), 111 - 117.

[4] Po-Yueh Chen and Hung-Ju Lin, "A DWT Based Approach for Image Steganography", International Journal of Applied Science and Engineering 4,(2006), 275-290

[5] R.Amirtharajan, Adarsh D, Vignesh V and R. John Bosco Balaguru, "PVD Blend with Pixel Indicator - OPAP Composite for High Fidelity Steganography", International Journal of Computer Applications 7(9),(October 2010),31–37.

6] Guorong Xuan; Jidong Chen; Jiang Zhu; Shi, Y.Q.; Zhicheng Ni; Wei Su," Lossless data hiding based on integer wavelet transform" , IEEE Workshop on Multimedia Signal Processing,vol 2,(2002).

[7] Saeed Sarreshtedari and Shahrokh Ghaemmaghami," high capacity Image Steganography in Wavelet Domain ", IEEE CCNC 2010 proceedings,(2010),1-5.

[8] Cheng jiang Lin, Bo Zhang,Yuan F. Zheng," Packed Integer Wavelet Transform Constructed by Lifting Scheme", IEEE Transactions on Circuits and Systems forVideo Technology, (Dec 2000), 1496 –1501.

[9] H. W. Tseng and C. C. Chnag, "High capacitydata hiding in jpeg compressed images," Informatica, vol. 15, no. I,2004.

[10] H. H. Zayed, "A High-Hiding Capacity Technique Hiding data in hnages Based on K-Bit LSB Substitution," The 30th International Conference on Artificial IntelligenceApplications (ICAIA - 2005) Cairo, Feb. 2005.

[11] A. R. Calderbank, 1. Daubechies, W. Sweldens and B. Yeo., "Wavelet transforms that map integers to integers". Appiedand Computational Harmonic Analysis, vol.5,.332-369,1998

[12] G. J. Simmons, "The prisoners' problern and the subliminal channel," in Proceedings of Crypto' 83, pp. 51-67, 1984.

[13] W. Bender, N. Morimoto, A. Lu, Techniques for data hiding, IBM Syst. J. 35 (3/4) (1996) 313–336.

[14] K.L. Chung, C.H. Shen, L.C. Chang, A novel SVD- and VQ-based image hiding scheme, Pattern Recognition Lett. 22 (9)

(2001) 1051–1058.

[15] Ran-Zan Wang, Chi-Fang Lin, Ja-Chen Lin, Image hiding by optimal LSB substitution and genetic algorithm, pattern Recognition 34 (3) (2001) 671–683.

[16] Yildiray Yalman, Ismail Erturk, "A New Histogram Modification Based Robust Image Data Hiding Technique", 24th Interational Symposium on Computer and Information Sciences, ISCIS 2009, 39-43 .

[17] Rengarajan Amirtharajan and John Bosco Balaguru Rayappan, "*Tri- Layer Stego for Enhanced Security – A Keyless Random Approach*" - IEEE Xplore, DOI, 10.1109/IMSAA.2009.

[18] J. Lillo M. Shih, "Generalizations of Pixel-Value Differencing Steganography for Data Hiding in bnages", Fundamenta aticae, vol. 83, noJ, pp. 319-335, 2008.

◈◈◈

TCP/IP Packet Reassembly and Scalable Segmentation Design for Network Monitoring Device

Mohsin Mulla [a], Bharathi M [b] & R Shankar [c]

[a, b] Dept. of Electronics and Communication, R.V. College of Engineering, Bangalore, India
[c] Whizchip Design Technologies Pvt Ltd, Bangalore , India

Abstract - Network monitoring device monitors the traffic from multiple users and has a capability to block or grant access to individual users. Reassembly and Segmentation modules are the features of the network monitoring device which reassembles the packets and segments the packets. Packet reassembly is necessary because today's communication systems use what is called packetized communication. Packets are segmented because they can travel across a packet switched network without tieing up a communication circuit, this increases both the reliability and the speed at which data can travel across a network. To monitor this network traffic it is important to assemble all the packets coming from different sources in order so that unauthorized access to the network can be blocked. Segmentation module is necessary at the output of network monitoring device to send constructed original message into segmented packets. To implement Reassembly and Segmentation modules control section fields are used. A control section field contains all the information regarding each packet which is very helpful for all the modules in the network monitoring device.

Key Words*: IP packet, Segmentation, Reassembly, Processing header.*

I. INTRODUCTION

Internet has made communication faster and easier to any corner of the world. Like every single innovation in science and technology, Internet has its downsides. These include risks like theft of personal information, spamming, virus threat, fraudulent transaction, inappropriate sites watched by the children at impressionable age etc. Therefore, it is crucial to monitor the network, in order to understand it and to react appropriately**.**

Internet Protocol (IP) contains address information for routing packets in Network Layer of TCP/IP model. IP, as an integral part of TCP/IP, is for addressing and routing packets. It provides the mechanism to transport datagram across a large network. The main purpose of IP is to handle all the functions related to routing and to provide a network interface to the upper-layer protocols, such as TCP from Transport Layer. Applications use this single protocol in the layer for anything that requires networking access. Network Layer is responsible for transmitting datagrams hop by hop, which sends from station to station until the messages reach their destination. Each computer should have a unique IP address assigned as an interface to identify itself from the network. When a message arrives from Transport Layer, IP looks for the message addresses, performs encapsulation and add a header end to become a datagram, and passes to the Data Link Layer. As for the same at the receive side, IP performs decapsulation and remove network layer header, and then sends to the Transport Layer. The IP implements datagram fragmentation, so that delivery of packets to the destination increases. Once the large packet is fragmented into smaller Packets then they can travel across a packet switched network without tieing up a communications link. Multiple conversations between different parties can therefore share a single communications link. If any single packet is lost, it can be retransmitted instead of having to start the entire conversation all over again. When a device receives an IP packet it examines the destination address and determines the outgoing interface to use. This interface has an associated Maximum transmission unit that dictates the maximum data size for its payload. If the data size is bigger than the Maximum transmission unit then the device must fragment the data. Reassembly module assembles the fragmented IP packets into one full IP packet so that it can be delivered to the higher layer protocol. Here packet is not being delivered to the next higher layer till entire packet arrives.

The hardware solution is available which is based on the partially packet reassembly for intrusion detection. When arrived each packet of fragment packets, this packet is merged into previously arrived packet. This method is divided into three steps: packet store, merge and pattern matching. In this method there is assumption that packets will arrives only in the ordered form [1].

The software solution is the simplest available for monitoring the network. Apart from being inexpensive and being flexible, this solution suffers from serious limitations. It needs experts to install / setup the device**,** it needs maintenance, and it does not present a business model for an expert organization due to very limited IP protection. Therefore there is a need for hardware device which overcome the drawbacks of the software solution.

II. NETWORK MONITORING DEVICE

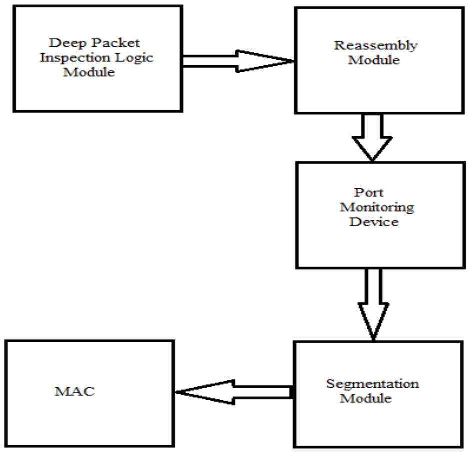


Figure 1: Monitoring Device

The Network monitoring Device is positioned between the switch and the router (which connects to the Internet). Monitoring device also has a capability to record the traffic which can be used for analysis.

Deep packet inspection logic (DPI) receives the packets from different source, this module inspects all the headers (Data link, Internet, Transport) and all the inspected fields are updated in the one processing header. This processing header is attached to each packet and is different for each packet. Deep packet inspection module forwards the packet with only processing header without any Data link, Internet and Transport headers to reassembly module.

Reassembly module can receive the packets from different sources and it must have the capability to handle the data, Here it analysis the processing header to differentiate the packets from different sources. Here original message is constructed by assembling packets belonging to the respective source only. Upon constructing the original message this block removes the processing headers received from the deep packet inspection block and attaches the only one processing header for original message which contains the useful information for the port monitoring device.

Port monitoring device analyses the processing header to decide whether the original message is allowed to pass or it should be blocked. Here entire message may be recorded for further analysis. If original message is blocked then it should inform the segmentation module so that it can erase the stored information related to blocked message. If the original message is allowed to pass then it should be forwarded to segmentation module.

Segmentation module upon receiving the original message divides the original message into fragments similar to fragments received at the reassembly side and before sending the fragments out it should attach all the headers like data link header, internet header, transport header to the respective fragment so that nobody knows the existence of this device.

III. CONTROL SECTION (PROCESSING HEADER)

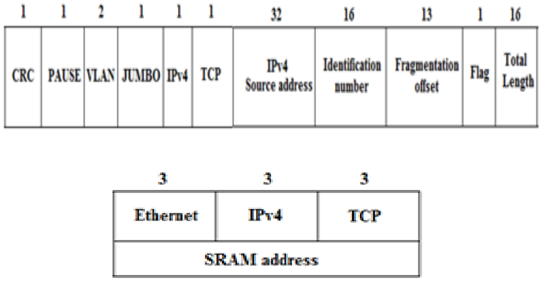
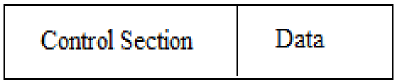


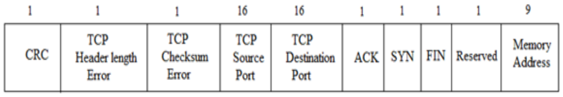
Figure 2: DPI generated Processing Header

Deep packet inspection logic analyses all the incoming packets from different sources. After analyzing it will update the control section fields appropriately and removes the incoming packets headers and stores in the memory, this memory address is all present in the control section. Updated information is used by Reassembly Module for arranging the packets. Control section contains fields like CRC error, PAUSE frame, VLAN frame, Total length field, fragmentation offset, IPv4 address, Jumbo frame, TCP flag, IP flag, identification number, flag, and Memory addresses. All these fields belong to standard header structure of IP, TCP packets [4].



Control Section	Data

Figure 3: Packet format



1	1	1	16	16	1	1	1	1	9
CRC	TCP Header length Error	TCP Checksum Error	TCP Source Port	TCP Destination Port	ACK	SYN	FIN	Reserved	Memory Address

Figure 4: Reassembly module generated processing Header

Reassembly module receives the packet as shown in figure 3. When the helps of control section it will arrange all the packets in order and constructs the original packet. To this packet it will attach a new control section which is shown in figure 4. Here control section contains the information regarding source port address, destination port address, acknowledge flag, synchronizing flag, finish flag and memory address of headers. This control information is very useful for the port monitoring device because with the help of this it will decide whether all allow the packet to get passed or to block permanently.

IV. PACKET SEGMENTATION AND REASSEMBLY

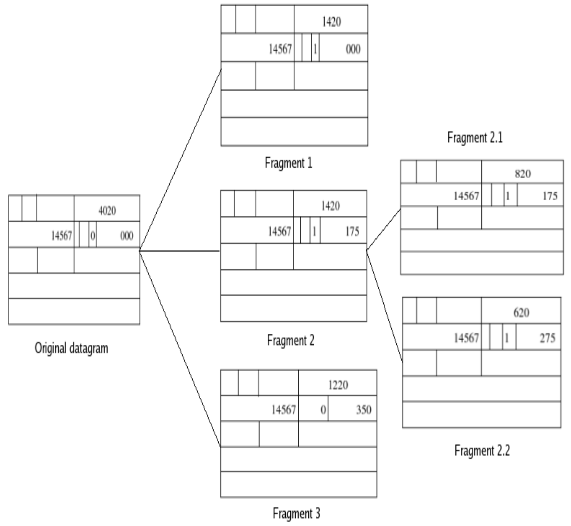


Figure 5: Detailed Fragmentation Example

The field that is related to fragmentation and reassembly of an Ipv4 datagram are the identification, flags, and fragmentation offset. Notice that the value of the identification field is same in all the fragments and also the value of the flags field with the more bit set for all the fragments except the last. Also, the value of the offset field for each fragment is shown in figure4.

The figure4 also shows what happens if a fragment is fragmented. In this case the value of the offset field is always relative to the original datagram. For example, the second fragment is itself fragmented later to two fragments of 800 bytes and 600 bytes, but the offset shows the relative position of the fragments to the original data. If each fragment follows the different path and arrives out of order, the final destination host can reassemble the original datagram from the fragment received by using the following strategy: reassemble the original datagram from the fragment received by using the following strategy:

- The first fragment has an offset field value to zero.
- Divide the length of the fragment by 8. The second fragment has an offset value equal to that result.
- Divide the total length of the first and second fragments by 8. The third fragment has an offset value equal to that result.
- Continue the process. The last fragment has a more bit value of 0.

When a receiver detects an IP packet where either of the following is true:

- “More fragment” flag set.
- “Fragment offset” field is non zero.

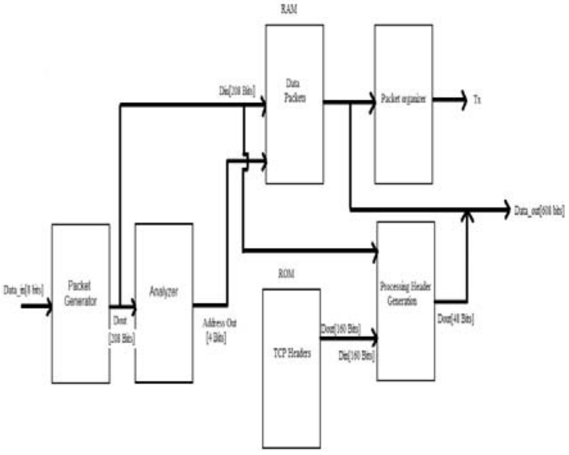


Figure 6: Reassembly Module

Then the receiver knows the packet is a fragment. When the receiver receives a fragment with the more fragments flag set 0 then it knows the length of the original data payload since the fragment offset multiplied by 8(bytes) plus the data length is equivalent to the original data payload size [4].

Packet generator receives data in the form of processing control information and data without any headers. When entire packet is received then it is delivered to the analyser and RAM memory. Analyser work is to separate the packets coming from different sources, it will enable the write signal of RAM according to it. Packet organizer work is to arrange the packets in order so the original packet is constructed and it also identifies the start of the packet and end of the packet then it raises the Tx signal which indicates packet is ready for the transmission. Processing header generation Block generates the processing header as shown in the figure 3 by accessing the TCP headers ROM and attaches with the constructed packet and send it to the Port Monitoring Device.

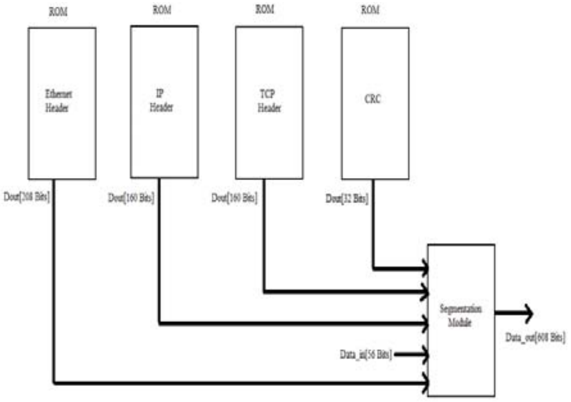


Figure 7: Segmentation Module

Segmentation module receives the data in the form of ROM memory addresses and data. Depending upon the address it will access the headers stored in the ROMs to attach with data to convert it into original segmented packet that was received at the input of the deep packet inspection logic block.

VI. SIMULATION RESULTS

IP packet Reassembly and Segmentation Design is done in Verilog. Implementation is done on the Spartan 3E FPGA kit. For implementing the design packets headers are length are maintained standard and data size is reduced because of hardware limitation of the FPGA kit.

Table 1

Results of Reassembly Module

Logic Utilization	Used	Available	Utilization
No. of slice registers	797	93120	0%
No. Of slice LUTs	314	46560	0%
No. Of fully used LUT-FF pairs	303	808	37%
No. Of bonded IOBs	267	240	111%
No. Of BUFG/BUFGCTRLs	1	32	3%

Table 2

Results of Segmentation module

Logic Utilization	Used	Available	Utilization
No. of slice registers	54	595200	0%
No. Of slice LUTs	28	297600	0%
No. Of fully used LUT-FF pairs	24	58	41%
No. Of bonded IOBs	666	600	111%
No. Of BUFG/BUFGCTRLs	1	32	3%

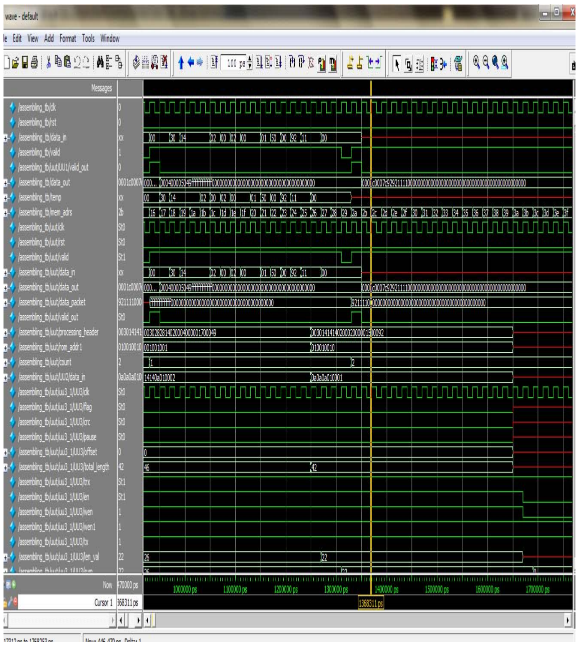

wave - default
le Edit View Add Format Tools Window
100 ps
Messages
/assembling_tb/clk 0
/assembling_tb/rst 0
/assembling_tb/data_in xx
/assembling_tb/valid 1
/assembling_tb/uut/UU1/valid_out 0
/assembling_tb/data_out 0001c0007
/assembling_tb/temp xx
/assembling_tb/mem_adrs 2b
/assembling_tb/uut/clk St0
/assembling_tb/uut/rst St0
/assembling_tb/uut/valid St1
/assembling_tb/uut/data_in xx
/assembling_tb/uut/data_out 0001c0007
/assembling_tb/uut/data_packet 921111000
/assembling_tb/uut/valid_out St0
/assembling_tb/uut/processing_header 003014141
/assembling_tb/uut/rom_addr1 010010010
/assembling_tb/uut/count 2
/assembling_tb/uut/UU2/data_in 0a0a0a010
/assembling_tb/uut/uu3_1/UU3/clk St0
/assembling_tb/uut/uu3_1/UU3/flag St0
/assembling_tb/uut/uu3_1/UU3/crc St0
/assembling_tb/uut/uu3_1/UU3/pause St0
/assembling_tb/uut/uu3_1/UU3/offset 0
/assembling_tb/uut/uu3_1/UU3/total_length 42
/assembling_tb/uut/uu3_1/UU3/trx St1
/assembling_tb/uut/uu3_1/UU3/en St1
/assembling_tb/uut/uu3_1/UU3/wen 1
/assembling_tb/uut/uu3_1/UU3/wen1 1
/assembling_tb/uut/uu3_1/UU3/tx 1
/assembling_tb/uut/uu3_1/UU3/len_val 22
Now 470000 ps
Cursor 1 368311 ps
1000000 ps
1100000 ps
1200000 ps
1300000 ps
1400000 ps
1500000 ps
1600000 ps
1700000 ps
1368311 ps

Figure 8: Reassembly Waveform

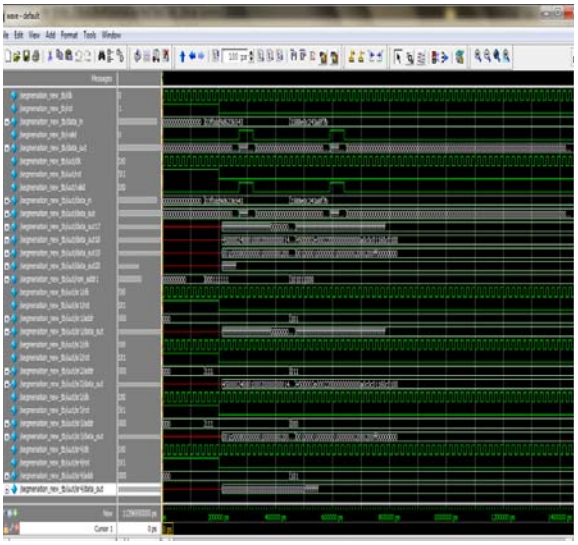


Figure 9: Segmentation Waveform

VII. CONCLUSION

This work explains the implementation of reassembly and segmentation module for TCP/IP packet in the internet. These two functions, although rarely mentioned in literature, are very important for efficient work of the network monitoring device. As we have seen from implementation results, slices count influence efficiency of the network monitoring device – The smaller the number is, the design is more efficient regarding resources and speed.

ACKNOWLEDGEMENT

I would like to thank R Shankar, CEO of Whizchip design technologies for their technical guidance and giving an opportunity to carry out the work.

REFERENCES

[1] Bo Heung Chung, Jae Deok Lim, Seung Ho Ryu, Young Ho Kim, Ki Young Kim "Fragment Packet Partial Re-assembly Method for Intrusion Detection", in proceeding 8th International conference on Advance Communication Technology 2006(ICACT), pp. 120 – 122 , Feb 2006.

[2] Patricia Gilfeather, Todd Underwood "Fragmentation and High Performance IP", IEEE International Conference on Cluster Computing, April 2001.

[3] David C. Black, Jack Donovan, Bill Bunton, Anna Keist "System on Chip Packet Processor for an Experimental Network Service Platform" IEEE/GLOBECOM, pp. 3933- 3937, 2003.

[4] Behrouz A. Forouzan and Sophia Chung Fegan, "Data Communication and Networking," Tata McGraw Hill Publications, 4th Edition, 2007.

[5] Kirill Levchenko, Ramamohan Paturi, George Varghese"On Difficulty of Scalably Detecting Network Attacks", in Proceedings Computer Communication Security Washington dc, October 2004.

Simulation of Proximity Coupled Feed Fractal Tree Antenna for WLAN

V. R Anitha & Sivasubramanyam Medasani

Dept. of Electronics & Communication Engineering, Sree Vidyanikethan Engineering College, Tirupati

Abstract – In radar and modern communication systems the demand on multi functional antennas is increasing. The requirements for these antennas are the abilities to have multiband, frequency independent or log-periodic behavior that has been attributed to the self-similar antenna geometry. Wireless applications, particularly with multiple resonances, put new demands on antennas pertaining to size, gain, efficiency, bandwidth, and more. One promising approach in this regard is to use fractal geometries to find the best distribution of currents within a volume to meet a particular design goal. For improving the inherently narrow bandwidth of patch fractal antenna proximity coupled feed is implemented. For this purpose it is very effective to use an electrically thick substrate. Several studies for feeding the microstrip antenna with a thick substrate have been reported, such as the L-shaped probe [1] and the capacitive probe-fed structure [2]. Recently the proximity-coupled microstrip antenna with a linear slot in ground plane has also been proposed [2][3]. For increasing the frequency bandwidth, a proximity-coupled feed is used and an H-shaped slot is placed in the ground plane. The % bandwidth is 42.5% which is maximum than probe and microstrip feed.

I. INTRODUCTION

The emergence of antennas with fractal geometries has given an answer to two of the main limitations stated by Werner (1999) of the classical antennas, which are the single band performance and the dependence between size and operating frequency. The term “fractal”, means broken or irregular fragments. It was originally coined by Mandelbrot (1983) to describe a family of complex shapes that possess an inherent self-similarity or self-affinity in their geometrical structure. Jaggard (1990) defined fractal electrodynamics as an area in which fractal geometry was combined with electromagnetic theory for the purpose of investigating a new class of radiation, propagation, and scattering problems. One of the most promising areas of fractal electrodynamics research is in its application to antenna theory and design. With the advance of wireless communication systems and their increasing importance, wideband and low profile antennas are in great demand for both commercial and military applications. Multiband and wideband antennas [5] are desirable in personal communication systems, small satellite communication terminals, and other wireless applications. Traditionally, a wideband antenna in the low frequency wireless bands can only be achieved with heavily loaded wire antennas, which usually means that different antennas are needed for different frequency bands. Recent progress in the study of fractal antennas suggests some attractive solutions for using a single small antenna operating in several frequency bands. The self-similar properties [6] of certain fractals result in a multiband behavior of the antennas while, the highly convoluted shape of these fractals makes possible the reduction in size, and consequently in mass and volume, of certain antennas as investigated by Puente (1998). These reductions can make possible to combine multimedia, communication and tele-detection functionalities in a reduced space like a handy phone, a wristwatch or a credit card e.g. a fractal antenna can provide GPS (Global Positioning System) services within a conventional mobile cellular phone. In the last few years, the fast growing development of mobile communication brought the need for devices that require their components to be ever smaller and lighter, capable of adjusting its frequency of operation and to operate in a multiband mode. Some recent results by Puente (1998) and Puente (2001) showed that fractal antennas have excellent multiband properties [7], [8] and low resonant frequencies. Radiation efficiency and impedance bandwidth decrease with the size of the antenna, making small antennas inefficient by nature, for these effects are accompanied by high currents in the conductors, high ohmic losses and large values of energy stored in the antenna near field. The inefficient performance of small antennas is summarized by the high values of its quality factor *Q* [9], as predicted by the fundamental limit and stated by Chu (1948) and McLean (1996). In order to meet the following attributes for antenna designs, i.e. the compact size, low profile, conformal and multiband or broadband, a number of approaches for designing multi-band antennas have been summarized by Maci (1997).

II. PROPOSED FRACTAL TREE ANTENNA DESIGN AND CONFIGURATION

The antenna geometry, fractal tree is a plane fractal from Rectangle and Triangle patches which is connected by using wire as shown in Figure3. The parent branch is one pair of rectangle and triangle. The child branches are splited from parent branch by 0.5 scale factor. The branch angle θ = 60^0. A 5-iterration tree is shown in the figure. Length of main stem L1 = 30mm, width of the stem W1 = 10mm, substrate height h = 3.2mm or 1.6mm, permittivity of the substrate is 2.2 Rogers RT/ Duroid 5880(tm) or 4.4 Epoxy FR4, length and width of each iteration is given in Table 1. The geometry shows multi-band and wide bandwidth behavior compared with the existing square patch of same dimensions. The antenna is considered to have infinite ground plane.

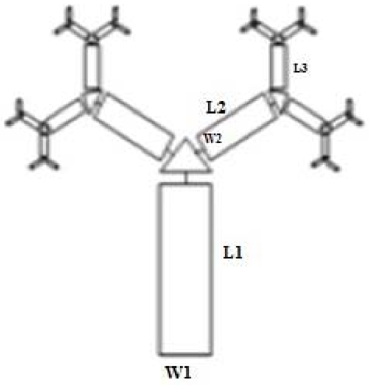


Fig. 1: Novel Design of Fractal Tree Antenna

Iteration Number	Length in mm	Width in mm
First	L1 = 30	W1 = 10
Second	L1 = 15	W1 = 5
Third	L1 = 7.5	W1 = 2.5
Fourth	L1 = 3.75	W1 = 1.25
Fifth	L1 = 1.875	W1 = 0.625

Table 1

III. PROXIMITY COUPLED FEED RESULTS

In this section the results of the new configurations are shown. The antenna is simulated for the proposed configuration. The effect on the shape of the radiating patch under the proposed feed scheme is investigated. The antenna can be fed using four different feeding methods, Coaxial, Microstrip, Aperture Coupled and Proximity Coupled. In this paper the Proximity coupled feeding is discussed. The simulated antenna shows wide band operation for WLAN for 1GHz-5GHz application. Five iterations were examined. The VSWR performance, Return loss, Radiation pattern and Gain for each iteration were simulated and discussed below. For improving the inherently narrow bandwidth of patch fractal antenna proximity coupled feed is implemented. The performance is demonstrated by simulation. The excitation method used is Proximity Coupled Feed Figure 2. For improving the inherently narrow bandwidth of a microstrip antenna, it is very effective to use an electrically thick substrate. Several studies for feeding the microstrip antenna with a thick substrate have been reported, such as the L-shaped probe [1] and the capacitive probe-fed structure [2]. Recently the proximity-coupled microstrip antenna with a linear slot in ground plane has also been proposed [2][3].

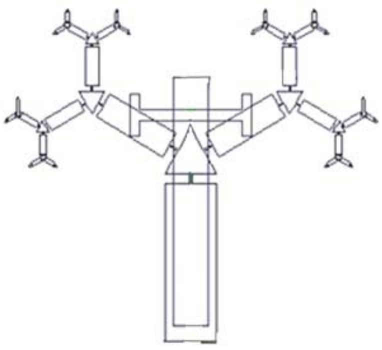


Fig. 2 : Proximity Coupled Novel Design of Fractal Tree Antenna

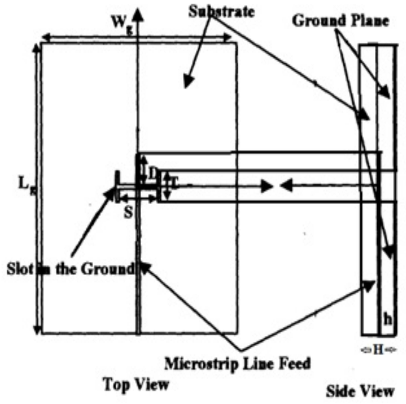


Fig. 3 : Geometry of microstip feed line and substrate

It should be noted that the proximity-coupled microstrip antenna without a slot in the ground plane is hard to achieve impedance matching. Here, we miniaturize a microstrip patch antenna with suitable gain and impedance bandwidth. For increasing the frequency bandwidth, a proximity-coupled feed is used and an H-shaped slot is placed in the ground plane. The structure consists of a two layers substrate, Fractal tree microstrip patch is mounted on a upper layer substrate and feed line is the between two layers. The substrate consists of Rogers ultralam 1300(tm) material (dielectric constant $\square_r$ = 3 and loss tangent=0.003) in the Upper and lower layer. The microstrip feed line is designed as a 50Ω line and symmetrically located above an H shaped slot. For the purpose of analyzing the antenna, finite ground plane of dimensions l00mm x 120mm is used. The open circuited tuning stub (D), slot parameters (S and T), and substrate thickness H and h is defined as follow: D=6mm, T=8mm, S=20mm, H=3.5mm and h=1.75mm. These parameters dominate the impedance and radiation characteristics of the antenna [4]. Figure 3 shows the geometry of substrate and feed line of microstrip antenna. Return loss, VSWR, VSWR bandwidth, and direction pattern is plotted in Figure 4, 5.

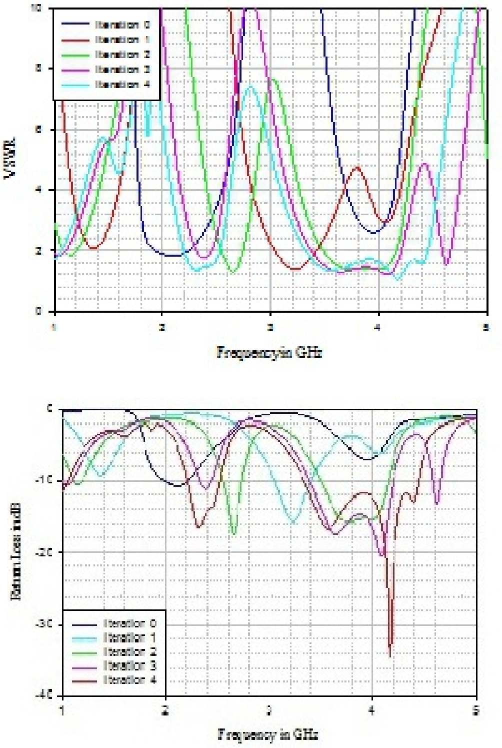


Fig. 4 : VSWR and Return Loss characteristic

Figure 4 shows the good result for fourth iteration. Bandwidth up to 9.5% for 2.4 GHz band and 31.5% for 3.6 GHz band, which meet the required bandwidth specification of 2.4/3.6 GHz WLAN standard. The radiation pattern of fractal antenna is nearly omnidirectional in azimuth plane throughout the operating frequency. Bandwidth can be enhanced by using other enhancement techniques [4]. The effect of change in feeding method on the characteristics of the antenna and the bandwidth along with corresponding change in the dimensions of key design parameters for each case are summarized in Table 2 and Table 3. The VSWR, Return Loss, Radiation Pattern and Gain graphs for High Bandwidth as shown in Figure 6, 7.

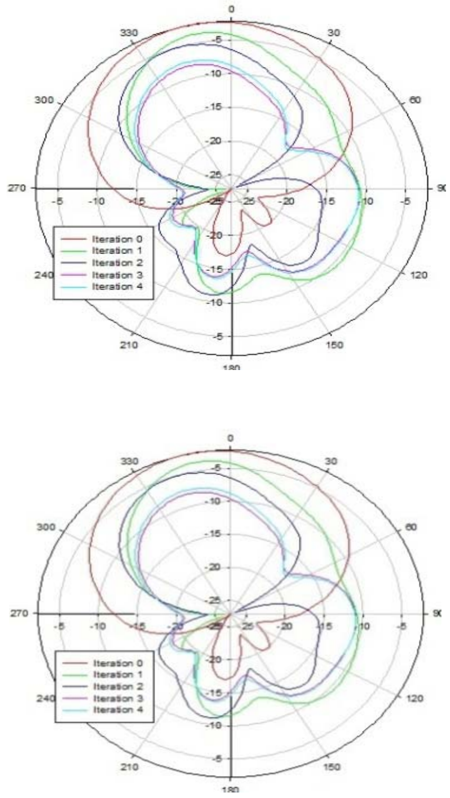


Fig. 5 : Radiation Pattern and Gain characteristic

To improve the impedance matching, the H-slot is placed along the patch diameter, without changing the main patch dimension. The proposed model reduces the resonant frequency more than other fractal geometries and it’s concluded that the miniaturization is better done, although the bandwidth impedance reduces but it could be compensated by parameters which have been stated in antenna configuration section. Due to proximity (electromagnetically) coupled structure when the fractal are applied to miniaturization, the impedance and gain bandwidth are suitable value and it could be applied in the array antenna design. Here we get 42.5% bandwidth which is maximum than probe and microstrip feed.

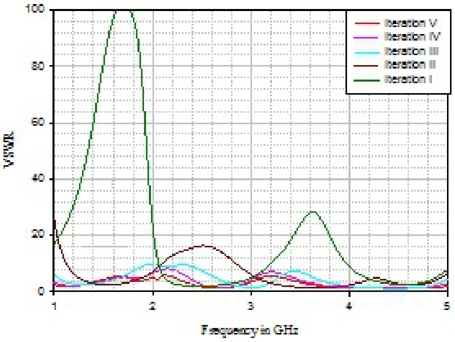


Fig. 6 : VSWR and Return Loss characteristic

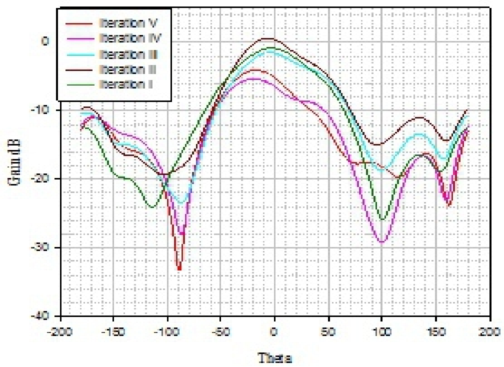
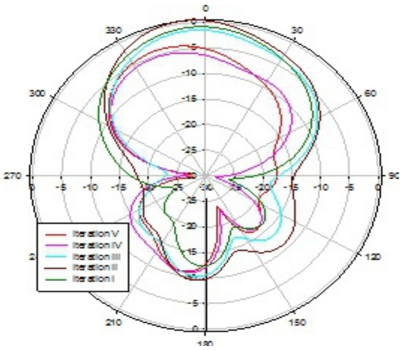


Fig. 7: Radiation Pattern and Gain characteristic



	H = 1.6				H =3.2			H = 3.5
	$\varepsilon_r = 2.2$		$\varepsilon_r = 4.4$		$\varepsilon_r = 2.2$	$\varepsilon_r = 4.4$		$\varepsilon_r = 3$
Resonance Frequency (GHz)	1.77	2.7	1.86	2.8	2.64	1.92	2.93	2.31
VSWR	1.76	1.08	1.35	1.45	1.23	1.09	1.24	1.4
% Bandwidth	2	16.5	8.5	10	22.5	13.5	17.5	16.5
Return Loss (dB)	-12	-28	-23	-15	-35	-27	-19	-35

Table 2: Parameters for $f_c = 2$GHz

	H = 1.6					H =3.2					H = 3.5	
	$\varepsilon_r = 2.2$			$\varepsilon_r = 4.4$		$\varepsilon_r = 2.2$		$\varepsilon_r = 4.4$			$\varepsilon_r = 3$	
Resonance Frequency (GHz)	1.68	2.54	3.9	1.22	2.8	2.6	4.1	1.9	2.94	3.5	2.32	4.16
VSWR	1.4	1.22	1.52	1.32	1.04	1.25	1.11	1.12	1.23	1.15	1.35	1.04
% Bandwidth	4	10.7	14	3.5	7	15	42.5	10	12.7	7	11	37.5
Return Loss (dB)	-15	-20	-14	-17	-13	-17	-35	-16	-17	-35	-17	-35

Table 3: Parameters for $f_c = 3$GHz

IV. CONCLUSIONS

A novel tree shaped fractal antenna is designed using rectangular, triangular microstrips and wired structures. It is observed that the resultant antenna is compact in size. The proposed novel design provides the bandwidth up to 42.5% using proximity coupled feed technique which is maximum when compared to probe and microstrip feed. The proposed antenna is simulated for 1-5GHz frequency, which finds its applications in WLAN. The novel fractal tree geometry describes a multiband behavior of fractal antenna. Changing the height of the substrate the fractal tree patch will control the frequency spacing and input impedance.

REFERENCES

[1] Joshua S. Petko and D. H. Werner. Dense 3-d fractal tree structures as miniature end-loaded dipole antennas. IEEE, 2002.

[2] B.Atrouz H. Kimouche eI, M.Bitchikh. Novel design of a fractal monopole antenna for wireless communications. IEEE transaction of Antenna Wave Propagation, 2008.

[3] A. Aggarwal and M. V. Kartikeyan. Pythagoras tree: A fractal patch antenna for multi-frequency and ultrawide bandwidth operations. Progress In Electromagnetics Research C, Vol. 16:25-35, 2010.

[4] V. R. Anitha Janakiraje S. Bhosale, Sanjay V. Khobragade. Proximity-coupled novel design of fractal tree antenna miniaturization. XXXth URSI General Assembly and Scientific Symposium, 2011.

[5] M. Sindoy, G. Ablat, and C. Sourdois, "Multiband and wideband propetties of printed fractal branched antennas," Electronics Letters, vol. 35, pp. 181-182, 1999.

[6] R. G. Hohlfeld and N. Cohen , "Self-Similarity and the Geometric Requirements for Frequency Independence in Antennae",Fractals, vol. 7, no. 1, pp. 79-84, 1999.

[7] C. Puente, J. Romeu, R. Pous, X. Garcia, F. Benitez,"Fractal Multiband antenna based on the Sierpinski gasket", IEE Electronic Letters, vol. 32, pp 1-2, 1996.

[8] C. Puente, J. Romeu, R. Pous, A. Cardana, "On the behavior of the Sierpinski multiband antenna",IEEE Transactions on Antennas and Propagation , vol. 46, pp 517-524, 1998.

[9] J. S. Mc Lean, "A Re-examination of the fundamental Limits on the Radiation Q of electrically small Antennas", IEEE trans. On Antennas and Propagation, vol. 69, pp- 672-676, 1996.

[10] J. Gianvittorio and Y. Rahmat-Samii, "Fractal Patch Antennas: Miniaturizing Resonant Patches," IEEE International Symposium on Antennas and Propagation and USNCNRSI National Radio Science Meeting MRSI Digest, Boston, Massachusetts, p. 298, 2001.

[11] J. P. Gianvittorio and Y. Rahmat-Samii," Fractals Antennas: A novel Antenna Miniaturization Technique, and Applications", IEEE Antennas and Propagation Magazine, vol. 44, pp 20- 36, 2002.

◈◈◈

Corner and Feature Points Based Image Mosaic Construction

Rajendranaidu Dindi & Chiranjeevulu Divvala

Dept. of Electronics & Communication Engineering, Sri Sivani College of Engineering,
Chilakapalem, Srikakulam, Andhra Pradesh, India

Abstract **–** Mosaic is that a picture or decoration made of small, usually, usually colored pieces of inlaid stone, glass, etc. in this method we use feature matching and image fusion process. the advantage of this method is reduce the rate of feature mismatches. we take a feature ,which is corners in two overlapping images, which are detected and matched in order to derive a transformation matrix to align the images to be combined. in order to guarantee a reduced rate of mismatching, a similarity definitude algorithm is employed. next we use weighted fusion process to get a smooth mosaic. Experimental results, which are conducted using indoor images, show that our approach can effectively obtain an accurate and seamless fused image**.**

Keywords *- 1. feature detection, 2. Feature matching, 3. image transformation and 4. Image fusion.*

I. INTRODUCTION

Image mosaic is a process that joins two or more images into a larger image and has been widely used in many arenas, such as photogrammetry, computer vision, and computer graphics [1][2][3]. Luck man et al. [1] described an image mosaic generation method for constructing a true-colour and high-spatial-resolution panoramic view of Antarctica using 1100 individually selected Antarctica scenes. Murray et al. [2] applied mosaics in a vision-based mobile robot system which combined two video sequences obtained by two cameras mounted on the robot into one to get a larger angle of view. Garcia et al. [3] developed a positioning system for an underwater vehicle that also employed image mosaic techniques. In general, it can be seen in these work that there are four steps involved in the construction of image mosaics: 1. feature detection, 2. Feature matching, 3. image transformation and 4. image fusion.

Feature detection finds the candidate features to be matched in two images. Geometrical features including lines, edges and corners are frequently used as detection targets [4][5]. Among these geometrical features, corners are widely used. There are broadly two approaches applied in corner detection, which are Moravec feature detection and Harris corner detection. Moravec feature detection calculates the quadratic sum of grey value differences between adjacent pixels with horizontal, vertical, diagonal and back-diagonal directions [6]. However, the response using Moravec detector is anisotropic and noisy. To overcome this problem, Harris et al. [7] introduced a method called Harris corner detector, which is shown to perform with good consistency on natural imagery.

The basic idea of feature matching is to declare that two features in separated images correspond to the same object in the scene. There is a number of possible feature matching methods where a common one adopted is the use of distance measures. Features are deemed similar if their statistical distance is less than or equal to a particular threshold [8]. Feature matching approaches can be further divided into two categories, pixel-based methods and histogram-based methods see for example, Kwok et al. [9]. Pixel-based matchings are simple while histogram-based matchings are more robust but the computation load is demanding.

The essence of image mosaic is to combine two or more images into a seemingly larger image. The images to be combined have to go through a transformation process in order that they can be closely aligned. The homograph transformation is an attractive candidate under most image mosaic construction situations [10]. This method requires matched features to lie on a common plane and the camera can be freely translated and rotated as the task requires.

The final stage in generating an image mosaic is to fuse the individual and aligned images into a larger one. In this process, minimum weighted distance (MWD) method [10] and overlapped area transition (OAT) method [11] are used widely. In MWD, overlapped images pixels are weighted and combined according to their distances to the overlap boundaries. On the other hand, the weights in OAT could be simply determined if the overlapping areas are parallel either in the horizontal or vertical direction.

In our work, the Harris corner detector is first applied to find corners, since this approach is straightforward, less time-consuming and provides satisfactory results. Then, in order to improve feature matching reliability, a combination of three similarity measures; the cosine similarity, the Euclidian distance and the similarity definitude, is proposed. The transformation matrix is then obtained by using all the correspondence features rather than four pairs of points because it can reduce the numerical error due to impressions in feature locations. Finally, we choose the overlapped area transition method for image fusion, as we find that this method not only can eliminate seams between two images, but also can alleviate alignment errors.

This paper is arranged as follows. The feature detection method is introduced in Section II; and the feature matching approach is presented in Section III. Then, transformation methodology and image fusion are introduced in Section IV and Section V respectively. Finally, experimental results are illustrated in Section VI and the paper is concluded in Section VII.

II. FEATURE DETECTION

In order to detect features in images, the Harris corner detection procedure is applied. This algorithm detects corners by considering the sum of squared differences (SSD) between the image patch over an area (u, v) (template) and shifting it by (x, y). Since computing SSD is computationally expensive, Harris and Stephens [7] presented the following equation instead to find corners in images:

$$M = \det(A) - k\mathrm{Tr}^2(A), \qquad (1)$$

where A is defined as Harris matrix, see equation (2), and det(A) is the determinant of matrix A, Tr(A) is the trace of matrix A. The value of *k* has to be decided empirically. In our experiment, we chose *k* to 0.04, and

$$\mathbf{A} = \begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix}, \qquad (2)$$

where I is the pixel value at each pixel coordinate. I_x and I_y is the image gradient along x and y axis. Two large eigenvalues for matrix M indicate a corner is detected. After applying Harris corner detector on images, corners found can be treated as feature points for mosaicing, Fig. 1

Illustrate the found corners marked by red plus signs.

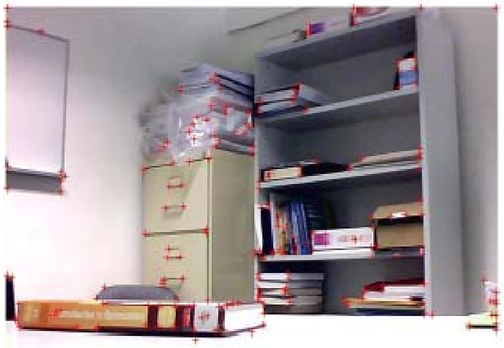


Fig. 1 : Harris corners detection

III. FEATURE MATCHING

After obtaining feature points from the above process, we proceed to pair these points in the template with their counterparts in the other image. There are several matching approaches that can be adopted for this purpose [9]. owever,experiment results show that they may be unreliable if only a single matching scheme is applied. Hence, Bei et al. [10] introduced a method which combined similarity matching approach and distance matching approach. The combination resulted in a reduce mismatched rate. Therefore, a feature matching approach based on cosine similarity and Euclidean distance is used in this work.

Cosine similarity is a measure of similarity between two vectors of n-dimensions by finding the cosine of the angle between them, which is defined as follows:

$$s = \frac{f_1 \cdot f_2}{|f_1||f_2|}, \qquad (3)$$

where f_1 and f_2 are the feature vectors, $f_1 \cdot f_2$ represents the dot product between two vectors, $|f_1|$ and $|f_2|$ are the magnitudes of vectors.

Euclidean distance, based on the multi-dimensional geometric distance, is also a conventional measure to estimate the difference between features. It can be expressed as follows:

$$D = \sqrt{\sum_{k=1}^{n}(X_k^1 - X_k^2)^2}, \qquad (4)$$

Where X_k^1 and X_k^2 are two feature vectors in the template image and the image to be matched.

In an ideal situation, if two feature points represent the same corner, the similarity s would be 1 and Euclidean distance D equals to 0. However, these values may not be attained in real-life images.

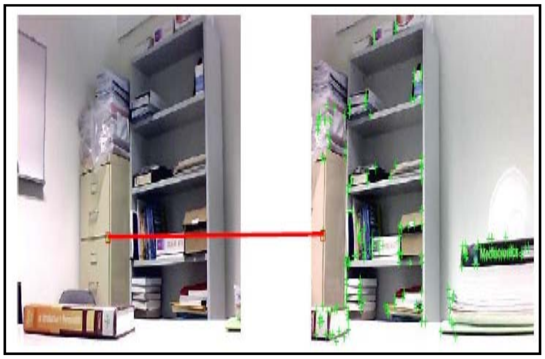
As we can see in Fig. 2, the similarity of matched points in Fig. 2(c) is 0.9754, while the climax of unmatched similarity in Fig. 2(d) is 0.9841 (linked by red lines in Fig. 2(a) and Fig. 2(b)). Apparently, the unmatched similarity is much higher than the matched one. Hence, we cannot solely use similarity value as threshold, and a more reliable criterion should be proposed. In [10], a similarity definitude, which represents the discrimination between the maximum similarity value and second maximum, was introduced. The basic idea of this method is that, in the matching process, if more than 2 similarity values are similar and close to 1, we cannot determine which pair is match. The similarity definitude equation can be expressed as follows.

$$U = 1 - \frac{S_1(x, x')}{S_2(x, x')}, \quad (5)$$

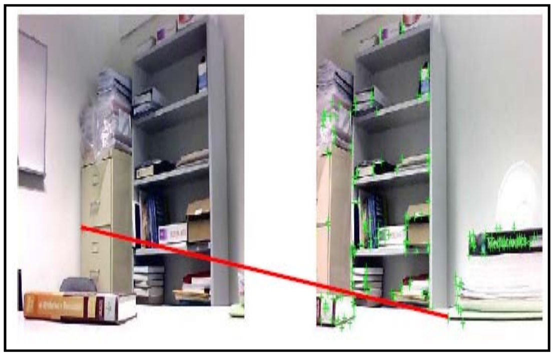
Where, $S_1(x, x')$ stands for the value of the maximum similarity value, $S_2(x, x')$ represents the second largest similarity.

Although this method improves feature matching accuracy, it still has a minor drawback. Specifically, in most situations,the cosine similarity value is close to 1 (higher than 0.95) if two candidate points are matched. Meanwhile, there are also several similarity value higher than 0.9, but are not matched. This makes the definitude value close to zero no matter matched points are found or not. For example, as is shown in Fig. 2, the definitude in Fig.2(c) is 0.04, while the counterpart in Fig.2(d) is 0.01. Since this value is very close to zero, it is difficult to choose a reliable threshold. A slightly change of definitude threshold will produce inaccurate matching. In our system, a definitude equation is given as:

$$U = 1 - \frac{S_2(x, x') - \epsilon}{S_1(x, x') - \epsilon}, \quad (6)$$



(a)



(b)

Fig. 2 : An example of matched and unmatched features (a) Images with matched features (matched features: red square, unmatched features: green asterisk), (b) Images without matched features

where € is a fixed value we chose. In our experiment, € = 0.95.

The main advantage of our algorithm is that the definitude value is no longer near zero; because we use the distance between the similarity value and a fixed number in the equation, instead of the similarity value itself. There is no denying that some of the correspondence pairs will be missed when the similarity definitude equation is applied in matching process. This deficiency may be attributed to the use of only four correspondence pairs in the conventional transformation process in aligning images. Therefore, the problem we concerned about is to find more than four correct pairs of correspondence points.

IV. TRANSFORMATION

After finding correct correspondence points in the template image and the image to be matched, the next process is to calculate the relationship between these two images. More specifically, an accurate transformation model should be constructed at this stage. In our paper, the homograph matrix is applied in transformation. The transformation matrix is:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (7)$$

Where x, y are coordinates in the template image and x′, y′ are coordinates in the image to be matched. m_{11}, m_{12} …, m_{23}, m_{33} are eight parameters to be estimated.

This kind of transformation is also called planar homography which has eight degrees of freedom [12]. The planar homography is appropriate under three conditions [13]: (i) a plane viewed under arbitrary camera motion; (ii) a scene viewed by a camera resolve round its optic center; (iii) a scene viewed by a camera zoom.

Since the transformation has eight degrees of freedom, the transformation matrix has eight parameters as shown in equation (7). Hence, if we want to solve this transformation matrix, only 4 pairs of correspondence points are required.However, if only 4 pairs of data are used to calculate the matrix, it is possible that several of these pairs are not correct, and lead to an incorrect matrix. In order to eliminate the error of transformation, all the correspondence points are used to solve the transform matrix by the least squares method which is a common practise for finding the best fit matrix. After this procedure, an accurate relationship is constructed between two images to be combined.

V. IMAGE FUSION

Since the two images to be spliced are obtained under different illumination conditions, slight color and intensity deviations exist between these two images. In order to eliminate obvious seams and make fusion more smooth, a weighted linear transition approach is introduced. In the overlapped area among two images, assume that the maximum and minimum value along each horizontal scan line is l_{max} and l_{min}. Then, a weight is obtained from:

$$w = \frac{l_{max} - l}{l_{max} - l_{min}}, \tag{8}$$

Where l is the x-axis value along each horizontal scan line.Itmeans that the smaller distance between the pixel in the overlapped area and the left-hand image center, w will have a larger value. After calculating the weight, the pixel intensity in the overlapped area can be estimated by:

$$I(x,y) = wI_1(x,y) + (1-w)I_2(x,y), \tag{9}$$

where $I_1(x, y)$ and $I_2(x, y)$ are the pixel intensity values of the template image and the image to be matched. With the help of this approach, the transition from the first image and the second image is smooth and without obvious seam.

VI. RESULTS

A. *Feature Match Results*

Before matching feature points, the histograms of two images are equalized to reduce the influence taken by illumination differences. In our test, the cosine similarity threshold is set to 0.95; the Euclidean distance threshold is set to 1000. In order to make sure large similarity discrimination, the similarity definitude value is set to 0.4. As is shown in Fig.3(a), if we simply use the first two thresholds, several pairs of corners are mismatched (linked by red lines). When the similarity definitude we introduced is applied in the system, most mismatched points disappear in Fig.3 (b). Admittedly, some of the correct matched points are eliminated in Fig.3 (b); however, there are sufficient correspondence points to generate the transformation matrix.

In our test, we collect 50 pairs of images in different locations, and then compare the matching results between the similarity definitude is applied and not used. The results are illustrated in Fig.4. The first box from left-hand represents mismatched points existed before the similarity definitude is applied, while the second one shows mismatched points existed after the similarity definitude is utilized. It is obvious that there is a considerable decrease of the number of mismatched features after our feature definitude being used. Moreover, the dispersion degree is much lower than before. The experimental result shows that 69% of mismatched points can be eliminated by the similarity definitude algorithm we introduced.

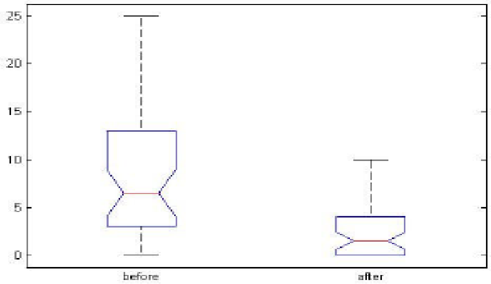


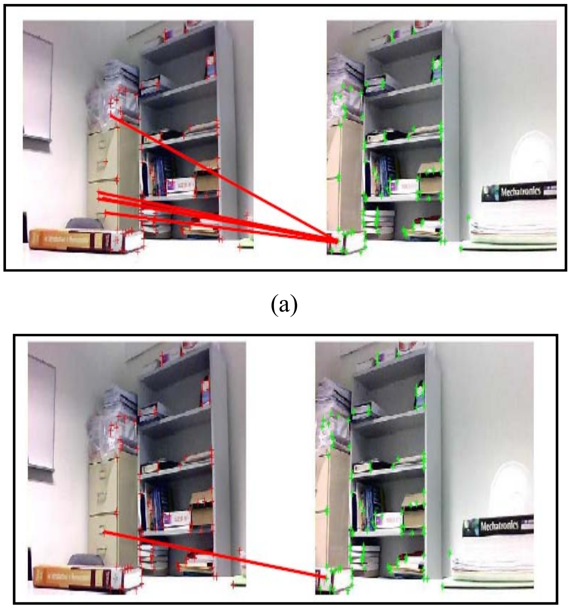
Fig. 4 : The box plot for feature matching results

B. *Image Transformation and Fusion Results*

With all the matched points we detected, we can find therelationship between these two images and fuse them. As we can see from the Fig. 5(a) and Fig. 5(b), we can combine two images together after matched points were detected. Because all the matched points are used to calculate the transformation matrix, although one pair of correspondence features is mismatched in Fig. 3(b) (at the drawer handle in the left-hand image and on the book label in the right-hand image), the result is still satisfactory as illustrated in Fig. 5(b). After comparing the results in Fig. 5(a) and Fig. 5(b), we can notice that there is no obvious abrupt changes between two images after image fusion algorithm being applied. Moreover,in Fig. 5(a), an apparent alignment error can be observed at the pile of books and papers; and the wall beside the bookcase and behind the table seems like not on the same plane. However, after the image fusion process, such problems have been solved effectively as illustrate in Fig. 5(b).

VII.CONCLUSION

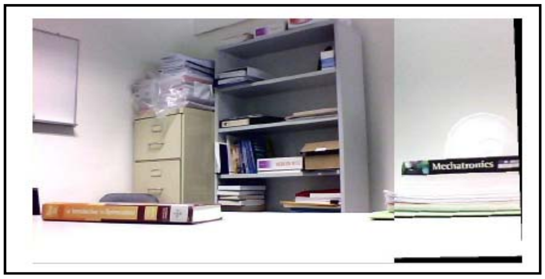
An image mosaic method based on feature match and planar homography is introduced in this paper. Also, we propose a new feature definitude algorithm aimed at reducing mismatched error. In comparison with methods using similarity and Euclidean distance, this method can effectively reduce mismatch rate during the feature matching process. Moreover, we find that the linear transition image fusion approach not only can solve seams problem, but also can alleviate the displacement problem arising from transformation rounding errors. Experimental results show that our method is effective for image mosaic.



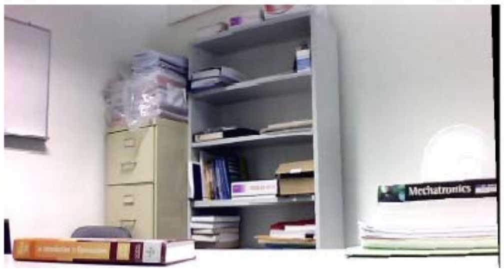
(a)

(b)

Fig. 3. (a) An example of feature match without similarity definitude method, (b) An example of feature match with similarity definitude method.



(a)



(b)

Fig. 5. (a) Image mosaic without image fusion technology we proposed, (b) Image mosaic with image fusion technology we proposed.

REFERENCES

[1] R. Bindschadler, P. Vornberger, A. Fleming, A. Fox, J. Mullins, D. Binnie, S. Paulsen, B. Granneman, and D. Gorodetzky, “The Landsat image mosaic of Antarctica,” Remote Sensing of Environment, vol. 112, no. 12,pp. 4214–4226, 2008.

[2] D. Murray and J. Little, “Using real-time stereo vision for mobile robot navigation,” Autonomous Robots, vol. 8, no. 2, pp. 161–171, 2000.

[3] R. Garcia, J. Batlle, X. Cufi, and J. Amat, “Positioning an underwater vehicle through image mosaicking,” in Proc. IEEE International Conference on Robotics and Automation, 2001. Proceedings 2001 ICRA, vol. 3, 2001.

[4] X. Dai and S. Khorram, “A feature-based image registration algorithm using improved chain-code representation combined with invariant moments,”IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 5, pp. 2351–2362, 1999.

[5] I. Zoghlami, O. Faugeras, and R. Deriche, “Using geometric corners to build a 2D mosaic from a set of images,” in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Citeseer, 1997, pp. 420–425.

[6] Y. Zhiqian and W. Hao, “An image mosaic algorithm of pathological section based on feature points,” in Proc. International Conference on Information Engineering and Computer Science, 2009, pp. 1–3.

[7] C. Harris and M. Stephens, “A combined corner and edge detector,” in Proceedings of Fourth Alvey Vision Conference, Manchester, UK, 1988, pp. 147–151.

[8] S. Antani, R. Kasturi, and R. Jain, “A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video,” Pattern Recognition, vol. 35, no. 4, pp. 945–965, 2002.

[9] N. M. Kwok, M. G. Carmichael, Q. P. Ha, and K. C. Tan, “Statistical decision based gray-level image feature matching,” in Proc. the Eighth International Conference on Intelligent Technologies (InTech’07), Sydney, Australia, 2007, pp. 269–274.

[10] L. Bei and Z. Haizhen, “An algorithm of fabric image mosaic based on sift feature matching,” in Proc. International Conference on Artificial Intelligence and Computational Intelligence, AICI ’09, vol. 3, Shanghai, China, 2009, pp. 435–438.

[11] L. Wei, S. Jin, and C. Wengang, “Image mosaic technology based on overlapped area linear transition method,” in Proc. 2nd International Congress on Image and Signal Processing, Tianjin, China, 2009, pp.1–3.

[12] D. Capel and A. Zisserman, “Computer vision applied to super resolution,”IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 75–86, 2003.

[13] R. Hartley and A. Zisserman, Multiple view geometry in computer vision. Cambridge University Press, 2003.

◈◈◈

A Novel Method for Watermarking using Opaque & Translucent Methods in Videos

Kuna Dhilli[1], Y. Rajyalakshmi[2] & Syam Babu Darsi[3]

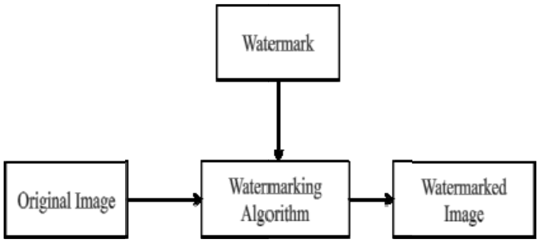
[1&2]Dept. of E.C.E., SSCE, Chilakaplem Jn, Srikakulam, India,
[3]Dept. of E.C.E., AVET, Tagarapuvalasa Jn, Srikakulam, India

Abstract – Watermarking is a popular technique for discouraging illegal copyright and distribution of copyrighted digital image information. One of the important features of the watermarking technique is the lossless visible watermarking, which will preserve the quality of the watermark and watermarked image. A novel method for generic visible watermarking with a capability of lossless video image recovery is proposed. The method is based on the use of deterministic one-to-one compound mappings of image pixel values for overlaying a variety of visible watermarks of arbitrary sizes on video images. The compound mappings are proved to be reversible, which allows for lossless recovery of original images from watermarked images. The mappings may be adjusted to yield pixel values close to those of desired visible watermarks. Various types of visible watermarks, including opaque Monochrome and translucent full color ones, are embedded as applications of the proposed generic approach. A two-fold monotonically increasing compound mapping is created and proved to yield more distinctive visible watermarks in the watermarked image. Security protection measures by parameter and mapping randomizations have also been proposed to deter attackers from illicit image recoveries. Videos are playing key role in broadcasting now. Experimental results demonstrating the effectiveness of the proposed approach are also included in videos.

Keywords *- Lossless reversible visible watermarking, mapping randomization, one-to-one compound mapping, parameter randomization, translucent watermark.*

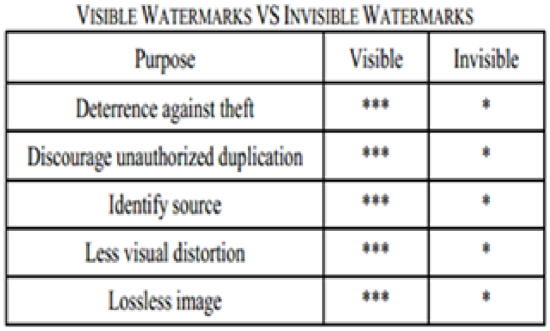
I. INTRODUCTION

Copy right Protection of intellectual properties has become an important topic with the advance of computer technologies and the proliferation of the Internet which have made reproduction and distribution of digital information easier than ever before. Digital Watermarking is one of the ways for copyright protection. Embedding of certain specific information about the copyright holder (company logos, owner-ship descriptions, etc.) into the media to be protected is called Digital Watermarking.



Digital watermarking methods for images are usually categorized into two types: *invisible* and *visible*. The invisible type of digital watermarking aims to embed copyright information imperceptibly into host media such that in cases of copyright infringements, the hidden information can be retrieved to identify the ownership of the protected host.

A watermarked image must be resistant to common image operations to ensure that the hidden information is still retrievable after such alterations. Methods of the visible type, on the other hand, yield visible watermarks which are generally clearly visible after common image operations are applied. In addition, visible watermarks convey ownership information directly on the media and can deter attempts of copyright violations. For the security reasons we are using visible watermarking in broadcasting.



VISIBLE WATERMARKS VS INVISIBLE WATERMARKS

Purpose	Visible	Invisible
Deterrence against theft	***	*
Discourage unauthorized duplication	***	*
Identify source	***	*
Less visual distortion	***	*
Lossless image	***	*

Number of "*" means the degree of importance

Watermarks' embedding, either visible or invisible, results in the degradation of the quality of the host media in general. A group of techniques, named *reversible* watermarking, allow legitimate users to remove the embedded watermark and restore the original content as needed. However, not all reversible watermarking techniques guarantee *lossless image recovery*, which means that the recovered image is identical to the original, pixel by pixel. Lossless recovery is important in many applications where serious concerns about image quality arise. Some examples include forensics, medical image analysis, historical art imaging, or military applications.

There are relatively few mentions of lossless visible watermarking in the literature compared with their invisible counterparts. In the past several lossless invisible watermarking techniques have been proposed. The most common approach is to compress a portion of the original host and then embed the com-pressed data together with the intended payload into the host. The other approach is to superimpose the spread-spectrum signal of the payload on the host so that the signal is detectable and removable. The third approach to embed a bit of information by manipulating a group of pixels as a unit. Although one may use lossless invisible techniques to embed re-movable visible watermarks, the low embedding capacities of these techniques obstruct the possibility of implanting large-sized visible watermarks into host media.

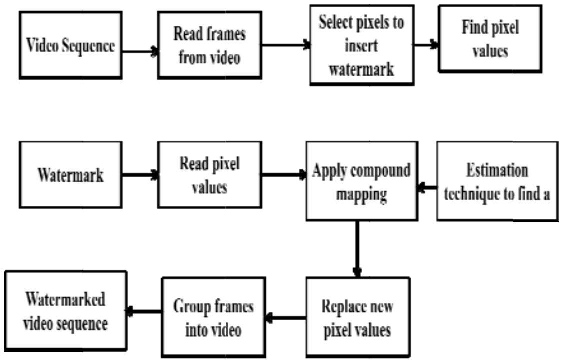
The most common approach for lossless visible watermarking is to embed a monochrome watermark using reversible and deterministic mappings of coefficients or pixel values in the watermark region. The other method for this is to embed a visible watermark by rotating consecutive watermark pixels. One advantage of these approaches is that watermarks of arbitrary sizes can be embedded into any host image. However, using these approaches, only *binary* visible watermarks can be embedded which is too restrictive since most company logos are colorful.

In this paper, a novel method for lossless visible watermarking is proposed by using appropriate *compound mappings* that allow mapped values to be controllable. For lossless recovery of the original image the mappings are proved to be *reversible*. There is a possibility of embedding different types of visible watermarks into cover images, an approach called *generic*. Two applications of the proposed method are demonstrated, where opaque monochrome watermarks and non-uniformly translucent full-color ones are respectively embedded into color images. The compound mappings which are more specific are also created and proved to be able to yield visually more distinctive visible watermarks in the water marked image.

In the remainder of this paper, the proposed method for deriving one-to-one compound mappings is described in section II. Related lemmas and theorems are also provided and security protection measures described. Applications of the proposed method for embedding opaque monochrome and translucent color watermarks into color images are described. The specific compound mapping for yielding more distinctive visible watermarks is described. Experimental results are presented to demonstrate the effectiveness of the proposed method finally; a conclusion with some suggestions for future work is also included.

II. PROPOSED NEW APPROACH TO LOSSLESS VISIBLE WATERMARKING

This section deals with the proposed approach to lossless reversible visible watermarking, based on which appropriate one-to-one compound mappings can be designed for embedding different types of visible watermarks into images. From the resulting watermarked image, the original image can be recovered without any loss by using the corresponding reverse mappings.



Watermark embedding process

Reversible One-to-One Compound Mapping

First, we propose a generic *one-to-one compound mapping f* for converting a set of numerical values, P={p1,p2,…….,pm} to another set Q={q1,q2……qm} such that the respective mapping from p_i to q_i for all i=1,2….., M is *reversible*.

Here, for the copyright protection applications investigated in this study, all the values p_i and q_i are image pixel values (grayscale or color values). The compound mapping f is governed by a one-to-one function F_X with one parameter x=a or b in the following way:

$$q = f(p) = F_b^{-1}(F_a(p)) \qquad (1)$$

Where F_x^{-1} is the inverse of which, by the one-to-one property, leads to the fact that if $F_a(p) = p'$, then $F_a^{-1}(p') = p$ for all values of a and p. On the other hand, $F_a(p)$ and $F_b(p)$ generally are set to be *unequal* if $a \neq b$. The compound mapping described by (1) is indeed *reversible*, that is p, can be derived exactly from q using the following formula:

$$p = f^{-1}(q) = F_a^{-1}(F_b(q)) \qquad (2)$$

as proved below.

Lemma 1 *(Reversibility of Compound Mapping):*

If $q = F_b^{-1}(F_a(p))$ for any one-to-one function F_x with a parameter x, then $p = F_a^{-1}(F_b(q))$ for any values of a, b, p and q.

Proof*:*

Substituting (1) into $F_a^{-1}(F_b(q))$, we get $F_a^{-1}(F_b(q)) = F_a^{-1}\left(F_b\left(F_b^{-1}(F_a(p))\right)\right)$.

By regarding $F_a(p)$ as a value c, the right-hand side becomes $F_a^{-1}(F_b(F_b^{-1}(c)))$, which, after F_b and F_b^{-1} are cancelled out, becomes $F_a^{-1}(c)$. But $F_a^{-1}(c) = F_a^{-1}(F_a(p))$, which just p is after F_a and F_a^{-1} are cancelled out. That is, we have proved $p = F_a^{-1}(F_b(q))$.

As an example, if $F_x(p) = xp + d$, then $F_x^{-1}(p') = (p' - d)/x$.

Thus

$$q = F_b^{-1}(F_a(p)) = F_b^{-1}(ap + d)$$
$$= (ap + d - d)/b = ap/b$$

and so, we have

$$F_a^{-1}(F_b(q)) = F_a^{-1}(b(ap/b) + d) = F_a^{-1}(ap + d)$$
$$= [((ap + d) - d)/a] = (ap/a) = p$$

as expected by Lemma 1.

Lossless Visible Watermarking Scheme based on Lemma 1, we will now derive the proposed generic lossless visible watermarking scheme in the form of a class of one-to-one compound mappings, which can be used to embed a variety of *visible watermarks* into images. The watermark can be removed to recover the original image losslessly. This makes the embedding reversible. For this aim, a preliminary lemma is first described as follows.

Lemma 2 *(Preference of Compound-Mapped Value):*

It is possible to use the compound mapping $q = F_b^{-1}(F_a(p))$ to convert a numerical value to another value close to a *preferred* value l.

Proof:

Let $F_x(p) = p - x$ where x is the parameter for F. Then $F_x^{-1}(p') = p' + x$. Also, let a=p-ε and b=l, where ε is a small value. Then, the compound mapping $F_b^{-1}(F_a(p))$ of p yields q as

$$q = F_b^{-1}(F_a(p)) = F_b^{-1}(p - a) = F_b^{-1}(\varepsilon)$$
$$= \varepsilon + b = \varepsilon + l$$

which means that the value q is close to the preferred value l. The above lemma relies on two assumptions. The first is that is a is close to p, or equivalently, that a=p- ε. The reason why we derive the above lemma for instead of for a=p is that in the reverse mapping we want to recover p from q *without knowing p*, which is a requirement in the applications of reversible visible watermarking investigated in this study.

Although the value of p cannot be known in advance for such applications, it can usually be estimated, and we will describe some techniques for such estimations in the subsequent sections.

The second assumption is that $F_x(p)$ yields a small value if a and p are close. Though the basic difference function $F_x(p) = p - x$ used in the above proof satisfies this requirement for most cases, there is a possible problem where the mapped value may exceed the range of valid pixel values for some values of a, b and p. For example, when a=255, b=255 and p=253, we have q=255-253+255=257>255. It is possible to use the standard modulo technique (i.e., taking q=257 mod 256=1) to solve this issue; however, such a technique will make q far from the desired target value of b, which is 255. Nevertheless, we will show in that using such a standard modulo function, $F_x(p) = (p \quad x)_{\mathrm{mod}\ 256}$, can still yield reasonable experimental results. Furthermore, we show in a more sophisticated one-to-one function that is free from such a wraparound problem.

Security Considerations

The Advance of computer technologies and the proliferation of the internet have made reproduction and distribution of digital information easier than ever before. Copyright protection of intellectual properties has, therefore become an important concern. As mentioned previously, although we want legitimate users to be able to recover the original image from a watermarked one, we do not want an attacker to be able to do the same. Herein, we propose some security protection measures against illicit recoveries of original images.

First, we make the parameters and in the above algorithms to be dependent on certain secret keys that are known only by the creator of the watermarked image and the intended receivers. One simple technique to achieve this is to use a secret key to generate a pseudo-random sequence of numerical values and add them to either or both of a and b for the pixels in the watermarking area. This technique is hereinafter referred to as *parameter randomization*.

Another way of security protection is to make the choices of the *positions* for the pixels to be dependent on a secret key. Specifically, we propose to process *two* randomly chosen pixels (based on the security key) in P simultaneously as follows. Let the two pixels be denoted as X1 and X2 with values p1 and p2, respectively.

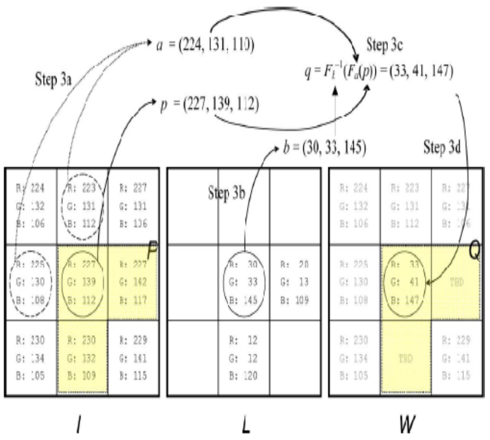


Fig. 1 : Illustration of mapping the center pixel of a 3 x 3 image. Only the mapping of the center pixel is shown for clarity; the east and south pixels are depicted as TBD (to be determined) in W.

The color estimates a1and a2 corresponding to X1 and X2, respectively, are individually derived as before using their respective neighbors. The parameters b1 and b2 are set to be the values l1 and l2 of the respective watermark pixels y1 and y2.

Then, instead of setting the values of the watermarked pixels Z1 and Z2 to be $q_1 = F_{b_1}^{-1}(F_{a_1}(p_1))$ and $q_2 = F_{b_2}^{-1}(F_{a_2}(p_2))$ as before, we *swap* the parameters and set

$$q_1 = F_{b_1}^{-1}(F_{a_2}(p_2)) \quad \text{and} \quad q_2 = F_{b_2}^{-1}(F_{a_1}(p_1)).$$

This parameter exchange does not affect the effectiveness of lossless recoverability, because we can now recover the original pixel values by the following compound mappings:

$$p_1 = F_{a_1}^{-1}(F_{b_2}(q_2)) \quad \text{and} \quad p_2 = F_{a_2}^{-1}(F_{b_1}(q_1)).$$

We will refer to this technique in the sequel as *mapping randomization.* We may also combine this technique with the above mentioned parameter randomization technique to enhance the security further. Last, the position in the image where a watermark is embedded affects the resilience of the watermarked image against illicit image recovery attempts. In more detail, if the watermark is embedded in a smooth region of the image, an attacker can simply fill the region with the background color to remove the watermark irrespective of the watermarking technique used. To counter this problem, an appropriate position should be chosen, using, for example, the adaptive positioning technique when embedding a watermark. However, for ease of discussions and comparisons, we always embed a watermark in the lower right-hand corner of an image in this study.

III. LOSSLESS VISIBLE WATERMARKING OF OPAQUE MONOCHROME WATERMARK

We describe now how we embed a losslessly-removable opaque monochrome watermark L into a color image I such that the watermark is *visually distinctive* in the watermarked image W, as an application of the proposed generic approach to lossless visible watermarking.

First, we denote the sets of those pixels in I corresponding spatially to the *black* and *white* pixels in L by P and P', respectively. An illustration of such areas of p and P' is shown in Fig. 2. We define Q and Q' in a similar way for the watermarked image W, which correspond to P and P', respectively. Then, we adopt the simple one-to-one function $F_a(p) = p - a$, and use the same pair of parameters a and b for *all* mappings of pixels in P. Also, we apply the "modulo-256" operation to the results of all computations so that they are within the valid range of color values. Our experiments show that this method still yields reasonable results.

For the values of parameters a and b, we set a to be the *average* of the color component values of the pixels in P'. This average value presumably is close to the value of pixel in P, satisfying the condition $a = p - \varepsilon$ mentioned previously. To ensure that the watermark is distinctive in W, we do not simply embed black values for pixels in watermarking area P (that is, we do not embed for l=0 for P), but set l to be a value which is distinctive with respect to the pixel colors in the surrounding region P'. To achieve this, we set $b = l = a + 128$, which is a value *distinctive* with respect to a. As a result, the value of a pixel in Q, according to Lemma2, becomes $q = F_b^{-1}(F_a(p)) = b + \varepsilon = a + 128 + \varepsilon$, meaning that the pixel values of Q are also distinctive with respect to those of the surrounding pixels in Q' as desired.

Since both a and b are derived from P' during watermark embedding, the exact same values of a and b can be derived during watermark removal because Q' is identical to P'. The original image can, therefore, be recovered without any loss using Algorithm 2.

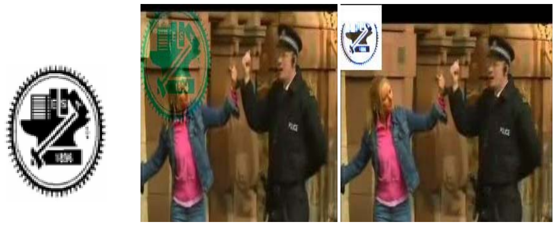


Fig. 2 : Illustration of images in watermark. (a) Logo (b) Opaque (c) Translucent

We embedded the watermark of Fig. 2(a) into the video frames, respectively through which we have demonstrated the effectiveness of the proposed method, in one of our experiments. We applied both the mapping randomization and the parameter randomization techniques described in Section 2, for security protection. Specifically, for the latter technique we added random integer values in the range of -12 to +12 to the parameter. The images recovered by using correct keys for the parameter b. That the embedded opaque watermarks are distinctive with respect to their surroundings and can be removed completely when the input key is correct. On the contrary, when the key was incorrect, the inserted watermark cannot be removed cleanly, with noise remaining in the watermarking area. The videos are converted to frames, frames are treated as image and same methods are applied.

IV. LOSSLESS VISIBLE WATERMARKING OF TRANSLUCENT COLOR WATERMARKS

As another application of the proposed approach, we describe now how we embed more complicated *translucent color watermarks*. A translucent color watermark used in this study is an arbitrary RGB image with each pixel being associated with an *alpha component value* defining its *opacity*. The extreme alpha values of 0 and 255 mean that the watermark pixel is *completely transparent* and *totally opaque*, respectively. A translucent full-color watermark is visually more attractive and distinctive in a watermarked image than a traditional transparent monochrome watermark, as mentioned previously. Such a kind of watermark can better represent trademarks, emblems, logos, etc., and thus is more suitable for the purpose of advertising or copyright declaration.

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

In this paper, a new method is proposed for reversible visible watermarking with lossless image recovery capability. The method uses one-to-one compound mappings that can map image pixel values to those of the desired visible watermarks. Relevant lemmas and theorems are described and proved to demonstrate the reversibility of the compound mappings for lossless reversible visible watermarking. The compound mappings allow different types of visible watermarks to be embedded, and two applications have been described for embedding opaque monochrome watermarks as well as translucent full-color ones. A translucent watermark is clearly visible and visually appealing, thus more appropriate than traditional transparent binary watermarks in terms of advertising effect and copyright declaration. The two-fold monotonically increasing property of compound mappings was defined and an implementation proposed that can provably allow mapped values to always be close to the desired watermark if color estimates are accurate. We have also described parameter randomization and mapping randomization techniques, which can prevent illicit recoveries of original images without correct input keys. Experimental results have demonstrated the feasibility of the proposed method and the effectiveness of the proposed security protection measures. Future research may be guided to more applications of the proposed method and extensions of the method to other data types other than bitmap images, like in JPEG images and MPEG videos.

REFERENCES

[1] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Information hiding—A survey,” Proc. IEEE, vol. 87, no. 7, pp. 1062–1078, Jul. 1999.

[2] N. F. Johnson, Z. Duric, and S. Jajodia, Information Hiding. Steganography and Watermarking—Attacks and Countermeasures,sBoston, MA: Kluwer, 2001.

[3] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” IEEE Trans. Image Process. , vol. 6, no. 12, pp. 1673–1687, Jun. 1997.

[4] M. S. Kankanhalli, Rajmohan,and K.R. Ramakrishnan, “Adaptive visible watermarking of images,” in Proc. IEEE Int. Conf. Multimedia Computing and Systems , 1999, vol. 1, pp. 568–573.

[5] Y.HuandS.Kwong,“Wavelet domain adaptive visible watermarking,” Electron. Lett. , vol. 37, no. 20, pp. 1219–1220, Sep. 2001.

[6] S. P. Mohanty, K. R. Ramakrishnan, and M. S. Kankanhalli, “A DCT domain visible watermarking technique for images,” in Proc. IEEE Int. Conf. Multimedia and Expo , Jul. 2000, vol. 2, pp. 1029–1032.

[7] G. Braudaway, K. A. Magerlein, and F. Mintzer, Protecting publicly available images with a visible image watermark,” in Proc. SPIE Int. Conf. Electronic Imaging , Feb. 1996, vol. 2659, pp. 126–133.

[8] Y. J. Cheng and W. H. Tsai, “A new method for copyright and integrity protection for bitmap images by removable visible watermarks and irremovable invisible watermarks,” presented at the Int. Computer Symp.—Workshop on Cryptology and Information Security, Hualien, Taiwan, R.O.C., Dec. 2002

[9] P. M. Huang and W. H. Tsai, “Copyright protection and authentication of grayscale images by removable visible watermarking and invisible signal embedding techniques: A new approach,” presented at the Conf. Computer Vision, Graphics and Image Processing, Kinmen, Taiwan, R.O.C., Aug. 2003.

[10] Y. Hu, S. Kwong, and J. Huang, “An algorithm for removable visible watermarking,” IEEE Trans. Circuits Syst. Video Te chnol. , vol. 16, no. 1, pp. 129–133, Jan. 2006.

[11] Y. Hu and B. Jeon, “Reversible visible watermarking and lossless recovery of original images,” IEEE Trans. Circuits Syst. Video Te chnol. , vol. 16, no. 11,pp. 1423–1429, Nov. 2006.

[12] B. Macq, “Lossless multiresolution transform for image authenticating watermarking,” presented at the European Signal Processing Conf., Tampere, Finland, Sep. 2000.

◈◈◈

Enhanced Power Transfer by Simultaneous AC–DC Transmission

Padmaja Tadiboina[1], Sambana Srikanth[2] & Ravikumar Tenali[3]

[1&2]BVC College of Engineering, Odalarevu, A.P, India
[3]Sri Venkateswara Institute of Science & Information Technology, A.P, India

Abstract – Several efforts have been done that allow full utilization of existing transmission facilities without decreasing system availability and security. One way of achieving this is simultaneous ac-dc transmission in which it is possible to load existing EHV lines very close to their thermal limits. The conductors are allowed to carry usual ac along with dc superimposed on it. The added dc power flow does not cause any transient instability. The advantages of this approach are no alterations of conductors, insulator strings, and towers of the original line are needed.

This paper presents the feasibility of converting a double circuit ac line into composite ac–dc power transmission line to get the advantages of parallel ac–dc transmission to improve stability and damping out oscillations. Simulation and experimental studies are carried out for the coordinated control as well as independent control of ac and dc power transmissions. Master current controller senses ac current and regulates the dc current orders for converters online such that conductor current never exceeds its thermal limit

Keywords *- Extra high voltage (EHV) transmission, flexible ac transmission system (FACTS), simultaneous ac–dc power transmission.*

I. INTRODUCTION

The present situation demands the review of traditional power transmission theory and practice, on the basis of new concepts that allow full utilization of existing transmission facilities without decreasing system availability and security.

To improve stability and to achieve power transmission close to its thermal limit [1]–[4]. There are 2 methods

1. The flexible ac transmission system (FACTS) concepts, based on applying state-of-the-art power electronic technology to existing ac transmission system,
2. Simultaneous ac–dc power transmission in which the conductors are allowed to carry superimposed dc current along with ac current.

In this paper, the feasibility study of conversion of a double circuit ac line to composite ac–dc line without altering the original line conductors, tower structures, and insulator strings has been presented. In this scheme, the dc power flow is point-to-point bipolar transmission system. Clerici *et al.* [7] suggested the conversion of ac line to dc line for substantial power up-grading of existing ac line. However, this would require major changes in the tower structure as well as replacement of ac insulator strings with high creepage dc insulators. The novelty of our proposed scheme is that the power transfer enhancement is achieved without any alteration in the existing EHV ac line. The main object is to gain the advantage of parallel ac–dc transmission and to load the line close to its thermal limit.

II. SIMULTANEOUS AC–DC POWER TRANSMISSION

Fig. 1 depicts the basic scheme for simultaneous ac–dc power flow through a double circuit ac transmission line. The dc power is obtained through line commutated 12-pulse rectifier bridge used in conventional HVDC and injected to the neutral point of the zigzag connected secondary of sending end transformer and is reconverted to ac again by the conventional line commutated 12-pulse bridge inverter at the receiving end. The inverter bridge is again connected to the neutral of zigzag connected winding of the receiving end transformer. The double circuit ac trans-mission line carriers both three-phase ac and dc power. Each conductor of each line carries one third of the total dc current along with ac current . Resistance being equal in all the three phases of secondary winding of zigzag transformer as well as the three conductors of the line, the dc current is equally divided among all the three phases. The three conductors of the second line provide return path for the dc current. Zigzag connected winding is used at both ends to avoid saturation of transformer due to dc current. Two fluxes produced by the dc current flowing through each of a winding in each limb of the core of a zigzag transformer are equal in magnitude and opposite in direction. So the net dc flux at any instant of time becomes zero in each limb of the core. Thus, the dc saturation of the core is avoided. A high value of reactor X_d is used to reduce harmonics in in dc current.

Assuming the usual constant current control of rectifier and constant extinction angle control of inverter [4], [8]–[10], the equivalent circuit of the scheme under normal steady-state op-erating condition is given in Fig. 2. The dotted lines in the figure show the path of ac return current only. The second transmission line carries the return dc current I_d , and each conductor of the line carries $I_{d/}3$ along with the current per phase.

.

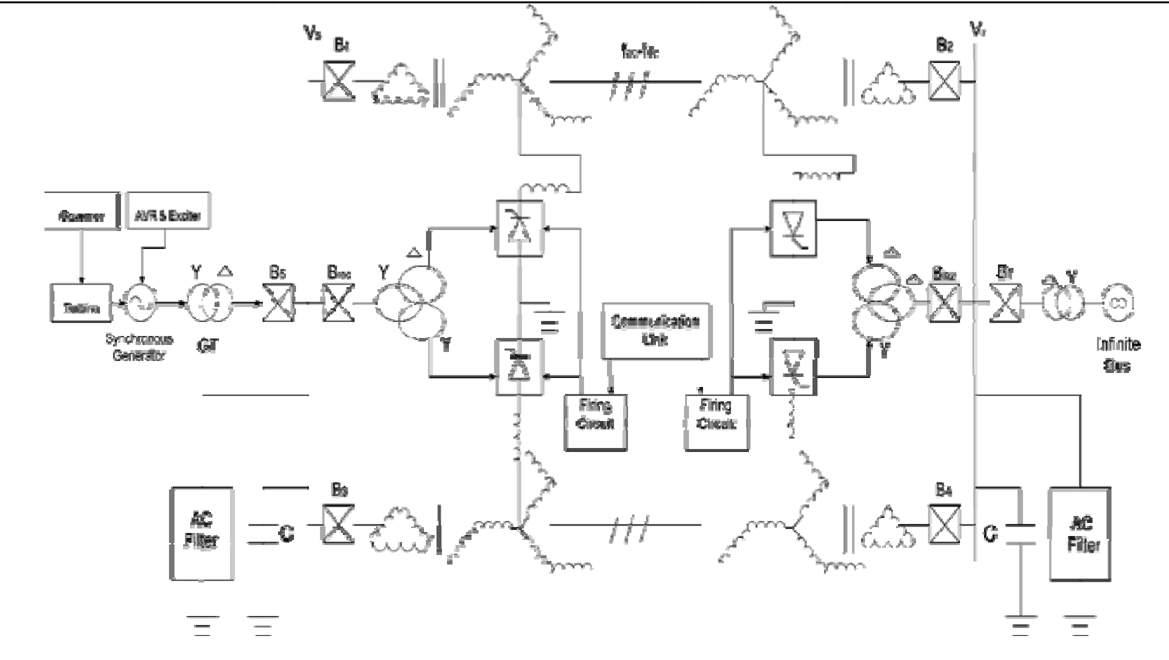


Fig. 1: Basic scheme for composite ac–dc transmission

.

V_{dro} and V_{dio} are the maximum values of rectifier and inverter side dc voltag es and are $(3\sqrt{2}/\pi)$ times converter ac input line-to-line voltage R,L,and C are the line parameters per phase of each line.

R_{cr},R_{ci} are commutating resistances and α, ϒ are firing and extinction angle control respectively

Neglecting the resistive drops in the line conductors and transformer windings due to dc current, expressions for ac voltage and current, and for active and reactive powers in terms of A, B, C, and D parameters of each line may be written as

$$E_s = AE_R + BI_R \tag{1}$$

$$I_s = CE_R + DI_R \tag{2}$$

$$P_s + jQ_s = -E_sE_R^*/B^* + D^*E_S^2/B^* \tag{3}$$

$$P_R + jQ_R = E_s^*E_R/B^* - A^*E_R^2/B^*. \tag{4}$$

Neglecting ac resistive drop in the line and transformer, the dc power P_{dr} and $P_{di}$of each rectifier and inverter may be ex-pressed as

$$P_{dr} = V_{dr}I_d \tag{5}$$

$$P_{di} = V_{di}I_d. \tag{6}$$

Reactive powers required by the converters are

$$Q_{dr} = P_{dr}\tan\theta_r \tag{7}$$

$$Q_{di} = P_{di}\tan\theta_i \tag{8}$$

$$\cos\theta_r = [\cos\alpha + \cos(\alpha + \mu_r)]/2 \tag{9}$$

$$\cos\theta_i = [\cos\gamma + \cos(\gamma + \mu_i)]/2. \tag{10}$$

$$P_{st} = P_s + P_{dr} \text{ and } P_{rt} = P_R + P_{di} \tag{11}$$

$$Q_{st} = Q_s + Q_{dr} \text{ and } Q_{rt} = Q_R + Q_{di}. \tag{12}$$

Transmission loss for each line is

I_a being the rms ac current per conductor at any point of the line, the total rms current per conductor becomes

$$I = [I_a^2 + (I_d/3)^2]^{1/2}$$

$$\text{Power loss for each line} = P_L \approx 3I^2R.$$
$$P_L = (P_S + P_{dr}) - (P_R + P_{di}).$$

The net current I_a in any conductor is offseted from zero. In case of a fault in the transmission system, gate signals to all the SCRs are blocked and that to the bypass SCRs are re-leased to protect rectifier and inverter bridges.

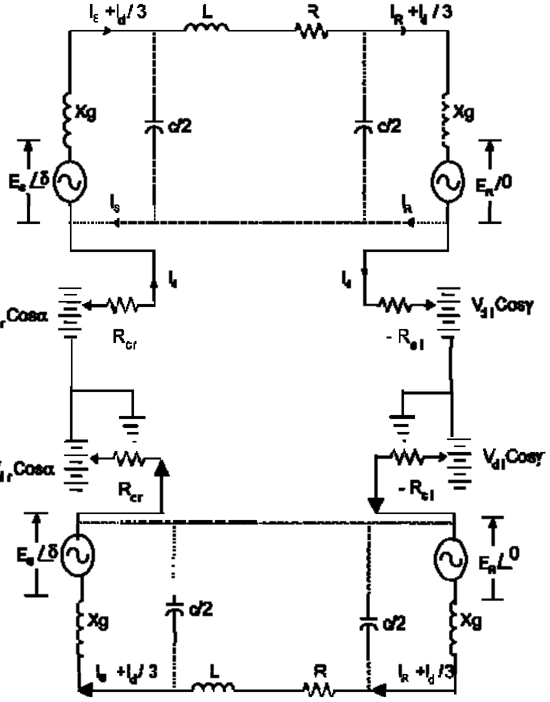


Fig. 2 : Equivalent circuit.

Now, allowing the net current through conductor equal to its thermal limit

$$(I_{th})$$

$$I_{th}= [I_a^2+(I_d/3)^{2]1/2} \qquad (14)$$

Let V_{ph} per-phase rms voltage of original ac line. Let also the $\mathbf{V_a}$ per-phase voltage of ac component of composite ac–dc line with dc voltage superimposed on it. As insula-tors remain unchanged, the peak voltage in both cases should be equal

$$V_{max} = \sqrt{2}V_{ph} = V_d + \sqrt{2}V_a. \qquad (15)$$

Each conductor is to be insulated for $\mathbf{V_{max}}$, but the line-to-line voltage has no dc component $V_{LL\,max} = \sqrt{6}V_a$. Therefore, conductor-to-conductor separation distance of each line is determined only by rated ac voltage of the line.

Allowing maximum permissible voltage offset such that the composite voltage wave just touches zero in each every cycle;

$$V_d = V_{ph}/\sqrt{2} \text{ and } V_a = V_{ph}/2. \qquad (16)$$

The total power transfer through the double circuit line before conversion is as follows:

$$P'\text{total} \approx 3V_{ph}{}^2 \mathrm{Sin}\delta_1/X \qquad (17)$$

$$P'_{total} = 2.M.SIL \qquad (18)$$

Where M is the multiplying factor, and its magnitude decreases with the length of line. The value of M can be obtained from the loadability curve [4].

The total power transfer through the composite line

$$P_{total} = P_{ac} + P_{dc} = 3V_L^2 \mathrm{Sin}\delta_2/X + 2V_dI_d. \qquad (19)$$

The power angle between the ac voltages at the two ends of the composite line may be increased to a high value due to fast controllability of dc component of power. For a constant value of total power, $\mathbf{P_{ac}}$ may be modulated by fast control of the current controller of dc power converters.

Approximate value of ac current per phase per circuit of the double circuit line may be computed as

$$I_a = V(\mathrm{Sin}\delta/2)/X. \qquad (20)$$

The rectifier dc current order is adjusted online as

$$I_d = 3\sqrt{I_{th}^{*2} - I_a^{*2}}. \qquad (21)$$

III. DESCRIPTION OF THE SYSTEM MODEL

The network depicted in Fig. 1 was studied using PSCAD/ EMTDC. A synchronous machine is feeding power to infinite bus via a double circuit, three-phase, 400-KV, 50-Hz, 450-Km ac transmission line. The 2750-MVA (5 550), 24.0-KV synchronous machine is dynamically modeled, a field coil on d-axis and a damper coil on q-axis, by Park's equations with the frame of reference based in rotor [4].

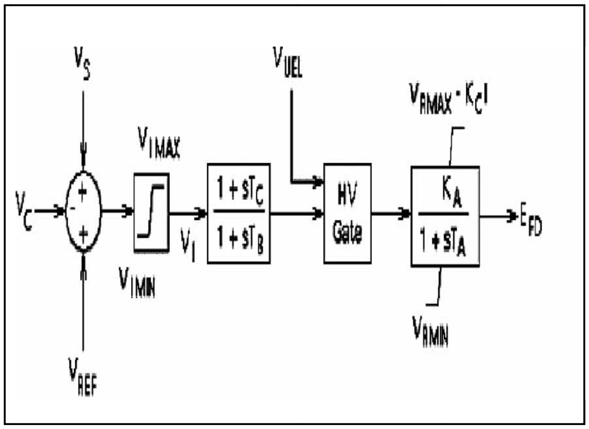


Fig. 3 : IEEE type AC4A excitation system.

IV. CASE STUDIES: COMPUTATIONS AND SIMULATION

A. AC Configuration Only

The loadability of Moose (commercial name), ACSR, twin bundle conductor, 400-kV, 50-Hz, 450-km double circuit line has been computed.

The parameters of the line are

$$\mathbf{z = 0.03252 + j0.33086\ \Omega/km/ph/ckt.}$$
$$\mathbf{y = j3.33797 \times 10^{-6}\ S/km/ph/ckt.}$$

Current carrying capacity of each subconductor= 0.9kA

$$\mathbf{I_{th} = 1.8\ kA/ckt.\ SIL = 511\ MW/ckt.}$$
$$\mathbf{M = 1.1,\ (From\ loadability\ curve[4];\ X = 74.4435\ \Omega/ph.}$$

Using (17)–(20), the computed power at receiving end and conductor current is

$$\mathbf{P'_{total} = 1124.2\ MW;\ \delta_1 \approx 30^\circ}$$
$$\mathbf{I_{ph/ckt} = 0.803\ kA.}$$

The computed and simulated results are found to be in close conformity.

The conductor current 0.805246 kA is much below the thermal limit 1.8 kA.

B. Conversion of the Conventional Double Circuit AC Line Into Composite AC–DC Power Transmission Line

Using (15) and (16)

$$V_a = 120\ kV/\ ph(208kV_{ll});\ V_d{=}160kv$$

The above ac voltage Va has been increased from 115.473 to 120 kV, and Vd has been decreased from 163.328 to 160.0 kV to have zero crossing in voltage wave.

δ₂ is assumed to be maximum 80 , which is commonly used in lines controlled by FACTS devices [13].

The computed *approximate* power transfers for converted line at various transmission angles are shown in Table I

TABLE I

COMPUTED RESULTS

Power Angle (δ) Degrees	30°	45°	60°	75°	80°
ac power(MW) $3V_a^2Sin\delta_1/X$	290	410	502.66	[illegible]	5[illegible]7.55
ac current I_a (kA) I_a V(Sinδ/2)/X	0.4166	0.6122	[illegible]	[illegible]	1.035
dc Current (kA) I_d= $3\sqrt{I_{th}^{*2} - I_a^{*2}}$	5.259	5.090	4.829	4.529	4.418
Dc Power P_{dc}=2Vd x Id (MW)	1684.8	1624.9	1545.5	[illegible]	1413.76
P_{total}=P_{ac}+P_{dc} (MW)	1971	2034	2048	2040	1935

The proposed composite ac-dc power scheme shown Fig.1 has been simulated in steady state mode as a real system using PSCAD software package. The wattmeter and ammeter readings at various points on the system are tabulated in Table-II

TABLE II

SIMULATED RESULTS

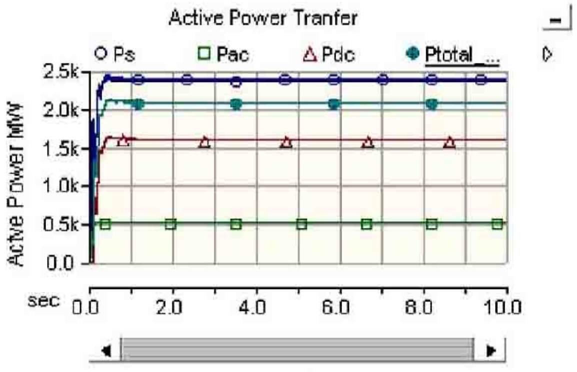
Power Angle (δ)	30°	45°	60°	75°	80°
Ps (MW)	2306	2371.0	2381.3	[illegible]	[illegible]
Pac (MW) Transfer	291.89	411.00	495.3	[illegible]	[illegible]
Pdc (MW) Transfer	1715.5	1657.0	585.8	[illegible]	[illegible]
Pac_line (MW) Pdc_loss(MW)	11.94 280.51	[illegible] [illegible]	[illegible] 341.17	[illegible] [illegible]	91.73 [illegible]
Ploss (total MW)	292.45	[illegible]	[illegible]	[illegible]	[illegible]
Pr (MW) Total Transfer	[illegible]	2051.14	[illegible]	[illegible]	[illegible]
Qs line (MVAR)	-13.78	64.98	.85.56	[illegible]	[illegible]
Qr_line (MVAR)	39.08	146.84 0	280.85	[illegible]	[illegible]
Qrec (MVAR)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Qinv (MVAR)	841.3	823.5	797.43	[illegible]	[illegible]
ac current Ia (kA)	0.41577	0.61123	[illegible]	[illegible]	[illegible]
dc current Id (kA) Cond. dc current Id/3 (kA)	[illegible] 1.74754	[illegible] 1.70153	[illegible] .6373	[illegible] [illegible]	[illegible] [illegible]
conductor current Isim(kA)	1.78587	1.78264	.78281	[illegible]	[illegible]
Increase of power transfer	[illegible]	82.49%	83.451%	[illegible]	77.5%

Fig 4 shows the time response of various system performances at power angle of 60^o

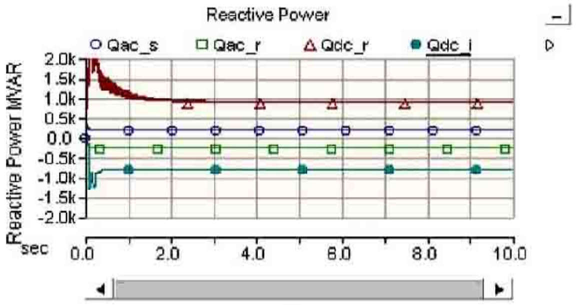
It has been observed from above tracing that system is stable even after superimposing dc on ac.

Experimental results for 3-A ac current with superimposed 2/3-A dc current in each line are depicted in Fig. 5.

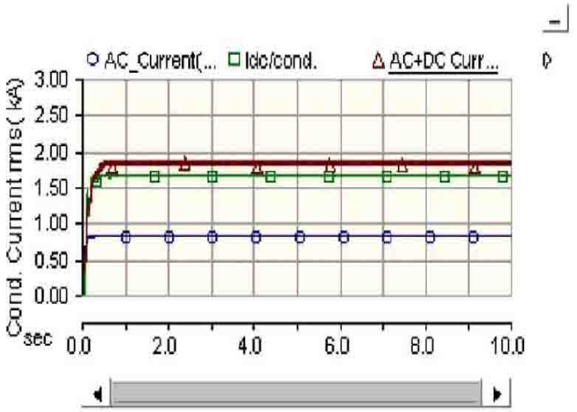
The shape and magnitude of primary current of the zig-zag connected transformer remains unchanged with and without injection of dc current Rectifier input ac current is stepped shaped as no filters has been connected on ac side .



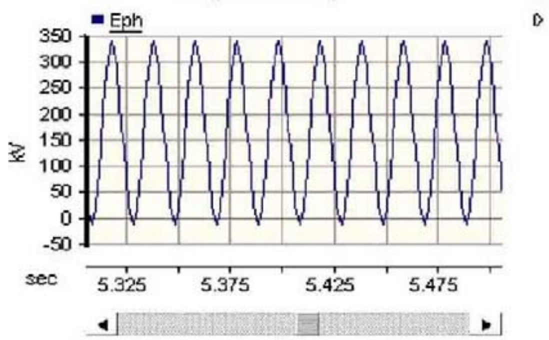
(a)



(b)



(c)



(d)

Fig. 4 (e) Sending end (Ps), ac (Pac), dc(Pdc), and total transfer (Ptotal_tr) power. (f) Line sending (Qac_s), receiving (Qac_s) end, rectifier (Qrec_dc), and inverter (Qinv_dc) reactive powers. (g) Ac, dc, and effective conductor current. (h) Phase voltage across insulator string.

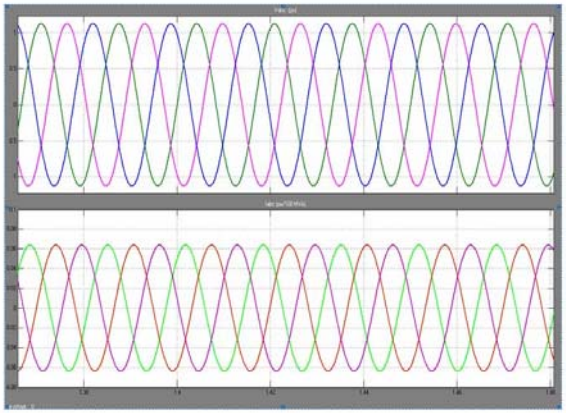


Fig 5 (a) Ac Input to the primary of Rectifier

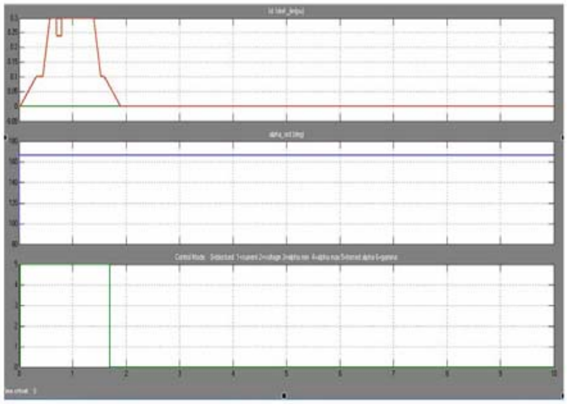


Fig 5 (b) DC Output at the Secondary of Rectifier

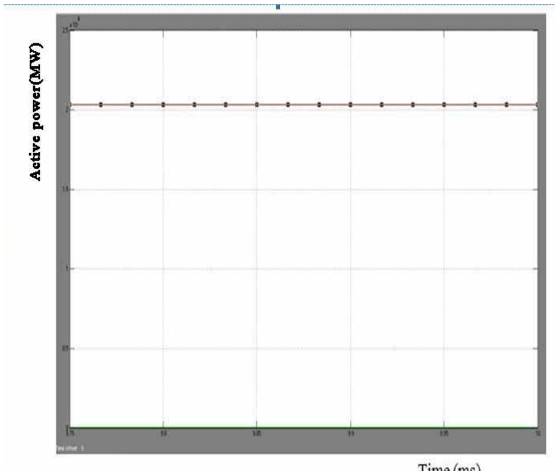


Fig 6 (a) AC Output

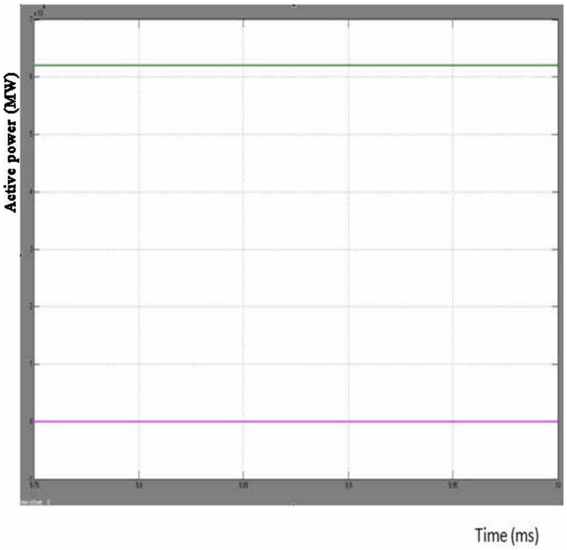


Fig 6 (b) Combined Ac-Dc Output

V. EXPERIMENTAL VERIFICATION

The feasibility of the conversion of the ac line to composite ac–dc line was verified in a laboratory-size model. The basic objective was to verify the operation of the transformers, particularly, the effect on core saturation, with the superimposed ac–dc current flow and the power flow control. The power transmission with and without dc component was found to be satisfactory. There was no saturation of the transformers core with and without dc component

VI. RESULTS

The ac input to the primary of the rectifier is as shown in fig:5(a).The dc output at the secondary of rectifier is shown infig:5(b).Without simultaneous ac-dc transmission the power transmitted is 2016 mega watts in fig:6(a). With simultaneous ac-dc transmission dc power is super imposed on ac conductors and the total power transmitted is 6015mega watts as shown in fig: 6(b). Thus the power enhancement is obtained by simultaneous ac-dc transmission without changing the size of conductors and insulator strings i.e No alterations of conduc-tors, insulator strings, and towers of the original line are needed. Substantial gain in the loadability of the line is obtained. Master current controller senses ac current and regulates the dc current orders for converters online such that conductor current never exceeds its thermal limit.

VII. CONCLUSION

The feasibility to convert ac transmission line to a composite ac–dc line has been demonstrated. For the particular system studied, there is substantial increase (about 83.45%) in the load-abilty of the line. The line is loaded to its thermal limit with the superimposed dc current. The dc power flow does not impose any stability problem. The advantage of parallel ac–dc transmission is obtained. Dc current regulator may modulate ac power flow. There is no need for any modification in the size of conductors, insulator strings, and towers structure of the original line. The optimum values of ac and dc voltage components of the converted composite line are 1/2 and ½ times the ac voltage before conversion, respectively

REFERENCES

[1] L. K. Gyugyi, “Unified power flow concept for flexible A.C. transmis-sion system,” Proc. Inst. Elect. Eng., p. 323, Jul. 1992.

[2] L. K. Gyugyi et al., “The unified power flow controller; a new approach to power transmission control,” IEEE Trans. Power Del., vol. 10, no. 2, pp. 1085–1097, Apr. 1995.

[3] N. G. Hingorani, “FACTS—flexible A.C. transmission system,” in Proc. Inst. Elect. Eng. 5th. Int. Conf. A.C. D.C. Power Transmission, London, U.K., 1991.

[4] P. S. Kundur, Power System Stability and Control. New York: Mc-Graw-Hill, 1994.

[5] K. P. Basu and B. H. Khan, “Simultaneous ac-dc power transmission,” Inst. Eng. (India) J.-EL, vol. 82, pp. 32–35, Jun. 2001.

[6] H. Rahman and B. H. Khan, “Enhanced power

transfer by simulta-neous transmission of AC-DC: a new FACTS concept," in Proc. Inst. Elect. Eng. Conf. Power Electronics, Machines, Drives, Edinburgh, U.K., Mar. 31–Apr. 2 2004, vol. 1, pp. 186–191.

[7] A. Clerici, L. Paris, and P. Danfors, "HVDC conversion of HVAC line to provide substantial power upgrading," IEEE Trans. Power Del., vol. 6, no. 1, pp. 324–333, Jan. 1991.

[8] Padiyar, HVDC Power Transmission System. New Delhi, India: Wiley Eastern, 1993.

[9] E. W. Kimbark, Direct Current Transmission. New York: Wiley, 1971, vol. I.

[10] J. Arillaga and N. R. Watson, Computer Modelling of Electrical Power Systems. Chichester, U.K.: Wiley, 2003.

[11] M. A. Chaudhry and D. P. Caroll, "Coordinated active and reactive power modulation of multiterminal HVDC system," IEEE Trans. Power App. Syst., vol. PAS-103, pp. 1480–1485, 1989.

[12] K. R. Padiyar, M. A. Pai, and C. Radhakrishna, "Analysis of D.C. link control for system stabilization," in Proc. Inst. Elect. Eng. Conf. Publ. No. 205, London, U.K., 1981, pp. 145–148.

[13] M. Stella, P. K. Dash, and K. P. Basu, "A neuro-sliding mode controller for STATCOM," Elect. Power Compon. Syst., vol. 32, pp. 131–147, Feb. 2004.

[14] M. Szechtman, T. Wees, and C. V. Thio, "First benchmark model for HVDC control studies," Electra, no. 135, pp. 54–67, Apr. 1991.

[15] PSCAD/EMTDC, User's Guide, Manitoba-HVDC Research Centre. Winnipeg, MB, Canada, Jan. 2003.

◈◈◈

Functional Verification of Secure Digital Host Controller

Jaikiran L[1], Prakash Biswagar[2] & Theiventhiran Murugan[3]

[1&2]ECE Dept., R V College of Engineering, R V Vidyanikethan Post Mysore Road, Bangalore-560059, India
[3]Adventura Technologies Pvt. Ltd, koramangala, Bangalore, India

Abstract – Portable storage devices are becoming popular and growing rapidly. These devices can store and acquire information wherever whenever you need. The important applications of portable storage devices are to make backup copies of important data, to share information between different computers or persons, to store digital pictures, music, games, power point presentations etc., and to secure information. These devices are very cost effective, easy to use, and extremely practical.

***Keywords** - SD Host Controller, DMA, ADMA, Verilog HDL.*

I. INTRODUCTION

With the increasing consumer digital content, demand for high capacity digital storage is increasing rapidly. Today, portable storage media's are widely used in all mobile phones, digital cameras, camcorders, and in many multimedia devices. Different memory formats like Flash, Secure Digital (SD), Compact Flash, Universal Serial Bus (USB), and Multimedia Card (MMC) are available in the market to store the digital contents. Of all these formats, SD provides many advantages over other formats. Secure Digital (SD) is a Non-volatile memory card format developed by the SD CARD Association (SDA) for use in portable devices. SD cards provide high storage capacity, higher transfer speed, and interoperability with Personal Computer (PC) - related devices and multimedia products. It is the leading standard for mobile phones, digital cameras, audio players, personal computers, printers, car navigation systems, electronic books, and many other consumer electronic devices. The previous versions of SD Host Controller were having a bus speed of upto 25Mbps. The Host Controller version 3.0 operates in UHS-I and provides bus speed up to 104Mbps.

SD standards are available in three capacity formats: SD, SDHC and SDXC. The Standard-Capacity (SDSC) card family, commonly termed SD, has an official maximum capacity of 2 GB, though some are available up to 4 GB. The High-Capacity (SDHC) card family has a capacity of 4 GB to 32 GB, eXtended-Capacity (SDXC) card family have a capacity starting above 32 GB with a maximum of 2 TB. The standards have the broad interoperability and compatibility needed to ensure support between multiple devices and for future applications.

II. ARCHITECTURE OF SD HOST CONTROLLER

The SD3.0 / SDIO3.0 / eMMC4.5 Host Controller (3MCR Host Controller) is a Host Controller with an ARM processor interface. This product conforms to SD Host Controller Standard Specification Version 3.00. The block diagram of Host Controller is shown below in figure 1.

The Host Controller handles SDIO/SD Protocol at transmission level, packing data, adding cyclic redundancy check (CRC), Start/End bit, and checking for transaction format correctness. The Host Controller provides Programmed IO method and DMA data transfer method. In programmed IO method, the ARM processor transfers data using the Buffer Data Port Register.

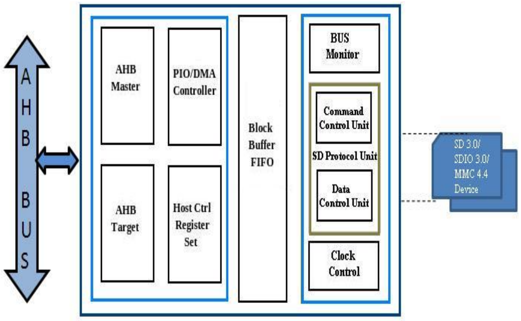


Fig.1: Block Diagram of SD Host Controller

Host controller support for DMA can be determined by checking the DMA support in the Capabilities register. DMA allows a peripheral to read or write memory without the intervention from the CPU. The Host Controller's AHB Host Controller system address register points to the first data address, and data is then accessed sequentially from that address.

A. *Host Controller:*

The Host Controller comprises of Host AHB interface, Host controller registers, Bus Monitor, Clock Generator, and CRC Generator and checker.

The host AHB interface acts as the bridge between AHB and Host Controller. The SD/SDIO controller registers are programmed by the ARM Processor through AHB slave interface. Interrupts are generated to the ARM Processor based on the values set in the Interrupt status register and Interrupt enable registers. Bus monitor will check for any violations occurring in the SD bus and time-out conditions. The Clock generation block will generate the SD clock depending on the value programmed by the ARM Processor in the Clock Control Register. The CRC7 and CRC16 generators calculate the CRC for command and Data respectively to send the CRC to the SD/SDIO and eMMC card. The CRC7 and CRC16 checker check for any CRC error in the Response and Data sent by the SD/SDIO card.

Register Map:

The standard register map is classified in 12 parts listed below. The Host Controller shall support byte, word and double word accesses to these registers. Reserved bits in all registers shall be fixed to zero. The Host Controller shall ignore writes to reserved bits; however, the Host Driver should write them as zero to ensure compatibility with possible future revisions to this Specification.

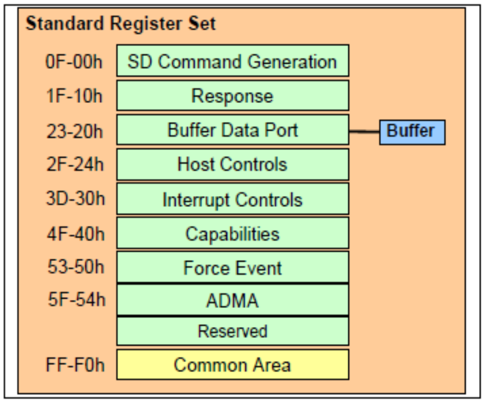


Fig. 2 : SD Standard Register Map

B. *Data FIFO*

The SD/SDIO Host Controller uses one 1k dual port fifo for performing both read and write transactions. During a write transaction (data transferred from ARM Processor to SD3.0 / SDIO3.0 / eMMC4.41 card), the data will be filled in to the first and second half of the FIFO alternatively. When data from first half of FIFO is transferring to the SD3.0 / SDIO3.0 / eMMC4.41 card, the second half of FIFO will be filled and vice versa. The two halves of the FIFO's are alternatively used to store data which will give maximum throughput. During a read transaction (data transferred from SD3.0 / SDIO3.0 / eMMC4.41 card to ARM Processor), the data from SD3.0 / SDIO3.0 / eMMC4.41 card will be written in to the two halves of the FIFO alternatively. When data from one half of the FIFO is transferring to the ARM Processor, the second half of the FIFO will be filled and vice versa and thereby the throughput will be maximum. If the Host controller cannot accept any data from SD3.0 / SDIO3.0 / eMMC4.41 card, then it will issue read wait to stop the data transfer from card or by stopping the clock.

C. *Command and Data Control logic:*

The DAT [0-7] control logic block transmits data on the data lines during write transaction and receives data from the data lines during read transaction. The DAT [0-7] control logic block transmits data in the data lines on posedge and negedge of the SD CLOCK during DDR mode of operation. The DATA [0-7] receiver block receives/ samples the data on the data lines in both posedge and negedge of the SD CLOCK during DDR mode of operation. The Command control logic block sends the command on the cmd line and receives the response coming from the SD3.0 / SDIO3.0 / eMMC4.41 card.

III. VERIFICATION ENVIRONMENT

The Design Under Test (DUT) i.e. SD Host Controller is verified using BFMs (Bus Functional Model) designed in Verilog HDL. The AHB master and AHB target models assist in generating AHB transactions along with arbiter. They emulate the function of an ARM processor. The organization of the verification environment is shown in above figure 3.

A. *AHB Master/Slave BFM & Arbiter:*

The Master BFM will drive the system BUS (AHB or AXI or OCP etc) interface signals. This is the processor read/write interface. The slave BFM is added to respond to the DMA master (Ex: ADMA descriptor read and DMA write/ read transfer). This is the DUT DMA interface which transfers data to and from system memory without the intervention of the processor.

Arbiter block takes care of the bus arbitration for the AHB bus when the core is the master. It takes the requests lines from the master core and gives grant according to the priority.

B. Bus Monitor & Score Board:

The bus monitors are used to check for protocol violations and to display information about the traffic.

AHB Bus Monitor:

AHB Bus monitor file will monitor the AHB Bus and store the data into its internal memory array. It will also monitor the contents driven through the BUS for ADMA operation, DMA and PIO mode of transaction.

SD Monitor:

SD Bus monitor file will monitor the SD Bus and store the data into its internal memory array. It decodes the CMD and DATA lines and fetches the information accordingly.

Score Board:

Score Board is used for data integrity check after data transmission. The expected data and the actual data are fed to the scoreboard for comparison. The test fails if there is data mis-compare or if transmitted data is not equal to the requested transfer.

C. SD Host command definitions (API File):

This file contains all the commands what the host driver will use for the device. The definition of each command shall invoke the system write and read tasks and pass the corresponding arguments.

ADMA descriptor calculation task is used to calculate the ADMA descriptor and program the same in the system slave memory for ADMA2 transaction.

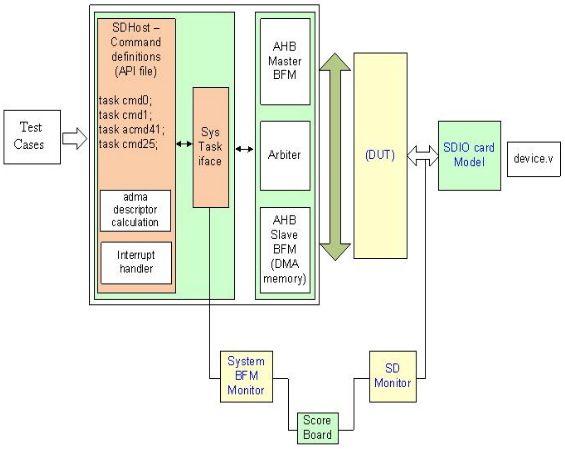


Fig. 3 : Verification Environment for SD Host Controller

Interrupt handler services all the host controller interrupts and will call the data transfer task and error recovery tasks automatically.

D. System Task Interface:

The main purpose of this file is to randomize the system write tasks (i.e., randomly call “sys_write32” or “sys_write16” or “sys_write8” depends on the number of bytes are passed in the “sys_write”). It also supplies the necessary information to the System BFM monitor for data integrity.

E. Test Cases:

The test cases covering different scenarios are written separately and referred with different names.

Perl Scripts are written to pick the test case as our wish, compare the files generated during simulations and running regression.

IV. DATA TRANSFER MODES

The SD Host Controller supports three modes to transfer Data between System and the card. The three modes include Non-DMA, DMA and ADMA mode.

In Non-DMA mode (also known as "programmed I/O" method) the Host Controller has a data buffer for data transfer. The Host Driver accesses internal buffer through the 32-bit Buffer Data Port register. Internally, the Host Controller maintains a pointer to control the data buffer. The pointer is not directly accessible by the Host Driver. Every time the Buffer Data Port register is accessed, the pointer is incremented depending on amount of data written to the buffer.

In DMA mode, the data transfer between system memory and SD card takes place without interruption of CPU execution. DMA shall support both single block and multiple-block transfers but does not support infinite transfers. The result of a DMA transfer shall be the same regardless of the system bus data transfer method. DMA had disadvantage that DMA Interrupt generated at every page boundary disturbs CPU to reprogram the new system address. This DMA algorithm forms a performance bottleneck by interruption at every page boundary.

A new DMA transfer algorithm called ADMA (Advanced DMA) is defined that adopts scatter gather DMA algorithm so that higher data transfer speed is available. The Host Driver can program a list of data transfers between system memory and SD card to the Descriptor Table before executing ADMA. It enables ADMA to operate without interrupting the Host Driver.

Furthermore, ADMA can support not only 32-bit system memory addressing but also 64-bit system memory addressing. The 32-bit system memory addressing uses lower 32-bit field of 64-bit address registers.

Block Diagram of ADMA2

Figure 4 shows block diagram of ADMA2. The Descriptor Table is created in system memory by the Host Driver. 32-bit Address Descriptor Table is used for the system with 32-bit addressing and 64-bit Address Descriptor Table is used for the system with 64-bit addressing. Each descriptor line (one executable unit) consists with address, length and attribute field. The attribute specifies operation of the descriptor line. ADMA2 includes SDMA, State Machine and Registers circuits. ADMA2 does not use 32-bit SDMA System Address Register (offset 0) but uses the 64-bit Advanced DMA System Address register (offset 058h) for descriptor pointer. Writing Command register triggers off ADMA2 transfer. ADMA2 fetches one descriptor line and executes it. This procedure is repeated until end of descriptor is found.

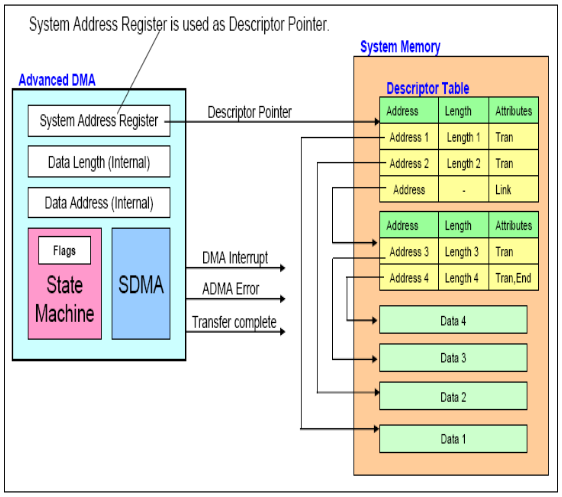
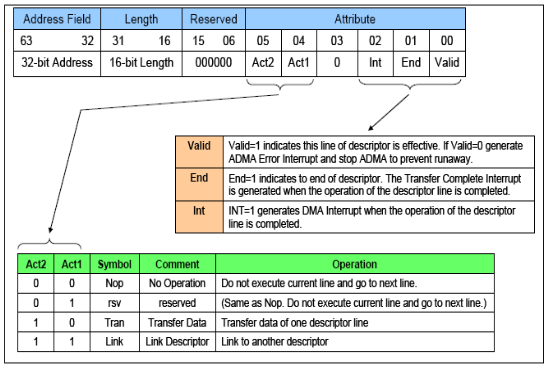


Fig. 4 : Block Diagram of ADMA2

Descriptor Table

Figure 5 shows the definition of 32-bit Address Descriptor Table. One descriptor line consumes 64-bit (8-byte) memory space. Attribute is used to control descriptor. 3 action symbols are specified. "Nop" operation skips current descriptor line and fetches next one. "Tran" operation transfers data designated by address and length field. "Link" operation is used to connect separated two descriptors. The address field of link points to next Descriptor Table.



Address Field	Length	Reserved	Attribute					
63 32	31 16	15 06	05	04	03	02	01	00
32-bit Address	16-bit Length	000000	Act2	Act1	0	Int	End	Valid

Valid	Valid=1 indicates this line of descriptor is effective. If Valid=0 generate ADMA Error Interrupt and stop ADMA to prevent runaway.
End	End=1 indicates to end of descriptor. The Transfer Complete Interrupt is generated when the operation of the descriptor line is completed.
Int	INT=1 generates DMA Interrupt when the operation of the descriptor line is completed.

Act2	Act1	Symbol	Comment	Operation
0	0	Nop	No Operation	Do not execute current line and go to next line.
0	1	rsv	reserved	(Same as Nop. Do not execute current line and go to next line.)
1	0	Tran	Transfer Data	Transfer data of one descriptor line
1	1	Link	Link Descriptor	Link to another descriptor

Fig. 5 : 32-bit Address Descriptor Table

ADMA2 States

Figure 6 shows state diagram of ADMA2. 4 states are defined; Fetch Descriptor state, Change Address state, Transfer Data state, and Stop ADMA state.

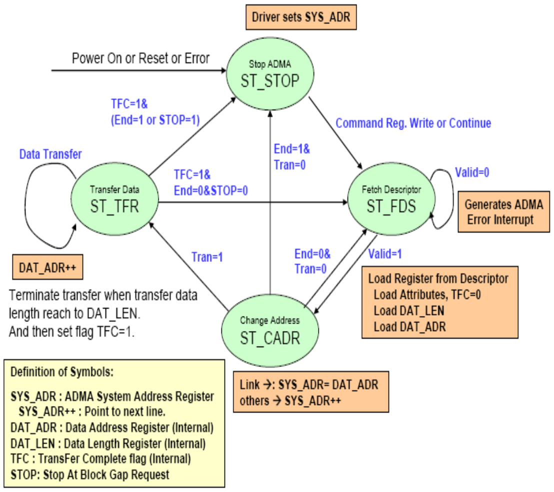
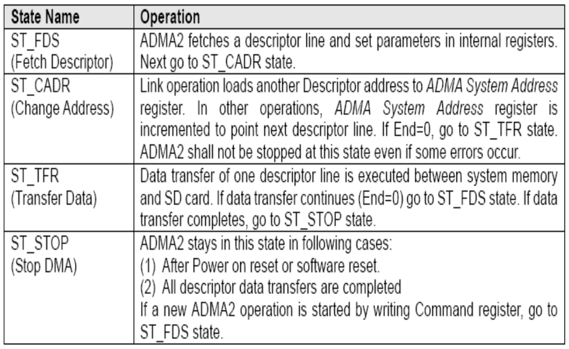

Driver sets SYS_ADR
Power On or Reset or Error
Stop ADMA
ST_STOP
TFC=1&
(End=1 or STOP=1)
Command Reg. Write or Continue
Data Transfer
End=1&
Tran=0
Valid=0
TFC=1&
End=0&STOP=0
Transfer Data
ST_TFR
Fetch Descriptor
ST_FDS
Generates ADMA Error Interrupt
Tran=1
End=0&
Tran=0
Valid=1
DAT_ADR++
Load Register from Descriptor
Load Attributes, TFC=0
Load DAT_LEN
Load DAT_ADR
Terminate transfer when transfer data length reach to DAT_LEN.
And then set flag TFC=1.
Change Address
ST_CADR
Definition of Symbols:
SYS_ADR : ADMA System Address Register
SYS_ADR++ : Point to next line.
DAT_ADR : Data Address Register (Internal)
DAT_LEN : Data Length Register (Internal)
TFC : TransFer Complete flag (Internal)
STOP: Stop At Block Gap Request
Link → SYS_ADR= DAT_ADR
others → SYS_ADR++

Fig. 6 : State Diagram of ADMA2

Operation of each state is explained in Table-1 below



State Name	Operation
ST_FDS (Fetch Descriptor)	ADMA2 fetches a descriptor line and set parameters in internal registers. Next go to ST_CADR state.
ST_CADR (Change Address)	Link operation loads another Descriptor address to *ADMA System Address* register. In other operations, *ADMA System Address* register is incremented to point next descriptor line. If End=0, go to ST_TFR state. ADMA2 shall not be stopped at this state even if some errors occur.
ST_TFR (Transfer Data)	Data transfer of one descriptor line is executed between system memory and SD card. If data transfer continues (End=0) go to ST_FDS state. If data transfer completes, go to ST_STOP state.
ST_STOP (Stop DMA)	ADMA2 stays in this state in following cases: (1) After Power on reset or software reset. (2) All descriptor data transfers are completed If a new ADMA2 operation is started by writing Command register, go to ST_FDS state.

V. RESULTS

Simulation report:

The Test cases were developed for verifying the different modes of Data Transfer. The three modes include non-DMA, DMA and ADMA mode. The test case is written for different block size and block count and waveforms are verified.

The figure shows the snap shot of verification results.

1. Simple non DMA data transfer operation:

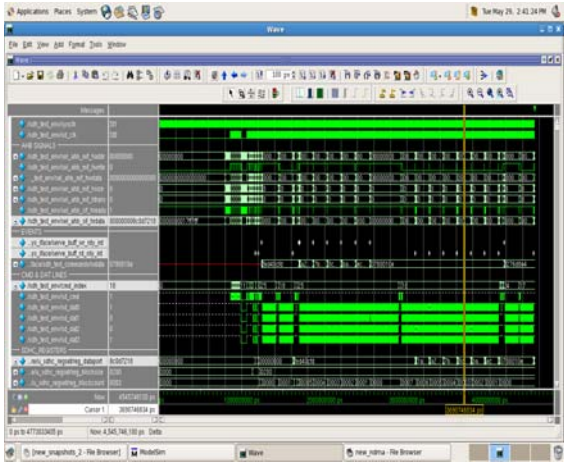


Fig. 7 : Waveform of Simple Non-DMA transfer

2. Simple DMA mode Data transfer operation:

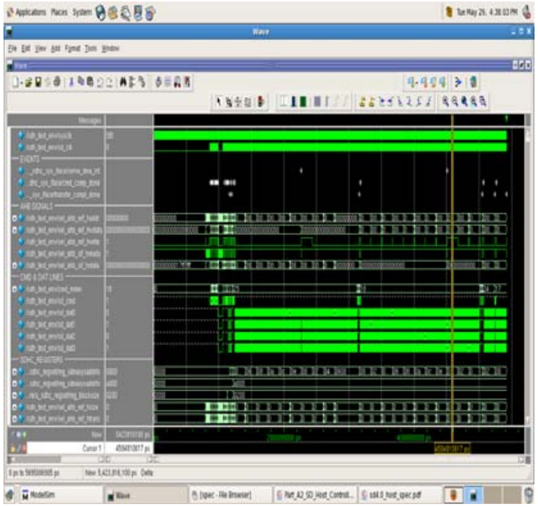


Fig. 8 : Waveform of Simple DMA mode transfer

3. Simple ADMA mode Data transfer operation:

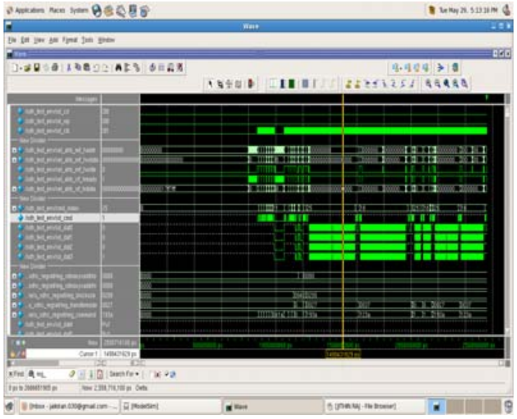


Fig. 9 : Waveform of Simple ADMA mode transfer

VI. CONCLUSIONS AND FUTURE WORK

We developed the verification environment and verified the Secure Digital Host Controller using BFMs (Bus Functional Model) designed in Verilog HDL. Test Cases covering different scenarios are written in Verilog HDL to verify various functionalities of Host controller. The functionalities include DMA, ADMA and Non DMA mode of Data transfer. Future improvement in the SD Host Controller is to add UHS-II mode which can provide bus speed up to 312 Mbps.

REFERENCES

[1] Technical Committee SD Association. (2011): SD Specifications Part A2 SD Host Controller Simplified Specification Version 3.00. Retrieved February 25, 2011 from SD website: http://www.sdcard.org.

[2] Technical committee SD association. (2011): SD Specifications Part 1Physical Layer Simplified Specification Version 3.01. Retrieved February 18, 2011 from SD website: http://www.sdcard.org.

[3] AMBA Advanced High Performance Bus (AHB) Protocol specification. http://www.arm.com/

[4] Chuan-Sheng Lin and Lan-Rong Dung, “A NAND Flash Memory Controller for SD/MMC Flash Memory Card”. IEEE Transactions on Magnetics, vol.43, No.2, February2007.

[5] P. Rashinkar, P. Paterson, Leena singh, “System-on-a-chip verification methodology and techniques,” Kluwer Academic Publishers, 2001.

A New Parallel Interference Cancellation Algorithm for RAKE Systems

G. Swarnalatha & C. Subhas

ECE Department, Sree Vidyanikethan Engineering College

Abstract – In order to suppres the multi-path interference (MPI) in the DS CDMA system, a new RAKE receiver based on parallel interference elimination is first proposed in this paper data symbol tentative decision is obtained by linear decision; the multi-path interference are evaluated by tentative decision and known user information. Then the performance over Rayleigh fading channel are analyzed and compared to conventional parallel interference cancellation (PIC) and RAKE receiver. It is shown that RAKE receiver performance can be improved greatly by using this method with simple structure and easy implementation.

Keywords - *parallel interference cancellation; RAKE receiver; multi-path interference; Rayleigh fading.*

I. INTRODUCTION

When the Signal through the wireless channel, it will produce multiple path fading inevitably. Turin analysed the effects of multiple path interference to the receiver in spread spectrum communication and disussed some anti-interference techniques[1]. The common methods used to suppress the multiple path interference have as the back detect progression and the Rake receiver etc. In the DS CDMA systems, the time domain Rake receivers are used to distinguish, correct and combine different time delay multi-path signal so the inter-symbol interferences are overcome and the path diversity is obtained. By the signal energy of the every path, the Rake receiver obtains diversity gain and as an effective anti multi-path technique, it has become the non-absent key technique in the Third Mobile Communication. Meanwhile, because of the multi-path transmission in channels, multi-path interferences exist between different paths in the identity user. Especially, when the self-correlation between spread codes is worse, the performance about conventional Rake receiver will degrade. So combining the Rake receiver to multi-user detector especially nonlinear multi-user detector has become the research hotspot recently[2~5]. Based on these research, a new parallel interference cancellation method is proposed in this paper in order to eliminate further multi-access and multi-path interference and improve the performance of the Rake receiver in CDMA systems.

II . THE SYSTEM MODEL

A. THE TRANSMITTER MODEL

Considering a single district with K users, the baseband signal for the kth user is given by

$$s_k(t) = \sqrt{2E_{ck}}\, d_k(t) c_k(t) \cos(\omega_c t + \theta_k) \tag{1}$$

Where E_{ck} , d_k (t) stand for the signal power and information stream for the Kth user respectively, ω_c is the carrier frequency, θ_k is the carrier frequency phase, c_k (t) is the signature wave for the Kth user, T_c is the chip width, T_s is the every symbol width for user information stream. So the spread gain is obtained: $N= T_s/T_c$

B. THE CHANNEL MODEL

Rake receivers are designed as taking out delay line model in frequency-selective channel, so the complex lowpass equivalent impulse response can be written as

$$h(t) = \sum_{l=1}^{L} \alpha_l e^{j\beta_l} \delta(t - lT_c) \tag{2}$$

Here, L is resolvable path number, α_l and β_l is the gain and phase of the lth resolvable path respectively, α_l obeys to Rayleigh distribution in the Rayleigh channel.

C. THE RECEIVING SIGNAL

For simplicity, a synchronous uplink CDMA system in the Rayleigh channel is considered, the received signal can be written as

$$r(t)=\sum_{k=1}^{K}\sqrt{E_c}\sum_{l=1}^{L}\sum_{j=-\infty}^{\infty}\alpha_{l,k}b_k(i)c_k(t-\tau_{l,k})\cos(\beta_{l,k}+\theta_k)+n(t) \tag{3}$$

In eq 3, $n(t)$ represents the Additive White Gaussian Noise(AWGN).

III. THE RECEIVER MODEL

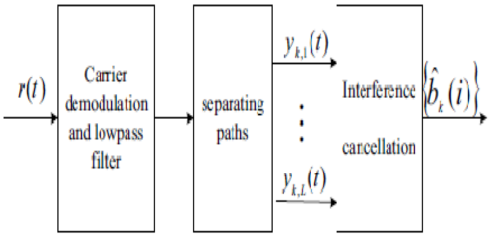

$r(t)$
Carrier demodulation and lowpass filter
separating paths
$y_{k,1}(t)$
$y_{k,L}(t)$
Interference cancellation
$\{\hat{b}_k(i)\}$

Fig. 1 : Rake parallel interference canceller receiver

The receiver diagram is shown in figure 1. After the received signal is demodulated by carrier correlation and low pass filtered, the paths are separated. So the output of the lth path through the correlator is

$$y_{k,l}(i)=\int_{iT}^{(i+1)T}r(t)c_k(t-iT)dt \tag{4}$$

Substituting eq 3 in eq 4

$$y_{k,l}(i)=N\sqrt{E_{ck}}\alpha_{k,l}b_k(i)+\sqrt{E_{ck}}\sum_{\substack{l'=1\\ l'\neq l}}^{L}\alpha_{k,l'}\cos\theta_{k,l'}+\sum_{\substack{k'=1\\ k'\neq k}}^{K}\sqrt{E_{ck'}}\sum_{l=1}^{L}\alpha_{k',l}b_{k'}(i)\rho_{k',k}+n_k(i)$$

$$=\sqrt{E_{ck}}\alpha_{k,l}b_k(i)+I^P+I^M+n_k(i) \tag{5}$$

Where:

$$\rho_{k',k}=\int_0^T c_k(t)c_{k'}(t)dt,\theta'_{k,l}=\theta_{k,l'}+\beta_{k,l'}$$

In the equation 5 the first item stands for signal, the second item is the multi-path interferences of the other (L-1) paths to the lth path, shortening for MPI, the third is the multiple access interference(MAI) of the other (K-1) users to the kth user, the last is the additive white Gaussian noise(AWGN). To traditional Rake receiver, the result using maximum ratio combining is:

$$y_k(i)=\sum_{l=1}^{L}\widetilde{\alpha}_{l,k}y_{k,l}(i) \tag{6}$$

In equation 6 $\widetilde{\alpha}_{l,k}$ is evaluation value for kth user in the lth path. When equation 6 is decided, the tentative test value about the symbol is obtained as

$$\widetilde{b}_k(i)=D(y_k(i)) \tag{7}$$

Here D(.) stands for decision function, which may be hard decision or soft decision. Traditional PIC adopts hard decision. In this thesis simple soft decision linear decision is used.In order to eliminate the MPI and MAI in the above equation, a new parallel interference cancellation arithmetic is proposed as shown in the below figure 5.2

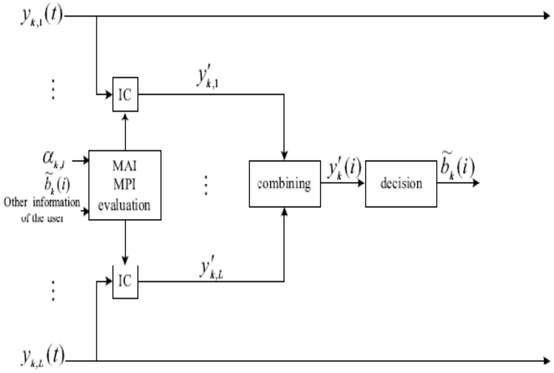


Fig. 2: A New PIC receiver

In the figure 2 taking the ith symbol decision of the kth user as example. It can be seen that interference estimators estimate the MAI and MPI of each channel path according to the tentative test value and other known user information (as the spread frequency codes of other users, signal amplitude, phrase etc) in the interference cancellation stage and subtract them, then the correlated outputs after interference cancellation for each path are combined, finally, the symbol decision result is given in the interference cancelling stage, MPI and MAI are reconstructed firstly, supposed that test value $\widetilde{b}_k(i)$ can be given by paper [9] and [10]

$$\widetilde{b}_k(i)\approx\begin{cases}\lambda_k(b_k(i))/2, & \lambda_k(b_k(i))\le 2\\ \operatorname{sgn}(\lambda_k(b_k(i))), & \lambda_k(b_k(i))\ge 2\end{cases} \tag{8}$$

Where $\lambda_k(b_k(i))=\ln\dfrac{P(y_{k,l}(i)|b_k(i)=+1}{P(y_{k,l}(i)|b_k(i)=-1}$

To the lth path, according to and other users information, reconstructed MPI and MAI can be written as respectively:

MPI: $\hat{I}^p=\widetilde{\alpha}_{l,k}\widetilde{b}(i)\cos(\widetilde{\theta}'_{l,k})$

MAI: $\hat{I}^{A} = \sum_{\substack{k'=1 \\ k'\neq k}}^{K} \sqrt{E_{ck}}\, \rho_{k',k} \tilde{b}_{k'}(i)$

When reconstructed interference being subtracted and it can be obtained

$$y'_{k,l}(i) = y_{k,l}(i) - \tilde{\alpha}_{k,l}(i)\tilde{b}_k(i)\cos(\tilde{\theta}'_{l,k}) - \sqrt{E_c}\sum_{\substack{k'=1 \\ k'\neq k}}^{K}\sqrt{E_{ck}}\,\rho_{k',k}\tilde{b}_{k'}(i) \tag{9}$$

Finally we obtain the received vector as

$$y_k'(i) = \sum_{l=1}^{L} \tilde{\alpha}_{k,l}\, y'_{k,l}(i) \tag{10}$$

The final decision for the ith symbol of kth user is

$$\hat{b}_k(i) = \mathrm{sgn}(y_k'(i)) \tag{11}$$

IV. SIMULATION RESULTS

The figure 3 is the BER performance of conventional PIC receiver and the proposed PIC in this thesis. It can be seen from Figure that when the signal noise ratio (SNR) is lower, the two kind receivers show little difference, but with the SNR increasing, the proposed PIC show superior performance than traditional PIC and the PIC system proposed in this paper appears the best performance.

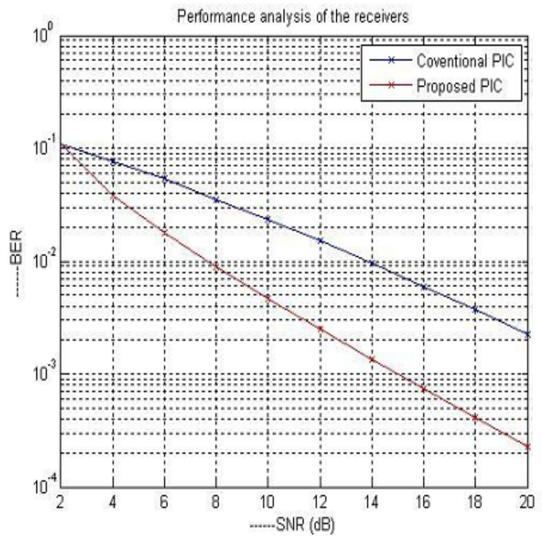


Fig. 3 : BER versus SNR for performance comparison of the two different receivers

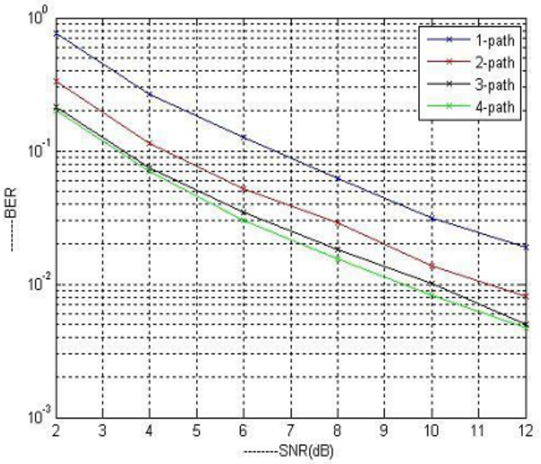


Fig. 4 : BER versus SNR for performance comparison for different values of L.

In Figure 4 the bit error ratios(BER) of the system for different path numbers are compared. It can be seen that when L=1, the system appears the worst performance because only MAI is eliminated. When L=3, the performance of the system is better than L=2; when L=4, the performance can not be improved obviously. So in MPI system, we usually set L=3. Because each path power is impossible same in practical channel, the receiver performance can be improve obviously only through eliminating those paths which SNR reach a certain threshold.

V. CONCLUSION

In this paper, a Rake PIC based on linear decision is proposed. The theory and simulation indicate that this technique can improve greatly the receiver performance in the multipath fading channel and have lower realization complexity, which will very suit for mobile single user receiver restricted by volume and hardware.

REFERENCES

[1] Theodore S. Rappaport Wireless communication principles and practice, Second edition. Pearson Education.

[2] J. G Proakis, Digital communication 3rd edition. New York : McGraw-Hill.

[3] Li-xin Song, zheng-qian Wan and Xiao-bo “A new parallel interference cancellation algorithm for RAKE systems,” IEEE international conference Genetic and Evolutionary computing, pp.806-809,2009.

[4] G.L.Turin. “Introduction to spread-spectrum antimultipath techniques and their

application to urban digital radio," Proc. IEEE, 1980, 68(3):328-353.

[5] Baltersee J, Fock G, Schulz-Rittich P. "A novel multipath interference cancellation scheme for Rake channel estimation," IEEE VTC, 2007 Spring 2:1493-1497.

[6] Junhui Zhao "Improved parallel interference cancellation algorithm for multicode CDMA system" IEEE International Conference Neural Networks&Signal Processing, 2003(11):1560-1564.

[7] Rong Hu ," Multipath interference cancellation and modified RAKE receiver "Jorunal of electronics and information,2002,24 (6):773-780.

[8] Junhui Zhao "An Improved RAKE Receiver Algorithm of the CDMA System ,"Journal of applied sciences, 2004,22(3):298-302.

[9] X.Wang,. "Iterative(Turbo) soft interference cancellation and decoding for coded CDMA, "IEEE Trans.Communication volume(47),1999(7):1046-1060.

[10] Xu Guoxiong,Gan Liangcai,Huang Tianxi. "A New parallel interference cancellation scheme for DS/CDMA system," Wuhan University Journal of Natural Science,2004,9(1):27-30.

[11] Rong Hu, Aiping Huang "Multipath Interference Cancellation and Modified RAKE Receiver," IEEE pp:1508-1512.

[12] "Multiusage detection in asynchronous Code Division Multiple Access communications", by M.K Varanasi and Behnaan aazhang, pp: 509-519;1990

[13] S.Haykins, Digital Communication, Wiley, 1988.

[14] A. Johansson and A. Svensson, "Successive interference cancellation in multiple data rate DS/CDMA systems," in Proc. IEEE VTC '95. Chicago, IL, July 26-28, 1995.

[15] J. Andrews, "Successive Interference Cancellation for Uplink CDMA," Ph.D. Dissertation, Stanford University, 2002.

[16] P. Patel and J. Holtzman, "Performance comparison of a DS/CDMA system using successive interference cancellation (IC) scheme and a parallel IC scheme under fading", Proceedings of IEEE ICC '94, pp. 510-514, 1994.

◈◈◈

An Adaptive Image Resolution Enhancement by using Multi-Resolution Transforms

Venkataramakrushna Durga & G.Parameswararao

Dept. Of E.C.E., Sri Sivani College of Engineering, Chilakapalem Jn, Srikakulam, Andhra Pradesh, India

Abstract – In this correspondence, the paper presents a regularized image sequence interpolation algorithm, an image resolution enhancement technique based on interpolation of the high frequency sub-band images obtained by discrete wavelet transform (DWT) and the input image. Edges in images convey a great deal of information, but wavelet transforms do not provide an economical representation. Thus, popular wavelet-based compression and restoration techniques perform poorly in the presence of edges. The edges are enhanced by introducing an intermediate stage by using stationary wavelet transform (SWT) which is the one of the Multi-Resolution Transforms (MRT). DWT is applied in order to decompose an input image into different sub-bands. Then the high frequency sub-bands as well as the input image are interpolated and are being modified by using high frequency sub-band obtained through SWT. Then all these sub-bands are combined to generate a new high resolution image by using inverse DWT (IDWT). The quantitative and visual results are showing the Enhancement of the edges of an input image by introducing an intermediate stage by using Multi-Resolution Transforms (MRT)

Keywords *- Discrete wavelet transform (DWT), Image interpolation, Image super resolution, Stationary wavelet transform (SWT), Multi-Resolution Transforms (MRT)*

I. INTRODUCTION

In his classic treatise on the human visual system, we focus on the importance of the representation of information for various image processing tasks. The way in which information is represented brings out certain types of features while hiding others. Signal compression and estimation applications also rely heavily on having an efficient representation of image data; we would like to approximate a signal with a small number of parameters. Therefore, we seek a transform which yields an efficient representation while bringing out the desired features of the signal. Resolution has been frequently referred as an important aspect of an image. Images are being processed in order to obtain more enhanced resolution. One of the commonly used techniques for image resolution enhancement is Interpolation. Interpolation has been widely used in many image processing applications such as facial reconstruction, multiple description coding, and super resolution. There are three well known interpolation techniques, namely nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation.

In this paper we mainly deal with image interpolation techniques to enhance the image quality in the sense of resolution. The terminology “resolution” represents the number of pixels in an image, which determines the physical size of the image, and at the same time it also represents the fidelity to high-frequency details in the image. By this reason, resolution is a fundamental issue in evaluating the quality of various image processing systems. Image interpolation is used to obtain a higher resolution image from a low resolution image, and therefore it is very important in multi-resolution or high-resolution image processing. For example, the spatial scalability function in MPEG-2 and wavelet-based image processing techniques require image interpolation techniques. On the other hand, high-resolution image processing applications such as digital HDTV, aerial photo, medical imaging, and military purpose images, need high-resolution image interpolation algorithms. Recently, it can also be used in changing the format of various types of images and videos, and in increasing the resolution of images.

Image resolution enhancement in the wavelet domain is a relatively new research topic and recently many new algorithms have been proposed. Discrete wavelet transform (DWT) is one of the recent wavelet transforms used in image processing. DWT decomposes an image into different sub-band images, namely low-low (LL), lowhigh (LH), high-low (HL), and high-high (HH). Another recent wavelet transform which has been used in several image processing applications is stationary wavelet transform (SWT). In short, SWT is similar to DWT but it does not use down-sampling, hence the sub-bands will have the same size as the input image

In this work, we are proposing an image resolution enhancement technique which generates sharper high resolution image. The proposed technique uses DWT to decompose a low resolution image into different sub-bands. Then the three high frequency sub-band images have been interpolated using bicubic interpolation. The high frequency sub-bands obtained by SWT of the input image are being incremented into the interpolated high frequency sub-bands in order to correct the estimated coefficients. In parallel, the input image is also interpolated separately. Finally, corrected interpolated high frequency sub-bands and interpolated input image are combined by using inverse DWT (IDWT) to achieve a high resolution output image

II. IMAGE RESOLUTION ENHANCEMENT

Image enhancement is the improvement of digital image quality (wanted e.g. for visual inspection or for machine analysis), without knowledge about the source of degradation. If the source of degradation is known, one calls the process image restoration. Both are *iconical processes*, viz. input and outputs are images.

Many different, often elementary and heuristic methods are used to improve images in some sense. The problem is, of course, not well defined, as there is no objective measure for image quality. Here, we discuss a few recipes that have shown to be useful both for the human observer and/or for machine recognition. These methods are very problem-oriented: a method that works fine in one case may be completely inadequate for another problem.

In image resolution enhancement by using interpolation the main point we have to concentrate is on its high frequency components (i.e., edges), this is nothing but which is due to the smoothing caused by interpolation. In order to increase the quality of the super resolved image, it is essential to preserve the edges. To achieve this work, DWT has been employed in order to preserve the high frequency components of the image. The redundancy and shift invariance of the DWT mean that DWT coefficients are inherently interpolable.

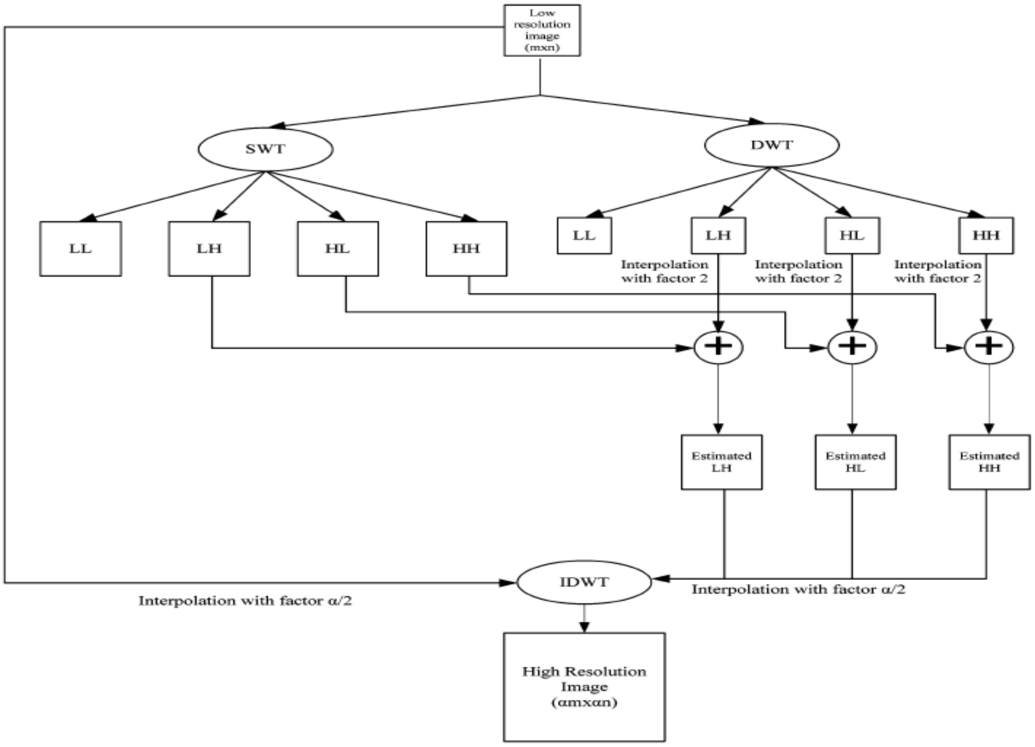

Low resolution image (mxn)
SWT
DWT
LL
LH
HL
HH
LL
LH
HL
HH
Interpolation with factor 2
Interpolation with factor 2
Interpolation with factor 2
Estimated LH
Estimated HL
Estimated HH
IDWT
Interpolation with factor α/2
Interpolation with factor α/2
High Resolution Image (αmxαn)

Fig. 1 Block diagram of the super resolved interpolation algorithm

In this correspondence, we are using one level DWT (with Daubechies 9/7 as wavelet function) to decompose an input image into different sub-band images. There are three high frequency sub-bands like LH, HL, and HH contain the high frequency components of the input image. In the proposed technique, bicubic interpolation with enlargement factor of 2 is applied to high frequency sub-band images. Down sampling in each of the DWT sub-bands causes information loss in the respective sub-bands. That is why SWT is employed to minimize this loss.

The SWT high frequency sub-bands and the interpolated high frequency sub-bands have the same size so that they can be added with each other. For the further higher enlargement the new corrected high frequency sub-bands can be interpolated. Also it is known that in the wavelet domain, the low resolution image is obtained by low-pass filtering of the high resolution image. In other words, low frequency sub-band is the low resolution of the original image. Therefore, instead of using low frequency sub-band, which contains less information than the original high resolution image, we are using the input image for the interpolation of low frequency sub-band image. Using input image instead of low frequency sub-band increases the quality of the super resolved image. Fig. 1 illustrates the block diagram of the proposed image resolution enhancement technique.

By interpolating input image by α/2, and high frequency sub-bands by 2 and α in the intermediate and final interpolation stages respectively, and then by applying IDWT, as illustrated in Fig. 1, the output image will contain sharper edges than the interpolated image obtained by interpolation of the input image directly. This is due to the fact that, the interpolation of isolated high frequency components in high frequency sub-bands and using the corrections obtained by adding high frequency sub-bands of SWT of the input image, will preserve more high frequency components after the interpolation than interpolating input image directly.

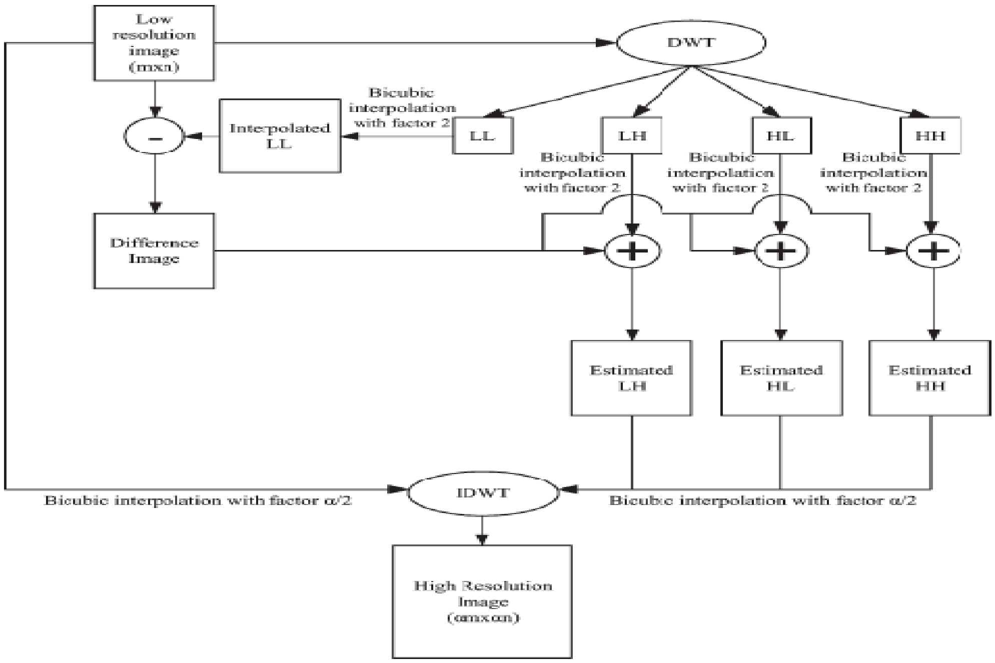

Low resolution image (mxn)
DWT
Bicubic interpolation with factor 2
Interpolated LL
LL
LH
HL
HH
Bicubic interpolation with factor 2
Bicubic interpolation with factor 2
Bicubic interpolation with factor 2
Difference Image
Estimated LH
Estimated HL
Estimated HH
Bicubic interpolation with factor α/2
IDWT
Bicubic interpolation with factor α/2
High Resolution Image (αmxαn)

Fig. 2 : Block diagram of the proposed algorithm.

We are proposing an image resolution enhancement technique which generates sharper high resolution image. The proposed technique uses DWT to decompose a low resolution image into different sub-bands. Then the three high frequency sub-band images have been interpolated using bi-cubic interpolation. The high frequency sub-bands obtained by SWT of the input image are being incremented into the interpolated high frequency sub-bands in order to correct the estimated coefficients. In parallel, the input image is also interpolated separately.

Finally, corrected interpolated high frequency sub-bands and interpolated input image are combined by using inverse DWT (IDWT) to achieve a high resolution output image. The proposed technique has been compared with conventional and state-of-art image resolution enhancement techniques.

We propose a resolution-enhancement technique using interpolated DWT high-frequency subband images and the input low-resolution image. Inverse DWT (IDWT) has been applied to combine all these images to generate the final resolution-enhanced image. In order to achieve a sharper image, we propose to use an intermediate stage for estimating the highfrequency subbands by utilizing the difference image obtained by subtracting the input image and its interpolated LL subband. The proposed technique has been compared with standard interpolation techniques, wavelet zero padding (WZP), where the unknown coefficients in high-frequency subbands are replaced with zeros, state-of-art techniques. The performance of the proposed technique over performs all available state-of-art methods for image resolution enhancement.

In all steps of the proposed satellite image resolution enhancement technique, Daubechies wavelet transform as mother wavelet function and bicubic interpolation as interpolation technique have been used.

III. RESULTS AND DISCUSSIONS

The Super resolved image of Baboon’s picture using proposed technique is shown Fig. 3 the image in (f) are much better than the low resolution image in (a), Bilinear interpolated image (b),Bicubic image (c) ,Super resolved image using WZP (d) and super resolved image by using the interpolation (e) . Note that the input low resolution images have been obtained by down-sampling the original high resolution images. In order to show the effectiveness of the proposed method over the conventional and state-of-art image resolution enhancement techniques, four well-known test images such as Baboon, Lena, and Barbara etc., with different features are used for comparison. Table I compares the PSNR performance of the proposed technique using bicubic interpolation with conventional and state-of-art resolution enhancement techniques: bilinear, bicubic, Super resolved image using WZP,and Super resolved image using bilinear interpolation preserving image interpolation. Additionally, in order to have more comprehensive comparison, the performance of the super resolved image by using SWT only (SWT-SR) is also included in the table. The results in Table I indicate that the proposed technique over-performs the aforementioned conventional and state-of-art image resolution enhancement techniques. Table I also indicates that the proposed technique over-performs the aforementioned conventional and state-of-art image resolution enhancement techniques

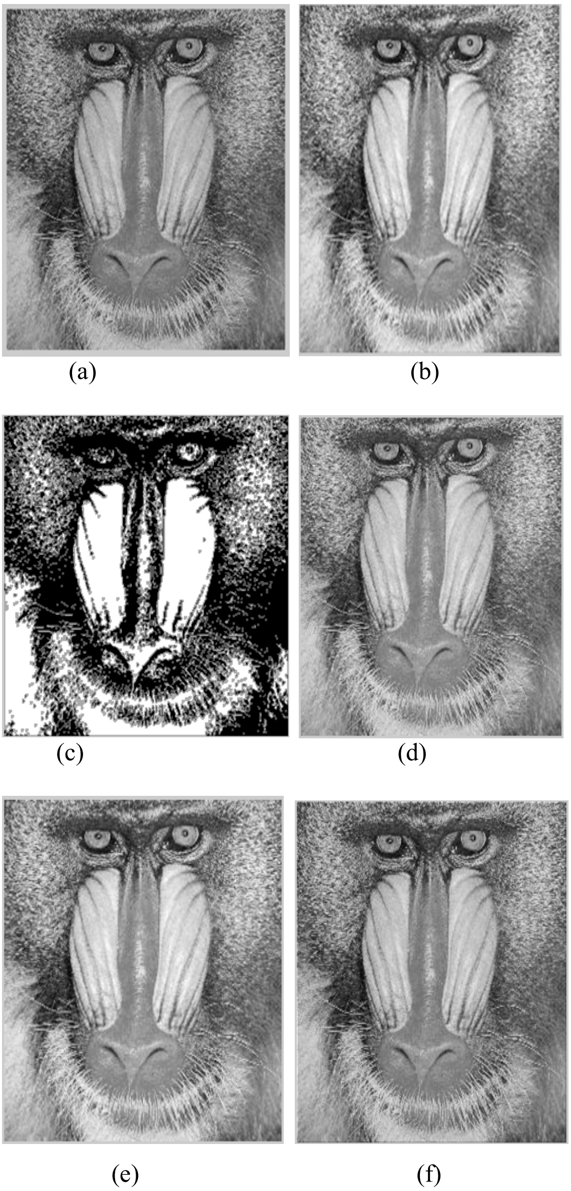


Fig.3 (a) Original low resolution Baboon’s image. (b) Bilinear interpolated image (c) Bicubic interpolated image. (d) Super resolved image using WZP. (e) Super resolved image using bilinear interpolation (f) Proposed technique.

IV. CONCLUSION

The proposed technique has been tested on well-known benchmark images, where their PSNR and visual results show the superiority of proposed technique over the conventional and state-of-art image resolution enhancement techniques. The image resolution enhancement technique based on the interpolation of the high frequency sub-bands obtained by DWT, correcting the high frequency sub-band estimation by using SWT high frequency sub-bands, and the input image. The proposed technique uses DWT to decompose an image into different sub-bands, and then the high frequency sub-band images have been interpolated. The interpolated high frequency sub-band coefficients have been corrected by using the high frequency sub-bands achieved by SWT of the input image. An original image is interpolated with half of the interpolation factor used for interpolation the high frequency sub-bands. Afterwards all these images have been combined using IDWT to generate a super resolved imaged. The proposed technique has been tested on well-known benchmark images, where their PSNR and visual results show the superiority of proposed technique over the conventional and state-of-art image resolution enhancement techniques.

TABLE I

	PSNR(dB)		
Techniques\Images	**Baboon**	**Lena**	**Barbara**
Bilinear	27.7120	29.2821	28.0735
Bicubic	29.6867	31.3053	30.1179
WZP	34.8653	36.4802	35.4011
Super resolved interpolated technique	35.0445	36.8774	35.7141
Proposed Technique	**35.5102**	**37.1994**	**36.0071**

PSNR (DB) RESULTS FOR RESOLUTION ENHANCEMENT FROM 128×128 TO 512×512 OF THE PROPOSED TECHNIQUE COMPARED WITH THE CONVENTIONAL AND STATE-OF-ART IMAGE RESOLUTION ENHANCEMENT TECHNIQUES

REFERENCES

[1] IMAGE Resolution Enhancement by Using Discrete and Stationary Wavelet Decomposition by Hasan Demirel and Gholamreza Anbarjafari, IEEE Transactions on Image Processing, Vol. 20, NO. 5, MAY 2011

[2] Y. Rener, J. Wei, and C. Ken, "Downsample-based multiple description coding and post-processing of decoding," in Proc. 27th Chinese Control Conf., Jul. 16–18, 2008, pp. 253–256.

[3] H. Demirel, G. Anbarjafari, and S. Izadpanahi, "Improved motionbased localized super resolution technique using discrete wavelet transform for low resolution video enhancement," in Proc. 17th Eur. Signal Process. Conf., Glasgow, Scotland, Aug. 2009, pp. 1097–1101.

[4] Y. Piao, I. Shin, and H. W. Park, "Image resolution enhancement using inter-subband correlation in wavelet domain," in Proc. Int. Conf. Image Process., 2007, vol. 1, pp. I-445–448.

[5] H. Demirel and G. Anbarjafari, "Satellite image resolution enhancement using complex wavelet transform," IEEE Geoscience and Remote Sensing Letter, vol. 7, no. 1, pp. 123–126, Jan. 2010.

[6] C. B. Atkins, C. A. Bouman, and J. P. Allebach, "Optimal image scaling using pixel classification," in Proc. Int. Conf. Image Process.,Oct. 7–10, 2001, vol. 3, pp. 864–867.

[7] W. K. Carey, D. B. Chuang, and S. S. Hemami, "Regularity-preserving image interpolation," IEEE Trans. Image Process., vol. 8, no. 9, pp. 1295–1297, Sep. 1999.

[8] S. Mallat, A Wavelet Tour of Signal Processing, 2nd ed. New York: Academic, 1999.

[9] J. E. Fowler, "The redundant discrete wavelet transform and additive noise,"Mississippi State ERC, Mississippi State University, Tech. Rep. MSSU-COE-ERC-04-04, Mar. 2004.

[10] X. Li and M. T. Orchard, "New edge-directed interpolation," IEEE Trans. Image Process., vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[11] K. Kinebuchi, D. D. Muresan, and R. G. Baraniuk, "Waveletbased statistical signal processing using hidden Markov models," in Proc. Int. Conf. Acoust., Speech, Signal Process., 2001, vol. 3, pp. 7–11.

[12] S. Zhao, H. Han, and S. Peng, "Wavelet domain HMT-based image super resolution," in Proc. IEEE Int. Conf. Image Process., Sep. 2003, vol. 2, pp. 933–936.

[13] A. Temizel and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning," Electron. Lett., vol. 41, no. 3, pp. 119–121, Feb. 3, 2005.

[14] A. Temizel and T. Vlachos, "Image resolution upscaling in the wavelet domain using directional

cycle spinning," J. Electron. Imag., vol. 14, no. 4, 2005.

[15] G. Anbarjafari and H. Demirel, "Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image," ETRI J., vol. 32, no. 3, pp. 390–394, Jun. 2010.

[16] A. Temizel, "Image resolution enhancement using wavelet domain hidden Markov tree and coefficient sign estimation," in Proc. Int. Conf. Image Process., 2007, vol. 5, pp. V-381–384..

[17] L. Yi-bo, X. Hong, and Z. Sen-yue, "The wrinkle generation method for facial reconstruction based on extraction of partition wrinkle line features and fractal interpolation," in Proc. 4th Int. Conf. Image Graph., Aug. 22–24, 2007, pp. 933–937.

◈◈◈

Ensuring Data Storage Security in Cloud Computing With Effect of Kerberos

Mehdi Hojabri

Abstract - cloud computing has been envisioned as the next-generation technology of IT enterprise.In other method is Internet based technology where the users can subscribe high quality of services from data and software that resides solely in the remote servers. This make many advantage and drawback for the users to create and store data in the remote servers thereby utilizing fewer resources in client system. And the other hand management of the data and software may not be fully trustworthy which possesses many security challenges.totaly the security is important aspect of quality of service.In this article, we concentrate on cloud data storage security by the Implementation of kerberos athentication service..

I. INTRODUCTION

As cloud computing continues to thrive and as more and more enterprises penetrate the cloud, security becomes a further pressing issue. Several trends are opening up the era of cloud computing, which is an internet base development and use of computer technology. The ever cheaper and more powerful processors, together with the software as a service(SaaS)computing architecture ,are transforming data center into pools of computing service on a huge scale. In this paper i survey the Kerberos (a.k.a.: Cerberus) effect in cloud computing server. Kerberos uses strong encryption and a complex ticket-granting algorithm to authenticate users on a network. Also of interest to many of users, Kerberos has the ability to distribute "session keys" to allow encrypted data streams over an IP network. Each user for connecting to the cloud at the first should make the profile and user ID. after that it must get the password. and also the information of all participating user such as User ID,hashed password will save in the large Data Base for more secure.all user are register with the kerberos server.in this method each user want connect to the cloud server at the first time he or she logs on to workstation and:

1. Send the Request for ticket granting ticket to the As.
2. As verifies user’s access right in data base,create ticket-granting ticket and session key.result are encryped using key derived from user password.
3. User will send the request cloud service granting ticket to TGS.

4 The TGS will send the Ticket+session key to the user.(it execute one per type of service).

5. Workstation sends ticket and authenticator to cloud server provider.
6. Server verifies ticket and authenticator match,then grant access to service.

In this paper i have try to assume each user for connecting and utilize the cloud server must create the profile and applay some private information for more secure.

II. PROBLEM STATEMENT

A representation network architecture for cloud data sorage with effect of kerberos is illustrated in Figure 1. six different network entities can be identified as follows:

User : user,who have data to be stored in the cloud and rely on the cloud for data computation,consist of both individual consumer and organization.and want access to cloud server for doing job with effect of kerberos service.

Cloud service provider(CSP) : Cloud service providers offer cloud solutions, like Google Apps, that are delivered electronically over the internet. Unlike a managed service provider, cloud service providers do not sell or install hardware – everything they offer is stored online and accessible securely from anywhere. There are many advantages to working with a cloud service provider like Cloud Sherpas when switching from your old email and collaboration software.

Kerberos operation : Kerberos is an authentication protocol for trusted hosts on untrusted networks. The Kerberos protocol is designed to provide reliable authentication over open and insecure networks where communications between the hosts belonging to it may be intercepted.the following requirement for kerberos is: Secure-Reliable-Transparent-Scalable

Authentication service(AS) : An authentication service that know the password of all user and stores these in a centralized database.in addition,the AS shares a unique secret key with each server.

Tickets granting service(TGS) : TGS provide and issue tickets to user who have been authentication to AS.

Data Base : The kerberos server must have the user ID(UID)and hashed password of all participating user in the data base.all user are register with the kerberos server.it make more security in cloud server.

III. ENSURING CLOUD DATA STORAGE WITH EFFECT OF KERBEROS

In cloud data storage system, users store their data in the cloud and for accessing must refer to cloud server provider. thus the correctness of the user being refer to the distributed cloud server must be guaranteed. the data stored in the cloud may be frequently, updated with user including, insertion, deletion, modification, appending, reordering, etc. to ensure this updating is under correctness user is important. so in this paper we introduce one model based on kerberos. in this model each user for gain the cloud server must be register and after added to the data base it can get some qualificatin and after that get the cloud server. and...

In this senario:

1. The client logs on the workstation and requests access a ticket-granting ticket on behalf of the user by sending its user's ID to the AS,together with TGS ID,indicating a request to use the TGS service.
2. The AS responds with a ticket that is encrypted with a key that is derived from user password. when this response arrives at the client, the client prompts the user for his or her password, generates the key, and attempt decrypt the incoming message if the correct password is supplied, the ticket is successfully recovered.

Because only the correct user should know the password, only the correct user can recover the ticket. Thus, we have used the password to obtain credentials from kerberos without having to transmit the password in plaintext.the ticket itself consist of the ID and network address of the user,and the ID of the TGS.this corresponds to the first scenario.the is that this ticket can be used by the client to recuest multiple cloud service granting tickets.so the ticket granting ticket is to be reusable.However,we do not wish an opponent capture the ticket and waits until the user has logged off his or her workstation.The opponent either gain access to that work station or configure his workstation with the same network address as that of the victim.

The ticket include a timestamp ,indicating the data and time for which the ticket wae issued,and alifetime,indicating the length of time for which the ticket is valid.thus,the client know has a reusable ticket and need not bother the user for a password for each new service request.

3. The client request a service-granting ticket on behalf of the user.For this purpose,the client transmits a message to the TGS containing the user's ID,the ID of the desire cloud service,and the ticket-granting ticket.
4. The TGS decrypt the incoming ticket and verifies the succss of the decryption by the presence of its ID.it check to make sure that the lifetime has not expired.Then it compares the user ID and network address with the incoming information to authenticate the user.if the uses is permitted access to V,the TGS issues a ticket to grant access to the requested cloud service provider.

The service-granting provider ticket has the same structure as the ticket-granting ticket.indeed,because the TGS is a server,we would expect that the same elements are needed to authenticate a client to the TGS and to authenticate a client to an application server.Again,the ticket contain a timestamp and lifetime.if the user wants access to the same cloud service at a later time,the client can simply use the previously acquired service-granting ticket and need not bother the user for a password.

Note that the ticket is encrypted with a secret key(K_v) known only to the TGS and the server,preventing alteration.

Finally, with a particular cloud service –granting ticket,the client can gain access to the corresponding service with step 5.

5. The user request access to service on behalf of the user. For this purpose the client transmits a message to the server containing the user's ID and the cloud service granting ticket,The server authentication by using the contents of the ticket.

The table 1 shows how to implement this scenario.

.

(A)Authentication Service Exchange:to obtain ticket-granting Ticket (1) C ⟶ AS: $ID_c\, ID_{tgs}\ \ TS_1$ (2) AS ⟶ C: $Ek_c[k_{c,tgs}\ \ ID_{tgs}\ \ TS_2\ \ lifetime_2\ \ Ticket_{tgs}$ $Ticket_{tgs}=Ek_{tgs}[k_{c,tgs}\ \ ID_cAD_c\ \ ID_{tgs}\ TS_2\ Lifetime_2]$
(B) Ticket-granting cloud service Exchange:to obtain cloud service-granting ticket (3) C ⟶ $TGS:ID_v\ \ Ticket_{tgs}\ \ Authenticator_c$ (4) TGS ⟶ $C:Ek_{c,tgs}\ [k_{c,v}\ \ \ ID_v\ \ TS_4\ \ Ticket_v]$ $Ticket_{tgs}=Ek_{tgs}[k_{c,tgs}\ \ \ ID_c\ \ \ AD_c\ \ ID_{tgs}\ \ TS_2\ \ Lifetime_2]$ $Ticket_v=Ek_v[k_{c,v}\ \ ID_c\ \ AD_c\ \ ID_v\ TS_4\ Lifetime_4]$ $Authenticator_c=EK_{c,tgs}[ID_c\ \ AD_c\ \ TS_3]$
(C) client/Server Authentication Exchange:to obtain cloud service (5) C ⟶ K: $Ticket_v\ \ Authenticator_c$ (6) K ⟶ $C:Ek_{c,v}[TS_5+1]$ (for mutual authentication) $Ticket_v=Ek_v\ [k_{c,v}\ \ ID_c\ \ AD_c\ ID_v\ \ TS_4\ \ Lifetime_4]$ $Authenticator_c=Ek_{c,v}\ [ID_c\ \ AD_c\ TS_5]$

Table1 : Summary of Kerberos mrssage exchange in cloud service

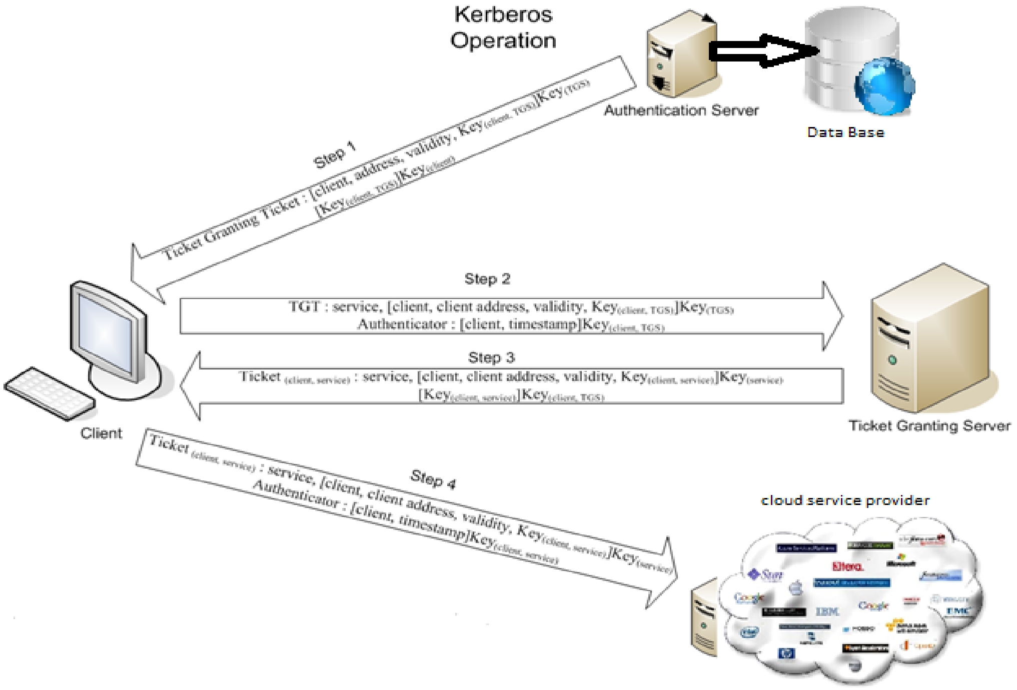

Kerberos Operation
Authentication Server
Data Base
Step 1
Ticket Granting Ticket : [client, address, validity, Key(client, TGS)]Key(TGS)
[Key(client, TGS)]Key(client)
Step 2
TGT : service, [client, client address, validity, Key(client, TGS)]Key(TGS)
Authenticator : [client, timestamp]Key(client, TGS)
Step 3
Ticket (client, service) : service, [client, client address, validity, Key(client, service)]Key(service)
[Key(client, service)]Key(client, TGS)
Client
Ticket Granting Server
Step 4
Ticket (client, service) : service, [client, client address, validity, Key(client, service)]Key(service)
Authenticator : [client, timestamp]Key(client, service)
cloud service provider

Fig.1: Cloud data storage architecture

IV. CONCLUSION

In this paper, we investigated the problem of data security in cloud data storage, which is essentially a distributed storage system. To ensure the correctness of users' data in cloud data storage, and correctness of users who can access to the cloud server , we proposed an effective and flexible distributed scheme with explicit dynamic data support, including kerberos and authentication service.kerberos provides a centralized authentication server whose function is to authenticate user to cloud server and cloud server to user.any user to access the cloud server first should make the profile and password.then it can use the cloud server with gain the qualify.as we know the unique attribute of network is security.so for making more secure network we must make the way for controlling the cloud system and storing the information of user's.we would like for cloud servers to be able to restrict access to authorized users and to be able authenticate request for service.as we know in an un protected network environment,any client can apply to any cloud server for service but kerberos operation with make use of DES,in arather elaborate protocol,to provide the authenticaion service.

REFERENCES

[1] MILL88 Miller, S,;Neuman,B.;Schiller,j.;and Saltzer,j." Kerberos Authentication and authorization System."Section E.2.1,Project Athena Technical plan,M.I.T.Project Athena,Cambridge,MA.27 October 1998.

[2] STE188 Steiner,j.:Neuman, C.;and Schiller,j. "Kerberos:An Authentication Service for Open Networked Systems."Proceeding of the Winter 1988 USENIX Conference,February 1988.

[3] KOHL89 Kohl,j."The Use of Encryption in Kerberos forNetworkAuthentication."Proceeding. Crypto 96,August 1996;published by Springer-Verlag.

[4] KOHL94 Kohl,j.;Neuman,B.;and Ts'o.T."The Evolution of the Kerberos Authentication Service."In Brazier.F.,and Johansen, D.Distributed Open Systems. Los Alamitos, CA:IEEE Computer Society Press,1994. Available at http://Web.mitedu/kerberos/ www.papers,html.

[5] Amazon.com, "Amazon Web Services (AWS)," Online at http://aws.amazon.com, 2008.

[6] N. Gohring, "Amazon's S3 down for several hours," Online at http://www.pcworld.com/ businesscenter/article/142549/amazons s3 down for several hours.html, 2008.

[7] A. Juels and J. Burton S. Kaliski, "PORs: Proofs of Retrievability for Large Files," Proc. of CCS '07, pp. 584–597, 2007.

[8] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc.of Asiacrypt '08, Dec. 2008.

[9] K. D. Bowers, A. Juels, and A. Oprea, "Proofs of Retrievability: Theory and Implementation," Cryptology ePrint Archive, Report 2008/175,2008,http://eprint.iacr.org/.

◈◈◈

Improvement in Semantic Security Policy Matching in Service Oriented Architecture

Madhu chauhan & Raghvendra Pratap Singh

Dept. of Computer Science & Engineering, Galgotia College of Engineering and Technology Greater Noida, India

Abstract - There are new security and privacy challenges, mainly related to resource sharing, interoperability and dynamicity among different providers in cloud computing. Policy specification languages address some of these challenges. Today, Service Oriented Architecture is the most relevant technology that facilitates such multi domain cooperation through service discovery, composition and orchestration .Presently a semantic security policy matching between web service and client by means of semantic annotations to WS-Policy documents, allows for matching of security requirements and capabilities based on their actual meaning ,that provide secure environment. But at some aspects it is not secure. We proposed a novel approach in our work, we insert a interceptor between web service and client this interceptor work before policy matching on semantic basis. In interceptor we store IP address of those clients whose policy successfully matched with the web service .When client try to access the web service first it check that its IP is present in interceptor or not ?if present then it directly access the web service otherwise policy matching is started, Due to this process if same client access one more time then there is no need of policy matching again and again.

Keywords - *Web service, WS-Policy , ontology, Semantic matching, Interceptor.*

I. INTRODUCTION

Today's techniques for guaranteeing adequate levels of confidentiality, integrity and authentication can be efficiently adopted by a Cloud provider [1]. Yet, when considering the interaction among different Cloud providers, there is a clear need to establish strong trust relationships on which to build an environment of shared resources, especially with partners unknown beforehand [2].Furthermore, the problem of interoperability among the different security systems and mechanisms used by each cloud provider must be faced. Indeed, cloud computing environments can be seen as multi-domain environments where each domain can use different security, privacy, and trust requirements, potentially leveraging on different mechanisms and protocols.

Web services are broadly used for different tasks such as giving programmatic access to Web applications, providing standardized interfaces for information sources, or enabling the creation of lightweight mashup applications. Furthermore

Web services are the most commonly used technology to implement service-oriented architectures (SOA).

The functional interface in form of input and output messages and their structure can be described using standards such as WSDL. Web service is characterized by its heterogeneity, dynamic, and loose coupling. These characteristics greatly facilitate the application integration based on heterogeneous platform, but along with introducing many distinct security problems at the same time. Security is one of the main reasons that hinder the wide scale deployment of web service [3].

A primary security problem of web service is how to precisely express and match the security policy of each web service entity that may be in different security domain. Here, we categorize the security policy into security requirements and security capabilities of a web service entity (including service requestor and service provider). To determine if a web service is suitable for a particular request, on the one hand, the functional aspects of the service should satisfy the request, on the other hand, the security requirements of the request need to be satisfied by the security capabilities of the service, whose security requirements also need to be satisfied by the security capabilities specified in the request. First, we define a general security ontology based on the abstraction of various existing security-related standards, protocols and notions of web service. And then, we present the definition method and matching algorithm of semantic security policy based on the security ontology. We definitely distinguish security requirements from security capabilities and define them respectively. Secondly the request of service requester sends to the interceptor then interceptor verify the service requester by its IP address.

Consider the fictitious company JassTrade that wants to implement an automated stock trading system [11]. The company maintains a repository of internal and external Web services with their corresponding WS-Policies [4]. The project manager assigns Alice the task to find a suitable service that delivers stock quotes. Alice is told that the service must be highly secure, for two reasons: (i) the stock quotes must be reliable and not modifiable by attackers, and (ii) no third party should be able to see which stocks are monitored by the company, as this could give hints about their secret trading strategy. Alice retrieves two Web services from the repository, that can deliver stock quotes. One is provided by Bob and supports the Basic256 algorithm suite as defined in the WS-Security Policy standard, and signed output messages. The other service is provided by Eve and supports the Basic128 algorithm suite and signed headers. In order to determine if one of the Web services fulfills Alice's requirement, she has to check if her requirements are a subset of the offered capabilities. This operation is called policy matching and requires domain-dependent knowledge to decide if a requirement (e.g. high security) is fulfilled by a number of capabilities (e.g. support for the Basic256 algorithm suite and signed messages). Without background knowledge a policy matching tool can only work with syntactical comparisons, which does not work in our example, where requirements and capabilities are specified on different levels of abstraction.

The rest of the paper is organized as follows. In section II, we present the related work . Section III , we present a motivating example to make a case for semantics in representing the security policy of web services . In section IV,we explain the proposed architecture and its components. Section V provides an overview of matching algorithm. Our conclusions and future work are outlined in section VI.

II. RELATED WORK

Policy-based management has been widely employed in enterprise information system [5] , where policies are often applied to automate network administration tasks. Multiple approaches for policy specification have been proposed that range from formal policy languages that can be directly processed by a computer, to rule-based policy notations, and to representation of policies as entries in a table consisting of multiple attributes. Within the last few years policy management has extended its original scope, going beyond traditional domains. In the Web service domain policies are used to express non-functional service properties: WS-Policy [6] is the specification used in this context. A policy is defined as a collection of alternatives and each alternative is a collection of assertions. Assertions specify characteristics that may be used for service selection such as requirements, capabilities or behaviours of a Web service. Policy assertions can represent both requirements on service requestors and the capabilities of the service itself. Requirements represent a demand on service requestors to follow a particular behaviour; capabilities are the service providers promises of behaviour. Among the others non functional properties of a service, WS-Policy may be used to express security requirements and security capabilities. For example, the use of a specific protocol to protect message integrity is a requirement that a service can impose on requestors.

On the other hand, the adoption of a particular privacy policy when manipulating requestors' data is a service capability. Policy matching in WS-Policy works at a syntactic level: it offers a domain-independent mechanism to find alternatives that are compatible to two policies. Two alternatives are compatible if, for each assertion in an alternative, there is a compatible assertion in the other one. Compatibility of assertions is defined exclusively according to the equality of their *qnames*, without any further mechanism involving their structure and content.

For this reason, several works in the literature [7],[8],[9],[10]have been trying to enhance WS-Policy with semantic annotations. In , WS-Policy assertions refer to policies expressed in OWL: however that work is not focused on policy matching, but on modeling policies as a set of rules, which have to be evaluated using an external rule-based system, requiring reasoning beyond OWL. In policies represented in WS-Policy are enhanced with semantics by using a combination of OWL and SWRL based rules to capture domain concepts and rules. In a lightweight approach to specify semantic annotations in WS-Policy is presented: it combines the syntactic matching with the semantic matching capability provided by OWL.

However, adopting WS-Policy as base language to express policies may limit the flexibility and the expressiveness of the policy itself; for this reason, several other approaches present ontology-based policy languages that are not based on WS-Policy, and are not specifically focused on the security domain. Among the most notably efforts in this domain it is worth citing Kaos, a policy management framework where concepts and policies themselves are represented using OWL, and Rei , a policy framework where OWL is extended with the expression of relations like rolevalue maps, making the language more expressive than the original OWL.

The work by Dumitru et al. combines WS-Policy with semantic Web services modeled with WSMO in two ways : (i) using policies as non-functional properties in WSMO descriptions, and (ii) using WSMO service descriptions as WS-Policy assertion. The second approach is related to ours. However it does not aim at adding semantics to existing policies, but introduces new assertions that need reasoning about WSMO services in order to decide their matching.

Anderson presents with WS-Policy Constraints a language to specify policy assertions with semantics defined by XPath expressions on Web service messages. Such assertions allow policy engines to check compliance of messages without relying on domain-dependent assertions with semantics defined in external standards. One of the drawbacks of WS-Policy Constraints is that it is dependent on the message structure and syntax and thus it is difficult to map between different vocabularies. Another problem is that for matching two policy assertions, it is necessary to check containment for XPath queries. The approach by Anderson is useful for testing message compliance and can be complemented by our proposed semantic annotations to support policy matching.

The computation methods for the similarity between Web services are studied and applied in many aspects.

Google, the Web service search engine, supports similarity search for Web services. Keyword search paradigm is insufficient for Web service search in that the underlying semantics cannot be captured. Motivated by the above fact, the technique to support similarity search for Web services was proposed by. In this approach, similar operations are determined mainly by using the association-rule-mining approach and a hierarchical clustering algorithm for parameter names of Web services. However, the association rules are not very effective when the associations among services are complex, and it is difficult to fully represent the causal relationships implied among them. Further, the reasoning, critical for automatic matchmaking and discovery of Web services, can`t be done straightforwardly. Similarities between Web services are applied to obtaining Web services communities .Proposes the nearest-neighbor approach to obtain the classes for the given services. The similarity measure just considers whether two services are similar, but does not explore how similar they are.

III. A SCENARIO FOR SEMANTICS IN POLICY MATCHING

WS-Policy provides a grammar for representing the nonfunctional attributes of entities in a Web services based XML environment. A policy is defined as a collection of alternatives and an alternative is defined as a collection of assertions. An assertion is used to represent a requirement, capability or a behavior of a Web service. An assertion can have arbitrary number of child assertions and attributes. Policy matching in WS-Policy operates on syntactic level, where pairs of policies are compared for structural and syntactic similarity to and is non-inferable, it is prone to result in policy conflict and get wrong matching result. In order to illustrate our viewpoint, let us consider the following example. Say that the security requirement and capability of a service requestor are as follows:

- *Security requirement*: Encryption is required, and the related data should be encrypted by 3DES;
- *Security capability*: Can supply identity assertion with SAML.

  Providing a service provider can satisfy the functional requirements of this service requestor, and its security policy is:

- *Security requirement*: The identity of the service requester must be authenticated;
- *Security capability*: Can support the encryption standard Xml-Enc.

In the above scenario, if we were to use WS-Policy to represent and match these policies on syntactic level, it would lead to false negative, because the matcher have no additional knowledge about the relationship between '3DES' and 'Xml- Enc', also 'SAML' and 'Identity Authentication'. In fact, the security policies of requestor and provider are compatible to each other. Say that now we create a security ontology to capture the following semantic information:

- Xml-Enc is an encryption standard and supports 3DES to execute symmetric encryption.
- SAML can supply identity assertion to authenticate the requestor.

When this additional domain information along with semantic reasoning is applied, the relationship between seemingly unrelated policies becomes apparent thereby making it possible to match them correctly. This example illustrates how semantic information captured as security ontology for Web Service can improve the quality of policy matching.

IV. SYSTEM ARCHITECTURE

In this section we discuss the overall architecture of the semantic security policy matching framework with interceptor between web service and client.

The basic logical view of the architecture, depicted in Figure1. The novelty is the presence of the *Interceptor*, and that of policy related information flows.

Service requester - Service requester is a client who wants to access the service from service provider. Service requester access service if its security policy match with the service provider security policy and it is also full fill the requirement of interceptor like its IP address present in interceptor.

Policy matching - In this section the security policy of service requester and service provider match on semantic basis. Security policy contains security issues like authentication, integrity, algorithm (encryption), credentials ,protocols etc.

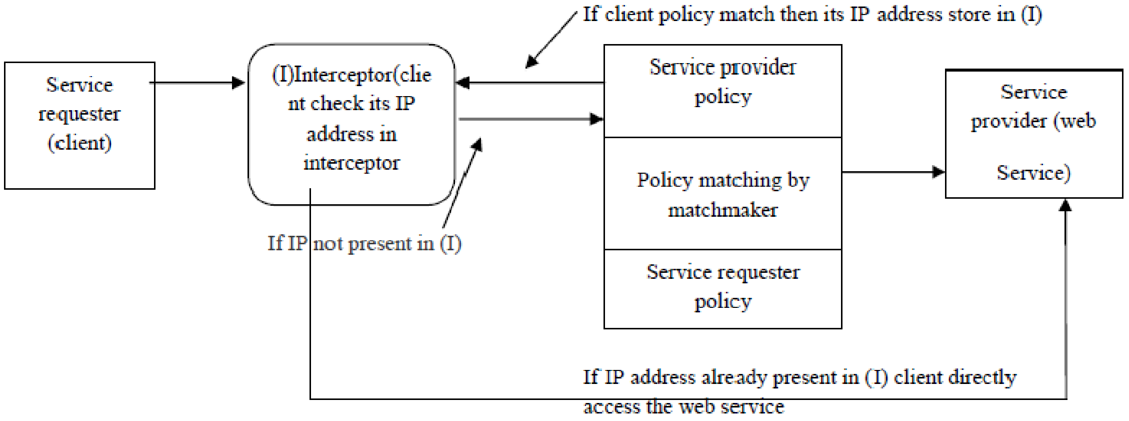

If client policy match then its IP address store in (I)
Service requester (client)
(I)Interceptor(client check its IP address in interceptor
Service provider policy
Policy matching by matchmaker
Service requester policy
Service provider (web Service)
If IP not present in (I)
If IP address already present in (I) client directly access the web service

Fig. 1 : Overall Architecture

Interceptor - In the field of software development, an interceptor pattern is a software design pattern that is used when software systems or frameworks want to offer a way to change, or augment, their usual processing cycle. In interceptor we store IP address of client whose policy sucessfully match with the web service .When client try to access the web service first it check that its IP is present in interceptor or not? if present then it directly access the web service otherwise policy matching is started, Due to this process if same client access one more time then there is no need of policy matching again and again.

Service provider - Service provider present on server and provide service to the client. Service provider provides service to those clients, which full fill the requirement of interceptor and policy matching means those clients whose policy matched already and there IP address store in interceptor.

V. MATCHING ALGORITHM

We provide some details on how the matching algorithm works. The process of making the match between two policies consists in searching a correspondence between requirements and capabilities [12]. Specifically, a) the requirements of a service are compared to the capabilities of the requestor, and b) the capabilities of the service are compared to the requirements of the requestor. In order for the comparison process to have a positive outcome, the two following conditions must hold:

- the capabilities expressed in the service's policy must meet the requirements expressed in the requestor's policy;
- the requirements expressed in the service's policy must be met by the capabilities expressed in the requestor's policy.

The matchmaking process breaks down in two steps: assigning the match level to each requirement-capability pair and assigning the match level between to overall set of requirements and the overall set of capabilities.

As for the first step, each requirement-capability pair can be assigned one out of a scale of four different match levels:

- Perfect Match;
- Close Match;
- Possible Match;
- No Match;

For each requirement, the objective is to find the capability that matches at best. In the second step the overall match between the two policies will be evaluated. The overall match is defined to be the minimum among the individual match levels evaluated in the first step for each requirement capability pair.

In the following a detailed description of the four levels of match is given.

Perfect Match : A perfect match occurs when the requirement and the capability both refer to the same

concept, or to two equivalent concepts. In particular, two cases are possible: a) the requirement and the capability refer to the very same semantic concept. If properties are also specified, their values are identical; b) the requirement and the capability refer to equivalent semantic concepts and, if specified, the properties are identical or equivalent.

Close Match : A close match occurs when the requirement is more general then the capability. Three cases are possible: a) the requirement specifies a more general semantic concept than the capability, i.e., a concept that is higher in the ontological hierarchy; b) the requirement and the capability refer to the same semantic concept, but more details are specified for the capability (using the ”property” construct); c) referring to the security ontology, the requirement is expressed in terms of security objective while the capability is expressed in terms of security concept that supports the specified objective.

Possible Match : A possible match is when the requirement is more specific than the capability. It can be presented as the opposite of the Close Match condition. Three cases are possible: a) the requirement specifies a more specific semantic concept than the capabilities, i.e., lower in the ontological hierarchy. b) the requirement and the capability refer to the same semantic concept, but the requirement is specified more in detail; c) with reference to the security ontology, the requirement is expressed as security concept while the capability is expressed in term of security objective supported by that concept. When the overall matching result gives a possible match level, a further negotiation step is needed to effectively verify the compliance with the requirement.

No Match : A no match occurs when the requirement and the capability have no chance to match. Two cases are possible: a) the requirement and the capability refer to semantic concepts that have no semantic relationship; b) the requirement and the capability refer to the same semantic concept but have a different specifications for their properties.

VI. CONCLUSION AND FUTURE WORK

Different from existing works, in this paper, we try to concern various security aspects of web service and propose a novel approach for more Security and time utilization between service requester(client) and service provider(web service) for this reason we add a interceptor before policy matching of service requester and service provider. In previous work many specification languages and ontologies are used for policy matching and security but in our work we add interceptor for more security and time utilization. Interceptor use according to your requirement in our work, In interceptor we store IP address of those clients whose policy successfully matched with the web service .When client try to access the web service first it check that its IP is present in interceptor or not ?if present then it directly access the web service otherwise policy matching is started, Due to this process if same client access one more time then there is no need of policy matching again and again.

In future we will build a new idea and add it in interceptor because we can use interceptor according to our requirements. Another plan for future is that Interceptor concept will be use in service oriented architecture.

REFERENCES

[1] S. Lakshminarayanan, “Interoperable security standards for web services,” IT Professional, vol. 12, no. 5, pp. 42 –47,Sep/Oct 2010.

[2] G. Peterson, “Don’t Trust. And Verify: A Security Architecture Stack for the Cloud,” IEEE Security & Privacy Magazine, vol. 8, no. 5, pp. 83–86, Sep. 2010.

[3] Anoop S, Theodore W, Karen S. Guide to Secure Web Service. National Institute of Standards and Technology Special Publication 800-95, 2007

[4] S. Speiser, “Semantic Annotations for WS-Policy,” in IEEEInternational Conference on Web Services (ICWS 2010).IEEE, 2010, pp. 449–456.

[5] T. Phan, J. Han, J. Schneider, T. Ebringer, and T. Rogers, “A survey of policy-based management approaches for Service Oriented Systems,” in Software Engineering, 2008. ASWEC2008. 19th Australian Conference on. IEEE, 2008, pp. 392–401.

[6] W3C, “Web services policy 1.5 - framework,” W3C Recommendation, sep 2007. [Online]. Available:http://www.w3.org/TR/ws-policy/.

[7] N. Sriharee, T. Senivongse, K. Verma, and A. Sheth, “On using ws-policy, ontology, and rule reasoning to discover web services,” in Intelligence in Communication Systems, no. May 2004. Springer, 2004, pp. 246–255.

[8] K. Verma, R. Akkiraju, and R. Goodwin, “Semantic matching of Web service policies,” in Semantic Web Policy Workshop (SDWP 2005), 2005.

[9] S. Speiser, “Semantic Annotations for WS-Policy,” in IEEE International Conference on Web Services (ICWS 2010). IEEE, 2010, pp. 449–456.

[10] H. Zheng-qiu, W. Li-fa, H. Zheng, and L. Hai-guang, “Semantic Security Policy for Web Service,” in 2009 IEEE International Symposium on Parallel and Distributed Processingwith Applications. Ieee, 2009, pp. 258–262.

[11] H. Zheng-qiu, W. Li-fa, H. Zheng, and L. Hai-guang, “Semantic Security Policy for Web Service,” in 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications. Ieee, 2009, pp. 258–262.

[12] G. Di Modica, O. Tomarchio” Semantic Security Policy Matching in service oriented architectures” in 2011 IEEE World Congress on Services.IEEE 2011,pp.399-405

◈◈◈

Copyright Protection And Authentication Using Water Marking Scheme

R. Deepthi[1], M.V. Rajesh[2] M. Ravi Kiran[3]

[1&2]Computer Science and Engineering, Aditya Engineering College, Surampalem, A.P. India
[3]Wipro Technologies, Banglore, India

Abstract - In the recent years, Digital Watermarking is used for copyright protection and authentication. In this paper, a new Dual Watermarking Scheme based on DWT-SVD is presented to improve the robustness and protection. Both Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) have been used as a mathematical tool to embed watermark in the image. In the proposed technique, two watermarks are embedded in the host image. First watermark is called primary watermark, which is a gray scale digital image. Second watermark is called secondary watermark which is a gray scale meaningful logo instead of randomly generated Gaussian noise type watermark. The secondary watermark is embedded into primary watermark and the resultant watermarked image is used as watermark for the host image. A reliable watermark extraction scheme is developed for the extraction of the primary as well as secondary watermark from the distorted image. Experimental evaluation demonstrates that the proposed scheme is able to withstand a variety of attacks. The secondary watermark is easy to detect in all the cases but sometimes primary one is severely distorted.

Keywords - *DWT, Dual Water Marking Scheme, Single Valued Composition, Embedded System*

I. INTRODUCTION

In the recent years, the advance of editing software and the popularity of the Internet, illegal operations, such as duplication, modification, forgery and others in digital media, have become easy, fast and difficult to prevent. Therefore, the protection of the intellectual property rights of digital media has become an urgent matter. The one of the solution for this is Digital Watermarking. It can be used for tracking the images that were illegally distributed. Watermarking, when complemented with encryption, can serve for many purposes, such as copyright protection, broadcast monitoring and data authentication. There are many watermarking algorithms proposed in literature. Some of them operate either in the spatial domain or in the frequency domain [1, 2, 3, 4]. In the recent years, a new transform is introduced for watermarking namely Singular Value Decomposition (SVD) [5, 6]. In SVD domain, a common approach is to modify the singular values by the singular values of a visual watermark. SVD is one of the most useful tools of linear algebra with several applications in image and signal processing. To improve the robustness and protection, the concept of dual watermarking [7, 8, 9] is introduced by Swanson *et al.*[7]. In all these scheme authors embed one visible and one invisible watermark. When the visible watermarked image is in question, the invisible watermark can provide rightful ownership. In this paper, we have presented a novel dual watermarking scheme to improve rightful ownership, protection and robustness. In the proposed scheme, both the watermarks are invisible. For embedding, both the images are transformed into wavelet domain. Further SVD transform is performed on both the images and sum up the singular values to find the new singular values.

II. SINGULAR VALUE DECOMPOSITION

Let A be a general real(complex) matrix of order $m \times n$. The singular value decomposition (SVD) of A is the factorization $A = U * S * V T$ (1) where U and V are *orthogonal(unitary)* and $S = diag(\sigma 1, \sigma 2, ..., \sigma r)$, where σi, $i = 1(1)r$ are the singular values of the matrix A with $r = min(m, n)$ and satisfying $\sigma 1 \geq \sigma 2 \geq ... \geq \sigma r$ (2) The first r columns of V the *right singular vectors* and the first r columns of U the *left singular vectors*. Use of SVD in digital image processing has some advantages. First, the size of the matrices from SVD transformation is not fixed. It can be a square or a rectangle. Secondly, singular values in a digital image are less affected if general image processing is performed. Finally, singular values contain intrinsic algebraic image properties.

III. WATERMARK EXTRACTION

The objective of the watermark extraction is to obtain the estimate of the watermark. For watermark extraction from watermarked image, both watermark and host images are needed. But in practical applications (in some cases), it is not possible to access the original image. We consider those applications where original host image is available. The extraction process is formulated as follows:

IV. EXTRACTING PRIMARY WATERMARK:

The extraction technique for primary watermark is given as follows. 1. Perform *l*-level wavelet transform on the host as well as watermarked image. Let us denote each sub-band with *f l θ* and ˜ *f l θ* for host and watermarked image respectively where *θ* ∈ *{* LL, LH, HL, HH *}* represents the orientation and *l* gives the level of the orientation. 2. The detail and approximation sub-images of the host as well as watermarked image is segmented into nonoverlapping rectangles of size *M*1 × *N*1 using ZIGZAG sequence. We denote these segmented rectangles by *f l,m θ* and ˜*f l,m θ* for host and watermarked images respectively, where *m* = 1*,* 2*, ...,* 2*M*+*N*−2*l*. 3. Perform SVD transform on all non overlapping rectangles of both the images, 4. Extract singular values of primary watermark from all non overlapping rectangles as *σm*5. Obtain all estimate of primary watermark as 6. Select the primary watermark estimate as detected watermark which has the greatest correlation coefficient and it is denoted by _*W*1*det*.

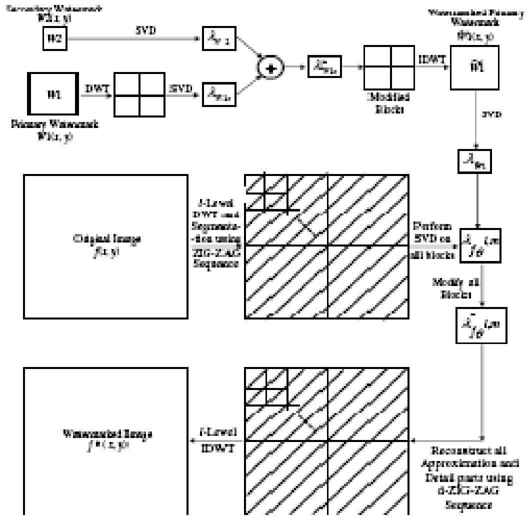


Fig. 1 : Proposed Embedding Algorithm.

V. RESULTS

We demonstrate our proposed algorithm using MATLAB. We have taken 8-bit gray scale Baboon image as original (host) image of size 512 × 512 and for primary and secondary watermark; we have used 8-bit gray scale lena image and IIT logo of sizes 128 × 128 and 64 × 64 respectively. For embedding the watermarked primary watermark into the host image, we have used 2-level of decomposition using Daubechies filter bank. We have embedded watermarked primary watermark 16 times in the host image. In the extraction, we are only selecting an image whose correlation coefficient is the greatest among all. This image is being used as the extracted primary watermark and this is used for extracting secondary watermark. In figures 2 and 3 all original, watermarked images and extracted watermarks are shown. To investigate the robustness of the algorithm, the watermarked image is attacked by Average and Mean Filtering, JPEG and JPEG2000 compression, Gaussian noise addition, Resize, Rotation and Cropping. After these attacks on the watermarked image, we compare the extracted watermarks with the original one. The watermarked image quality is measured using PSNR (Peak Signal to Noise Ratio). Watermarked Baboon and Lena images are having PSNR values 43.2093 and 41.9187 respectively. Extracted watermark after applying 11 × 11 averaging and median filtering are shown in figures 4 and 5 respectively. It can be observed that after applying these filters, images are very much degraded and huge data is lost but extracted primary watermark is still recognizable. The quality of extracted secondary watermark is not so good but still it is recognizable. Results for additive Gaussian noise are shown in figure 6 and it is very clear from the figure that this algorithm is stand with this attack also. In figure 6, watermarks extracted form 70% additive Gaussian noise are shown. Another most common manipulation in digital image is image compression. To check the robustness against image compression, the watermarked image is tested by JPEG compression attacks. Results for JPEG (50:1) and JPEG 2000 (75:1) are shown in figures 7 and 8 respectively. In figures 9, 10 and 11, results on resizing, rotation and cropping are given. First we reduce the size of watermark image to 128 × 128 and again expand to 512 × 512. For rotation, watermarked image is rotated by 30◦ and extract both the watermarks as shown in figure 10. For cropping, 50% area remaining attacked and extracted watermark images are shown in figure 11. To verify the presence of watermark, different measures can be used to show the similarity between the original and the extracted singular values. In our proposed algorithm,

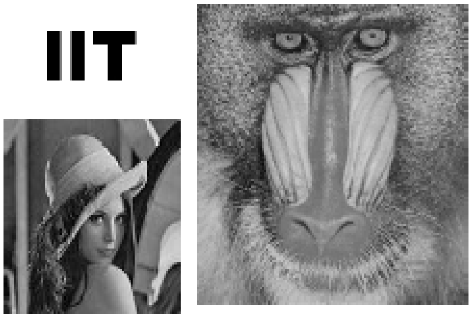


Fig. 2 : Original Images

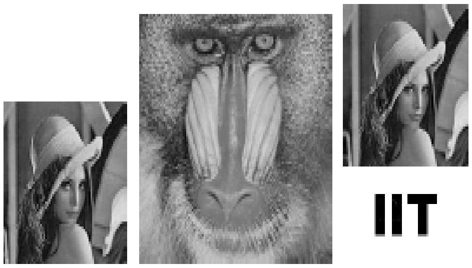


Fig. 3 : Watermarked and Extracted Watermark Images

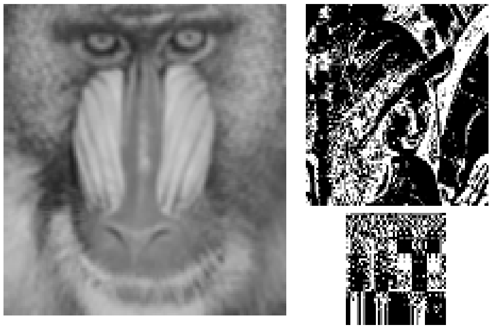


Fig. 4 : Average filtering attack

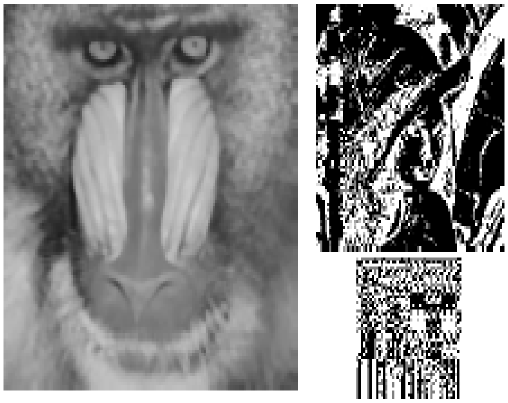


Fig. 5 : Median filtering attack

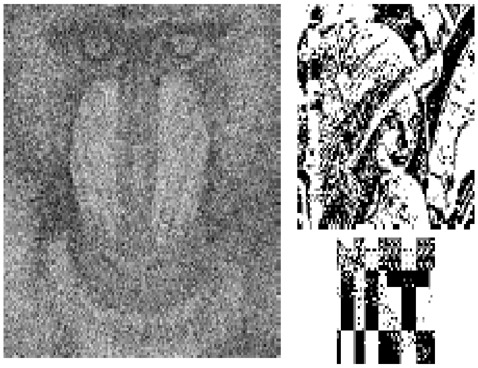


Fig. 6 : Additive Gaussian Noise attack
(Adding 70% noise)

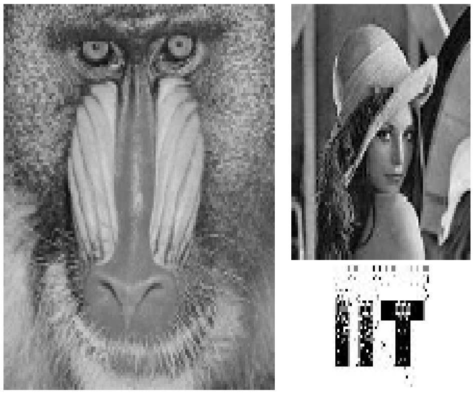


Fig.7 : JPEG Compression (50:1)

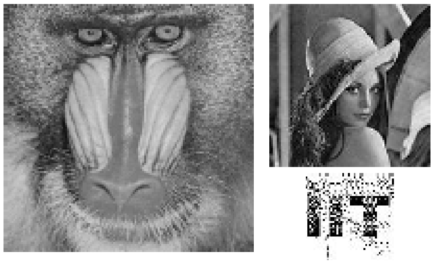


Fig. 8 : JPEG2000 Compression (75:1)

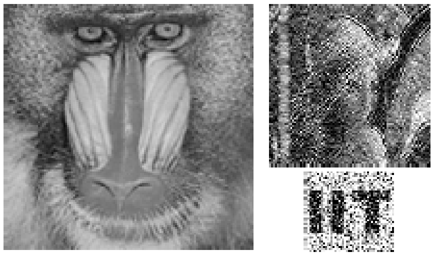


Fig. 9 : Resizing attack

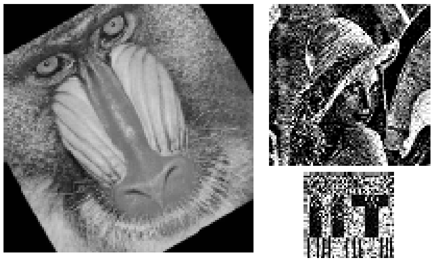


Fig. 10 : Rotation attack (30° rotation)

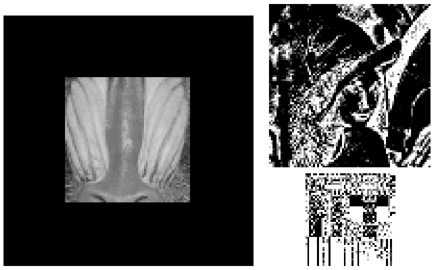


Fig. 11 : Cropping attack (50% area remaining)

VI. CONCLUSIONS

A dual watermarking scheme is presented in which the watermarks are either a gray scale image or visually meaningful gray scale logo instead of a noise type Gaussian sequence. For the extraction of watermark, a reliable watermark extraction scheme is constructed for both primary and secondary watermark. Robustness of this method is carried out by variety of attacks. The algorithm uses DWT variety of time-frequency decomposition for images, and modifies its singular value matrix with a watermark matrix before reconstituting the signal. The left and right singular vectors must be available at the receiver. The dual watermarking algorithm presented in this paper is useful for extraction of ownership of digital images and extraction of manipulations in the images.

REFERENCES

[1] X. Xia, C.G. Boncelet and G. R. Arce, "A multiresolution watermark for digital images," Proceedings of IEEE Int. Conf. Image Processing, Santa Barbara, CA, vol. 3, 1997, pp. 548-551.

[2] S.H. Wang and Y.P. Lin, "Wavelet tree quantization for copyright protection watermarking," IEEE Transcations on Image Processing, vol. 13, No. 2, 2004, pp. 154-165.

[3] Z. Dawei, C. Guanrong and L. Wenbo, "A chaosbased robust wavelet-domain watermarking algorithm," J. Chaos Solitons Fractals, vol. 22, 2004, pp. 47-54.

[4] P. Meerwald and A. Uhl, "A survey ofWavelet-Domain Watermarking Algorithms," Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents III, San Jose, CA, USA, vol. 4314, 2001.

[5] R. Liu and T. Tan, "An SVD-Based Watermarking Scheme for Protecting Rightful Ownership," IEEE Transactions on Multimedia, vol. 4, no. 1, 2002, pp. 121- 128.

[6] E. Ganic and A. M. Eskicioglu, "Robust Embedding of Visual Watermarks Using DWT-SVD," Journal of Electronic Imaging, vol. 14, no. 4, 2005.

[7] M. D.Swanson, B. B. Zhu and A. H. Tewfik, "Robust audio watermarking using perceptual masking," Signal processing Elesvier, vol. 66, 1998, pp. 337-355.

[8] S. P. Mohanty, K. R. Ramakrishnan and M. Kankanhalli, "A Dual Watermarking Technique for Images," Proceedings of the seventh ACM international conference on Multimedia, Orlando, Florida, USA, 1999, pp. 49-51.

[9] Y. Hu, S. Kwong and J. Huang, "Using Invisible Watermarks to Protect Visibly Watermarked Images," Proceedings of the International Symposium on Circuits and Systems, vol. 5, 2004, pp. 584-587.

[10] B. B. Zhu, M. D. Swanson and A. H. Tewfik, "When seeing isnt believing," IEEE Signal Processing Magazine, vol. 21, 2004 pp. 40-49. 531Authorized licensed use limited to: Jawaharlal Nehru Technological University. Downloaded on July 18, 2009 at 09:00 from IEEE Xplore. Restrictions apply.

Efficient Distribution Based Method in Decision Trees for Uncertain Data

S.Ramachandra[1], M.V.Jagannatha Reddy[2] & M. Roshini[3]

[1&2]CSE Dept., MITS, Madanapalle, [3]CSE Dept., MTIET, Palamaneru

Abstract - The conventional decision tree classifier will work with known data or certain data. We extend such classifiers to handle with unknown data or uncertain information. The uncertainty may arise in many applications during information collection process. Examples quantization/equipment errors, datastaleness and multiple repeated measurements.With uncertainty the value of data item is often represented not only by single value but it represented by multiple values forming a probability distribution. We discover that the accuracy of the decision tree classifier can be much improved when the complete information of the data item rather than considering the abstracting uncertain data by Statistical derivation such as median and median. We enhance such classical decision tree building algorithms to handle data tuples with uncertain values. The experiments have been conducted to show that the resulting classifier is more accurate more than when averages are used. Here processing the pdf is more costly than processing the single values. We propose a series of pruning techniques that can have drastic improve construction efficiency.

Keywords *- uncertain data, decision trees, classification, data mining.*

I. INTRODUCTION

Classification is a classical problem in machine learning and data mining [1].there are some training data tuple and each having a class and it is represented by a feature vector. Here the constraint is to build an model to predicts the class label of an unknown test tuple based on the feature vector of an tuples. Among these one of the most power full classification models are decision tree model. The decision trees are very popular because they are easy to understand.Many algorithms, such as ID3 [2] and C4.5 [3], have been developed forbuilding decisiontree. These algorithms are widely brought and used in a wide range ofapplications such as image recognition, medical diagnosis [4], and credit rating of loan applicants, scientific tests, fraud detection, and target marketing. Intraditional decision tree classification

In conventional decision tree classification an attribute of a tuple is categorical or numerical.But the data uncertainty in common in many applications. The value of a feature/attribute is thus bestcaptured not by a single point value, but by a range of values giving rise to a probability distribution. A simple way to handle data uncertainty is to abstract probability distributions by summary statistics such as means and variances. We call this approachaveraging. An another way is to handle the complete information carried by the probability distribution to build a decision tree is called efficient distribution based .In this paper we study how to constructing decision tree classifiers on data with uncertain numerical attributes. Ourgoal are 1) To develop an algorithm for constructing decision trees for uncertain data by using the efficient Distribution-based approach,2)to verify whether the Efficient distribution based approach could lead to a higher classification accuracy than with averaging approach, and 3)to build a theoretical foundation on pruning techniques and that can greatly improve the computational efficiency of the Efficiency Distribution-based algorithms.

A. Measurement Errors

Due to measurements errors data may often imprecise. An example A new tympanic thermometer is analyzed and tested experimentally. An electrically calibrated pyro electric detector of special configuration is employed to determine a person’s body temperature. An energy-storage, power-supply-isolated capacitor is used as the electrical heating reference. The new thermometer design has accuracy within ∓0.1 °C with a 90% confidence and is immune to ambient temperature, detector aging, and parameter variations. An equivalent-circuit model is established in the analysis to account for the heat exchanges among the tympanum, the surroundings, and the detector as well as for the electro thermal behavior of the detector. The model provides effective simulation of the thermometer with PSPICE. Critical parameters governing the accuracy and the limitation of the tympanic thermometer are also pointed out by the simulation.

B. Data Staleness

Consider some applications the data values and the recorded information is always not stable, it is always stale.one example is vehicle tracking system. Where it can measure the movements of vehicles [5], from which the uncertain data may arises. A typical uncertainty model requires knowledge about the moving speed of the device and whether its movement is restricted or unrestricted. A 2D probability density function is defined over a bounded region to model such uncertainty.

C. Repeated Measurements

Perhaps the most common source of uncertainty comes from repeated measurements. For example, a patient's body temperature could be take multiple times during a day; an anemometer could record wind speed once every minute; the space shuttle has a large number of heat sensors installed all over its surface. When we inquire about a patient's temperature, or wind speed or the temperature of a certain section of the shuttle, which values shall we use? Or, would it be better to utilize all the information by considering the distribution given by the collected data values?

As a more elaborate example, consider the"BreastCancer"data set reported in [6]. This data set contains a number of tuples. Each tuple corresponds to a microscopic image of stained cell nuclei. A typical image contains 10-40 nuclei. One of the features extracted from each image is the average radius of nuclei. We remark that such a radius measure contains safe few sources of uncertainty: 1)an average is taken from a large number of nuclei from an image, 2) the radius of an (irregularly shaped) nucleus is obtained by averaging the length of the radial line segments defined by the centroid of the nucleus and a large number of sample points on then nucleus' perimeter, and 3) a nucleus' perimeter was outlined by a user over a fuzzy 2D image. From points 1 and 2, we see that a radius is computed from a large number of measure-mints with a wide range of values. The source data points thus form interesting distributions. From point 3, the fuzziness of the 2D image can be modeled by allowing a radius measure be represented by a range instead of a concrete point value.

From the above examples we see that in many applications, information cannot be ideally represented by point data. More often, a value is best captured by a range possibly with a pdf.Our concept of uncertainty refers to such ranges of values. Again, our goal is to investigate how decision trees are built trees are built over uncertain data. Our contribution includes the following:

1. To device a basic algorithm for building decision trees to an uncertain data sets,
2. A study comparing the classification accuracy achieved by the Averaging approach and the Efficient-Distribution-based method,
3. A Set of mathematical theorems that allows the significant pruning of the large search space of the best split point determination.
4. Finally a performance analysis on the various algorithms through a set of experiments.

II. PROBLEM DEFINITION

In this section defines the problem of decision tree classification on behalf of uncertain data. Let us first see the conventional decision tree breafly.Later we discuss how data tuples handles the uncertainty.

A. Traditional Decision Trees

In our model, a dataset consists of d training tuples, $\{t_1,t_2,\dots.t_d\}$, and k numerical feature attributes,$A_1\dots.$ A_k. The domain of attribute A_j is dom(Aj).Each tuple ti is associated with a feature vector Vi = (vi.1, $v_{i.2},\dots\dots$ $v_{i.k}$) and a class label c_i, where $v_{i,j}$εdom(A_j) and ci ε C, the set of all class labels. The classification problem is to construct a model M that maps each feature vector ($v_{x,1}\dots..$ $v_{x,k}$) to a probability distribution P_x on C such that given a test tuple t_0 = ($v_{0,1}\dots\dots$ $v_{0,k}$, c_0), P0 =M($v_{0,1}\dots..$ $v_{0,k}$) predicts the class label c0 with high accuracy. We say that P0 predicts c0 if c_0 = argmaxc C $P_{0(c)}$.

In this paper we study binary decision trees with tests on numerical attributes. Each internal node n of a decision tree is associated with an attribute A_{jn} and a split point z_nεdom(A_{jn}), giving a binary test $v_0,j_n \leq z_n$. An internal node has exactly 2 children, which are labeled "left" and "right", respectively. Each leaf node m in the decision tree is associated with a discrete probability distribution Pm over C. For each c εC, Pm(c) gives a probability reflecting how likely a tuple assigned to leaf node m would have a class label of c.

To determine the class label of a given test tuple t0 = ($v_{0,1}\dots..$ $v_{0,k}$), we traverse the tree starting from the root node until a leaf node is reached. When we visit an internal node n, we execute the test v0;jn _ z_n and proceed to the left child or the right child accordingly. Eventually, we reach a leaf node m. The probability distribution Pm associated with m gives the probabilities that t0 belongs to each class label c εC. For a single result, we return the class label c εC that maximizes Pm(c).

B. Handling Uncertainty Information

Under our uncertainty model, a feature value is represented not by a single value, $v_{i,j}$, but by a pdf, $f_{i;j}$. For practical reasons, we assume that $f_{i;j}$ is non-zero only within a bounded interval [$a_{i;j}$; $b_{i;j}$]. (We will briefly discuss how our methods can be extended to handle pdf's with unbounded domains in Section VII-C.) A pdf$f_{i;j}$could be programmed analytically if it can be specified in closed form. More typically, it would be implemented numerically by storing a set of s sample points x ε[$a_{i;j}$,$b_{i;j}$]with the associated value $f_{i;j}(x)$, effectively approximating $f_{i;j}$by a discrete distribution with s possible values. We adopt this numerical approach for the rest this paper. With this representation, the amount of information available is exploded by a factor of s. hopefully; the richer information allows us to build a better classification model. On the down side, processing large numbers of sample points is much more costly. In this paper we show that accuracy can be improved by considering uncertainty information. We also propose pruning strategies that can greatly reduce the computational effort.

For each internal node n (including the root node), to determine Φ_n(c; t_x,w_x), we first check the attribute $A_{jn}$and split point $z_n$of node n. Since the pdf of $t_x$under attribute A_{jn}spans the interval [$a_{x;jn}$,$b_{x;jn}$], we compute the "left" probability $p_L = \int_{ax,jn}^{Zn} fx$,jn(t) dt (or p_L = 0 in case z_n<a_x;j_n) and the "right" probability p_R = 1 - p_L. Then, we split t_x into 2 fractional tuples t_L and t_R. Tuples t_L and t_R inherit the class label of t_x as well as the pdf's of tx for all attributes except Ajn.The tuple t_L is assigned a weight of $w_L = w_x . p_L$ and its pdf for A_{jn} is given by

$$F_{L,jn}(x) = \begin{cases} f_{x,jn}(x)/w_L & \text{if } x \ \varepsilon \ [a_{x,jn},z_n] \\ 0 & \text{otherwise} \end{cases}$$

the tuple tR is assigned a weight and pdf analogously. We define $\Phi_n(C;t_j,w_x)=P_L.\Phi_{nL}(c;t_L,w_L)+p_R.\Phi_{nR}(C;t_R.w_R)$ where n_L and n_R are the left child and the right child of node respectively

For every leaf node m, recall that it is associated with a probability distribution P_m over C. We define Φ_m(c; t_x;w_x) = w_x _P_m(c). Finally, for each class c, let P(c) = Φ_r(c; t_0, 1.0), where r is the root node of the decision tree. Obtained this way, each probability P(c) indicates how likely it is that the test tuple t_0 has class label c. These computations are illustrated in Figure 1, which shows a test tuple t_0 with onefeature whose pdf has the domain [-2:5, 2]. It has a weight of 1.0 and is first tested against the root node of the decision tree. Based on the split point -1, we find that p_L = 0.3 and p_R = 0.7. So, t_0 is split into two tuples t_L and t_R with weights w_L = 0.3 and w_R = 0.7. The tuple t_L inherits the pdf from t_0 over the sub domain [-2.5,-1], normalized by multiplying by a factor of 1=w_L. Tuple t_R inherits the pdf from t_0 in a= similar fashion. These tuples are then recursively tested down the tree until the leaf nodes are reached. The weight distributed in such a way down to each leaf node is then multiplied with the probability of each class label at that leaf node. These are finally summed up to give the probability distribution (over the class labels) for t_0, giving P(A) = 0:59; P(B) = 0:41.

If a single class label is desired as the result, we select the class label with the highest probability as the final answer. in the example in Figure 1, the test tuple is thus classified as class "A" when a single result is desired.

The most challenging task is to construct a decision tree based on tuples with uncertain values. it involves finding a good testing attribute $A_{jn}$and a good split point $z_n$for each internal node n, as well as an appropriate probability distribution P_m over C for each leaf node m. we describe algorithms for constructing such trees in the next section.

III. ALGORITHMS

Here we discuss two approaches for handling uncertain data. the first approach, called "averaging", transforms an uncertain dataset to a point-valued one by replacing each pdf with its mean value. More specifically, for each tuple $t_i$and attribute $A_{j,}$ we take the mean value[1] $v_{i,j}= \int_{ai,j}^{bi,j} xfi,j(x)dx$ as its representative value. The feature vector of t_i is thus transformed to ($v_{i,1}$,.........,$v_{i,k}$).a A decision tree can then be built by applying conventional tree construction algorithm.

To exploit the full information carried by the pdf's our second approach, called "Distribution-based", considers all the sample points that constitute each pdf. The challenges here is that a training tuple can now "pass" a test at a tree node probabilistically when its pdf properly contains the split point of the test, also, a slight change of the split point modifies that probability, potentially altering the tree structure. We present details of the tree-construction algorithms under the two approaches in the following subsections.

Table I – Examples Tuples

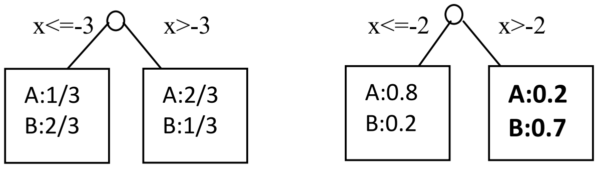
Tuple	Class	Mean	Probability distribution				
			-5	-1.0	0.0	+1.0	+5
1	X	+3.0		8/11			3/11
2	X	-3.0	1/9	8/9			
3	X	+3.0		5/8		1/8	2/8
4	Y	-3.0	5/19	1/19		13/19	
5	Y	+3.0			1/35	30/35	4/35
6	Y	-3.0	3/11			8/11	

A. *Averaging*

A straight-forward way to deal with the uncertain information is to replace eachpdf with its expected value, thus effective converting the data to tuples to point-valued tuples. This reduces the problem back to that for point-valued data, hence traditional decision tree algorithms such as ID3 and C4.5 [3] can be reused. We call this approach AVG (for averaging). we use an algorithm based on C4.5. Here is a brief description.

AVG is a greedy algorithm that builds a tree top-down. When processing a node, we examine a set of tuples S. The algorithm starts with the root node and with S being the set of all training tuples. At each node n, we first check if all the tuples in S have the same class label c. if so, we make n a leaf node and set $P_n(c) = 1, P_n(c') = 0$ $c' \neq c$. Otherwise you select an attribute $A_{jn}$and split point $z_n$and divide the tuples into two subsets: "left" and "right " all tuples with $v_{i,jn} <= z_n$are put in the "left" subset L; the rest go to the "right" subset R. If either L or R is empty (even after exhausting all possible choices of $A_{jn}$and z_n),it is possible to use the available attributes to further discern the tuples in S. If neither L nor R is empty, we make n an interval node and create child nodes for it. We recursively invoke the algorithm on the "left" child and the "right" child, passing to them the sets L and R, respectively.

To build a good decision tree, the choice of $A_{jn}$and $z_n$is crucial. At this point, we may assume that the selection is performed by the black box algorithm Best Split, which takes a set of tuples as parameter, and returns the best choice of attribute and Split point for this tuples. We will examine this Black Box in details. Typically, BestSplitis designed to select the attribute and split the point that minimizes the degree of dispersion. The minimization is taken over the set of all possible attributes A_j(j=1 ...k), considering all possible split points in dom(A_j). Given a set S = {t_1 ,..............t_m} of m tuples with point values there are only m-1 ways to partition S into two non-empty L and R sets. for each attribute A_j, the split points to consider are given by set of values of tuples under attribute A_j, i.e., {$v_{1,j}$,.....$v_{m,j}$}..Among this values, all but the largest one gives valid split points,(the largest one gives an empty R set, so invalid.)



(a) Averaging (b) Distribution-based

Fig. 2 : Decision tree built from example tuples in table 1

for each point of the (m-1)k combinations of attributes (A_j) and split points (z), we divide the set S into the "left" and "right" subsets L and R. we then compute the entropy for each such combination:

$$H(z,A_j) = \sum_{X=L,R} \frac{|X|}{|S|} \left(\sum_{c \in C} -pc/x \log_2 Pc/x\right) \quad (1)$$

where pc=X is the fraction of tuples in X that are labelled c. We take the pair of attribute Aj^* and split point z^* that minimises H(z;Aj) and assign to node n the attribute Aj with split point z*.[3]

The resulting decision tree is shown in Figure 2(a). Since the left subset has 2 tuples of class B and 1 tuple of class A, the left leaf node L has the probability distribution PL(A) = 1=3 and PL(B) = 2=3 over the class labels. The probability distribution of class labels in the right leaf node R is determined analogously. Now, if we use the 6 tuples in Table I as test tuples4 and use this decision tree to classify them, we would classify tuples 2, 4, 6 as class “B” (the most likely class label in L) and hence misclassify tuple 2. We would classify tuples 1, 3, 5 as class “A”, thus getting the class label of 5 wrong. The accuracy is 2=3.

3) To alleviate the problem of over-fitting, we apply the techniques of pre-printing and post-pruning (see [9], [3] for details). 4 In practice and in the following experiments, disjoint training sets and testing sets are used. In this hand-crafted example, however, we use the same tuples for both training and testing just for illustration.

B. *Efficient Distribution-Based*

For uncertain data, we adopt the same decision tree building framework as described above for handling point data. After an attribute Ajn and a split point zn has been chosen for a node n, we have to split the set of tuples S into two subsets L and R. The major difference from the point-data case lies in the way the set S is split. Recall that the pdf of a tuple t_i 2 S under attribute A_{jn} spans the interval [$a_{i,jn}$; $b_{i,jn}$]. If $b_{i;jn}$ _ z_n, the pdf of t_i lies completely on the left of the split point and thus ti is assigned to L. Similarly, we assign ti to R if $z_n < a_{i,jn}$. If the pdf properly contains the split point, i.e., $a_{i,jn} <= z_n < b_{i,jn}$, we split t_i into two fractional tuples $t_L$and t_R in the same way as described in Section III-B and add them to L and R, respectively. We call this algorithm UDT (for Uncertain Decision Tree).

Again, the key to building a good decision tree is a good choice of an attribute A_{jn} and a split point z_n for each node n. With uncertain data, however, the number of choices of a split point given an attribute is not limited to m - 1 point values. This is because a tuple t_i’spdf spans a continuous range [$a_{i,j}$; $b_{i,j}$]. Moving the split point from ai;j to bi;j continuouslychanges the probability $p_L = \int_{ai,jn}^{zn} fi,jn(x)dx$ (and likewisefor p_R). This changes the fractional tuples t_L and t_R, andthus changes the resulting tree. If we model a pdf by s samplevalues, we are approximating the pdf by a discrete distributionof s points. In this case, as the split point moves from oneend-point $a_{i;j}$ to another end-point $b_{i;j}$ of the interval, theprobability p_L changes in s steps. With m tuples, there arein total ms sample points. So, there are at most ms 1possible split points to consider. Considering all k attributes,to determine the best (attribute, split-point) pair thus requireus to examine k(ms- 1) combinations of attributes and splitpoints.Comparing to AVG, UDT is s time more expensive.

Let us re-examine the example tuples in Table I to see how the distribution-based algorithm can improve classification accuracy. By taking into account the probability distribution, UDT builds the tree shown in Figure 3 before pre-pruning and post-pruning are applied. This tree is much more elaborate than the tree shown in Figure 2(a), because we are using more information and hence there are more choices of split points. The tree in Figure 3 turns out to have a 100% classification accuracy! After post-pruning, we get the tree in Figure 2(b). Now, let us use the 6 tuples in Table I as testing tuples4 to test the tree in Figure 2(b). For instance, the classification result of tuple 3 gives P(A) = 5/8*0:80 + 3/8 * 0:212 = 0:5795 and P(B) = 5/8*0:20+3/8*0:788 = 0:4205. Since the probability for "A" is higher, we conclude that tuple 3 belongs to class "A". All the other tuples are handled similarly, using the label of the highest probability as the final classification result. It turns out that all 6 tuples are classified correctly. This hand-crafted example thus illustrates that by considering probability distributions rather than just expected values, we can potentially build a more accurate decision tree.

C. Experiments on Accuracy

To explore the potential of achieving a higher classification accuracy by considering data uncertainty, we have implemented AVG and UDT and applied them to 10 real data sets (see Table II) taken from the UCI Machine Learning Repository. These datasets are chosen because they contain mostly numerical attributes obtained from measurements. For the purpose of our experiments, classifiers are built on the numerical attributes and their "class label" attributes. Some data sets are already divided into "training" and "testing" tuples. For those that are not, we use 10-fold cross validation to measure the accuracy.

The first data set contains 640 tuples, each representing an utterance of Japanese vowels by one of the 9 participating male speakers. Each tuple contains 12 numerical attributes, which are LPC (Linear Predictive Coding) coefficients. These coefficients reflect important features of speech sound. Each attribute value consists of 7–29 samples of LPC coefficients collected over time.

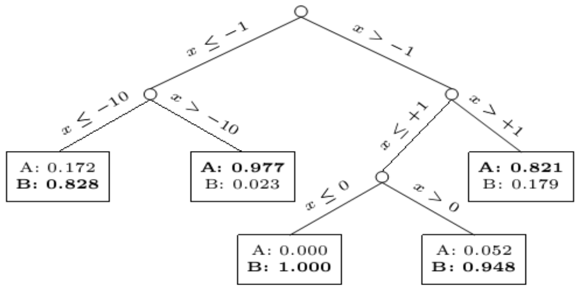


Fig. 3 : Example decision tree before post-pruning

TABLE II
SELECTED DATASETS FROM THE UCI MACHINE LEARNING REPOSITORY

Data Set	Training Tuples	No. of Attributes	No. of Classes	Test Tuples
JapaneseVowel	270	12	9	370
PenDigits	7494	16	10	3498
PageBlock	5473	10	5	10-fold
Satellite	4435	36	6	2000
Segment	2310	14	7	10-fold
Vehicle	846	18	4	10-fold
BreastCancer	569	30	2	10-fold
Ionosphere	351	32	2	10-fold
Glass	214	9	6	10-fold
Iris	150	4	3	10-fold

The other 9 data set contains point-valued without uncertainty. We model uncertainty information by fitting appropriate error models on to the point data. For each tuple $t_i$and for each attribute A_j , the point value $v_{i;j}$ reported in a dataset is used as the mean of a pdf$f_{i;j}$, defined over an interval [$a_{i,j}$;$b_{i,j}$]. The range of values for A_j (over the whole data set) is noted and the width of [$a_{i,j}$,$b_{i,j}$] is set to w. |A|, where |A| denotes the width of the range for A_j and w is a controlled parameter. To generate the pdf$f_{i;j}$, we consider two options. The first isuniform distribution, which implies $f_{i;j}$ (x) = ($b_{i;j}$-$a_{i;j}$)$^{-1}$. The other option is Gaussian distribution 5 , for which we use 1/4($b_{i,j}$,$a_{i,j}$) as the standard deviation. In both cases, the pdf is generated using s sample points in the interval. Using this method (with controllable parameters w and s, and a choice of Gaussian vs. uniform distribution), we transform a data set with point values into one with uncertainty.

The results of applying AVG and UDT to the 10 datasets are shown in Table III. As we have explained, under our uncertainty model, classification results are probabilistic. Following [3], we take the class label of the highest probability as the final class label. We have also run an experiment using C4.5 [3] with the information gain criterion. The resulting accuracies are very similar to those of AVG and are hence omitted. In the experiments, each pdf is represented by 100 sample points (i.e., s = 100), except for the "Japanese Vowel" data set. We have repeated the experiments using various values for w. For most of the datasets, Gaussian distribution is assumed as the error model. Since the data sets "Pen Digits", "Vehicle" and "Satellite" have integer domains, we suspected that they are highly influenced by quantisation noise. So, we have also tried uniform distribution on these three datasets, in addition to Gaussian[6] For the "Japanese Vowel" data set, we use the uncertainty given by the raw data (7–29 samples) to model the pdf.

From the table, we see that UDT builds more accurate decision trees than AVG does for different distributions over a wide range of w. For the first data set, whose pdf is modelled from the raw data samples, the accuracy is improved from 81.89% to 87.30%; i.e., the error rate is reduced from 18.11% down to 12.70%, which is a very substantial improvement. Only in a few cases (marked with "#" in the table) does UDT give slightly worse accuracies than AVG. To better show the best potential improvement, we have identified the best cases (marked with "*") and repeated them in the third column of the table. Comparing the second and third columns of Table III, we see that UDT can potentially build remarkably more accurate decision trees than AVG. For example, for the "Iris" data set, the accuracy improves from 94.73% to 96.13%. (Thus, the error rate is reduced from 5.27% down to 3.87%.)

Using Gaussian distribution gives better accuracies in 8 out of the 9 datasets where we have modelled the error distributions as described above. This suggests that the effects of random noise dominate quantisation noise. The exception is "Pen Digits". As we have pointed out, this dataset contains integral attributes, which is likely subject to quantisation noise. By considering a uniform distribution as the error model, such noise is taken into consideration, resulting in a high classification accuracy.

TABLE III

ACCURACY IMPROVEMENT BY CONSIDERING THE DISTRIBUTION

Data Set	AVG	Best Case	UDT			
			W=1 %	W=5 %	W=10 %	W=20 %
Japanese Vowel	81.89	87.30	*87.30 (the distribution is based on samples from raw data)			
Pen-digit	90.87	96.11	91.66	92.18	93.79	95.22
Page Block	95.73	96.82	*96.82	96.32	95.74	94.87
Satellite	84.48	87.73	85.18	87.1	*87.73	86.25
Segment	89.37	92.91	91.91	*92.91	92.23	89.11
Vehicle	71.03	75.09	72.44	72.98	73.18	*75.09
BreatCancer	93.52	95.93	94.73	94.28	95.51	*95.93
Ionosphere	88.69	91.69	89.65	88.92	*91.69	91.6
Glass	66.49	72.75	69.6	*72.75	70.79	69.69
Iris	94.73	96.13	94.47#	95.27	96	*96.13

IV. PRUNING ALGORITHMS

Although UDT can build a more accurate decision tree, it is not as efficient as AVG. As we have explained, to determine the best attribute and split point for a node, UDT has to examine k(ms-1) split points, where k = number of attributes, m = number of tuples, and s = number of samples per pdf. (AVG has to examine only k(m-1) split points.) For each such candidate attribute Aj and split point z, an entropy $H(z,A_j)$ has to be computed (see(1)). Entropy calculations are the most computation-intensive part of UDT. Our approach to developing more efficient algorithms is to come up with strategies for pruning candidate split points and entropy calculations.

A. Pruning Empty and Homogeneous Intervals

Recall that the Best Split function in UDT is to solve the optimisation problem of minimising $H(z,A_j)$ over all attributes A_j and all possible split points in dom(A_j). Let us first focus on finding the best split point for one particular attribute A_j. (Note that there may be more than one best split point, each giving the same entropy value. Finding any one of them suffices.) We then repeat the process to find the best split point for every other attribute. The attribute with the best split point giving the lowest entropy is taken as the result of Best Split.

We define the set of end-points of tuples in S on attribute A_j as $Q_j = \{q \mid (q = a_{h,j})V(q = b_{h,j})$ for some $t_h \in S\}$. We assume that there are v such end-points, $q_1,q_2,.....,q_v$sorted in ascending order. Within $[q_1,q_v]$, we want to find anoptimal split point for attribute A_j.

Definition 1: For a given set of tuples S, an optimal split point for an attribute A_j is one that minimises $H(z,A_j)$. (Note that the minimisation is taken over all z 2 $[q_1;q_v]$.)

The end-points define v 1 disjoint intervals: $(q_{i,}q_{i+1})$ for i = 1,......,v-1. We will examine each interval separately. For convenience, an interval is denoted by (a; b].

Definition 2 (Empty interval): An interval (a,b] is empty if $\int_a^b fh,j(x)dx = 0$ for all t_h ∈S.

Definition 3 (Homogenous interval): an interval (a,b] is homogenous if there exists a class label c ∈ C such that $\int_a^b fh,j(x)dx \neq 0$ for all t_h ∈S.

Intuitively, an interval is empty if no pdf's intersect it; an interval is homogeneous if all the pdf's that intersect it come from tuples of the same class.

Definition 4: (Heterogeneous interval): An interval (a,b] is heterogeneous if it is neither empty nor homogeneous.

V. EXPERIMENTS ON EFFICIENCY

The algorithms described above have been implemented [8]in Java using JDK 1.6 and a series of experiments were performed on a PC with an Intel Core 2 Duo 2.66GHz CPU and 2GB of main memory, running Linux kernel 2.6.22 i686. Experiments on the accuracy of our novel distribution-based UDT algorithm have been presented already in Section IV-B. In this section, we focus on the pruning effectiveness of our pruning algorithms and their run-time performance. The data sets used are the same as those used in Section IV-B. The same method is used to synthesise data uncertainty. Only Gaussian distribution is used for the experiments below. We use the parameters s = 100 (no. of sample points per pdf) and w = 10% (width of the pdf's domain, as apercentage of the width of the attribute's domain) as the baseline settings.

For the data set "Japanese Vowel", since its uncertainty is taken from raw data (7–29 samples per pdf), we cannot control its properties for sensitivity studies. So, it is excluded from Figures 8 and 9. The bars for "Japanese Vowel" in Figures 6 and 7 are given for reference only.

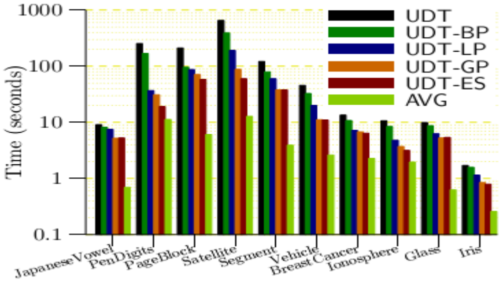


Fig. 6: Execution time

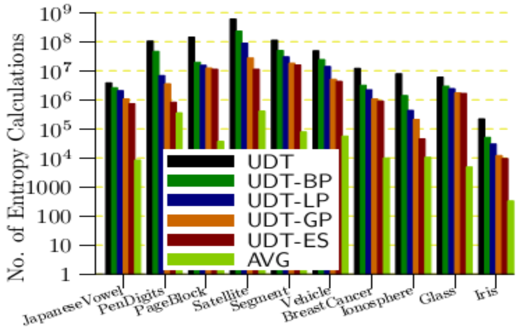


Fig. 7 : Pruning effectiveness

VI. CONCLISION

We have extended the model of decision tree classification to accommodate data tuples having numerical attributes with uncertainty describe by arbitrary pdfs. We have modified classical decision tree building algorithms to build decision tree for classifying such data. Then we found empirically that when suitable pdfs are used, exploiting data uncertainty leads to decision tree with remarkably higher accuracies. Performance is a issue, though because of increase amount of information to the processed, as well as more complicated entropy competitions involved. We have devised so many pruning techniques to improve tree construction efficiency. Our algorithms have been experimentally verified to be highly effective. Therefore execution times are of an order of comparably to classical algorithms. Other techniques namely pruning by bounding and end point sampling, are novel. Although our novel techniques are primarily design to handle uncertain data.

REFERENCES

[1] R. Agrawal, T. Imielinski, and A. N. Swami, "Database mining: A performance perspective," IEEE Trans. Knowl. Data Eng., vol. 5, no. 6, pp. 914–925, 1993.

[2] J. R. Quinlan, "Induction of decision trees," Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.

[3] ——, C4.5: Programs for Machine Learning. Morgan Kaufmann, 1993, ISBN 1-55860-238-0.

[4] C. L. Tsien, I. S. Kohane, and N. McIntosh, "Multiple signal integrationby decision tree induction to detect artifacts in the neonatal intensivecare unit," Artificial Intelligence in Medicine, vol. 19, no. 3, pp. 189–202, 2000

[5] O. Wolfson and H. Yin, “Accuracy and resource consumption in tracking and location prediction,” in SSTD, ser. Lecture Notes in Computer Science, vol. 2750.Santorini Island, Greece: Springer, 24-27 Jul. 2003, pp. 325–343.

[6] W. Street, W. Wolberg, and O. Mangasarian, “Nuclear feature extraction for breast tumor diagnosis,” in SPIE, vol. 1905, San Jose, CA, U.S.A., 1993, pp. 861–870.

[7] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees. Wadsworth, 1984.

[8] L. Breiman, “Technical note: Some properties of splitting criteria,” Machine Learning, vol. 24, no. 1, pp. 41–47, 1996.

[9] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997, ISBN 0070428077.

[10] A. Asuncion and D. Newman, “UCI machine learning repository,” 2007.

◈◈◈

A Study of Various Image Compression Techniques

B. Eswar Reddy[1], M. Veeresha[2] & A. Nagaraja Rao[3]

[1]CSE Dept, JNTUA, Anantapur, A.P, India, [2]Dr.K.V.S.R.I.T,Kurnool,A.P, India,[3]MITS,Madanapalli,A.P, India

Abstract - This paper addresses the area of image compression as it is applicable to various fields of image processing. On the basis of evaluating and analyzing the current image compression techniques this paper presents the Principal Component Analysis approach applied to image compression. PCA approach is implemented in two ways – PCA Statistical Approach & PCA Neural Network Approach. It also includes various benefits of using image compression techniques.

I. INTRODUCTION

1.1 Image

An image is essentially a 2-D signal processed by the human visual system. The signals representing images are usually in analog form. However, for processing, storage and transmission by computer applications, they are converted from analog to digital form. A digital image is basically a 2- Dimensional array of pixels. Images form the significant part of data, particularly in remote sensing, biomedical and video conferencing applications. The use of and dependence on information and computers continue to grow, so too does our need for efficient ways of storing and transmitting large amounts of data.

1.2 Image Compression

Image compression addresses the problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage/transmission requirements. Compression is achieved by the removal of one or more of the three basic data redundancies:

1. Coding Redundancy 2. Interpixel Redundancy 3. Psychovisual Redundancy

Coding redundancy is present when less than optimal code words are used. Intermixed redundancy results from correlations between the pixels of an image. Psycho visual redundancy is due to data that is ignored by the human visual system (i.e. visually non essential information). Image compression techniques reduce the number of bits required to represent an image by taking advantage of these redundancies. An inverse process called decompression (decoding) is applied to the compressed data to get the reconstructed image. The objective of compression is to reduce the number of bits as much as possible, while keeping the resolution and the visual quality of the reconstructed image as close to the original image as possible. Image compression systems are composed of two distinct structural blocks : an encoder and a decoder.

Image f(x,y) is fed into the encoder, which creates a set of symbols form the input data and uses them to represent the image. If we let n1 and n2 denote the number of information carrying units(usually bits) in the original and encoded images respectively, the compression that is achieved can be quantified numerically via the compression ratio, CR = n1 /n2 the encoder is responsible for reducing the coding, interpixel and psychovisual redundancies of input image. In first stage, the mapper transforms the input image into a format designed to reduce interpixel redundancies. The second stage, qunatizer block reduces the accuracy of mapper's output in accordance with a predefined criterion. In third and final stage, a symbol decoder creates a code for quantizer output and maps the output in accordance with the code. These blocks perform, in reverse order, the inverse operations of the encoder's symbol coder and mapper block. As quantization is irreversible, an inverse quantization is not included.

1.3 Benefits of Compression

- It provides a potential cost savings associated with sending less data over switched telephone network where cost of call is really usually based upon its duration.
- It not only reduces storage requirements but also overall execution time.
- It also reduces the probability of transmission errors since fewer bits are transferred.

- It also provides a level of security against illicit monitoring.

II. IMAGE COMPRESSION TECHNIQUES

The image compression techniques are broadly classified into two categories depending whether or not an exact replica of the original image could be reconstructed using the compressed image . These are:

1. Lossless technique
2. Lossy techniqhe

2. 1 Lossless compression technique

In lossless compression techniques, the original image can be perfectly recovered form the compressed (encoded) image. These are also called noiseless since they do not add noise to the signal (image).It is also known as entropy coding since it use statistics/decomposition techniques to eliminate/minimize redundancy. Lossless compression is used only for a few applications with stringent requirements such as medical imaging.

Following techniques are included in lossless compression:

1. Run length encoding
2. Huffman encoding
3. LZW coding
4. Area coding

2.2 Lossy compression technique

Lossy schemes provide much higher compression ratios than lossless schemes. Lossy schemes are widely used since the quality of the reconstructed images is adequate for most applications .By this scheme, the decompressed image is not identical to the original image, but reasonably close to it. In this prediction – transformation – decomposition process is completely reversible .The quantization process results in loss of information. The entropy coding after the quantization step, however, is lossless. The decoding is a reverse process. Firstly, entropy decoding is applied to compressed data to get the quantized data. Secondly, dequantization is applied to it & finally the inverse transformation to get the reconstructed image.

Major performance considerations of a lossy compression scheme include:

1. Compression ratio
2. Signal - to – noise ratio
3. Speed of encoding & decoding.

Lossy compression techniques includes following schemes:

1. Transformation coding
2. Vector quantization
3. Fractal coding
4. Block Truncation Coding
5. Subband coding

2.3 Lossless Compression Techniques

2.3.1 Run Length Encoding

This is a very simple compression method used for sequential data. It is very useful in case of repetitive data. This technique replaces sequences of identical symbols (pixels) ,called runs by shorter symbols. The run length code for a gray scale image is represented by a sequence { Vi , Ri } where Vi is the intensity of pixel and Ri refers to the number of consecutive pixels with the intensity Vi as shown in the figure. If both Vi and Ri are represented by one byte, this span of 12 pixels is coded using eight bytes yielding a compression ration of 1: 5.

2.3.2 Huffman Encoding

This is a general technique for coding symbols based on their statistical occurrence frequencies (probabilities). The pixels in the image are treated as symbols. The symbols that occur more frequently are assigned a smaller number of bits, while the symbols that occur less frequently are assigned a relatively larger number of bits. Huffman code is a prefix code. This means that the (binary) code of any symbol is not the prefix of the code of any other symbol. Most image coding standards use lossy techniques in the earlier stages of compression and use Huffman coding as the final step.

2.3.3 LZW Coding

LZW (Lempel- Ziv – Welch) is a dictionary based coding. Dictionary based coding can be static or dynamic. In static dictionary coding, dictionary is fixed during the encoding and decoding processes. In dynamic dictionary coding, the dictionary is updated on fly. LZW is widely used in computer industry and is implemented as compress command on UNIX.

2.3.4 Area Coding

Area coding is an enhanced form of run length coding, reflecting the two dimensional character of images. This is a significant advance over the other lossless methods. For coding an image it does not make too much sense to interpret it as a sequential stream, as it is in fact an array of sequences, building up a two dimensional object. The algorithms for area coding try to find rectangular regions with the same characteristics. These regions are coded in a descriptive form as an element with two points and a certain structure. This type of coding can be highly effective but it bears the problem of a nonlinear method, which cannot be implemented in hardware. Therefore, the performance in terms of compression time is not competitive, although the compression ratio is.

2.4 Lossy Compression Techniques

2.4.1. Transformation Coding

In this coding scheme, transforms such as DFT (Discrete Fourier Transform) and DCT (Discrete Cosine Transform) are used to change the pixels in the original image into frequency domain coefficients (called transform coefficients).These coefficients have several desirable properties. One is the energy compaction property that results in most of the energy of the original data being concentrated in only a few of the significant transform coefficients. This is the basis of achieving the compression. Only those few significant coefficients are selected and the remaining are discarded.The selected coefficients are considered for further quantization and entropy encoding. DCT coding has been the most common approach to transform coding.It is also adopted in the JPEG image compression standard.

2.4.2 Vector Quantization

The basic idea in this technique is to develop a dictionary of fixed-size vectors, called code vectors. A vector is usually a block of pixel values. A given image is then partitioned into non-overlapping blocks (vectors) called image vectors. Then for each in the dictionary is determined and its index in the dictionary is used as the encoding of the original image vector. Thus, each image is represented by a sequence of indices that can be further entropy coded.

2.4.3 Fractal Coding

The essential idea here is to decompose the image into segments by using standard image processing techniques such as color separation, edge detection, and spectrum and texture analysis. Then each segment is looked up in a library of fractals. The library actually contains codes called iterated function system (IFS) codes, which are compact sets of numbers. Using a systematic procedure, a set of codes for a given image are determined, such that when the IFS codes are applied to a suitable set of image blocks yield an image that is a very close approximation of the original. This scheme is highly effective for compressing images that have good regularity and self-similarity.

2.4.4 Block truncation coding

In this scheme, the image is divided into non overlapping blocks of pixels. For each block, threshold and reconstruction values are determined. The threshold is usually the mean ofthe pixel values in the block. Then a bitmap of the block is derived by replacing all pixels whose values are greater than or equal (less than) to the threshold by a 1 (0). Then for each segment (group of 1s and 0s) in the bitmap, the reconstruction value is determined. This is the average of the values of the corresponding pixels in the original block.

2.4.5 Sub band coding

In this scheme, the image is analyzed to produce the components containing frequencies in well-defined bands, the sub bands. Subsequently, quantization and coding is applied to each of the bands. The advantage of this scheme is that the quantization and coding well suited for each of the sub bands can be designed separately.

III. PCA (PRINCIPAL COMPONENT ANALYSIS)

3.1 Introduction

In statistics, PCA is a technique for simplifying a dataset by reducing multidimensional datasets to lower dimensions for analysis. PCA is a standard technique commonly used for data reduction in statistical pattern recognition and signal processing. PCA has been called one of the most valuable results from applied linear algebra. It is used abundantly in all forms of analysis from neuroscience to computer graphics, because it is a simple non- parametric method of extracting relevant information from confusing datasets.

PCA is also called the KARHUNEN-LOEVE Transform (KLT, named after Kari Karhunen & Michel Loeve) or the HOTELLING Transform. Its general objectives are:

1. Data reduction
2. Interpretation

There are basically two approaches for performing PCA. They are classical statistical method and artificial neural network method.

3.2 PCA Classical Statistical Method

It involves finding eigen values and corresponding eigen vectors of the data set using covariance matrix. The corresponding eigen values of the matrix gives an indication of amount of information the respective principal components represent. The methodology for calculating principal component is given by the following algorithm. Let X1,X2,……..Xm are the sub images of dimension N. The corresponding algorithm is described as follows:

1. Computation of the global mean (X) from sub images.

   X = 1 / M Σ Xi

2. Subtraction of the mean from each sub image to generate the mean removed image.

   Øi = Xi - X

3. Formation of the matrix using mean removed sub image of (M X N) dimension

   A = [Ø 1 Ø2 … ……ØM]

4. Computation of the sample covariance matrix (C) of dimension (N X N)

5. Computation of the Eigen values of the covariance matrix. Computation of Eigen values is performed by jaccobian iteration method.

   C : λ1 > λ2 > ……… >λN

6. Computation of the eigen vectors for the eigen values

   C: u1 , u2, …………..,uN

7. Dimensionality reduction step. Keep only the Eigen vectors corresponding to K largest eigen values. These Eigen values are called as "principal components".

The above said steps are needed to generate the principal components of the image. Corresponding eigen vectors are uncorrelated and have the greater variance. In order to avoid the components that have an undue influence on the analysis, the components are usually coded with mean as zero and variance as one. This standardization of the measurement ensures that they all have equal weight in the analysis.

3.3 PCA Neural Network

Artificial neural network are model that attempt to achieve performance via dense inter connection of simple computational elements. The most important property of a neural network is the ability to learn from its environment. PCA is a powerful linear block transform coding in which, an image is subdivided into non-overlapping blocks of N×N pixels which can be considered as N-Dimensional vectors with N = n × n. A linear Transformation, which can be written as an M ×N – dimensional matrix W with M ≤ N, is performed on each block with the M rows of W, wi being the basis vectors of the transformation. An adaptive principal component extraction (APEX) is used to decorrelate the principal components. The main difference between this APEX architecture and the existing PCA networks lies in the additional lateral connections at the outputs of the network.

3.4 Applications of PCA in Computer Vision Representation

When using these sort of matrix techniques in computer vision, representation of images should be considered. A square, N by N image can be expressed as an N2 – dimensional vector.

X = (x1 x2 x3 …………………xN2)

Where the rows of pixels in the image are placed one after the other to form a one dimensional image. E.g. The first N elements x1 – xN will be the first row of the image, the next N elements are the next row, and so on. The values in the vector are the intensity values of the image, possibly a single greyscale value.

- PCA to find patterns

Say we have 20 images Each image is N pixels high by N pixels wide. For each image we can create an image vector as described in the representation section. Which gives a starting point for our PCA analysis. It turns out that these axes works much better for recognizing faces, because the PCA analysis has given the original images in terms of the differences and similarities between them. The PCA analysis has identified the statistical patterns in the data.

- PCA for Image Compression

If we have 20 images each with N2 vectors and 20 dimensions. Now, PCA can be implemented on this set of data. 20 Eigen vectors will be obtained because each vector is 20 –dimensional. To compress the data, choose the data using only 15 Eigen vectors. This gives a final data set with only 15 dimensions, which has saved ¼ of the space. However, when the original data is reproduced, the images have lost some of the information. This compression is said to be lossy because the decompressed image is not exactly the same as the original.

IV. CONCLUSION

This paper presents various types of image compression techniques. There are basically two types of compression techniques. One is Lossless Compression and other is Lossy Compression Technique. Comparing the performance of compression technique is difficult unless identical data sets and performance measures are used. Some of these techniques are obtained good for certain applications like security technologies. Some techniques perform well for certain classes of data and poorly for others. PCA (Principal Component Analysis) also found its applications as image compression. PCA can be implemented in two forms i.e. either statistical approach or neural network approach. The PCA Neural Network provides new way of generating codebook based on statistical feature of PCA transformational coefficients. It leads to less storage of memory and reduction of calculation.

REFERENCES

[1] Subramanya A, “Image Compression Technique,” Potentials IEEE, Vol. 20, Issue 1, pp 19-23, Feb-March 2001,

[2] David Jeff Jackson & Sidney Joel Hannah, “ Comparative Analysis of image Compression Techniques,” System Theory 1993, Proceedings SSST ’93, 25th Southeastern Symposium,pp 513-517, 7 –9 March 1993.

[3] Hong Zhang, Xiaofei Zhang & Shun Cao, “ Analysis & Evaluation of Some Image Compression Techniques,” High Performance Computing in Asia. Pacific Region, 2000 Proceedings, 4th Int. Conference, vol. 2, pp 799-803,14-17 May, 2000

[4] Ming Yang & Nikolaos Bourbakis ,“An Overview of Lossless Digital Image Compression Techniques,” Circuits & Systems, 2005 48th Midwest Symposium ,vol. 2 IEEE ,pp 1099-1102,7 – 10 Aug, 2005

[5] Milos Klima, Karel Fliegel, “Image Compression Techniques in the field of security Technology: Examples and Discussion,”Security Technology, 2004, 38th Annual 2004 Intn. Carnahan Conference, pp 278-284,11-14 Oct., 2004

[6] Ismail Avcibas, Nasir Memon, Bulent Sankur, Khalid Sayood, “ A Progressive Lossless / Near Lossless Image Compression Algorithm,”IEEE Signal Processing Letters, vol. 9, No. 10, pp 312-314, October 2002.

[7] Dr. Charles F. Hall, “ A Hybrid Image Compression Technique,” Acoustics Speech & Signal Processing, IEEE International Conference on ICASSP’ 85, Vol. 10, pp 149-152, Apr, 1985

[8] Wen Shiung Chen, en- HuiYang & Zhen Zhang, “ A New Efficient Image Compression Technique with Index- Matching Vector Quantization,” Consumer Electronics, IEEE Transactions, Vol. 43, Issue 2, pp 173- 182, May 1997.

[9] David H. Kil and Fances Bongjoo Shin, “ Reduced Dimension Image Compression And its Applications, ”Image Processing, 1995,

Proceedings, International Conference,Vol. 3 , pp 500-503, 23-26 Oct.,1995

[10] C.K. Li and H.Yuen, “A High Performance Image Compression Technique For Multimedia Applications,” IEEE Transactions on Consumer Electronics, Vol. 42, no. 2, pp 239-243, 2 May 1996.

[11] Shi-Fei Ding, Zhong –Zhi Shi,Yong Liang , Feng- Xiang Jin, “ Information Feature Analysis and Improved Algorithm of PCA,” Proceedings of the 4th International Conference on Machine Learning and Cybernetics, Guangzhou, pp 1756-1761 , 18-21 August,2005

[12] Vo Dinh Minh Nhat, Sung Young Lee, “Two-Dimensional Weighted PCA algorithm for Face Recognition,” Proceedings 2005 IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp 219-223, June 27-30,2005, Espoo, Finland

◈◈◈

Moving Object Detection Using Scaled Feature Tracking

Ashok Kumar Patel & Sanjay Kumar

Dept of Electrical Engineering, NIT Raipur, Chhattisgarh, India

Abstract - Every object has features like color, shape, texture and so on. In which shape is most important feature to identify the object. Edges are used to create shape of an image. Similarly color can identify an object when its background is of different color. Texture can also identify the object of an image. Edges with direction are used to characterize the texture of an image. These features can be used to identify an object when it is moving as the object move some of the shapes, color and texture of an image is changed. I used LoG algorithm for edge detection, histogram analysis for color detection and rangefilter for texture detection.

Keywords - *Object feature, color, shape, texture, edges, log algorithm, histogram, rangefilter etc.*

I. INTRODUCTION

The feature is defined as a function of one or more measurements, each of which specifies some quantifiable property of an object, and is computed such that it quantifies some significant characteristics of the object. General features of an image are color, shape and texture. Feature calculated at each pixel is color, subdivision of image is shape and entire image is texture. In which shape is most important to identify an object in an image. Edges are used for shape feature extraction.

Moving object detection is used very much for surveillance purposes. It is also the first step for further processing of image mining. There are two methods of detection of moving object, one is using background detection [1] and second one is temporal frame differencing [2]. In background detection algorithm first background is calculated using different method such as Gaussian background estimation; median filter back ground estimation and so on. While in temporal frame differencing method two consecutive frames are differenced using pixel wise.

In ideal edge is discontinuity. There are various methods for edge detection, in which Sobel, Prewitt, Roberts, Canny and Laplacian of Gaussian (LoG) are commonly used. Here we use LoG algorithm for edge detection.

The Laplacian of an image highlights regions of rapid intensity change and is therefore often used for edge detection. The Laplacian ***L(x,y)*** of an image with pixel intensity values I(x,y) is given by

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

As it is second derivative so it is very sensitive to noise so image is often Gaussian smoothed before applying the Laplacian filter.

So the LoG function centered on zero and with Gaussian standard deviation σ has the form

$$LoG(x,y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Generally we say that the color is the pixels value of an image. It has different format RGB, CMYK, YUV, grayscale, binary (black and white) and so on. Histogram is used for the analysis of the color of an image. It present the number of occurrence of color value in the range of color present in image. The horizontal axis shows the range of color while vertical axis shows the number of occurrence.

For n bit image the range of color is 0 to 2^n-1

Then histogram

$$h(r_k) = n_k$$

Where

r_k : the kth gray level

n_k : the number of pixels in the image having gray level r_k

$h(r_k)$: histogram of a digital image with gray levels r_k

Texture is spatial arrangement of color or intensity in an image. There is two approach for texture analysis, structured approach and statistical approach. Statistical approach is easier and generally used. Range filter, standard deviation filter and entropy filter are some of the texture analyzing filter. Here we use range filter for texture analysis.

If the chosen neighborhood is n by n

Then for each pixel i of the image the result is

$$Val_i=max_{nxn}-min_{nxn}$$

Where

***max_{nxn}* :** is the maximum value for n by n neighborhood an

***min_{nxn}* :** is the minimum value for n by n neighborhood

II. MY ALGORITHM

I want to present three algorithms for feature tracking. The first algorithm is tracking edges changes, the second algorithm is tracking histogram variation and the third is structure variation.

A. First algorithm

This algorithm reads the two consecutive frames and converts them into grayscale. Then find the edges in each images and subtract the previous frame with the next frame. The algorithm is as follow:

1. Read two consecutive frames
2. Convert them into grayscale image
3. Find the edges using LoG edge detection algorithm
4. Subtract the previous image with the current image

   A and **B** are Logical images

   So apply **and(~A,B)**
5. Apply filter2 on the result image

B. Second algorithm

This algorithm reads the two consecutive frames and converts them into grayscale. Then find the histogram of each image. Finally subtract the histogram of previous image with the current image and check for the threshold. If change is more than threshold then current image with changed intensity value is checked in the previous image. The algorithm is as follow:

1. Read two consecutive frames
2. Convert them into grayscale image
3. Find the histogram of each image
4. Subtract the current image with previous image
5. Select a threshold value of change
6. For each pixel of current image

   If its intensity is changed intensity

   And it is not same in the previous image

   Than remain same as in current image

   Else

   Change to zero
7. Apply Morphologically open binary image on the result image

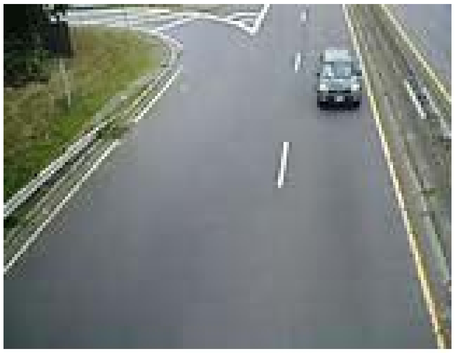
C. Third algorithm

This algorithm reads the two consecutive frames and converts them into grayscale. Then find the structure of each image. Finally subtract the current image with the previous image. The algorithm is as follow:

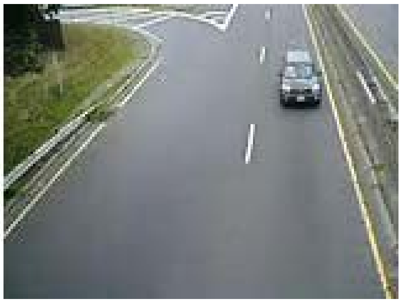
1. Read two consecutive frames
2. Convert them into grayscale image
3. Find the histogram of each image
4. Subtract the current image with previous image
5. Apply median filter 2 on the result image

III. EXPERIMENT AND RESULT

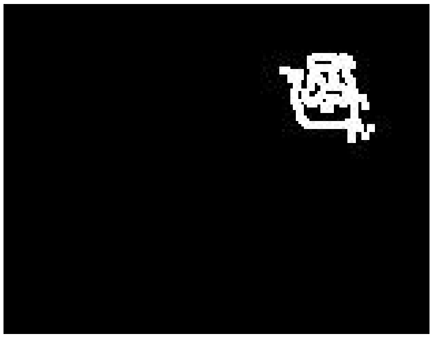
All the three algorithms are sufficient to detect the moving object. While not any two algorithm produces same result. I have applied these algorithms on the AVI file present in the demo of Matlab named as ‘viptraffic.avi’ and select two consecutive frames numbered as 108 and 109.



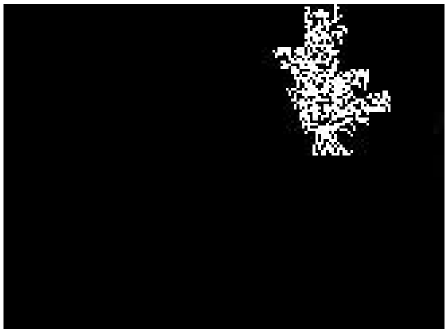
(a) Frame 108 of vip traffic



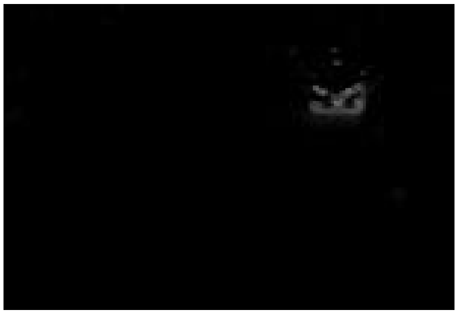
(b) Frame 109 of vip traffic



(c) by edge tracking



(d) By color detection



(e) By structure detection

IV. CONCLUSION AND FUTURE SCOPE

From the result we can see that the entire three algorithms are sufficient for detection of moving object. In future we can try combined feature like edge and structure, edge and color, or structure and color to detect. They may produce more good result for boundaries of object as here it is distorted.

REFERENCES

[1] Arnab Roy, Sanket Shinde and Kyoung-Don Kang for “An Approach for Efficient Real Time Moving Object Deatection” ESA 2010: 157-162. 31,

[2] Ka Ki Ng and Edward J. Delp for “Object Tracking Initialization Using Automatic Moving Object Detection” Proc. SPIE 7543, 75430M (2010)

[3] http://homepages.inf.ed.ac.uk/rbf/HIPR2/log.htm.

[4] http://en.wikipedia.org/wiki/Image_texture

[5] http://www.mathworks.in/help/toolbox/images/f11-27972.html.

◈◈◈

Voxel Based Morphometry Revisited

Asra Anjum & Meghana Nagori

Dept of Computer Science and Engineering, Govt.Engineering College Aurangabad (M.S), Aurangabad (M.S) India

Abstract - The methodology of Voxel Based Morphometry (VBM) deals with volumetric comparison of grey matter between two groups. The result obtained is Statistical Parametric Map (SPM) showing regions where grey matter concentration differs significantly among the groups. This Paper reviews the Voxel Based Morphometry methodology, with particular emphasis on segmenting gray matter from fmri images to demonstrate grey matter volume differences in patients with schizophrenia and healthy controls. Voxel Based Morphometry was applied to evaluate grey matter volume in 8 patients with schizophrenia and 8 healthy control subjects. Statistical Parametric Map showing regions where gray matter differs among the groups also produced.

Keywords *- VBM, Grey Matter, Schizophrenia, SPM, healthy controls.*

I. INTRODUCTION

Morphometric studies of brain which deals with quantitative volumetric region of interest shows that schizophrenia is related with abnormalities in brain structure. These findings are of great significance in understanding of neurobiology of schizophrenia as in ([5],[6],[7]).The recently developed VBM technique explores differences in local composition of brain tissues by comparing on voxel by voxel basis as in([2],[3],[4]). Voxel-based morphometry of MRI data deals with normalising all the images to some target template, extraction of grey matter from normalised images, smoothing and making inferences from statistical analysis. The output is a statistical parametric map with difference in gray matter concentration between groups as in [1].

II. SUBJECTS AND METHODS

A. Subjects

Healthy comparison subjects and schizophrenic/ schizoaffective male and female adults between the ages of 18 and 70 were recruited for this study. All subjects had regular hearing levels (no more than loss in either ear), had sufficient eyesight or were correctable to be able to see visual display, were fluent in English, and were able to perform the cognitive tasks in this study. No female subjects were pregnant and female subjects of childbearing potential received a urine or blood pregnancy test before the MRI. There could be no contradictions to MRI scanning including a cardiac pacemaker, metal fragments in eye, skin, body; heart valve replacement, brain clips, venous umbrella, being a sheet-metal worker or welder, aneurysm surgery, intracranial bypass, renal, aortic clips; prosthetic devices such as middle ear, eye, joint, or penile implants, joint replacements; hearing aid, neurostimulator, insulin pump; shunts/stents, metal mesh/coil implants; metal plate/pin/screws/wires, or any other metal implants; permanent eyeliner or permanent artificial eyebrows or significant claustrophobia. Control subjects were excluded if they had a current or past history of a major neurological, psychiatric, medical illness; previous head injury; substance or alcohol dependence; and IQ less than 75 (as measured by the North American Adult Reading Test (NAART)) Subjects with schizophrenia or schizoaffective disorder meeting DSM-IV criteria were allowed in the study.

B. Image and Statistical Analysis

Image analysis was performed using statistical parametric mapping software (SPM5; Welcome Department of Cognitive Neurology, Institute of Neurology, London, UK, http:// www. fil.ion. ucl.ac.uk/ spm/) with the VBM toolbox version VBM5.1 Published by Christian Gaser in MATLAB R2009a , which involved spatial normalization, segmentation, and spatial smoothing .(Gaussian kernel of 12 mm full width at half maximum for GM images). Modulated images were used for statistical analyses.Fig.1 shows the segmented images for one healthy control and Fig 2 shows segmented images for one patient with schizophrenia. Based on the general linear model, statistical parametric maps were created to identify brain regions with significant changes in grey matter in patients relative to controls. The smoothed GM images were analyzed with 2 *t* test. An absolute threshold of

0.1, implicit masking and age as covariate was used in the model. Small volume correction of 30 mm sphere at global maxima performed.

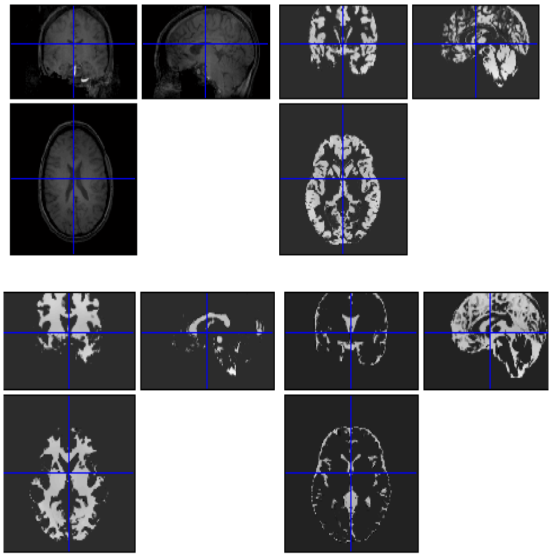


Fig. 1 : VBM result for healthy control (top left original image, top right GM tissue, bottom left WM tissue, bottom right CSF tissue)

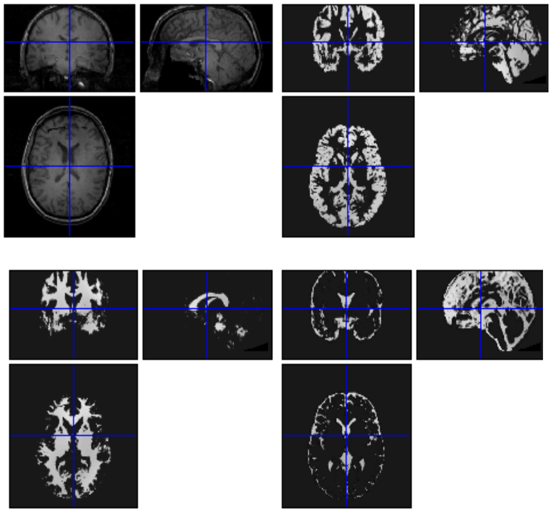


Fig. 2 : VBM result for schizophrenia (top left original image, top right GM tissue bottom left WM tissue, bottom right CSF tissue)

III. RESULT

The results were successfully obtained after performing the VBM steps on fmri images of patients with schizophrenia and healthy control subjects. To obtain grey matter volume the stat tool of VBM (Read raw volumes GM/WM/CSF/total volume) was used. The mean volumes of grey matter for patients with schizophrenia and healthy control obtained are in Table1.

Using the glass brain the result of SPM analysis are displayed in three orthogonal planes as in Fig.3 for schizophrenia < healthy controls .Result for schizophrenia > control as shown in Fig.4.The result demonstrated that schizophrenia patients shows reduced grey matter regions as compared to healthy controls.

Voxel wise coordinates of significant regions are as shown in Table.2

TABLE I

GM volume evaluation

Tissue Volumes(cm3)	**Mean for schizophrenic patients**	**Mean for normal control**
GM Volume	720.60	856.23

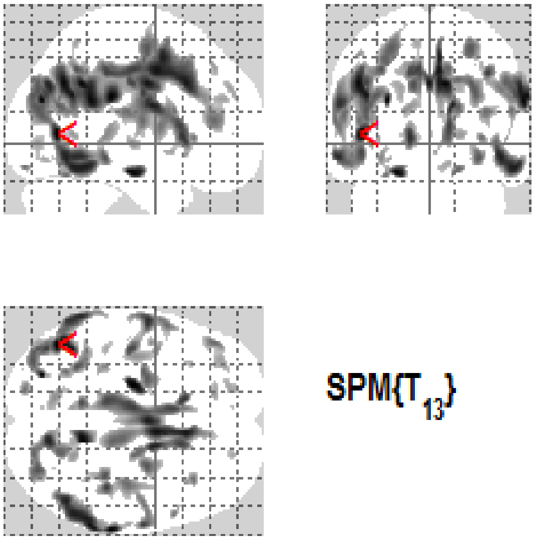


Fig. 3 : SPM (T) maps for grey matter schizophrenia<controls

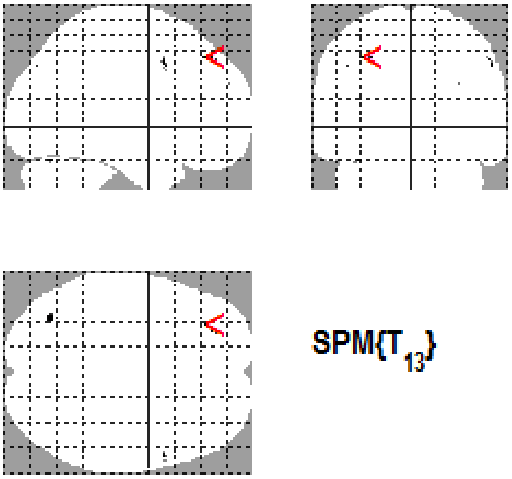


Fig. 4 : SPM (T) maps for grey matter schizophrenia>controls

TABLE 2

Related regional findings in patients with schizophrenia and healthy controls

Anatomical region	**Brodmann Area**	**X**	**Y**	**Z**
Temporal lobe	37	-45	-65	5
	39	-42	-56	30
	37	-61	-51	-7
	19	-42	-78	21
	39	-42	-73	30
	39	-42	-75	16
	37	-54	-60	-12
Occipital lobe	19	-27	-76	26
	19	-36	-82	28
	19	-36	-82	28
	19	-27	-74	-10
	18	-47	-78	0
	18	-43	-78	-9
Limbic Lobe	30	-18	-61	5
	31	-18	-64	18
Posterior Lobe	-	-29	-59	-19

REFERENCES

[1] Voxel-Based Morphometry—The Methods John Ashburner and Karl J. Friston NeuroImage **11,** 805–821 (2000)

[2] Daniel Mietchen and Christian Gaser; Computational morphometry for detecting changes in brain structure due to development, aging, learning,disease and evolution; Frontiers in Neuroinformatics www.frontiersin.org August2009 | Volume 3 | Article 25

[3] Simon S. Keller1 & Neil Roberts2, Measurement of brain volume using MRI:software, techniques, choices and prerequisites, Journal of Anthropological Sciences the JASs is published by the Istituto Italiano di Antropologia www.isita-org.com Vol. 87 (2009), pp. 127-151.

[4] Robyn Honea, B.Sc. Tim J. Crow, M.B., Ph.D.,Dick Passingham, Ph.D.Clare E. Mackay, Ph.D.,Regional Deficits in brain Volume in schizophrenia A Meta Analysis of Voxel Based Morphometry studies ,Am J Psychiatry 2005.

[5] Wright IC,Rabe-Hesketh S,WoodruffPW,David AS,Murray RM, Bullmore ET (2000) Meta-analysis of regional brain volumes in schizophrenia.Am J Psychiatry 157:16–25

[6] Shenton ME, Dickey CC, Frumin M,McCarley RW (2001) A re- view of MRI findings in schizophrenia. Schizophr Res 49:1–52

[7] Lawrie SM, Abukmeil SS (1998) Brain abnormality in schizo- phrenia – A systematic and quantitative review of volumetric magnetic resonance imaging studies. Br J Psychiatry 172: 110–120

◈◈◈