ROBUST LDP BASED FACE DESCRIPTOR

Mahadeo D. Narlawar and Jaideep G. Rana
Department of Electronics Engineering, Jawaharlal Nehru College of Engineering,
Aurangabad-431004, Maharashtra, India

ABSTRACT

This paper presents a novel LDP based image descriptor which is more robust to temporal face changes. LDP is a framework to encode directional pattern based on local derivative variations, hence LDP is highly directional. However texture based features extracted globally tend to average over the image area. Hence this paper proposes to divide the face image into multiple regions and perform LDP separately on each region. The procedure consists of using the texture descriptor to build several local descriptions of the face and combining them into a global description. Block based processing ensures robustness against pose, facial expressions and other temporal face changes. Also, a modification in Chi Square Distance is proposed to suit the spatial histogram matching of faces. Extensive experimental results on FERET, CAS-PEAL, CMU-PIE, and Extended Yale B databases show that the LDP based face descriptors consistently performs better than LDP extracted globally for both face identification and face verification under various conditions.

KEYWORDS: *Local Derivative Pattern, Local Binary Pattern, Spatially Enhanced Histogram, Face Recognition, Chi Square Distance*

I. INTRODUCTION

Automatic face detection has been an active research topic in the field of computer vision. In recent days, increased processing power makes it possible to use face recognition in various real time applications. Application of face recognition ranges from security, surveillance to fraud detection etc. A key issue in face recognition is finding efficient face descriptors for face appearance. Representation issues include: what representation is desirable for the recognition of a pattern and how to effectively extract the representation from the original input image. An efficient descriptor should be of high ability to discriminate between classes, has low intra-class variance, and can be easily computed. Holistic methods such as Principle Component Analysis, Linear Discriminant Analysis, and Hidden Markov Model etc. are widely used. There are a different set of feature based approaches which are gaining popularity in recent times. Lately it's been learned that local methods are robust to temporal face changes like facial expressions, rotation and scale etc. To use face recognition systems in real life scenarios temporal face changes must have to be taken in to account. This paper presents a novel descriptor based on local derivative pattern extracted from local facial regions.

One of the earliest features based on local regions is Eigen faces – Eigen Features [4]. Eigen faces is fast, simple and practical. It's on based Principle Component Analysis, the goal is to represent image in lower dimensions without losing much information and then restructuring it. Extension to Eigen faces in the form of Eigen Features was proposed by Pentland et al [7]. It's a hybrid approach in which features are obtained in local regions independently. Elastic Bunch Graph Matching (EBGM) [9] describes faces using Gabor filter responses in certain facial landmarks and a graph describing the spatial relations of these landmarks. The validity of the component based approach is also attested by the study conducted by Heisele et al [10]. in which a component-based face recognition system clearly outperformed global approaches on a test database containing faces rotated in depth [10].One of the local descriptors is local feature analysis (LFA) proposed by Penevet al. [8]. In LFA, a dense set of

local-topological fields are developed to extract local features. Through discovering a description of one class objects with the derived local features, LFA is a purely second-order statistic method.
The recent local binary pattern (LBP) features are first order derivative features containing texture information [1], [16]. The idea behind LBP is face can be seen as a collection of large number of micro pattern. It achieves better performance than PCA, EBGM and other earlier methods. However it's been observed that first order features do not contain all the information. Hence higher order local derivative pattern (LDP) are proposed, which can capture more discriminative information than LBP.
Local derivative patterns are proposed by Zhang et all, it's a general framework to encode directional derivative variations. Following section explains LBP and LDP in detail.

II. LDP BASED FACE RECOGNITION

2.1. Local Binary Pattern

Local Binary Pattern (LBP) was introduced by T. Ojala [1]. LBP is gradient based local gray scale texture measure and used to model the texture of an image. The original LBP operator labels the pixels of an image by thresholding the 3x3 neighbourhood of each pixel with the value of the central pixel. The binary results of all 3x3 pixels are concatenated to generate a binary string. The decimal equivalent value of this binary string is LBP of the given pixel.
Let Z_0 be the central pixel and Z_1 to Z_8 be the pixels around the central pixel as shown in figure 1.

Z1	Z2	Z3
Z8	Z0	Z4
Z7	Z6	Z5

Figure 1. 3x 3 neighbourhoods around Z_0

Now, a thresholding function f(.,.) can be defined as

$$f(I(Z0), I(Z1)) = \begin{cases} 0, & I(Zi) < I(Z0) \\ 1, & I(Zi) \geq I(Z0) \end{cases} \quad (1)$$

Where i = 1, 2, 3, …., 8
Figure 2 shows an example of the LBP where value of central pixel is used as a threshold.
A normalized histogram can optionally be calculated and used as the features. The histogram contains the information about the edges and spread of the discontinuities over the image.

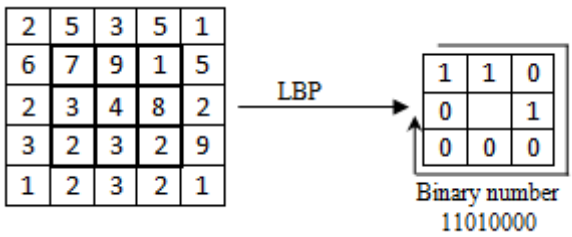

Figure 2. Illustration of basic LBP operator for the region highlighted in black

LBP is very efficient in terms of computational speed and complexity [16]. Also the memory requirements of LBP are very less. LBP has been proved to be very useful in many of the studies [2]. LBP is very popular due to easy to compute features and used in many face recognition applications.

2.2. Local Derivative Pattern

Local Derivative Pattern was proposed by Baochang Zhang [3] for face recognition with high order local pattern descriptor. It encodes directional feature pattern based on local derivative variations. It can capture more detailed information than the first order LBP. LDP is a micro pattern representation which can also be modelled by histogram to preserve the information about the distribution of the LDP micro patterns. LBP is always considered first-order local pattern operator, because LBP encodes all-direction first-order derivative binary result whereas LDP encodes the higher-order derivative information. So it contains more discriminative features than LBP.

Given an image, the first-order derivatives along 0°, 45°, 90° and 135° directions are denoted as $I'_{\alpha}(Z)$ where, α = 45°, 90° and 135° . Let Z_0 be a point in I(Z), and Zi, I = 1, …, 8 be the neighbouring point around Z_0 (see Fig. 1). The four first-order derivatives at Z = Z_0 can be written as

$$I'_{0^0}(Z_0) = I(Z_0) - \mathrm{I}(Z_4)$$
$$I'_{45^0}(Z_0) = I(Z_0) - \mathrm{I}(Z_3)$$
$$I'_{90^0}(Z_0) = I(Z_0) - \mathrm{I}(Z_2)$$
$$I'_{135^0}(Z_0) = I(Z_0) - \mathrm{I}(Z_1) \quad (2)$$

The second order directional derivative in α direction at Z = Z0 can be defined as

$$LDP^2_{\propto}(Z_0) = \{f(I'_{\propto}(Z_0), I'_{\propto}(Z_1)), f(I'_{\propto}(Z_0), I'_{\propto}(Z_2)) \dots, f(I'_{\propto}(Z_0), I'_{\propto}(Z_8))\} \quad (3)$$

Finally, 8-bit LDP can be derived by concatenating the derivatives for all directions.

$$LDP^2(Z) = \{LDP^2_{\propto}(Z) | \propto = 0^0, 45^0, 90^0, 135^0\} \quad (4)$$

Hence a 32 bit binary number is generated as a result of comparing two derivatives for two neighbouring pixels.

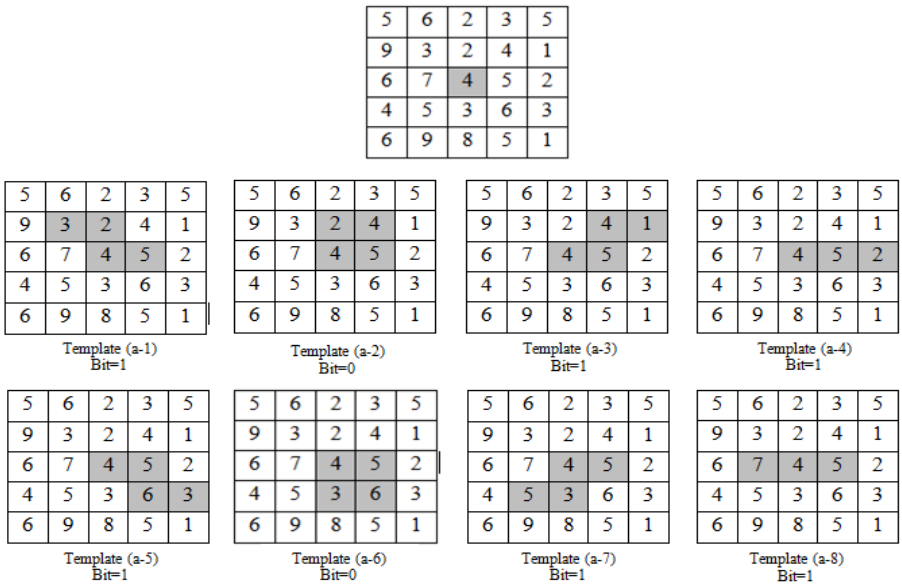

Figure 3. 3 point and 4 point LDP Templates

LDP can be learned as illustrated in figure 3. It can be classified as either a 3 point or 4 point template. For a 3 point template assigns a "0" to a monotonically increasing or decreasing pattern and "1" is assigned to turning point. Similarly for a 4 point template a gradient turning pattern is labelled as "1" and monotonically increasing or decreasing pattern is labelled as "0".

2.3. Face Description with LDP

In this work, LDP operator defined in previous sub section is used for face description. The procedure consists of using the texture descriptor to build several local descriptions of the face and combining

them into a global description. Instead of striving for a holistic description, this approach was motivated by two reasons: the local feature based or hybrid approaches to face recognition have been gaining interest lately [6], [8] which is understandable given the limitations of the holistic representations. These local features based and hybrid methods seem to be more robust against variations in pose or illumination than holistic methods.

Another reason for selecting the local feature based approach is that trying to build a holistic description of a face using texture methods is not reasonable since texture descriptors tend to average over the image area. This is a desirable property for ordinary textures, because texture description should usually be invariant to translation or even rotation of the texture and, especially for small repetitive textures, the small-scale relationships determine the appearance of the texture and thus the large-scale relations do not contain useful information. For faces however, the situation is different: retaining the information about spatial relations is important.

This reasoning leads to the basic methodology of this work. The facial image is divided into local regions and texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face. Following figure for an example of a facial image divided into rectangular regions.

Figure 4. Various Image Regions

The basic histogram can be modified to spatially enhanced histogram which encodes both the appearance and spatial relations of facial image. As m x n regions is present, histograms of each region can be computed independently. The resulting histograms can be combined to obtain the spatially enhanced histogram. Regions need not to be rectangular and equal in size. Even all the face region coverage is not compulsory. Only few of the blocks can be used to obtain LDP histogram to save the time. Overlapping regions can increase the robustness of the method.

In such a way, a face is modelled at multiple levels; in-region LDP models the pixel level directional appearance pattern. Block level histogram models the spatial relationship helping to increase the recognition rate. Concatenation of all histogram provides a global outlook.

Many histogram matching techniques are proposed like histogram intersection, chi square distance etc. The original Chi Square distance is proposed as –

$$x^2(P,Q) = \frac{1}{2}\sum_i \frac{(P_i - Q_i)^2}{(P_i + Q_i)} \tag{5}$$

In this paper we use weighted Chi Square Distance method to match the histograms. The Weights are assigned to individual blocks of facial region. The weights are defined to increase robustness against the temporal face changes.

III. EXPERIMENTS

Several experiments with benchmark databases are conducted. Also for evaluating tests under rotated and scaled faces, separate datasets were generated. Method is tested for online operation also.

3.1. Previous Experiments

Previous experiments for high order LDP are conducted by Zhang et all [3]. The authors extend the LDP to feature images containing wider range of appropriate discriminative features could achieve a higher level of system performance.

The experiments used a large data from FERET [11], CAS-PEAL [12], CMU-PIE [13], Extended Yale B [14], and FRGC [15] databases. The authors considered LBP, 2nd order LDP, 3rd order LDP and 4th

order LDP on gray scale as well as Gabor images. The extensive experiments concluded that 3rd order LDP performs consistently better using gray scale as well as Gabor images [3].

3.2. Experimental Setup

To compare results with previous implementation and other standard algorithms, publically available benchmark datasets are used.
Commonly available databases are FERET, CAS-PAL, CMU-PIE, Yale B and FRGC.
FERET is a very popular database used for face recognition purpose. It contains the faces under various expressions, lightning conditions and aging. Fa containing 1,196 frontal images of 1,196 subjects was used as the gallery set, while Fb (1,195 images with expression variations), Fc (194 images taken under different illumination conditions), Dup I (722 images taken later in time between one minute to 1,031 days), and Dup II (234 images, a subset of Dup I taken at least after 18 months) were used as the Probe sets.
CAS – PEAL database have images for various poses and illumination conditions for 1040 subjects. It contains slightly more images than 9000. It focuses on faces having various accessories for different lightning conditions.
CAS-PIE database have images for various poses and illuminations. This database contains the face images of 68 subjects. This is the only database suitable for our experimental purpose.

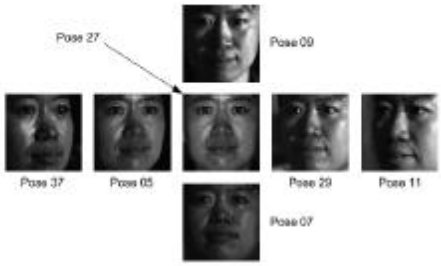

Figure 5. Samples from CAS-PIE database

Experimental implementation is carried out using MATLAB to operate on stored and online camera as well. Face pre-processing is used to nullify the variations in illuminations. In online version faces are captured in controlled environment within specified distanced from camera. The performance is measured in terms of false positives and true positives; a lower true negative makes system more reliable.

3.3. Parameters for LDP

Experiments using various window sizes are conducted, results do not vary to a large extent on basis of window size. Optimal window size can be selected according to face area. The selection of the window size is crucial and decides recognition rate against the scale. Feature vector length is directly proportional to window size. A window size of 15 x 18 was selected; it provides a good trade off between recognition rate and feature vector length. The histogram bin length was used to 64 bins. The overall performance is also robust against the selection of no of regions. A 7 x 7 combination was used; you may change the number of regions depending upon the image resolution. There are no thumb rules for selecting number of regions; however each region should logically describe some face feature.
Various distance measures were compared and it was noticed that Weighted Chi Square distance performs consistently better across most of the datasets. Weights were selected experimentally. Selection of weights is tedious and crucial factor affect the recognition rate very heavily. Higher weights are assigned to those face regions which are least affected by temporal changes.

(a) (b)

Figure 6. (a) A facial image divided into 7X7 windows. (b) The weights set for the weighted x^2 dissimilarity measure. Black squares indicate weight 0.0, dark grey 1.0, light grey 2.0 and white 4.0.

Lighter windows represent higher weight and darker windows represent lower weight. Lowest weighted windows can even be discarded to speed up the operation.

IV. RESULTS

The authors of [3] Zhang et all concluded that 3rd order LDP performs consistently better than LBP and 2nd order LDP. They have covered very large dataset for the experimentation. Also, 3rd order LDP performs better than 4th order LDP. Hence it can be concluded that higher order LDP captures more detailed information than first order LBP and second order LDP. Higher order LDP can be used with gray scale images as well as Gabor feature images. And it has been seen than Gabor feature images prove to be more accurate than gray scale images. Hence for evaluation of face descriptors we have used 3rd order LDP.

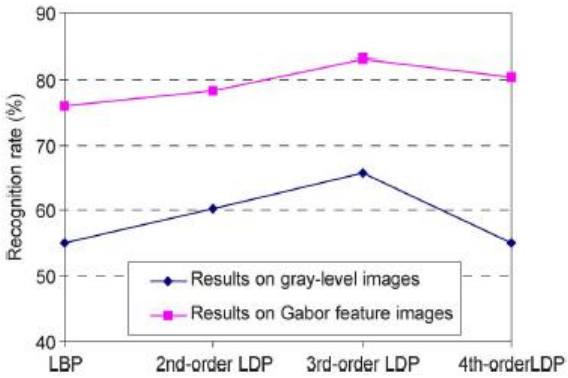

Figure 7. Comparative rank one identification accuracies of LDP and LBP on the CAS-PEAL database.

Table1. Recognition result on different texture descriptors.

Method	Fa	Fb	Dup I	Dup II
Difference Histogram	0.87	0.12	0.39	0.25
Homogeneous Texture	0.86	0.04	0.37	0.21
Texton Histogram	0.97	0.28	0.59	0.42
LDP Descriptor	0.94	0.56	0.71	0.50

Above table presents the results obtained using different texture descriptors. For easiest face set Fa, all the methods perform well. The performance is more or less similar; a large variation in performance can be seen for rest of the sets. Fb is most difficult set of all containing different facial expressions. The performance of LDP based face descriptor is highest among all.

The improvement can be justified by the fact that LDP descriptor is more tolerant as compared to rest of the texture descriptors. It should be noted that the performance is dependent on the configuration we select.

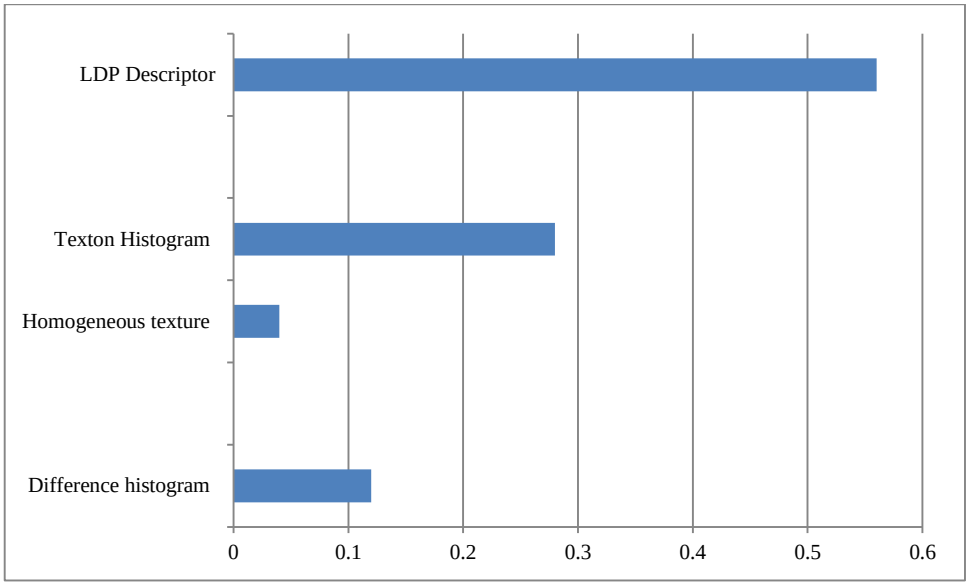

Figure 8. Comparison of Various Texture Descriptors for Fb dataset

V. CONCLUSION

The paper reaffirms that local methods perform far better than conventional global holistic approaches. The proposed approach is hybrid in nature; local pixel level appearance is encoded in LDP. Histogram provides a way to encode spatial in region relations. Finally concatenation provides the global descriptors. Results conclude effectiveness of modified Chi Square Distance especially in cases of partial occlusions or pose/illumination variations also weighted Chi Square Distance measure makes it possible to use the method for online operation. For speeding up the entire processing we can only consider the partial face region for matching purpose.

The contributions of the paper includes 1) Proposing LDP based face descriptors 2) Weighted Chi Square Distance for Spatially Enhanced Histogram Matching 3) Evaluation of the method using benchmark datasets.

VI. FUTURE WORK

Experiments carried out to cover most of the databases; however additional tests on more challenging cases would help to identify further improvements in proposed approach. As per authors of LDP, higher level LDP performs better than second order LDP which needs to be evaluated in this method also. Currently, the window size is fixed and not dependent on face orientation which degrades the recognition rate. Hence face orientation and scale should be considered as priory to make window adaptively. Matching performance can also be re-evaluated with the help of appearance based classifiers.

REFERENCES

[1]. T. Ojala, M. Pietikainen, and D. Harwood. “A comparative study of texture measures with classification based on featured distributions,”. Pattern Recognition, 29(1):51-59, 1996

[2]. T. Ahonen, A. Hadid, and M. Pietikäinen, “Face description with local binary patterns: Application to face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[3]. Baochang Zhang, Yongsheng Gao, ”Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High Order Local Pattern Descriptor,” IEEE Transactions on Image Processing, Vol 19, Feb2010.

[4]. M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[5]. K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," Journal of the Optical Society of America, vol. 14, pp. 1724–1733, 1997.

[6]. J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131–137, Jan 2004.

[7]. A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigen spaces for face recognition," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 84–91.

[8]. P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," Network – Computation in Neural Systems, vol. 7, no. 3, pp. 477–500,August 1996.

[9]. L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 775–779, 1997.

[10]. B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus Global approach," Computer Vision and Image Understanding, vol. 91, no. 1–2, pp. 6–21, 2003.

[11]. P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 10, pp. 1090–1104, Oct.2000.

[12]. W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations, "IEEE Trans. Syst., Man., Cybern. A, Syst. Humans, vol. 38,no. 1, pp. 149–161, 2008.

[13]. T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25,no. 12, pp. 1615–1618, Dec. 2003.

[14]. K. C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 5, pp. 684–698, May 2005.

[15]. P. J Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, C. Jin, K. Hoffman, J. Marques, M. Jaesik, and W. Worek, " Overview of the face recognition grand challenge," in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, 2005, pp.947-954.

[16]. T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary pattern," IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.7, pp. 971-987, Jul. 2002.

AUTHORS

Mahadeo D. Narlawar was born in Pandharkawada, Maharashtra, India in 8th June 1978. He received the Bachelor in Industrial Electronics degree from the University of Amravati, Amravati, Maharashtra, India in 2000. In July 2001, he joined Vivekanand Education Society's Polytechnic Mumbai-71 Maharashtra, India as a lecturer in Industrial Electronics Department. He is currently perusing Master degree in Electronics from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India. His research interest includes image processing, digital systems.

Jaideep G. Rana was born in Aurangabad, Maharashtra, India, in 25th November 1964. He received the Bachelor in Electronics and Telecommunication degree from the University of Dr. Babasaheb Ambedkar Marathwada Aurangabad, Maharashtra India in 1987, and the Master in Electronics degree from the University of Dr. Babasaheb Ambedkar Marathwada Aurangabad, Maharashtra India, in 1996. In 20th July 1987, he joined Jawaharlal Nehru Engineering Collage Aurangabad as a lecturer in Electronics and Telecommunication Department. He is currently working as a professor. His research interests include digital systems, image processing, and switching theory.