Signal Processing 93 (2013) 1557–1565

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Efficient face recognition using wavelet-based generalized neural network

Poonam Sharma [a,*], K.V. Arya [b], R.N. Yadav [c]

[a] Department of Computer Science & Engineering, Madhav Institute of Technology & Science, Gwalior, India
[b] Department of Information & Communication Technology, ABV-Indian Institute of Information Technology & Management, Gwalior, India
[c] Department of Electronics & Communication Engineering, Maulana Azad National Institute of Technology, Bhopal, India

ARTICLE INFO

Article history:
Received 1 February 2012
Received in revised form
8 September 2012
Accepted 16 September 2012
Available online 2 October 2012

Keywords:
Gabor wavelets
Local Gabor binary pattern
Histogram sequence
Generalized neural network

ABSTRACT

This paper presents an efficient face recognition method where enhanced local Gabor binary pattern histogram sequence has been used for efficient face feature extraction and generalized neural network with wavelet as activation function is being used for classification. In this method the face is first decomposed into multiresolution Gabor wavelets the magnitude responses of which are applied to enhanced local binary patterns. The efficiency has been significantly improved by combining two efficient local appearance descriptors named Gabor wavelet and enhanced local binary pattern with generalized neural network. Generalized neural network is a proven technique for pattern recognition and is insensitive to small changes in input data. The proposed method is robust-to-slight variation of imaging conditions and pose variations. Performance comparison with other existing techniques shows that the proposed technique provides better results in terms of false acceptance rate, false rejection rate, equal error rate and time complexity.

1. Introduction

Face recognition is a very challenging research area in computer vision and pattern recognition due to variations in facial expressions, poses and illumination. Several emerging applications, from law enforcement to commercial tasks, demand the industry to develop efficient and automated face recognition systems. Although, many researchers have worked on the problem of face recognition for many years [1,2], still several challenges need to be solved. Difference in illumination of the scene, changes in pose, orientation and expression are examples of some of the issues to be dealt carefully. Also when size of face database increases the recognition time becomes a big constraint. Several face recognition methods have been proposed during the last thirty years [1] which are classified as holistic methods [2], local feature-based methods [3] and hybrid methods [4]. In this paper, we are concentrating on the local feature based methods, and show that face recognition performance can significantly be improved by combining two of the most successful local appearance descriptors, Gabor wavelet [5] and local binary pattern (LBP) [6,7]. LBP is basically a fine scale descriptor that captures small texture details. Local spatial invariance is achieved by locally pooling (histogramming) the resulting texture codes. As LBP is very resistant to light changes, it is a good choice for coding fine details of facial appearance and texture. In contrast, Gabor features [5] encode facial shape and appearance information over a range of coarser scales.

In this work, we developed an efficient face recognition approach based on local Gabor binary pattern histogram sequence (LGBPHS) [8], which is robust to the variations in imaging conditions, and offers higher recognition efficiency as compared to the existing state-of-art methods. It has been

* Corresponding author. Tel.: +91 9826090830.
E-mail addresses: enggpoonam@yahoo.com (P. Sharma), kvarya@iiitm.ac.in (K.V. Arya), yadavrn@manit.ac.in (R.N. Yadav).

0165-1684/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2012.09.012

shown through extensive experiments and comparison with existing methods that the proposed method exhibits better efficiency. The reduced feature vectors are separately normalized and then concatenated into a single combined feature vector and classification is done by generalized neural network. Concatenation of features is done to reduce the complexity of neural network. The main constraint of the basic neural network is the high recognition time with increase in the size of database which has been improved in the recent years in several neural networks. Thus, generalized neuron model based neural network is used in this work which demonstrates better computational power [9]. Generalized mean neural (GMN) network provides fast convergence and are insensitive to small changes in input data [9].

The rest of this paper is organized as follows. In Section 2, we describe the related work done in the field of face recognition. Section 3 describes the proposed methodology. The detailed comparative analyses of experimental results have been presented in Section 4. Section 5 concludes the paper.

2. Related work

Local appearance based features are more stable to global changes such as illumination and pose variation. A large number of spatial local descriptor based algorithm have been developed in the last decade. Jahan et al. [10] have presented the implementation of face recognition using neural network (recognition classifier) on low resolution images. In this implementation, they first convert the original images is pre-processed by average filtering and lighting normalization. The pre-processed image is given to the input of neural network classifier, which uses back propagation algorithm to recognize the familiar faces. The efficiency of their method degrades drastically in the presence of pose variation as is the case of FERET database [11]. Huang et al. [4] extracted multi-pose face image features based on Gabor wavelets with different orientations and scales, then the mean and standard deviation of the filtered image were computed as features for face recognition. These features were fed up into support vector machine (SVM) for face recognition which shared successful face recognition over a wide range of facial variations in colour, position, scale, orientation, 3D pose, and expression in images from stereo-pair database. Aroussi et al. [12] presented a novel face recognition approach based on steerable pyramid (SP) decomposition, where multi-orientation information in face images was captured by computing the derivatives in different directions. LBP was applied to get local steerable pyramid binary pattern (LSPBP) feature vector. For classification, similarity between LSPBP feature vector and nearest neighbor is computed. Although the recognition accuracy was better as compared to face recognition using LBP on ORL database [13] but it was inefficient in terms of time complexity. In [14] Aroussi et al. modified the method for face recognition by dividing the SP coefficients into smaller subblocks, and the feature was extracted using mean, variance and entropy of each block. The proposed method proved to be efficient than LSPBP in terms of recognition, but for larger dataset the efficiency is reduced as all the subbands are equally considered although some subbands contains more useful information than others. Bashyal and Vinayagamoorthy [15] used Gabor filter kernels for feature extractions and principal component analysis (PCA) for lowering the dimensionality. Linear vector quantization (LVQ) was used to classify the images with less efficiency and high time complexity. Ahonen et al. [16] considered both shape and texture information to represent the face images. The face area was first divided into small regions, LBP histogram of each region is extracted, and concatenated into a single, spatially enhanced feature histogram efficiently representing the face image. The recognition is performed using a nearest neighbour classifier in the computed feature space with Chi-square density [16] as a dissimilarity measure. It is proved that in this method some of the facial regions contribute more information than others in terms of extra-personal variance. It is shown that on FERET database [11] the method is superior to PCA and elastic bunch graph matching. Drawback of this approach lies in the length of the feature vector which makes the recognition process very slow.

Shan et al. [17] applied simple LBP for representing salient micro-patterns of face images. Template matching with weighted Chi-square statistic and SVM are adopted to classify facial expressions. It proves that LBP features are robust to low-resolution images. Anam et al. [18] developed face recognition system for personal identification and verification using genetic algorithm and back-propagation neural (BPN) network. Here the extracted features are saved into memory and recognition is performed using genetic algorithm. The advantage of this technique is that no extra learning process is required, only by saving the face information of the person, and appending the person's name in the learned database completes the learning process. The extracted features are fed to the multilayer neural network, and the network is trained to create a knowledge base for recognition. BPN is less efficient in terms of time complexity. Rizon et al. [19] have developed a computational model to identify unknown faces by applying eigenfaces. This significant feature vector is used to identify an unknown face by using the back-propagation neural network that utilized Euclidean distance for classification and recognition. The classification and recognition using back-propagation neural network showed better results to classify face images with very small database, but the performance degrades with increase in database size. Sahoolizadeh et al. [20], have shown that Gabor wavelet representation of face images is an effective approach for both facial action recognition and face identification. Gabor wavelet faces are combined with extended neural network feature space classifier but exhibits poor efficiency if the face databases have uncertainties. Huang et al. [21] modelled a face image as a "histogram sequence" by concatenating the histograms obtained from all subregions. For recognition, histogram intersection is used to measure the similarity of different LGBPHS, and the nearest neighbourhood is exploited for final classification. The performance is appreciable, but it has higher computation time complexity due to very large feature vector.

Lei et al. [22] applied extended Gabor volume LBP to depict neighboring changes in spatial, frequency and

orientation simultaneously, but the feature vector provides overcomplete and redundant information. It uses weighted histogram intersection for matching. It increases the computational cost and is practically impossible to implement because the feature vector require more memory space to be stored. Zhang et al. [23] developed face recognition algorithm using linear derivative pattern to improve local spatial information. This algorithm is claimed to have better efficiency than existing variants of LBP, but the feature vector size is very high as compared to the LBP. It also provides redundant information along with the spatial information, without any description about the matching algorithm used. Lu and Zhao [24] proposed face recognition using dominant singular value decomposition (SVD) representation. In this approach, each gray scale face image is decomposed into three subsets of bases by SVD, subset of bases are selected on the basis of mean cumulative energy, used according to the discriminative and reconstructive power of the subband and nearest neighbor classifier is used for classification. The drawback of the method is higher computational complexity, and efficiency depends on the proper selection of the factor by which dominant space is to be deflated as it is sensitive to changes in illumination, pose and expression.

Zhou et al. [25] developed manifold elastic net (MEN) for dimensionality reduction with application to face recognition. The local features of the face images is encoded by part optimization of every manifold via patch alignment framework, and then unified by whole alignment. Then classification error is minimized. The main advantage is that the local geometry and discriminative information of intra-class samples is well preserved, and reduces overfitting as compared to other dimensionality reduction methods but needs to be improved in selection of optimal sparsity. Guan et al. [26] proposed an efficient and state-of-the-art optimization algorithm for Frobenius norm based NMF which approximates a non-negative matrix by the product of two non-negative matrix factorization (NeNMF). It uses Nesterov's optimal gradient method to optimally minimize one matrix factor with another fixed. It converges much faster with the use of Lipchitz constant [26] to decide the step size. The results were tested on synthetic datasets and proved its effectiveness in terms of efficiency, convergence time and clustering datasets. The same can be used for dimensionality reduction for face recognition. Thus a proper combination of feature extraction, dimensionality reduction and classification needs to be developed for face recognition. Guan et al. [27] proposed an online robust stochastic approximation-negative matrix factors (ORSA-NMF) algorithm which is trained in an incremental manner with an individual or chunks of images at a time. It converges at a rate of $O(1/\sqrt{k})$ and also reduces space complexity. It uses optimized approximation method to optimize the bases. But the efficiency depends on the selection of learning rate. The method works well for ORL database but efficiency reduces with change in pose or unconstrained environment.

Candes et al. [28] considered face recognition as a practical low rank and sparse separation problem. They used robust principal component pursuit for dimensionality reduction and for removing the errors due to change in illuminations. This method has a limitation that the number of images per subject must be large, cropped with background eliminated and aligned. Zhou and Tao [29] introduced go decomposition (GoDec) algorithm which decomposes image matrix to efficiently and robustly estimate the low rank part L which contains the real face and the sparse part S representing the shadow and light images of a matrix $X=L+S+G$ with noise G. GoDec has low computational cost, fast speed and nearly optimal error bounds as compared to robust PCA and SVD based approximation. The method works well for illumination variation and presence of noise but it is ineffective to the changes in pose. Liu et al. [30] considered face recognition as a subspace clustering problem. They proposed a low rank representation function that represents a set of face images as a linear combination of bases defined in a dictionary. They classified face images from outliers but needs images to be cropped and aligned before being applied. The efficiency is better as compared to robust PCA and sparse representation but the selection and development of a generalized dictionary is complex process. Similar to robust PCA and GoDec, MahNMF [31] estimates the low-rank part and sparse part of a non-negative matrix. They used rank-one residual iteration method for fast convergence and Nesterov's smoothing method to optimize the algorithm. This algorithm successfully handles noise and outliers in different images. Peng et al. [32] considered face recognition as a sequence of convex problem by using optimization techniques for rank optimization. This algorithm has demonstrated better accuracy and efficiency for realistic misalignments and corruptions but the efficiency reduces for the major change in pose.

3. The proposed method

In this section, we first describe the motivation and basic idea of the proposed method, then present the steps used for face recognition. In this approach, a face image is converted to a feature vector in the form of histogram sequence. Initially the face image is normalised and transformed to obtain Gabor features [33–35] in frequency domain by applying multiscale and multiorientation Gabor filters [8]. The Gabor wavelet transform provides an effective way to analyse images, and has been elaborated as a frame for understanding the orientation and spatial frequency of the image. It is a good approximation to the sensitivity profiles of neurons found in visual cortex of higher vertebrates. The Gabor filters used are defined as in [17]

$$\psi_{\mu,\nu}(z)=\frac{||K_{\mu,\nu}||}{\sigma^2}e^{\left(-||k_{\mu,\nu}||^2||z||^2/(2\sigma^2)\right)}\left[e^{ik_{\mu,\nu}z}-e^{-\sigma^2/2}\right] \quad (1)$$

where μ and ν define the orientation and scale of the Gabor filters, $z=(x,y)$ respectively, $||\cdot||$ denotes the norm operator and the wave vector $K_{\mu,\nu}=K_\nu e^{i\emptyset_\mu}$, where $K_\nu=K_{\max}/\lambda^\nu$, $\emptyset_\mu=\pi\mu/8.\lambda$ is the spacing factor between filters in the frequency domain and σ=variance of the Gabor filter image. The Gabor image is obtained by the convolution of the face image with the Gabor filter.

Let $f(x, y)$ be the face image, its convolution with a Gabor filter $\psi_{\mu,v}(z)$ is defined as follows

$$G_{\psi f}(x,y,\mu,v) = f(x,y) * \psi_{\mu,v}(z) \quad (2)$$

where $*$ denotes the convolution operator. Here Gabor filter use five frequencies $v \in\{0,\ldots,5\}$ and eight orientations $\mu \in\{0,\ldots,8\}$. Gabor features are generated by convolving the image with each of the 40 Gabor filters, then Gabor magnitude pattern (GMP) is used to generate the Gabor features. Note that, because the phase information of the transform is generally time-varying hence, only its magnitude is explored. Thus, for each Gabor filter, one magnitude value is computed at each pixel location, which results in 40 GMPs. In order to enhance the local information in the GMPs, we encode the magnitude values with LBP operator.

Ojiala et al. [6] introduced the LBP operator to summarize local gray-level structure. The operator takes a local neighbourhood around each pixel, thresholds the pixels of the neighbourhood at the value of the central pixel, and uses the resulting binary-valued image patch as a local image descriptor. It was originally defined for eight neighbours, giving 8 bit codes based on the 8 pixels around the central one. The original LBP operator [6,7] labels the pixels of an image by thresholding the 3×3 neighbourhood of each pixel $f_p(p=0,1,\ldots,8)$ with the center value $f_c$and considering the result as a binary number [8] computed as

$$S\left(f_p - f_c\right) = \begin{cases} 1, & f_p \geq f_c \\ 0, & f_p < f_c \end{cases} \quad (3)$$

Then, by assigning a binomial factor $2^p$for each$S\left(f_p - f_c\right)$, the LBP pattern at the pixel is achieved as

$$LBP = \sum_{p=0}^{7} S\left(f_p - f_c\right) 2^p \quad (4)$$

which characterizes the spatial structure of the local image texture. As an example the LBP calculation is shown in Fig. 1.

LBPs are resistant to lighting effects as they are invariant to monotonic gray-level transformations, and they have high discriminative power for texture classification [7]. However, they tend to be sensitive to noise, especially in near-uniform image regions as the threshold is exactly at f_c. The grayscale values in the grayscale images uses 8 bits for representing the brightness of any pixel leading to 2^8 different binary representations. An LBP is called *uniform* if it contains at most one 0-1 and one 1-0 transition when viewed as a circular bit string in the binary representation of gray levels. The LBP code shown in Fig. 1 is uniform as it has one 0-1 and 1-0 transitions. By observing all the 256 binary representations, we found

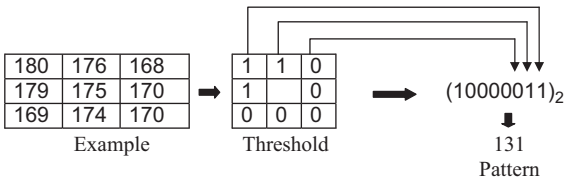

Fig. 1. Example of LBP calculation.

that only 58 combinations do satisfy the uniformity condition stated above. It is observed that nearly 90% of all image neighbourhoods are uniform. In methods based on histogramming LBP's, the number of bins can thus be significantly reduced without much loss of information by assigning all non-uniform patterns to a single bin. The enhanced LBP can be obtained by obtaining the difference in the magnitude of the centre pixel and the neighboring pixel and by calculating the binary pattern based on Eq. (3). The descriptors are then concatenated to form a global description of the face. We denote the transform result of (μ,v) GMP as$G_{\mathrm{lg}\ bp}(x,y,\mu,v)$, which composes the (μ,v) in local Gabor binary pattern (LGBP) map at location (x,y) of the image.

The histogram H for any image $f(x,y)$ for a particular frequency μ and orientation v of GMP and having the gray levels in the range $[0, L-1]$ is obtained by concatenating histogram values at all grayscale levels and is computed as follows:

$$H_{\mu,v} = \{\, h_i\} \quad \forall\, i = 0,1,\ldots,L-1 \quad (5)$$

where L is the number of grayscale levels in the face image, and h_i represents the number of pixels in image $f(x,y)$ having the grayscale value i , and can be defined for all gray level values as follows:

$$h_i = \begin{cases} \sum_{x,y} 1, & \text{If } (f(x,y) = i) \\ 0, & \text{If } (f(x,y) \neq i) \end{cases} \quad (6)$$

Now, assuming that each *LGBP Map* is divided into m regions $R_0, R_1,\ldots,R_{m-1}$. The feature vector can now be defined in terms of set of histograms and is described below:

$$\mathcal{F} = (R_0, R_1,\ldots,R_{m-1}) \quad (7)$$

where R_0 can be defined as:

$$R_0 = \{H_{\mu,v}\} \quad \forall \mu, v$$

The dimension of the feature vector in (7) is 40 times more than that of the uniform LBP feature. For example, let us assume that each LGBP image of size 177×216 is divided into $m=9$ regions. The size of a single region is 59×72. Then the enhanced LBP is applied on each region generating the feature vector of a dimension of 59. Hence, the LGBP feature vector is represented as $5\times 8\times 59\times m = 2360\, m$. The processing of such a high dimensional feature vector for matching is time consuming. Therefore, there is a need to reduce the dimension for time-effective matching process. Several global methods can be used to reduce dimension such as PCA and randomly down-sampling. However, these methods fail to take into account the spatial information. To cope with this problem, this work apply reduction dimension independently on each of the region using mapping.

The outputs of the enhanced LGBPHS for each image are $[1\times 720]$ elements (18 elements for each Gabor, i.e., two elements per sub-region) of 40 Gabor's form a row vector. The pictorial representation of the feature extraction have been given in Fig. 2. The values of the row vector are normalised to limit the values between 0 and 1. This is applied as the inputs to the neural network, and the network is trained with few images of all the subjects in the face dataset. Neural network is a very old technique,

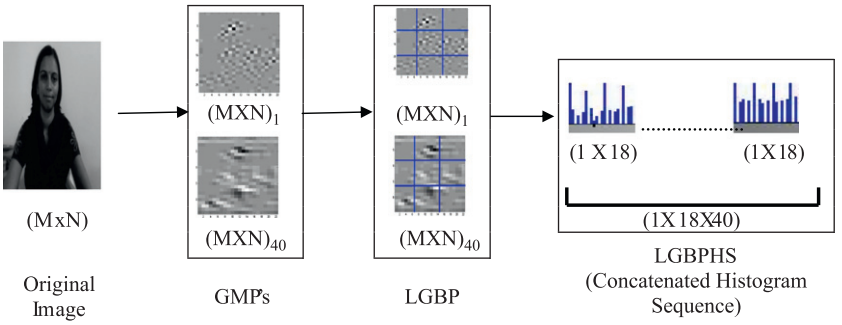

Fig. 2. Face representation using LGBPHS approach for representative image from ORL database.

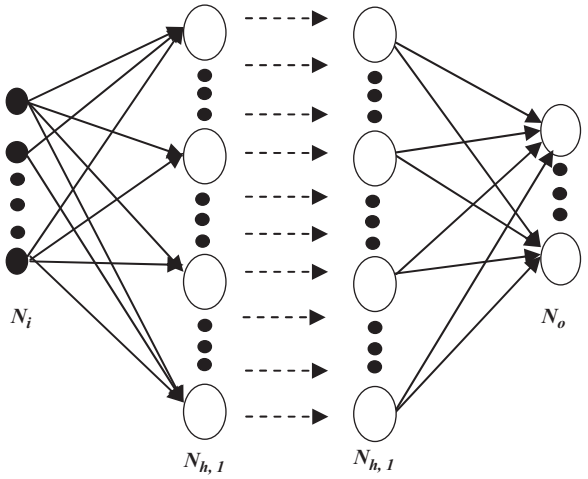

Fig. 3. The basic neural network.

but generalized mean neural network is a new advancement in the field of neural network which have shown good results for different applications of signal processing, but not been used for image processing applications. Neural networks have been extensively studied for applications related to face recognition. The higher order neurons [18] have high computational power and generalization ability but are difficult to train with explosion of the number of inputs. GMN has high computational power and fast convergence [9]. GMN is motivated by the idea of a generalized mean of the input signals. The *GM* of *N* input signals $x_j(j=1,2,\ldots,N,N\epsilon I)$ can be given as [9]

$$GM=\frac{1}{N}\left[\sum_{j=1}^{N}x_j^r\right]^{1/r} \tag{8}$$

where $r(r\,\epsilon\,\Re^+)$ is the generalization parameter and gives the various means (arithmetic mean, geometric mean and harmonic mean) depending on the value of *r*. Inspired by the importance of the flexibility of the above equation, the aggregation function of the GMN can be defined as [9]

$$y(x,w)=\left[\sum_{j=1}^{N}w_jx_j^r+w_0\right]^{1/r} \tag{9}$$

where w_j is the adaptive weight parameter corresponding to each x_j, and w_0 is bias of the neuron. From (3) we find that

$$y(x,w)=\left[\sum_{j=1}^{N}w_jx_j^r+w_0\right] \quad \text{for } r=1 \tag{10}$$

which is the output of the McCulloch–Pitts model [9].

Thus the perceptron is a special case of GMN model. The physical architecture of the GMN network used in this work is shown in Fig. 3.

In the proposed scheme, we have used a multilayer feedforward network using GMN model. The activation function used in the neural network is based on the Gabor wavelet transform as in Eq. (1) to enhance the recognition efficiency.

Algorithm 1.

For all the images available in the dataset do steps 1 to 5:

Step 1:	Extract the face from the image.
Step 2:	Apply the Gabor wavelet transform on the face image achieved in step 1.
Step 3:	Divide the Gabor wavelet image into nine subsections and apply local binary pattern on each sub section.
Step 4:	Prepare the feature vector by concatenating partial feature vectors obtained from each Gabor face image.
Step 5:	Apply feature vector obtained in step 4 to GMN network.[1]
Step 6:	Repeat steps 2 to 5 for all training images.
Step 7:	For the images not included in training set repeat steps 1 to 5 and tested using the trained neural network.

[1] The GMN network converges the network using Eq. (9) and gradient decent rule (back propogation algorithm). Activation function is the wavelet function given in Eq. (1). Training set includes the images 2 to 8 per subject chosen randomly.

Based on the experimentation the observations are made for false acceptance rate (FAR) and false rejection rate (FRR). The percentage of imposter acceptance is called as FAR. To compute FAR the proposed system is trained with the faces from the one database and tested for faces from other database and vice-versa. There have been the cases when genuine subjects have been rejected and the percentage of such rejections is termed as FRR. GMN network is trained by one database and tested on the same database to get false rejection ratio.

4. Experimental results and discussion

The experiments have been conducted using the ORL, FERET and labeled faces in the wild (LFW) [36] face database. ORL database [13] contains 10 different images (92×112) each for 40 distinct subjects. Images in this database were taken at different times varying the lighting, facial expression and facial details (glasses/no glasses). FERET database [11] consists of 14,051 eight-bit grayscale images of human heads with views ranging from frontal to left and right profiles. F_a containing 1196 frontal images of 1196 subjects is used as the gallery set, while F_b contains 1195 images with expression variations, F_c contains 194 images taken under different illumination conditions, Dup I includes 722 images of different subjects taken later in time between 1 min to 1031 days, and Dup II contains 234 images, a subset of Dup I taken at least after 18 months were used as the probe sets. Thus all the three databases provide entirely different datasets in terms of variation in pose, illumination, and expression.

Two sets of images are used, one for the training of the neural network and another set of images for testing. In this experiment 10 images of 40 subjects are used. Five images out of ten images per subject are randomly chosen to train the neural network and rest five images are used for testing the neural network. Thus, 400 images are taken out of which 200 (5 for each subject) are used for training the GMN network. The trained network is tested for the remaining 200 images.

The normalized images from a database are applied to the Gabor wavelets with 5 frequencies and 8 orientations which gives 40 Gabor jets for each image pixel. The GMP's so obtained are divided into 3×3 subregions and uniform LBP is applied to each subregion. Applying uniform LBP and then taking mean and variance of the LBP of each subregion of the image gives a row vector of 18 elements. Uniform LBP is applied to all the 40 Gabor jets, which generates a row vector 720 elements for each image. This row vector for each image is applied to train the GMN network. In GMN network, 200 input image patterns are used. The GMN network architecture used in this experiment consists of 720 input neurons, 10 hidden neurons and 40 output neurons. The threshold value of momentum and learning rate are taken as 0.9 and 0.15, respectively, by trial method. In the recognition step, the identity of human face is correctly determined when network output error is less than 10^{-6}.

It is possible to have two types of errors when operating under the verification task. To evaluate the efficiency of proposed method, we calculate its errors with different values of training and testing set and also at different error threshold. When the individual makes an errant claim as to their identity, but the returned similarity score is higher than the preset threshold. This is called a "false accept." The percentage of times that a false accept occurs across all individuals is called the "false accept rate." The second type of error is when the individual makes a proper claim as to their identity, but the returned similarity score is lower than the preset threshold. This is called a "false reject." The percentage of times that a false reject occurs across all individuals is called the "false reject rate". Subtracting this rate from 100% (100%—false reject rate) gives us the *probability of verification*. To visualize the effectiveness of the proposed method, the receiver operating curve (ROC), which represents the genuine acceptance and false acceptance rate, is shown in Fig. 4.

The experiment has also been performed for varying number of images (from 5 to 1) in the training set and the observed values of FAR and FRR and efficiency for ORL and FERET face datasets are given in Tables 1 and 2, respectively.

It explains that as the FAR increases FRR decreases. It is due to the fact that if the system is trained with more number of images of a subject, then it is more likely to accept an errant image as the image of the subject for which it is trained and vice versa. This reduces efficiency also. The efficiency is higher in FERET as compared to ORL. The ROC curves of ORL, FERET (F_a, F_b, F_c) and LFW dataset are shown in Fig. 4.

It is clear from Fig. 4 that the proposed scheme is efficient as the obtained ROC curve is approaching the ideal curve for ORL, FERET (F_a, F_b, F_c) and LFW dataset. The genuine acceptance rate reduces for F_b and F_c as there is variation in expression and illumination. Ideal ROC curve is a straight line touching genuine acceptance rate of 1 for false acceptance rate from 0 to 1.

Equal error rate (EER) is the common threshold used to evaluate the performance of a face recognition system. EER is the point on the receiver operating characteristics curve at which false-acceptance and false-rejection errors are equal. EER point can be obtained easily by drawing

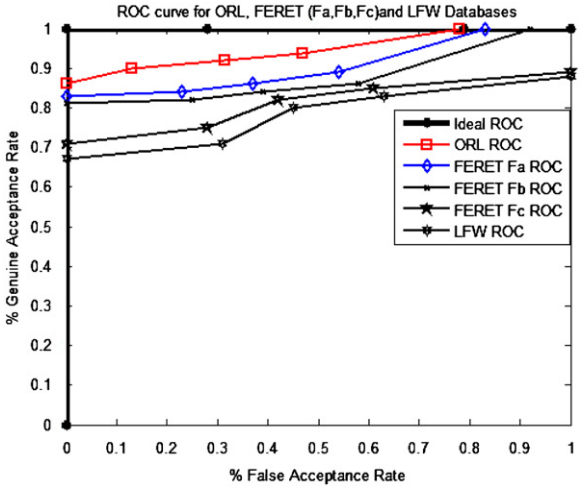

Fig. 4. ROC curve for ORL and FERET test set.

Table 1
Comparisons of FAR and FRR for ORL database with 40 subjects.

Training set/testing set	FAR	FRR	% Efficiency
5/5	3.986	0.1500	99.85
4/5	1.8125	0.4500	99.55
3/5	0.5000	0.83333	99.16
2/5	0.1500	1.3750	98.73
1/5	0	1.500	98.5

Table 2
Comparisons of FAR and FRR for FERET database with 40 subjects.

Training set/testing set	FAR	FRR	% efficiency
5/5	4.125	0	100
4/5	3.833	0.0833	99.02
3/5	3.450	0.200	99.80
2/5	3.250	0.250	99.75
1/5	3.125	0.395	99.61

Table 3
Comparisons of FAR and FRR for different error thresholds for ORL databases.

Error threshold	FAR	FRR	% efficiency
10^{-6}	0	4.95	95.05
10^{-5}	1.55	4.30	95.70
10^{-4}	2.65	4.10	95.90
10^{-3}	2.85	3.75	96.25
10^{-2}	3.07	3.55	96.45
10^{-1}	3.15	3.10	96.90
0.2	3.30	2.60	97.40
0.3	3.50	2.09	97.91
0.4	4	1.20	98.80
0.5	4.80	0.40	99.60

Table 4
Comparisons of FAR and FRR for different thresholds for FERET databases.

Threshold	FAR	FRR	% efficiency
10^{-6}	0	4.98	95.02
10^{-5}	1.32	4.25	95.75
10^{-4}	1.45	3.53	96.47
10^{-3}	1.91	2.43	97.57
10^{-2}	2.07	2.38	97.62
10^{-1}	2.85	1.83	98.17
0.2	3.45	1.28	98.72
0.3	4.40	0.83	99.17
0.4	4.68	0.26	99.74
0.5	4.86	0.04	99.96

FAR and FRR probability density against threshold. The idea of using the EER is to find FAR and FRR in order to obtain the decision of the recognition biometric systems. The values of FAR, FRR and recognition accuracy for varying threshold values (from 10^{-6} to 0.5) on ORL face database and FERET database are given in Tables 3 and 4 respectively.

By compiling the values given in Table 3, the EER for ORL database is observed to be 0.023 which occurs at equal values of FAR and FRR, i.e., 2.2%. The values of FAR and FRR against error threshold have been plotted in Fig. 5. The values of FAR and FRR against various threshold for FERET database have been represented in Fig. 6. It is observed that FAR and FRR are equal at a percentage 3.17% and thus the corresponding value of EER is 0.087.

The performance on the basis of the recognition efficiency of the proposed method has also been compared with different holistic methods such as, PCA, linear discriminant analysis (LDA) and binary LDA [12]. Table 5 presents a comparative analysis of the results obtained for

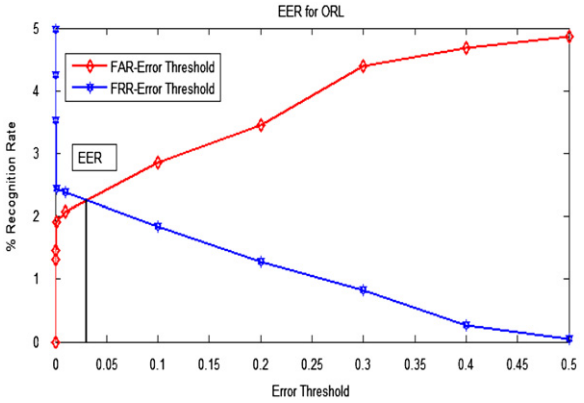

Fig. 5. EER curve for ORL database.

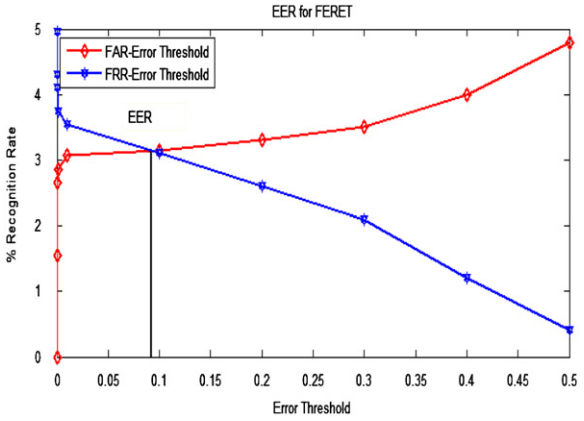

Fig. 6. EER curve for FERET database.

all methods on ORL database. The experiments of these schemes have been done on ORL database. But we have also tested them on FERET database which have slight pose variations. The results of our approach for FERET database are also given in Table 5. Aligned dataset of LFW [36] was used for evaluating the proposed method. For comparison, we computed different methods on LWF database are shown in Table 5.

Based on the experiments done on PCA [12], low resolution single neural network (LR-SNN-T) [10], MEN [25], NeNMF [26], ORSA-NMF [27] and our approach the testing time comparison of different techniques is provided in Table 6. It is clear that the proposed method outperforms the other algorithms in terms of recognition efficiency and testing time for ORL, FERET and LFW databases.

5. Conclusions

This paper proposes an efficient face recognition method, using LGBPHS and GMN neural network, which have shown better results even for slight appearance variations due to lighting and expression. The LGBPHS has proven the capability to provide the significant features of the image as input to the neural network. Thus, the efficiency for recognition has improved and time complexity has been reduced

Table 5
Comparison of recognition efficiency with other schemes.

Methods	Recognition efficiency %		
	ORL database	FERET database	LFW
PCA [12]	87.5	-.-.-.-	23.9
LDA [12]	75.6	-.-.-.-	29.7
LSPBP [12]	96	-.-.-.-	65.4
LR-SNN-T [10]	90.25	-.-.-.-	78.6
PCA (Euclidean) [37]	-.-.-.-	72.3	55.7
Automatic fiducial point detection method [37]	-.-.-.-	98	81.6
GV-LBP-TOP [22]	-.-.-.-	80.2	71.5
E-GV-LBP [22]		80.72	72.4
Gabor+LDP [23]	-.-.-.-	92.33	85.6
Block based Steerable pyramid [14]	99	94.89	79.98
MEN [25]	-.-.-.-	96.50	85.8
NeNMF [26]	92.36	95.4	82.7
ORSA-NMF [27]	94.5	96.9	85.3
Proposed method	**96.25**	**98.56**	**86.9**

-.-.-.- indicates that PCA (Euclidean), automatic fiducial point detection method and MEN have not used ORL database and PCA [12], LDA [12], LSPBP [12], LR-SNN-T[10] have not used FERET database in their experiments. All the results are evaluated on the same platform for LFW for comparison.

Table 6
Comparison of results in terms of testing time for ORL database.

Method	Testing time (s)
PCA	8.9
LR-SNN-T	38.12
MEN	8.6
NeNMF	8.2
ORSA-NMF	6.88
Proposed method	**6.17**

a lot. The efficiency of the proposed method is due to the use of multi-resolution and multi-orientation Gabor decomposition, the LBP, and the local spatial histogram modelling. Experimental results of the proposed method on the ORL, FERET and LFW face database have evidently illustrated the effectiveness and robustness of the slight variations of lighting, expression, and pose. Future efforts will be focused on expression recognition, especially for pose and occlusion variations in unconstrained environment.

Acknowledgement

This work was supported by the Department of Science and Technology, New Delhi, India under Technology System Development Scheme DST/TSG/ICT/2011/56-G.

References

[1] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: A literature survey, ACM Computing Surveys 35 (4) (2003) 399–458.
[2] M.A. Turk and A.P. Pentland, Face recognition using Eigenfaces, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991, pp. 586–591.
[3] L. Wiskott, J.M. Fellous, N. Kruger, C.V.D. Malsburg, Face recognition by elastic bunch graph matching, IEEE Transaction on Pattern Analysis and Machine Learning 19 (7) (1997) 775–779.
[4] J. Huang, B. Heisele, and V. Blanz, Component-based face recognition with 3D Morphable models, in: International Conference on Audio- and Video-Based Biometric Person Authentication, 2003, vol. 2688, pp. 27–33.
[5] T.S. Lee, Image representation using 2D Gabor wavelets, IEEE Transaction on Pattern Analysis and Machine Learning 18 (1996) 959–971.
[6] T. Ojala, M. Pietikainen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, Pattern Recognition 29 (1996).
[7] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transaction on Pattern Analysis and Machine Learning 24 (7) (2002) 971–987.
[8] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition, in: Proc. of the Tenth IEEE International Conference on Computer Vision, 2005, pp. 786–791.
[9] R.N. Yadav, N. Kumar, P.K. Kalra, J. John, Learning with generalized-mean neuron model, Neurocomputing 69 (16–18) (2006) 2026–2032.
[10] Z. Jahan, M.Y. Javed, and Q. Usman, Low resolution single neural network based face recognition, in: Proceedings of the Fourth International Conference on Computer Vision, Image and Signal Processing, 2007, vol. 22, pp. 189–193.
[11] P.J. Phillips, The FERET database and evaluation procedure for face recognition algorithm, Image and Vision Computing 16 (5) (1998) 295–306.
[12] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine, Local steerable pyramid binary pattern sequence LSPBPS for face recognition method, Signal Processing 5 (4) (2009) 281–284.
[13] ORL Database at URL: 〈www.uk.research.att.com/facedatabase.html〉.
[14] M.E. Aroussi, M.E. Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine, Local appearance based face recognition method using block based steerable pyramid transform, Signal Processing 91 (2011) 38–50.
[15] S. Bashyal, G.K. Venayagamoorthy, Recognition of facial expressions using Gabor wavelets and learning vector quantization, Engineering Applications of Artificial Intelligence 21 (2008) 1056–1064.
[16] T. Ahonen, A. Hadid, M. Pietikäinen, Face recognition with local binary patterns, Computer Vision, ECCV 2004 Proc., Lecture Notes in Computer Science 3021, Springer, 2004, pp. 469–481.
[17] C. Shan, S. Gong, and P.W. McOwan, Robust facial expression recognition using local binary patterns, in: Proceedings of IEEE International Conference on Image Processing (ICIP), 2005, vol. 2, pp. 370–373.
[18] S. Anam, Md. Shohidul Islam, M.A. Kashem, M.N. Islam, M.R. Islam, and M.S. Islam, Face recognition using Genetic algorithm and back propagation neural network, in: Proceedings of the International MultiConference of Engineers and Computer Scientists, March 18–20, 2009, vol. 1, pp. 811–814.

[19] M. Rizon, M.F. Hashim, P. Saad, S. Yaacob, Face recognition using eigenfaces and neural networks, American Journal of Applied Sciences 2 (6) (2006) 1872–1875.
[20] H. Sahoolizadeh, D. Sarikhanimoghadam, H. Dehghani, Face detection using Gabor wavelet and neural network, World Academy of Science, Engineering and Technology 35 (2008) 552–555.
[21] Z. Huang, W. Zhang, H. Huang, and L. Hou, Using Gabor filters features for multi-pose face recognition in color images, in: Proceedings of Second IEEE International Symposium on Intelligent Information Technology Application, 2008, pp. 399–402.
[22] Z. Lei, S. Liao, R. He, M. Peitikainen, and S.Z. Li, Gabor volume based local binary pattern for face representation and recognition, in: Proceedings of IEEE International Conference on Automatic Face Gesture Recognition, 2008, pp. 1–6.
[23] B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor, IEEE Transactions on Image Processing 19 (2) (2010).
[24] J. Lu, Y. Zhao, Dominant singular value decomposition representation for face recognition, Signal Processing 90 (2010) 2087–2093.
[25] T. Zhou, D. Tao, X. Wu, Manifold elastic net: a unified framework for sparse dimension reduction, Data Mining and Knowledge Discovery 22 (3) (2011) 340–371.
[26] N. Guan, D. Tao, Z. Luo, B. Yuan, NeNMF: An optimal gradient method for non-negative matrix factorization, IEEE Transactions on Signal Processing 60 (6) (2012) 2882–2898.
[27] N. Guan, D. Tao, Z. Luo, B. Yuan, Online nonnegative matrix factorization with robust stochastic approximation, IEEE Transactions on Neural Networks and Learning Systems 23 (7) (2012) 1087–1099.
[28] E. Candes, X. Li, Y. Ma, J. Wright, Robust principal component analysis? Journal of the ACM 58 (3) (2011).
[29] T. Zhou and D. Tao, GoDec: Randomized low-rank & sparse matrix decomposition in noisy cases, in: Proceedings of 28th International Conference on Machine learning, Bellevue, WA, USA, 2011.
[30] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, IEEE Transactions on Pattern Analysis & Machine Intelligence 99 (2012).
[31] N. Guan, D. Tao, Z. Luo, MahNMF: Manhattan non-negative matrix factorization, Journal of Machine Learning Research (2012).
[32] Y. Peng, A. Ganesh, J. Wright, Y. Ma, RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated Images, IEEE Transactions on Pattern Analysis & Machine Intelligence 99 (2012).
[33] C. Liu, H. Wechsler, Independent component analysis of Gabor features for face recognition, IEEE Transactions on Neural Networks 14 (4) (2003) 919–928.
[34] L. Shen, L. Bai, A review on Gabor wavelets for face recognition, Pattern Analysis and Applications 9 (2) (2006) 273–292.
[35] Al.A. Bhuiyan, C.H. Liu, On face recognition using Gabor filters, Proc. of World Academy of Science, Engineering and Technology 22 (2007).
[36] G.B. Huang, M. Ramesh, T. Berg, and E.L. Miller, Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments, University of Massachusetts, Amherst, Technical Report, pp. 07–49, 2007.
[37] M.T. Harandi, M.N. Ahmadabadi and B.N. Araabi, Multiresolution local binary pattern for face recognition, in: Proceedings of International Conference on Computer Theory and Applications (ICCTA'08), 2008.