Statistical Models for Face Recognition System With Different Distance Measures

R.Thiyagarajan thiyagucdm@gmail.com
Department of Electronics and Instrumentation Engineering
Annamalai University
Annamalai Nagar, India

S.Arulselvi arulselvi_2k3@yahoo.co.in
Department of Electronics and Instrumentation Engineering
Annamalai University
Annamalai Nagar, India

G.Sainarayanan sai.jgk@gmail.com
R&D ICTACT, Chennai
India

Abstract

Face recognition is one of the challenging applications of image processing. Robust face recognition algorithm should posses the ability to recognize identity despite many variations in pose, lighting and appearance. Principle Component Analysis (PCA) method has a wide application in the field of image processing for dimension reduction of the data. But these algorithms have certain limitations like poor discriminatory power and ability to handle large computational load. This paper proposes a face recognition techniques based on PCA with Gabor wavelets in the preprocessing stage and statistical modeling methods like LDA and ICA for feature extraction. The classification for the proposed system is done using various distance measure methods like Euclidean Distance(ED), Cosine Distance (CD), Mahalanobis Distance (MHD) methods and the recognition rate were compared for different distance measures. The proposed method has been successfully tested on ORL face data base with 400 frontal images corresponding to 40 different subjects which are acquired under variable illumination and facial expressions. It is observed from the results that use of PCA with Gabor filters and features extracted through ICA method gives a recognition rate of about 98% when classified using Mahalanobis distance classifier. This recognition rate stands better than the conventional PCA and PCA + LDA methods employing other and classifier techniques.

Keywords: Face Recognition, Statistical Models, Distance measure methods, PCA/LDA/ICA.

1. INTRODUCTION

Human face detection and recognition is an active area of research spanning several disciplines such as image processing, pattern recognition and computer vision with wide range of applications such as personal identity verification, video-surveillance, facial expression extraction, advanced human and computer interaction. The wide-range variations of human face due to pose, illumination, and expression, result in a highly complex distribution and deteriorate the recognition performance. Hence, there is a need to develop robust face recognition algorithm. Block diagram of a typical face recognition system is shown in Fig. 1. In this block diagram, the preprocessing stage is a filtering process used to reduce noise and dependence on precise registration. Classification is usually one among number of standard methods like minimum distance classifiers, artificial neural networks, etc. Feature extraction is the area that tends to differentiate. This paper addresses the feature extraction method using gabor wavelets and PCA and ICA on face images. A good survey of face recognition system is found in [1]. The methods for face recognition can be divided into two different classes: geometrical feature matching and template matching.

FIGURE 1. Block diagram of a typical face recognition system.

In the first method, some geometrical measures about distinctive features such as eyes, nose, mouth and chin were extracted. These extracted features were used for the recognition. In the second method, the face image is represented as a two-dimensional array of intensity values and this is compared to a single or several templates representing a whole face. The choice of features for the recognition with the classification method accounts for a good recognition results. These features are first extracted from the raw data using some feature extraction methods. These extracted features should be of lesser dimension and provide good discriminating ability.

In the proposed work, gabor wavelets were used for filtering in the pre processing stage. This results in gabor feature vector representing the face image. Gabor transformed face images exhibit strong characteristics of spatial locality, scale, and orientation selectivity. These images can, thus, produce salient local features that are most suitable for face recognition. The input face image is treated as a two – dimensional pattern of intensity variations. In the face based approach, this two dimensional data helps to define the face image as feature vectors. The well known method for feature extraction is the PCA method. Reconstruction of human faces using PCA was first done by Kirby and Sirovich [2] and recognition of human faces was done by Turk and Pentland [3]. The recognition method, known as eigenface method defines a feature space which reduces the dimensionality of the original data space. This reduced data space is used for recognition.

But the two most common problems with the PCA method are its poor discriminator power for data set from the same class and large computational load. These limitations with the PCA method is overcome by combining Linear Discriminant Analysis (LDA) method which is one of the most important feature selection algorithms in appearance based methods [4]. With direct implementation of LDA method, the dataset selected should have larger number of samples per class for LDA method to extract the discriminating features. With the small sample size problem, the LDA method will results in poor discriminating features. Hence, many LDA based approach first use the PCA to project an image into a lower dimensional space otherwise called as face space. Then perform the LDA to maximize the discriminatory power.

The second problem in PCA-based method is that it will de-correlates the input data using second order statistics, generating a compressed data with minimum mean squared projection error, but it will not account for the higher order dependencies in the input data projected onto them. To overcome this problem ICA is used which minimizes both second order and higher order dependencies in the input and attempts to find the basis along which the data (when projected onto them) are statistically independent[5]. ICA is used to reduce redundancy and represent independent features explicitly. These independent features are most useful for subsequent pattern discrimination and associative recall [6].

In this proposed work, the gabor wavelets are used in the pre processing stage. This gabor filtered face images results in a large gabor feature vector. This larger dataset is reduced to a smaller dimension using the PCA algorithm which extracts the principle features (top eigen vectors). The reduced dataset is further processed using the two feature extraction techniques namely LDA and ICA. This results in the most discriminating and the most independent features of the input data respectively. The classification is done using the minimum distance measure methods like ED, CD and MD methods. The results are compared and tabulated for the resulted recognition rate. In comparison with the conventional use of PCA and LDA alone, the proposed method gives a better recognition rate. The algorithm is tested with the use of ORL face database with 200 training and 200 test samples. The feasibility of the new algorithm has been demonstrated by experimental results. Encouraging results were obtained for the proposed method compared to conventional methods.

This paper is organized as follows: Section 2 deals with the basics of Gabor wavelets and Gabor features. Section 3 deals background of PCA and Eigenfaces. Section 4 deals with LDA. Section 5 details about ICA. Section 6 deals with the various types of distance measure method. The proposed method is reported in section 7. Experimental results are presented in Section 9 and finally, Section 10 gives the conclusions.

2. GABOR WAVELET

The prime motive behind the feature based approach is the representation of face image in a very compact manner, thus reducing the memory used. Such methods gains importance as the size of the database is increased. This feature based method in a face recognition system aims to find the important local features on a face and represent the corresponding features in an efficient way. The selection of feature points and representation of those values has been a difficult task for a face recognition system. Physiological studies found that simple cells present in human visual cortex can be selectively tuned to orientation as well as spatial frequency. J.G.Daugman[7] has worked on this and confirmed that the response of the simple cell could be approximated by 2D Gabor filters.

Because of their biological relevance and computational properties, Gabor wavelets were introduced into the area of image processing [8,9]. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit desirable characteristics of spatial locality and orientation selectivity, and are optimally localized in the space and frequency domains. These drawbacks were overcome by Gabor wavelets. The Gabor wavelet is well known for its effectiveness to extract the local features in an image. It has the ability to exhibit the desirable characteristics of capturing salient visual properties such as spatial localization, orientation selectivity, and spatial frequency selectivity. This method is robust to changes in facial expression, illumination and poses. Applying statistical modeling methods like PCA, LDA, ICA for the Gabor filtered feature vector, results in a crisp gabor feature vectors. In order to consider all the gabor kernels, all the features were concatenated to form a single gabor feature vector. This high dimensional gabor vector space is then reduced using the PCA. Further on applying the ICA and LDA on the reduced data, the more independent and discriminating features were obtained.

The Gabor wavelets (kernels, filters) can be defined as follows [7], [8]:

$$\Psi_{\mu,v}(z) = \left\|k_{\mu,v}\right\|^2 \exp\left(-\frac{\left\|k_{\mu,v}\right\|^2 \|z\|^2}{2\sigma^2}\right)\left(\exp\left(ik_{\mu,v}z\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \qquad (1)$$

Where µ and v in Eqn. 1 define the orientation and scale of the Gabor kernels, z = (x,y) and ||.|| denotes the norm operator, and the wave vector $k_{\mu,v}$ is defined as follows:

$$k_{\mu,v} = k_v \exp\left(i\Phi_\mu\right) \qquad (2)$$

where $k_v = k_{max} / f_v$ and $\varphi_\mu = \pi\mu/8$, k_{max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain [9]. The Gabor kernels in Eq. 1 are all self-similar since they can be generated from one filter, the mother wavelet, by scaling and rotation via the wave vector $k_{\mu,v}$. Each kernel is a product of a gaussian envelope and a complex plane wave, while the first term in Eq. 2 determines the oscillatory part of the kernel and the second term compensates for the DC value. The effect of the DC term becomes negligible when the parameter σ, which determines the ratio of the gaussian window width to wavelength, has sufficiently large values. In most cases the use of gabor wavelets of five different scales, $v \in \{0,\ldots, 4\}$ and eight orientations, $\mu \in \{0,\ldots,7\}$is used for representational bases. Fig. 2 shows the Gabor kernels at five scales and eight orientations, with the following parameters: $\sigma = 2\pi$, $k_{max} = \pi/2$ and $f = \sqrt{2}$.

2.1 Gabor Wavelet Representation

The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor kernels as defined by Eq. 3. Let I(x,y) be the gray level distribution of an image, the convolution of image I and a Gabor kernel $\psi_{\mu,v}$ is defined as follows.

$$O_{\mu,v}(z) = I(z) * \Psi_{\mu,v}(z) \tag{3}$$

where z = (x,y) and * denotes the convolution operator, and $O_{\mu,v}(z)$ is the convolution result corresponding to the gabor kernel at orientation μ and scale v. Therefore, the set $S = \{O_{\mu,v}(z)$, for $\mu \in \{0,\ldots7\}$ and $v \in \{0,\ldots4\}\}$ forms the gabor wavelet representation of the image I(z). Applying the convolution theorem, we can derive each $O_{\mu,v}(z)$ from Eq. 3 via the Fast Fourier Transform (FFT):

$$Ғ\{ O_{\mu,v}(z) \} = Ғ\{I(z)\}\, Ғ\{ \psi_{\mu,v}(z)\} \tag{4}$$
$$O_{\mu,v}(z) = Ғ^{-1}\{ Ғ\{I(z)\}\, Ғ\{ \psi_{\mu,v}(z)\}\} \tag{5}$$

where Ғ and $Ғ^{-1}$ denote the Fourier and inverse Fourier transform, respectively. Fig.2 and 3 show the Gabor wavelet representation (the real part of gabor kernels with five scales and eight orientations and the magnitude of gabor kernels at five different scales). The input frontal face image as shown in Fig 4 is preprocessed using these kernels and the resultant convolution output of the image and the kernels are as shown in Fig. 5 (Real part of the convolution output of a sample image) and Fig. 6 (Magnitude of the convolution output of a sample image). To encompass different spatial frequencies, spatial localities, and orientation selectivities into a single augmented feature vector, concatenate all these gabor representations.

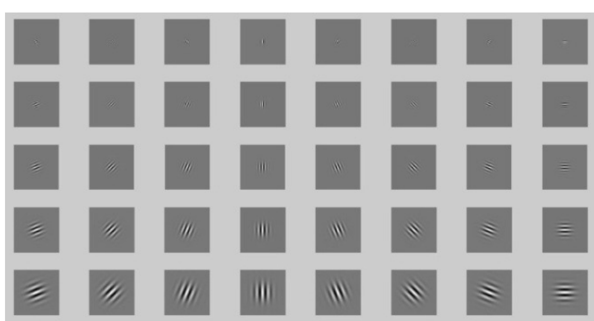


FIGURE 2. Gabor Kernels (Real part with five scales and eight orientations)

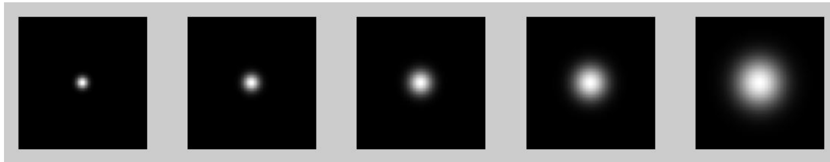


FIGURE 3. Magnitude of Gabor Kernel at five different scales.

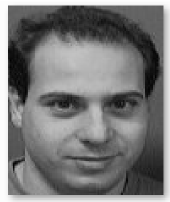


FIGURE 4. Sample frontal image

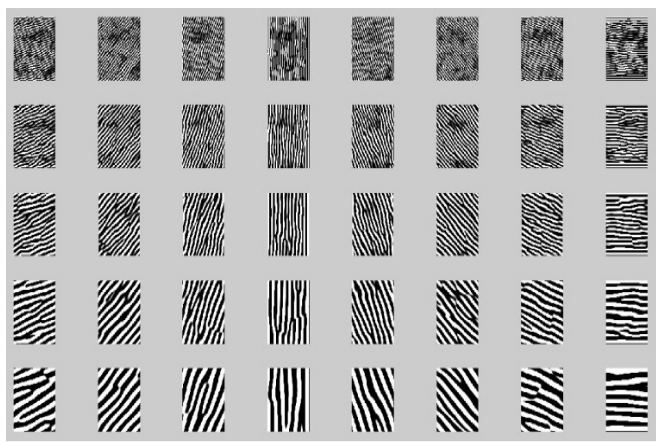


FIGURE 5. Real part of the convolution output of the sample image

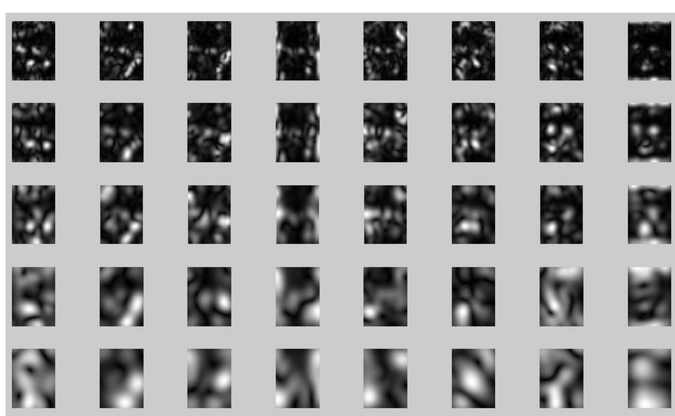


FIGURE 6. Magnitude of the convolution output of the sample image.

Before the concatenation, each $O_{\mu,v}(z)$ is down sampled by a factor ρ to reduce the space dimension. Then constructing a vector out of the $O_{\mu,v}(z)$ by concatenating its rows (or columns). Now, let $O_{\mu,v}(\rho)$ denote the vector constructed from $O_{\mu,v}(z)$ (down sampled by ρ), the augmented gabor feature vector X(ρ) is then defined as follows:

$$X(\rho) = (O(\rho)t_{0,0} \;\; O(\rho)t_{0,1} \; ..... \; O(\rho)t_{4,7}) \tag{6}$$

where t is the transpose operator. The augmented gabor feature vector thus encompasses all the elements (down sampled) of the gabor wavelet representation set, S = { Oμ,v (z) : μ ∈ {0,…..7} , v ∈ {0,…..4}} as important discriminating information. Fig. 6 shows (in image form rather than in vector form) an example of the augmented Gabor feature vector, where the down sampling factor is 64, i.e. ρ = 64.

3. PRINCIPLE COMPONENT ANALYSIS

The Principal Component Analysis (PCA) is a widely accepted and a successful technique that has been used for image compression and image recognition .The purpose of this method is to reduce the large dimensionality of the data (observed variables) to a reduced dimensionality of features (independent variables), which describe the data economically with a strong correlation between data and features [2]. These characteristic features are called eigen faces in the face recognition domain or principle components generally. In a face recognition system, PCA can transform each original image of the training set into a corresponding eigenface. An important feature of PCA is that reconstruction of original image is possible by combining the all eigenfaces from the training set. These eigenfaces are nothing but characteristic features of the training faces. Therefore summing up all the eigen faces in right proportion results in the original face image.

By using all the eigenfaces extracted from original images, exact reconstruction of the original images is possible. But for practical applications only certain part of the eigenfaces is used. Then the reconstructed image is an approximation of the original image. However losses due to omitting some of the eigenfaces can be minimized. This happens by choosing only the most important features (eigenfaces)[2].

3.1 Mathematics of PCA

A 2-D face image can be represented as 1-D vector by concatenating each row (or column) into a long thin vector.

1. Assume the training sets of images represented by $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_m$, with each image Γ(x,y) where (x,y) is the size of the image represented by p and m is the number of training images. Converting each image into set of vectors given by (m x p).
2. The mean face Ψ is given by:
$$\Psi = \frac{1}{m}\sum_{i=1}^{m}\Gamma_i \tag{7}$$
3. The mean-subtracted face is given by(Φ_i):
$$\Phi_i = \Gamma_i - \Psi \tag{8}$$
where i = 1, 2, 3….m. and A = [$\Phi_1, \Phi_2 \ldots \Phi_m$,] is the mean-subtracted matrix with size Amp.
4. By implementing the matrix transformations, the vector matrix is reduced by:
$$C_{mm} = A_{mp} \times A_{pm}^{T} \tag{9}$$
where C is the covariance matrix
5. Finding the eigen vectors V_{mm} and eigen values λ_m from the C matrix and ordering the eigen vectors by highest eigen values.
6. With the sorted eigen vector matrix, Φm is adjusted. These vectors determine the linear combinations of the training set images to form the eigen faces represented by U_k as follows

$$U_k = \sum_{n=1}^{m} \Phi_n V_{kn}$$,where , k = 1, 2…m. (10)

7. Instead of using m eigen faces, m′ eigen faces (m′<< m) is considered as the most significant eigen vectors provided for training of each individual.
8. With the reduced eigen face vector, each image has its face vector given by

$$W_k = U_k^T(\Gamma - \Psi)$$, k = 1, 2… m′. (11)

9. The weights form a feature vector given by

$$\Omega^T = [w_1, w_2, ... w_{m'}]$$ (12)

10. These feature vectors are taken as the representational basis for the face images with reduced dimension.
11. The reduced data is taken as the input to the next stage for extricating discriminating feature out of it.

4. LINEAR DISCRIMINANT ANALYSIS

The goal of PCA on choosing a lower dimensional basis is to minimize the reconstruction error. This is not the major concern in pattern recognition applications, whose goal is to maximize the recognition rate. The Linear Discriminant Analysis (LDA) provides a procedure to determine a set of axes whose projections of different groups have the maximum separation. The LDA searches for those vectors in the underlying space that best discriminates among classes (rather than those that best describe the data).

When the training data set are labeled for each identity, supervised training techniques like LDA are more profitable for feature extraction compared with methods of unsupervised learning. By applying the supervised learning, illumination variation and pose variation will be removed and retaining the identity information. The LDA provides a procedure to determine a set of axes whose projections of different groups have the maximum separation. This procedure can be described as follows.

Suppose that the sample consists of p face images from where m_j images are of individual j, for j = 1... g.

Let µ be the mean feature vector of images from individuals j. The samples between individual covariance matrix is defined as

$$S_B = \sum_{i=1}^{k} m_i(\mu_i - \mu)(\mu_i - \mu)^T \quad (13)$$

$$S_W = \sum_{i=1}^{k} (X_k - \mu)(X_k - \mu)^T \quad (14)$$

Where μ_i denotes the mean image of the class selected (individual) and µ denotes the mean image of the entire data set, and p denotes the number of images in the entire data set.

The eigenvectors of LDA are called "fisherfaces". LDA transformation is strongly dependent on the number of classes (c), the number of samples (m), and the original space dimensionality (d). It is possible to show that there are almost (c-1) nonzero eigenvectors. (c-1) being the upper bound of the discriminant space dimensionality. For Sw to be a non singular matrix, it is required to have (d+c) samples in the training set [10]. It is impossible to guarantee this condition in many real applications. Consequently, an intermediate transformation is applied to reduce the dimensionality of the image space. To this end, PCA is used. Now LDA derives a low dimensional representation of a high dimensional face feature vector of PCA. From Eqn. 15, the covariance matrix C is obtained as follows

$$C = S_W^{-1} * S_B \quad (15)$$

The coefficients of the covariance matrix gives the discriminating feature vectors for the LDA method. The face vector is projected by the transformation matrix WLDA. The projection coefficients are used as the feature representation of each face image. The matching score between the test face image and the training image is calculated as the distance between their coefficients vectors. A smaller distance score means a better match. For the proposed work, the column vectors w_i (i = 1, 2…c-1) of matrix W are referred to as fisherfaces.

5. INDEPENDENT COMPONENT ANALYSIS

The problem in PCA-based method is that it will decorrelates the input data using second order statistics, generating a compressed data with minimum mean squared projection error. But it will not account for the higher order dependencies in the input and attempts to find the basis along which the data is projected onto them. To overcome this problem, Independent Component Analysis(ICA) is used which minimizes both second order and higher order dependencies in the input and attempts to find the basis along which the data (when projected onto them) are statistically independent[11,12]. ICA of a random vector searches for a linear transformation which minimizes the statistical dependence between its components [13]. The principle feature vectors of the frontal image is obtained using PCA as given in Eqn. 12. This principle components are used as the input for the ICA and the most independent features are obtained. Since it is difficult to maximize the independence condition directly, all common ICA algorithms recast the problem to iteratively optimize a smooth function whose global optima occurs when the output vectors U are independent . Thus use of PCA as a pre-processor in a two-step process allows ICA to create subspaces of size m by m. In [14], it is also argued that pre-applying PCA enhances ICA performance by (1) discarding small trailing eigenvalues before whitening (linear transformation of the observed variable) and (2) reducing computational complexity.

Let the selected principle features be the input to ICA as given in Eqn. 12. i.e Ω be of the size m by z, containing the first m eigenvectors of the face database. The rows of the input matrix to ICA are variables and the columns are observations, therefore, ICA is performed on ΩT. The m independent basis images in the rows of U are computed as

$$U = W * \Omega T \tag{16}$$

where W is the weight matrix from the PCA. Then, the n by m, ICA coefficients matrix B for the linear combination of independent basis images in U is computed as follows

Let C be the n by m matrix of PCA coefficients. Then,

$$C = I*\Omega \text{ and } I = C * \Omega T \tag{17}$$

From U = W* ΩT and the assumption that W is invertible,

$$\Omega T = W\text{-}1 * U \tag{18}$$

Therefore,

$$I = (C* W^{-1})* U \tag{19}$$

$$= B * U$$

Each row of B contains the coefficients for linearly combining the basis images to comprise the face image in the corresponding row of I. Also, I is the reconstruction of the original data with minimum squared error as in PCA.

6. CLASSIFIER METHODS

Image classification systems analyze the numerical properties of various image features and organize data into categories. Classification includes a broad range of decision-theoretic approaches to the identification of images (or parts thereof). All classification algorithms are based on the assumption that the image in question depicts one or more features and that each of these features belongs to one of several distinct and exclusive classes. In practice, the minimum (mean) distance classifier works well when the distance between means is large compared to the spread (or randomness) of each class with respect to its mean.

The minimum distance classifier is used to classify unknown image data to classes where minimum distance between the image data and the class in multi-feature space exists. The distance is defined as an index of similarity so that the minimum distance is identical to the maximum similarity. There exist various types of distance classifier techniques. Three different types of distance classifiers namely Euclidean distance, Cosine distance and Mahalanobis distance methods were considered for the proposed approach.

Euclidian Distance $d(x, y) = kx - yk2$ (20)

Cosine Metric $d(x, y) = x.y/|x||y|$ (21)

Mahalanobis Metric $d(x, y,C) = squrt((x - y)TC^{-1}(x - y))$ (22)

where the variables x and y in the above set of equations represents the train and the test features respectively with k and K1 as constant and C is the covariance matrix of the training set.

7. PROPOSED METHOD

The block diagram of the proposed face recognition system is as shown in Fig. 7. This proposed work uses the ORL database acquired at the Olivetti Research Laboratory in Cambridge, U.K. The database is made up of 400 face images that correspond to 40 distinct subjects. Thus, each subject in the database is represented with 10 facial images that exhibit variations in terms of illumination, pose and facial expression. The images are stored at a resolution of 92 × 112 and 8-bit grey levels. Out of these 400 frontal images, 200 images (5 from each class) were considered for training and the remaining 200(remaining 5 from each class) were used for testing. Figs. 8&9 show the scaled training and test samples.

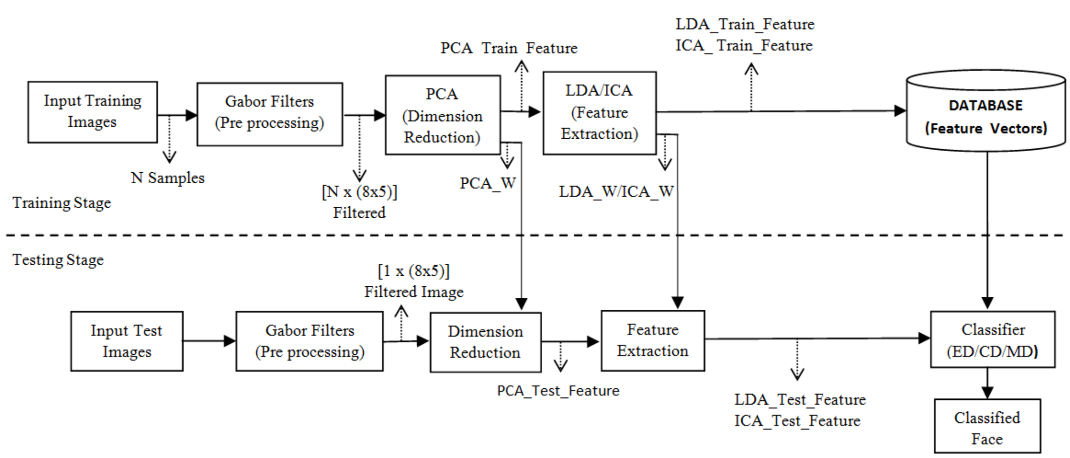


FIGURE 7. Proposed Block diagram

The system follows the image based approach and it consists of two stages namely the training and the recognition stage. In the training stage, gabor wavelet is used in the preprocessing stage which is robust to changes in illumination, pose and expression. To facilitate the convolution of the input image with gabor filters, the input images are scaled to 128x128 using a bicubic interpolation. The gabor kernels as defined in Eqn. 7 uses five different scales and eight orientations which results in 40 different filters. The input image is convoluted with the gabor kernel and the convoluted real and magnitude responses are shown in Fig. 5 and 6 respectively. This convoluted feature vector is then down sampled by a factor of $\rho = 64$. To encompass all the features produced by the different gabor kernels, the resulting Gabor wavelet features are concatenated to derive an augmented gabor feature vector.

The resulting high dimension gabor feature vector is taken as input to the next stage. This stage uses the PCA for reducing the dimension of the feature vector and extract the principle features. Here the eigen values and the corresponding eigen vectors were calculated. The eigen values are sorted in the ascending order and the top eigen values (n-1) are used and the corresponding eigen vectors were selected for representation of feature vectors (PCA_Feat). Also the weight matrix is computed as PCA_W. This eigen projection is then used as input to the next stage of ICA using the FastICA algorithm. The algorithm takes the projected data and gives the independent features ICA_Feat and the independent weight matrix ICA_W. This is used to find the test_ICA features.

FIGURE 8. Scaled Train Samples

The same procedure is applied for feature extraction using LDA. Here PCA_Feat is used as input to the LDA block. The between class (SB) and within class scatter matrix (SW) is obtained using this projection matrix as given in Eqn 13 and 14. LDA gives the projected weight matrix LDA_W, which is used to find the LDA_test_features.

FIGURE 9. Scaled Test Samples

In the recognition stage, the test samples are preprocessed as done in the training stage. With the weight matrix from the PCA stage and the test image, the test features (PCA_Test_Feature) were obtained. This PCA_Test_Feature when convoluted with ICA_W gives the test features (ICA_test). The same procedure is used to obtain the test features for LDA method (LDA_test and LDA_W). The classification stage uses different distance measure methods. Here the training and test feature vectors are considered. For a given sample image, test feature vector is found and the distance between this test feature vector and the all the training feature vectors are calculated. From the distance measure, the index with minimum distance represents the recognized face from the training set.

8. RESULT AND DISCUSSION

The effectiveness of the proposed system is tested with ORL Face database and the results are compared for the following experimental setups.

1. Gabor with PCA.
2. Gabor with PCA and LDA.
3. Gabor with PCA and ICA.

For each of the above said feature extraction method, the extracted features is used as it is, followed by classification using these three distance measures. The number of features considered for recognition is varied and the corresponding recognition rate is given in Table 1-3. The successfulness of the proposed method is compared with some popular face recognition schemes like Gabor wavelet based classification [15], the PCA method [16], the LDA method [17], and ICA based systems[18].

S.No	Distance	Features				
		40	50	66	100	200
1.	Euclidean	87.0	89.3	93.2	93.7	94.0
2.	Cosine	87.5	91.0	95.0	95.9	96..0
3.	Mahalanobis	89.3	92.5	96.1	96.9	97.0

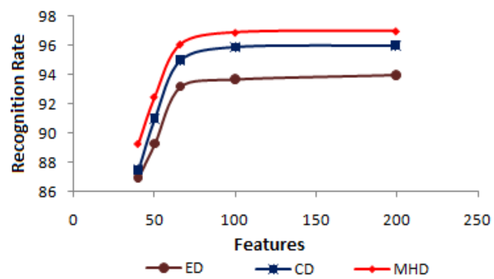
TABLE 1: Features Vs Recognition rate for different distance measure methods (Gabor wavelets and PCA method for feature extraction)

S.No	Distance	Features				
		40	50	66	100	200
1.	Euclidean	87.8	90.1	94.3	95.2	95.3
2.	Cosine	88.4	91.4	93.6	96.8	97.1
3.	Mahalanobis	90.0	93.5	94.5	97.1	98.0

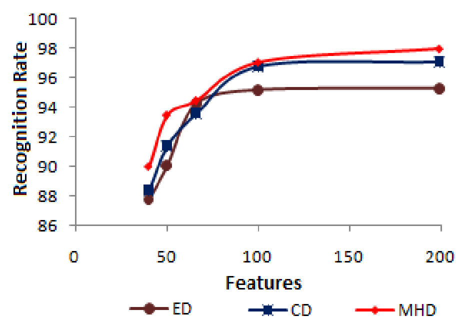
TABLE 2: Features Vs Recognition rate for different distance measure methods (Gabor wavelets and PCA method for feature extraction)

S.No	Distance	Features				
		40	50	66	100	200
1.	Euclidean	89.8	91.6	94.3	95.7	96.3
2.	Cosine	90.1	91.5	93.5	96.8	97.3
3.	Mahalanobis	90.5	92.6	94.7	96.8	98.6

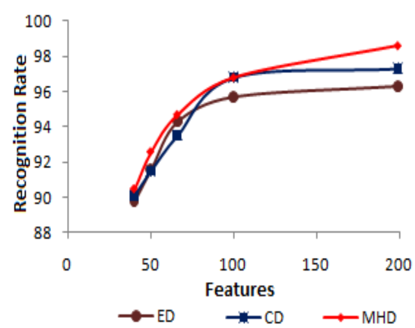
TABLE 3:. Features Vs Recognition rate for different distance measure methods (Gabor wavelets and PCA along with ICA for feature extraction)



10(a)



10(b)



10(c)

FIGURE 10. Performance characteristics of the three system considered with different classifiers.

10(a) – G + PCA   10(b) – G + PCA + LDA   10(c) G + PCA + ICA

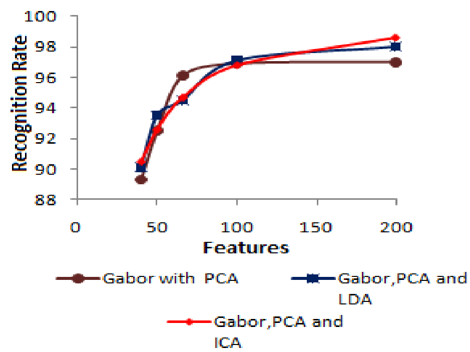


FIGURE 11. Performance characteristics of the proposed system with other methods using Mahalanobis Distance Classifier

9. CONCLUSION

In this proposed work, a Gabor feature based method is introduced for face recognition system. For different scales and orientations of the Gabor filter, the input image is filtered. Using PCA, the high dimensionality of the Gabor feature vector is reduced followed by feature extraction using ICA and LDA. Three different types of distance measure methods namely Euclidean Distance, Cosine Distance, and Mahalanobis Distance are used for classification. From the simulation results, it has been found that the recognition rate for the selected database is high with features extracted from Gabor filters based on ICA than with LDA and simple PCA methods. This recognition rate is obtained by varying the number of features at the PCA stage. The results are tabulated in the Table .1 to 3. For all these three systems, the number of PCA features assumed for further processing is 40, 50, 66, 100, and 199.

From the results, it is obvious that as the number of features selected in PCA increase, then the more discriminating features is obtained from the LDA method and the more independent features is obtained from the ICA method. This helps to increase the recognition rate percentage for the proposed system. But this in turn increases the computational load. Also from the results, when compared to other classifiers tested in the proposed system, Mahalanobis distance classifier does a better classification for the Gabor features based on ICA method.

10. REFERENCES

1. W. Zaho, R.Chellappa, P.J.Philips and A.Rosenfeld, "Face recognition A literature survey," ACM Computing Surveys, 35(4):399– 458,2003.

2. M. Kirby and L. Sirovich, "Application of the karhunenloeve procedure for the characterization of human faces", IEEE Trans. Pattern Analysis and Machine Intelligence, 12(1):103-108,1990.

3. M. Turk and A. Pentland, "*Eigenfaces for recognition*" J. Cognitive Neuroscience, 3: 71-86,1991.

4. D. L. Swets and J. J. Weng, "*Using discriminant eigenfeatures for image retrieval*", IEEE Trans. PAMI., 18(8):831-836,1996.

5. P. Belhumeur J. P. Hespanha and D. J. Kriegman, "*Eigenfaces vs. Fisherfaces; Recognition using class specific linear projection*", Pattern Recognition and Machine Intelligence, 17(9):711-720, 1997.

6. Chengjun Liu, and Harry Wechsler, "*Independent Component Analysis of Gabor Features for Face Recognition*", IEEE Transactions on Neural Networks, 14(4):919-928, 2003.

7. J. G. Daugman, "*Two dimensional spectral analysis of cortical receptive field profile*", Vision Research, 20: 847-856, 1980.

8. S. Marcelja, "*Mathematical description of the responses of simple cortical cells*" Journal Opt. Soc. Amer.,70:1297–1300, 1980.

9. M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, Wurtz R.P., and W. Konen, "*Distortion invariant object recognition in the dynamic link architecture*" IEEE Trans. Computers, 42:300–311, 1993.

10. Vitomir Struc, Nikola pavesi, "*Gabor-Based Kernel Partial-Least-Squares Discrimination Features for Face Recognition*", Informatica, 20(1):115–138, 2009.

11. M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski, "*Face recognition by independent component analysis*", IEEE Trans Neural Networks, 13:1450–1464, 2002.

12. Kresimir Delac, Mislav Grgic, Sonja Grgic , "*Independent comparative Study of PCA, ICA, and LDA on the FERET Data Set*", Wiley periodicals, 15:252-260, 2006.

13. P. Comon, "*Independent component analysis, a new concept*" Signal Processing, 36:287–314, 1994.

14. C. Liu and H. Wechsler, "*Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition*", In Proceedings of the International Conference on Audio and Video Based Biometric Person Authentication, Washington, DC, 1999.

15. A. J. O'Toole, H. Abdi, K. A. Deffenbacher and D. Valentin, "*A low- dimensional representation of faces in the higher dimensions of the space*", J. Opt. Soc. Am., A, 10:405-411,1993.

16. M. H. Yang, N. Ahuja, and D. Kriegman, "*Face recognition using kernel eigenfaces*," Proc. IEEE Int. Conf. Image Processing, 2000.

17. A. Hossein Sahoolizadeh, B. Zargham, *"A New feace recognition method using PCA, LDA and Neural Networks*", proceedings of word academy of science, Engineering and Tech, 31,2008.

18. Marian Stewart Bartlett, Javier R. Movellan, and Terrence J. Sejnowski,"*Face Recognition by Independent Component Analysis*", IEEE Trans on Neural Networks,13(6):1450-1464,2002.