# Performance Evaluation of Face Recognition Algorithms

**S. Suganya and D. Menaka**
*Department of Electronics and Communication Engineering*
*Sri Venkateswara College of Engineering*
*Irungattukottai, Tamil Nadu, India*
sugan.ss@gmail.com, menaka@svce.ac.in

***Abstract —*** Biometric-based techniques have emerged for recognizing individuals instead of using passwords, PINs, smart cards, plastic cards, tokens etc for authenticating people. Automated face recognition has become a major field of interest. In this field several facial recognition algorithms have been explored in the past few decades. A face recognition system is expected to identify faces present in images and videos automatically. The input to the facial recognition system is a two dimensional image, while the system distinguishes the input image as a user's face from a pre-determined library of faces. Finally, the output is a discerned face image.

This paper deals with the comparison of two popular dimensionality reduction algorithms such as PCA and LDA. Here, our main goal is to evaluate the performance of Principal Component Analysis and Linear Discriminant Analysis for large training data set. Finally, we concluded that LDA outperforms PCA for the large samples of training set.

***Keywords-*** Principal Component Analysis, Linear Discriminant Analysis, Eigenfaces, ORL, Face recognition.

*****

## I. Introduction

Over the past few years, the user authentication is increasingly important because the security control is required everywhere. Traditionally, ID cards and passwords are popular for authentication. Recently, biological authentication technologies across voice, iris, fingerprint, palm print, and face, etc are playing a crucial role and attracting intensive interests for many researchers. Among them, face recognition is an amicable alternative because the authentication can be completed in a hands-free way without stopping user activities. Also, the face recognition system is economic with the low-cost of cameras and computers. A face recognition system is expected to identify faces present in images and videos automatically. It can operate in either or both of two modes: (1) face verification/authentication and (2) face identification/recognition. Face verification involves a one-to-one match that compares a query face image against a template face image whose identity is being claimed. Face identification involves one-to-many matches that compare a query face image against all the template images in the database to determine the identity of the query face.Automatic face recognition by computer can be divided into two approaches [1, 2], namely, content-based and face-based. Under this face-based approach, face is matched through identifying its underlying statistical regularities. Principal Component Analysis (PCA) [7, 8, 10,] has been proven to be an effective face-based approach. Sirovich and Kirby [10] first proposed using Karhunen-Loeve (KL) transform to represent human faces. In their method, faces are represented by a linear combination of weighted eigenvector, known as eigenfaces. Turk and Pentland [8] developed a face recognition system using PCA However common PCA-based methods suffer from two limitations, namely, poor discriminatory power and large computational load. The Eigenface is the first method considered as a successful technique of face recognition. The Eigenface method uses Principal Component Analysis (PCA) to linearly project the image space to a low dimensional feature space [3], [4]. The Fisherfaces is an enhancement of the Eigenface method. The Eigenface method uses PCA for dimensionality reduction, thus, yields projection directions that maximize the total scatter across all classes. It extracts the principal components of the multi-dimensional data. The first principal component is the linear combination of the original dimensions that has the highest variability. Instead, the Fisherfaces method uses Fisher's Linear Discriminant Analysis (FLDA or LDA) which maximizes the ratio of between-class scatter to that of within-class scatter. LDA is widely used to find linear combinations of features while preserving class separability. Unlike PCA, LDA tries to model the differences between classes. PCA and LDA are very different: LDA is a supervised learning technique that relies on class labels, whereas PCA is an unsupervised technique.

## II. Basic Steps Of Face Recognition

Any image or face has size n x m pixels which require n.m dimensional space. This space is too large and needs to be reduced for better recognition which is achieved by dimensionality reduction techniques [1]. For better performance we have implemented these two algorithms with several pre-processing factors such as gray scale conversion and modified histogram equalization before recognition algorithms[5]. The aim of this paper is to study the performance of the PCA and LDA with respect to recognition percentage and recognition time. The experiments are based on ORL database. The main objective of these techniques is to enhance the discriminative information contained in the facial images.

### A. Grayscale Conversion

In order to retain as much as information of images, the color images are converted into grayscale images. Pixels in grayscale images are stored as 8–bit integer to represent color into black and white.

### B. Histogram Equalization

An image histogram is a graphical representation of the tonal distribution in a digital image. It is usually done on low contrast images in order to enhance image quality and to improve face recognition performance. The PCA technique converts a two dimensional image into a one dimensional vector. This vector is then decomposed into uncorrelated principal components (known as Eigen faces) in other words, the technique selects the features of the image which vary the most from the rest of the image. In the process of decomposition, a large amount of data is discarded since 90% of the total variance in the face is contained in 5-10% of the components. Each face image is represented as a weighted sum (feature vector) of the principal components, which are stored in a one dimensional array. A probe image is compared against a gallery image by measuring the distance between their respective feature vectors. For PCA to work well the probe image must be similar to the gallery image in terms of size, pose and illumination.

### C. Feature Extraction and Recognition

Once the face equalization has been completed, the feature extraction and recognition of the face can take place. In feature extraction, a mathematical representation called a biometric template or biometric reference is generated, which is stored in the databases. Facial recognition algorithms differ in the way they transform a face image pixels into a simplified mathematical representation in order to perform the recognition task. It is important for successful recognition that maximal information is retained.

## III. Principal Component Analysis

Principal Component Analysis (PCA) is probably the most widely used subspace projection technique for face recognition. PCA basis vectors are computed from a set of training images. As a first step, the average image is computed and subtracted from the training images, creating a set of data samples. These data samples are then arrayed in a matrix X, with one column per sample image. $XX^T$ is then the sample covariance matrix for the training images, and the principal components of the covariance matrix are computed. Typically, only the N eigenvectors associated with the largest Eigen values are used to define the subspace, where N is the desired subspace dimensionality.

### A. Eigenface Method

The Eigen face is the first successful technique of face recognition. The Eigen face method uses PCA to linearly project the image space to a low dimensional feature space called Eigen face approach. Basically, eigen face is the eigenvector obtained from PCA. In face recognition, each training image is transformed into a vector by row concatenation. The covariance matrix is constructed by a set of training images. This idea is first proposed by Sirovich and Kirby. After that, Turk and Pentland developed a face recognition system using PCA. The significant features (eigenvectors associated with large eigen values) are called eigen faces. The projection operation characterizes a face image by a weighted sum of eigen faces. Recognition is performed by comparing the weight of each eigen face between unknown and reference faces.

### B. Dimensionality Reduction

We know from linear algebra theory that for a PxQ matrix, the maximum number of non-zero eigen values that the matrix can have is min(P-1,Q-1).Since the number of training images (P) is usually less than the number of pixels (M*N), the most non-zero eigen values that can be found are equal to P-1. So we can calculate Eigen values of A'*A instead of A*A'. It is clear that the dimensions of A*A' is much larger that A'*A. So the dimensionality will decrease.

### C. Algorithm steps for PCA

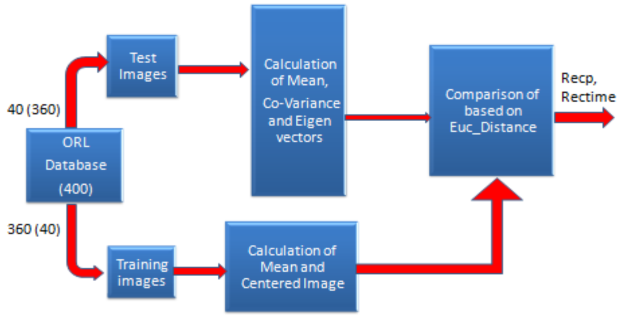


Fig.1 Block Diagram of PCA

The main idea of PCA is to find the vectors that best account for the distribution of face images within the entire image space. First divide the database into test set and training set. Here the training database contains more no of samples compared to test database. After getting the test image the following steps should be taken as follows.

1. Calculate the Average Mean of test image by subtracting test image from each image of training samples

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} Xi \qquad (1)$$

2. The pixels of each sample will be placed in column wise order to make a co-variance matrix and calculate the co-variance matrix by using the formula

$$C = \frac{1}{N}\sum_{i=1}^{N}(Xi-\overline{X})(Xi-\overline{X})^T \qquad (2)$$

3. From the Co-Variance matrix we can find the Eigenvectors and its corresponding Eigenvalues by

$$CV = \lambda V \qquad (3)$$

Where V and $\lambda$ are the Eigenvector and Eigenvalue respectively.

4. Sort the eigenvector according to their corresponding eigen values from high to low.

5. In the testing phase each test image should be mean centered, now project the test image into the same eigen space as defined during the training phase
6. This projected image is now compared with projected training image in eigen space.
7. Images are compared with similarity measures (Euclidean Distance). The training image that is closest to the test image will be matched and used to identify.

### *D. Euclidean Distance*

Euclidean distances between the projected test image and the projection of all centered training images are calculated. Test image is supposed to have minimum distance with its corresponding image in the training database.

```
temp = ( norm( ProjectedTestImage - q ) )^2;
Euc_dist = [Euc_dist temp];
[Euc_dist_min , Recognized_index] = min(Euc_dist);
```

## IV. LINEAR DISCRIMINANT ANALYSIS

Linear Discriminant Analysis (LDA) is commonly used technique for data classification and dimensionality reduction. It easily handles where the within-class frequencies are unequal and their performances has been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. The difference between LDA and PCA is that PCA does feature classification and LDA does data classification. In PCA, the shape and location of the original data sets changes when transformed to a different space whereas LDA doesn't change the location but only tries to provide more class separability and draw a decision region between the given classes.

LDA is a supervised dimensionality reduction technique. It projects high-dimensional data onto a lower dimensional space by maximizing the separation of data points from different classes and minimizing the dispersion of data from the same class simultaneously, thus achieving maximum class discrimination in the dimensionality-reduced space. Linear Discriminant analysis or Fisher faces method overcomes the limitations of eigenfaces method by applying the Fisher's linear discriminant criterion. This criterion tries to maximize the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples. Fisher discriminants group images of the same class and separates images of different classes.

As with eigen space projection, training images are projected into a subspace. The test images are projected into the same subspace and identified using a similarity measure. Unlike the PCA method that extracts features to represent face images, the LDA method tries to find the subspace that best discriminates different face classes.

### *A. Main Goal of LDA*

1. It perform dimensionality reduction while preserving as much of the class discriminatory information as possible.
2. It seeks to find directions along which the classes are best separated.
3. It takes into consideration the scatter within-classes but also the scatter between-classes.
4. It has capable of distinguishing image variation due to identity from variation due to other sources such as illumination and expression.

### B. *Algorithm for LDA*

Suppose there are C known pattern classes $w_{1,}w_{2,}\dots.w_c$ and N training samples $X = [x_j^i]$, i=1,2,…. $I_c$, j=1,2,….c is a set of samples with(mxn) dimension. $I_j$ is the number of training samples of class j and satisfies $\sum_{i=1}^{c} I_j = N$

1. Calculate the average matrix X of the N training image using

$$\overline{X} = \frac{1}{N}\sum_{j=1}^{c}\sum_{i=1}^{I_j} X_j^i \qquad (4.6)$$

2. Compute the mean $\overline{A_i}$ of $i^{th}$ class by

$$\overline{X_i} = \left(\frac{1}{I_j}\right)\sum_{i=1}^{I_j} X_j^i \qquad (5)$$

3. Calculate the image between-class scatter matrix by

$$S_b = \frac{1}{N}\sum_{j=1}^{c}(\overline{X_j} - \overline{X})(\overline{X_j} - \overline{X})^T \qquad (6)$$

4. Calculate the image within-class scatter matrix by

$$S_w = \frac{1}{N}\sum_{j=1}^{c}\sum_{i=1}^{I_j} (\overline{X_j} - \overline{X})(\overline{X_j} - \overline{X})^T \qquad (7)$$

We use best eigenvector corresponding to the maximum eigen value by

$$\frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{n} \lambda} \qquad (8)$$

5. Find the optimal projection W so that the total scatter of the projected samples of the training images is maximized. The objective function of LDA is defined by

$$W = \arg\max = \left|\frac{W^T S_b W}{W^T S_W W}\right| \qquad (9)$$

6. For test image we project the test matrix onto the eigenvectors matrix to find the new matrix of dimension (m × k): $B_j = A_T V$

7. Calculate the face distance between two arbitrary feature matrix $B_i$ and $B_j$ is defined by

$$d(B_j, B_i) = \sum_{n=1}^{k} Y_n^j - Y_{n2}^i \tag{4.13}$$

If $d(B, B_i) = \min d(B_j, B_i)$ and $B_j \in \omega_k$ identify the class and B is a test sample, then the resulting decision is B $\in \omega_k$

C. *Cosine Similarity Measures*

The Discrete Cosine Transform is the most widely used lossy image compression technique, which forms basis for international standard compression algorithm called "JPEG". Directly using face images having high redundant and correlation data causes heavy burden and complexity in terms of storage and processing speed. Therefore, a 2D-DCT is applied to an image. It segments the image into a non overlapping blocks and DCT is applied to each block separately. This results a transformed image with same dimensions as input image.

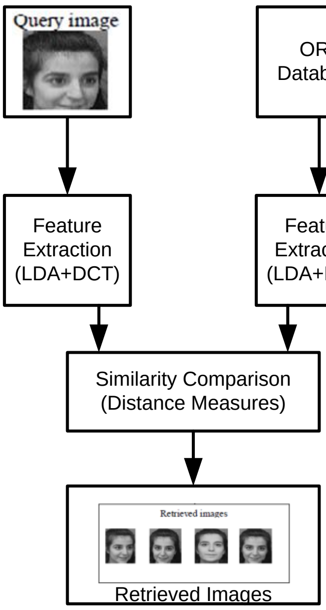


Fig.2 Block Diagram of LDA based Face Recognition

D. *Implementation Steps for LDA*

The implementation has been done with the following steps

1. All the images in the database are loaded.
2. The images are partitioned into training images and testing images.
3. The training image feature vectors are computed using LDA feature extraction and then a subspace is created using training data.
4. The computed training and test image feature vectors are matched using similarity matrix such as cosine similarity measure.
5. Then results are evaluated and performance metrics are presented.

## V. Results and Discussion

The ORL database is used for in this work to evaluate the performance of both PCA and LDA algorithms. First of all, the image pre-processing steps are carried out for improving performance of algorithms. Then by applying principle component analysis and linear discriminant analysis, face recognition is done. Further the performance of PCA and LDA based algorithms was evaluated with respect to face recognition rate and verification rate.

A. *ORL Database*

The Fig. 3 shows the some samples of ORL database. There are 40 classes with each class consisting of 10 different images. Out of 10 images, any no. of images can be selected as samples for recognition. The no. of samples can vary from 1 to 9. The LDA is tested for one fixed no. of sample and then the recognition rate is tested for different no. of samples. The no. of samples are considered as training image set and the remaining samples are considered as test image set.

Fig. 3 Sample images from ORL Database

B. *Principal Component Analysis*

Fig. 4 and Fig. 5 shows the ROC and CMC curve of PCA when no. of samples is considered as 3. The ROC plots the false accept rate (FAR) of a 1:1 matcher versus the false reject rate (FRR) of the matcher. We show that the CMC is also related to the FAR and FRR of a 1:1 matcher, i.e., the matcher that is used to rank the candidates by sorting the scores. The recognition percentage is calculated as 90% and other performance metrics observed are given below.

<u>Identification experiments</u>

- The rank one recognition rate equals (in %): 66.07%

<u>Verification/authentication experiments:</u>

- The equal error rate equals (in %): 5.03%
- The minimal half total error rate equals (in %): 4.72%
- The verification rate at 1% F AR equals (in %): 86.79%
- The verification rate at 0.1% FAR equals (in %): 66.79%
- The verification rate at 0.01% FAR equals (in %): 45.00%

International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 2 Issue: 1 135 – 140

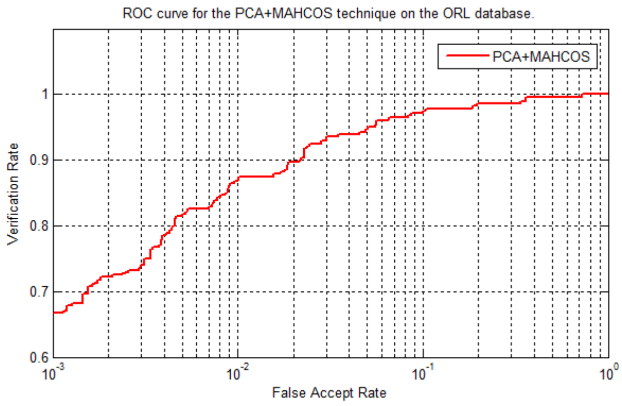


Fig.4 ROC curve for the PCA when no. of samples is 3.

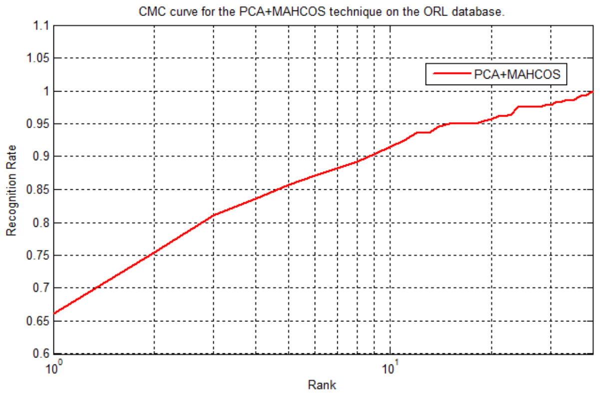


Fig.5 CMC curve for the PCA when no. of samples is 3.

The recognition percentage remains same at 97.5% when the no. of samples are 7 and above. But it improves recognition time significantly. It shown in the table that the recognition time is reduced more than half when the no. of samples are 9 instead of 7. The recognition percentage with respect to no. of samples are shown in Fig. 3.12.

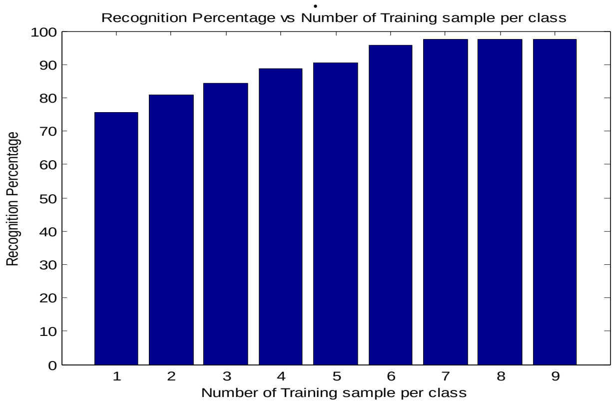


Fig.6 Recognition Percentage and Time for different no.of samples

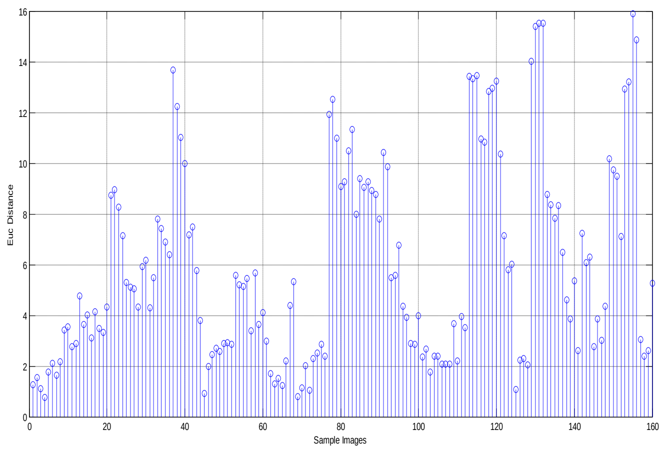


Fig.7 Euclidean distance when no. of samples = 4

### *C. Linear Discriminant Analysis*

The Cumulative Match Curve (CMC) is used as a measure of 1: m identification system performance. It judges the ranking capabilities of an identification system. The Receiver Operating Characteristic curve (ROC curve) of a verification system, on the other hand, expresses the quality of a 1:1 matcher. The ROC and CMC curves has been plotted to show the performance of LDA for face recognition on ORL database.

Fig. 6 and Fig. 7 shows the ROC and CMC curve of LDA when no. of samples is considered as 4. The recognition percentage is calculated as 90% and other performance metrics observed are given below.

Identification experiments:

- The rank one recognition rate equals (in %): 90.00%

Verification/authentication experiments:

- The equal error rate equals (in %): 3.75%
- The minimal half total error rate equals (in %): 2.95%
- The verification rate at 1% FAR equals (in %): 93.75%
- The verification rate at 0.1% FAR equals (in %): 82.08%
- The verification rate at 0.01% FAR equals (in %): 74.58%

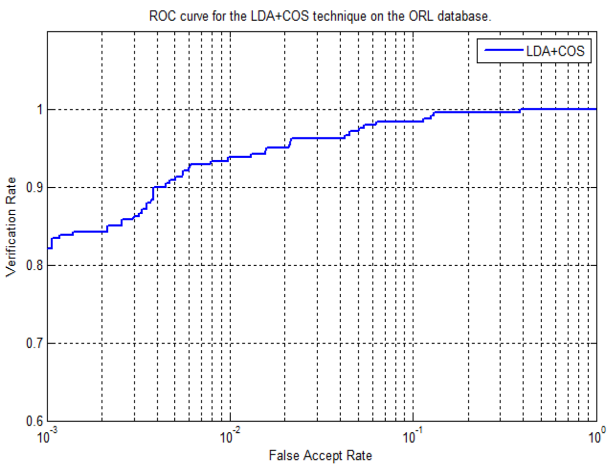


Fig.8 ROC curve for the LDA when no. of samples is 4.

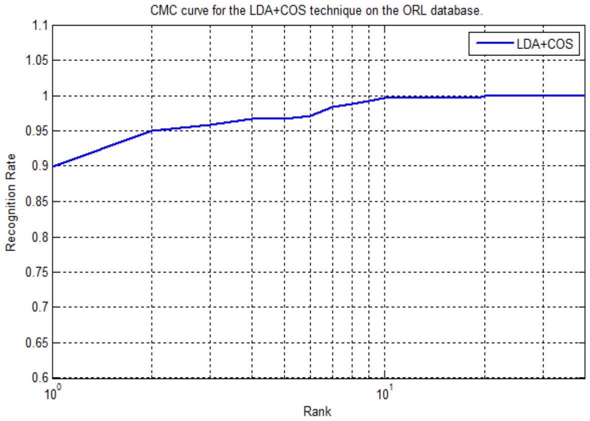


Fig.9 CMC curve for the LDA when no. of samples is 4.

The recognition percentage for different no. of samples is shown in Fig. 4.9. It clearly shows that the LDA has problem when sample size is small. With 2 samples the recognition percentage is more than the 3 samples.

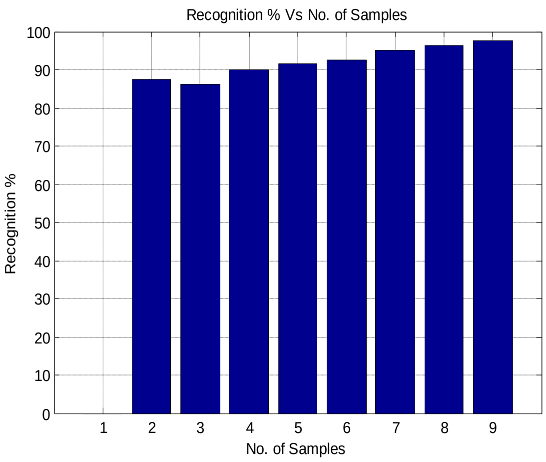


Fig.10 Recognition percentage Vs No. of samples for LDA

## VI. CONCLUSION

The major focus in this paper has been to evaluate the performance of two important face recognition algorithms namely Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). These algorithms are implemented in MATLAB and the performance is tested with ORL database. The recognition rate and verification rate of these two algorithms is mainly tested in this work. The major drawback of applying LDA is that it may encounter the small sample size problem[11]. When the small sample size problem occurs, the within-class scatter matrix becomes singular. Since the within-class scatter of all the samples is zero in the null space of $S_w$, the projection vector that can satisfy the objective of an LDA process is the one that can maximize the between-class scatter. Actually LDA outperforms PCA when training set is large whereas PCA outperforms LDA when training set is small. This is clearly shown in the results of recognition percentage. The recognition percentage is quite high with LDA compared to PCA for the same no. of samples.

## VII. REFERENCES


[1] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, Vol. 83, No. 5, 705-740, May 1995.

[2] G. Chow and X. Li, "Towards a system for automatic facial feature detection," *Pattern Recognition*, Vol. 26, No. 12, 1739-1755, 1993.

[3] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma and N. Otsu, "Face recognition system using local autocorrelations and multiscale integration," *IEEE Trans. PAMI*, Vol. 18, No. 10, 1024-1028, 1996.

[4] K. M. Lam and H. Yan, "Locating and extracting the eye in human face images", *Pattern Recognition*, Vol. 29, No.5 771-779, 1996.

[5] Ming-Hsuan Yang, David J. Kriegman and Narendra Ahuja, "Detecting Face in Images: A Survey," *IEEE Trans. On Pattern Analysis and Machine Intelligence, V*ol.24, No.1, Jan-2002.

[7] M. Kirby and L. Sirovich," Application of the Karhunen-Loeve procedure for the characterization of human faces, "*IEEE Trans. PAMI*., Vol. 12, 103-108, 1990.

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

[9] L.Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Am. A*, Vol. 4, No. 3, 519-524, 1987.

[10] Aitao Zhao, Pong Chi Yuen and James T. Kwok, "A Novel Incremental Principal Component Analysis and Its application for Face Recognitions," *IEEE Trans. On Systems, Man, and Cybernetics-Part B: Cybernatics Vol.36, No.4, Aug-2006.*