

## FACE RECOGNITION USING PCA, LDA AND ICA APPROACHES ON COLORED IMAGES

Önsen TOYGAR<sup>1</sup> Adnan ACAN<sup>2</sup>

<sup>1,2</sup> Computer Engineering Department, Eastern Mediterranean University,  
Famagusta, Turkish Republic of Northern Cyprus, Mersin 10, Turkey.

<sup>1</sup>E-mail: [onsen.toygar@emu.edu.tr](mailto:onsen.toygar@emu.edu.tr)

<sup>2</sup>E-mail: [adnan.acan@emu.edu.tr](mailto:adnan.acan@emu.edu.tr)

### ABSTRACT

*In this paper, the performances of appearance-based statistical methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) are tested and compared for the recognition of colored face images. Three sets of experiments are conducted for relative performance evaluations. In the first set of experiments, the recognition performances of PCA, LDA and ICA are demonstrated. The effect of illumination variations is evaluated in the second set, whereas input images are partially occluded in the third set of experiments. The results show that PCA is better than LDA and ICA under different illumination variations but LDA is better than ICA. On the other hand, LDA is more sensitive than PCA and ICA on partial occlusions, but PCA is less sensitive to partial occlusions compared to LDA and ICA sensitivity. That is, PCA performance is better than LDA and ICA while performance of ICA performance is better than LDA on partial occlusions.*

**Keywords:** Face Recognition, PCA, LDA, ICA

### I. INTRODUCTION

In recent years, there is a great interest of many researchers on the face recognition problem. Among these researchers are the engineers, neuroscientists, and psychophysicists studying this popular problem in different fields and in different points of view. There are several application areas of face recognition in our real life such as identification of personnel using credit cards, passport checks, entrance control, criminal investigations, etc.

Principal Component Analysis (PCA) is one of the most popular appearance-based methods used mainly for dimensionality reduction in compression and recognition problems. PCA is known as Eigenspace Projection which is based

on linearly projecting the image space to a low dimensional feature space that is called eigenspace. It tries to find the eigenvectors of the covariance matrix that correspond to the directions of the principal components of the original data.

Another powerful dimensionality reduction technique is the Linear Discriminant Analysis (LDA) which is also known as Fisher's Discriminant Analysis. LDA searches for a linear transformation such that the feature clusters are most separable after the transformation which can be achieved through scatter matrix analysis.

Linear Discriminant Analysis deals directly with discrimination between classes, whereas PCA deals with the data in its entirety for the principal

*Received Date : 15.10.2002*

*Accepted Date: 26.12.2002*

components analysis without paying any particular attention to the underlying class structure.

The third appearance-based statistical method is Independent Component Analysis (ICA) which is a special case of redundancy reduction technique and it represents the data in terms of statistically independent variables. ICA is a method for transforming an observed multidimensional random vector into components that are statistically as independent from each other as possible.

In this paper, we analyzed the comparative performance of PCA, LDA, and ICA on face recognition problem using colored images. The FastICA algorithm was used as an independent component analysis. It is one of the best known MATLAB program that implements the fast "fixed-point algorithm" for independent component analysis. In the implementation of the three methods, the colored face images were splitted on R, G, B color components. Each component was processed separately and the result of each component is compared with each other. The test image is recognized under two different conditions. In the first condition, for each component, a training image that mostly resembles to the test image is found. The second condition checks if the result of each component is below a threshold value. The test image is recognized if all component results of the first condition point to the same training image and if the second condition shows that all the results are below a threshold value. If any one of the above conditions is not satisfied, then we conclude that the test image is not recognized.

Various experiments were done using different test images which are variations of the training images. Manually, the R, G, B values were changed separately with different percentages and some test images were occluded from left, right, up and down sides. These illumination changes and occlusions were tested on PCA, LDA and ICA algorithms.

The organization of the paper is as follows. Section II reviews face recognition problem. Section III, IV and V give brief information on PCA, LDA and ICA, respectively. The face database used, the experiments done, and the

results obtained are described in section VI. Finally, section VII concludes the paper.

## 2. FACE RECOGNITION PROBLEM

Face recognition is one of the most popular problems in the field of image analysis and understanding. Identifying a person from an unknown face is usually done by comparing the unknown face with the known faces from a face database.

The interest of researchers and engineers in face recognition problem has grown rapidly in the recent years since there is a wide range of commercial and law enforcement applications on face recognition [1,2]. The increasing need for surveillance-related applications, especially due to drug traffic and terrorist activities, has a great impact on the growth of interest in the field of face recognition.

Some of the application areas of face recognition includes personnel identification of credit cards, driver's licence, passport checks, entrance control, computer access control, criminal investigations, crowd surveillance, witness face reconstruction, and ATM machines.

The people interested in face recognition problem are not only the engineers who work in the area of machine learning, but also psychophysicists and neuroscientists are studying on this problem to understand human recognition mechanisms. The studies and findings of psychophysicists and neuroscientists [1,2] help the engineers who are designing and/or implementing algorithms or systems for machine recognition of faces.

The solution to the face recognition problem is mainly divided into three parts [1]. Segmentation of faces from cluttered scenes or face detection is the first stage of the solution. Then, the feature extraction should be done from the face region and finally a decision should be taken. The result of the decision process is either recognition, verification or categorization of an unknown (test) face by comparing the face with a database of faces. However, solving the problem is not easy since there are technical problems common

to all face recognition problems such as the lack of robustness to illumination and pose variations.

In the literature, the approaches to the solution of face recognition problem are divided into three types, namely frontal, profile and view-tolerant algorithms depending on both the kind of imagery (facial views) available, and on the employed recognition algorithms. In the *frontal recognition* approach, which is the classical approach, a preprocessing step finds and extracts facial features in head-on 2D face images which then are matched against the features of a face database. The frontal images contain inherently more discrimination power than facial profiles, but the analysis of such head-on images is computationally much more complex and analytically more sensitive to variation of illumination and pose [3].

*Profile recognition* approaches as stand-alone systems, include an initial search over a large face database to index candidates who then are subject to a more accurate but computationally expensive frontal recognition. These approaches are very practical, easy to analyze and therefore they allow fast algorithms and show sufficiently high number of details to support face recognition. On the other hand, *view-tolerant recognition* approaches perform recognition in a more sophisticated fashion by taking into consideration some of the underlying physics, geometry and statistics. They employ various techniques to correct for perspective or pose-based effects due to illumination and 3D nature of the head. In addition to these, there are also hybrid approaches which combine more than one approach and try to overcome the shortcomings of the individual approaches [3].

### 3. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a dimensionality reduction technique which is used for compression and recognition problems. It is also known as Eigenspace Projection or Karhunen-Loeve Transformation [4,5,6]. PCA projects images into a subspace such that the first orthogonal dimension of this subspace captures the greatest amount of variance among the images and the last dimension of this subspace captures the least amount of variance among the images [7]. The main goal of PCA is the

dimensionality reduction, therefore the eigenvectors of the covariance matrix should be found in order to reach the solution. The eigenvectors correspond to the directions of the principal components of the original data, their statistical significance is given by their corresponding eigenvalues [6].

The steps in finding the principal components can be summarized as follows:

- Collect  $x_i$  of an  $n$  dimensional data set  $x, i=1,2, \dots, m$
- Mean correct (center) all the points: Calculate mean  $m_x$  and subtract it from each data point,  $x_i - m_x$
- Calculate the covariance matrix:  $C = (x_i - m_x)(x_i - m_x)^T$
- Determine eigenvalues and eigenvectors of the matrix  $C$ .
- Sort the eigenvalues (and corresponding eigenvectors) in decreasing order.
- Select the first  $d \leq n$  eigenvectors and generate the data set in the new representation.
- The projected test image is compared to every projected training image by using a similarity measure. The result is the training image which is the closest to the test image.

### 4. LINEAR DISCRIMINANT ANALYSIS

Linear Discriminant Analysis (LDA) is a dimensionality reduction technique which is used for classification problems. LDA is also known as Fisher's Discriminant Analysis and it searches for those vectors in the underlying space that best discriminate among classes (rather than those that best describe data as in PCA) [7,8,9].

LDA creates a linear combination of independent features which yields the largest mean differences between the desired classes. The basic idea of LDA is to find a linear transformation such that feature clusters are most separable after the transformation which can be achieved through scatter matrix analysis [9]. The goal of LDA is to maximize the between-class scatter matrix measure while minimizing the within-class scatter matrix measure [8].

The basic steps in LDA are as follows:

- Calculate within-class scatter matrix,  $S_w$ :

$$S_w = \sum_{j=1}^C \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T$$

where  $x_i^j$  is the  $i$ th sample of class  $j$ ,

$\mu_j$  is the mean of class  $j$ ,

$C$  is the number of classes,

$N_j$  is the number of samples in class  $j$ .

- Calculate between-class scatter matrix,  $S_b$

$$S_b = \sum_{j=1}^C (\mu_j - \mu)(\mu_j - \mu)^T$$

where  $\mu$  represents the mean of all classes.

- Calculate the eigenvectors of the projection matrix

$$W = \text{eig}(S_w^{-1} S_b)$$

- Compare the test image's projection matrix with the projection matrix of each training image by using a similarity measure. The result is the training image which is the closest to the test image.

## 5. INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis (ICA) is a statistical method for transforming an observed multidimensional random vector into its components that are statistically as independent from each other as possible [11]. ICA is a special case of redundancy reduction technique and it represents the data in terms of statistically independent variables.

ICA of a random vector consists of searching for a linear transformation that minimizes the statistical dependence between its components [13]. The goal of ICA is to provide an independent image decomposition and representation. In other words, the goal is to minimize the statistical dependence between the basis vectors. ICA method can be distinguished from other methods since it looks for components that are both statistically

independent and nongaussian [15]. An important principle of ICA estimation is the maximum nongaussianity. The idea is that according to the central limit theorem, sums of nongaussian random variables are also closer to gaussian. Therefore, if we take a linear combination of the observed mixture variables (which, because of the linear mixing model, is a linear combination of the independent components as well), this will be maximally nongaussian if it equals one of the independent components [15].

Compared with the other statistical methods, ICA provides a more powerful data representation than PCA because PCA considers the second order moments only and it uncorrelates data, while ICA accounts for higher order statistics and it identifies the independent source components from their linear mixtures [14].

The basic steps to derive the independent components are as follows:

- Collect  $x_i$  of an  $n$  dimensional data set  $x$ ,  $i=1,2, \dots, m$

- Mean correct all the points: Calculate mean  $m_x$  and subtract it from each data point,

$$x_i - m_x$$

- Calculate the covariance matrix:  $C = (x_i - m_x)(x_i - m_x)^T$

- The ICA of  $x$  factorizes the covariance matrix  $C$  into the following form:

$$C = F \Delta F^T \quad \text{where } \Delta \text{ is a diagonal real positive matrix.}$$

- $F$  transforms the original data  $x$  into  $Z$  such that the components of the new data  $Z$  are independent:  $X = F Z$

ICA transformation  $F$  can be derived by the algorithms such as Fixed-Point Algorithm, Natural Gradient Algorithm, Extended Bell-Sejnowski Algorithm and EASI Algorithm [17].

## 6. EXPERIMENTS AND RESULTS

The experiments were done using PCA, LDA and ICA for colored face images. The images were obtained from Libor Spacek's Collection of Facial Images [10]. This database includes 7900 colored images of faces of 395 individuals. Each individual has 20 image samples in the database. The database consists of male and female images

of various racial origins. The images are mainly of first year undergraduate students between 18-20 years old, but some older individuals are also present. Some of the individuals have glasses and some of the male individuals have beards. The image format is 24-bit color jpeg, in other words, 200x180 array of pixels and each pixel is represented by 24 bits of RGB color values. The images were recorded with an S-VHS camcorder camera and the lighting is artificial, mixture of tungsten and fluorescent overhead.

Three sets of experiments were conducted. In these experiments, general performance of the three algorithms, the effect of illumination and partial occlusions was studied on different number of images. Besides these three sets of experiments, the effect of increasing the number of training images was also investigated. According to this, the three sets of experiments were conducted firstly with 25 colored face images (five individuals with five different samples) and then with 50 images (ten individuals with five different samples). The colored face images were splitted into R, G, B color components and each color component was processed separately and then their results were combined together to obtain the actual recognition results. The same training face images were used in PCA, LDA and ICA methods. The methods were implemented in the same way as described in sections III, IV and V. Euclidean distance was used as a similarity measure to compare the projections of the test images with the projections of the training images both in PCA, LDA and ICA methods. These three set of experiments are described below in detail.

### Experiment I

The training face images and all the other face images in the database were tested using PCA, LDA and ICA. For an accurate performance analysis of PCA, LDA and ICA methods, the same training and testing images were used for all three techniques.

The experiment was conducted with 25 and 50 training images separately to see the effect of using different number of training images. The results demonstrating the recognition performance of the three techniques are presented in Table 1.

In the first set of experiments using 25 training images, among 100 tests using PCA, LDA and ICA, 93 of them were successful with PCA, 85 with LDA and 36 with ICA. That means 93% of the images were exactly recognized by PCA whereas the success rate of LDA is 85% and the success rate of ICA is 36%. In this category of experiments, PCA recognition rate is better than the recognition rate of LDA and ICA, and LDA's performance is better than ICA. However, the second set of experiments, with training images increased to 50, demonstrates that different success percentages with respect to the first set of experiment percentages were obtained. The results show that whenever the number of training images is increased, the thresholds must be increased to obtain better results. Whenever the thresholds are changed in PCA and LDA for 50 training images, the results obtained become the same as the results of 25 training images. But for ICA, changing the thresholds for 50 training images does not affect the results. That is, the results of 50 images are all same and worse than the results of 25 images.

**Table I.** Experiment I Results: PCA, LDA and ICA Performance

Number of Training Faces	Method	Threshold	Success Rate (Percentage)
25 (5x5)	PCA	3000	93
	LDA	500	85
	ICA	60	36
50 (10x5)	PCA	3000	79
		4000	89
		5000	93
	LDA	500	45.5
		800	67
		1000	79.5
		1200	85
	ICA	60	32
		100	32
		500	32

So increasing the number of training images does not have a great impact on the performance of PCA and LDA, but for ICA, the performance decreases a little bit.

## Experiment II

In the second category of experiments, different R, G, B variations of the original training and test images were carried out to test the robustness of PCA, LDA and ICA to partial illumination variations. The R, G, B values of images were increased by 20% and 50% separately and the results were recorded as shown in Table 2.

This experiment was done with 25 and 50 training images separately. For 25 training images, the results are summarized below. It is seen that 20% increase in Red, Green and Blue values does not have a great impact on PCA and LDA performances compared to the experiments on the original images. Increasing Red and Green values have similar effects on the results for both PCA and LDA methods. However, for the values of Blue color, LDA performance decreases about 10-23 % compared to Red and Green color performance. On the other hand, the results of ICA shows that increasing the R,G, B values by 20% decreases the success rate by 16%.

When R, G, B values are increased by 50%, the recognition rates of PCA decrease by at least 51% compared to the results of 20% increase. The success rates of LDA are decreased by at least 42%. However, the success rates of ICA are either the same as the cases of 20% increase or the rates are increased by 1-13% when R,G,B values are increased by 50%.

Whenever the number of training images was increased from 25 to 50, the success rates are decreased compared to the original rates. An increase of 20% on R,G,B values does not have a great impact on PCA, but for LDA and ICA the situation is different. However, the success rates of LDA for Red and Blue values are increased by 2.5-17.5%. But for the Green values, the rate decreases by 42%. On the other hand, ICA success rates are all decreased by 10% for Red, Green and Blue values.

**Table II.** Experiment II Results: R,G, B Variations with 25 and 50 Training Faces

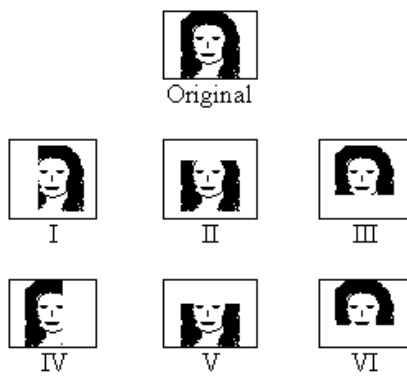
Method	Color Component	20% Increase		50% Increase	
		25 images	50 images	25 images	50 images
PCA	R	93	92	42	6.5
	G	91	92	20	6.5
	B	89	91.5	22	15
LDA	R	75	77.5	0	11
	G	85	43	0	1
	B	62	79.5	20	5.5
ICA	R	20	10	20	10
	G	20	10	21	9.5
	B	20	10	33	10

Whenever the R,G,B values are increased by 50%, PCA success rates decrease by 7-35.5%. LDA rates for Blue component decreases by 14.5% but for Red and Green values, the rates increase by 1-11%. ICA success rates decrease by 10-23%.

All these results show that, in general, increasing the illumination or in other words R,G,B values by 20% on PCA, LDA and ICA does not have a great impact on success rates, but 50% increase decreases the success rates effectively.

## Experiment III

Various experiments were done with partial occlusion of faces. Partial occlusion of original faces was obtained for six different cases. The training face images were cut from left, right, up and down parts and six different databases were formed with these training images. The six different partial occlusions applied are shown in Figure 1.



**Figure 1.** Face Images of Six Partial Occlusion Cases

Case I : The left part of the image is cut vertically until the left ear of the face is lost.

Case II : The upper part of the image is cut horizontally until the hairs are lost, but the forehead will be included.

Case III : The bottom part of the image is cut horizontally until the bottom edge of the chin is reached.

Case IV : The right part of the image is cut vertically until the right edge of the right eye is reached.

Case V : The upper part of the image is cut horizontally until the upper part of the eyebrows is reached.

Case VI : The bottom part of the image is cut horizontally until the bottom edge of the lip is reached.

The experiments were done with 25 and 50 images for each case using PCA, LDA and ICA. The recognition results, using the closest distance image measure are given in Table 3. Using 25 training images with partial occlusions show that PCA is not very sensitive to partial occlusion as LDA and ICA. In most of the cases, PCA estimates the correct face, but LDA is more sensitive to occlusion and it may not even estimate the correct face. ICA is better than LDA, but the results show that ICA is also sensitive to partial occlusions. Whenever the number of training images is increased to 50, PCA success rates are decreased but according to LDA and ICA rates, PCA is still less sensitive to occlusions. In general, the sensitivity of the three algorithms on partial occlusions does not change whenever the number of training images is increased. PCA's success rates are better than that of LDA, and ICA is better than LDA on partial occlusion results.

**Table III. Experiment III Results : Estimates with Partial Occlusions**

Method	No. of Training Faces	Case I Percentage	Case II Percentage	Case III Percentage	Case IV Percentage	Case V Percentage	Case VI Percentage
PCA	25	96	60	100	100	32	100
	50	33	37	90	58	21	71
LDA	25	0	0	0	20	0	0
	50	4	1	7	0	1	1
ICA	25	28	24	4	20	20	20
	50	10	20	2	10	20	10

## 7. CONCLUSIONS

A comparative performance analysis of Principal Component Analysis, Linear Discriminant Analysis and Independent Component Analysis techniques was conducted on face recognition using colored images. The performances were obtained using different number of training images and three sets of experiments were employed for relative performance evaluations of PCA, LDA and ICA methods. In the first set of experiments, the recognition performances of PCA, LDA and ICA were demonstrated on the original colored images. The effect of illumination variations were demonstrated in the second set of experiments by increasing the R, G, B color component values. The input images were partially occluded in the third set of experiments.

The results of the experiments show that PCA is better than LDA and ICA in general and under different illumination variations. It is also demonstrated that LDA is more sensitive than PCA and ICA on partial occlusions, but PCA is less sensitive to partial occlusions compared with ICA sensitivity. That is, PCA's success rates are better than that of LDA and ICA on partial occlusions, and ICA is better than LDA on partial occlusion results.

On the other hand, increasing the number of training images does not have a great impact on PCA and LDA in general, but for ICA, the performance decreases. For the illumination changes and partial occlusions, increasing the number of training images decreases the performance rates. The reason of this performance decrease may be because of the abundance of the training images or the difference between the samples of the training and test images.

## REFERENCES

- [1] Zhao W., Chellappa R., Rosenfeld A., and Phillips P. J.. "Face recognition: A literature survey" *Technical Report CAR-TR-948*, CS-TR-4167, N00014-95-1-0521, October 2000.
- [2] Chellappa R., Sirohey S., Wilson C.L., Barnes C.S., "Human and Machine Recognition of Faces: A Survey", in *Proceedings of IEEE*, May 95, Vol. 83, pp 705-740.
- [3] Fromherz T., Stucki P., Bichsel M., "A Survey of Face Recognition", *MML Technical Report*, No 97.01, Dept. of Computer Science, University of Zurich, 1997.
- [4] Turk M. and Pentland A., "Face recognition using eigenfaces", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, 1991.
- [5] Belhumeur N., Hespanha J. and Kriegman D., "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *Proc. ECCV*, 45--58, 1996.
- [6] Calvo R., Partridge M., and Jabri M., "A Comparative Study of Principal Components Analysis Techniques", *In Proc. Ninth Australian Conf. on Neural Networks*, Brisbane, QLD., pp. 276-281, 1998.
- [7] Yambor W.S., "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms", *Technical Report CS-00-103*, Computer Science Department, Colorado State University, July 2000.
- [8] Martinez A.M. and Kak A.C., "PCA versus LDA", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No.2, pp. 228-233, 2001.
- [9] Yang J., Yu Y., Kunz W., "An Efficient LDA Algorithm for Face Recognition", *The Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV2000)*, 2000.
- [10] Spacek L., "Description of Libor Spacek's Collection of Facial Images", 1996, online: <http://cswww.essex.ac.uk/allfaces/index.html>
- [11] Hyvärinen, A., "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, pp. 626-634, 1999.
- [12] Giannakopoulos, X., Karhunen, J. & Oja, E., "A Comparison of Neural ICA Algorithms Using Real-World Data", *International Joint Conference on Neural Networks (IJCNN'99)*, Washington, D.C., USA, July 10-16, 1999.



- [13] Comon P., "Independent component analysis. a new concept ?", *Signal Processing*, Elsevier, vol. 36, pp. 287--314, Special issue on Higher-Order Statistics, April 1994.
- [14] Liu C., Wechsler H., "Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition", *Second Int'l Conf. on Audio- and Video-based Biometric Person Authentication*, AVBPA'99, Washington D. C., March 22-24, 1999.
- [15] Hyvarinen A., Karhunen J., Oja E., *Independent Component Analysis*, John Wiley & Sons, 2001.
- [16] Baek K., Draper B.A., Beveridge J. R. and She K., "PCA vs.ICA: A Comparison on the FERET Data Set", *International Conference on Computer Vision, Pattern Recognition and Image Processing in conjunction with the 6th JCIS*, Durham, North Carolina, March 8-14, 2002.
- [17] Giannakopoulos X., "Comparison of Adaptive Independent Component Analysis Algorithms". *Master's thesis*, Helsinki Univ. of Technology, Finland, March 1998.



**ADNAN ACAN** received his B.Sc, M.Sc, and Ph.D. degrees form Middle East Technical University in Electrical and Electronic Engineering Department, Ankara, Turkey. He served as a visiting scholar at The Ohio State University in Electrical Engineering Department for one year. Since 1995, he is an assistant professor in Computer Engineering Department of Eastern Mediterranean University, North Cyprus. His current research interests include soft computing, artificial life, graph algorithms, and VLSI layout design.



**Önsen TOYGAR** received the B.S. and M.S. degrees in computer engineering from Eastern Mediterranean University, Famagusta, Turkish Republic of Northern Cyprus, in 1997 and 1999, respectively. She is currently pursuing a Ph.D. degree in computer engineering from Eastern Mediterranean University. Her current research interests include image processing, pattern recognition, neural networks, fuzzy systems and genetic algorithms.