Principle Component Selection for Face Recognition Using Neural Network

Ibrahim A. Saleh
College of Computer Sciences and Mathematics
University of Mosul, Iraq

ABSTRACT

Face Recognition is an emerging field of research with many challenges such as large set of images,. Artificial Neural Network approach is one of the simplest and most efficient method to overcome these obstacles in developing a system for Face Recognition.. This research deals with both face extraction and recognition, Firstly, Eigenfaces are eigenvectors of covariance matrix, representing given image space. Any new face image can then be represented as a linear combination of these Eigenfaces which can be found by Principal Component Analysis (PCA) for face extraction,and by Recurrent (Time Cycling) Back Propagation artificial neural network for face recognition. The whole system was performed by training using 120 color images (40 human faces with 3 poses) and testing using 40 color images. The images were taken from Collection of Facial Images: Faces95 by Computer Vision Science Research Projects. The results indicated that the proposed method lends itself to good extraction and classification accuracy relative to existing techniques.

Keywords: Face Recognition, Back propagation Artificial Neural Network.

<div dir="rtl">

اختيار المكونات الاساسية لتمييز الوجه باستخدام الشبكة العصبية

إبراهيم صالح

كلية علوم الحاسوب والرياضيات، جامعة الموصل

تاريخ القبول: 2008/12/04   **تاريخ الاستلام: 2008/11/12**

الملخص

يعد تمييز الوجه احد المجالات البحثيه التي تواجه الكثير من العقبات وتشكل تحدي لدى الكثير من الباحثين والشبكات العصبيه الاصطناعية تعتبر واحده من الطرق الكفوءه للتغلب على هذه العقبات في مجال تمييز في مجال تمييز الوجه يتعامل البحث مع كلا من المكونات الاساسيه (Principal Component)PCA Analysis (PCA، ومن ثم تمييز الوجه باستخدام الشبكه العصبيه الاصطناعيه (Recurrent (Time Cycling) Back Propagation ANN).حيث تم اختبار كفاءة النظام المقترح باستخدام قاعدة البيانات (faces95) التي تتكون من مجموعتين، المجموعة الاولى مجموعة التدريب بواقع (120) صوره ملونه بمعدل (40) صوره لكل وجه بثلاث اوضاع ، اما المجموعة الثانية هي مجموعة الاختبار التي تضم (40) صورة .النتائج التي حصلنا عليها من النظام المقترح لتمييز الوجه كانت جيده. حيث ان نتائج التجارب التي انجزت على النظام المقترح لتمييز الوجه تؤكد ان الطرق المستخدمه لاستخلاص وتمييز الوجه جيده ودقه تصنيفها عالية مقارنة بالتقنيات المتاحه حاليا.

الكلمات المفتاحية: تمييز الوجه، الشبكة العصبية الاصطناعية ذات الانتشار العكسي.

</div>

1. Introduction

Face recognition is a long standing and well studied problem in computer vision. In their recent work has proposed a recognition strategy using interest points extracted from the detected faces. The features of an interest point used in this strategy are Lowe.s SIFT features [1]. The faces are represented using a set of keypoints; and then a matching algorithm is applied to find the similar faces in the test data using a few

training faces.Face recognition may seem an easy task for humans, and yet computerized face recognition system still can not achieve a completely reliable performance. The difficulties arise due to large variation in facial appearance, head size, orientation and change in environment conditions. Such difficulties make face recognition one of the fundamental problems in pattern analysis. In recent years there has been a growing interest in machine recognition of faces due to potential commercial application such as film processing, law enforcement, person identification, access control systems, also Face recognition system could be applied to: Airport surveillance, Private surveillance, Access control for PCS in a corporate surveillance, Added security for ATM transactions, Mugshot matching for law enforcement agencies, Improve Human-Computer interface[6].

There are many methods for face recognition. These methods are namely correlation, Eigenface methods, Template matching, Bunch graph matching [8]. Template matching is represented as a two-dimensional intensity value, which is compared using a suitable metric such as Euclidean distance with a single template representing the whole face. This technique is effective only when the test images have the same scale, orientation, as training images. But this technique is cumbersome and time consuming and not at all robust. Elastic Bunch graph matching method gives appreciable results for less distortion invariant object recognition, if data base size is moderate. The correlation method is the simplest method for image classification, where the test set is classified by assigning it to the label of the closest point in the learning set. Here distances are measured in the image space. This technique has several disadvantages, first is, if the trained and test images are taken under varying moderate lighting conditions, then the corresponding points in the image may not be tightly clustered. Secondly, it requires large storage and is computationally more expensive. Hence an alternative method for dimension reduction scheme is used. The most commonly used technique for dimension reduction is Principal Component Analysis (PCA), which chooses a dimension reducing linear projection that maximizes the scatter of all projected samples. [8]. Techniques based on Principal Components Analysis (PCA) popularly termed **eigenfaces,** have demonstrated excellent performance [12].

This research introduces a simple algorithm for face recognition, that satisfies the requirements and also significantly outperforms PCA-based methods and Recurrent (Time Cycling) Back Propagation Artificial Neural network on face recognition datasets. In this research, we propose a human face recognition system that can uses available information and extracts more characteristics for face classification purpose by extracting feature domains from input images. In this paper Principal Component Analysis (PCA) feature domains have been used for extracting features from input images. which produce the best result for human face recognition. Finally Recurrent (Time Cycling) Back Propagation Artificial Neural network is used as the classifier.

2. Face Recognition Design

The face recognition system has been designed to perform recognition on images. Figure (1) presents a block diagram of the face recognition system that includes three major tasks[6,7,10]:

• **Face Detection** : The ultimate goal of the face detection is finding an object in an image as a face candidate that its shape resembles the shape of a face.

• **Feature Extraction** : The key issue of any recognition system is feature extraction. Feature extraction abstracts high level information about individual patterns to facilitate recognition. Selection of feature extraction method is probably the single most

important factor in achieving high recognition performance. In order to design a good face recognition system, the choice of feature extractor is very crucial. To design a system with low to moderate complexity the feature vectors should contain the most pertinent information about the face to be recognized. Face recognition system should be capable of recognizing unpredictability of face appearance and changing environment.

• **Classifier**: Comparison of the face to a database of known faces.

Input image → [Face Detection] → [Feature Extraction] → [Neural Network Classifier] → [Decision Strategy] → Out put person
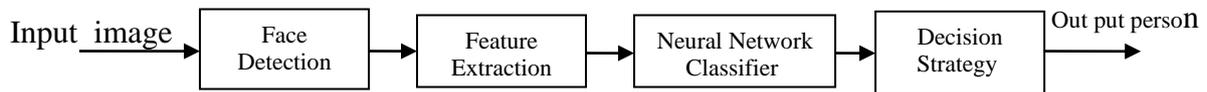
Figure 1 : Face Recognition System

3. Proposed Face Recognition System

The architecture of the proposed system is depicted in figure (1). The face recognition system developed comprises three major processing modules which are:

3.1. Face Detection

The problem of face recognition is all about face detection, before face recognition is possible, one must be able to reliably find a face and its landmarks. Most face detection systems attempt to extract a fraction of the whole face, thereby eliminating most of the background and other areas of an individual's head such as hair that are not necessary for the face recognition task. With static images, this is often done by running a '*window'* across the image. [3] A manual face detection system is implemented by measuring the facial proportions of the average face. To detect a face, a human operator would identify the locations of the subject's eyes in an image and using the proportions of the average face, the system would segment an area from the image. In the ideal frontal view segmented facial image for face recognition, the lower edge of each eye is 27% from the top of the image and the left and right eyes are 20% and 80% from the left border of the image respectively[2], see Figure (2). Operator instructed to click under a subject's left and right eye. However, just use a single statistic (vector between lower edge of eyes) so as not to lose the natural variation between human faces.
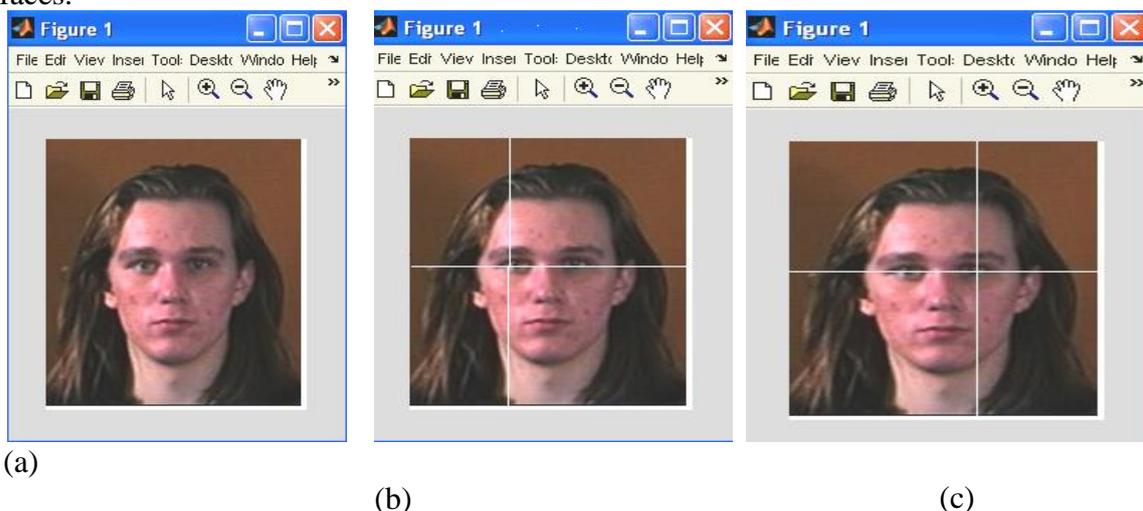

(a)
(b)   (c)

**Figure 2: Manual Face Detection a- Original Image b- Click Under Eye on left ,
c- Click Under Eye on Right**

3.2. Feature Extraction

Feature selection in pattern recognition involves the derivation of certain features from the input data in order to reduce the amount of data used for classification and provide discrimination power. Due to the measurement cost and classification accuracy, the number of features should be kept as small as possible. A small and functional feature set makes the system work faster and use less memory. On the other hand, using a wide feature set, may cause "curse of dimensionality" which is the need for exponentially growing number of samples [2,5]. Feature extraction methods try to reduce the feature dimensions used in the classification step. There are especially two methods used in pattern recognition to reduce the feature dimensions; Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [5]. The advantage of PCA comes from its generalization ability. It reduces the feature space dimension by considering the variance of the input data. The method determines which projections are preferable for representing the structure of the input data. Those projections are selected in such a way that the maximum amount of information (i.e. maximum variance) is obtained in the smallest number of dimensions of feature space. In order to obtain the best variance in the data, the data is projected to a subspace(of the image space)which is built by the eigenvectors from the data. In that sense, the eigenvalue corresponding to an eigenvector represents the amount of variance that eigenvector handles [5].

In the proposed system we use **Principle Component Analysis** (PCA) to extract feature from the derived subimages. Therefore this approach can extract characteristics of face images for classification purpose.

3.2.1. Principle Component Analysis

PCA aims to determine a set of orthogonal vectors that optimally represent the distribution of the data. Any face images can then be theoretically reconstructed by projections onto the new coordinate system. In search of a technique that extracts the most relevant information in a face image to form the basis vectors.

3.2.1.1. PCA In Statistics :

Principal components analysis (PCA) is a technique that can be used to simplify a dataset; more formally it is a transform that chooses a new coordinate system for the data set such that the greatest variance by any projection of the data set comes to lie on the first axis (then called the first principal component), the second greatest variance on the second axis, and so on. PCA can be used for reducing dimensionality in a dataset while retaining those characteristics of the dataset that contribute most to its variance by eliminating the later principal components [1] . PCA aims at

- It reducing the dimensionality of the data set.
- It eliminates those components, which have the least variation in the data set.
- It speeds up the computational time

3.2.1.2. PCA In Mathematic [9,11]:

A two dimention facial image can be represented as one dimension vector by concatenating each row (or column) into a long thin vector. Let's suppose we have M vectors of size N (= rows of image x columns of image) representing a set of sampled images. p^j's represent the pixel values.

$$xi = [p1 \ldots pN]^T, \quad i = 1, \ldots, M \qquad \ldots(1)$$

The images are mean centered by subtracting the mean image from each image vector. Let m represent the mean image.

$$m = \frac{1}{M} \sum_{i=1}^{M} x_i \qquad \qquad ...(2)$$

And let w_i be defined as mean centered image

$$w_i = x_i - m \qquad \qquad ...(3)$$

Our goal is to find a set of e_i's which have the largest possible projection onto each of the w_i's. We wish to find a set of M orthonormal vectors e_i for which the quantity

$$\lambda_i = \frac{1}{M} \sum_{n=1}^{M} (e_i^T w_n)^2 \qquad \qquad ...(4)$$

is maximized with the orthonormality constraint

$$e_1^T e_k = \delta_{lk} \qquad \qquad ...(5)$$

It has been shown that the e_i's and λi's are given by the eigenvectors and eigenvalues of the covariance matrix

$$C = WW^T \qquad \qquad ...(6)$$

where W is a matrix composed of the column vectors w_i placed side by side. The size of C is $N \; x \; N$ which could be enormous. For example, images of size $64 \; x \; 64$ create the covariance matrix of size $4096 \; x \; 4096$. It is not practical to solve for the eigenvectors of C directly. A common theorem in linear algebra states that the vectors e_i and scalars λi can be obtained by solving for the eigenvectors and eigenvalues of the $M \; x \; M$ matrix W^TW. Let d_i and μ_i be the eigenvectors and eigenvalues of W^TW, respectively.

$$W^TWd_i = \mu_i d_i \qquad \qquad ...(7)$$

By multiplying left to both sides by W

$$WW^T(Wd_i) = \mu_i(Wd_i) \qquad \qquad ...(8)$$

which means that the first M - 1 eigenvectors ei and eigenvalues λi of WW^T are given by Wd_i and μ_i, respectively. Wd_i needs to be normalized in order to be equal to e_i. Since we only sum up a finite number of image vectors, M, the rank of the covariance matrix cannot exceed M - 1 (The -1 come from the subtraction of the mean vector m).

The eigenvectors corresponding to **nonzero eigenvalues** of the covariance matrix produce an orthonormal basis for the subspace within which most image data can be represented with a small amount of error. The eigenvectors are **sorted from high to low** according to their corresponding eigenvalues. The eigenvector associated with the largest eigenvalue is one that reflects the greatest variance in the image. That is, the smallest eigenvalue is associated with the eigenvector that finds the least variance. They decrease in exponential fashion, meaning that the roughly 90% of the total variance is contained in the first 5% to 10% of the dimensions.

A facial image can be projected onto M' ($<< M$) dimensions by computing

$$\Omega = [v_1 v_2 .... v_{M^i}] \qquad \qquad ...(9)$$

where $vi = e^T_i \; w_i$. v_i is the ith coordinate of the facial image in the new space, which came to be the principal component. The vectors e_i are also images, so called, *eigenimages*, or *eigenfaces*

3.3. Classifier of Faces by Recurrent (Time Cycling) Back Propagation Artificial Neural Network

Neural networks have been employed and compared to conventional classifiers for a number of classification problems. The results have shown that the accuracy of the neural network approaches equivalent to, or slightly better than, other methods. Also, due to the simplicity, generality and good learning ability of the neural networks, these types of classifiers are found to be more efficient [12]. Due to the above reasons

Recurrent (Time Cycling) Back Propagation ANN used as classifier and it serve as an excellent candidate for pattern applications and attempts have been carried out to make the learning process in this type of classification faster.
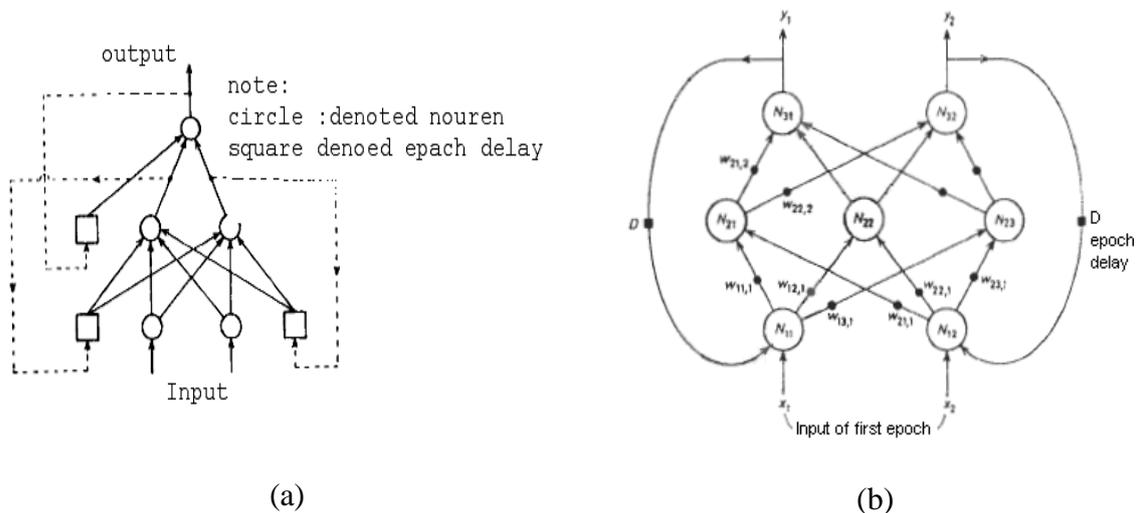
- **Recurrent (Time Cycling) Back Propagation Artificial Neural Network Structure [4]**

A recurrent structure can be introduced into back propagation neural networks by feeding back the network's output to the input after an epoch of learning has been completed. This recurrent feature is in discrete steps (cycles) of weight computation. This arrangement allows the employment of back propagation with a small number of hidden layers (and hence of weights) in a manner that effectively is equivalent to using m-times that many layers if m cycles of recurrent computation are employed.

A recurrent (time cycling) back propagation network is described in Figure 3 a,b) The delay elements (D in Figure (4)) in the feedback loops separate between the time-steps (epochs, which usually correspond to single iterations). At end of the first epoch the outputs are fed back to the input. Alternatively, one may feed back the output-errors alone at the end of each epoch, to serve as inputs for the next epoch. The network of Figure (4b) receives inputs x1 and x2 at various time steps of one complete sequence (set) that constitutes the first epoch (cycle). The weights are calculated as in conventional back-propagation networks and totaled over all time steps of an epoch with no actual adjustment of weights until the end of that epoch. At each time step the outputs y1 and y2 are fed back to be employed as the inputs for the next time step. At the end of one complete scan of all inputs, a next epoch is started with a new complete scan of the same inputs and time steps as in the previous epoch.

Equation 10 used to calculate output of Recurrent (Time Cycling) Back Propagation, (y_i) denoting the output and v_j being the outputs of the hidden layers neurons, x_i being the external input, g denoting an activation function.

$$y_i = g(x_i + \sum_j w_{ij}v_i) \qquad\qquad …(10)$$

(a)   (b)

Figure3: a - Fully Recurrent (Time Cycling) Back Propagation ANN

b- Recurrent (Time Cycling) Back Propagation ANN

4. Extract Feature from Database for Classifier

If the dimension of the input vector is too large, the network can be quite complex and therefore difficult to train and may take more time for classification; hence

it is required to reduce the input vector dimension. In our research we used Principal Component Analysis (PCA) technique for dimension reduction in face recognition

4.1. Face Database

face image databases (containing both training (120 faces image) and test (40 faces image) data) used form the Collection of Facial Images : Faces95 database is one that was created by Computer Vision Science Research Projects on Face Recognition [13] , see figure (4). This database widely used by researchers to test face detection and recognition systems, see appendix (B).

Database Description

- Number of individuals: 72, 40 individual are used in this research.
- Image resolution: 180 by 200 pixels
- Contains images of male and female subjects

Figure 4: Samples of Database using in Proposed Face Recognition System

4.2 Input and Output Data of Feature Extraction (PCA)

Image resolution for database is 180 by 200 pixels and segmented images is 73 x 65 pixels. The column matrix of all images is converted from (73 x 65) to a vector (4745, 1). This vector is used as input matrix. The size of the input matrix depends on the number of poses 'n' of 'N' persons. If database has 'n' poses of 'N' persons, then size of the input matrix becomes (4745, n × N). The first n columns represents the n poses of 1st person, 2nd n columns represents the n poses of 2nd person and so on. In our research 3 poses of 40 persons will be taken for training then, the size of the input matrix will become (4745, 3 × 40) or (4745, 120). The output from Feature extraction PCA is 20 values represent Eigen vectors of the covariance matrix of the training database.

4.3. Output Data from Classifier

The target matrix is to identify the person to whom that test vector belongs. If N persons are to be identified, then size of target vector is (N, 1). If there are 120 image in the input matrix, then size of target matrix is (40, 120 (where 120 represent 3 poses of 40 persons). The target matrix elements are all zeros except one element whose value is 1, which indicates the position of the corresponding person.

4.4 . PCA Extraction for Face Recognition System

The following algorithm are involved for extracting principal components algorithm of the input vector to the classifier [8].

1. First the preprocessing of the matrix is done so that the mean of all the elements of the matrix is zero and the standard deviation is one. This can be obtained as follows:

$$P_n = (P - meanP)stdP \qquad\qquad \ldots(11)$$

where, P is the input matrix whose principal components are to be derived; mean P, mean of all the elements of the matrix P; std P, standard deviation of the matrix P;

and P_n, the matrix derived from the *P* matrix whose mean is zero and standard deviation is one.

2. Singular value decomposition is used to compute the principal components. Singular value decomposition of a matrix is done as follows:

$$[u, d, v] = svd(p_n, 0) \qquad \qquad …(12)$$

where the operator svd computes the singular value decomposition of the matrix P_n and produces a pseudo-diagonal matrix *d* with non-negative diagonal elements in decreasing order, unitary matrics *u* and *v*, such that $Pn = u \times d \times v^T$, v^T being the transpose form of *v*. The diagonal elements are the principal singular values of the matrix P_n.

3. Compute the variance of each principal component by squaring each principal components and dividing by (*Q*-1).

$$var = diag(d)^{2/(Q-1)} \qquad \qquad …(13)$$

where the operator diag produces an array whose elements are the diagonal elements of d, which are the principal components, and *Q* is the number of columns of the matrix P_n.

4. Compute total variance and fractional variance. Total variance is the sum of all the variance and fractional variance is obtained by dividing each variance by total variance.

$$total - variance = sum(var) \qquad \qquad …(14)$$

$$frac\_var = \frac{var}{total\_variance} \qquad \qquad …(15)$$

5. Define a fractional number min-frac and find the components of the frac_var which is more than min_ frac.

$$greater = (frac\_var > min\_frac) \qquad \qquad …(16)$$

$$size\_pc = sum(grater) \qquad \qquad …(17)$$

6. Reduction of the size of the unitary matrix u by taking only first size_pc columns and deleting the other columns. The resultant matrix is transposed.

$$u = u(:,1:size\_pc) \qquad \qquad …(18)$$

7. Then the matrix P_n is transformed by multiplying it with the unitary matrix *u*, so that the resultant transformed matrix only consists of principal components.

The above algorithm is used to find out the principal components of the input matrix to the neural network. Now the input matrix consists of only these principal components. The size of the input matrix is reduced from (4745, 120) to (20, 120).

Principle Component Analysis (PCA) is programmed depended on the above algorithm and used Matlab system. See Figure (5).

5. Design Recurrent (Time Cycling) Back Propagation Artificial Neural Network

face recognition problem has been solved using a recurrent back propagation neural network. The task is to teach the neural network to recognize 40 faces, output is used to recognized one face from 40 faces (40 output node)

The neural network consists of three layers with 20 neurons input (represent PCA) , 5 neurons hidden , 40 neurons output. The neural network is as it is a recurrent network, such that its outputs y1 … y40 are fed back as additional inputs at the end of each iteration. The structural diagram of neural network is given in Figure (6).

To train the network to produce error signal we will use 120 face images (40 person in 3 poses). To check whether the network has learned to recognize errors we will use 40 face images. To minimize the error-energy at the output layer, weight

setting is as in regular Back-Propagation. To train the network to recognize faces we applied 20 values represents PCA to the input of the network. Additional inputs were initially set equal to zero and in the course of the training procedure were set equal to the current output error.

6. Experimental Results

To check the utility of our proposed algorithm experimental studies are carried out on the Collection of Facial Images : Faces95 databases (containing both training and test data). 120 face images from 40 individuals in 3 poses Faces95 database have been used to evaluate the performance of the proposed method. None of the 40 samples are identical to each other. They vary in position, rotation, scale and expression. In this database each person has changed his face expression in each of 40 samples.

A PCA feature domains and the Recurrent (Time Cycling) Back Propagation neural network has been developed. In this example, for the PCA feature vector has been created based on the 20 largest PCA number for each image. A total of 120 images have been used to train and another 40 for test. Recognition rate of training data set is 99% and 91.89 % was obtained for test data set using this proposed technique, see table(1).

Table (1) Face Recognition Accuracy Rate

Number of faces image (person)		Number of poses		Accuracy Rate %	
Training data set	Testing data set	Training data set	Testing data set	Training data set	Testing data set
40 (120 face image (3 poses * 40)	40	3	At any poses	99%	91.89

Begin

Enter Segmented image from Face Detection step with size of (73 x 65) for 120 (3 poses x 40 persons

Calculate mean for image by using matlab function **mean** to computing the average face image

Calculate deviation of each image from mean image by computing the difference image for each image in the training set

Calculate eigenvalues and eigenvector : This is programmed by using matlab function .   Syntax   : [V,D] = eig(A)

Sorting and eliminating eigenvalues :   All eigenvalues of matrix X are sorted and those

Calculate eigenvectors of covariance matrix

All centered images are projected into facespace by multiplying in Eigenface basis's. Projected vector of each face will be its corresponding feature vector.

Save Resulte in three Variables1) Eigenface: Eigen vectors of the covariance matrix of the training database, 2)
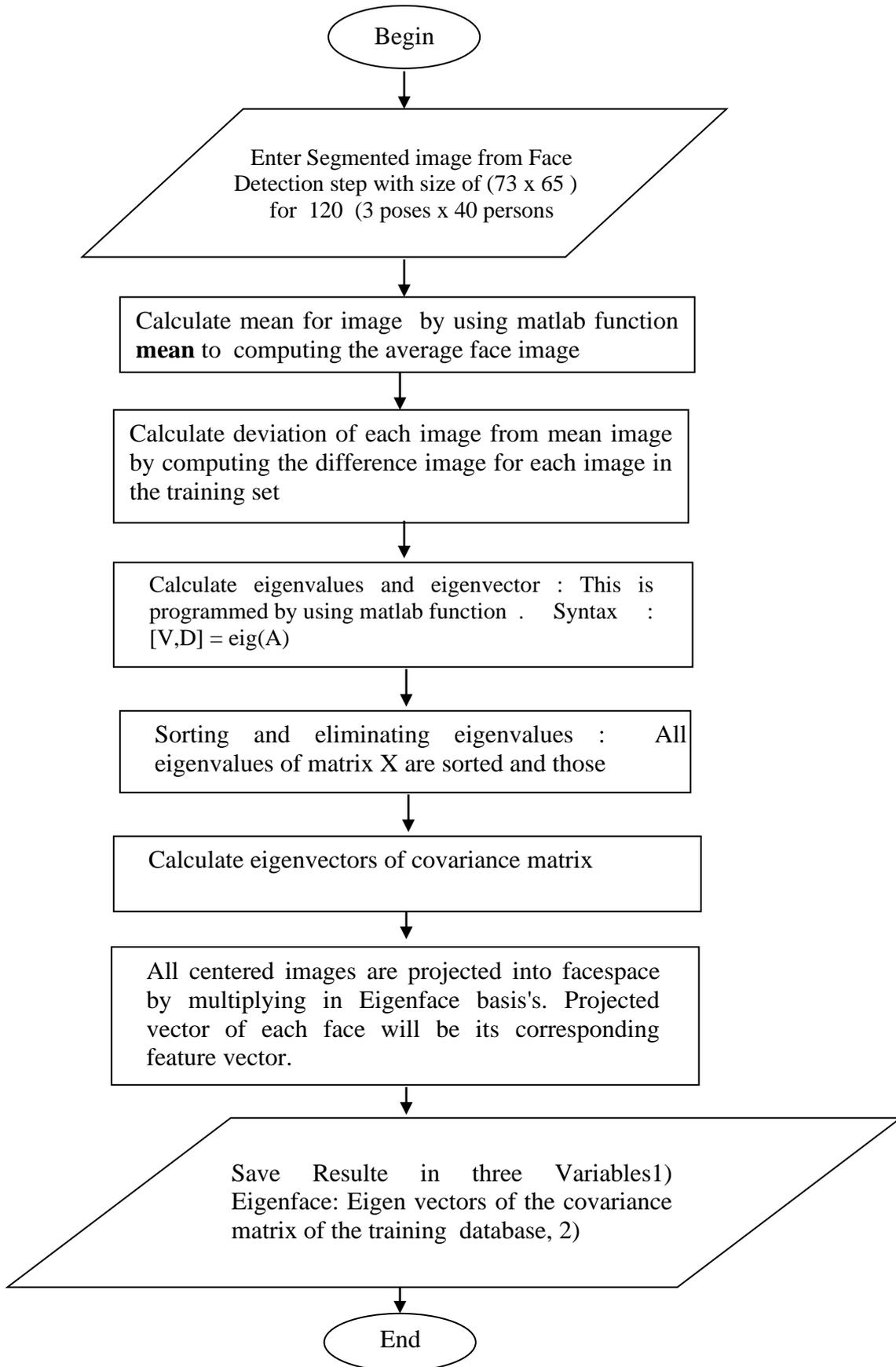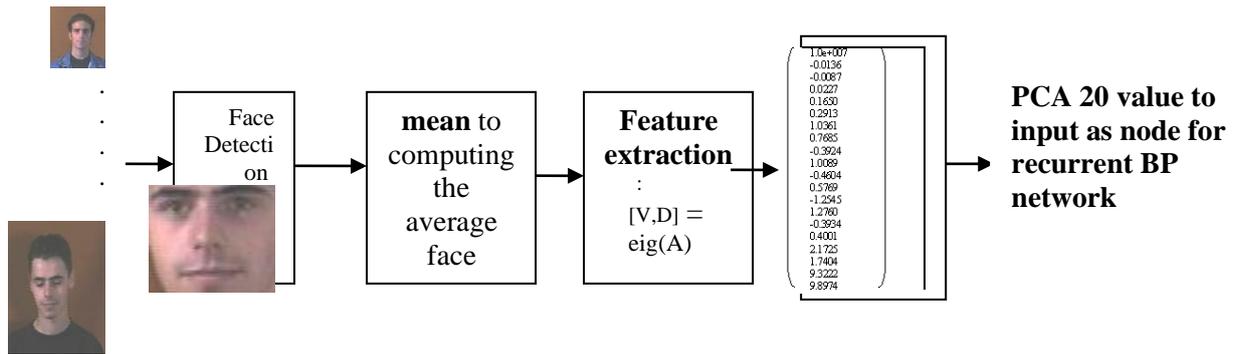
End

Figure 5: Flow chart for PCA Feature Extraction

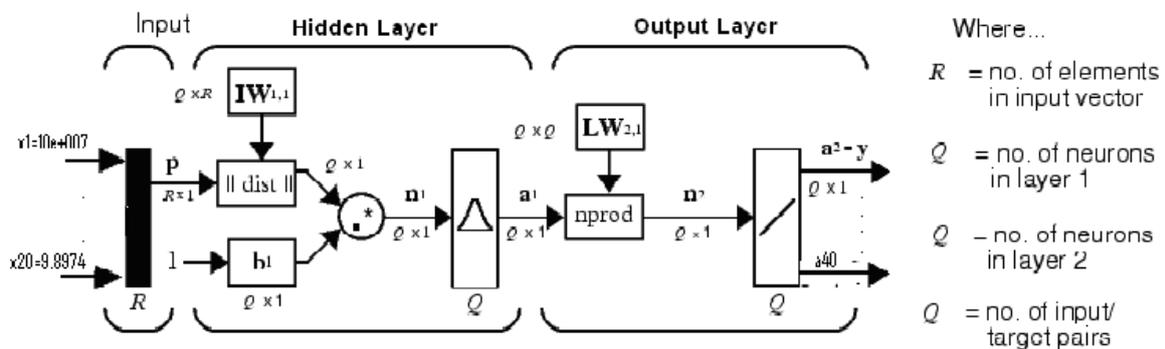**Flow chart for PCA Feature Extraction**

Figure (6) recurrent back propagation neural network

7. Conclusion

This paper presented a novel method for the recognition of human faces in 2-Dimensional digital images. It employs the Recurrent (Time Cycling) Back Propagation neural networks and PCA feature domains. The highest recognition rate for training 99% and testing 91.89 with the the Faces95 database was obtained using this proposed algorithm indicates the usefulness of the proposed technique.

S Future Work

For our future work, we are planning to apply the genetic algorithm on a number of interest points of some faces and determine the best features for face. Then using only these selected features, and using recurrent ANN for classifier.

REFERENCES

[1]   Aditya, K., 2002, "Face Recognition using Eigenface   Approach" , Master thesis. Department of Computer Science and Engineering Indian Institute of Technology, Bombay Mumbai.
www.cse.iitb.ac.in/~adityak/seminar.ps

[2]   Balasuriya, L. S.,   2000. "Frontal View Human Face Detection and Recognition", Thesis Department of Statistics and Computer Science, University of Colombo.
www.dcs.gla.ac.uk/~sumitha/papers/FaceRecognition.pdf

[3]   Brunelli, R. and Poggio, T. 1993, "Face Recognition: Features versus Templates*".* IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042-1052

[4]   Daniel G., 2007, Principle of Artificial Neural Network, World Scientific.

[5]   GÜL, A. B. .2003, "Holistic Face Recognition by Dimension Reduction ", Master thesis school of Natural and Applied Sciences of the Middle East Technical University
www. etd.lib.metu.edu.tr/upload/1056738/index.pdf

[6]   Haddadnia, J. , Faez, K., Ahmadi, M. "N-Feature Neural Network Human Face Recognition", Electrical and Computer Engineering Department, University of Windsor,Windsor, Ontario, Canada, N9B 3P4.
http://www.uwindsor.ca/units/eng-electric/electric2.nsf/inToc
/F7F0FDECEB796F7C85256FA3006D38DA

[7]   Haddadnia, J., Faez, K., Moallem,P. , 2000,   "Neural Network Based Face Recognition with Moments   Invariant", IEEE Int. Conf. On Image Processing*,* Vol. I, pp. 1018-1021,

[8]   Karibasappa ,K. 2004, " Face Recognition by ANN using Wavelet Transform Coefficients" , IE(I) Journal-CP, Vol 85, May 2004.

[9]   Kyungnam ,K. "Face Recognition using Principle Component Analysis" , Department of Computer Science, University of Maryland, College Park MD 20742, USA.
 www.umiacs.u**md**.edu/~**k**n**k**im/**K**G_VISA/PCA/FaceRecog_PCA_**K**im.pdf

[10]   Raphael C.m,1999. "Real Time Face Recognition using Eigenfaces", Thesis , Dept of Computer sciences and Electrical Eng. ,University of Queenland.

[11]   Turk , M.,Pentland A.P.,1991. "Face Recognition Using Eigenfaces", IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591.

[12]   Turk , M. ,Pentland A. P., 1991. " Eigenfaces for Recognition". Journal of Cognitive Neuroscience, No. 3, Vol. 1.

[13]   Collection of Facial Images : Faces95, Computer Vision Science Research Projects, Designed and maintained by Dr Libor Spacek. Updated Friday, 16-Feb-2007 , Http://cswww.essex.ac.uk/mv/allfaces/faces95.html.

Appendix (A)
MATLAB Function Reference   (eig)
Find eigenvalues and eigenvectors
Syntax   :   [V,D] = eig(A)
produces matrices of eigenvalues (D) and eigenvectors (V) of matrix A, so that A*V = V*D. Matrix D is the canonical form of A--a diagonal matrix with A's eigenvalues on the main diagonal. Matrix V is the modal matrix--its columns are the eigenvectors of A.

Appendix (B)
Collection of Facial Images : Faces95
Collection of Facial Images : Faces95 by Computer Vision Science Research Projects, Designed and maintained by Dr Libor Spacek. Updated Friday, 16-Feb-2007, Http://cswww.essex.ac.uk/mv/allfaces/faces95.html.
Acquisition conditions : Using a fixed camera, a sequence of 20 images per individual was taken. During the sequence the subject takes one step forward towards the camera. This movement is used to introduce significant head (scale) variations between images of same individual. There is about 0.5 seconds between successive frames in the sequence.
Database Description :Number of individuals: 72 , Image resolution: 180 by 200 pixels (portrait format) , Contains images of male and female subjects
Variation of individual's images : Backgrounds: the background consists of a red curtain. Background variation is caused by shadows as subject moves forward. **Head Scale:** large head scale variation , **Head turn,tilt and slant:** minor variation in these attributes , **position of face in image** : some translation , **image lighting variation:** as subject moves forward, significant lighting   changes occur on faces due to the artificial lighting arrangement , **expression Variation:** some expression variation, **additional comment:** there is no hair style variation as the images were taken in a single session.