Principal Component Analysis: An Efficient Facial Feature Extraction Technique

[1] Drakshayani Desai, [2]Dr. Ramakrishna Seemakurti.
[1] *Research scholar, ECE Dept, JJTU, Rajasthan, India, 333001(Associate Professor, SVIT, Secunderabad-500 003)*

[2] *Research Guide (Pricipal,, SVIT, Secunderabd, India, 500003) (Approved Research Guide, JJTU, Jhunjhunu-333001, Rajasthan, India)*

Abstract:
Biometric access systems came into existence long back. Methods such as voice based access, finger print recognition, password key systems, iris recognition, DNA and face recognition are a few important modes of biometric access systems. This paper highlights on Face Recognition using Principal Component Analysis(PCA).The principal component analysis is a kind of algorithms used in biometrics. It is a statistical technic which converts a set of observations that are orthogonally correlated variables into a set of values of linearly uncorrelated variables. The system compares the given individual to all the other individuals in the database and gives a ranked list of matches of the images. Principal component analysis (PCA) is the simplest form of true eigenvector based multivariate analysis. Mathematically, it is an orthogonal linear transformation that transforms the data to a new coordinate system. The use of Eigen faces is commonly called as Principal Component Analysis. With PCA, the image which are used for verification and the identification they should be of same size and they are normalized to line-up the eyes and mouth of the subjects within the image. Using PCA, dimension of data is reduced and precisely so as to decompose the face structure into orthogonal and uncorrelated components know as Eigen faces.

1. INTRODUCTION:

Biometrics is derived from Greek words “bio” meaning “life” and metrics meaning “to measure”. Biometrics refers to the identification or recognition of a person based on their physiological and/or behavioral characteristics. Several detection and recognition based biometrics have evolved based on various unique aspects of human body. In many real time applications the human face recognition is most important. Face detection and recognition system has the various applications in the various fields. For identifying the person from the digital images or from the video input this system is used. Face recognition systems is the part of facial image processing applications. The face Recognition system uses the biometric information of the humans for recognizing the face and it is applicable easily instead of fingerprint, iris, signature etc because it does not need human interaction. There are various applications of this system such as it can be used for prevention of the crimes, for video surveillance system, for person verification, and other security activities.

There are several methods to detect the face in a given image such as artificial neural network, Eigen face method, support vector machine etc. The proposed face recognition system “Principal Component Analysis” is a combination of face detection and recognition techniques in digital image processing. This principal component analysis algorithm is used to classify given images with known structured properties called as “Eigen Vectors”. These Eigen Vectors are used commonly in most of the computer vision applications. Principal component analysis (PCA), also known as Karhunen-Loeve expansion, is a classical feature extraction and data representation technique widely used in the areas of pattern recognition and computer vision such as face recognition [4]. The tactic of the Eigen face method consists of extracting the characteristic features on the face and representing the face in a linear combination of the so called ‘Eigen faces’ that can be obtained from the feature extraction process [5]. The principal components of the faces in the training set are calculated using Eigen face.

Figure below show the face recognition system.

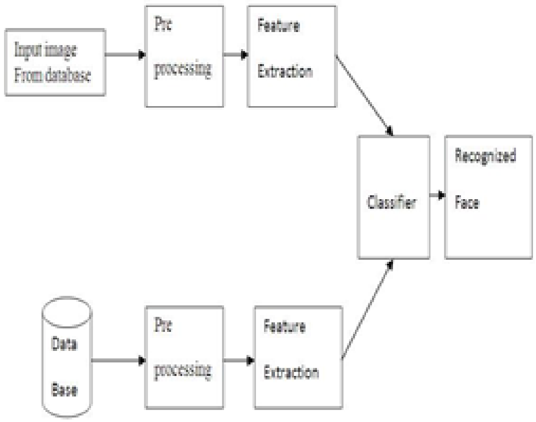

Figure 1.1 face recognition system.

It includes three main parts pre-processing, feature extraction and classification. Preprocessing step removes noise from the image. Feature extraction step extracts the prominent features of the face such as position of eyes, nose, chicks and chin. The classifier compares extracted features with database images, depending on the comparison result face is declared as recognized or un-recognized.

2. Principal Component Analysis

The PCA method was developed in 1991 by Karl Pearson. This Eigen face technique is used for recognition of the images of the faces and for the compression of images. The prediction of the images can be done with the help of Principal Component Analysis algorithm. Similarly the redundancy is removed from the images, feature extraction and the data compression of the images is done with the help of PCA.

Some terms associated with Principal Component Analysis:

A] Training Set

Training set is formed by combining the different set of face images. These images are collected from different type of sources. These collected images contain the different facial expressions and pose. These images are used for the detection and recognition purpose.

B] Eigen Face

These are set of the facial features in the form of the vectors. It is a weighted combination of some distinct component of base faces. Every image in the training set has its own contribution on making the Eigen faces. Facial expression which occurs in the Eigen faces are deviates from the original images because of their facial expression and pose. Eigen vectors are formed by converting the captured image matrix into the corresponding vector form. These calculated Eigen vectors must satisfy the Eigen value equation. The size of this Eigen vector is less(since the image is compressed) but there is no loss of data in these Eigen vector.

C] Weight Vector

For the recognition of the face, the weight of the known Eigen face (stored in image database) is calculated. When the new face image is to be recognized then the weight associated with that Eigen face is calculated. By comparing these calculated weighs with the weights of the known face images, we can recognize whether the face is known face or unknown face.

D] Euclidian Distance

Euclidian Distance is a distance between the two points in the vector plane. By using the Pythagorean formula we can calculate this Euclidian distance. The Euclidian distance between two points is a length of the segment connecting these two points. Here we are calculating the Euclidian distance between the input image and the training face. If the Euclidian distance is less than the threshold value then the face is determined as known face or if this distance is above the threshold value then the determined face is unknown.

The procedures of Principal Component Analysis consist of two phases, training step and recognition step.

1. *Training Step:* This step is a process to get Eigen face from training image which previously has been changed into data matrix. Samples of data, on which the system needs to recognize, are used to create an Eigen Matrix which transforms the samples in the image space into the points in Eigen face.

2. *Recognition Step:* This step is a process to get Eigen face from test image which previously has been changed into data matrix. These results were then compared with results from training phase to get minimum difference.

2.1 PCA Algorithm

To extract the relevant features of facial images, Principal Component Analysis (PCA) method is used. face recognition based on PCA is generally referred to as the use of Eigen faces. Eigen face is a Principal Components of the face, or equivalently, the Eigen vectors of the covariance matrix of the set of the training images, where an image with n by n pixels is considered as a point in n2 dimensional space [5].The PCA algorithm involves the following steps:

Step1: Prepare the training faces and obtain face images I1, I2, I3, I4 , ... IM (training faces). The face images must be centered and of the same size.

Step 2: Prepare the data set of each face image Ii in the database and is transformed into a vector and placed into a training set S.

$S = \{\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \ldots\ldots\ldots\ldots\Gamma_M\}$

Each image is transformed into a vector of size MN × 1 and placed into the set. For simplicity, the face images are assumed to be of size N × N resulting in a point in N^2 dimensional space. An ensemble of images, then, maps to a collection of points in this huge space.

Step 3: Compute the average face vector The average face vector (Ψ) has to be calculated by using the following formula:

$$\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$$

Step 4: The average face vector Ψ is subtracted from the original faces Γ_i and the result stored in the variable Φ_i.

$$\Phi_i = \Gamma_i - \Psi$$

Step 5: Calculate the covariance matrix C in the following manner,

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T$$

$$= AA^T \qquad (N^2 \times N^2\ matrix$$

Where, $A=[\Phi_1, \Phi_2, \Phi_3, \Phi_4, \ldots\ldots\ldots\Phi_M$

$$(N^2 \times M\ matrix)$$

Step 6: Calculate the Eigen vectors and Eigen values of the covariance matrix.

The covariance matrix C in step 5 has a dimensionality of N^2 x N^2, so one would have N^2 Eigen face and Eigen values. For an 256 × 256 image 65,536 Eigen faces can be calculated. Computationally, this is not very efficient as most of those Eigen faces are not useful for our task. In general, PCA is used to describe a large dimensional space with a relative small set of vectors [3].

Compute the Eigen vectors μ_i of AA^T

The matrix AA^T is very large.

Step 6.1: consider the matrix

$$L = A^T A\ (M \times M \text{ matrix})$$

Step 6.2: Compute Eigen vectors v_i of $L = A^T A$

$$A^T A v_i = \mu_i v_i$$

the relationship between μ_i and v_i

$$A^T A v_i = \mu_i v_i$$

$$AA^T A v_i = \mu_i A v_i$$

$$C A v_i = \mu_i A v_i \quad [\ \text{since } C = AA^T]$$

$C\mu_i = \mu_i A v_i$ where $\mu_i = A v_i$ thus, $C = AA^T$ and $L = A^T A$ have the same eigenvalues and their eigenvectors are related as follows: $\mu_i = A v_i$

Note 1: $C = AA^T$ can have upto N^2 eigenvalues and eigenvectors.

Note 2: $L = A^T A$ can have upto M eigenvalues and eigenvectors.

Note 3: The M eigenvalues of $C = AA^T$ (along with their corresponding eigenvectors) correspond to the M largest eigenvalues of $L = A^T A$ (along with their corresponding eigenvectors).

Where v_i is an eigenvector of $L = A^T A$. From this simple proof it can see that $A v_i$ is an eigenvector of $C = AA^T$.

The M eigenvectors of $L = A^T A$ are used to find the M eigenvectors of μ_i of C that form our Eigen face basis:

$$u_i = \sum_{i=1}^{M} v_i \Phi_i$$

Where, μ_i are the Eigenvectors i.e. Eigen faces.

Step 7: keep only K eigenvectors (corresponding to the K largest eigenvalues)
Eigen faces with low eigenvalues can be omitted, as they explain only a small part of Characteristic features of the faces.

2.2 Flowchart of Principal Component Analysis

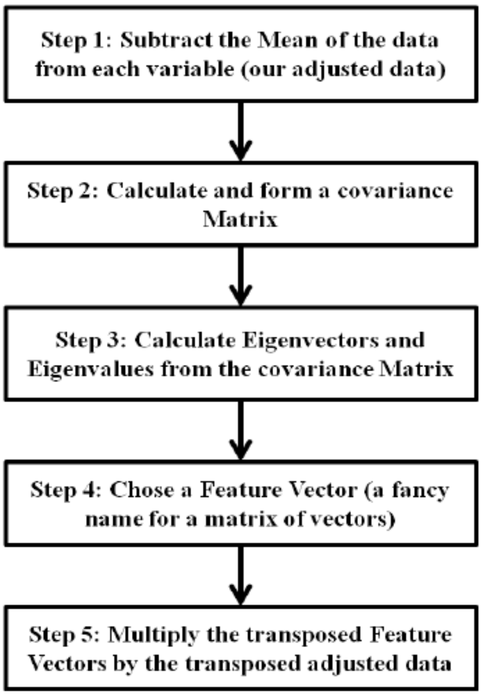

Figure. 2 Flowchart of PCA

3. Advantage and disadvantage of PCA

The advantages of Principal Component Analysis are its:

1) Low noise sensitivity.
2) Reduced complexity.
3) Smaller database representation.
4) Less memory requirement.

Disadvantages of PCA are:

1) The covariance matrix is difficult to be evaluated in an accurate manner.
2) Even the simplest invariance could not be captured by the PCA unless the training data explicitly provides this information.

Table 1. The features of PCA are shown in the table 1

Features	Principal Component Analysis
Discrimination between classes	PCA manages the entire data for the principal components analysis without taking into consideration the fundamental class structure.
Applications	PCA applications in the significant fields of criminal investigation are beneficial
Computation for large datasets	PCA does not require large computations
Direction of maximum discrimination	The directions of the maximum discrimination are not the same as the directions of maximum variance as it is not required to utilize the class information such as the within class scatter and between class scatter
Focus	PCA examines the directions that have widest variations
Supervised learning technique	PCA is an unsupervised technique.
Well distributed classes in small datasets	PCA is not as powerful as other Methods.

4. Result Analysis

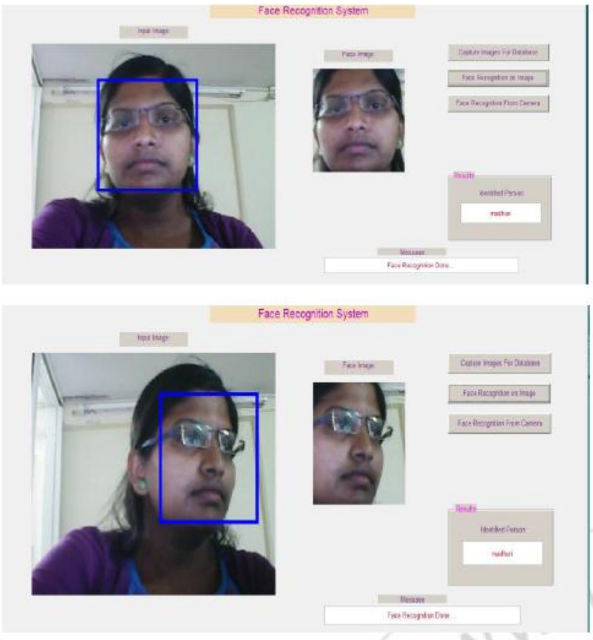

Figure 3. Face Recognition using PCA

Here sixty face images of three different person with different expressions are taken. Out of that sixty images of the face of each person ten images are placed as the training images and ten images are placed as the testing images of each person. And the face recognition is done on the testing images by using Principal Component Analysis Algorithm. The accuracy achieved in recognizing the first person faces of different expression is 80%. The accuracy achieved in recognizing the faces of second person with different expression is 90%.And the accuracy achieved in recognizing the faces of third person with different expression is 80% as shown in the table below

Input Image sets	Training Images	Testing Images	Matched Images	Accuracy %
1	10	10	8	80%
2	10	10	9	90%
3	10	10	8	80%

Table 2:Face recognition result

5. Conclusions

The PCA method is an unsupervised technique of learning that is mostly suitable for databases that contain images with no class labels. A detailed description of the PCA technique utilizing in face recognition has been provided. As mentioned above, the PCA method's advantages and disadvantages have also been explained in this study.

REFERENCES

[1] DashunQue, Bi Chen, Jin Hu "A Novel Single training sample face recognition algorithm based on Modular Weighted (2d)2 PCA" School of Information Technology, Wuhan University of Technology, Wuhan 430063, P. R. China.

[2] J. Ashok, V. Shivashankar and P. Mudiraj, "An Overview of Biometrics," *International Journal*, Vol. 2, 2010.

[3] F. Ã–zen, "A Face Recognition System Based on Eigen faces Method," *Procedia Technology*, Vol. 1, 2011, pp. 118-123

[4] W. Miziolek and D. Sawicki, "Face Recognition: PCA or ICA," *Przeglad Elektrotechniczny*, Vol. 88, 2012, pp. 286-288.

[5] C. Li, Y. Diao, H. Ma and Y. Li, "A Statistical PCA Method for Face Recognition," *in Intelligent Information Technology Application*, 2008, pp. 376-380.

[6] H. Duan, R. Yan and K. Lin, "Research on Face Recogni-tion Based on PCA," *in Future Information Technology and Management Engineering*, 2008, pp. 29-32.

[7] T. F. Karim, M. S. H. Lipu, M. L. Rahman and F. Sultana, "Face Recognition Using PCA-based Method," *in Ad-vanced Management Science (ICAMS), 2010 IEEE Inter-national Conference on*, 2010, pp. 158-162.

[8] Z. Haiyang, "A Comparison of PCA and 2DPCA in Face Recognition," *in Electrical Power Systems and Com-puters: Springer*, Vol. 99, 2011, pp. 445-449. doi:10.1007/978-3-642-21747-0_55

[9] B. Poon, M. A. Amin and H. Yan, "PCA based Face Recognition and Testing Criteria," *in Machine Learning and Cybernetics, 2009 International Conference on*, 2009, pp. 2945-2949.

[10] Y. Wen and P. Shi, "Image PCA: A New Approach for Face Recognition," in Acoustics, Speech and Signal Pro-cessing, *IEEE International Conference on*, 2007, pp. 1241-1244.

[11] A. Boualleg, C. Bencheriet and H. Tebbikh, "Automatic Face Recognition Using Neural Network-PCA," *in In-formation and Communication Technologies*, 2006, pp. 1920-1925.

[12] Z. Wang and X. Li, "Face Recognition Based on Im-proved PCA Reconstruction," *in Intelligent Control and Automation (WCICA)*, 2010 8th World Congress on, 2010, pp. 6272-6276

[13] J. Meng and Y. Yang, "Symmetrical Two-Dimensional PCA with Image Measures in Face Recognition," *Int J Adv Robotic Sy*, Vol. 9, 2012.

[14] D. S. Pankaj and M. Wilscy, "Comparison of PCA, LDA and Gabor Features for Face Recognition Using Fuzzy Neural Network," *in Advances in Computing and Infor-mation Technology: Springer*, 2013, pp. 413-422.

[15] M. Safayani, M. T. Manzuri Shalmani and M. Khademi, "Extended Two-dimensional Pca for Efficient Face Rep-resentation and Recognition," *in Intelligent Computer Communication and Processing*, 2008. ICCP 2008. 4th International Conference on, 2008, pp. 295-298.

[16] Faces," Vol. 4, No. 3, 1987, pp. 519-524. doi:10.1364/JOSAA.4.000519

[17] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp. 71-86. doi:10.1162/jocn.1991.3.1.71

[18] "Citation Report Title=(PCA face recognition)" Vol. 2013, 2013.

[19] K. S. Sodhi and M. Lal, "Face Recognition Using PCA, LDA and Various Distance Classifier," *Journal of Global Research in Computer Science*, Vol. 4, 2013, pp. 30-35.

[20] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min and W. Worek, "Overview of the Face Recognition Grand Challenge," *in Computer vision and pattern recognition*, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, pp. 947-954.

[21] D. Srinivasulu Asadi, Ch. DV Subba Rao and V. Saik-rishna "A Comparative Study of Face Recognition with Principal Component Analysis and Cross-Correlation Technique," *International Journal of Computer Applications* Vol. 10, 2010.