

Identification of Best Suitable Samples for Training Database for Face Recognition using Principal Component Analysis with Eigenface Method

Divyakant T. Meva
Research Scholar
Saurashtra University,
Rajkot, India

C. K. Kumbharana, Ph.D.
Associate Professor,
Dept. of Computer Science,
Saurashtra University
Rajkot, India

ABSTRACT

Security is one of the most important aspects in today's computer environment. Especially, person authentication now a day is necessary to maintain security of computer based systems. Biometric authentication methods are becoming popular since last decade. Face recognition is one of the mature and popular biometric authentication methods. Today, with this paper, discussion on identifying best suitable samples for generating training database has been done. PCA based face recognition approach using Eigenface method has been discussed for the said purpose.

General Terms

Face recognition

Keywords

Face recognition, PCA, Eigenface, Euclidean distance

1. INTRODUCTION

Biometric authentication systems are one of the most popular and interesting area for research in today's digital world. Biometric authentication methods can be divided with two categories:

1. Methods based on physical traits of human
2. Methods based on behavioral traits of human

Physical biometrics methods include fingerprint, face, hand geometry, ear shape, palmprint, iris and retina scanning. While behavioral biometrics methods include voice, gait, keystroke recognition etc. [1]

Face recognition methods can be implemented with two different ways: Identification and Verification. Application of biometrics at Airports used for surveillance is an example of identification. While applying biometrics at border of the country is an example of verification. Face recognition can be used on both the ways. As face recognition is one of the mature technologies, we cannot doubt on its applicability [2].

1.1 Different Face Recognition Techniques

Zhao et al. in their survey categorized face recognition techniques in three ways:

1. Appearance based techniques
2. Feature based techniques
3. Hybrid techniques

Appearance base technique uses holistic texture features and they are applied to either whole face or specific regions of the face image. Feature based technique uses geometric facial

features like mouth, eyes, cheeks etc. and geometric relationship between these features. As human being, we are used to for matching face as a whole with holistic approach as well as with the help of features of the faces, which can be considered as hybrid techniques [2].

In the appearance base approach, one of the widely used methods is techniques based on Principal Component Analysis (PCA). First successful experiment of face recognition was made by Turk and Pentland in 1991 [3]. Moghaddam and Pentland in 1997 [4] extended this method to Bayesian approach. Face recognition system using Linear Discriminant Analysis (LDA) has been successful. LDA training is performed with Scatter Matrix Analysis. Another technique was suggested by Lin et al. in 1997 [5]. They proposed system based on Probabilistic Decision Based Neural Network (PDBNN).

The earlier methods of face recognition were working with structural matching. These methods were based on geometry of local features. They were using the features like width of the head, distance between eyes, and the features of mouth, nose and chins [6]. Hidden Markov Model (HMM) based methods use lines of pixels covering eye, nose, mouth, chins etc [7].

This approach used holistic methods and local features both. Modular Eigenface approach uses global Eigenfaces and local eigenfeatures both [8].

2. FACE RECOGNITION USING PCA WITH EIGENFACE METHOD

Principal component analysis transforms a set of data obtained from possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of components can be less than or equal to the number of original variables. The first principal component has the highest possible variance, and each of the succeeding components has the highest possible variance under the restriction that it has to be orthogonal to the previous component. It is required to find the principal components; here eigenvectors of the covariance matrix of facial images [9].

(1) First step

In the first step, it is required to form a training data set. 2D image can be represented in 1D vector with concatenation of rows. Image will be transformed in a vector with length $N=m*n$.

$$I = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \xrightarrow{\text{CONCATENATION}} \begin{bmatrix} x_{11} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{2n} \\ \vdots \\ x_{mn} \end{bmatrix} = x.$$

Create a matrix of learning images X with M vectors of length N. Then center the matrix. Determine vector of mean values and subtract that vector from each image vector.

Average vectors are arranged to form a new training vector with size (NxM).

(2) Second step

Second step will calculate covariance matrix C, and find its eigenvectors and eigenvalues. Covariance vector C will have dimension NxN. From this matrix, we can get N eigenvectors and eigenvalues. Rank of covariance matrix is limited by number of images in learning set. Eigenvector associated with highest eigenvalue reflects the highest variance and one associated with the lowest eigenvalue, the smallest variance.

The vectors should be sorted by eigenvalues so that first vector corresponds to the highest eigenvalue. These vectors will be normalized next. This will create a new matrix where each vector is a column vector. The dimension of this matrix will be NxN, where N represents desired number of eigenvectors. Each original image can be reconstructed by adding mean image to weighted sum of all vectors.

(3) Third step

Third and last step is the recognition of faces. Image of the person required to find in training database is transformed into a vector P, reduced by the mean value and projected with a matrix of eigenvectors (Eigenfaces).

Classification is done by identifying distance person's eigenvector and each vector of matrix Y. Euclidean distance is the most common method. Other methods can also be applied.

If the minimum distance between test face and training faces is higher than threshold value, then person will be unknown otherwise it will be known one.

3. EXPERIMENT MODEL

Here with our system, the objective is to find out samples for training set and testing set from enrolled samples.

The initial database is created by taking 30 objects with each enrolled with 10 samples of face. Total no. of samples is 300. The problem here is how to identify samples for training database and testing database. We planned to keep 2 samples with training database and 8 samples in testing database for each person.

As shown in fig. 1 , initially two sample images of each person are included in train database. Features are extracted from images by applying Eigenface method. Rest 240 images are added in test database. Now, by applying Euclidean distance method for Eigenface approach, minimum distance will be calculated. The image with minimum distance from train database sample with test sample will be identified. And success rate will be counted. In this system, we have taken five rounds of the above discussed flow of process. Initially, sample image 1 and 2 of each person were placed in train

database and rest of the images was placed in test database. In the second round, images 3 and 4 were placed in train database and compared with rest test samples to identify success rate. Round three contained sample 5 and 6, round 4 contained samples 7 and 8 and round five contained samples 9 and 10 as train database. The results of this approach to find best images suitable to add in train database are shown in section 4.

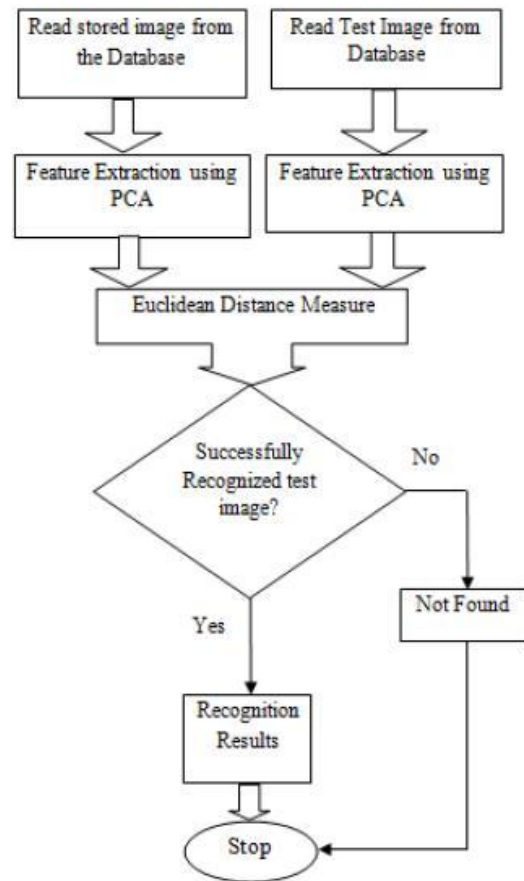


Fig. 1: PCA base face recognition model

4. EXPERIMENT RESULTS

The discussed method in section 3 generated following results:

Considering a database of 60 sample facial images in Train Database of 30 persons (2 for each person) and taking database of 240 sample facial images in Test Database of 30 persons (8 for each person)

No. of persons = 30

No. of images in Train Database = 60 (Two for each person)

No. of images in Test Database = 240 (Eight for each person)

Table 1: Success rate of five cases

Case number	Success rate in %	Samples identified successfully (Out of 240)
1	63.75	153

2	57.50	138
3	74.58	179
4	70.41	169
5	62.08	149

Case 1 – Image 1, 2 as train database and rest as test database

Case 2 – Image 3, 4 as train database and rest as test database

Case 3 – Image 5, 6 as train database and rest as test database

Case 4 – Image 7, 8 as train database and rest as test database

Case 5 – Image 9, 10 as train database and rest as test database

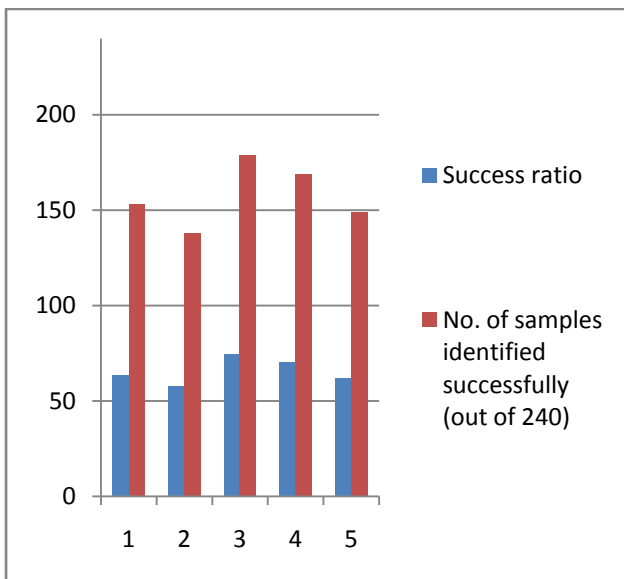


Fig.2: Success ratio and no. of samples identified correctly

The experiment shows that sample 5 and 6, if considered as train database samples results in greater GAR.

5. CONCLUSION

From the experiments, we came to know that this method can be implemented in real life environment to create train database with best suitable samples to generate best results in authentication. This method can be applied to other techniques of face recognition also.

6. REFERENCES

- [1] Divyakant Meva, CK Kumbharana, Amit Kothari, “ The study of adoption of neural network approach in fingerprint recognition”, International Journal of Computer Applications, Volume 40– No.11, February 2012, pp. 8-11
- [2] Divyakant Meva, CK Kumbharana, “Study of different trends and techniques in face recognition”, International Journal of Computer Applications, Volume 96– No.8, June 2014, pp. 1-4
- [3] Turk, M. and Pentland, A., “Eigenfaces for recognition”, J. Cogn. Neurosci. 3, 1991, 72–86
- [4] Moghaddam, B. and Pentlad, A., “Probabilistic visual learning for object representation”, IEEE Trans. Patt. Anal. Mach. Intell. 19,1997, 696– 710
- [5] Lin, S. H., Kung, S. Y., And Lin, L. J., “Face recognition/ detection by probabilistic decision based neural network”, IEEE Trans. Neural Netw. 8, 1997, 114–132
- [6] Kelly, M. D., “Visual identification of people by computer”, Tech. rep. AI-130, Stanford AI Project, Stanford, CA.
- [7] Samaria, F. And Young, S., “HMM based architecture for face identification”, Image Vis. Comput. 12, 1994, 537–583.
- [8] Pentland, A., Moghaddam, B., Starner, T., “View-based And modular eigenspaces for face recognition”. Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 1994.
- [9] Marijeta Slavković, Dubravka Jevtić, “Face recognition using Eigenface approach”, Serbian Journal of Electrical Engineering, Vol. 9, No. 1, February 2012, 121-130