International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Improvement in Recognition Rate by Using Linear Regression with Principal Component Analysis

Tanvi Ahuja[1], Vinit Agarwal[2]

[1]M.Tech, Suresh Gyan Vihar University, Jaipur, India

[2]Assistant Professor, Suresh Gyan Vihar University, Jaipur, India

Abstract: ***In this paper we propose a face recognition system which uses the combination of Regression and PCA (Principal Component Analysis). We use regression for classification of Eigen vectors generated and PCA for extraction of facial features. The results obtained are more accurate than the previous approaches.***

Keywords: Face Recognition, Eigen Vectors, PCA, Linear Regression, Euclidean Distance

1. Introduction

Face recognition is one of the most germane applications of image analysis. It's a factual challenge to formulate an automated system which equals human propensity to recognize faces. A facial recognition system is a computer application for automatically identifying or verifying person from a digital image using webcams or video camera. At present times face recognition has become an imperative technology in the field of biometrics. The main propulsive force behind the analysis of face recognition technologies is high degree of accuracy of recognizing face.

In this research we mainly focus on two approaches: First is Principal Component Analysis which is used here for feature extraction. PCA is used an approach for detecting the similarities of differences in the data. In the view of fact that it is not easy to find out the patterns in data because data is of high dimensionality. So we use PCA to reduce the dimensions of data. Data here means the images taken for research purpose. The main advantage of using PCA is that there is no actual loss of information during reduction in dimensions. In this we have to follow some steps: First is acquisition of data, then we have to calculate the mean of all images. Then the calculated mean is subtracted from all images and a correlation matrix is generated. In next step eigen values and eigen vectors of correlation matrix is calculated. At last feature vectors from the generated values are extracted.

Another main technique used in our research is Regression. In simple terms regression is calculation of the nearest similar value. In the field of statistics regression is finding the relationship between one dependent variable and independent variable. So we have used this technique for classification of feature vectors.

Research in the field of programmed face recognition was begun in the 1960's by ***Bledsoe.*** After him the first fully programmed face recognition system was produced by ***Kanade*** in 1977. In 1991, ***Mathew Turk*** and ***Alex Pentland*** presented their research on 'Eigenface'. ***Imran Naseem*** and ***Roberto Tongeri*** in ***2010[1]*** present an innovative way of face identification by devising the pattern recognition issue in terms of linear regression. In ***2012***, ***Dulal Chakraborty***, ***Sanjit K. Saha*** and ***Md. Al-Amin Bhuiyan*** present their paper namely ***'Face Recognition using PCA and Eigen vector' [2].*** The research presented by them shows a system that has the ability to recognize a person's face by correlating facial structure to that of a known person by using the frontal view. In ***2013***, a research was forwarded by ***Krishnapriya B*** and ***Jude Hemanth*** which overcome the problem of illumination by using PCA+LRC[3].

2. Proposed Methodology

In the proposed work we have used the two main algorithms for face recognition: One is the Principal Component Analysis and second one is Linear Regression. In our approach we have taken ORL Database of 400 images of 40 persons with 10 images per person. These persons in 400 images have different pose, expressions or some images are with glasses. All the 10 images of per person have little variation in angle of rotation. In the proposed research we have performed experiments on several images and all the experiments are performed by using MATLAB 7.10.0.

3. Proposed Algorithm:

In this we have proposed two algorithms. One is for loading of database into the matrix and another is for face recognition:

A. Algorithm for Loading the Database into the Matrix:

Step 1: Zeros or Empty Vector is declared in accordance with the size of image data.
Step 2: Reshape the matrix into the vector with elements collected from the various subjects.
Step 3: Convert the matrix into unsigned 8 bit numbers
Step 4: Provide each image a separate space into the dataset.

B. Algorithm for Face Recognition

Step 1: Mean and variance of the loaded dataset are made almost equal using mean and variance normalization technique. Mean and variance of one of the face image from the collections is treated as preferred mean and variance.

Step 2: Calculated mean is subtracted column by column from the matrix and then reshape the matrix into the vector whose elements are collected column by column from the matrix.

Step 3: Correlation matrix is generated by using the collected vectors. Evaluation of Regression Coefficients is done by computing the Eigen vectors from the corresponding Eigen values of the correlation matrix.

Step 4: Eigen vectors calculated are reshaped into the matrix of the original size. They are called Eigen faces.

Step 5: Normalization of orthogonal Eigen faces is done by normalizing the vectors to make them orthonormal.

Step 6: Feature vectors are calculated known as Principal Components for every face image matrix as the inner product of Eigen basis vectors and the reshaped face image matrix.

Step 7: Mean vector is calculated from the collected feature vectors of the same face is treated as the template assigned to that corresponding face of person. This is repeated for other persons also. As a result, one template is assigned to every person and they are stored in the database.

Step 8: To classify the unknown image as one among the four categories, template is computed as the inner product of Eigen basis vectors (Eigen faces) with reshaped normalized unknown image. The template thus obtained is compared with the group of templates stored in the database using Euclidean distance.

Step 9: Template equivalent to the minimum Euclidean distance is selected and the person corresponding to that template is confirmed as the identified person.

Implementation Details

In this section we are going to discuss about the implementation steps which we have used in our research:

- Firstly the database of 400 images is loaded in the 2×2 matrix. Then randomly an image is selected to be searched. Then from the matrix of 400 images the image to be searched is excluded and again a new database with 399 images is formed.
- Mean of the 399 images are calculated and then subtracted from the images and a correlation matrix is generated.
- Eigen vector of the correlation matrix is calculated. Then suppose we have taken 20 images for which Eigen vectors are calculated and signature for images is calculated.
- Then Eigen vector for the image to be searched is calculated and results for matching are obtained with minimum Euclidean distance. The image with nearest distance is given as output.
- In this work we also evaluated the recognition time for each image and the overall recognition time for 399 images.

We have followed certain steps which are discussed below:

The first step is image acquisition, in this image is loaded into the database. Next step is pre-processing and edge detection, in which detection and removal of noise is performed. After pre-processing eigen vectors are calculated. With the help of these eigen vectors regression coefficients are computed. Then in next step normalization of eigen faces is performed which are obtained after the regression. In the last step feature vectors are detected by using PCA and minimum Euclidean distance is evaluated. On the basis of minimum distance image is identified. If image is wrongly identified then the process is stopped and all the steps are repeated. In the flowchart presented in figure (1) below we have shown the steps which we have followed throughout the research

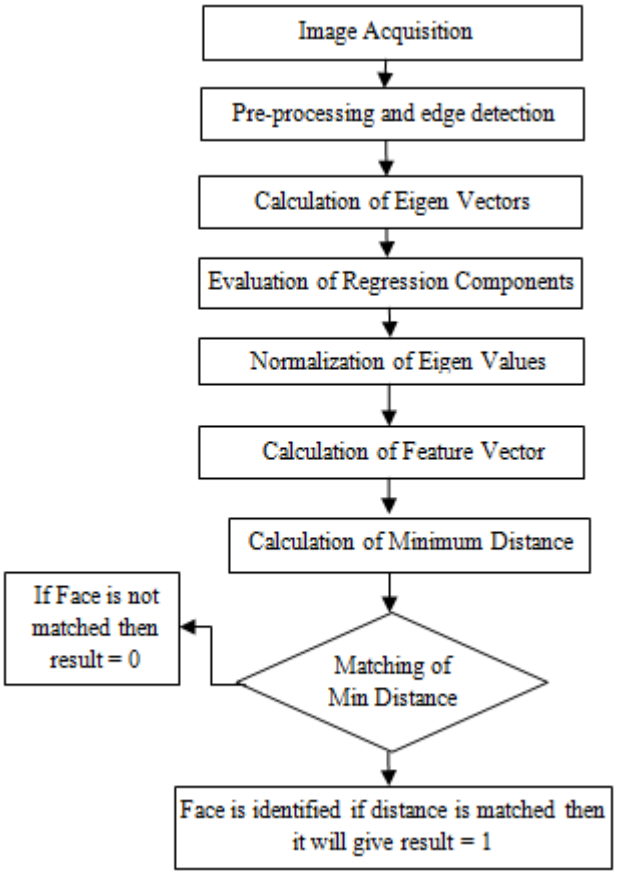

Figure 1: Flowchart representing the proposed approach

4. Experimental Results

When we run our codes on the command prompt window we got the following results. The results we get are based on the images having different pose, facial expressions, angle of rotation etc:

Figure 2: Image showing recognition procedure's success

Figure 3: Result showing success when there is variation in the image taken from closed view and distant view

Figure 4: Results obtained when compared face is tilted and obtained resultant face is straight.

The results shown above are obtained when we performed testing on different image. We have performed testing on 400 images by randomly dividing into five subsets. Subset 1 containing 20 images of two persons with 10 images person. Subset 2 comprises of 40 images with 10 images of 4 persons. Subset 3 contains 70 images of 7 different persons with 10 images per person. Subset 4 contains 120 images of 12 subjects and Subset 5 comprises of 150 images. Recognition rate and testing time of experiments are shown in the Table 1 below:

Table 1: Results obtained from proposed algorithm after testing 400 images

Subsets taken for Testing	*Recognition Rate*	*Testing Time(in seconds)*
Subset 1	100%	0.7419
Subset 2	97.14%	0.8615
Subset 3	84%	0.8126
Subset 4	95%	0.9132
Subset 5	88%	0.9868
Mean	**92.8%**	**0.8632**

As we can see in the table above the mean recognition rate obtained after applying the proposed approach is 92.8%. In our approach we estimate testing time as an extra factor. On an average the time required for testing the images is 8 to 9 seconds.

The graph in figure 2 which is screened below shows comparison of our approach with the previous approaches. As we can see in the graph in figure 2 the cone in yellow colour represents the recognition rate obtained from proposed approach i.e., LR+PCA. The other three cones represents the three different approaches which are previously used. The results obtained by using our approach are more accurate then the previous approaches. There is a sharp increment of 13.58%.

5. Conclusion

In this research a new approach for face recognition is proposed. The algorithm proposed is a combination of linear regression and principal component analysis. The proposed LR+PCA approach provides remarkable increase in recognition rate. This approach can be widely used for pattern recognition like iris recognition, ear recognition etc.

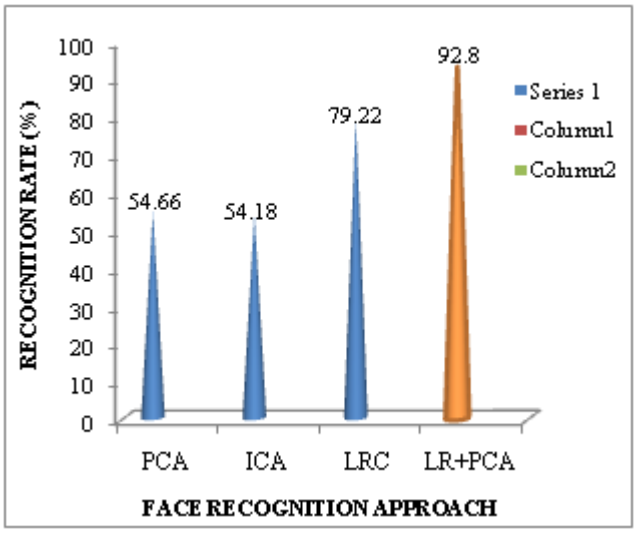

Figure 2: This graph shows comparative results of our approach with previous approaches

References

[1] Imran Naseem, Roberto Togneri, Senior Member, IEEE, and Mohammed Bennamoun," Linear Regression for Face Recognition", in IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 32, NO. 11, NOVEMBER 2010.

[2] Dulal Chakraborty, Sanjit K. Saha and Md. Al-Amin Bhuiyan," Face Recognition using PCA and Eigen vector" International Journal of Computer Applications (0975 – 8887) Volume 50 – No.10, July 2012.

[3] Krishnapriya B and Jude Hemanth," Comparison of Principal Component Analysis and Modified Principal Component Analysis for Face Recognition" International Conference on Advances in Electrical and Electronics Engineering (AEEE), 13th Jan 2013.

[4] A. K. Jain, R. P. W. Duin, and J. C. Mao, "Statistical pattern recognition: a review," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4–37, 2000.

[5] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance", In 3rd International Conference on Audio- and Video-Based Biometric Person Authentication 2001, Halmstad, Sweden, pages 90–95, 2001.

[6] A. Gunjan Dashore and Dr. V. Cyril Raj," An Efficient Method for Face Recognition Using Principal Component Analysis (PCA)", in 2012

[7] Coo ray Saman and O'Connor," Facial Features and Appearance-Based Classification for Face Detection in Color Images", 2004.

[8] Kresimir Delac, Mislav Grgic and Sonja Grgic," Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set", 2006.

[9] Jan Flusser, Usman Ullah Sheikh, Mohammad Khansari, Siti Mariyam Shamsuddin and Kourosh Jafari-Khouzani," Rotation and noise invariant near infrared face recognition by means of Zernike moments and spectral regression discriminant analysis", in Journal of Electronic Imaging 22(1), 013030 (Jan–Mar 2013).

[10] Kuldeep Singh Sodhi and Madan Lal," Comparative Analysis of PCA-based Face Recognition System using different Distance Classifiers", in 2013