LDA AND PCA Mashup for Face Recognition System

Harmeet Singh

University Of Missouri, Kansas City

Hsp79@mail.umkc.edu,harmeet.singh57@gmail.com

I Harmeet Singh declares that I am the sole author of this paper and the references used for writing this paper are clearly cited in reference section.

Abstract

In this paper, we propose a methodology for the real-time effective face recognition system. Since these days nature of application domain requires real time result and better accuracy, it poses a serious challenge. To address this challenge, we studied various classification techniques, namely, support vector machine (SVM), linear discriminant analysis (LDA), K-nearest neighbor (KNN) and Principle component (PCA). To speed up data retrieval we propose a feature reduction technique using Principle component analysis (PCA) to facilitate near real time face recognition along with better accuracy. We apply LDA after we reduce the number of features by PCA. Hence, we test new classification approach with combination of PCA with LDA over At & t dataset and demonstrate the effectiveness of LDA with PCA.

1. Introduction

With the development of digital technology these days, we need to consider the future of computing in our daily activities and how it will affect our lives[1,2]. So we would like to develop the face recognition system with an advanced intelligent services and with better accuracy. **Facial recognition** (or **face recognition**) is a biometric method of identifying an individual by comparing live capture or digital image data with the stored record for that person. **Facial recognition systems** are commonly used for security purposes but are increasingly being used in a variety of other applications [2]. **Biometric** verification is any **means** by which a person can be uniquely identified by evaluating one or more distinguishing biological traits. Unique identifiers include fingerprints, hand geometry, earlobe geometry, retina and iris patterns, voice waves, DNA, and signatures. So we have used multi classifier system (Principal component analysis (PCA) and Linear discriminant analysis(LDA)) [3,4] for the best accuracy over training dataset.

One of the most popular techniques for linear transformation in feature space is **Principal Component Analysis** (PCA) [1,3,4]. PCA reduces the dimensions by rotating feature vectors from a large highly correlated feature space (image space) to a smaller feature space (face space) that has no sample covariance between the features.

After applying PCA to reduce the face space to a lower dimensional manifold, a **Linear discriminant classifier** is typically used. The LDA is a powerful technique for predicting seen data; however, it cannot predict unseen data. Furthermore, LDA may not work for non-linearly separable dataset. LDA tries to achieve a projection that best discriminates between the different subjects.

The **Training set** is defined to be all the images of subjects that are available for constructing the face space. So here over for this multi classifier system we have used a standard **At &T** database. We propose to utilize multiple classifier systems or ensembles in the biometric problem of 2-D face recognition. We randomly sample from the acquired images of a subject to construct face spaces.

The organization of this paper is as follows. Method used and its approach work is discussed in section 2. Our results and various discussions for face recognition system is described in section 3. The paper is concluded with discussions in section 4. The reference papers used for writing this paper are listed in section 5. The table with results comparison is shown in section 6.

2.Method used and its Implementation

First, we extract features for images. Second, we train classifier for training images and generate model for classes. Finally, these classification models will be used to predict test images. For the classifications, we have used multi classification approach using both principal component analysis and linear discriminant analysis. Then Scores of PCA and LDA are combined using minimum, maximum and Average rules and the scores belonging to each test sample are averaged of the scores obtained on comparing the test sample with all the templates of a particular identity. Below we have elaborated the methodology used.

2.1. Feature extraction (step1)

A two-dimensional face image will be represented as a vector. This vector will be formed by concatenating each row (or column) of the image. For example, 20×22 a two-dimensional image will be represented by a vector of 440 dimensions.

2.2. Classifications (step 2)

In this paper, we study various classifiers, namely, Principal component analysis (PCA) and Linear Discriminant Analysis (LDA). These are discussed in detail as follows.

2.2.1. PCA

The raw feature vectors are a concatenation of the gray-level pixel values from the images. Let us assume there are m images and n pixel values per image. Let Z be a matrix of (m,n), where m is the number of images and n is the number of pixels (raw feature vector). The mean image of Z is then subtracted from each of the images in the training set, $\Delta Zi = Zi - E[Zi]$. Let the matrix M represent the resulting "centered" images; $M = (\Delta Z1, \Delta Z2, ..\Delta Zm)T$. The covariance matrix can then be represented as:

$$\Omega = M.M^T . \Omega$$

is symmetric and can be expressed in terms of the singular value decomposition $\Omega = U.\Lambda.UT$, where U is an m x m unitary matrix and $\Lambda = diag(\lambda 1, ..., \lambda m)$. The vectors U1, ..., Um are a basis for the m-dimensional subspace. The covariance matrix can now be re-written as

$$\Omega = m\ i=1\ \zeta i.Ui$$

The coordinate $\zeta i, i \in 1, 2, ...m$, is called the ζth i principal component. It represents the projection of ΔZ onto the basis vector U. The basis vectors, Ui, are the principal components of the training set. Once the subspace is constructed, recognition is done by projecting a centered probe image into the subspace, and the closest gallery image to the probe image is selected as the match.

2.2.2 Linear Discriminant Analysis(LDA)

LDA tries to achieve a projection that best discriminates between the different subjects. PCA can be used to reduce the dimensionality before applying LDA. The Fisher face is constructed by defining a d dimensional subspace in the first d principal components [7,8]. Fisher's method [8] finds the projecting vectors W, such that the basis vectors in W maximize the ratio of the determinant of the inter-class scatter matrix S_b and the determinant of the intra-class scatter matrix S_w.

$$W = \text{argmax } |WT\ S_{bW}\ |WT\ S_{wW}|$$

Let us define the number of subjects to be m and the number of images (samples) per subject available for training to be s_i, where i is the subject index. Then Sb and Sw can be defined as:

$$S_w = m\ i{=}1\ \ x_k{\in}X_i\ (x_k - \mu_i)(x_k - \mu_i)^T$$

$$S_b = m\ i{=}1\ s_i(\mu i - \mu)(\mu i - \mu)^T$$

and where µi is the mean of vector of samples belonging to the class (or subject) i, µ is the mean vector of all the samples. S_w may not be well estimated if the number of samples is too small.

2.3 Fusion at Score level(step 3)

Finally, once we get all the scores of PCA and LDA, these are combined using minimum, maximum and Average rules. In our data set the combination of PCA and LDA works best with average of both as the accuracy comes best with this average rule. Then in our second part the scores belonging to each test sample are averaged of the scores obtained on comparing the test sample with all the templates of a particular identity. Here PCA with multi-instance fusion tends out to come with best accuracy.

3.Results

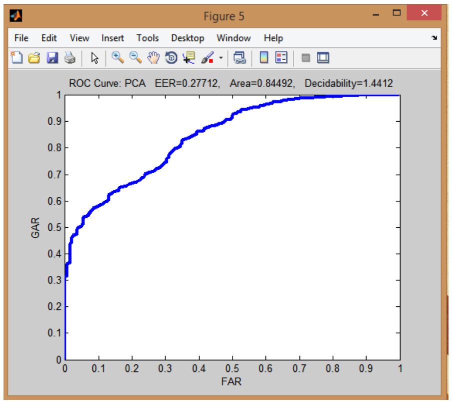
Here we have used the AT&T database which is freely downloadable and contains a set of 400 face images (http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html).
There are ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses).

For classification, we applied PCA,LDA and PCA+LDA. In this study, we report our results in terms of ROC curves depicting accuracy as follows.

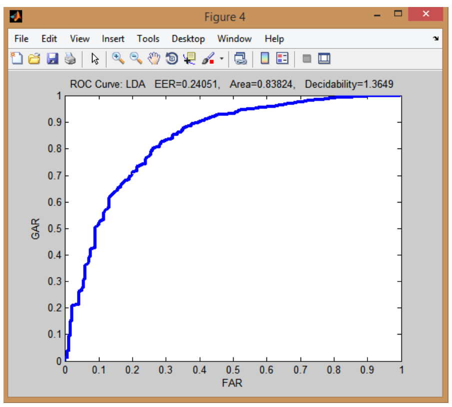
- Figure 1 depicts the ROC curve for performance of PCA over mentioned datasets. Here, accuracy means percentage of correctly classifying face images. Here accuracy tends to be 84.492%.

- Figure 2 depicts the ROC curve for performance of LDA over mentioned datasets. Here accuracy tends to be 83.824%.

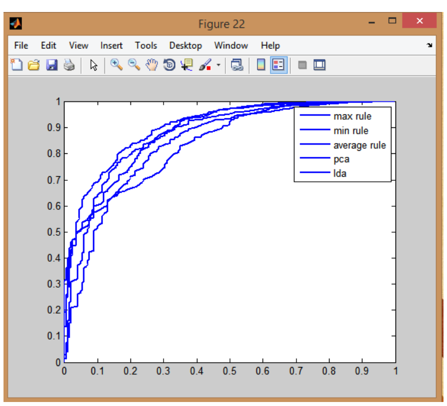
- Figure 3 depicts the ROC curve for performance of overall curves using minimum, maximum and average sum rule scores of PCA and LDA. Here accuracy tends to be 83.824%.



Figue1-ROC curve using PCA



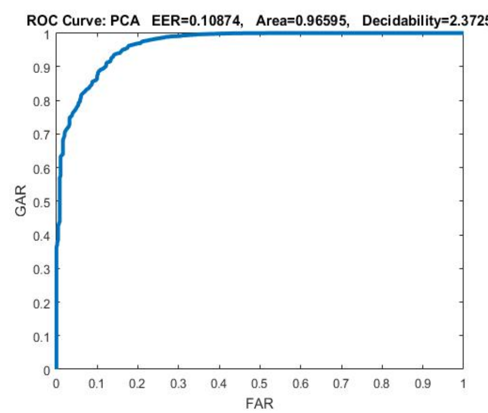
Figue2-ROC curve using LDA



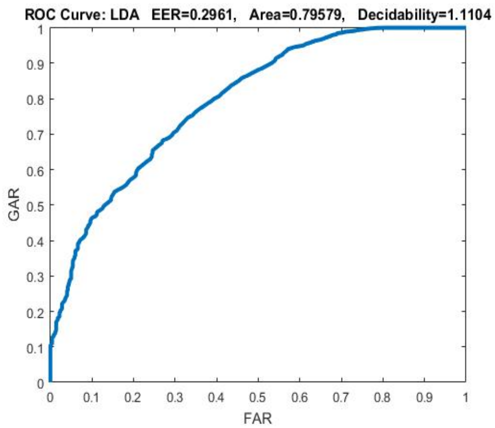
Figue3- Overall ROC curves

In second part of the project as we compared the scores belonging to each test sample which are the average scores of obtained test sample with all the templates of a particular identity. In this study, we report our results in terms of ROC curves depicting accuracy as follows.

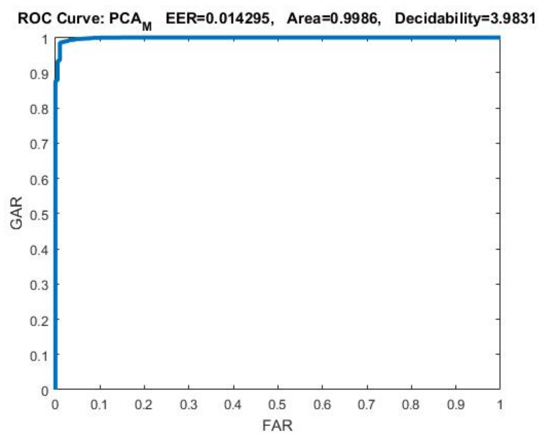
- Figure4 depicts the ROC curve for performance of PCA without multi-instance fusion over mentioned datasets. Here, accuracy means percentage of correctly classifying face images. Here accuracy tends to be 96.595%.
- Figure5 depicts the ROC curve for performance of LDA without multi-instance fusion over mentioned datasets. Here accuracy tends to be 79.575%.
- Figure6 depicts the ROC curve for performance of PCA with multi-instance fusion over mentioned datasets. Here accuracy tends to be 99.86%.
- Figure7 depicts the ROC curve for performance of LDA with multi-instance fusion over mentioned datasets. Here accuracy tends to be 89.619%.
- Figure8 depicts the overall combination of all ROC curves.



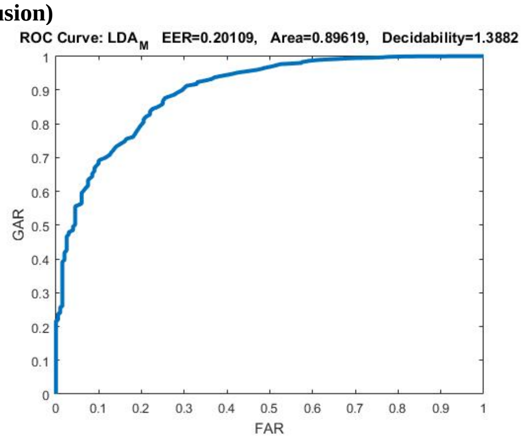
Figue4-ROC curve using PCA (without multi-instance fusion)



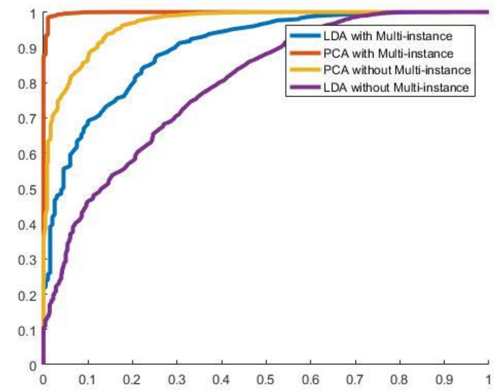
Figue5-ROC curve using LDA(without multi-instance fusion)



Figue6-ROC curve using PCA(with multi-instance fusion)



Figue7-ROC curve using LDA(with multi-instance fusion)



Figue8- Overall ROC curves

In conclusion, in this study, we can say that for face recognition, LDA and PCA are effective classification techniques. Furthermore, to speedup retrieval for LDA we advocate to the usage of PCA.

4.Conclusion

In this paper, we propose a near real time face recognition system for the best accuracy which tended almost 99.86% for PCA with multi-instance fusion. For this, first, we apply various classification techniques, namely, PCA and LDA. Finally, we show that PCA with multi instance fusion will be most effective as depicted in the table.

The research reported in this paper is part of project done at the University of Missouri, Kansas City [29], [30]. We majorly conducted research Biometrics of which Face Recognition system is a part of. And we also believe that biometrics in general and automatic face recognition in particular will have many applications in surveillance and privacy preserving surveillance [2] [4], [5].

5.References

[1] Nitesh V. Chawla and Kevin W. Bowyer, Designing Multiple Classifier

Systems for Face Recognition, University of Notre Dame, 2006.

[2] Pallabi Parveen and Bhavani Thuraisingham, Face Recognition using Multiple Classifiers,University of Texas at Dallas, 2008.

[3] M. Turk and A. Pentland, “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[4] G. Shakhnarovich and G. Moghaddam, Handbook of Face Recognition, ch. Face recognition in subspaces. Springer-Verlag, 2004

[5] F.Zuo and P.H.N.de With, Fast facial feature extraction using a deformable shape model with Haar-wavelet based local texture attributes, to be published in Proc. ICIP, 2004.

[6] F.Zuo and P.H.N.de With, Two-stage face recognition incorporating individual-class discriminant criteria, Proc. WIC 2004, pp. 137–144, 2004.

[7] Jones, MJ; Viola, P., "Fast Multi-view Face Detection", IEEE Conference on Computer Vision and Pattern Recognition

[8] B. Thuraisingham and P. Pallabi, Data Mining for Biometrics Applications, UTD Technical Report, 2006.

[9] B. Thuraisingham, Privacy Preserving Data Mining: Developments and Directions, Journal of Database Management, Special Issue in Data Management for National Security, January 2005.

[10] J. R. Beveridge, K. She, B. Draper, and G. Givens, “A nonparametric statistical comparison of principal component and inear discriminant subspaces for face recognition,” in IEEE Conference on Computer Vision and Pattern Recognition, pp. 535 – 542, 2001.

6.Appendix

The table below clearly depicts the results obtained and validate our classifier.

S.No	Experiment	EER	Area Under Curve	Decidability
1	PCA only	0.27712	0.84492	1.4412
2	LDA Only	0.24051	0.83824	1.36349
3	MAX of PCA & LDA	0.20058	0.8738	1.567
4	MIN of PCA & LDA	0.21724	0.88119	1.6952
5	Average of PCA & LDA	0.18152	0.9025	1.8499
6	LDA with multi-instance fusion	0.20109	0.89619	1.3882
7	PCA with multi-instance fusion	0.014295	0.9986	3.9831

Table depicting actual result values

The above results show that PCA with multi instance fusion tends to show the best results with an accuracy of 99.86%. The ROC curves generated used are already shown above in section 3 and all the formulas used are already defined in section2.