A Realtime Face Recognition system using PCA and various Distance Classifiers

Project Report
for the course
CS676: Computer Vision and Image Processing

by

Deepesh Raj
Roll: Y7027134

under the guidance of

Dr. Amitabha Mukerjee
and
Dr. Prithwijit Guha

Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur
Spring 2011

A Realtime Face Recognition system using PCA and various Distance Classifiers

Spring, 2011

Abstract

Face recognition is an important application of Image processing owing to it's use in many fields. The project presented here was developed after study of various face recognition methods and their efficiencies. An effective and real time face recognition system based on OpenCV and C++ is developed in the project. The system was tested on YALE Face database B and ORL Face Database. The recognition produced using 3 different matching techniques are compared and the results have been presented. The correct recognition rate achieved using the Mahalanobis distance is 92.3% in comparison to the 73.1% for the normal PCA with euclidean distance. The system responds to the face recognition queries in less than 0.2 seconds.

Keywords: Real time, Face Recognition, PCA, eigen faces, Mahalanobis distance, Manhattan distance, Yale face database, ORL facedatabase.

1 Motivation

Face recognition has been a sought after problem of biometrics and it has a variety of applications in modern life. The problems of face recognition attracts researchers working in biometrics, pattern-recognition field and computer vision . Several face recognition algorithms are also used in many different applications apart from biometrics , such as video compressions , indexings etc. They can also be used to classify multimedia content, to allow fast and efficient searching for material that is of interest to the user. An efficent face recognition system can be of great help in forensic sciences, identification for law enforcement, surveillance , authentication for banking and security system, and giving preferential access to authorised users i.e. access control for secured areas etc. The problem of face recognition has gained even more importance after the recent increase in the terrorism related incidents. Use of face recognition for authentication also reduces the need of remembering passwords and can provide a much greater security if face recognition is used in combination with other security measures for access control. The cost of the license for an efficient commercial Face recognition system ranges from 30,000 $ to 150,000 $ which shows the significant value of the problem.

Though face recognition is considered to be a very crucial authentication system but even after two decades continuous research and evolution of many face recognition algorithms , a truely robust and efficient system that can produce good results in realtime and normal conditions is still not available. The Face Recognition Vendor Test (FRVT) [8] that has been conducted by the National Institute of Standards and Technology (NIST), USA, has shown that the commercial face recognition systems do not perform well under the normal daily conditions. Some of the latest face recognition algorithm involving machine learning tools perform well but sadly the training period and processing time is large enough to limit its use in practical applications. Hence there is a continuous strife to propose an effective face recognition system with high accuracy and acceptable processing time.

2 Literature Review

The project on face recognition had helped the author to have a detailed survey of a number of face recognition algorithms along with their advantages and limitiations. Some of the important methods studied will be described in this section. Face recognition systems architecture broadly consists of the three following tasks:

1. Acquisition(Detection,Tracking of face-like images)
2. Feature extraction (Segmentation,alignment & normalization of the face image)
3. Recognition

2.1 Face Detection Approaches

Some of the main face detection methods are discussed here.
1) *Knowledge based methods* are developed on the rules derived from the researchers knowledge of human faces. Problem in this approach is the difficulty in translating human knowledge into well-defined rules.
2) *Featured-based methods*: Invariant features of faces are used for detecting texture, skin color. But features from such algorithm can be severely corrupted due to illumination, noise and occlusion.
3) *Template matching*: Input image is compared with predefined face template. But the performance here suffers due to variations in scale, pose and shape.
4) *Appearance-based method*: In template matching methods, the templates are predefined by experts. Whereas, the templates in appearance based methods are learned from examples in images. Statistical analysis and machine learning techniques can be used to find the relevant characteristics of face and non-face images.

2.2 Face Recognition Approaches

LFA [2] method of recognition analyzes the face in terms of local features, e.g., eyes, nose, etc. referred to as LFA kernels. LFA technique offers better robustness against local variations on the facial image in carrying out a match, but does not account for global facial attributes. Neural Network [4] are based on learning of the faces in an example set by the machine in the 'training phase and carrying out recognition in the 'generalization phase. But in order to succeed in a practical set-up, the examples should sufficiently large in number to account for variations in real life situations. Model Matching methods of face recognition (like Hidden Markov Model (HMM)[6]) train a model for every person during model learning and choose the best matching model, given a query image. Here also a big realistic representative models is necessary for good results. A recognition system based on sparse representation [1] computed by l^1 -minimization works with the basic idea of casting the recognition as a sparse representation problem. The main concern in this approach is the presence of sufficiently large number of features and correct computation of sparse representation. It is a robust and scalable algorithms for face recognition based on linear or convex programming.

3 Proposed System

We propose a real time face recognition system based on PCA and Mahalanobis distance . The main challenge for a face recognition system is of effective feature extraction. The proposed system utilizes the Eigen face method is information reduction for the images. There is an incredible amount of information present even in a small face image. A method must be able to break down pictures so as to effectively represent face images rather than images in general. 'base faces are generated and then image being analyzed can be represented by the system as a linear combination of these base faces. Each face that we wish to classify can be projected into face-space and then analyzed as a vector. A k-nearest-neighbor approach, a neural network or even a simple Eucledian distance measure can be used for classification.

The proposed system uses Principal Component analysis for feature extraction and various distance classifiers such as the Eucledian distance, the Manhattan distance and the Mahalanobis distance. The technique used here involves generating the 'eigen faces' then projecting training data into face-space to be used with a predetermined classification method and evaluation of a projected test element by projecting it into face space and comparing to training data.

3.1 Steps Involved

The various steps to calculate eigenfaces are:

A. *Prepare the data* A 2-D facial image can be represented as 1-D vector by concatenating each row (or column) into a long thin vector. Lets suppose we have M vectors of size N (= rows x columns of image) representing a set of sampled images Then the training set becomes: $\Gamma_1, \Gamma_2, \Gamma_3.....\Gamma_M$

B. *Subtract the mean* The average matrix Ψ has to be calculated, then subtracted from the original faces (Γ_i) and the result stored in the variable Φ_i

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

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

C. *Calculate the co-variance matrix* In the next step the covariance matrix A is calculated according to:

$$A = \Phi^T\Phi$$

D. *Calculate the eigenvectors and eigenvalues of the covariance matrix* In this step, the eigenvectors (eigenvectors) X_i and the corresponding eigenvalues λ_i should be calculated.

E. *Calculate eigenfaces*

$$[\Phi]X_i = f_i$$

where X_i are eigenvectors and f_i are eigenfaces.

F. *Classifying the faces* The new image is transformed into its eigenface components. The resulting weights form the weight vector Ω_k^T :

$$\Omega_k = \Omega_k^T(\Gamma_k - \Psi)$$

where k = 1,2,3,4.. and $\Omega_k^T = [\Omega_1\Omega_2....\Omega_M]$

The Euclidean distance between two weight vectors $d(\Omega_i, \Omega_j)$ provides a measure of similarity between the corresponding images i & j.

Figure 1: Sample Eigen face images of normal face images

3.2 Various Distance metrics

Let X, Y be eigenfeature vectors of length n. Then we can calculate the following distances between these feature vectors

3.2.1 Mahalanobis Distance

Mahalinobis space is defined as a space where the sample variance along each dimension is one. Therefore, the transformation of a vector from image space to feature space is performed by dividing each coefficient in the vector by its corresponding standard deviation. This transformation then yields a dimensionless feature space with unit variance in each dimension. If there are two vectors x and y in the unscaled PCA space and corresponding vectors m and n in Mahalinobis space. First, we define $\lambda_i = \sigma_i^2$ where λ_i are the PCA eigenvalues, σ_i^2 is the variance along those dimensions and σ_i is the standard deviation. The relationship between the vectors are then defined as:

$$m_i = \frac{x_i}{\sigma_i} \quad n_i = \frac{y_i}{\sigma_i}$$

$$d(x,y) = \sqrt{\sum_{i=1}^{k}(m_i - n_i)}$$

Where λ_i is the i^{th} Eigenvalue corresponding to the i^{th} Eigenvector.

3.2.2 Manhattan Distance

Also known as the L1- norm or the Manhattan Distance or the City Block Distance. It is defined as follows:

$$d(x,y) = |x - y| = \sum_{i=1}^{k} |x_i - y_i|$$

3.2.3 Euclidean Distance

Also known as the L2-norm or Euclidean Distance. It is defined as follows:

$$d(x,y) = ||x-y||^2 = \sum_{i=1}^{k}(x_i - y_i)^2$$

3.3 Decision on the test

Having calculated the distance between the two feature vectors, the training image closest to the given test image is returned as the result of the query. If the subject of the test image and the subject of the training image closest to the given test image are the same then a correct match is said to have occured or else it is taken as an incorrect match. The above approach is tested on the whole database mentioned below, and the results are presented.

4 Codes

The face recognition systems presented here can extract the features of face and compare this with the existing database. The faces considered here for comparison are still faces. The codes for the project are provided in the folder code.tar.gz folder. A README.txt file inside it provides the instructions of compiling and running the system. The code is written in C++ and utilizes the OpenCV libraries.

5 Database

The developed face recognition codes have been tested against the following two standard databases.A total of about 1200 face images of 78 test subjects with varying illumination and pose variations were used in the project. The databases are described in the following sections.

5.1 The Yale Face Database B

Figure 2: Sample images from Yale Face Database B

The cropped Yale B database used contains 856 grayscale face images in raw PGM format of 38 individuals.[9]. The image resolution is 168(width) x 192 (height) pixels. There are 31 images per subject, one per different facial expression or configuration: center-light, with glasses, happy, left-light, with no glasses, normal, right-light, sad, sleepy, surprised, and wink. Sample images are shown in the fig 5. The file names of the images in this database have been named in a special order mentioning the pose and illumination detail. The first part of the filename begins with the base name 'yaleB' and is followed by the two digit number signifying the subject number and then by the two digit number signifying the pose. The rest of the filename deals with the azimuth and elevation of the single light source direction. An example image yaleB03_ P06A+035E+40.pgm belongs to subject 3 seen in pose 6, and the light source direction with respect to the camera axis is at 35 degrees azimuth (A+035) and 40 degrees elevation (E+40). Here a positive azimuth implies that the light source was to the right of the subject while negative means it was to the left. Positive elevation implies above the horizon, while negative implies below the horizon. A complete description of the pose and illumination is provided on the web protal of the database.[10] The acquired images are 8-bit (gray scale) captured with a Sony XC-75 camera (with a linear response function) and stored in PGM raw format.

5.2 The AT & T Database of Faces (ORL)

The "AT & T Database of Faces" was formerly "The ORL Database of Faces" [11]. It consists of face images 40 distinct subjects with 10 images per subject. 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). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). A preview image of the Database of Faces is available on the official website . A sample preview of images of four different subjects are also shown in figure 3

The files are in PGM format, and can conveniently be viewed on LINUX systems using the default image viewer. The size of each image is 92(width) x 112 (height) pixels, with 256 grey levels per pixel. The images are organised in 40 directories (one for each subject), which have names of the form sX, where X indicates the subject number (between 1 and 40). In each of these directories, there are ten different images of that subject, which have names of the form Y.pgm, where Y is the image number for that subject (between 1 and 10).

Figure 3: sample images from the AT & T Database of Faces

6 Results

6.1 Overall Results

The results of running the codes on the two databases are described below in Figure 4

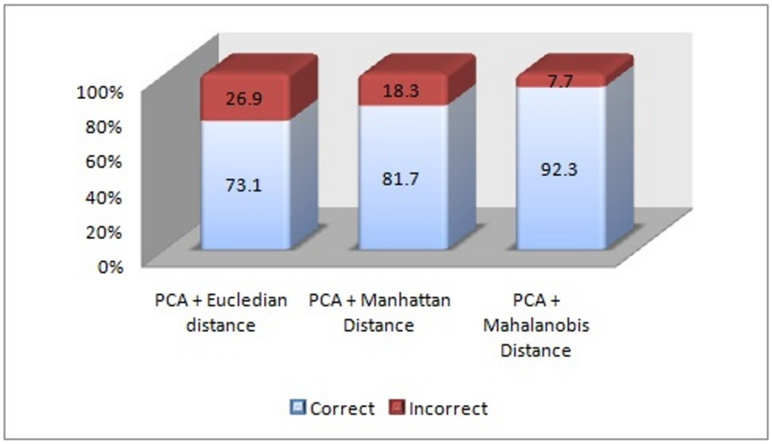

Figure 4: overall Recognition rate (in %) on total of 1176 face images

6.2 The Yale Face Database B

The results of running the codes on the Yale Face Database B are described below in Figure 5 and Table 1

Method Used	Correct	Incorrect	Recognition accuracy
PCA + Eucledian distance	574	282	67.1%
PCA + Manhattan Distance	677	179	79.1%
PCA + Mahalanobis Distance	785	71	91.7%

Table 1: Results on The Yale Face Database B

6.3 AT& T Face Database

The results of running the codes on the AT& T Face Database are described below in Figure 6 and Table 2

Method Used	Correct	Incorrect	Recognition accuracy
PCA + Eucledian distance	286	34	89.4%
PCA + Manhattan Distance	284	36	88.8%
PCA + Mahalanobis Distance	301	19	94.1%

Table 2: Results on The AT& T Face Database

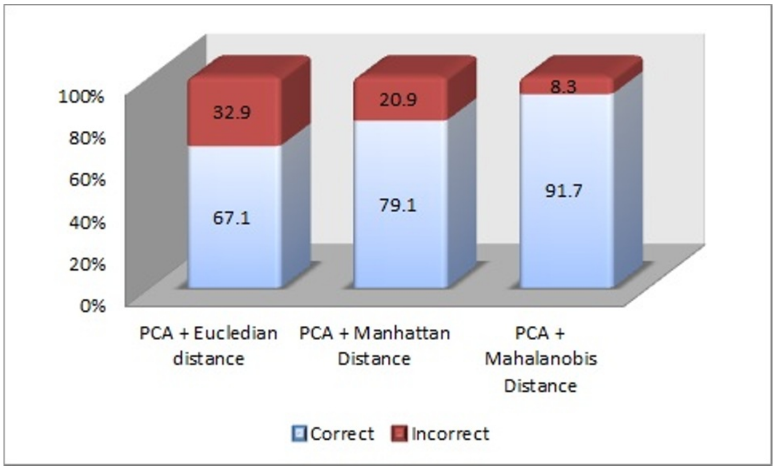

Figure 5: Recognition rate (in %) on Yale Face Database B on total of 856 images

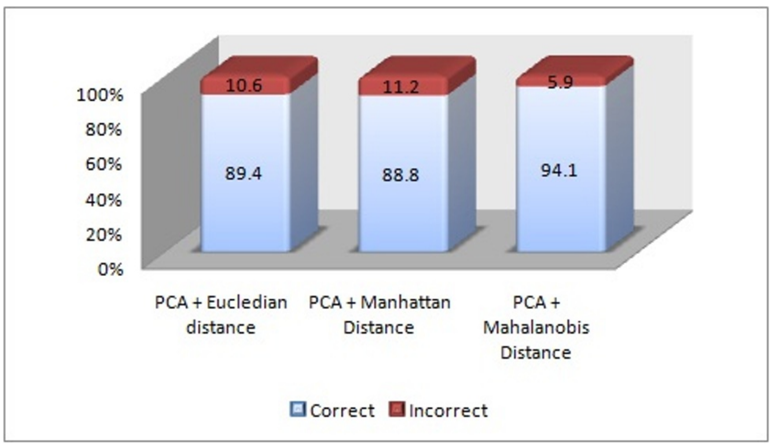

Figure 6: Recognition rate (in %) on AT& T Face Database on total of 320 images

6.4 A Realtime Application : Process time

The processing time of the presented system was also measured. On a Linux based Ubuntu (10.04) operating system with Core 2 duo 2.4GHz Intel processor and 4GB memory, running on our database of 1200 images containing images size 168 x 192 pixels from 78 subjects under various lightning conditions , facial expressions etc of the Yale B database, our OpenCV based C++ implementation of the proposed system using Mahalanobis distance takes 2 seconds for training on the face images and 3 seconds for testing all the images of the database in order to recognise the test subject.

Once trained the system responds to single face recognition queries in less than 0.2 seconds. The system completed a query stream of 900 test images in 3 seconds , taking a per query time slot of just 3 milliseconds.

7 Conclusion

This project on Face recognition had given us an opprtunity to study about many popular methods used in the field of face recognition. The elaborate literary survey provided us with the pros and cons of many recognition systems and the trade-off associated with them. We also came to know that combining two or more techinques can improve the accuracy of system greatly.

In this project we have developed a PCA based face recognition system for feature extraction and matching using various distance classifiers. The Distance classifiers used are Eucledian distance, Manhattan Distance and Mahalanobis distance. The results for all three have been presented. The results clearly shows that a recognition system based on Mahalanobis distance performs far better than the conventional Eucledian distance based classifier. The run time of the codes is also considerably fast with a single query response time of less than 0.2 seconds.

8 Acknowledgement

I extend my sincere thanks to Dr Amitabha Mukerjee and Dr. Prithwijit Guha for their suggestions and guidance in completing the project. The project deliverables were timely scheduled so as to encourage a steady development and thereby leading to the timely completion of the project. The lectures of the course Computer vision and image processing provided a good understanding of the various concepts involved in face recognition. The course assignments were also useful in giving a good first hand experience of the OpenCV libraries which were used in the project codes.

References

[1] Wright, J. and Yi Ma and Mairal, J. and Sapiro, G. and Huang, T.S. and Shuicheng Yan , "Robust Face Recognition via Sparse Representation," *IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2009.* pp.597 -604, 2009

[2] Fazl-Ersi, E.; Tsotsos, J.K.; , "Local feature analysis for robust face recognition," *IEEE Symposium on Computational Intelligence for Security and Defense Applications , page.1-6* July 2009

[3] Delac K., Grgic M., Grgic S., "Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set, *International Journal of Imaging Systems and Technology*, Vol. 15, Issue 5, 2006, pp. 252-260

[4] D. Bryliuk and V. Starovoitov, "Access Control by Face Recognition using Neural Networks and Negative Examples, *2nd International Conference on Artificial Intelligence*, pp. 428-436, Sept 2002.

[5] Wright, J. and Yi Ma and Mairal, J. and Sapiro, G. and Huang, T.S. and Shuicheng Yan "Sparse Representation for Computer Vision and Pattern Recognition" 2010 *Proceedings of the IEEE,* volume 98 pp 1031 -1044

[6] Massimo Tistarelli, Manuele Bicego, Enrico Grosso, Dynamic face recognition: From human to machine vision, Image and Vision Computing, Volume 27, Issue 3, *Special Issue on Multimodal Biometrics - Multimodal Biometrics Special Issue,* Pages 222-232, February 2009

[7] Kar, S. and Hiremath, S. and Joshi, D.G. and Chadda, V.K. and Bajpai, A. "A Multi-Algorithmic Face Recognition System", International Conference on Advanced Computing and Communications, 2006. ADCOM 2006. pp 321 -326, 2006

[8] Face Recognition Vendor Test: Overview and summary.
http://www.frvt.org/

[9] The Extended Yale Face Database B,
http://vision.ucsd.edu/ leekc/ExtYaleDatabase/ExtYaleB.html

[10] Description page for the Yale Face Database B,
http://vision.ucsd.edu/ leekc/ExtYaleDatabase/Yale%20Face%20Database.htm

[11] The AT& T Database of Faces formerly "The ORL Database of Faces"
http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[12] Nicolas Morizet, Frdric Amiel, Insaf Dris Hamed, Thomas Ea A Comparative Implementation of PCA Face Recognition Algorithm ,ICECS'07