Face Recognition Using Principal Component Analysis

Md. Kamrul Hasan, Amjad Khan, Zafar Toshpulatov and Désiré Sidibé

***Abstract* — The topic of face recognition has received considerable concentration in recent years and has become one of the sophisticated parts of image analysis and pattern recognition research. In image processing, the face recognition is a classical problem and also has various applications. There are numerous methods that had been implemented for the face recognition. However, in this project, the Principal component Analysis (PCA) based image recognition has been studied and implemented in MATLAB work-space according to given guidelines. Recently, the PCA has been extensively employed for face recognition algorithms that not only reduces the dimensionality of the image, but also retains some of the variations in the image data. The implementation has also been tested on the given data set of faces that was captured during the lectures. The overall accuracy for the detection of the unknown face is 55.84 percent for the given set of face image.The results reflect acceptable accuracy of the PCA based implementation.**

I. INTRODUCTION

The face recognition is a classical problem in image processing. It has numerous applications in various fields. For instance, from identity verification to surveillance and monitoring the face recognition has vital role in digital media. It is the general opinion that advances in computer vision research will provide useful insights to neuroscientists and psychologists into how human brain works, and vice versa [1]. The face recognition had been solved by various methods [2]. The Principal Component Analysis (PCA) is the one of the methods that provide a linear transformation solution the stated problem. The PCA had been proposed first by an English mathematician Karl Pearson [3]. It is a statistical technique that utilizes the orthogonal transformation to convert a set of observations of correlated variables into a set of uncorrelated variables that are called principal components. It is also a useful tool to reduce a complex data set into a lower dimensions by removing the redundancy in the data. This technique uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The resulting vectors that are the principle components are an uncorrelated orthogonal basis set as shown in Fig.1. The number of distinct principal components in the PCA is equal to the smaller of the number of original variables or the number of observations minus one. The transformed vector can be define as the first principal component that has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components.

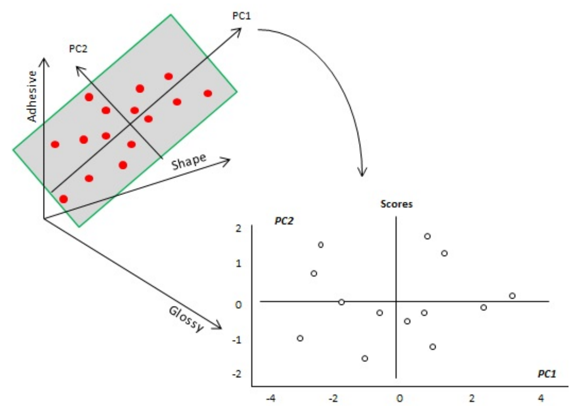

Fig. 1. Graphical representation of an example of PCA.

The PCA had been utilized by various authors to solve image recognition problem. In this assignment, the PCA based face recognition is studied and implemented. The accuracy of the face recognition also determined. The reset of the report is divided in different sections. Section II is presenting the implemented methodology. Section III discussed the obtained results. Finally, the report is concluded in Section IV.

II. METHODOLOGY

The implemented method is adopted as discussed in the project guidelines. The captured data set of different faces of various subjects has been resized to 240×320. Five different views of each subject are captured. Then five points are selected and their x and y locations are extracted from each image as features. These five points are located as: two points from the center of eyes, one point from the center of the nose and two points are extracted from the ending edges of the lips as shown in Fig. 2.

Fig. 2. Extracted positions of five facial points as features.

Once these features have been extracted from all the images than a transformation is applied by employing Singular Value Decomposition (SVD) to get a mapped matrix of 64×64 with predetermined locations of all the features in the data set. Upon applying transformation, each image in the data set is converted into single vector and then each vector is placed into a matrix such that each vector is acted as a row of the matrix as given in the Eq. 1.

$$D = \begin{bmatrix} I_1(1,1) & I_1(1,2)..... & I_1(M,N) \\ I_2(1,1) & I_2(1,2)..... & I_2(M,N) \\ . & & \\ . & & \\ . & & \\ . & & \\ I_n(1,1) & I_n(1,2)..... & I_n(M,N) \end{bmatrix} \tag{1}$$

Now the co-variance of the matrix D is determined by using the Eq. 2, where N is number of images in the data set and $\tilde{D}$ is given in Eq. 3. The overall process for the training of data set for recognition of face is shown in Fig. 3.

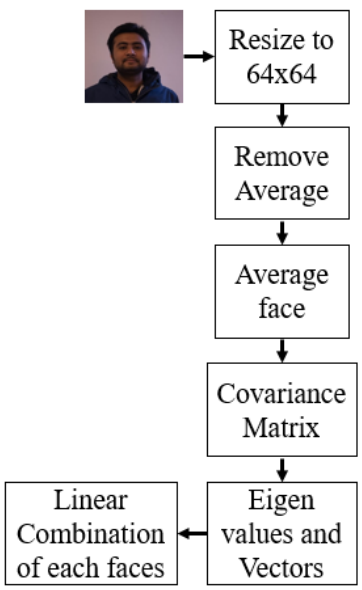

Fig. 3. The block diagram of training process.

Due to symetric nature of the co-variance matrix C, the diagonalization is applied to it. Then the eigenvectors are obtained that are called principal components of matrix C.

$$C = \frac{1}{N}\tilde{D}\tilde{D}^T \tag{2}$$

$$\tilde{D} = D - mean(D) \tag{3}$$

As each principal component is representing an image in the data set, therefore, each image can be obtained back by converting the principal component. The conversion is a reverse process of concatenation and converted images are called eigenfaces. The k principal components are put into d projection matrix Φ, here d is the size of matrix D and each column in Φ matrix represents a principal component. Therefore, any image I_i can be represented as a vector X_i and can be projected in PCA space by computing ϕ_i as given in Eq.4.

$$\phi_i = \Phi . X_i \tag{4}$$

At the final stage, the images are trained by projecting them to the PCA space. Then in the testing part, the test image is passed through the implemented algorithm in order to get the its match from the the trained database. The overall implemented method is described graphically in Fig.4.

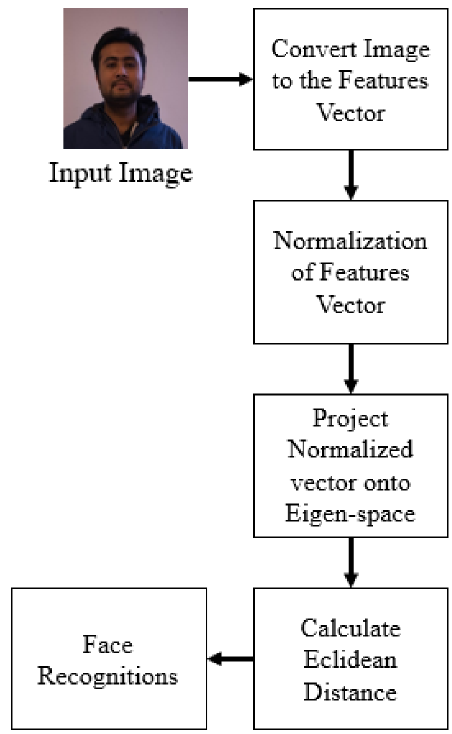

Fig. 4. Block diagram of implemented algorithm.

III. Results

As all the implementation has been performed in MATLAB environment, therefore, for the ease of access to the developed code the Graphical User Interface (GUI) has also been built. On back hand side various functions such as $Normalization - for - Test - Image()$, $Normalization - for - Training - Image()$ and $PCA - SVD - Based - Recognition()$ has been generated and these function has been integrated with GUI main function named, $gui - pca - final()$. In order to arrange the data set, various directories has been created such as $ActualColorImage240x320withFeature$ (contains the all the original RGB images), $ImageanddataTobeNormalizedfortesting$ (contains the

test images and it is called by Normalization-for-Test-Image()), $ImageanddataTobeNormalizedforTraining$(contains the training data set and called by Normalization-for-Training-Image()), $TestImageAfterNormalization$ (Normalized (64×64) test image set and called by gui-pca-final()) and $TrainingImageAfterNormalization$ (contains the normalized (64×64) training image set and called by gui-pca-final()).
Upon execution of gui-pca-final() the interface is shown in Fig. 5. From this GUI, the user can select the search image from the generated list of the GUI on the left hand side as given in Fig. 5. Then selected image is compared within the trained data set and exact match can be shown in GUI display area and at the same the calculated recognition accuracy of the algorithm is prompted on the bottom of the GUI as shown in Fig. 6.

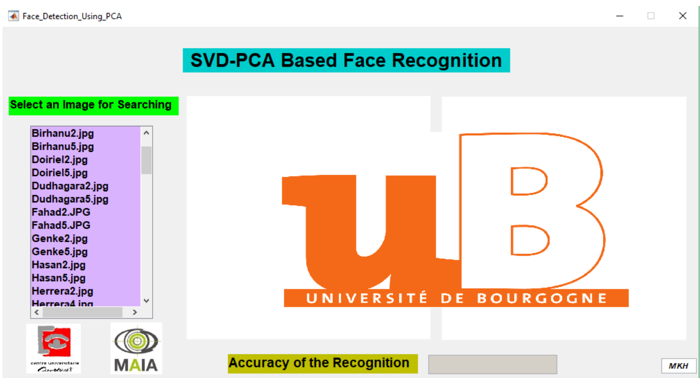

Fig. 5. A preview of developed GUI.

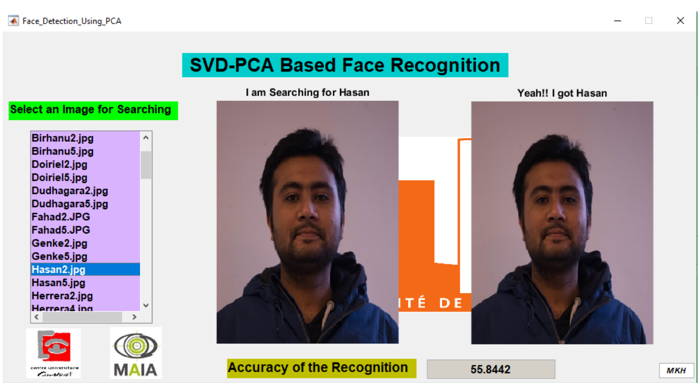

Fig. 6. A preview of GUI with good matched faces.

As from the reflected results, the implemented algorithm shows an accuracy of 55.84 percent. However, on some images within the data set the recognition results are not acceptable as shown in Fig. 7. The reason that has been analyzed during the testing phase that brings unacceptable results. The reason is simple that PCA recognition is based on only the Euclidean distance among feature set.

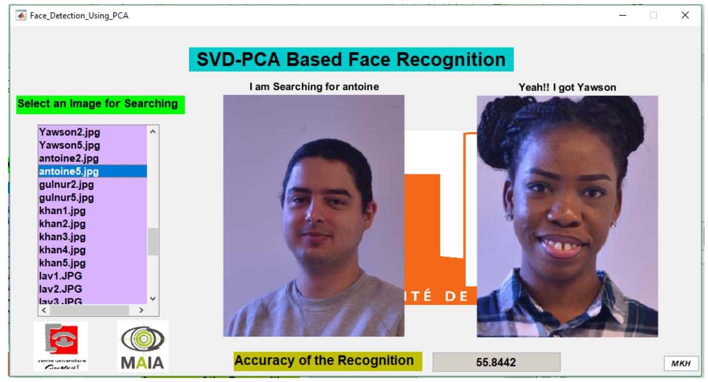

Fig. 7. A preview of GUI with bad matched faces.

IV. Conclusions and Future Improvements

In this project, a PCA and SVD based implementation for face recognition has been studied by practically implementing the algorithm. From the obtained results analysis, it is found that only five facial features are not sufficient for the PCA based recognition although less features provide less computational complexity. To achieve higher accuracy and perfection in recognition, the number of features can be increased without taking caring of the computational complexity of the system. The computational complexity can be improved nowadays by adding the high performance GPU based computer.

Acknowledgment

Authors would like to thank Dr. Désiré Sidibé for his guidance and support throughout the study period and other faculty members as well as the University of Burgundy for facilitating to learn new techniques and technologies. Authors are also very much grateful to the European Union for funding MAIA studies under the Erasmus Mundus scholarship program. Last but not the least, authors would like to thank all the volunteers who participated in this project while face data collection.

References

[1] K. D. Mislav Grgic, "FACE RECOGNITION HOMEPAGE." http://www.face-rec.org. [Online; accessed 10-Jan-2018].

[2] W. Wjcik, K. Gromaszek, and M. Junisbekov, "Face recognition: Issues, methods and alternative applications," in *Face Recognition - Semisupervised Classification, Subspace Projection and Evaluation Methods* (S. Ramakrishnan, ed.), ch. 02, Rijeka: InTech, 2016.

[3] K. P. F.R.S., "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901.