Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance

Nawaf Hazim

Related papers

Face Recognition: A Literature Review
Nawaf Hazim

Face Image Recognition Based on Linear Discernment Analysis and Cuckoo Search Optimization with …
Journal of Computer Science IJCSIS

Automatic Facial Expression Recognition Based on Hybrid Approach
Ali Bermani

Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance

Nawaf Hazim Barnouti

Al-Mansour University College
Baghdad, Iraq

Wael Esam Matti

Al-Mansour University College
Baghdad, Iraq

Sinan Sameer Mahmood Al-Dabbagh

Al-Mansour University College
Baghdad, Iraq

Mustafa Abdul Sahib Naser

Al-Mansour University College
Baghdad, Iraq

Abstract—In this paper, an automatic face recognition system is proposed based on appearance-based features that focus on the entire face image rather than local facial features. The first step in face recognition system is face detection. Viola-Jones face detection method that capable of processing images extremely while achieving high detection rates is used. This method has the most impact in the 2000's and known as the first object detection framework to provide relevant object detection that can run in real time. Feature extraction and dimension reduction method will be applied after face detection. Principal Component Analysis (PCA) method is widely used in pattern recognition. Linear Discriminant Analysis (LDA) method that used to overcome drawback the PCA has been successfully applied to face recognition. It is achieved by projecting the image onto the Eigenface space by PCA after that implementing pure LDA over it. Square Euclidean Distance (SED) is used. The distance between two images is a major concern in pattern recognition. The distance between the vectors of two images leads to image similarity. The proposed method is tested on three databases (MUCT, Face94, and Grimace). Different number of training and testing images are used to evaluate the system performance and it show that increasing the number of training images will increase the recognition rate.

Keywords—Face Detection; Face Recognition; PCA; LDA; Viola-Jones; Feature Extraction; Distance Measurement; MATLAB; MUCT; Face94; Grimace

I. INTRODUCTION

Face detection is among the important advanced topics in computer vision and pattern recognition communities and it is the first important step for facial analysis methods and among the most important issues in computer vision like face recognition, facial expression, head tracking, face verification. With the arrival of the internet and low price digital cameras, in addition to impressive image editing software such as adobe Photoshop, average users have more access to the tools of digital doctoring than in the past. The objective of face detection would be to determine if there are any faces in the image, then return the location and the bounding box of each face in the image regardless of illuminations, occlusions, facial pose, orientation and expression. Automatic human detection and tracking is an essential and challenging field of research and offers many application areas [1]. Tracking is

regarded as a challenging step of tracking system, which localizes and associates the feature across a series of frames. Face recognition has attached much more attention because of its great potential in numerous applications (security, criminal justice system, surveillance, human-computer interactions, image database investigation, smart card application, multimedia environments with adaptive human-computer interface, video indexing and civilian applications) [2] [17].

The expanding use of computer vision in replacing human beings, surveillance, has started the research in the field of face detection. Earlier research is biased to human recognition rather than tracking. Tracking the movement of human beings raised the requirements for tracking. Tracking movements are of high interest in identifying the activities of individual and knowing the attention of individual [1]. The performance of different faces based applications, from standard face recognition and verification to the latest face clustering, retrieval and tagging, depends on efficient and accurate face detection. Face detection is an important part of face recognition system simply because it has the ability to focus computational resources on the important part of an image containing face.

Face recognition involves recognizing individuals with their intrinsic facial characteristic. Compared to other biometrics, face recognition is more natural, non-intrusive and can be used without the cooperation of the individual. Face recognition system can be used in two modes: verification and identification. Face verification system (one-to-one matching) involves confirming or denying the identity claimed by a specific individual. Face identification system (one-to-many matching) attempts to find the identity of a given individual against all image templates in face individual database [3].

Face recognition methods can be divided into appearance-based or model-based methods. Appearance-based (Holistic) face recognition legally attempt to identify faces using global representations based on the entire image rather than local facial features. An image is considered as a high dimensional vector. Statistical methods are frequently used to gain a feature space from the image distribution. The sample image is compared to the training set. Appearance-based methods can be classified as either linear or non-linear. Linear

appearance-based methods perform a linear dimension reduction [4]. The face vectors are projected to the basis vectors; the projection coefficient are used as the feature representation for each face image through the projection of the face image vector onto the basis vectors. Linear methods are Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA). Non-linear appearance-based methods are usually more complicated than linear methods. Direct non-linear manifold schemes are explored to learn this non-linear manifold. Linear subspace analysis is an approximation of a non-linear manifold. Kernel PCA (KPCA) is widely used. Model-based face recognition scheme aims to construct a model of the human face that can capture facial variations. It can be either 2-Dimensional or 3-Dimensional. These models are frequently morphable. Morphable model make it possible for classifying faces even when pose changes are present. Model-Based methods are Elastic Bunch Graph Matching (EBGM) or 3D Morphable Models [5]. Hybrid method is a combination of appearance-based and model-based methods. Regardless of the method, the most important concern in face recognition is dimensionality. Suitable methods are needed to reduce the dimension of the studded space. Working on higher dimensions causes overfitting, where the system starts to memorize. Computational complexity is also an important problem when working on large database.

II. PROPOSED METHODOLOGY

Fig. 1. Face Recognition Proposed Methodology Process

Face recognition is a complex image processing problem in real world applications. In this work, details are provided for the method and training process of the proposed face detection and recognition system. Technologies characteristics and features make face recognition important and better performer depending on the application. Face recognition basically divided into three steps begins with face detection continue with feature extraction and end up with distance measurement process. Three benchmark databases (MUCT, Face94 and Grimace) are used to test the system performance. MUCT database contains 3.755 faces with 76 manual landmarks, Face94 contains 153 images each with a resolution of 180*200 pixel, and Grimace database contains face images with 20 individuals each having 20 images. Viola-Jones face detection method is used to detect and crop face region in each face database. The linear appearance-based method PCA-LDA is used for feature extraction and dimension reduction. Finally, Square Euclidean Distance measurement is used. The distance between two images is a major concern in pattern recognition. Image similarity is the distance between the vectors of two images. Figure 1 shows the proposed methodology process.

III. FACE DATABASE

Numerous face databases are available for face recognition researchers. These databases differ in size, scope and purpose. It is recommendable to use a standard test face recognition database for researchers to be able to directly compare the final results. The photographs in many of these databases are acquired by small research teams specifically to study face recognition. MUCT, Face94, and Grimace databases are used in this work. Table I shows the important features of different face recognition databases.

TABLE I. FACE DATABASES FEATURES

Database	Format	Individuals	Image Size
MUCT	RGB	276	480 * 640
Face94	RGB	153	180 * 200
Grimace	RGB	20	180 * 200
ORL	Gray	40	92 * 112
FERET	Gray/RGB	1199	256 * 384
UMIST	Gray	20	92 * 112
Indian	RGB	40	640 * 480

A. MUCT Database

The Milborrow/University of Cape Town (MUCT) database contains 3.755 faces with 76 manual landmarks. The database is created to diversity lighting, age, and ethnicity. In this database, all images captured in December 2008 are from the individuals around the University of Cape Town campus. The individuals in this database are university students, parents, high school teachers, and employees, each individual is photographed using five webcams, which makes the database useful for applications that require multiple occurring views of the individual [6]. Figure 2 shows a sample images of MUCT database.

Fig. 2.   MUCT Database Sample Images

B. Face94 Database

The Face94 contains 153 images each with a resolution of 180*200 pixel and the directories comprise images of male and female individuals in separate directories (20 females, 113 males, and 20 male staffs). The images are mainly from first year underground students. The majority of the individuals are between 18 and 20 years old. The lighting is artificial and some of the images are captured with glasses, and a mixture of tungsten and fluorescent overhead [7]. Figure 3 shows a sample images of Face94 database.

Fig. 3.   Face94 Database Sample Images

C. Grimace Database

The Grimace database contains face images with 20 individuals each having 20 images with a resolution of 180*200 pixel with a small head scale variation. This database contains images of both female and male individuals. The lighting of the images minimally varies. The images of the individuals from various facial origins have major expression variations with breads and glasses [8]. Figure 4 shows a sample images of Grimace database.

Fig. 4.   Grimace Database Sample Images

IV.   FACE DETECTION

Face detection is generally considered as a certain case of object-class detection and it's a popular topic in biometrics research. Face detection is the first step of face recognition system. Objects can be detected using one of the face detection methods. Then feature extraction and distance measurement methods can be applied to the system. In object-class detection, the task is to find the location of all objects in an image that belong to a given object [3]. Face detection is not simple because it carries lots of variations of appearance in images, such as facial expression, pose variation, image orientation, occlusion and illuminating condition. In this work, Viola-Jones face detection method is used.

The Viola-jones object detection method suggested by Paul Viola and Michael Jones in 2001. This method has the most impact in the 2000's and known as the first object detection framework to provide relevant object detection that can run in real time. Viola-Jones requires full view frontal upright faces [9]. At a high level, the method read an input image with a window looking for human face features. When enough features are found, then this window type of the image is reported to be a face [10]. In order to bring different size faces, the window must be scaled and the process is repeated. For each window scale involves through the method separately of the other scales. This method happens to be rather time consuming resulting from the calculation of the different images size. To decrease the number of features each window have to check and each window is passed through levels. Early levels include less features to check and are much easier to pass but later levels end up having more features and are more demanding. At each level, the evaluation of features for that levels are collected and whether if the collected value does not pass the threshold, the level is failed and this window will be not recognized as a face. The Viola-Jones face detection method is divided into three main parts (Integral image, classifier learning with AdaBoost and attentional cascade structure) that make it possible to build a successful face detection that can be used on real time application.

A. Creating an Integral Image

An image representation called the integral image. Integral image also known as a summed area table. Integral image is computed as a pre-processing step. The first step of Viola-Jones method is to convert the input face image into an integral image. This can be done by making each pixel equal to the entire summation of all pixels above and to the left of the concerned pixel [9]. The integral image can be calculated as shown in the equation below:

$$I(x,y) = \sum_{x' \leq x, Y' \leq y} O(x',y') \qquad (1)$$

Where I is the integral image and O is the original image.

To complete the summation of any rectangular area by using the integral image is extremely efficient. The summation of pixels in rectangle area $Z = [A1, A2, A3, A4]$ can be calculated as shown in the equation below:

$$I(x, y) = \sum_{(x,y) \in Z} O(A4) + O(A1) + O(A2) + O(A3) \qquad (2)$$

Features will be calculated in constant time considering that the summation of the pixels can be computed in the constituent rectangles in constant time. Viola-Jones have noticed that a detector with a basic resolution of 24*24 pixels offers positive results [10].

B. AdaBoost Training

AdaBoost is a machine learning boosting method capable of finding a highly accurate hypothesis by combining many week hypothesis each with average accuracy. The AdaBoost method is generally viewed as the first step straight into more practical boosting methods [9].

C. Cascade Structure

Cascade of gradually more complex classifiers achieves even better detection rates. The concept of the Viola-Jones face detection method is to scan the detector frequently by the same image each time with a new size. Regardless of whether an image should contain one or more faces, there is no doubt that an excessive large amount of the evaluated sub windows might still be non-faces [10]. The Cascade classifier consist of levels each containing a strong classifier. The responsibility of each level is to evaluate if a given sub-window is actually non-face or maybe a face. The implementation contains 22 levels with early levels containing much less features and later levels containing more in depth detailed features. Typically, early levels are passed more frequently with later levels being more demanding.

V. Feature Extraction

Feature extraction involves reducing the amount of resources required to describe a large amount of data. Feature extraction from given data is a critical problem for the successful application of machine learning. In this work PCA and LDA are used as feature extraction and dimension reduction method from the original face images. PCA and LDA produce feature vectors in a reduced dimension.

A. Principal Component Analysis (PCA)

Principal component analysis (PCA) is one of the most important methods used in pattern recognition and compression. PCA is feature extraction and dimension reduction method [11]. PCA is a common statistical method using a holistic approach to find patterns in high dimensional data. The purpose of PCA is derived from the information theory approach, which break down facial images into small sets of characteristic feature images called Eigenfaces which used to represent both existing and new faces [12]. In PCA method, the 2-Dimensional face image matrices must be transformed into a 1-Dimensional vector. The 1-Dimensional vector can be either row or column vector. As a result, the image representation leads to a high dimensional space [13].

PCA Method Steps are as Follows:

1) Training set of total M images are used to compute the Average Mean as shown in the equation below:

$$average = \frac{1}{M} \sum_{n=1}^{M} TrainingImages\ (n) \qquad (3)$$

2) Original image will be subtracted from the Average Mean as shown in the equation below:

$$Sub = TrainingImages - average \qquad (4)$$

3) Calculate the Covariance Matrix as shown in the equation bellow:

$$Covariance = \sum_{n=1}^{M} Sub(n)\ Sub^{T}(n) \qquad (5)$$

4) Calculate the Eigenvalues and Eigenvectors of the Covariance Matrix.

5) Sort and choose the best Eigenvalues. The highest Eigenvalues that belong to a group of Eigenvectors is chosen, these M Eigenvectors describe the Eigenfaces. Given that new faces are encountered, the Eigenfaces can be updated or recalculated accordingly.

6) Project the training samples onto Eigenfaces.

B. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is also known as fischerface method used to overcome drawback the PCA of its application kept in small image database. It is achieved by projecting the image onto the Eigenface space by PCA after that implementing pure LDA over it to classify the Eigenface space projected data [12]. LDA searches for those vectors in the underlying space that best discriminate among classes. LDA group images of the same class and separates images of different classes. As Mathematically two measures are defined (within-class scatter matrix and between class scatter matrix) [14]. For all samples of all classes the between-class scatter matrix SB and the within-class scatter matrix SW are defined as shown in the equations below:

$$SB = \sum_{n=1}^{N} \sum_{m=1}^{M_n} (S_m^n - u_n)\ (S_m^n - u_n)^{T} \qquad (6)$$

$$SW = \sum_{n=1}^{N} (u_n - u)\ (u_n - u)^{T} \qquad (7)$$

Where S_m^n is the mth sample of class n, u_n is the mean of class n, N is the number of classes, M_n is the number of samples in class n and u is the mean of all classes. Then subspace for LDA is spanned by a set of vectors $W = [W1, W2, W3, \dots, W_M]$.

$$W = \arg\max = mod\ (W^{T} * SB\ W / W^{T} * SW\ W) \qquad (8)$$

The goal is to maximize the between class measure while minimizing the within class measure. Figure 5 shows that when maximize the ratio of between class variance to within class variance will find a good class separation. To do this we maximize ratio $det|SW|/det|SB|$ to prove that if SW is non-singular matrix. The with class scatter matrix represent how face images are distributed closely with-in classes and between class scatter matrix describe how classes are separated from each other. When face images are projected into the discriminant vector W. Face Images should be

distributed closely with-in classes and should be separated between classes as much as possible. In other words, these discriminant vectors minimize the denominator and maximize the numerator [15].
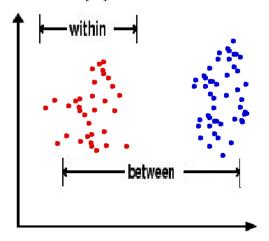

Fig. 5. Class Separation in Linear Discriminant Analysis

VI. DISTANCE MEASUREMENT

Once the features are extracted and selected using PCA-LDA, the next step is to measure the distance between images. Most face recognition methods from the last decade help in deciding according to the distance measurement. The distance between two images is a major concern in image recognition and computer vision. The final step of face recognition is measuring the distance between two images. Image similarity is the distance between the vectors of two images. The distance among feature space representations are used as the basis for recognition decisions [16]. One way or another, distance measurement has a big impact in face recognition area. Distance measurement methods are used in many areas like finance, data mining, voice recognition and signal decoding.

Euclidean distance is used for distance measurement between images. Euclidean Distance is defined as the straight line distance between two points, which examines the root of square differences between the coordinates of a pair of objects [16]. Euclidean Distance can be calculated using the equation below:

$$EuclideanDistance(X, Y) = \sqrt{\sum_{n=1}^{No.of\ Images} (X_n - Y_n)^2} \quad (9)$$

Without the square roots, we can obtain the Square Euclidean Distance (SED) measurement. The standard Euclidean Distance can be squared in order to place progressively greater weight on objects that are farther apart. In this case, the equation becomes as shown below:

$$SquareEuclidean(X, Y) = \sum_{n=1}^{No.of\ Images} (X_n - Y_n)^2 \quad (10)$$

VII. RESULT AND DISCUSSION

In this analysis, three databases (MUCT, Face94, and Grimace) are used to evaluate the system performance. In MUCT database, 8 individuals with 1 to 3 training and testing images for each individual is used. While, in Face94 and Grimace databases, 8 individuals with 1 to 4 training and testing images for each individual are used. The simulation of the proposed methodology was performed using MATLAB software package The analysis shows that increasing the number of training images will increase the recognition rate. Viola-Jones method is used for face detection on each database. This method achieved high detection rate and all images are detected and cropped in the three databases. Figure 6 shows a sample image detection and cropping using Viola-Jones method.
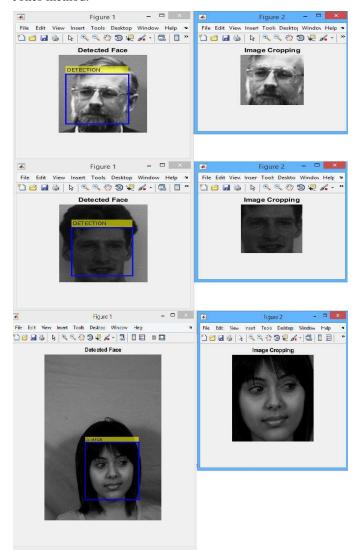

Fig. 6. Image Detection and Cropping Using Viola-Jones Method

Figure 7 shows MUCT images database after detection and cropping. Figure 8 shows Face94 images database after

detection and cropping. Figure 9 shows Grimace images database after detection and cropping. PCA-LDA are applied on the detected cropped images for feature extraction and dimension reduction. Different number of training and testing images are used in each database. Square Euclidean distance is used to measure the distance between two images. Face94 and Grimace databases with 1 to 4 images shows high recognition rates, while MUCT database with 1 to 3 images shows low recognition rates. To avoid this problem, the number of images in the database must be increased to become 1 to 8 images for each individual. Table II shows the recognition rate of MUCT with 1 to 8 images. Table III shows the recognition rates of Face94 and Grimace databases.

TABLE II. THE MUCT DATABASE RECOGNITION RATES

Test Image	Train Image	Recognition Rates MUCT
7	1	60.71 %
6	2	66.67 %
5	3	82.5 %
4	4	84.38 %
3	5	85 %
2	6	87.5 %
1	7	87.5 %

TABLE III. FACE94 AND GRIMACE DATABASE RECOGNITION RATES

Test Image	Train Image	Recognition Rates Face94	Recognition Rates Grimace
3	1	100 %	91.67 %
2	2	100 %	100 %
1	3	100 %	100 %

Fig. 7. Detection and Cropping on MUCT Database

Fig. 8. Detection and Cropping on Face94 Database
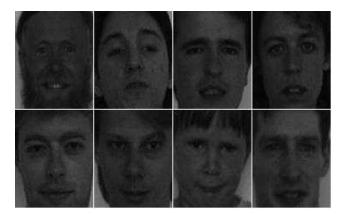
Fig. 9. Detection and Cropping on Grimace Database

VIII. CONCLUSION

The purpose of this work was to implement an automatic face recognition system based on appearance-based methods. Face detection using Viola-Jones method is used to detect and crop faces in each database. Viola-Jones method show high detection rates. MUCT, Face94, and Grimace databases are used, each with 8 individuals and 1 to 3 images are choosing for each individual in MUCT database, 1 to 4 images are choosing for each individual in Face94 and Grimace databases. PCA-LDA is used for feature extraction and dimension reduction. PCA-LDA implementation was successful. Square Euclidean Distance is used to measure the distance between two images, which leads to find image similarity. Face94 and Grimace databases using different number of testing and training images shows high recognition rates, while MUCT database shows low recognition rates. In MUCT database, increasing the number of images to become 1 to 8 images for each individual shows increasing the recognition rates. The recognition time was acceptable and takes few seconds. The results show increasing in recognition rates when increase the number of training images.

REFERENCES

[1] Hatem, H., Beiji, Z., Majeed, R., Lutf, M. and Waleed, J., 2015. Face Detection and Pose Estimation Based on Evaluating Facial Feature Selection. *International Journal of Hybrid Information Technology*, 8(2), pp.109-120.

[2] Bakshi, U. and Singhal, R., 2014. A survey on face detection methods and feature extraction techniques of face recognition. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(3), pp.233-237.

[3] Rath, S.K. and Rautaray, S.S., 2014. A Survey on Face Detection and Recognition Techniques in Different Application Domain. *International Journal of Modern Education and Computer Science*, 6(8), pp.34-44.

[4] Fernandes, S. and Bala, J., 2013. Performance Analysis of PCA-based and LDA-based Algorithms for Face Recognition. *International Journal of Signal Processing Systems*, 1(1), pp.1-6.

[5] Gayathri, S., Mary Jeya priya, R., and Dr.Valarmathy, S., 2014. Face Recognition by Using Distance Classifier Based On PCA and LDA,

International Journal of Innovative Research in Science, Engineering and Technology, 3(3), pp.1121-1126.

[6]   Milborrow, S., Morkel, J. and Nicolls, F., 2010. The MUCT landmarked face database. *Pattern Recognition Association of South Africa*, *201*(0).

[7]   Abdullah, M., Wazzan, M. and Bo-saeed, S., 2012. Optimizing face recognition using PCA. *International Journal of Artificial Intelligence & Applications,* 3(2), pp.23-31.

[8]   Mandal, T. and Wu, Q.J., 2008, December. Face recognition using curvelet based PCA. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.

[9]   Cha, Z. and Zhengyou, Z., 2010. A survey of recent advances in face detection. *Microsoft Research, Microsoft Corporation*.

[10]   Hefenbrock, D., Oberg, J., Thanh, N.T.N., Kastner, R. and Baden, S.B., 2010, May. Accelerating viola-jones face detection to fpga-level using gpus. In *2010 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines* (pp. 11-18). IEEE.

[11]   Kaushik, Sangeeta, Dubey, R. B., and Abhimanyu Madan. 2014. Study of Face Recognition Techniques, *International Journal of Advanced Computer Research*, 4(17), pp.939-949.

[12]   Singh, A., Singh, S.K. and Tiwari, S., 2012. Comparison of face Recognition Algorithms on Dummy Faces. *The International Journal of Multimedia & Its Applications*, *4*(4), pp.121-135.

[13]   Barnouti, Nawaf Hazim. 2016. Face Recognition Using Eigen-Face Implemented On DSP Processor, *International Journal of Engineering Research and General Science*, 4(2), pp.107-113.

[14]   Bhattacharyya, S.K. and Rahul, K., 2013. Face Recognition by Linear Discriminant Analysis. *International Journal of Communication Network Security, ISSN*, 2(2), pp.2231-1882.

[15]   Kamerikar, Umesh Ashok, and Dr. M.S. Chavan. 2014. Experimental Assessment of LDA and KLDA for Face Recognition, *International Journal of Advance Research in Computer Science and Management Studies*, 2(2), pp.137-146.

[16]   Gawande, M.P. and Agrawal, D.G., Face recognition using PCA and different distance classifiers. *IOSR Journal of Electronics and Communication Engineering*, 9(1), pp.1-5.

[17]   Barnouti, Nawaf Hazim. 2016. Improve Face Recognition Rate Using Different Image Pre-Processing Techniques, *American Journal of Engineering Research (AJER)*, 5(4), pp.46-53.