ISSN (e): 2250 – 3005 || Volume, 08 || Issue, 11|| November – 2018 ||
International Journal of Computational Engineering Research (IJCER)

# Advance Face Recognition through PCA and Mahalanobis Distance

Vaishali[1], Banita[2]
[1] M. Tech.., Research Scholar, Department of Computer Science and Engineering, PDM University, Bahadurgarh
[2]Assistant Professor, Department of Computer Science and Engineering, PDM University, Bahadurgarh
Corresponding Author: Vaishali

**ABSTRACT**
Face recognition in pictures or videos is trending research area in the field of biometric system and application. In so many it is very easy to see the video surveillance cameras and camcorders for security and safety. It is very widespread that it has been recognized that face recognition plays an important role in the monitoring system as it does not require the subject of cooperation. The real benefits of face-to-face recognition of other biometric technologies are unique and responsive. Just like humans, the face is a very dynamic aspect of the object, making face detection a problem of computer vision. In this sense, identifying precision and speed is a major issue. The purpose of this paper is to evaluate the detection of different faces. The identification method provides a complete solution for recognizing front images and better recognition, a better response rate is the first step in video surveillance. The solution is based on tests conducted on various rich databases. As for the theme, she presents emotion, race and light. We also learned that with two or more mechanisms, the precision of the system can be much better. In this paper we developed a PCA-based identification system that extracts and uses functions to respond to different site information.
**Keyword:** Face Recognition, Biometric Identification, PCA, Mahalanobis Distance, AT & T Database, MATLAB

---

Date of Submission: 15-12-2018 Date of acceptance: 31-12-2018

---

## I. INTRODUCTION

The face recognition system consists of two parts: hardware and software. This system is used for automatic user identification or password confirmation. The input uses digital images or video frames from the same video. National agencies and some private organizations use such face recognition systems, particularly for identifying cameras as input parameters or for biometric identification to verify identity using 3D cameras and scanners. The system must recognize the face of the image, take it out of the image and check it. There are many ways to check, but the most popular one is the recognition of facial features. The face has about 80 characteristic parameters, some of which are the width of the nose, the space between the eyes, the tip of the eye, the shape of the zygote bone and the width of the jaw.

A properly connected neural network examines the small windows of the image and determines if each window contains a face. The system arbitrates between multiple networks to improve the performance of a single network. Forming a neural network for face detection is a challenge because it is difficult to characterize the prototype "non-face". Unlike face recognition, the categories to be transmitted are different faces, and the two categories. That need to be distinguished in face detection are "face up image" and "non-face image". It is easy to obtain a representative sample of an image containing a face, but it is much more difficult to obtain representative samples without a face. The size of the second type of training set can grow very quickly.

## II. LITERATURE REVIEW

[1] Dr. Asmahan M. Altaher [2014] proposed that Face recognition is the ability of categorize a set of images based on certain discriminatory features. Classification of the recognition patterns can be difficult problem and it is still very active field of research. The paper introduces conceptual framework for descriptive study on techniques of face recognition systems. It aims to describe the previous researches have been study the face recognition system, in order scope on the algorithms, usages, benefits , challenges and problems in this felids,

the paper proposed the face recognition as sensitive learning task experiments on a large face databases demonstrate of the new feature. The researcher recommends that there's a needs to evaluate the previous studies and researches, especially on face recognition field and 3D, hopeful for advanced techniques and methods in the near future.

[2] Priyanka [2015] proposed that Face Recognition is having the importance to provide biometric authentication with easy image acquisition that can be used for online and offline applications. There are number of existing approaches for biometric facial recognition and classification. In this paper, some of the common and reliable approaches for facial recognition are explored. These approaches include PCA, LDA, KDA, Neural Network etc. The paper has also discussed the basic model of facial recognition and explained each stage of this model. It also includes different methods of feature extraction to describe the facial components.

[3] PratibhaSukhija et.al [2016] proposed that the face recognition is one of the most challenging aspect in the field of image analysis. Face recognition has been a topic of active research since the 1980's, proposing solutions to several practical problems. Face recognition is probably the biometric method that is used to identify people mainly from their faces. However, the recognition process used by the human brain for identifying faces is very challenging. In this paper, a Genetic Algorithm (GA) based approach is proposed for face recognition. The proposed algorithm recognizes an unknown image by comparing it with the known training images stored in the database and gives information regarding the person recognized. The proposed algorithm is then compared with other known face recognition algorithms viz: Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA) algorithms. It has been observed that the recognition rate of the proposed algorithm is better.

[4] Mahmoud Afifi et.al [2017] proposed that The Viola-Jones face detection algorithm was (and still is) a quite popular face detector. In spite of the numerous face detection techniques that have been recently presented, there are many research works that are still based on the Viola-Jones algorithm because of its simplicity. In this paper, we study the influence of a set of blind pre-processing methods on the face detection rate using the Viola-Jones algorithm. They focus on two aspects of improvement, especially badly illuminated faces and blurred faces. Many methods for lighting invariant and de-blurring are used in order to improve the detection accuracy. They want to avoid using blind pre-processing methods that may obstruct the face detector. To that end, they perform two sets of experiments. The first set is performed to avoid any blind pre-processing method that may hurt the face detector. The second set is performed to study the effect of the selected pre-processing methods on images that suffer from hard conditions. They present two manners of applying the pre-processing method to the image prior to being used by the Viola-Jones face detector. Four different datasets are used to draw a coherent conclusion about the potential improvement caused by using prior enhanced images. The results demonstrate that some of the pre-processing methods may hurt the accuracy of Viola-Jones face detection algorithm. However, other pre-processing methods have an evident positive impact on the accuracy of the face detector. Overall, they recommend three simple and fast blind photometric normalization methods as a pre-processing step in order to improve the accuracy of the pre-trained Viola-Jones face detector.

[5] Hajime Nada et.al [2018] proposed that Face detection has witnessed immense progress in the last few years, with new milestones being surpassed every year. While many challenges such as large variations in scale, pose, and appearance are successfully addressed, there still exist several issues which are not specifically captured by existing methods or datasets. In this work, they identify the next set of challenges that requires attention from the research community and collect a new dataset of face images that involve these issues such as weather-based degradations, motion blur, focus blur and several others. They demonstrate that there is a considerable gap in the performance of state-of-the-art detectors and real-world requirements. Hence, in an attempt to fuel further research in unconstrained face detection, they present a new annotated Unconstrained Face Detection Dataset (UFDD) with several challenges and benchmark recent methods. Additionally, they provide an in-depth analysis of the results and failure cases of these methods.

**Table 2.1:** Schemes based on different technique of face recognition

| Sr.No | Author Name | Objective | Outcome |
|---|---|---|---|
| 1 | Chen, Zhang, Zhang | To find 3-D model of face | Mixture based feature model + 3D shape model |
| 2 | Cristinacce, Cootes, Scott | To find Gaussian Noise | Boosted Haar wavelet-like features and classifiers |
| 3 | Smeraldi, Bigun | To provide the Filter for SVM | Gabor responses of the complete retinal field + SVM, 30 dimensional Gabor response of each point + SVM |
| 4 | Feris, Gemmell, Toyama, Krüger | To find the template matching | Template matching using Hierarchical Gabor Wavelet Network (GWN) representation of features, Template matching using Hierarchical Gabor Wavelet Network (GWN) representation of faces |
| 5 | La, Yuen, Chen, Lao, Kawade | Color segmentation (skin, lip) + edge map | Vertical projection of thresholded image obtained from the coarse level |
| 6 | Shakunaga, Ogawa, Oki | Find the principal component Analysis(PCA) | PCA on canonical positions of features structural matching has been extracted. |
| 7 | Ryu, Oh | To find PCA analysis | PCA on coordinates of the feature edge map + MLP for template matching,<br>Vertical and horizontal projections of face edge map |
| 8 | Shih, Chuang | To find edge of face | Edge projections + geometric model of facial features |
| 9 | Arca, Campadelli, Lanzarotti | To give geometrical Heuristics | Geometrical heuristics, Color segmentation (Skin and lip) + SVM |
| 10 | Zobel, Gebhard, Paulus, Denzler | To find the Probabilistic model of face | Geometrical heuristics on DCT coded images + Probabilistic model, based on coupled structured representation of feature locations |
| 11 | Gourier, Hall, Crowley | To find Gaussian Cluster | Gaussian derivatives (Gx, Gy, Gxx, Gxy, Gyy) + clustering to 10 centroids |
| 12 | Antonini, Popovici, Thiran | To find Corner detection | Feature extraction using PCA and ICA projections of windows surrounding the corner points + SVM for template matching |

## III. PROPOSED SYSTEM

We propose a real time face recognition system based on PCA and Mahalanobis. The main task of face recognition is to extract effective features. The proposed system uses the Eigen face method to reduce image information. Even small faces have incredible information. This method should be able to break the image, generally to represent the face image effectively rather than the image. Once the basic surface is created, the system can express the image to be analyzed as a linear combination of these basic planes. You can display each plane you want to classify in a parallel space and analyze it as a vector. Approach k - neighbors, neural networks, or even Euclidean simple distance measurements can be used for classification. The proposed system uses basic analysis of components to extract classifications of distances different from features such as Euclidian Distance, Manhattan Distance, and Mahalanobis Distance. The technique used here involves generating the "positive face", displaying the training data in the face region used with the predetermined classification method, dropping it in the face region and comparing the training data.

### 3.1 Steps Involved

The different steps for calculating the eigenface are as follows. A. Data Preparation a two-dimensional face image can be represented as a one-dimensional vector device by connecting each row (or column) to an elongated vector. We denote the vector N size (=

Column x rows of the image) representing the image set in which the sample was taken, assuming Γ1, Γ2, Γ3 ... ΓM, subtract the average. To be calculated, it is subtracted from the original face (Γ i) and the results stored in the variable.

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

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

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

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

D. Calculation of eigenvalues and eigenvalues of the dispersion matrix in this step we need to compute the eigenvalues of the self-vectors (eigenvectors) Xi and λi. E. Eigen mode calculation

$$[\Phi]X_i = f_i$$

Xi is an eigenvector and fi is a unique surface. Classification of faces new images are converted to their Eigen-face components. The obtained weights are weight vectors TT k:

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

Where k = 1,2,3,4 and kT k = [1122 .... MM] theEuclidean distance between two weight vectors (d, i, jj) provides a similarity scale between the corresponding images i & j.

**Fig 1:**Image Segregation with Eigen vector

**Various Distance metrics**

Let X and Y are vectors specific to n lengths. Next, we can compute the next space between these vectors.

**3.2 Mahalanobis Distance**

The MahénNobis space is defined as the area where sample changes are along each dimension. Consequently, vector transformation from image space to feature space is performed by dividing each factor of the vector by the corresponding standard deviation. This transformation provides a dimensionless space with changes in units of each dimension. If there are x and y curves in the unconverted PCA space, the corresponding vectors m and n are in the Mahalanobisspace. First, λi = σ2 is defined. Here, are 'i' is the e-par value, σ2i is the amount of change along these dimensions, and σ i is the standard deviation. Next, the relation between the vectors is defined

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

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

Where,λi is the $i_{th}$ Eigen value corresponding to the $i_{th}$ Eigenvector.

**Manhattan Distance**

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

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

**Euclidean Distance**

ALSO L2-norm or Euclidean Distance

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

**Decision on the test**

After calculating the distance between the two unique vectors, the nearest training image is returned to the test image assigned 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, it is said that a correct match has occurred, otherwise it is considered an incorrect match. The above policy will be tested across the following databases and results will be displayed.

**Codes**
The face recognition system here can extract facial features and compare them with the current database. The face discussed here for comparison is still a face. The project code is in the code.tar.gz folder. The README.txt file contains procedures for grouping and executing the system. This code is written in MATLAB.

**Database**
The face recognition code created for the two standard databases was tested. A total of about 1200 face images were used for 78 people tested using various variations and configuration variables in the project. The database is described in the next section.

The Yale Face Database B

**Fig 2:** Sample Image from Yale Face Database B

The Yale B database used contains 386 initial PGM formats and 856 grayscale faces. [9] The resolution of the image is 168 (width) × 192 (height) pixels. There are 31 themes in the eyeglassed light center, happy, light left, no glasses, nature, light, sad, sleepy, surprised, and wink: 31 themes for pictures, each expression of face and different shape. Sample images, by special arrangement, the file name of the image of the label in this database refers to the situation and lighting details. It is shown in Figure 2. The first part of the filename begins with the primary name "Yale-B" followed by a two-digit number followed by a two-digit number to indicate the number object and the situation. The rest of the file name deals with the azimuth and elevation directions for the direction of the individual light sources. Belong to topic 3 in the direction of the light source for the camera axis, for example the image of Yale DB. PGM visible situation 6, and the azimuth angle 35 ° (A + 0 35) and the high 40 ° (E + 40) Here, the positive azimuth indicates that the light source, which means that it was left, was on the right side of the topic, negative. The positive height is above the horizontal line and the negativeness is below the horizontal line. A complete description of the Internet's state and lighting is provided by the database. [10] The obtained image was taken with the camera Sony XC-75 with 8 bit (grayscale) (with linear response function) and stored in PGM raw format.

**3.3The AT & T Database of Faces (ORL)**
"Face AT & T database" was formerly "face ORL database" [11]. The image consists of 40 different themes and 10 images per subject. In some subjects, the images are taken at different times, the lighting is different depending on facial expressions (open / closed eyes, not smile/smile), face details (eyeglasses / without glasses). All images were captured on a dark homogeneous background with vertical position (with some tolerance for some lateral motion) subjects. A preview image of the face database can be obtained on the official website. Figure 3 shows a typical preview of the images of four different topics. The file is in PGM format and can be easily displayed on the LINUX system using the default image viewer. Each image is 92 (width) × 112 (height) pixels, 256 gray levels per pixel. The image is organized into 40 directories (one for each subject) and contains Sx format names. Here X is the subject (1 to 40). In each of these directories, there are ten kinds of images of the subject including the name of Y. Here, Y is the image number (1 to 10) of the subject.

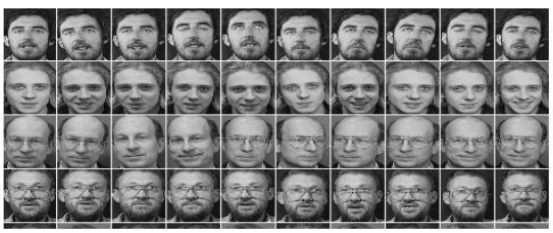

**Fig 3:**AT & T Database of Faces

## IV. RESULTS

The results of running the codes on the two databases are described below in table below.

**The Yale Face Database B**

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

| Model Used | Correct | Incorrect | Recognition Accuracy |
|---|---|---|---|
| **PCA+ Euclidian Distance** | 574 | 282 | 67.1% |
| **PCA+ Manhattan Distance** | 677 | 179 | 79.1% |
| **PCA+ Mahalanobis Distance** | 785 | 71 | 91.7% |

**AT& T Face Database**

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

| Model Used | Correct | Incorrect | Recognition Accuracy |
|---|---|---|---|
| **PCA+ Euclidian Distance** | 286 | 34 | 89.4% |
| **PCA+ Manhattan Distance** | 284 | 36 | 88.8% |
| **PCA+ Mahalanobis Distance** | 301 | 19 | 94.1% |

### 4.1 A Real-time Application: Process time

We measured the system processing time provided. In the windows operating system (10.04) with 2.4 GHz Core 2 duo processor and 4 GB memory, the 1200 image database contains 168 x 192 pixels consisting of 78 subjects under different lightning environments, expressions In other Yale B databases, MATLAB of the proposed system using the Mahalanobis distance is 2 seconds for facial training and 3 seconds for all database images after training, within 0.2 seconds Respond to individual face recognition queries. The system completed the query stream of 900 test images in 3 seconds, taking into account the 3ms query time slot.

## V. CONCLUSION

This face recognition project gave me the opportunity to find out many of the common methods used for face recognition. Detailed literature surveys gave positive and eight negatives from many relevant identification and differentiation systems. In addition, we also learned that by combining two or more Teckins, the accuracy of the system can be greatly improved. This investigationdevelop a PCA-based face recognition system that extracts features and uses them to match different spatial matrices. The spaces used are Eucledian Distance, Manhattan Distance and Mahalanobis Distance. The results clearly show that recognition systems based on Mahalanobis distance exert better performance than conventional Eucledian-based workbooks. The code runtime is also fast and the response time is less than 0.2 seconds.The purpose of this research investigation is to evaluate different faces. Identification method provides complete solution for image face recognition and better identification. The better step of video surveillance is the correct answer. This solution is based on very rich database testing.

## REFERENCES


[1]. Mr. Dinesh Chandra Jain Dr. V. P. Pawar Research Schoolar in Comp-Science Associate Professor , School of Comp- Science SRTM University – Nanded (M.S) SRTM University – Nanded (M.S) "A Novel Approach For Recognition Of Human Face Automatically Using Neural Network Method" Volume 2, Issue 1, January 2012 ISSN: 2277 128X.

[2]. Sujata G. Bhele1 and V. H. Mankar2 "A Review Paper on Face Recognition Techniques" ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 8, and October 2012.

[3]. MamtaDhanda Assist. Prof (Computer Department), Seth Jai Prakash MukundLal Institute of Engineering and Technology, Radaur "Face Recognition Using Eigenvectors From Principal Component Analysis" International Journal of Advanced Engineering Research and Studies E-ISSN2249–8974.

[4]. M.Nandini, P.Bhargavi, G.RajaSekhar Department of E-Conm, SreeVidyanikethan Engineering College "Face Recognition Using Neural Networks" International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013 1 ISSN 2250-3153.

[5]. Michel F. Valstar, TimurAlmaev, Jeffrey M. Girard, Gary McKeown, Marc Mehu, Lijun Yin, MajaPantic and Jeffrey F. Cohn: "FERA 2015 - Second Facial Expression Recognition and Analysis Challenge" 978-1-4799-6026-2/15/$31.00 ©2015 IEEE.

[6]. Marian Stewart Bartlett, Member, IEEE, Javier R. Movellan, Member, IEEE, and Terrence J. Sejnowski, Fellow, IEEE: "Face Recognition by Independent Component Analysis" IEEE Transactions on Neural Networks, Vol. 13, No. 6, November 2002.

[7]. Volker Blanz and Thomas Vetter, Member, IEEE "Face Recognition Based on Fitting a 3D Morphable Model" IEEE Transaction On Pattern Analysis And Machine Intelligence, Vol. 25, No. 9, September 2003.

[8]. Juwei Lu, Student Member, IEEE, Konstantinos N. Plataniotis, Member, IEEE, and Anastasios N. Venetsanopoulos, Fellow, IEEE "Face Recognition Using Kernel Direct Discriminate Analysis Algorithms" IEEE Transactions On Neural Networks, Vol. 14, No. 1, January 2003.

[9]. WU Xiao-Jun Josef Kittler YANG Jing-Yu Kieron Messer Wang Shi-Tong Department of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, China: "A New Kernel Direct Discriminant Analysis (KDDA) Algorithm for Face Recognition" Win2PDF available at http://www.daneprairie.com.

[10]. Juwei Lu, K.N. Plataniotis, A.N. Venetsanopoulos Bell Canada Multimedia Laboratory, "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms" IEEE Transactions on Neural Networks in December 12, 2001. Revised and re-submitted in July 16, 2002. Accepted for publication in August 1, 2002.

[11]. Felix Juefei-Xu, KhoaLuu , MariosSavvides , Tien D. Bui, and Ching Y. Suen"Investigating Age Invariant Face Recognition Based on Periocular Biometrics" 978-1-4577-1359-0/11/$26.00 ©2011 IEEE.

[12]. GuodongGuoGuowang Mu Yun Fu BBN Technologies Thomas S. Huang UIUC "Human Age Estimation Using Bio-inspired Features" 978-1-4244-3991-1/09/$25.00 ©2009 IEEE

[13]. Thomas SerreLior Wolf TomasoPoggio "Object Recognition with Features Inspired by Visual Cortex" Brain and Cognitive Sciences Department Massachusetts Institute of Technology Cambridge, MA 02142
[14]. Anjali1, Asst. Prof. Arun Kumar2 "Face Recognition And Age Progression Using Aaia Algorithm" International Journal for Technological Research in Engineering Volume 2, Issue 11, July-2015.
[15]. Young H. Kwon∗ and Niels da Vitoria Lobo "Age Classification from Facial Images" School of Computer Science, University of Central Florida, Orlando, Florida 32816 Received May 26, 1994; accepted September 6, 1996.
[16]. AsumanGünay and Vasif V. Nabiyev "Age Estimation Based on AAM and 2D-DCT Features of Facial Images" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 2, 2015"
[17]. XinGeng, Kate Smith-Miles, Zhi-Hua Zhou "Facial Age Estimation by Learning from Label Distributions" ARC (DP0987421), NSFC (60905031, 60635030), JiangsuSF (BK2009269), 973 Program (2010CB327903) and Jiangsu 333 Program".
[18]. Dr. Asmahan M Altaher, "Face Recognition Techniques - An evaluation Study," Volume: 6 Issue: 4 Pages: 2393-2397 (2015) ISSN: 0975-0290.
[19]. Priyanka, "A Study on Facial Feature Extraction and Facial Recognition Approaches," IJCSMC, Vol. 4, Issue. 5, May 2015, pg.166 – 174.
[20]. PratibhaSukhija, Sunny Behal, Pritpal Singh, "Face Recognition System Using Genetic Algorithm," International Conference on Computational Modeling and Security (CMS 2016), Procedia Computer Science 85 (2016) 410 – 417.
[21]. Mahmoud Afifi et.al, "Can We Boost the Power of the ViolaJones Face Detector Using Pre-processing? An Empirical Study," pp.1-14, 2017.
[22]. Hajime Nada et al., "Pushing the Limits of Unconstrained Face Detection: a Challenge Dataset and Baseline Results," IEE, 2018.

Vaishali " Advance Face Recognition through PCA and Mahalanobis Distance . "International Journal of Computational Engineering Research (IJCER), vol. 08, no. 11, 2018, pp 26-32.