



Data Retrieval using Face Recognition- A Review

Yogesh Chavan¹, Akash Kajrolkar², Tushar Gaikwad³, Neha Chankhore⁴
Student^{1,2,3,4}, Assistant Professor⁴

Department of Information Technology Engineering
Vidyalankar Institute of Technology, Mumbai, India

Abstract:

Face is a complex multidimensional structure and needs a good computing technique for recognition. An ideal face contains various facial feature e.g. eyes, nose, mouth, eyebrow. In this project face recognition is done by Local Binary Pattern Histogram (LBPH). Local Binary Pattern Histogram (LBPH) is a robust method used as feature extraction techniques for face recognition. LBPH is one of the easiest face recognition algorithms. The basic form of the LBPH is to minimize the dimensionality of the data set that consist of a large number of the correlated variables, while maintain as possible of the variation present in the set of data. It can represent local features in the images. Face will be categorized as known or unknown face after matching with the present database. If the user is new to the face recognition system then his/her template will be stored in the database and also all his documents will be stored in another database. If the user is a registered user then we can retrieve his/her documents using face recognition.

Keywords: Local Binary Pattern Histogram (LBPH), Face Recognition, Correlated Variable

I. INTRODUCTION

Face recognition is a biometric personal identification based on the human physical characteristics that is a human face. This method has the ability to identify a person from still image by using human face. For the accurate recognition, algorithm and reference database needs to be concerned. Face recognition is a useful identification method that overcomes the others by several advantages. Face recognition is performed by independently matching templates of facial regions such as eyes, nose, mouth etc. . This method can be done without any device participation from user, and the images can be captured from distance. Furthermore, a necessary device of this method is any kind of camera according to a system demand. Thus, the face recognition provides an inexpensive and reliable identification system, which has feasibility to apply in many purposes such as surveillance, access control system, etc. Face recognition starts with the detection of face patterns in sometimes cluttered scenes, proceeds by normalizing the face images to account for geometrical and illumination changes, possibly using information about the location and appearance of facial landmarks, identifies the faces using appropriate classification algorithms, and post processes the results using model based schemes and logistic feedback

Face Recognition can be of two types:

Feature Based (Geometric)

Template Based (Photometric)

In geometric or feature-based methods, facial features such as eyes, nose, mouth, and chin are detected. Properties and relations such as areas, distances, and angles between the features are used as descriptors of faces. Although this class of methods is economical and efficient in achieving data reduction and is insensitive to variations in illumination and viewpoint, it relies heavily on the extraction and measurement of facial features. Unfortunately, feature extraction and measurement techniques and algorithms developed to date have not been reliable enough

to cater to this need. In contrast, template matching and neural methods generally operate directly on an image based representation of faces, i.e., pixel intensity array. Because the detection and measurement of geometric facial features are not required, this type of method has been more practical and easier to implement when compared to geometric feature based methods.

II. AIM AND OBJECTIVES

Aim:

The aim of our project is to retrieve stored data from database using face Recognition. Various algorithms and methodologies are studied and hardware resources are done to achieve the goal. This kind of face recognition system can be widely used in our daily life in different sectors. This will help various organizations such as passport offices, banks, government offices to recognize a particular user and retrieve his/her information and also their documents. Using this system the user will not have to provide his/her documents at various organizations. This will also help the organizations to identify fake identity as well as fake documents.

Objectives:

Trying to find a face within a large database of faces. In this approach the system returns a possible list of faces from the database and hence we can retrieve the required documents. Real time face recognition: Here, face recognition is used to identify a person on theSpot as it can be used for security purpose to identify fake user as well as fake information given by the user. This system can also be used at ATM's to avoid forgery.

III. LITERATURE SURVEY

Principal component Analysis: The Principal component analysis (PCA) is a one of the powerful statistical technique used to

analyse the data in classes. The basic form of the PCA is to minimize the dimensionality of the data set that consist of a large number of the correlated variables, while maintain as possible of the variation present in the set of data. The mathematical formula of PCA is clinging in the standard deviation, and the eigenvectors, and the eigenvalues .The final process of facial recognition is distance measures. Different methods used to estimate the distance between the test image and data base of images. The classification satisfied with the minimum distance between the training faces and the test image. Many of approaches are introduced to optimize the distance of image recognition. The PCA algorithm reduces effectively the dimensions of the human face images also keep its status of identifying information. The mathematical form of PCA takes three primary steps. The first step is creating a transformation matrix using the training of the human face images. The second step, the training face images are arranged into the matrix vectors. Ultimately, the test face image is recognize by introducing it in the subspace of Eigen faces and compared with the trained face images in the subspace of Eigen face domain[1]. The major advantage of PCA is using it in Eigen face approach which helps in reducing the size of the database for recognition of a test images. The images as stored as their feature vectors in the database which are found out projecting each and every trained image to the set of Eigen faces obtained. PCA is applied on Eigen face approach to reduce the dimensionality of a large data set[2].

Independent Component Analysis (ICA): Its minimizes both second-order and higher-order dependencies in the input. It keeps the assumption of linearity but abandons all other that PCA uses. Although the amplitude spectrum is captured by second-order statistics in PCA, there remains the phase spectrum that lies in higher-order statistics. It is believed that this high-order statistics (i.e. the phase spectrum) contains the structural information in images that drives human perception .ICA attempts to find the basis along which the data (when projected onto them) are statistically independent. ICA is a way of finding a linear non-orthogonal coordinate system in any multivariate data. The directions of axes of this coordinate system are determined by both the second and higher order statistics of the original data. The goal is to perform a linear transform, which makes the resulting variables as statistically independent from each other as possible[2].

Linear Discriminant Analysis (LDA):Both PCA and ICA do not use face class information. Linear Discriminant Analysis (LDA) finds an efficient way to represent the face vector space by exploiting the class information. It differentiates individual faces but recognizes faces of the same individual. LDA searches for vectors in the underlying space that best discriminate among classes. For all the samples of all classes, two measures are defined[2].

Local Binary Pattern Histogram(LBPH):**Local Binary Pattern (LBP)** is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. It was first described in 1994 (LBP) and has since been found to be a powerful feature for texture classification. It has further been determined that when LBP is combined with histograms of oriented gradients (HOG) descriptor, it improves the detection performance considerably on some datasets. Using the LBP

combined with histograms we can represent the face images with simple data vector. LBPH uses 4 parameters. The four parameters are

- **Radius:** the radius is used to build the circular local binary pattern and represents the radius around the central pixel. It is usually set to 1.
- **Neighbors:** the number of sample points to build the circular local binary pattern. Keep in mind: the more sample points you include, the higher the computational cost. It is usually set to 8.
- **Grid X:** the number of cells in the horizontal direction. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8.
- **Grid Y:** the number of cells in the vertical direction. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8.

LBPH is one of the easiest face recognition algorithms. It can be represent local feature in the images. It is possible to get result (mainly in a controlled environment).It is robust against monotonic grayscale transformation.

IV. PROBLEM STATEMENT

Here a face recognition system is developed that is used for authentication purpose to retrieve particular users documents stored in database. Principal component analysis (PCA) is used to develop this system. In this system we can store all the documents of users which can be later retrieved using face recognition. Recognition includes comparing and matching faces with the databases in which faces are available.

V. USEFULNESS

Biometrics is the measurement and statistical analysis of people's physical and behavioural characteristics. The technology is mainly used for identification and access control, or for identifying individuals that are under surveillance. The basic premise of biometric authentication is that everyone is unique and an individual can be identified by his or her intrinsic physical or behavioural traits. By using biometrics, it is possible to confirm or establish an individual's identity based on "who she is," rather than by "what she possesses" (e.g., an ID card) or "what she remembers" (e.g., a password). There are many types of biometric system like fingerprint recognition, face detection and recognition, iris recognition etc., these traits are used for human identification in surveillance system, criminal identification. Advantages of using these traits for identification are that they cannot be forgotten or lost. These are unique features of a human being which is being used widely.

VI. PROPOSED SYSTEM

This system will have following working features:

- 1.Admin login: Here the admin will be authenticated using face recognition as well as login id and password for cross verification.
- 2.After admin login, users face will be captured using camera and will be categorized as known or unknown face after matching with the present database.

3.If the user is new to the face recognition system then his/her template will be stored in the database and also all his information and documents will be stored in another database. The following four-stage process illustrates the way face recognition systems operates:

Capture - A physical or behavioral sample is captured by the system during enrollment

Extraction - Unique data is extracted from the sample and a template is created

Comparison - The template is then compared with a new sample

Matching - The system then decides if the features extracted from the new sample are matching or not

4.If the user is a registered user then we can retrieve his/her documents using face recognition.

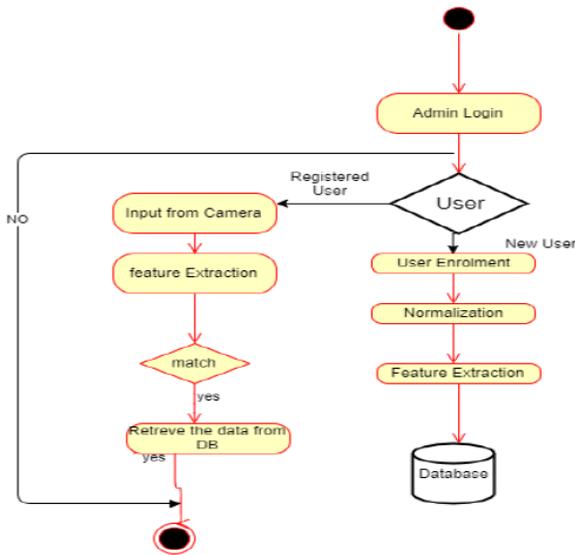


Figure.2. Flowchart

VII. SCOPE

This system will help various organizations such as passport offices, banks, government offices to recognize a particular user and retrieve its information and also their documents. Using this system the user will not have to provide his/her documents at various organizations. This system is used for security purpose to identify fake user as well as fake information given by the user. This system can also be used at ATM’s to avoid forgery. In order to prevent the frauds of ATM in India, it is recommended to prepare the database of all ATM customers with the banks in India & deployment of camera and face recognition software at all ATMs. So, whenever user will enter in ATM his photograph will be taken to permit the access after it is being matched with stored photo from the database.

VIII. METHODOLOGY

The platform used to develop this system is python. The diverse application of the Python language is a result of the combination of features which gives this language an edge over others. Some of the benefits of programming in Python include:

Presence of Third Party Modules: The Python Package Index (PyPI) contains numerous third-party modules that make Python

capable of interacting with most of the other languages and platforms.

Extensive Support Libraries: Python provides a large standard library which includes areas like internet protocols, string operations, web services tools and operating system interfaces. Many high use programming tasks have already been scripted into the standard library which reduces length of code to be written significantly.

Open Source and Community Development: Python language is developed under an OSI-approved open source license, which makes it free to use and distribute, including for commercial purposes.

The algorithm used to develop face recognition is Principal Component Analysis (PCA).Principal Component Analysis (PCA) is a method to efficiently represent a collection of sample points, reducing the dimensionality of the description by projecting the points onto the principal axes, where an orthonormal set of axes points in the direction of maximum covariance in the data. These vectors best account for the distribution of face images within the entire image space. PCA minimizes the mean squared projection error for a given number of dimensions, and provides a measure of importance for each axis.

IX. CONCLUSION

This system will help various organizations such as passport offices, banks, government offices to recognize a particular user and retrieve its information and also their documents. Using this system the user will not have to provide his/her documents at various organizations. This system is used for security purpose to identify fake user as well as fake information given by the user.

BLOCK DIAGRAM

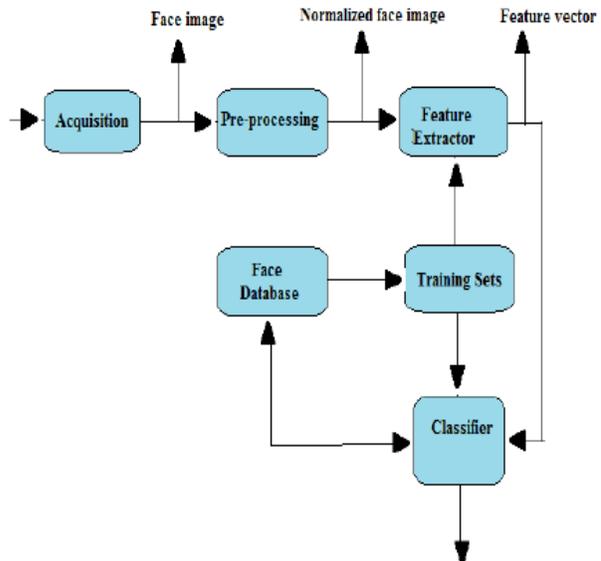


Figure.2. Face recognition block diagram

X. REFERENCES

[1]. 2017 International Conference on Current Research in Computer Science and Information Technology (ICCI), Slemani Iraq "Face Rate Using Different Classifier Method Based On PCA ".

[2]. 47th International Symposium ELMAR-2005, 08-10 June 2005, Zadar, Croatia Appearance based Statistical Methods for Face Recognition Panos Liatsis Information and Biomedical Engineering Centre, City University London, UNITED KING3.

[3]. IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2016), December 23-25, 2016, Jaipur, India "Face Detection and Recognition". Divya Meena Department of Computer Science and Engineering Central.

[4]. 2013 Fourth International Conference on Emerging Security Technologies. "Age Sensitivity of Face Recognition Algorithms" DK H PHM Yassin, S Hoque, and F. Deravi School of Engineering and Digital Arts, University of Kent, Canterbury, United Kingdom.

[5]. Proceedings of the International MultiConference of Engineers and Computer Scientists 2016 Vol I, IMECS 2016, March 16 - 18, 2016, Hong Kong. "Improved Method on PCA Based Human Face Recognition for Distorted Images" Bruce Poon, M. Ashraful Amin, and Hong Yan.

[6]. 2016 23rd Iranian conference on biomedical engineering, tehran, iran 23-25 november 2016 "A new hybrid face recognition algorithm based on discrete wavelet transform and direct LDA".

[7]. School of communication and information engineering, xian university of post & telecommunication, Xian, china, jie Kang, xin LIN, "research on multi-scale pca algorithm for face recognition"

[8]. 3rd IEEE international conference on "computational intelligence and communication technology"

[9]. Dept of electronic and communication engineering, R.J.I.T Takenpur, gwalior, aparna.ras90@gmail.com "low resolution face recognition technique".

[10]. Kresimir delac, croatian telecom, savska 32, zagreb, CROATIA "appearance based statistical method for face recognition"

[11]. School of information science and engineering, northeastern university, shenyang 110819 "A comparative study of face recognition algorithm on R1 face database"