Face Recognition For Automatic Voters' Registration And Accreditation System Using Principal Component Analysis (PCA): An Evaluation

[1]**Mary S. Ele,**
Department of Statistics, Cross River University of Technology,
Calabar, Nigeria

[2]**Emereonye, Goodluck Ikwudiuto**
Department of Computer Science, Akanu Ibiam Federal Polytechnic Unwana, Afikpo, Nigeria

[3]**Ele, Sylvester I.**
Department of Computer Science, University of Calabar
Calabar, Nigeria.

Corresponding Author: Ele, Sylvester I. myyrs2015up@gmail.com

***Abstract*—This research paper describes a method for implementing an automatic voters' registration and accreditation system that will incorporate face recognition technology using Principal Component Analysis (PCA) algorithm. The voter's identity will be matched using the facial recognition system to circumvent the fake voting. Due to the invasive nature of face recognition, it remains the most focused research area and has become an evolving universal biometric solution as it requires almost no effort at the user end when it is compared to biometric. The system adopted an information theory approach whereby the face images was decomposed into a small set of characteristic feature images known as Eigenfaces, which is the principal components of the original training set of the face images. PCA is considered here above other approaches because of its low noise sensitivity; decreased requirements for capacity and memory and increased efficiency.**

Keywords—PCA, Face Recognition, Voters' Registration, Artificial Neural Network.

1. Introduction

Being one of the most successful applications of the image processing, face recognition has a vital role in technical field especially in the field of security. Human face recognition is an important field for verification purpose especially in the case of election's voters' registration and accreditation system.

Voting is a right and taking part in the government of your country is the cornerstone of democracy and without a comprehensive, credible, trustworthy and reliable election roll, this right cannot be exercise. Instituting voting integrity is a foundation stone of modern democracy that promotes belief, trust and honesty in elected governments and upholds confidence and faith in elected leaders. Independent states are more and more turning to biometric voting systems to help support fair and credible elections devoid of fraud and illegal practices to ensure the protection of democratic philosophy, beliefs and ideology.

The results of recent biometric voting exercises in countries such as Nigeria in 2015 general election have brought us to the realization of the fact that governments expect fast, accurate, and reliable voter registration at the polls under any conditions that help to maintain the integrity and credibility of the electoral process and reduce distrust and abnormalities.

Many African countries, including Nigeria are still finding themselves confronted with difficulties in registering and authenticating voters. Nigeria, for example, have for so many years now been seeking to ensure voter equality, based on the principle of one man, one vote, in other words that everyone's vote should count equally. It is a well known fact that a democratic, reliable and fraud free electoral process is an important factor in establishing lasting peace, security and stability in a country. Consequently, Nigeria and many other African countries are witnessing a growing desire to digitize and secure data and documents related to civil identity. No fewer than 25 sub-Saharan African countries (e.g. Sierra-Leone, Democratic Republic of Congo, Zambia, Malawi, Rwanda, Senegal, Somaliland, Mali, Togo, Ghana etc.) have already held elections employing a biometric voter register (Piccolino, 2015).

Prior to the 2015 general elections, a number of technologically based reforms such as Biometric

Register of Voters, Advanced Fingerprints Identification System and smart card readers were embarked upon. It was used for the first time in the electoral process of Nigeria and it remains one of the greatest technological innovations of the 2015 general elections. The smart card reader is a technological device setup to authenticate and verify on Election Day a Permanent Voter Card (PVC) issued by Independent National Electoral Commission (INEC).

Despite the use of voters' Card Readers in the 2015 general election in Nigeria, the entire electoral process clearly demonstrated a lack of credibility, distrust, fraud and other irregularities. Registration of voters as provided by the independent National Electoral Commission (INEC) was deficient in several respects.

- Some polling units did not have the name of any voter on the register (the so-called zero units), notwithstanding, there were people in those units who had authentic and valid voter's card.
- Several voters who turned up with valid voter's cards did not have their names on the voters' registers.
- Names of voters were swapped between units and wards that were unrelated.

There were several cases of either mismatched or missing voters' photographs.
Notwithstanding the number of systems available today, face detection remains a challenge because of dissimilarity in facade, illumination and expression. Problems occur mostly when searching for an image in the database containing side view face images. The proposed system, when completed will serve as a useful reference material in the field of machine language. The independent National Electoral Commission (INEC), the Nigerian electorates, the Nigerian democratic system, and the Nigerian state in general, including all other developing countries who wish to adopt a more sophisticated technology for the conduct of its elections. Moreover, researcher in the field of Artificial Intelligence, especially in the area of machine learning and artificial Neural Networks in particular will benefit from the algorithm that will be developed in this proposed research work, by either improving on it or use it directly to implement other systems.

According to Kar, Debbarma, Saha and Pal (2012), authentication is a significant issue in system control in computer based communication. Human face reorganization is an important biometric verification and has widely used in many applications such as video monitoring system, human computer interaction, door control system and network security. Nigeria is barely 18 years in full Democratic practice. The Electoral System and process, like any other African country, is characterized with flaws, fraud and many other irregularities. The 2019 general election is by the corner, and the country needs an enhanced registration and accreditation system with security to guarantee a fairly credible and malpractice free election that will restore a bit of confidence to our democratic structure.

1.2. Timely Justification of the Research

This research is appropriate at this time in various ways. Image recognition, and computer vision in broader sense, is integral to a number of emerging technologies, from high-profile advances like driverless cars and facial recognition software to more ordinary but still important developments, like building smart factories that can spot defects and irregularities on the assembly line, or developing software to allow insurance companies to process and categorize photographs of claims automatically. Today, data scientists are confronted with the challenges of categorizing and searching content in cases where metadata is sparse or nonexistent.

The recent development of a number of APIs that aim at allowing institutions to assemble insights from images without requiring in-house computer vision or machine learning proficiency makes this research more appropriate. Example of this technologies include

i. The Google Cloud Vision is Google's visual recognition API, which based on the open-source TensorFlow framework making use of the REST API
ii. Another API is the Clarif.ai, an upstart image recognition service that also uses a REST API and
iii. The IBM Watson Visual Recognition, which is part of the Watson Developer Cloud, with a large set of built-in classes, but is really built for training custom classes based on images you supply.

With these in mind, it is evident that the study and application of principal component analysis (PCA) is appropriate because more indebt study is required and more efficient algorithms are needed to develop more solution – based system to meet the demand of the world teeming industrial and economic populace. The general architecture of typical voters' registration system is illustrated in figure 1.

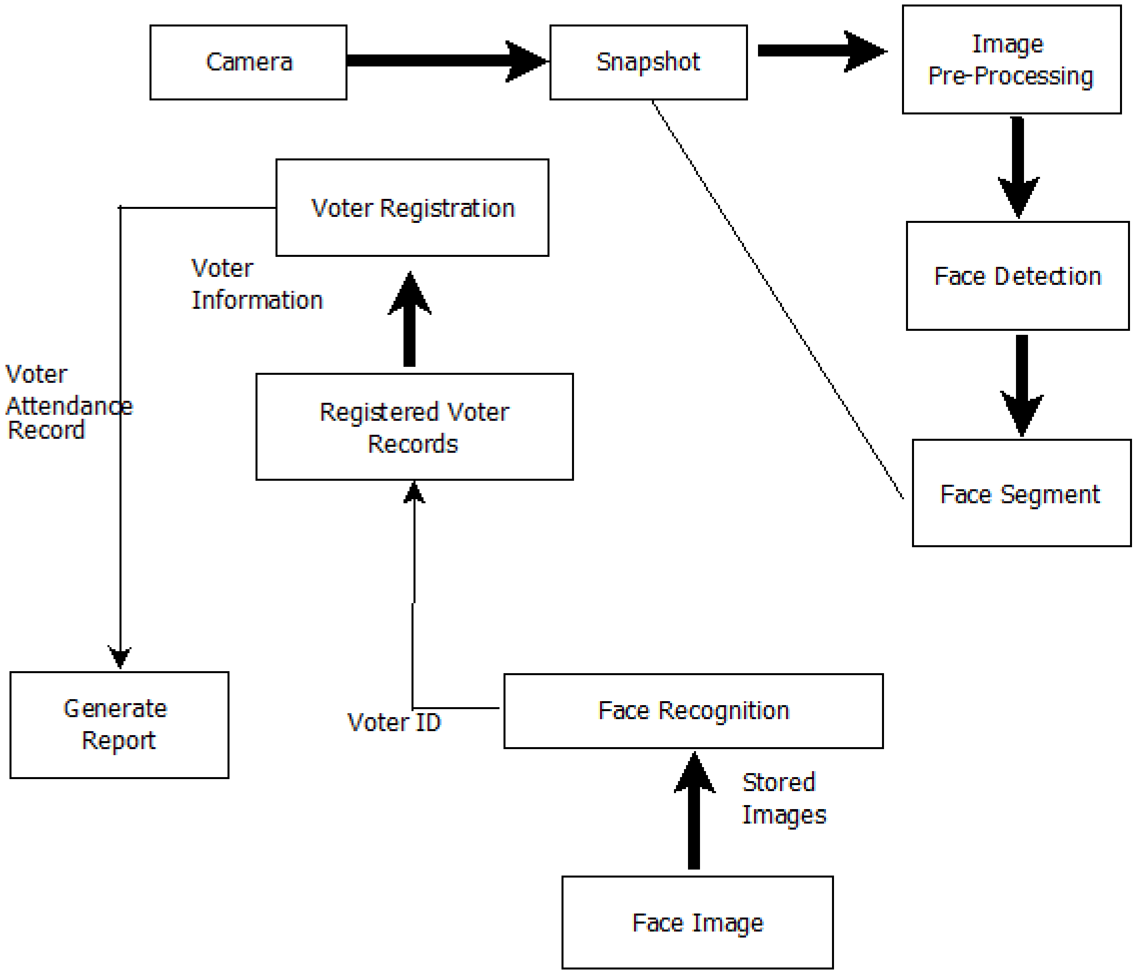


Figure 1: Architecture of Voters' Registration and Authentication System
Source: Author (modified)

2.0 Review of Related Work

Several academic works has been done in the area of Artificial Neural Network (ANN) in face recognition and detection, image processing, and pattern recognitions. Agrawal et al (2017) studied on Webcam Based Facial Recognition Using OTP Verification for Voting System. In their study, Agrawal et al developed a system that captured the face of a voter and match with the existing faces in the stored databases. After the verification of the valid face identify, an OTP is generated and send to the voters register mobile and then he will use that provided OTP and then only his validation is succeeded and allowed to vote. Their adopted algorithm was based on information theory which decomposes the face images. It then forms the minute set of the characteristics features images which are called as "Eigen faces". The algorithm has two set of image blocks, first one is training set image block and second one is test set image block. In training set image block,

Thakare et al (2016) developed a framework for face detection and recognition for automatic attendance system. Their proposed system framework takes the participation naturally utilizing face identification and recognition. This participation is recorded by utilizing a camera connected as a part of front of classroom that is continuously catching pictures of students, detect the faces in image and contrast the distinguished appearances and the database and mark the attendance.

Khan and Khiyal (2017) studied on a Side-View Face Detection Using Automatic Landmarks. Their study focuses on side-view profiles. And to take into consideration extreme pose variations, landmarks were used. Commonly used features for landmarks are eyes, nose, lips, ear lobes, edges of the mouth, arch corners of the eyebrows, chin and nasion. In order to broaden the mapping area, ears as a feature was also included. For feature extraction, they registered the images using labeled landmarks. The aim of their research was to automatically generate 15 landmarks and then employ a distance based indexing method based on nearest neighbor to overcome the problem identified in their study. Their image extraction was done using a baseline algorithm and along with nearest neighbor searching scheme with the following steps.

1. Face Detection

2. Calculating Value differences
3. Normalize the result
4. Converting the image into binary form
5. Extraction of connected components
6. Automatic Landmark detection
7. Storing landmark position in Array
8. Nasal marking and Edge Detection for face image
9. Height density search
10. Slope calculation on the graph and
11. Similar pictures search using nearest neighbor

Joseph and Zacharia (2013) researched on Automatic Attendance Management System Using Face Recognition. Their proposed system consisted of a high resolution digital camera to monitor the classroom or office room. It was embedded on a micro-controller based motor system which enables it to rotate in left & right directions. The data or images obtained by the camera are sent to a computer programmed system for further analysis. The images obtained are then compared with a set of reference images of each of the employees or students & mark the corresponding attendance. The system also provides for continuous monitoring of the classroom by an operator if needed. The camera module can be a wireless or wired system.

MATLAB was used to develop the Face Recognition module. Initially the images of each of the users are provided for the MATLAB and generate a set of facial features using of the feature extraction methods namely, PCA (Principal Component Analysis). Images of human face was extracted using MATLAB's Image Acquisition Toolbox, in which a camera is configured, accessed and brought one frame at a time into MATLAB's workspace for further processing using MATLAB's Image Processing Toolbox. Joseph and Zacharia used a PIC micro-controller along with as servo motor mechanism as Micro-controller based embedded hardware system.

Per Parmar and Mehta (2013), face recognition systems have been carried out for almost 50 years now. It is one of the researches area in pattern recognition & computer vision as result of its many practical applications in the area of biometrics, Information security, access control, law enforcement, smart cards and surveillance system. Agrawal et al (2017) state several system of voting such as; the Direct Recording Electronic (DRE) voting, Electronic Vote Collector (EVC), Online Voting System with Multi Security using Biometric and Steganography, finger print recognition, and remote voting through internet. According to Gupta et al (2010), DRE systems completely eliminate paper ballots from the voting process. In this system, when the voter‟s selection is complete, DRE systems will typically present a curt of the voter‟s selections, giving them a last chance to make changes. Subsequent to this, the ballot is “cast” and the voter is free to leave. In EVC, the voters drop their votes on their own personal computers, even as a mobile device pass close the machines and collect their stored votes, with the help of management software working in a stationary server. Highly Protected Online Voting System with Multi Security using Biometric and Steganography is meant to merge the secret key with the cover image on the basis of core image (Agrawal et al, 2017). The fingerprint recognition or fingerprint verification is an automated method of verifying a match between two Human fingerprints. It looks at the patterns set up on a fingertip. Fingerprint authentication would possibly be a good choice for a touch sensing voting systems, where you can give users adequate enlightenment and guidance, and where the system operates in a controlled environment (Ravi and Wilson, 2010).

2.1 Principal Component Analysis (PCA) for Face Recognition

PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension (Smith, 2002). According to Paul and Sumam (2012), PCA is a statistical approach used for reducing the number of variables in face recognition. In PCA, every image in the training set is represented as a linear combination of weighted eigenvectors called eigenfaces, Vyas (2016) sees PCA as a holistic based Statistical method which is used to extract the feature from face image and to decrease the large dimensionality of the data to the smaller dimensionality of feature space, then classification is done using Euclidian distance classifier to recognize the face. Principal component analysis (PCA) involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components (Gunturk, Batur and Altunbasak, 2003).

PCA is a popular technique, to derive a set of features for both face recognition. Any particular face can be economically represented along the eigen pictures coordinate space, and approximately reconstructed using a small collection of Eigen pictures (Latha, Ganesan and Anadurai,). Abdi and Williams (2010), defined PCA as a statistical procedure that uses an orthogonal transformation to convert a sets of observations of possibly correlated variable into a set of values of linearly uncorrelated variables called principal components. Based on Jolliffe (2002), PCA can be thought of as fitting an *n*-dimensional ellipsoid to the data, where each axis of the ellipsoid represents a principal component. If some axis of the ellipsoid is small, then the variance along that axis is also small, and by omitting that axis and its corresponding principal component from our representation of the dataset, we lose only a commensurately small amount of information.

As noted by (2016), every image in the training set is represented as a linear combination of weighted eigenvectors called eigenfaces. These eigenvectors are obtained from covariance matrix of a training image set. Face recognition technique is divided into three categories. Feature based, holistic and hybrid approach where holistic approach is very popular approach for face recognition (Jafri and Hamid 2009). Principal Component Analysis reduces the dimension of the image form 2 dimensional to single dimensional vector, containing row and column. It calculates the mean and average matrix. The average matrix calculated is then subtracted from original face image and stored in a variable. The covariance matrix is computed. Thereafter, the Eigen vector and Eigen values of covariance matrix is calculated. This leads to the generation of the Eigen faces. The new Eigen face is stored as Eigen component. At this point, the Euclidean distance Classifier measure the weight between two weight vectors and match the vectors with training set and recognizes the image.

2.2 **Covariance**

Covariance is a property that gives us the amount of variation of two dimensions from their mean with respect to each other. This property is used on the DCT-matrix so that a form of distance measure is performed on the image pixel values thus providing their relative intensity measures ASHOK and RAJAN (2010). It provides a measure of the strength of the correlation between two or more sets of random variates. Let *X* and *Y* be random variables (discrete or continuous!) with means μx and μy . The covariance of X and Y given as $cov(X,Y)\ or\ \partial XY$ is defined as:

$$cov(X,Y) = \partial XY\ E\ [(X - \mu x)(Y - \mu y)]$$

If *X* and *Y* are discrete random variables with joint support *S*, then the covariance of *X* and *Y* is:

$$cov(X,Y) = \sum_{x,y} \sum_{\in s} (x - \mu X)(y - \mu Y) f(x,y)$$

If *X* and *Y* are continuous random variables with supports S_1 and S_2, respectively, then the covariance of *X* and *Y* is: $cov(X,Y) = \int_{s_2} \int_{s_1} (x - \mu X)(y - \mu Y) f(x,y)\, dy\, dx$

In this article, one of the foremost steps in our algorithm is the covariance matrix. To calculate eigenvectors to generate the eigenfaces, we need to calculate the covariance matrix.

3.0 The PCA Face Recognition Algorithm

One of the simplest and far most efficient and effective PCA approaches used in face recognition systems is the eigenface approach. This approach converts faces into a small set of critical and important characteristics, called eigenfaces, which are the main components of the initial set of training set (learning images). The entire recognition process is broadly divided into two steps: an Initialization process and Recognition process (Turk and Pentland, 1991). The initialization process involves acquire the initial set of face images called as training set, calculate the Eigenfaces from the training set and computation of distribution in this M-dimensional space for each known person by projecting his or her face images onto this face-space. While the recognition process involves the calculate a set of weights based on the input image and the M eigenfaces; determine if the image is a face at all (known or unknown); if it is a face, then classify the weight pattern as either a known person or as unknown and then update the eigenfaces or weights as either a known or unknown (Paul and Sumam, 2012). The entire process architecture is represented in figure 2.

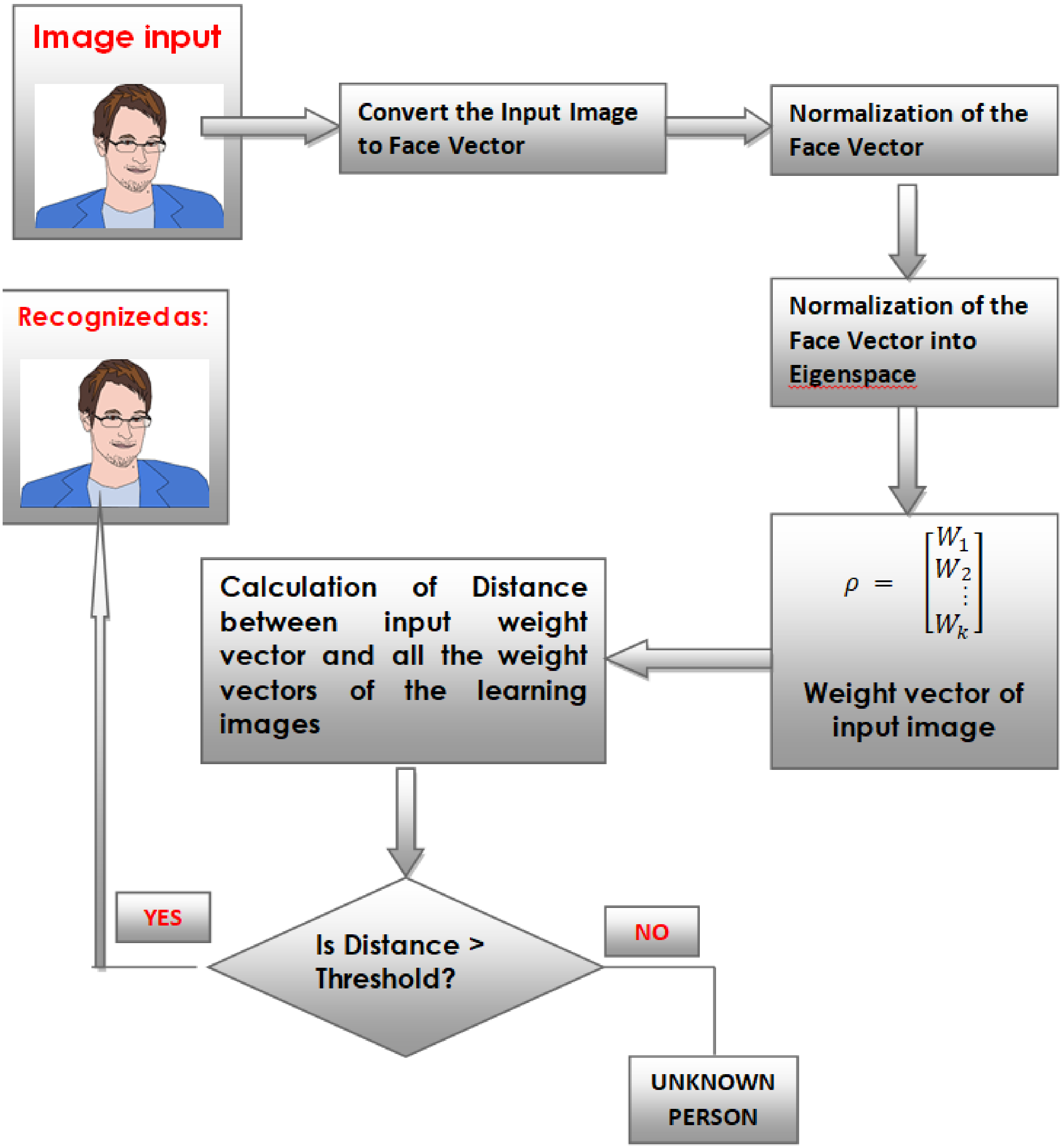


Figure 2: PCA Face Recognition Process
Source: Author (modified)

3.1 Description of Principal Component Analysis (PCA) Algorithm

Figure 2: A training set consisting of a total of M images
Source: https://www.gettyimages.com/photos/human-face

The PCA technique in the Eigenface Algorithm explains that the number of selected eigenvectors (Eigenfaces) forming up the face space must not be greater than the number of original face images (Alabi, Akanbi and Ibrahim, 2015). The description of the PCA algorithm is shown in steps as follows:

Step 0: Creation and loading of Training set.
Here we use a set of image of 103 X 78 pixels as shown above. Since an image may also be considered as a vector of dimension M × N, our image of size 107 × 78 becomes a vector of dimension 8, 034, that is a point in 8,034 dimensional space.

Step 1: Preparation of training faces. Obtaining a face images I_1, I_2, I_3, I_4, ...I_5, which is the training images.

Step 2: Convert face images in Training Set to Face Vectors. Each training image I_i in the dataset is transformed into a vector and placed into a training set Y = {Y_1, y_2, y_3, y_4, ... y_n}. in our dataset, M = 50 and each image is transformed into a vector of size MN X 1as show in figure 3 and placed in the set.

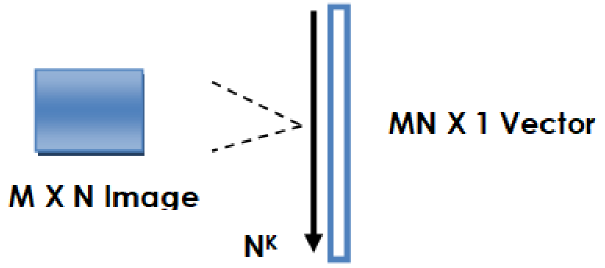


Figure 3: Conversion of M X N Image into MN X 1 Vector

Step 3: Normalization the Face vector. Here, the average face vector (Ψ) is to be calculated with the formula:

$$\varphi = \frac{1}{M}\sum_{n=1}^{M} y_n \ .$$

Step 4: Subtract the mean (average) face vector φ from each face vectors y_i to get the normalized face vectors and result stored in variable β,

$$\beta = y_i - \varphi.$$

Step 5: Calculate the covariance matrix. To calculate eigenvectors, we need to calculate the covariance matrix $\boldsymbol{C}$.

$\boldsymbol{C} = A.A^T \quad (N^2\ X\ N^2\ matrix)$,
Where:

$$A = [\beta_1, \beta_2, \beta_3, \ \beta_4, \dots . \beta_M]\ (N^2\ X\ M\ matrix).$$

$$C = \frac{i}{M}\sum_{n=1}^{M} \beta_n \beta_n^T$$

There is need for dimensionality reduction. The matrix $A.A^T$ is very large, because with the dimension $N^2\ X\ N^2$, means that there would be N^2 eigenfaces and eigenvalues, implying that for 200 X 200 image you must calculate a 40,000 X 40, 000 matrix and compute 40, 000 eigenfaces. What happens in this case is that the system will slow down terribly or run out of memory, because the computation requires are enormous. The solution is dimensionality reduction. This is achieved by calculating eigenvectors from a covariance with reduced dimensionality. This leads us to the next step.

Step 6: Reduction of Dimensionality of Training Set. To reduce calculations and effect of noise on the needed eigenvectors, u_i , calculate them from a covariance matrix, AA^T of reduced dimensionality.

$$u_i = \sum_{i=1}^{M} v_i \beta_i$$

Where $u_i$the eigenvectors (eigenfaces).

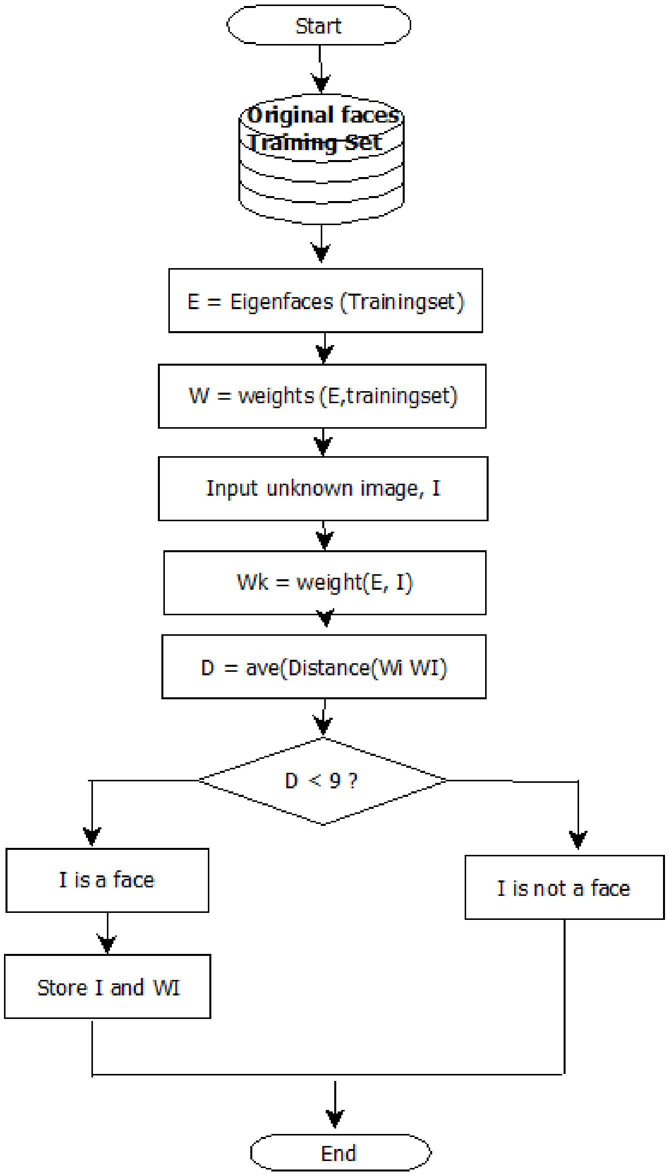
Step 7: Select K best eigenvectors, such that K < M and can represent the whole Training set.

Step 8: Convert lower dimensional K eigenvectors to original face dimensionality. Selected K eigenfaces must be in the original dimensionality of the face vector space.

Selected K eigenfaces

Step 9: Represent each face image a linear combination of all K eigenvectors. Each face from Training Set can be represented as a weighted sum of the K Eigenfaces + the mean face.

The flowchart for eigen-based face recognition algorithm is shown in figure 3.



Functional Principle of Eigen based face Recognition algorithm

Conclusion

The result of this study and other several statistical research in the area of facial recognition technology deploying the Principal Component Analysis algorithm have revealed the significance of using this algorithm for identifying and verifying facial features. Some of the findings which include reduction of noise since the maximum variation; smaller database representation given that only the trainee images are stored in the form of their projections on a reduced form.

References

Abdi, H. and Williams, L.J. (2010).Principal Componet Analysis. Wiley Interdisciplinary Reviews: Computational Statistics. 2 (4): 433–459.

Agrawal, A. S., Gupta B. I, Maurya A. R, and Phadnis H. B. (2017). Webcam Based Facial Recognition Using OTP Verification for Voting System. International Journal for Research in Engineering Application & Management (IJREAM). Volume 03, Issue 03.

Alabi A. A., Akanbi L. A. and Ibrahim A. A. (2015). Performance Evaluation of the Eigenface Algorithm on Plain-Feature Images in Comparison with Those of Distinct Features. American Journal of Signal Processing, Volume 5 Issue 2, pp. 32-39.

ASHOK, J. and RAJAN, E. G. (2010). Principal Component Analysis Based Image Recognition. International Journal of Computer Science and Information Technologies, Volume 1 Issue 2, pp. 44-50.

Gunturk, B.K.., Batur, A.U. and Altunbasak, Y. (2003). Eigenface-domain super-resolution for face recognition. IEEE Transactions of . Image Processing. vol.12, no.5.pp. 597-606.

Jafri, R. and Hamid R. A. (2009). A survey of face recognition techniques. JIPS, Volume 5, Issue 2 pp. 41-68.

Jolliffe, I.T. (2002). Principal Component Analysis, Series: Springer Series in Statistics, 2ed, Springer, NY, 2002, XXIX.

Joseph, J. and Zacharia, K. P. (2013). Automatic Attendance Management System Using Face Recognition. International Journal of Science and Research (IJSR). Volume 2 Issue 11.

Kar, N., Debbarma, M. K., Saha, A. and Pal, D. R. (2012), Study of Implementing Automated Attendance System Using Face Recognition Technique. International Journal of Computer and Communication Engineering, Vol. 1, No. 2, July 2012.

Khan, H. H., and Khiyal, M. S. H. (2017). Side-View Face Detection Using Automatic Landmarks.Journal of Multidisciplinary Engineering Science Studies (JMESS). Vol. 3 Issue 5.

Latha, P., Ganesan, L. and Anadurai, S. (). Face Recognition using Neural Networks. Signal Processing: An International Journal (SPIJ) Volume (3) : Issue (5).

Parmar, D. N. and Mehta, B. B. (2013). Face Recognition Methods & Applications. International .Journal Computer Technology & Applications, Volume 4 Issue 1.

Paul, L. C. and Sumam, A. A. (2012). Face Recognition Using Principal Component Analysis Method. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 9.

Piccolino, G. (2015). What other African Elections tell us about Nigeria‟s Bet on Biometrics. Available at www.washingtonpost.com/blogs/monkey-cage/wp/2015/03/10/what-otherafrican elections-tell-us-about-nigerias-bet-on-biometrics/. Accessed 18, June.

Smith, L. I. (2002). A tutorial on Principal Components Analysis. A tutorial on Principal Components Analysis.

Thakare, N., Shrivastava, M. and Kumari, M. (2016). Face Detection And Recognition For Automatic Attendance System *(IJCSMC) Volume. 5, Issue. 4.*

Turk, M. and Pentland, A. (1991). Face Recognition using Eigenfaces. Conference on Computer Vision and Pattern Recognition, 3 – 6 June 1991, Maui, HI , USA, pp. 586 – 591.

Vyas, R. A. (2016).Feature Extraction Technique of PCA for Face Recognition With Accuracy Enhancement. International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 4 Issue: 11.