Real Time Face Recognition System based on Pattern Matching

Temilola M. Adepoju[1], Adebanjo Adekiigbe[2], Adedayo Sobowale[1] and Blessing S. Akinola[1]

[1]Department of Computer Engineering Technology, Federal Polytechnic Ede, Osun State Nigeria
[2] Department of Computer Science, Federal Polytechnic Ede, Osun State Nigeria
*atemilola@gmail.com

Abstract—*Face recognition is an interesting area of research today for the fact that face plays an important role in security of life and properties. Research has been done in areas of security such as biometric using finger print, still manipulation or hacking of biometric information is possible. Face recognition system can be a profound solution to these problems (manipulation and hacking) at any sensitive point such as airport, and access control at corporate environment. Therefore, this study developed a face recognition system based on pattern matching for access control. The input images were enhaced using histogram equalization while Principal Component Analysis (PCA) algorithm was employed to extract face features. The extracted features of the registered image from the database were used to compare the real time image using minimum Euclidean Distance (ED). The developed system was implemented on white and black faces. The white faces were obtained from exiting database (OUR database for face detection) while the black faces were acquired locally with a digital camera, to populate a database which was later used to compare real time faces. The recognition accuracy of the developed system with the OUR database and locally populated database are94% and 98% respectively. This show that for any real time scenario the system can recognize human faces correctly without mixed-up. Therefore security of any point where thhe system is being implemented is 98% guaranteed.*

Keyword—Accuracy, Computational Time; Biometric; face pattern matching; Principal Component Analysis (PCA); Euclidean Distance (ED

1 INTRODUCTION

Given the requirement for determining people's identity, the obvious question is what technology is best suited to supply this information? (Madalina and D'Amico, 2004). There are many ways that humans can identify each other, and so is for machines. There are many different identification technologies available, many of which have been in commercial use for years. The most common person verification and identification methods today are Password/PIN known as Personal Identification Number, systems (Madalina and D'Amico, 2004). The problem with this or other similar techniques is that they are not unique, and is possible for someone to forget, loose or even have it stolen to or by someone else.

In order to overcome these problems considerable interest was given to “biometrics” identification systems, which uses pattern recognition techniques to identify people based on their characteristics (Omidiora 2013). Although traditional biometric methods of identification such as fingerprints, Iris scans and voice recognition are viable, they are not always the best, depending on where they will be used (Vanaja, Waghmare, and Chirchi, 2011). In applications such as Surveillance and monitoring of public places for instance, such methods would end up failing because they are time consuming and inefficient especially in situations where there are many people involved. Also, the cost of implementation is also a hindrance as some components might have to be imported. Biometric authentication is a profound way to deal with these difficulties rather than using the traditional password systems. Potentially, biometric systems can be employed in all applications that need authentication mechanism, and so in all applications that today use passwords, PINs, ID cards, or the like (Madalina and D'Amico, 2004).

Biometrics is the science of automated recognition of persons based on one or more physiological or behavioral characteristics possessed by such individuals (Vanaja, Waghmare, and Chirchi, 2011). These physiological characteristics include face, fingerprints, iris, retinal features, hand geometry, and ears. Behavioral characteristics include handwritten signature, voice, keystrokes (typing), and gait (how a person walks). The current state of technology presents several uses of the biometrics and each has its own advantages and disadvantages according to the requirements on biometric identifiers.

A practical biometric system should have an acceptable recognition accuracy, speed with reasonable resource requirements. It should be harmless to users, be accepted by the intended population, and be sufficiently robust to various fraudulent methods. Over the years, fingerprints have been one of the most widely used and accepted biometric (Vanaja, Waghmare, and Chirchi, 2011). This is evidenced with the Chinese who have used fingerprints to sign documents for over 1000 years (Max, 2002)

However, human face plays an important role in our social interaction. Using the human face as a key to security, biometric face recognition technology is of significant attention in this research, as it presents a wide range of applications in Airport Surveillance, Private Surveillance, Access Control for PCs in a Corporate Environment, Added Security for ATM Transactions and law enforcement. As compared with other biometrics systems such as fingerprint or palmprint and iris, face recognition has distinct advantages because of its non-contact requirement process (Max, 2002). Face images can be captured from a distance without physical contact the person being identified, and the identification does not require interacting with the person.

In addition, face recognition serves the crime deterrent purpose because face images that have been recorded and archived can later help identify a person. Consequent to the advent of various high-tech devices and ideas, hacking has become increasingly high and as a result several security devices can be hacked or guessed by others thereby making individuals private and personal information vulnerable.

Biometrics such as fingerprints provides an alternative method but can be forged (gummy fingers). Hand geometry is not distinctive enough to be used in large scale applications likewise hand-written signatures can be forged (Vanaja, Waghmare, and Chirchi, 2011). However, face is unique among all humans. Thus, there is a need for an affordable and mobile system similar to human eye to identify a person.

2 LITERATUR REVIEW

2.1 Principal Component Analysis (PCA)

PCA is a well - known technique in statistics. It picks out patterns in variables. As a result, patterns rather than all data are studied. It reduces the dimensionality of data without a significant loss of information. It can be used for prediction, redundancy removal, feature extraction, data compression, and so on. PCA is a suitable tool for feature extraction because it identifies the variation between human faces, which may not be immediately obvious (Omidiora, 2006). PCA does not attempt to classify faces using familiar geometrical differences, such as length of the nose or the width of eyebrow but a set of human faces is analysed using PCA to determine which variables account for the variation in the face images (Omidiora, 2013). In face feature, these variables are called eigenfaces because when plotted they display a ghostly resemblance to human faces. Eigenface deals with projecting an image space linearly to a low dimensional feature space such that recognition of the faces is done in this low dimensional feature space (Cendrillon and Lovell, 2000). Mathematically, eigenface adopts the PCA. The eigenface finds the principal components of the distribution of faces or the eigenvectors of the covariance matrix of the set of face images. According to the significance of the eigenvalues, the eigenvectors are ordered and each one accounting for a different amount of the variations among the face images. These eigenvectors can be considered as a set of features that together characterise the variations between face images.

PCA Method Steps are as follows (Cendrillon and Lovell, 2000).

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

$$average = \frac{1}{M}\sum_{n=1}^{M} Training\ \text{Im}ages \quad (1)$$

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

$$Sub = Training\text{Im}ages - Average \quad (2)$$

- Calculate the Covariance Matrix as shown in the equation bellow:

$$Co\text{var}iance = \sum_{n=1}^{M} Sub(n)Sub^{T}(n) \quad (3)$$

- Calculate the Eigenvalues and Eigenvectors of the Covariance Matrix.
- 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.
- Project the training samples onto Eigenfaces.

2.2 Euclidean Distance

Euclidean distance or Euclidean metric is the "ordinary" straight-line distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space. The Euclidean distance between points p and q is the length of the line segment connecting them ($\overline{pq}$). In Cartesian coordinates, if $p = (p_1, p_2, \dots\dots p_n)$ and $q = (q_1, q_2, \dots\dots q_n)$ are two points in Euclidean *n*-space, then the distance (d) from p to q, or from q to p is given by the Pythagorean formula: The position of a point in a Euclidean *n* space is a Euclidean vector. So, p and q are Euclidean vectors, starting from the origin of the space, and their tips indicate two points. The Euclidean norm, or Euclidean length, or magnitude of a vector measures the length of the vector as seen in Equation 4

$$\|p\| = \sqrt{p_1^2 + p_2^2 + \dots\dots + p_n^2} = \sqrt{p.p} \quad (4)$$

The distance between any two points on the real line is the absolute value of the numerical difference of their coordinates. It is common to identify the name of a point with its Cartesian coordinate. In one dimension, there is a single homogeneous, translation-invariant metric (in other words, a distance that is induced by a norm), up to a scale factor of length, which is the Euclidean distance.

2.3 Review of Related Work on Face Recognition

Cohen, Nicu, Larry, Ashutosh, and Thomas (2003) developed a facial expression recognition from video sequences in their work, used dynamic classifiers, such as Hidden Markov Models. This method is proposed for person-dependent systems, as it is more sensitive to temporal pattern changes, in the case of videos. Studies by the same authors also recommend using static classifiers, such as Tree Augmented Naive Bayes, for person-independent scenarios.

Shamla and Kalpana (2011) proposed a label Self-Organizing Map (SOM) to measure image similarity. To achieve their set goal, they fed Facial images associated to the regions of interest into the neural network. At the end of the learning step, each neural unit is turned to a particular Facial image prototype. Facial recognition is then performed by a probabilistic decision rule. This scheme offers them very promising results for face identification when dealing with illumination variation, facial poses and expressions. The paper presented a novel Self-Organizing Map (SOM) for face recognition. Eventually, they had good features extracted due to SOM's topological ordering. The Facial Analytics results for the 400 images of AT&T database used reflects that the face recognition rate using one of the neural network algorithms SOM is 92.40% for 40 persons.

Also, Deepesh Raj (2011), developed an effective and real time face recognition system based on OpenCV and C++. The system uses Principal Component analysis for feature extraction and various distance classifiers such as the Euclidean distance, the Manhattan distance and the Mahalanobis distance. The technique used here involves generating the `eigen faces' then projecting training data into face-space to be used with a PCA classification method and evaluation of a projected test element by projecting it into face space and comparing to training data. The system was tested on YALE Face database B and ORL Face Database. The recognition produced using different matching techniques are compared and the results were been presented. The correct recognition rate achieved using the Mahalanobis distance is 92.3% and 73.1% for the normal PCA with Euclidean distance.

Ojo, J. A. and Adeniran, S. A. (2011) employed two dimensional discrete wavelet transform (2D DWT) to extract features from face images; and hidden Markov model (HMM) for training, recognition and classification. The algorithms were tested with a subset of the AT&T database. The recognized images were verified for correctness, 80% correct classification (Hit) occurred while 20% were misclassified. The rest of the images that were not in the training set were used to test the false acceptance rate (FAR) i.e. the ratio of the numbers of images falsely accepted to the total number of images tested and 0.02 FAR occurred.

In this paper, PCA will be used to extract specific features from acquired images and Eucleadean distance to classify and obtain similarity of the images.

3 METHODOLOGY

The complete framework for the face recognition which involves feature extraction and pattern matching is presented in Figure 1.

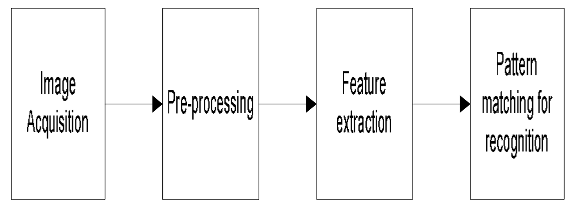

Figure 1: Framework of the developed system

The first step is the acquisition of black face images using conditions such as facial pose, angle and expression. The pre-processing stage was performed using histogram equalization while the extraction of specific features was crried out using PCA as stated in equations (1) to (3). Based on the extracted features, mean Euclidean distance was used for identification of few black faces in real time following equation (4). Figure 2 shows the step by step procedure to achieve the developed system.

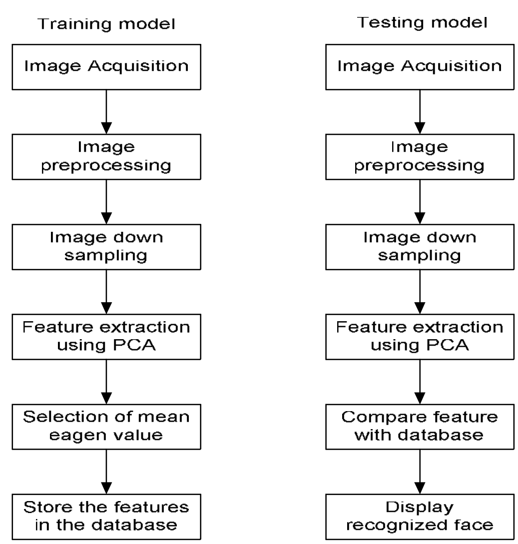

Figure 2: Developed face recognition model

Given a set of N face images (training images) to be tested or recognized, feature extraction is first performed. The eigenface from the training set is calculated and only the M images that correspond to the highest eigenvalues is kept. These M images define the “facespace”. As new faces are encountered, the “eigenfaces” can be updated or recalculated accordingly. The corresponding distribution in M dimensional weight space for each known individual is calculated by projecting their face images onto the “face space”. Determining whether the image corresponds with face in the database or not by checking the closeness of the image to the face space and the face with most likelihood is identified to be the one representing the face. Euclidean distance was used to test if the face is in the training set or database. If the likelihood is within the stated distance, the face is recognized to be in the training set or in the database. The testing of the system was carried out such that real time face images were used to evaluate the performance with respect to sensitivity and accuracy using, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) as performance indices. The study was developed in a MATLAB 8.0.0 (R2013a) environment.

Sample of white and black faces for the training ang testing of the system is presented in Figure 3(a) and Figure 3(b) respectively.

The following metrics are used for analyzing and evaluating the recognition accuracy of the system;

i. True Positive (TP) when a trained face is correctly classified with the same face on database,
ii. False positive (FP) when a trained face is wrongly matched with another in the database,
iii. True Negative (TN) when a face that does not exist in the database is not recognized.
iv. False Negative (FN) when an untrained face is not identified by the system as existing in the database
v. The recognition accuracy is obtained by using the equation 5

$$\text{Recognition} \quad accuracy = \frac{TP + TN}{Total \ \text{Image}}$$

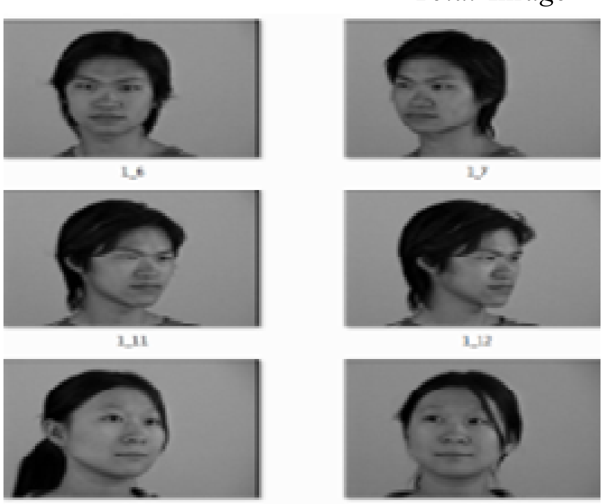

Figure 3(a): Sample of white faces from OUR database

Figure 3(b): Samples of locally acquiredblack face

4 RESULT DISCUSSION

The recognition system is used to detect or reject the identity that a user claims. Figure 4 depicts the interface of the MATLAB application and its response after an individual has been verified by the system. Several faces of different persons were trained by the system and subsequently real time faces of such persons were used to verify the system’s ability to recognize the face-image of the same person taken under various conditions. As seen in Figure 4, description of the system after the face image has been called in for the verification process was shown. The real-time face was used to verify the individual and as seen in the Figure 4.2 the person was granted access because a face whose Euclidean distance falls in range has been detected. However, Figure 5 depicts a face image that was wrongly classified by the system. It is worthy to be noted that the threshold value determines how strong the system is in rejecting false values.

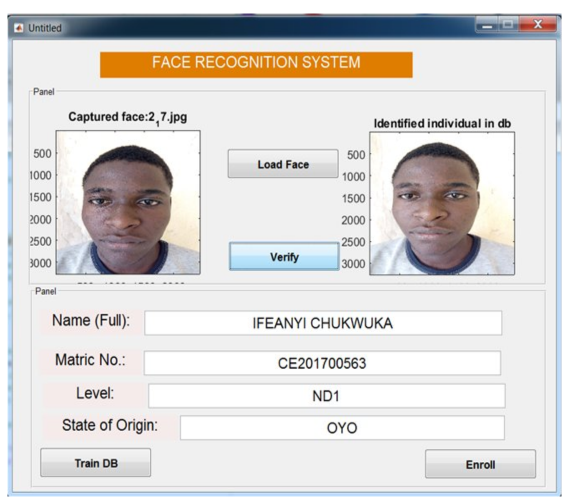

Figure 4: Correctly recognized Face

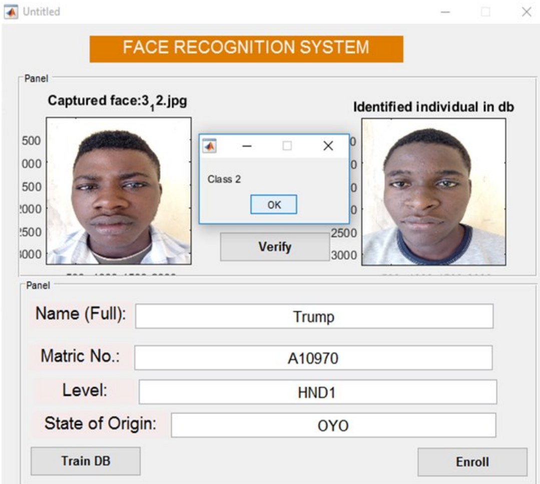

Figure 5: Wrongly recognisedf Face

Summary of the developed system performance evaluation is presented in Table 1. Fifty (50) faces each were used for testing the system from the two database. For white face database forty-seven (47) were correctly recognized and three (3) were wrongly recognized. The system was implemeted on locally acquired balck faces and forty-nine (49) faces were correctly recognized while one (1) is wrongly recognized. It was observed that the developed system achieved 94% and 98% recognition accuracy with white face databse and locally acquired black face database respectively.

Table 1: Performance Summary of the Developed System

Metrics	**Results (Locally Acquired Database)**	**Results (OUR Database)**
Image size (Pixels)	50 × 50	50 × 50
True Positive (TP)	49	47
True Negative (TN)	0	0
False Positive (FP)	1	2
False Negative (FN)	0	1
Correctly Recognized	49	47
Incorrectly Recognized	1	3
Recognition Accuracy (%)	98	94

Some of the existing algorithms adopted by different researchers that revealed the year and results obtained are presented in Table 2. The results were compared with the developed system result and it can be deduced that the system performed better.

Table 2: Comparison of existing results with developed system result

S/N	Authour	Algorithm Used	Result Obtained
1	Maha (2005)	2D Discrete Cosine Transform (2D-DCT) and Principal Component Analysis (PCA)	96.5%.
2	Jawad, Syed, and Farrukh (2008)	Neural Networks	81.36%
3	Surya and Pritee (2012)	Discrete Cosine Transform and Nearest Neighbor Discriminant Analysis	98.5%
4	Zahraddeen , Fatma, Abdulganiyu and Mustafa (2016)	Discrete Cosine Transform	93.4%
5	**Developed Real Time Face Recognition System based on Pattern Matching**	**Principal Component Analysis and Euclidean Distance**	**94% (OUR database) 98% (Local database)**

5 CONCLUSION

The results obtained by implementing the the pattern matching face recognition system on both white and black faces show that the system is reliable and can be adopted at any access control point. The pattern matching face recognition system can also be used to improve security in term of verification.

REFERENCES

[1] J. F. Hair, G. T. Hult, and M. C. M. Ringle, *A primer on partial least squares structural equation modeling*, Second Edi. Los Angeles: SAGE Publications, Inc., 2017.

[2] S. A. O. S.A. Aduloju, A.L. Awoponle, “Recapitalization, mergers, and acquisitions of the Nigerian insurance industry,” *J. Risk Financ.*, vol. 9, no. 5, pp. 449–466, 2008.

[3] R. H. Mcguckin, “On productivity and plant ownership change : New evidence from the longitudinal research database,” *RAND J. Econ.*, vol.

26, no. 2, pp. 257–276, 1995.
[4] B. E. H. A. N. Kim and V. Singal, “Mergers and Market Power : Evidence from the airline industry,” *Am. Econ. Assoc.*, vol. 83, no. 3, pp. 549–569, 1993.
[5] T. C. and M. J. Piskorski, “Power imbalance , mutual dependence , and constraint absorption : A closer look at resource dependence theory,” *Sage Publ. Inc.*, vol. 50, no. 2, pp. 167–199, 2005.
[6] J. Pfeffer, “Size and composition of corporate boards of directors: The organization and its environment,” *Sage Publ. Inc.*, vol. 17, no. 2, pp. 218–228, 1972.
[7] O. Williamson, *The economic institutions of capitalism*, First. New York: The Free Press, 1985.
[8] David R. King and R. J. S. and I. Kesner, “Intensive firms performance implications of firm resource interactions in the acquisition of R & D-intensive firms,” *INFORMS*, vol. 19, no. 2, pp. 327–340, 2008.
[9] J. P. Antony and S. Bhattacharyya, “Measuring organizational performance and organizational excellence of SMEs – Part 1: A conceptual framework,” *Meas. Bus. Excell.*, vol. 14, no. 2, pp. 3–11, 2010.
[10] J. Hales, *Accounting and financial analysis in the hospitality industry*. New York: Butterworth-Heinemann, 2005.
[11] A. AbdulRasheed, I. S. Babaita, and M. A. Yinusa, “Fraud and its implication for bank performance in Nigeria,” *Int. J. Asian Soc. Sci.*, vol. 2, no. 4, pp. 382–387, 2012.
[12] H. Fauzi, “The Performance Implications of Fit among Environment , Strategy , Structure , Control System and Social Performance,” *Issues Soc. Environ. Account.*, vol. 3, no. 2, pp. 117–142, 2010.
[13] R. Simons, *Performance Measurement and Control Systems for Implementing Strategy Text and Cases*, Fourth. New Jersey, USA: Pearson education limited, 2013.
[14] Z. Hoque, “A contingency model of the association between strategy, environmental uncertainty and performance measurement: Impact on organizational performance,” *Int. Bus. Rev.*, vol. 13, no. 4, pp. 485–502, 2004.
[15] J. F. Henri, “Performance measurement and organizational effectiveness: Bridging the gap,” *Manag. Financ.*, vol. 30, no. 6, pp. 93–123, 2004.
[16] A. Neely, *Business Performance Measurement; Unifying theories and integrating practice*, Second. New York: Cambridge University Press, 2007.