ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 11(10), DOI: 10.17485/ijst/2018/v11i10/121126, March 2018*

Kinect User Authentication using Multimodal Biometric

A. Saravana Priya[1] and Rajeshwari Mukesh[2]

[1]Department of Computer Science and Engineering St. Peters University, Avadi – 600054, Chennai, India; priyaayyavu@yahoo.co.in
[2]Department of Computer Science and Engineering Hindustan Institute of Technology and Science, Padur, Kelambalam - 603103, Chennai, India; rajeswarim@hindustanuniv.ac.in

Abstract

Objectives: Facial Recognition system has been the most popular non-intrusive authentication system. The proposed system offers to improve the accuracy of facial recognition using multimodal authentication with a 3d depth sensor. **Methods/Statistical Analysis**: To increase the reliability and robust of the system, we superimpose skeletal biometric features with facial features where data combined using 3D Imaging Technology. Statistical approaches are used for feature transformation using PCA and LDA Algorithm. **Findings**: To evaluate the system, we collected nearly 1000 individuals with two samples for each, captured in indoor environments under various light effects. The accuracy of the system is built with 95% confidence interval and based on a classifier, three performance curves Receiver Operating Characteristics (ROC), Cumulative Match Score Curves (CMC) and Expected Performance Curves (EPC) are used for performance evaluation. The FRR of the proposed multimodal fusion approach is 0.12 at 0.001 minimal FAR with 95% accuracy. **Applications/Improvements**: Multimodal biometric is required for strong authentication for the surveillance of smart home, cities and health care sectors where security plays a vital role. This fusion technique improves the reliability, latency, and accuracy of the system in recognizing and authenticating the user.

Keywords: 3D Imaging Technology, Face Features, Multimodal Authentication, Skeletal Biometric

1. Introduction

Any home, corporate sector or government has a great responsibility to secure with advanced software and hardware to protect their infrastructure from malicious attack. Existing technology has the ability to protect our system from the intruders. Many researchers and research institutes are developing strong and confidential security systems. The person who accesses the security information has to be authenticated beforehand by taking adequate security measures. Person authentication is done by providing identity to each of them using several authentication mechanisms. Biometric authentication has been widely regarded as the most fool proof - or at least the hardest to forge or spoof[1,2]. The main functioning of Biometric Recognition System involves two parts; enrolment and test. During the enrollment process, the template is stored in a database. And during the test process, the individual's data is compared with the acquired templates that are stored in the database[3]. The Matching program evaluates template with input, estimating the distance between these two using suitable algorithms. This is considered as the output for a specified purpose. Biometric identifiers quantifiable attributes used to portray the person and label their features[4]. Physiological characteristics are related to the shape of the body. Face

*Author for correspondence

recognition, body tracking, voice recognition, iris scans, finger prints, finger pressure plays an important role in authenticating the person. One of the most popular 3-D camera and a biometric tool is the Microsoft Kinect Sensor for Windows. Microsoft Kinect Sensor for Windows gets you the latest in human computing technologies, empowering the development of applications that are built to interact humans with the computers using natural gestures and speech[5]. In the proposed biometric system, more than 1000 real-time data are collected for user authentication. Dataset of facial features and Anthropometric dataset are acquired using the Kinect sensor. The system uses pattern analysis methods, to relate similar segments as well as to distinguish the segments which show unique identities among the individuals. Original images are of high dimensional data that has to be reduced to low dimensions, to visualize for analysis. The PCA analysis is used to reduce multi-co linearity and distinguish the biometric features in the set. Face feature vector and skeleton feature vector are partitioned into training images and testing images. The template of the skeletal features is built using training data and evaluation of the data by using test data. The evaluation is done by using logistic regression classification technique where the data are statistically analyzed[6]. This fusion of techniques validates the user as genuine or an imposter, by comparing the original biometric data with the templates stored in the database. The accuracy of the combined user recognition system is improved by a lot factor than the accuracy of the facial recognition system alone.

2. Existing Work

Many researchers have come up with user identification and authentication using the Kinect sensor. The static and dynamic data are captured and derived with appropriate solution. For example, Fusiello[7] used a new person identification method that uses motion and anthropometric biometric acquired from Kinect. System has EER of 13% and an average CMC Rank-1 identification rate of 90%. Anderson and Araujo[8] used person identification using anthropometric and gait data from Kinect sensor, performed data analysis by applying KNN, SVM, and MLP with 87.7% accuracy. Elena[9] Human Classification Using Gait Features are done by physical and behavioral features aiming at identifying the more relevant parameters for gait description for people recognition in 3D Model. Changjun[10] used PCA and logistic regression classifier for face recognition with an accuracy rate of 93.33%. Naseer[11] used Linear Regression Classification Algorithm for Face recognition and for continuous occlusion in the image rectification using distance based Evidence Fusion. Accuracy of the biometric system is achieved to 96% in Multimodal user authentication using Kinect sensor[12]. Luca[13] used different fusion technique for face and fingerprint fusion at match score level resulting FAR at 0% and FRR at 0.6%. Marina[14] used Face, Ear and signature fusion at rank score level using logistic regression with EER =1.12%. EIJI Hayashi[15] used SVM classifier to classify the Body length and gestures resulting in 1.6% of Equal Error Rate. Chaudhary[16] integrated multimodal biometric system, Face, Iris and voice using sum rule techniques. Fusion did at the score level resulting in 92% accuracy and 2% FAR. Suman[17] used Linear Discriminate Analysis Algorithm for Face Recognition performed with ORL face database. This LDA method overcomes the PCA limitation in multiple classifications. Shih_Ming[18] proposed Kernel Linear Regression classification Algorithm for Face Recognition at low resolution. This Kernel projection helps in minimizing the reconstruction Error under variable illumination.

3. Proposed Methodology – Kinect Study

The latest Microsoft Kinect v2 camera detector and the SDK 2.0 take natural user interactions with computers to the next degree. It accelerates the development of applications that react to movement, gesture, and voice offering greater overall precision, reactivity, and intuitive capabilities. Various applications are built using Microsoft Kinect for Windows in places like smart home, commercial sectors and in healthcare finding new ways for better human life. Kinect uses low-cost infrared pro-

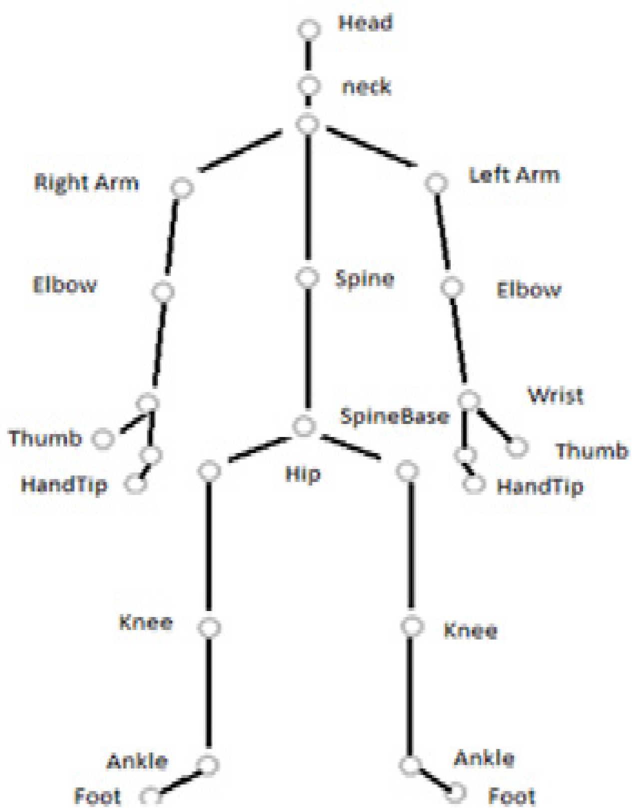

Figure 1. Skeletal Joints and Skeletal Dimensions.

jector and depth-sensing camera for human detection. Microsoft developed Kinect Sensor Camera and Kinect Software Development Kit (SDK) which can be applied to build an Intrusion detector using camera capturing the user in three-dimensional [3D] space. 3D images can be represented as the range of images that are robust to color changes in the object and illumination variations[19]. Sensor, at Home; it replaces the watch dog activity by sensing the unidentified person visiting home. Initially, the Kinect can recognize the human whoever stands in the field view of the camera rather than any object are set to ready for tracking the user. Depth data plays a vital role in Human Recognition that project the tracking object or image and sets a boundary from the background interference. In Figure 1, Kinect skeleton tracking starts from center of the head to foot, by identifying their skeleton joints or dimensions where it tracks 25 skeletal joint points all over the body. Kinect acts as a security the camera to detect the person when a human appears in front of camera and captures the skeleton as a biometric trait. First, Face Tracking SDK contains a face tracking engine, which can analyze the input from the Kinect camera, it can detect the head pose and face features like eyes, nose, mouth, cheeks, and jawbones depending on the points that can be tracked, and generate an information to the application in real time[20]. Second, anthropometric measurement is taken from each individual and registered in the database. Generally, human features show minimum variations in each and almost match with body segment of others. Skeleton data are not consistent and shows a significant increase in errors as sample size increases. This integrated system is built and implemented to the clients in an unobtrusive way which talks about the new approach for authentication.

4. Proposed System using Biometrics

4.1 Image Acquisition and Preprocessing

The samples are collected from various places and selected for analysis. The data are collected in an unobtrusive way where respondent find easy with sensor tracking. Data is collected with a variety of samples containing persons of different age, gender, tall person, short person, bald head person, including the different clothing style. Using Kinect depth sensor, the image captured under any multiple light conditions. The image of two traits (Skeleton body and Face) is acquired from Kinect sensor where the distance of the person from Kinect sensor is 5 feet to 6 feet depending on the person's height. Raw skeleton data and Face feature data produce background noise at the boundary edges. Tracking with the sensor is improved by applying extended Kalman filter to 3D data[21]. This filter smoothens out the jitter, adds the capability to continue tracking for a short period of time when the object moves out of range of the sensor. The Kalman filter is an efficient recursive filter that controls the noise data by capturing the skeleton features and face features in Kinect.

4.2 Feature Extraction

Feature extraction plays an important role to discriminate individuals from their surroundings. The system extracts skeleton feature joints and Face Features points separately. All the skeleton joints and Face features points are captured in three dimensional (3D) spaces. In X, Y, Z Space co-ordinates X and Y are the distance between the joints and where Z is the distance from the sensor or depth value. A line can be drawn to get the human skeleton picture in 3D space.

$$Đ=\sqrt{(x1-x2)^2+(y1-y2)^2+(z1-z2)^2} \quad (1)$$

Where distance d is measured for skeletal features points, calculates the segment length to distinguish the head, shoulder, hand length, hip width and leg length as a whole body. Similarly, face features are taken as input and the distance between the feature points are segmented and distance between the segments are calculated to distinguish eyes, nose, cheek, jaws, and mouth in one form. The distance between the feature points can be computed using the Euclidean norm (i.e.) by finding the square root of the sum of squared differences of the coordinates.

5. Analyzing the 3D Space Data

5.1 Principal Component Analysis

The Skeleton Image and Face Image are the high dimensional dataset and difficult to visualize the data. PCA is the distinct linear feature extraction method where new features are constructed using the principal components and compress the high dimensional training sample data into total scatter matrix as a set containing the sum of the within-scatter matrix and between-scatter matrix. In the equation, say, point, represent a 3-dimensional pixel with co-ordinates () be p(). Also co-ordinates is bound by space such that and it also represents face feature = 1, 2, 3 ... F. Segment S, represented as a distance () between any two features and deduced by Euclidean Theorem. Now the data (w) can be represented by a set of distance of several segments or classes () of = 1, 2, 3 ..., C[10,22,23].

$$w=\sum_{i=0}^{F}\sum_{i=0}^{C} d_{ij} \quad (2)$$

Now for sample data of k = 1, 2… N can be represented as,

$$X=\sum_{k=0}^{N} w_k \quad (3)$$

Where is the set of all segment lengths in a given image.between-class scatter matrix and within-class scatter matrix can be calculated as,

$$S_b=\sum_{i=0}^{C} P_{m_i}(x_i-x_0)(x_i-x_0)^T \quad (4)$$

Where the average is vector of the class and is the overall average vector of the training sample.is the priori probability of class

Then within-class matrix,

$$S_w=\sum_{i=0}^{c} P_{m_i} Z_i \quad (5)$$

Then covariance matrix calculated as

$$Z_i=E\{(X-x_i)(X-x_i)^T|X\epsilon w_i\} \quad (6)$$

Total scatter matrix gives the scattered features in the feature space,

$$S_t=S_b+S_w=\frac{1}{N}\sum_{j=1}^{N}(X_j-x_0)(X_j-x_0)^T \quad (7)$$

Optimal projection matrix is obtained from the determinant of total scatter matrix which shows largest Eigen values and of the matrix is associated with Eigenvectors.

6. Classifying the 3D Space Data

6.1 Logistic Regression Analysis

In the simplest form, Logistic Regression can be said as a form of Linear Regression with the addition of a discrete mapping function layer operated on the data set mapping features. Typically, LR is used for predicting the probability. But it can also be used for classification. It can be used to classify the data in discrete values on a category axis which can be an outcome of probabilistic function[10, 24]. The Logistic regression curve is shown in Figure 2.

Consider N, number of images in the training dataset, with each of them representing multiple classes, C. Then the entire matrix can be represented as

$$X_i = \sum_{j=1}^{c} w_j \tag{8}$$

Where i = 1, 2 ... N

Now consider learning function P(Y|X) where Y is a set of discrete values such that Y = []. In this paper, we are proposing this function be limited to Boolean values 1 and 0 for classification purposes

Hence we can rewrite as

$$P(Y = 1|X) = \frac{1}{1 + \exp(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)} \tag{9}$$

$$P(Y = 0|X) = \frac{\exp(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)}{1 + \exp(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)} \tag{10}$$

By probability outcome, sum of (9) and (10) is equal to 1

The expression P(Y|X) can be used to classify datasets. Any value of X can generally be classified based on such that it maximizes. In other words,

$$1 < \frac{P(Y = 0|X)}{P(Y = 1|X)} \tag{11}$$

This further leads to

$$1 < \exp(\beta_0 + \sum_{i=1}^{n} \beta_i X_i) \tag{12}$$

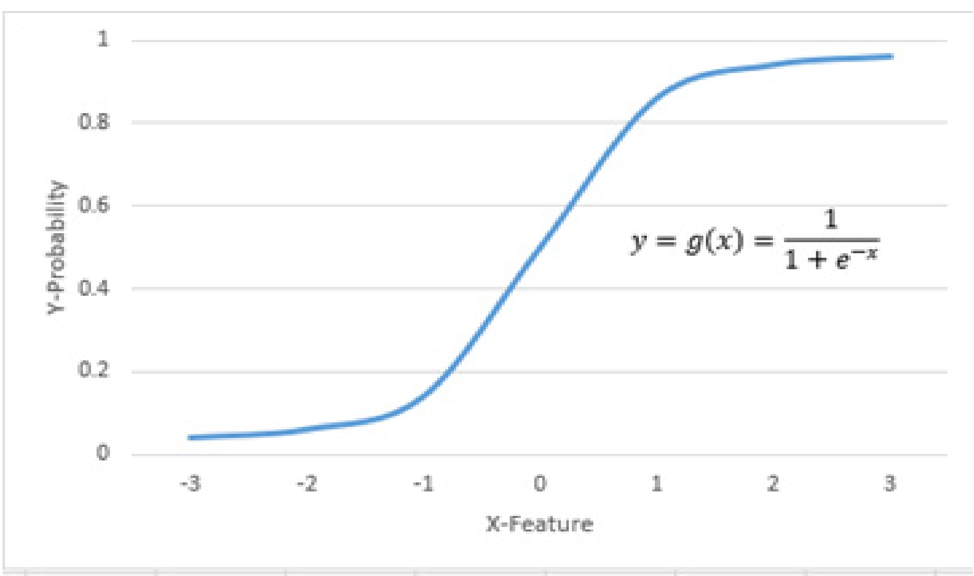

Figure 2. Logistic Regression Curve.

7. Experimental Results

The proposed system is experimented to check the efficiency and reliability of the system by training samples and test data. The individuals are taken as training samples having two sets of data containing face feature points and skeleton feature points each in a separate database. These 3D face Images are initially collected by extracting facial features and its relative distance of feature points to form a face pattern. Similarly, skeleton Images are collected by extracting skeletal joints and distance calculated to determine the segment length. In PCA method, Eigen-faces are used for feature extraction where the maximal variance of the feature is determined by maximum attained Eigen values from the eigenvectors. Next, face detection is done by processing the location and its spatial extent within the image. These face images are stored in the database as a training dataset and template created to train the face images using test data under different lightning conditions and orientations. This illumination variations, shades, lightning effects occurs during face detection are normalized using histogram equalization and logarithmic transform methods. Next, logistic regression is used as a classifier on a transformed data which maximize the conditional likelihood of samples. The test user can be recognized by attaining the closest match among the samples. We focused on skeleton tracking system where every user generally distinguished over their peers. The sample data collection of Human body features and Face features are shown in Figure 3.

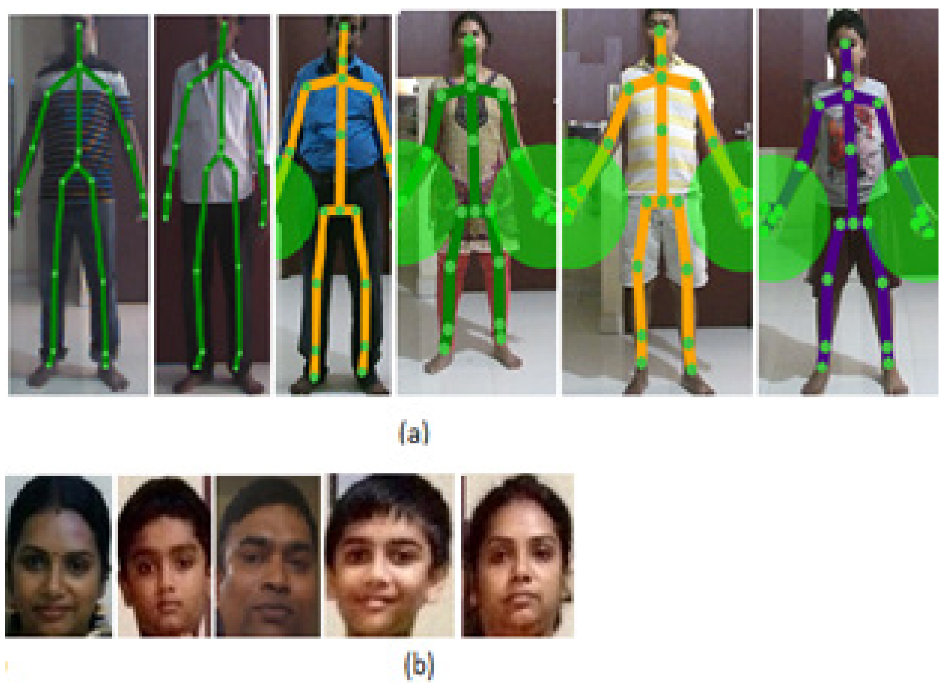

Figure 3. (a) Sample Data of Body based User Authentication (b) Sample Data of User Authentication using Face Features.

8. Performance of the Classifier

To evaluate and compare the proposed biometric authentication system, three performance curves are depicted using PCA and Logistic Regression. The diagnostic test evaluation is done, to show the true positive rate (Sensitivity) against the false positive rate (1-Specificity) using ROC curve. Skeletal feature vector(Sq) and Facial feature vector(Fq) are the inputs taken and compared

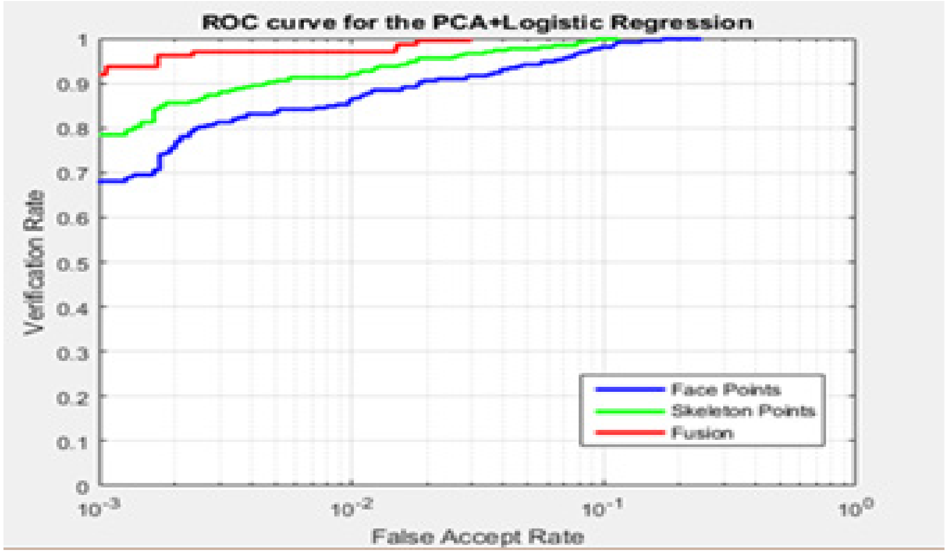

Figure 4. ROC Curves of the proposed authentication system.

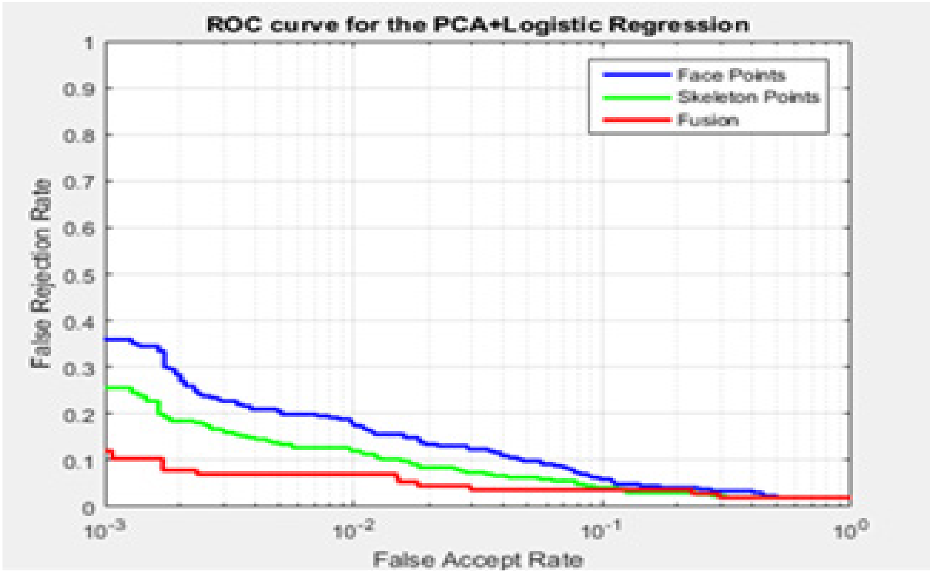

Figure 5. ROC Curves of the proposed authentication system - FAR Vs FRR.

with their trained template(s) St and Ft for the test user M to determine the user passed or rejected in the biometric authentication. In any feature recognition technique, two types of errors are important to consider that shows the robustness of the system. First, the number of imposters incorrectly recognized as a genuine user from the total number of users (Type I Error or FPR). Second, the number of genuine users incorrectly rejected as imposters from the total number of users treated as genuine (Type II Error or FNR). The ROC plot is used to show the performance of the system by comparing actual value and the predicted value. User recognized as genuine for the first time itself indicates 100% accuracy of the system. The accuracy of the system can be evaluated by considering true positive rate. Similarly, Cumulative Match Score Curves (CMC), is the plotted between the rank and Human recognition rate and Expected Performance Curves (EPC) is plotted

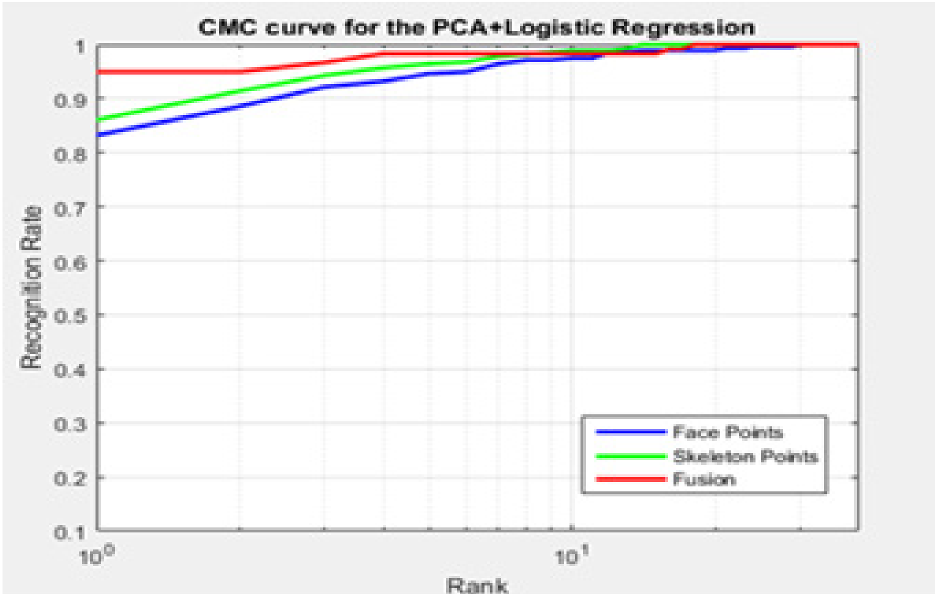

Figure 6. CMC Curves of the proposed multimodal authentication system using PCA + Logistic Regression.

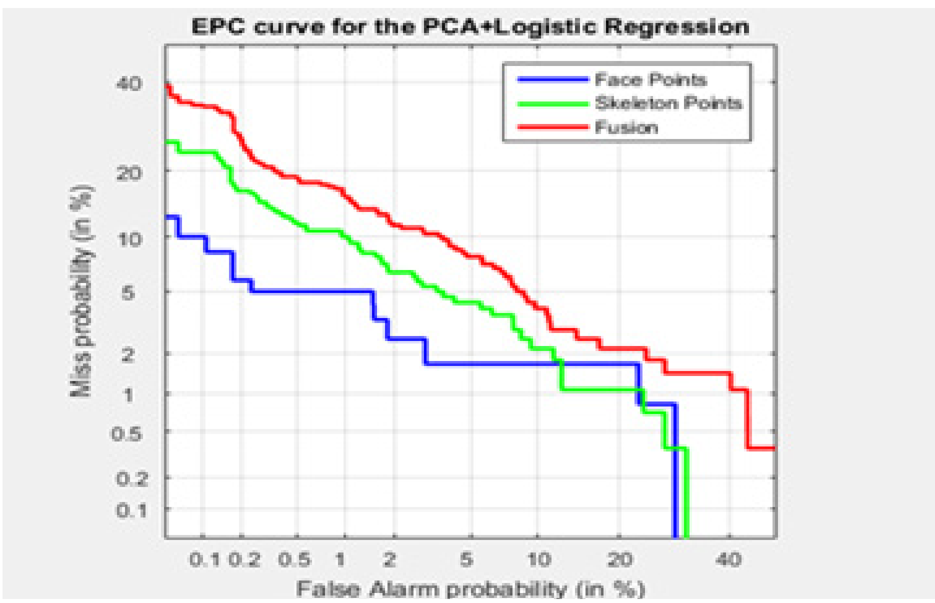

Figure 7. EPC Curves of the proposed multimodal authentication system using PCA + Logistic Regression.

between the alpha and the error rate. The performances of the proposed system are analyzed based on these metrics.

$$\text{TPR}=\frac{\text{TruePositive(n)}}{\text{TruePositive(n)}+\text{FalseNegative(n)}} \quad (13)$$

$$\text{FPR}=\frac{\text{FalsePositive(n)}}{\text{FalsePositive(n)}+\text{TrueNegative(n)}} \quad (14)$$

$$\text{FNR}= 1-\text{TPR} \quad (15)$$

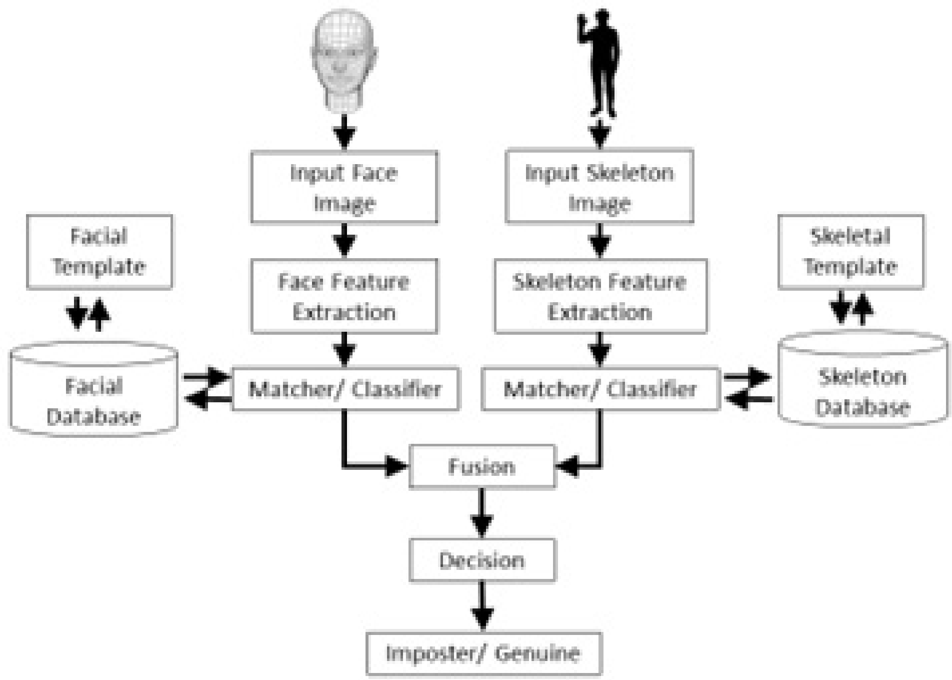

Figure 8. Multimodal biometric authentication using Kinect in proposed system.

It is observed from the Figure 4 that the proposed multimodal biometrics authentication system provides better performance than their individual counter parts such as facial points and skeleton points. While considering the fusion of facial points and skeleton points, the verification rate at 0.001 FAR is 92%. For the same FAR, the facial points and skeleton points produce only 68% and 78% accuracy respectively. It is observed from the Figure 5 that the FRR of the proposed fusion approach is very low in comparison with facial points and skeleton points. The FRR of the proposed multimodal fusion approach is 0.12 at 0.001 FAR. For the same FAR, the obtained FRR of the face and skeleton points are 0.26 and 0.37 respectively. Figure 6 shows the Cumulative Match Score Curves (CMC) based on the data computed using face points, skeleton points, and fusion. The CMC is used as a measure of 1:n identification system performance. It judges the ranking capabilities of an identification system. The CMC curves drawn for the proposed multimodal biometrics authentication system using the fusion of face and skeleton points outperform their individual performance. The recognition accuracy of the system using fusion approach is 95% whereas it is 86% and 83% for facial and skeleton features respectively. Figure 7 shows the Expected Performance Curve (EPC) enables us to analyze the performance of a model to a criterion, decided by the application. The two criteria used to analyze the proposed identification system are miss probability and false alarm probability. Figure 8 shows the EPC curve for the proposed multimodal biometrics authentication system. It is observed from the EPC curve of fusion approach, in Figure 8, that the increase in false alarm probability, the probability of missed genuine user, increases greatly in comparison with facial points and skeleton points. The probability of misses for the proposed fusion approach at 10%, false alarm probability is only 4%. For facial points and skeleton points approach, it is 1.8% and 2.1% respectively. The flow Diagram of Multimodal biometric authentication using Kinect in Proposed System is shown in Figure 8.

9. Conclusion

Microsoft Kinect is less sensitive and less expensive sensor device which provides an accurate solution for recognition and detection problems. To improve the human recognition in Kinect, a novel, unobtrusive fusion technique is applied using PCA and Logistic regression analysis. With this new technology, many applications are built for improved security in various areas like forensic, government, many business sector, and smart home and change it for a better human life; integrating the Microsoft Kinect with the existing authentication mechanism can produce >99% accuracy.

10. References

1. Munse BC, Temlyakov A, Qu C, Wang S. Heidelberg: Springer-Verlag Berlin: Person Identification Using Full-Body Motion and Anthropometric Biometrics from Kinect Videos. European Conference on Computer Vision. 2012; p. 1-10.
2. Andersson VO, Araujo RM. Person identification using anthropometric and gait data from Kinect sensor. TX: Austin: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015 January; 1:425-431. Crossref. PMid:25411178 PMCid:PMC4298551
3. Gianaria E, Grangetto M, Lucenteforte M, Balossino N. Human Classification Using Gait Features. Switzerland: Springer International publishing. 2014; p. 1-13. Crossref.
4. Zhou C, Wang L, Zhang Q, Wei X. Face Recognition Based on PCA and logistic regression analysis. Elsevier Gm BH. 2014; p. 1-3.
5. Naseem I, Togneri R, Bennamoun M. Linear Regression for Face Recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence. 2010 November; 32(11):600-11. Crossref. PMid:20603520
6. Hayashi E, Maas M, Hong JI. Wave to Me: User Identification Using Body Lengths and Natural Gestures. ACM Association for Computing Machinery copyright CHI. 2014 April; p. 1-10.
7. Kinect for Windows Developer. Date accessed: 3/12/2013: Available from: https://www.microsoft.com/en-us/download/details.aspx?id=36998
8. Gaur S, Shah VA, Thakker M. Biometric Recognition Techniques: A Review. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2012 October; 1(40):1-10.
9. Jain A, Hong L, Pankati S. Biometric Identification. Communication of the ACM Association for Computing Machinery. 2000; 43(2): 91-98. Crossref.
10. Anil JK, Arun R. Introduction to Biometrics. Jain, AK; Flynn; Ross, a Handbook of Biometric Springer. 2008; p. 1-22. PMCid:PMC2773880
11. Kaschte B. Biometric Authentication systems today and in the future. Computer Science Department world. 2005 October; p. 1-13.
12. Kinect v2. Date accessed: 12/01/2018: Available from: https://en.wikipedia.org/wiki/Kinect
13. Ross A, Jain AK. Information fusion in biometrics. Pattern Recognition Letter. 2003 September; 24(13): 2115-25. Crossref.
14. Savage R, Clarke N, Li F. Multimodal Biometric Authentication Using Kinect Sensor. International Journal of Control Theory and Applications. 2017 November; 10:1-12.
15. Khalajzadeh H, Mansouri M, Teshnehlab M. Face Recognition using Convolutional Neural Network and simple Logistic Classifier. Online conference on soft Computing in Industrial Applications.WSC17. 2012; p. 1-10.
16. Draper NR, Smith H. Applied Regression Analysis. Wiley Series in probability and statistic third Edition. 1998 April; p. 736.
17. Chaudhary S, Nath R. A New Multimodal Biometric Recognition System Integrating Iris, Face and Voice. International Journal of Advanced Research in Computer Science and Software Engineering. 2015 April; 5 (4):1-6.
18. Monvar M, Gavrilova ML. Multimodal Biometric System Using Rank-Level Fusion Approach. IEEE Transaction on Systems, Man, Cybernetics- Part B: Cybernetics. 2009 August; 39(4):867-78. Crossref. PMid:19336340
19. Zhao W, Chellappa R, Krishnaswamy A. Discriminant analysis of principal components for face recognition. Third IEEE International Conference on Automatic Face and Gesture Recognition. 1998; p. 1-6. Crossref. PMCid:PMC1364323
20. Eftekhari A, Forouzanfar M, Moghaddam HA, Alirezaie J. Block-wise 2D kernel PCA/LDA for face recognition. Information Processing Letters. 2010; 110(17):761-66. Crossref.

21. Bhattacharyya SK, Rahul K. Face Recognition by Linear Discriminate Analysis. International Journal of Communication Network Security. 2013; 2(2): 1-5.
22. Huang SH, Yang JF. Kernel Linear Regression for Low Resolution Face Recognition under variance Illumination. IEEE Transaction ICASSP International Conference on Acoustics, Speech and Signal Processing. 2012; p. 1-4.
23. Xia L, Chen CC, Aggarwal JK. Human Detection using Depth Information by Kinect. Computer Vision and Pattern Recognition Workshop IEEE Computer Society Conference. 2011; p. 15-22. Crossref.
24. Marcialis GL, Roli F. Heidelberg: Springer-Verlag Berlin: Serial Fusion of Fingerprint and Face Matchers. M. Haindl. 2007; p. 1-10. Crossref.