An Authentication System using Audio-Visual Modality

Satyajit nayak[1], Sunil Panigrahi[2] and Gati Krushna Nayak[3]
[1,3]Associate Professor, Department of Computer Science Engineering, Gandhi Institute For Technology (GIFT), Bhubaneswar
[2]Assistant Professor, Department of Computer Science Engineering, Gandhi Engineering College, Bhubaneswar

Publishing Date: January 27, 2017

Abstract—Biometrics is an emerging technology in this era, which has been widely used in many application such as secured access to a computer and any other system, criminal identification, person authentication, etc. Face recognition is a biometric method of identifying a person by comparing the data with the stored information of that person. But the recognition of face can be affected by illumination variation, facial expression and other issues. Therefore, the authentication using only face image might be difficult for the system. To overcome this challenge, biometric authentications have to rely on more than one method. In this paper, we consider audio-video person authentication based on two sources of information. Single modality evidence has limitations in both security and robustness. Therefore, here, audio recognition is added with visual recognition. Facial features, as well as speech is extracted separately from audiovisual data and integrate both the modality for secure user authentication. Mel Frequency Cepstral Coefficients (MFCC) is used as the speech feature. Viola-Jones and Scale-Invariant Feature Transform (SIFT) algorithm are used for visual feature extraction. After the extraction of audio and visual features, the feature selection method is employed. The two-phase algorithm consisting of statistical analysis, Analysis of Variance (ANOVA) with Incremental Feature Selection (IFS) is proposed to select significant features from audio-visual data. The Audio and Video processing is done in two separate phases using machine learning algorithms. The results of both the modalities are then combined at the decision level based on majority voting. It has been observed that multiple modalities of both audio visual information give immensely good results compared to a standalone single modality.

Keywords: ***A****uthentication, MFCC, Viola-Jones, SIFT, ANOVA, IFS, SVM.*

I. INTRODUCTION

Authentication is the prime need of today's digital world and the biometric authentication is the most successful approach of it. Facial recognition systems are commonly used for authentication purposes but now increasingly being used in a variety of other applications like smart phones. Image tagging on face book and other social networking as well as personalized marketing are the common application of face recognition. But it also encounters many issues in handling large amount of facial variations due to head pose, light intensity and different facial expressions [1] etc. therefore, the recognition rate may degrade. Authentication is a very important aspect for mobile devices, social network or any other financial system. In today's growing world improved and secured customer services is the most significant tool for growth and development. Speech is one of the ancient ways to express ourselves and speech recognition develops methodologies that enable recognition of spoken words into text. There are many real world applications where speech recognition is applied to authenticate, especially for remote access of a system. But speech recognition quality may degrade because of noisy environment and channel variation. Single modality based authentication is accepted everywhere for practical deployment. If the single modality based evidence is considered such as only speech or only face, it encounters many issues in both security and robustness. A possible way to dealing with the limitations of single modality is to integrate information from more than one modality. This proposed method introduces a multi-modal authenticating system that is based on audio and visual information of a person. This is an approach to improve the authenticity using audiovisual information where speaker speaks a particular word which is password and with this, speakers facial biometrics are also added to that system. The system is designed to prevent unauthorized access of any system, social network and web security etc. There are many research papers where voice/speaker recognition is added with face recognition in the area of biometric authentication. Here we address some issues of authentication using speaker recognition.

Identify applicable funding agency here. If none, delete this.

. Illness of a person such as a cold, throat infection can change a person's voice. So, in that case making absolute identification is difficult or impossible.

- Speaker recognition does not provide very good accuracy. In the maximum situation like noisy environment and channel variations, it gives low accuracy.

Therefore, we consider speech recognition with facial recognition to generate a multi-modal system. The proposed system

decomposes the information from video data into two components: speech recognition and face recognition. The face recognition subsystem is responsible for detecting the face and facial part from video using the Viola-Jones[2] and SIFT[3] algorithm. Similarly, in speech recognition module, MFCC[4] is extracted as a speech features and select the efficient features using Analysis of variance (ANOVA)[5]. Both the recognition task is carried out using multiclass Support Vector Machine (SVM)[6].

Motivations:

- Combining evidence from two modalities will increase the authenticity and integrity of the system.
- Adding facial features to the speech recognition is beneficial because the biometric feature is difficult to share and cannot be forgotten and lost.

Applications:

- Improve authentication of remote access to any system.
- The multimodal biometric system can be used for online transaction.
- Increase the security of payment gateways in samartphone.
- To improve security of any system by adding biometric features like eyes, nose and lip detection to speech recognition.
- In social media application, data protection, security in smarrphone, it can also be used.

Issues from the literature survey and addressed in this paper:

- Only facial biometric based authentication may encounter many limitations and lack of robustness beacuse of lighting issue, facial expression. Audio signal may degrade because of noisy environment, chanel variation .
- Multiple face profile may face the large image variation problem.
- In facial recognition single module of face has been used for visual feature with audio features.
- Frame level fusion mismatch the audio and visual frame. . Feature fusion is partially valid because there is different data rate of audio and visual data and differing segmentation.
- Feature matrix may contain redundant information and irrelevant feature.

Major contributions of this paper are: . Develop facial and speech features based multimodal authentication system.

- In speech recognition MFCC is used to extract speech features.
- Face recognition model is developed based on two types of facial features: extract features from Viola Jones and SIFT algorithm.
- ANOVA statistical analysis is used for analyzing statistical significance of each feature vector. After ANOVA, IFS is implemented here for selecting statistical significant features from speech and facial feature matrix.
- Final decision is taken on the basis of majority voting of three individual systems, speech recognition, face recognition using Viola Jones and face recognition using SIFT algorithm.

The paper is organized as follows: Section II provides the related study of audio visual based multi model authentication system, proposed methodology is discussed in section III, Section IV and V give the experimental result and integration method. Result analysis is provided in Section VI followed by the conclusion in Section VII.

II. RELATED STUDY

To prevent any unauthorized access to any private data, security is a prime concern. Biometric features are very sensitive data and it's use is subject to the privacy protection law [7]. Taking this issue authors proposed a multi-modal biometric authentication system using encrypted biometric for edge-centric cloud environment. They used two approaches for authentication; encrypted speech and face in a cloud environment. In the speech, the author used MFCC and perceptual linear prediction (PLP) coefficients for encrypting the biometrics through edges. Here they have taken final decision based on majority voting. Authentication through a single modality is unreliable and not very secure in a real-world application. To enhance the authentication author proposed the android based multi modal biometric identification of a system [8]. Face and voice biometric recognition techniques were used to identify a person. In both the voice and face recognition, face feature vector and voice feature vector were extracted separately and were integrated at the feature level. Finally the authentication of the designed system was done by binary classification of the SVM classifier [8]. The paper [9] proposed a system based on multi modal biometric face recognition. Here 2D Discrete Wavelet Transform (DWT) as well as data fusion (DF) was used and they have applied the data fusion at the score level of the system algorithm. Principle Component Analysis (PCA) and Local Binary Pattern (LBP) method were used for image feature extraction, also third feature vector of images are transformed into DWT. All these features are used for recognition of face using K-Nearest Neighbor (KNN) algorithm. This paper [10] combined face and behavioral biometric that requires speech to enhance the security of a system. Here for extracting features MFCC was used and for training the designed system Fuzzy logic and Genetic algorithm were used which also helps in optimizing the feature values. Ultimately they have done the fusion of face and speech by integrating their feature values.

To enhance multi modal biometric authentication, face and speaker recognition were combined to achieve better performance [11]. Multi modal person authentication system combined information about the face and voice of a person. Multiple face profiles were extracted from the user in this Audio-video user authentication system [12] MFCC speech

coefficients were used with transformed face profiles (TFP) feature vectors to form an audio-visual feature vector. The nearest neighbor classifier was used to process the audiovisual feature. In [13] the author presented an audio-visual person authentication system that consists of several novel visualized speech features from the spoken password and multiple face profiles. In the paper [14], speech signal and facial movement had been analyzed by the author. They have proposed a multi-modal person authentication system and fused the lip movement as well as speech signal for better authentication purpose which improves performance of the system. The human face is a three-dimensional (3D) surface and has a detailed anatomical structure, so 3D face information improves the performances of the recognition system. In [15], the authors have proposed a feature warping technique based on thin-plate-spline (TPS) analysis for 3D audio-video person authentication. TPS features include speaking face such as expression lines, gestures, and wrinkles. The system enhanced the performance of the system against imposter and spoof attacks. Experiments has been done using VidTIMIT and AVOZES data corpus. This paper [16] also proposed an acoustic and visual feature based person authentication using distinguish live synchronous audio-video recordings from replay attacks that use audio with a still photo.

III. Proposed Methodology

Multimodal authentication includes two separate recognition process and after that combine the both in a decision level. The proposed methodology includes following steps:

Database collection: Audio-video dataset of an individual has been collected in a lab environment. Speech Recognition;

- Audio Feature: Audio feature mel frequency cepstral coefficient (MFCC) is used for speech recognition.
- Audio feature selection: ANOVA and IFS are used for selecting significant speech feature.
- Classification: Multiclass SVM is used for classification and recognition of speech.

Facial feature Recognition:

- Face recognition: Face and facial part identification from a video is done using Viola Jones and SIFT algorithm.
- Facial feature selection: ANOVA+IFS is also used here for selecting facial feature.
- Classification: Multiclass SVM had been used for recognition of a person based on facial part.

Integration: At the decision phase majority voting is used on the basis of three individual decisions. Face recognition provides two individual decisions about the identity of a user by applying Viola-Jones as well as SIFT features. The third decision is come out from speech recognition module. Majority voting of these three decision makes the final decision of user authentication. In this majority voting, for taking a decision, one positive result has to come from audio speech recognition and for face recognition, any one positive result from the two modalities. Flowchart of multi-modal authentication is depicted in figure 1.

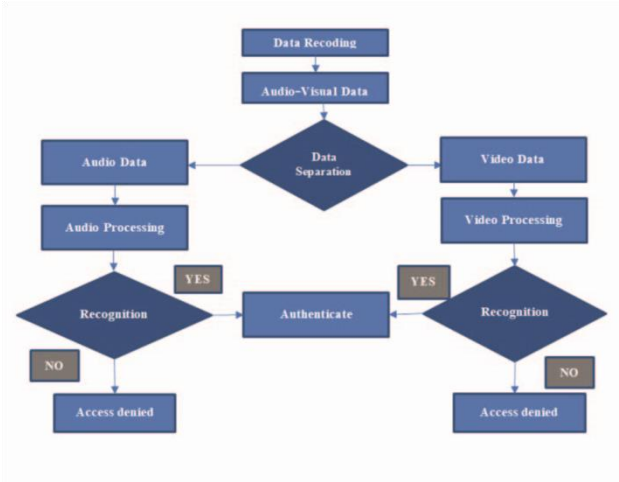

Fig. 1. Flowchart of multi-modal authentication.

1) Database description: Database had been created by us under lab environment since no database has been found for our work after survey. Our database consists of 10 different speakers (5 male and 5 female). Total 1000 (10 connected digits * 10 speakers * 10 utterances) utterances. The recoding was done in a noise-free environment. The recording type is mono channel and done with the help of a microphone and video camera. There are certain constrains we had to follow to get better result in face recognition, those are-

- No head movement while recording.
- Upright head position.
- One person at a time.

A. MFCC

Acoustic feature extraction is the successful method in the domain of speech research. MFCC is the widely used acoustic feature extraction proposed in 1980 by David and Mermelstein 1980 [4]. In this, analyze each speech signal into short time duration by splitting the signal into several frames instead of analyzing the total signal at once. After framing the signal a window function is multiplied with each frame of the speech signal. Discrete Fourier Transform (DFT) is used for extracting spectral information for a discrete frequency band. Here FFT is used to convert the signal from time domain to frequency domain for preparing the next stage of mel frequency warping. After getting FFT, the Mel filter bank includes following calculations.

Mel scale (Davis et al., 1980) is defined as:

$$m_f = 1125ln(1 + \frac{f}{700}) \quad (1)$$

Where, is the actual frequency in Hz. In the final step the log Mel spectrum is converted back to time domain. The result is called the MFCCs [17].

B. Viola-Jones

The first step in the face identification is that the face has to be detected from the larger image or video. The Viola-Jones object detection framework facilitates Haar-Like features [19] to be extracted from a face image as the initial step. The reason Haar-Like features are favored over the raw pixel value of the image as input to a learning algorithm is to reduce the inclass variability while increasing the out-of-class variability compared to the raw data and thus making classification easier [19]. A Haar-like feature considers neighboring rectangular regions in the image that is targeted for facial detection. It sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, if we have an image with a human face, it is common that the region of the eyes is darker than the region of the cheeks. Therefore, a common Haar-like feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region [19].

C. SIFT

SIFT [3] was developed for object recognition as well as matching features from an image against a database of features extracted from that images. The main steps in the SIFT algorithm are:

- Scale-space extrema detection: Potential interests points are identified by using a Difference of Gaussian function to identify points in the image that are invariant to scale and orientation.
- Key point localization: Key points are located precisely so that only stable and salient features are retained.
- Orientation Assignment: To make an image rotation invariant, assignment of orientation has to be carried out. A histogram having 36 bins crossing over 360 degrees is made. The highest peak in the histogram and peak greater than 80 percent is considered to measure the assignment of orientation.
- Key point Descriptor: A key point with16x16 neighbourhood is considered. It is parted into 16 sub-blocks of 4x4 sizes. For each sub-block, 8 bin orientation histogram is produced and 128 feature vectors are obtained.

D. Feature selection

The objective of feature selection is to select the efficient and important features that can increase recognition rate. In this work two phase algorithms is used to rank the feature and select the feature subset. To rank the features using statistical analysis, Analysis of Variance (ANOVA) is used for feature relevance criterion along with incremental feature selection (IFS) technique which are explained below.

1) The analysis of variance (ANOVA): ANOVA is a statistical method to test the difference in means between groups [5,18]. It computes significant difference in a scale-level dependent variable to a nominal-level variable having 2 or more categories.

$$F(\xi) = \frac{sB^2(\xi)}{sW^2(\xi)} \tag{2}$$

[5] Where $sB^2(\xi)$ is the sample variance between groups and $sW^2(\xi)$ is sampled variance within groups [10]. The decision is made using F-statistics, in one way ANOVA, the F-statistics is a ratio calculated by following equation-

$$F = \frac{Variation\ between\ sample\ means}{Variation\ within\ the\ samples} \tag{3}$$

2) Incremental Feature Selection (IFS): IFS [18] method evaluates performance of the system using different feature subset. The feature with the highest F value in the ranked feature set is the first feature. After that second highest F value is added and continues with this procedure. Until all the candidate features are added, this process continues from the higher F value to the lower F value. A new feature subset is produced using this procedure. The F statistics calculate the statistical information such as mean and variance.

E. Support Vector Machine

SVM [Aravind Ganapathiraju at all.] is a machine learning technique that estimates decision surfaces directly rather than modelling a probability distribution across the training data [6]. SVM classifies the non-linear output vector to a higherdimensional feature space. We are using different kernel function of SVM classifier such as radial basis function, polynomial and linear and select the best kernel function of the classifier. The model has been selected based on feature selection and best kernel function.

IV. Experimental Result and Analysis

A. Speech Recognition

19 dimensional MFCC features are extracted from speech signal. ANOVA test is carried out to rank the feature based on the F statistics which is discussed above. Rank the MFCC cepstral coefficients according to ANOVA test. The performance of the system is calculated by total number of words correctly recognized during testing phase. [17]. In this experiment the performance of the individual feature can be evaluated and also observe the performance using multiclass SVM when concatenating with other features. Incremental feature selection (IFS) technique gradually concatenates the features. From the experiment it has been observed that SVM classifier with kernel function rbf gives the highest accuracy for 12 (MFCC coefficients 1,2,7,3, 5, 4, 10, 13, 11, 18, 7, 6) cepstral features. Performance of speech recognition with feature selection technique using SVM classifier is 95.03 %.

TABLE I
SPEECH RECOGNITION USING MFCC, ANOVA AND SVM

Exp. No.	Kernel Function	No.of features	Accuracy
1	RBF	12	95.05
2	Linear	18	82.75

3	Polynomial	13	92.78

B. Facial feature Recognition using Viola-Jones

Frames are extracted from video of a person and from all these frame facial features are detected. All the detected face are collected and saved into folders for the next stage,

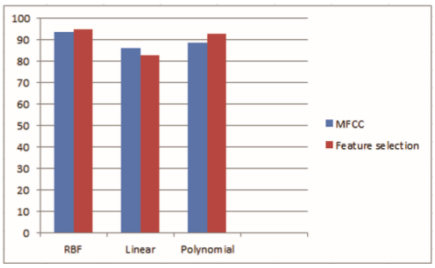

Fig. 2. Performance comparison using MFCC and feature selection method for audio recognition.

Fig. 3. Extraction of frames.

feature extraction. Our corpus consists of 1000 faces, 100 faces per person (10*100=1000). Viola-Jones and SIFT algorithms are used for facial feature extraction. Figure 2 shows the detected face from video using viola jones. Facial part such as eyes, nose, and lip are detected from face using viola jones algorithm which are depicted in figure 3 The first step of face recognition is to detect the face from the larger image or video. The Viola-Jones object detection framework facilitates Haar-Like features to extract face from an image as the initial step [19,20,21]. Therefore recognition of a person is based on their biometric features which were collected from their video stream. Training and testing is carried out using multiclass SVM machine learning algorithm. Table 2 gives the experimental result of speech and face recognition using SVM.

C. Facial feature recognition using SIFT

SIFT keypoints of objects from face are first extracted and stored in a database. A face is recognized by individually

Fig. 4. Face detection using Viola Jones.

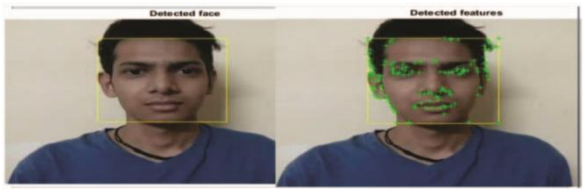

Fig. 5. Feature Extraction using SIFT.

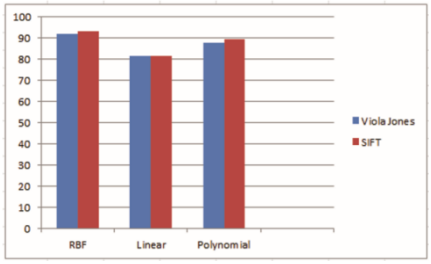

Fig. 6. Performance comparison using Viola Jones and SIFT for visual recognition.

comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of keypoints that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches. Such features are extracted in four steps**.** In first step, orientation and scale invariant features at all scale are finding out using a DOG filters.

- An extended model is used in the next step to determine the appropriate points based on different sustainability criteria.
- In third step, based on the local gradient of the image at target points, one orientation is assigned to each feature point.
- Finally, the information contained in the gradient function around the feature points is encoded. Each feature is considered as a vector in a 128-dimensional space identified over the key point neighborhood.

  Feature selection method ANOVA+IFS is employed here to remove redundant features and improve recognition rate using significant features. after this analysis 12 keypoint features are selected for facial recognition using SIFT and for Viola-Jones matrix dimension has also been reduced.

TABLE II
FACE RECOGNITION ACCURACY USING VIOLA JONES, SIFT AND SVM

Exp. No.	Kernel Function	Viola Jones	SIFT
1	RBF	92.15 %	93.05 %
2	Linear	81.56 %	81.75 %
3	Polynomial	87.66 %	89.45 %

V. INTEGRATION

In this work we have recognized audio and video of a person in a separate module using SVM classifier. Here three individual model is developed. For facial features recognition two model using Viola-Jones as well as SIFT are developed and for speech recognition one model using MFCC with feature

selection ANOVA+IFS is generated. Therefore, three individual recognition model is established. If two of the recognition model gives positive response then system provide authentication of that person and from two of the positive response one have to come from speech recognition. We set threshold for each individual recognition model based on the their recognition rate. If the results is greater or equal to that threshold then only we consider that result as a positive response.

VI. Result Analysis

The experiment has done separately for both audio and video. For speech recognition average 95.03 % accuracy is achieved using MFCC and feature selection technique. We have used Viola-Jones and SIFT algorithm for facial features recognition of a person from video stream. Feature selection technique is proposed to select important feature which improve recognition accuracy. In both the cases SVM classifier is used. If we analyze the system individually, speech and face recognition can give good recognition rate in a lab environment. But when we talk about authentication of a system, single modality can be easily spoofed by the intruder. Combination of audio and visual systems helps to overcome the shortcomings of individual recognition. Multi modal authentication system is more secure as compared to single modality authentication system. Taking these points of view our proposed multi modal authentication system is designed. Proposed system combined audio and visual recognition at the decision level and recognizes the speech based on audio visual features. Speech is recognized and considered as password; along with speech persons biometric information are also added. Therefore authentication system becomes more secure having biometric information with speech based password. Comparison of results for audio and visual recognition are depicted in figure 2 and 6.

VII. Conclusion

Audio-visual authentication systems are still an important subject of research. In this paper, work has been focused on establishing a multi-modal user authentication system based on face and speech recognition. MFCC is used to extract the features from speech and Viola- Jones and SIFT algorithm are used to extract the face and facial part from video. After the face has been detected and extracted, it undergoes face preprocessing steps. The face pre-processing procedure consists of eye detection, nose detection , lip detection. ANOVA statistical algorithm is proposed for both audio and visual feature selection and dimension reduction. At the decision phase when two systems recognize the audio and visual biometric feature then only system authenticates. All these experiments are executed in a lab environment; in a real time environment the recognition rate may be dropped. This multimodel system increases the security of any authentication system. The research on the multi-modal authentication system can be enhanced and extended to cover more scenarios with better results.

References

[1] R.L.Hsu and A.K.Jain, Face Modeling for Recognition, Proceedings Intl Conf. Image Processing, ICIP, Greece, Oct. 7-10, 2001.

[2] YI-QING WANG , “An Analysis of the Viola-Jones Face Detection Algorithm,” *IPOL Journal Image Processing On Line* , 4 (2014), pp. 128148. https://doi.org/10.5201/ipol.2014.104.

[3] Ridhi Jindal, Dr. Sonia Vatta , SIFT: Scale Invariant Feature transform (Review),” International Journal of Advance research, Ideas and Innovations in Technology , Volume 1, Issue 1, October 2014.

[4] Davis ,S. B. , Mermelstein, P.,”Comparison of Parametric Representation for Monosyllabic Word Recognition in Continuously Spoken Sentences,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), pp. 357-365.(1980).

[5] Ding H, Feng PM, Chen W, Lin H , Identification of bacteriophage virion proteins by the ANOVA feature selection and analysis. MolBioSyst 10(8):22292235 (2014).

[6] Aravind Ganapathiraju, Jonathan E. Hamakerand Joseph Picone ”Applications of Sup-port Vector Machines to Speech Recognition” *IEEE TRANSACTIONS ON SIGNAL PRO-CESSING*, VOL.52, NO. 8, (2004).

[7] Zulfiqar Ali ; Misbah M. Baig; Shahzain Mehboob , “Edge-centric multimodal authentication system using encrypted biometric templates,” *Future Generation Computer Systems*, Elsevier (2018).

[8] Xinman Zhang, Yixuan Dai1, Xuebin Xu; Misbah M. Baig; Shahzain Mehboob , “Android-based multimodal biometric indentification system using feature level fusion,” *International Symposium on Intelligent Signal Processing and Communication Systems*, November 6-9, 2017, China.

[9] Bilal Ahmed, Taner Cevik, Nazife Cevik , “Data Fusion-Based Multimodal Biometric System for Face Recognition Using Manhattan Distance Penalty Weight,” *International Journal Of Electronics, Mechanical And Mechatronics Engineering* , Vol.7 Num.3 - 2016 (1441-1452).

[10] Nancy Bansal, Amit Verma, Iqbaldeep Kaur, ”Multimodal Biometrics by Fusion for Security Using Genetic Algorithm,” *4th IEEE International Conference on Signal Processing, Computing and Control (ISPCC 2k17)*, Sep21-23, 2017, Solan, India.

[11] Lalitha S, Ashwini V, Madhusudhan.K.N, Sachin B S, ”Person Authentication Using Face And Voice Modalities,” *International Journal of Advances in Science Engineering and Technology*, ISSN: 2321-9009 Volume- 1, Issue- 2, Oct-2013.

[12] Amitava Das & Vaibhav Bedia, ”Audio-Visual Person Authentication with Multiple Face-profiles and Compressed-Feature-Dynamics Signatures of Spoken Passwords,” *IEEE.MMSP09*, Rio de Janeiro, Brazil. 978-1-4244-4464-9/09, October 5-7, 2009.

[13] Amitava Das, Ohil K. Manyam, Makarand Tapaswi, ”Audio-Visual Person Authentication with Multiple Visualized-Speech Features and Multiple Face Profiles,” *IEEE*, 978-0-7695-3476-3/08 2008 DOI 10.1109/ICVGIP.2008.106.

[14] Ming Liu, Thomas S. Huang, ”VIDEO BASED PERSON AUTHENTICATION VIA AUDIO/VISUALASSOCIATION,” *IEEE* 1424403677/06.

[15] Girija Chetty and Michael Wagner, ”Audio-Video Person Authentication based on 3D Facial Feature Warping,” *Proceedings of the Digital Imaging Computing: Techniques and Applications ,IEEE*, (DICTA 2005) 0-7695-2467-2/05 2005.

[16] Girija Chetty and Michael Wagner,””Liveness Verification in AudioVideo Authentication,” *INTERSPEECH 2004 ICSLP 8th International Conference on Spoken Language Processing*, ICC Jeju, Jeju Island, Korea October 4-8, 2004.

[17] Soni,B., Debnath,S. , Das,p.k. (2016), Text-dependent speaker verification using classical LBG, adaptive LBG and FCM vector quantization, International Journal of Speech Technology September 2016, Volume 19, Issue 3, pp 525536.

[18] Saswati Debnath, Pinki Roy (2018) Speaker Independent Isolated Word Recognition based on ANOVA and IFS ICCMS 2018, January 810, 2018, Sydney, Australia 2018 Association for Computing Machinery. ACM ISBN978-1-4503-6339-6/18/0115.00 https://doi.org/10.1145/3177457.3191708.

[19] Lienhart, Rainer, and JochenMaydt. An Extended Set of Haar-like Features for Rapid Object Detection. Intel Labs, Intel Corporation, Santa Clara, n.d. Web. 17 June 2014.
[20] Gregori, Eric. Introduction To Computer Vision Using OpenCV. Embedded Vision Alliance.Embedded Vision Alliance, n.d. Web.9 Jan. 2014.Color FERET Database.N.p., n.d. Web. 18 Feb. 2014.
[21] Wilson, Phillip I., and John Fernandez, Facial feature detection using haar classifiers. Texas A & M University, Apr. 2006. Web. 17 June 2014.