IJCTA, 10(9), 2017, pp. 881-888 ISSN: 0974-5572

Person identification using combined Face and Speech for the reduction of FAR and FRR

Dinesh Kumar D.S.* and P.V. Rao**

ABSTRACT

Biometric Identification deals with person identification based on physiological and behavioural features.Since, single modal systems suffered from a lot of disadvantages such as non-universality,noise in sensed data, spoof attacks etc.To overcome these difficulties, a novel method of multimodal system combining two or more biometric traits is used. The proposed system combines the features of both face and speech for person identification. In this paperhistogram method is used for face feature extraction and LMS algorithm for speech features extraction which provides better recognition in terms of reduced FAR0.00 & FRR of 0.05.

Keywords: FAR, FAR, LMS Algorithm

1. INTRODUCTION

Biometrics plays an important role in providing secure authentication techniques in various fields based on measurable physiological and individual characteristics of a person that can be automatically verified. Depending on the application, a biometric system may operate in verification mode or identification mode. Multimodal biometric integrates different biometric systems for verification in making a personal identification.

2. RELATED WORK

Biometric Identification deals with the person identification usingvarious biometric identities [1, 2]. There are several biometric identities that have been commonly used. These techniques are broadly divided into two categories. The physiological biometrics that deals with static characteristics that includes faces [3], ear, hand geometry, iris etc. The second category is the behavioural biometrics that includes the active features such as speech, signature, handwriting etc. The lower recognition efficiencies of single modal systems resulted in a motivation towards the multi-modal recognition systems [4, 5, 6, 7, 8, 9, and 10]. Here the motivation was to combine two or more modalities in order to achieve higher recognition efficiencies and also to overcome the limitations of single modal systems.

S. Palanivel and B.Yegnanaraya, proposed a method on automatic multimodal person identification system for combining face, speech and visual speech. Here auto associate neural network models are used to extract the features of face, speech and visual speech.Motion information is used to locate face region and eyes. This method produces an equal error rate of about 0.45%for 50 subjects [9].

Mahesh. P.K. and M.N. Shanmukhaswamy, proposed a biometric authentication system for face and palmprint identity verification and this gives an accuracy of the overall accuracy of the System is more than 97%, FAR & FRR of 2.4% & 0.8% respectively.[10].

* Associate Professor (RS-VTU), Dept. of ECE, VTU Research Scholar, Bangalore, Karnataka, India, *Email: crp.kumar@gmail.com*

** Professor, Dept. of ECE, Vignanabharathi Institute of Technology, Ghatkesar, Hyderabad, India, *Email: pachararao@rediffmail.com*

Changhan Park and Joonki Paik, extracts the features of speech and face using PCA and HMM algorithm. The disadvantages of a single mode system were eliminated and FRR is reduced to 0.0001% [11].

Shaveta, Naveen Kumari,proposed a Biometric Authentication Using Face, Speech and Fingerprint Feature [12].here the FAR and FRR values are reduced.

3. PROPOSED METHOD FOR COMBINED FACE AND SPEECH

3.1. Face feature extraction

The features of face are extracted using the following steps. Initially face image is feed to the algorithm. In the pre-processing step the input image is applied to resized, color space conversions. The most common way to change the size of an image is to resize or scale an image.The input image is resized to a new vale i.e. 250 × 250, it is necessary to display the image pixels within the range. The resized image is applied color space conversion; in this the input image is in color format it must be converted to grayscale image. The converted gray scale image has intensities ranging from 0-255. The gray image is applied to the segmentation process. The gray image is applied to the histogram.

Histogram of an imageis a graph that describes the occurrence of pixel intensities. In an image histogram, the x axis shows the gray level intensities and the y axis shows the frequency of these intensities.

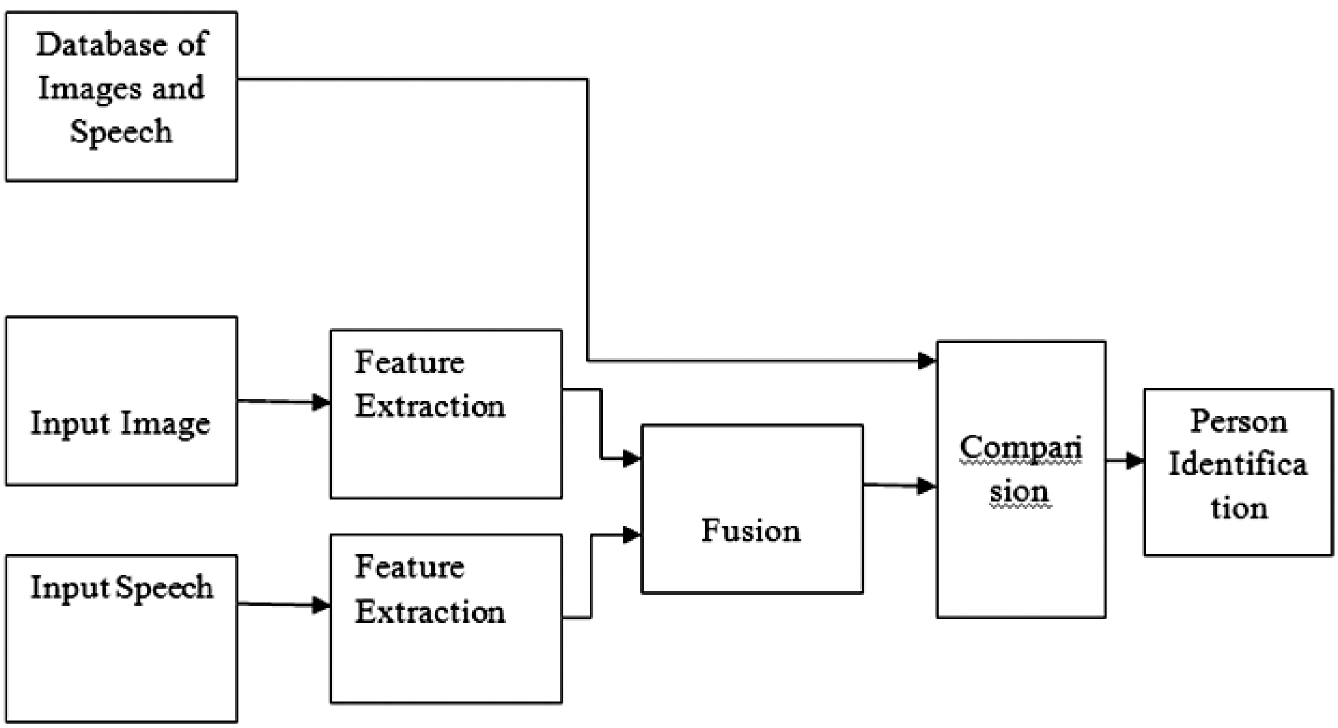

Figure 1: Block diagram of proposed method for combined face and speech

3.2. Speech feature extraction

Speech features are extracted using the following steps. Here the input audio signal is applied to adaptive filter to remove the noise present in the audio signal. Initially input audio signal is in time domain and it must be converted to the frequency domain after filtering. The frequency domain signal is applied to normalized Fast Fourier Transform to normalize the harmonic component present in input signal. Extracting the power spectral density values from the normalized FFT signal and these values are verified with the database. If the input values are matched with database it displays authorized or if it’s not matched displays unauthorized as shown in figure 2.

4. FLOW CHART OF FUSION OF FACE AND SPEECH

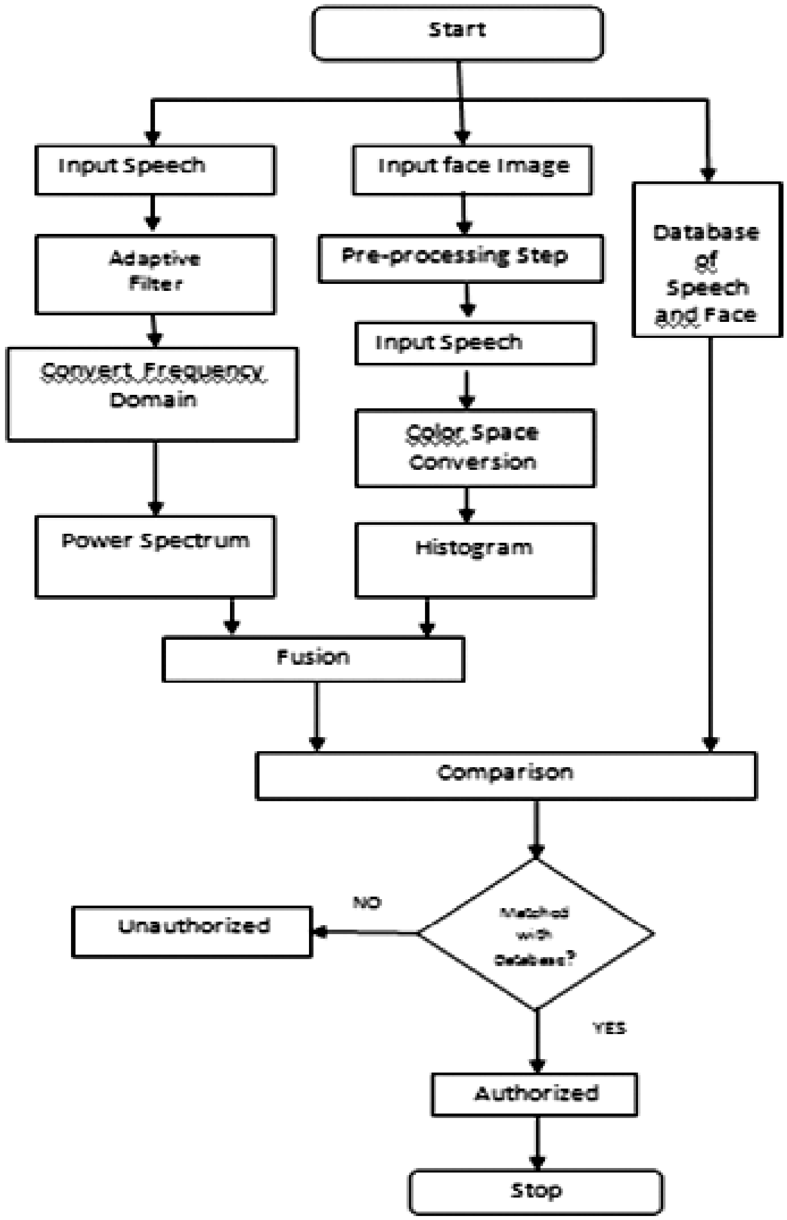

Figure 2: Flow chart of combined face and speech

5. ADAPTIVE FILTER

Adaptive filter is a linear filter which completely removes the noise in the given image, by minimizing the error between the given images and desired output image. The choice of the weight vector W is based on the statistics of the image. Basically, the objective is to optimize the filter response with respect to a prescribed criterion, so that the output contains minimal contribution from noise and interference. There are three prominent criteria's for choosing the optimum weights, and they are –

1. Minimum Mean Square Error
2. Maximum Signal to interference ratio
3. Minimum Variance

The following Fig. 3 shows the diagram of a typical adaptive Filter.

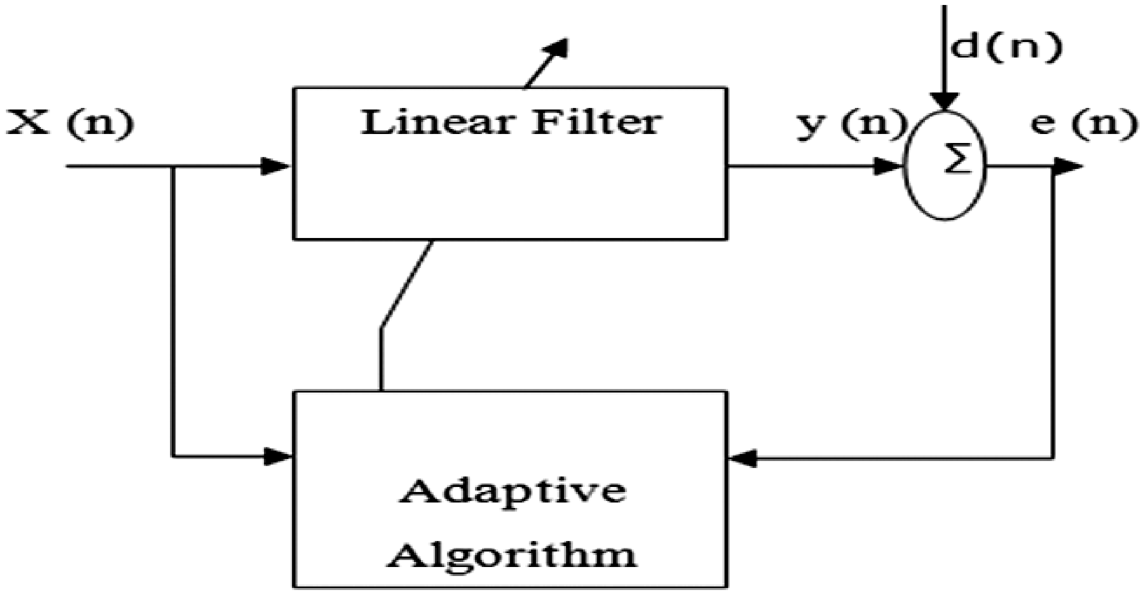

Figure 3: Adaptive Filter

Where,

- x(n) is the input signal to a linear Filter at time n
- y(n) is the corresponding output signal
- d(n) is an additional input signal to the adaptive filter
- e(n) is the error signal that denotes the difference between d(n) and y(n)

An adaptive algorithm adjusts the coefficients of the linear Filter iteratively to minimize the power of *e(n)* is shown in the above Fig. 3. For different applications the input and output signals *x(n), y(n), d(n) and e(n)* are chosen in several ways. Adaptive Filter is performed on the degraded image that contains original image and noise.

6. LEAST MEAN SQUARES (LMS) ALGORITHM

The Least Mean Square (LMS) algorithm is an adaptive algorithm, which uses a gradient-based method of steepest decent. This algorithm uses the estimates of the gradient vector from the available data. LMS incorporates an iterative procedure that makes successive corrections to the weight vector in the direction of the negative of the gradient vector which eventually leads to the minimum mean square error.

The LMS algorithm can be summarized in following equations

$$\text{Output: } y(n) = W^H x(n) \tag{1}$$

$$\text{Error: } e(n) = d^*(n) - y(n) \tag{2}$$

$$\text{Weight: } W(n+1) = W(n) + \mu \; x(n) \; [\; e^*(n) \;] \tag{3}$$

Where μ is the step-size parameter and controls the convergence characteristics of the LMS algorithm, e^2 (n) is the mean square error between the output y(n) and the desired output d(n). The successive corrections of the weight vector eventually leads to the minimum value of the mean squared error.

7. IMPLEMENTATION RESULTS

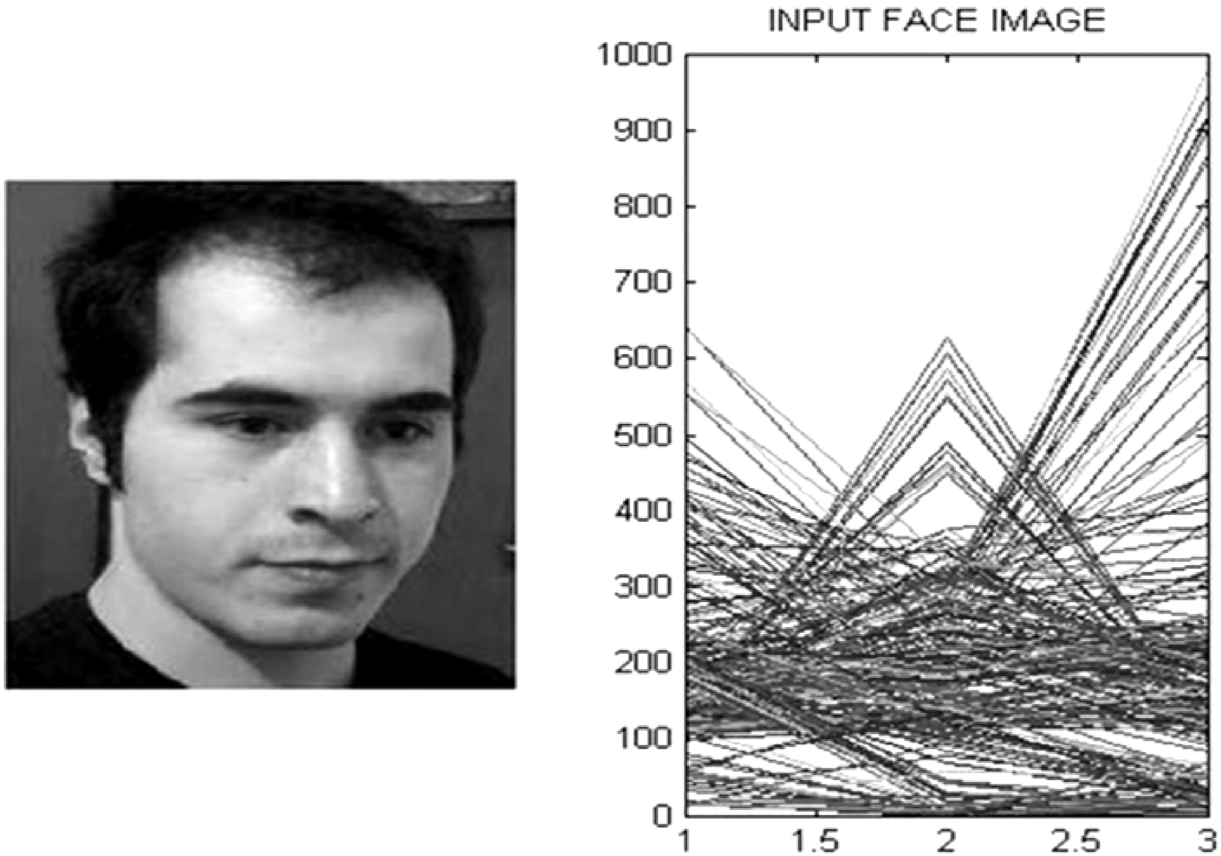

Figure 4: Input face image

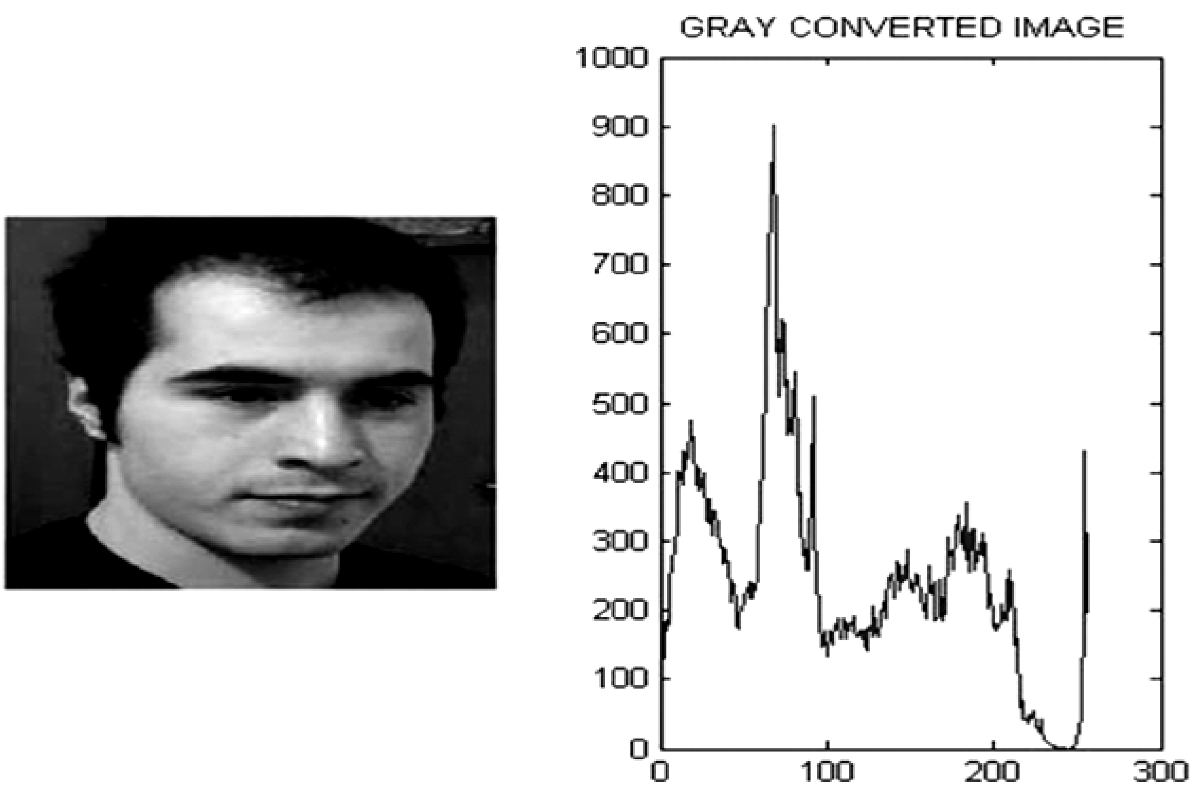

Figure 5: Gray converted image

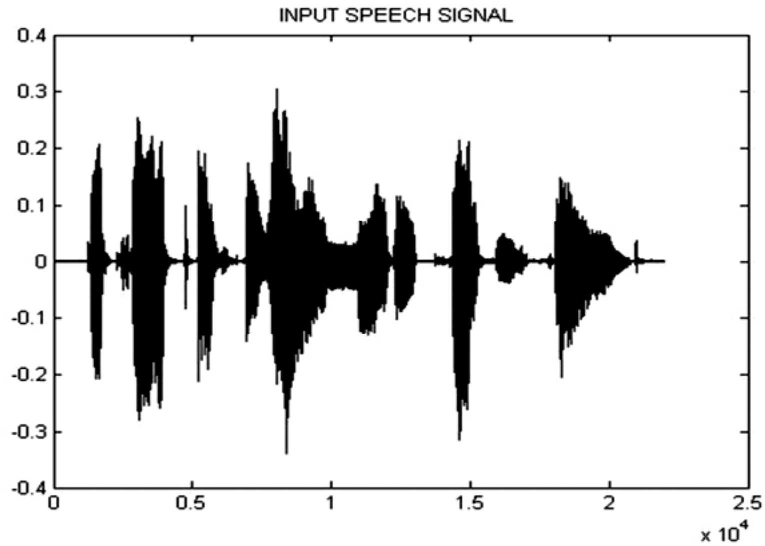

Figure 6: Input speech signal

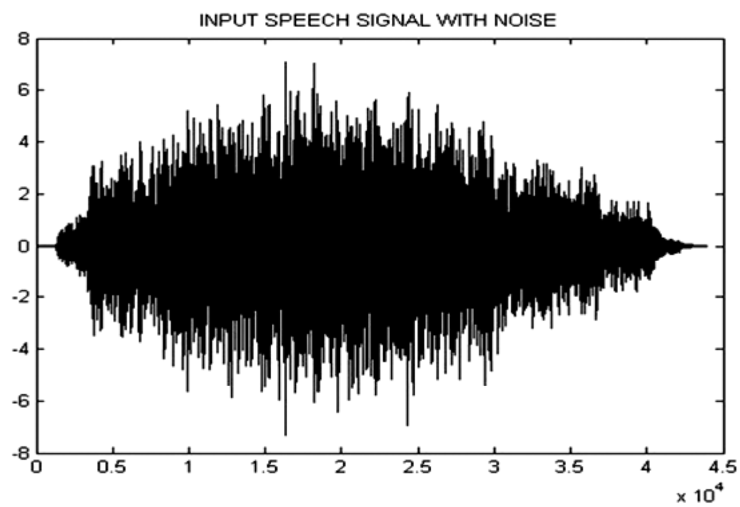

Figure 7: Input speech signal with noise

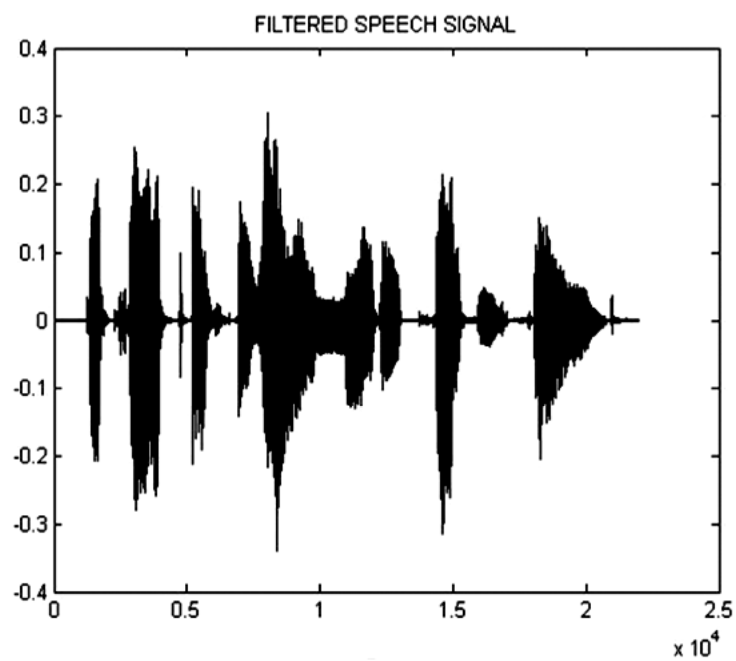

Figure 8: Filtered speech signal

Recognition rate = Number of databases matched/Total number of databases*100

FAR: When an unauthorised person is accepted as authorised, then it is known as False Acceptance Ratio (FAR).

$$FAR = \frac{WA}{N}$$

WA = No. of wrong person Accepted.

N = No. of times fraud person trials.

FRR: When an authorised person is rejected as unauthorised, then it is termed as false rejection rate.

$$FRR = \frac{PR}{N}$$

8. CONCLUSION

In this proposed work the biometric traits like face and speech are combined to achieve better recognition rate with reduced false acceptance rate (FAR).The simulation is done using MATLAB R2012a. In the proposed method, histogram based feature extraction technique is used for face recognition and spectrum based feature extraction technique is used for speech recognition. Theresulting system (multi-modal) considered here providebetter performance. The proposed multi model system has been verified with a database having 20 people's speech and face. Simulation results shows that the proposed system is veryefficient with recognition rate of 95% with FAR of 0.00 and FRR of 0.05.

REFERENCES

[1] Dineshkumar D.S, Dr P.V. Rao, analysis and design of principal component analysis and hidden markov model for face recognition

[2] Souheil Ben-Yacoub, Yousri Abdeljaoued & Eddy Mayoraz, Fusion of Face and Speech Data for Person Identity Verification, IEEE Transactions On Neural Networks, Vol 10, No 5, 1065, September 1999.

[3] Anil Jain, Lin Hong, SharathPankanti, Biometric Identification, Communications of the ACM, Volume 43, Issue 2, pp. 90 – 98, February 2000.

[4] Ching-Han Chen, Chia-Te Chu, Combining Multiple Features for High Performance Face Recognition System, International Computer Symposium(ICS2004) Taipei, (Dec 2004), pp. 387-392.

[5] Robert Snelick, Mike Indovina, James Yen, Alan Mink, Multimodal Biometrics: Issues in Design and Testing, ICMI'03, Canada, (Nov 5-7. 2003), pp. 68-72.

[6] A. Ross, A. Jain, Information fusion in biometrics, Pattern Recognition Letters, (2003), (24), 2115- 2125.

[7] Andrew L. Rukhin, Igor Malioutov, Fusion of Biometric Algorithm in the Recognition Problem, Pattern Recogition Letters, (2001), 299-314.

[8] R.W. Frischholz and U. Dieckmann, Bioid: A Multimodal Biometric Identification System, IEEE Computer, (Feb 2000), (33), 64–68.

[9] S. Palanivel, B. Yegnanarayana, "Multimodal person Authentication using Speech, Face and Visual speech", in Computer Vision and Image understanding, 20 January 2007.

[10] Mahesh P.K., M.N. Shanmukha Swamy, "A Biometric Identification System based on the Fusion of Palmprint and Speech Signal", in 978-1-4244-8594-9,2010 IEEE.

[11] Changhan Park, Joonki Paik, "Multi-Modal Human Verification Using Face and Speech",in ISBN 978-3- 902613-03-5,I-Tech,Vienna, Austria, June 2007.

[12] Shaveta, Naveen Kumari, "Multimodal Biometric Authentication Using Face, Speech and Fingerprint Feature" ISSN: 2277 128X , Volume 5, Issue 8, August 2015

[13] R. Brunelli and D. Falavigna, "Person identification usingMultiple cues," *IEEE Trans. Pattern Anal. Machine Intell.*,vol. 17, pp. 955–966, Oct. 1995.

[14] B. Duc, G. Maitre, S. Fischer, and J. Bigun, "Person Authentication by fusing face and speech information," in *1st Int. Conf. Audio- Video- Based Biometric PersonAuthentication AVBPA'97*, J. Bigun, G. Chollet, and G.Borgefors, Eds. Berlin, Germany: Springer-Verlag,Mar. 12–14, 1997, vol. 1206 of Lecture Notes in Computer Science, pp. 311–318.

[15] E. Bigun, J. Bigun, B. Duc, and S. Fischer, "Expertconciliation for multi modal person authentication systems By Bayesian statistics," in *Proc. 1st Int. Conf. Audio-Video-Based Biometric Person Authentication AVBPA'97*. Berlin, Germany: Springer-Verlag, Lecture Notes in ComputerScience, 1997, pp. 291–300.

[16] L. Hong and A. K. Jain, "Integrating faces and fingerprint for Personal identification," *IEEE Trans. Pattern Anal. MachineIntell.*,vol. 20, 1997.

[17] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "OnCombining classifiers," *IEEE Trans. Pattern Anal. MachineIntell.*,vol. 20, pp. 226–239, 1998.

[18] A. K. Jain, L. Hong, and Y. Kulkarni, "A multimodalbiometric system using fingerprints, face and speech," in*Proc. 2nd Int. Conf. Audio-Video Based Biometric PersonAuthentication*, Washington, D.C., Mar. 22–23, 1999,pp. 182–187.

[19] T. Choudhury, B. Clarkson, T. Jebara, and A. Pentland,"Multimodal person recognition using unconstrained audioand video," in *Proc. 2nd Int. Conf. Audio-Video BasedPerson Authentication*, Washington, D.C., Mar. 22–23,1999, pp. 176–180.

[20] S. Ben-Yacoub, "Multimodal data fusion for personauthentication using SVM," in *Proc. 2nd Int. Conf. Audio-Video Based Biometric Person Authentication*, Washington,D.C., Mar. 22–23, 1999, pp. 25–30.