Design of Multimodal Biometric System for Strengthening the Security Systems

Anurag Bajpai[1] Umang Varshney[2] Deepam Dubey[3]
[1,2,3]Department of Electronics Design & Technology
[1,2,3]NIELIT, Gorakhpur, India

Abstract— This paper is about biometric system with the use of iris and speech recognition technology. It primarily deals with enhancing the accuracy of the voice based biometrics system with including one more parameter i.e. iris recognition. MFCC (Mel Frequency Cepstral Coefficients) is used for feature extraction of the voiced signal and its matching is done with the help of ANN (Artificial Neural Networks). Hough transform is used for feature extraction of iris and pattern matching of iris feature vector is done with the Hamming distance calculation.

Key words: MFCC (Mel Frequency Cepstral Coefficients), ANN (Artificial Neural Networks), Hamming Distance, Hough Transform

I. INTRODUCTION

Biometrics refers to metrics related to human characteristics. Biometrics authentication or realistic authentication is used in computer science as a form of identification and access control [1]. With the help of speech a person can interact with the machine, e.g. computer rather than using any interface(s). The spectral properties of human voice vary with person to person such as the pitch, loudness and frequency. So it is useful to reject unauthorized users. The iris of a person is same throughout the life so it can be used for the second biometrics parameter of the security system.

It has always been a bit of a tedious task to implement an efficient biometrics system mainly due to the following reasons:

- There is a voiced part and an unvoiced part in the speech.
- Degradation in speech signal due to noise.

II. BLOCK DIAGRAM OF PROPOSED MODEL

The proposed model consists of two subsystems, first model for the feature extraction of the voice features and matching of the features and the second model will have the feature encoding and the pattern matching of the iris data.

After the combined results of both the subsystems the system will decide if the given data is matched to the previous recorded authorized user's data.

The proposed block diagram is shown in the figure.

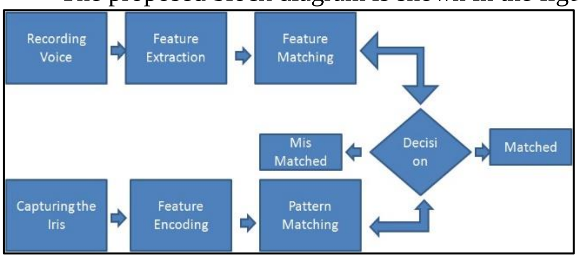

Fig. 1: Block diagram for proposed model

III. VOICE SUBSYSTEM

The voice subsystem consists of three blocks which are as follows:

Fig. 2: Block diagram for voice subsystem

A. Recording of the voice

The voice signal is recorded with the help of voice recorder ISD 1820 sound recorder .The sampling rate of the recording system is 16000 samples/sec to avoid aliasing effect in the recorded signal.

B. Feature Extraction

Feature extraction is done to capture the spectral characteristics of the speech signal. Those features are extracted which are somewhat invariant to changes in the speaker. Some of the feature extraction techniques are

- Power Spectral Analysis
- Linear Predictive Analysis (LPC) [4]
- Mel Frequency Cepstral Coefficients (MFCC) [2]
- Mel Scale Cepstral Analysis[3]
- Linear Predictive Cepstral Coefficients (LPCC)
- Relative Spectral Filtering of Log domain coefficients (RASTA) [5]

In this system MFCC is used for the feature extraction of the recorded voice.

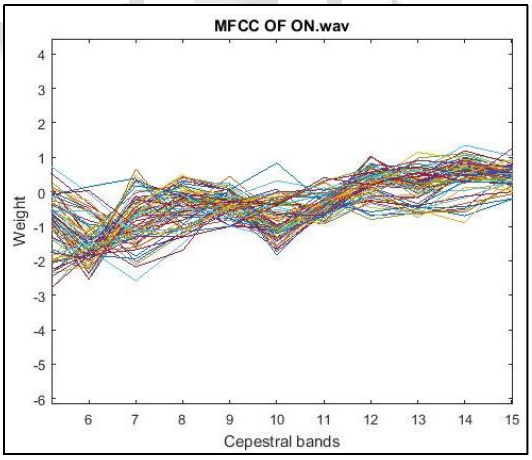

Fig. 3: MFCC of the voice

C. Feature Matching (Pattern Matching)

The second step after feature extraction of voice signal is to train the model with the help of a training vector. For training and testing purposes we can use a variety of models. Some of them are:

- Hidden Markov Models (HMM) [6]
- Artificial Neural Networks (ANN) [7]
- Correlation Method
- Gaussian Mixture Model (GMM) [9]

Artificial Neural Networks (ANN) is used in the pattern matching of the extracted voice feature vector. For learning of the neural network some algorithms which are used in simulation are as follows:

- Polak-Ribiere Conjugate Gradient (CGP)
- Resilient Back propagation (RP)
- Conjugate Gradient with Powell \ Beale Restarts
- Scaled Conjugate Gradient Back Propagation

IV. IRIS SUBSYSTEM

The second subsystem of the designed model is to take the image of the iris of the person and do feature encoding and pattern matching to compare the test iris image to the trained iris image. The flow chart for the iris subsystem is as follows:

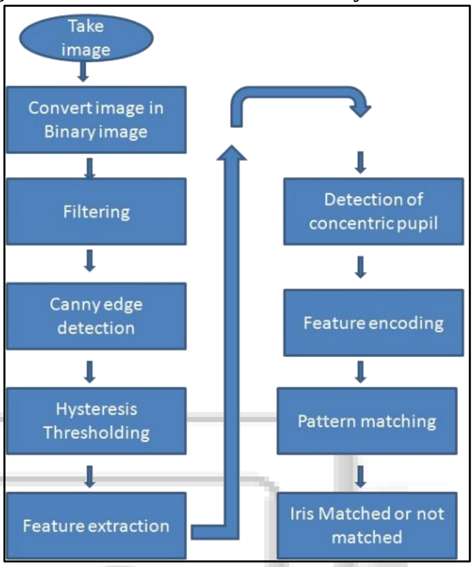

Fig. 4: Flow chart for iris recognition

The iris feature extraction and matching is done in the following steps:

1) Step 1: The iris of the person is captured through the camera.
2) Step 2: The image is converted into the binary image. As we are doing the feature extraction on the basis of intensity, we can also convert the image into the gray scale image.
3) Step 3: Now the canny edge detection of the image is done with the help of canny algorithm with use of non-maximum suppression [10].

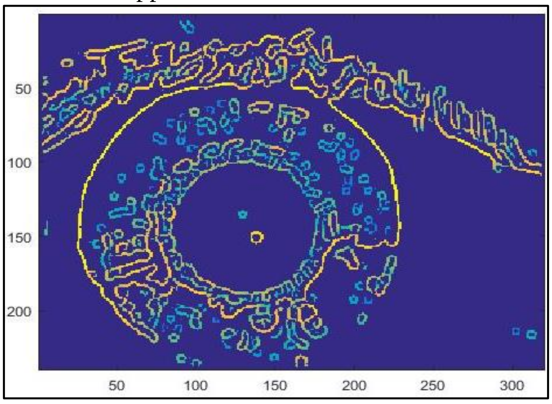

Fig. 5: Non maximum suppression

4) Step 4: After the non maximum suppression the hysteresis thresholding is done of the image.

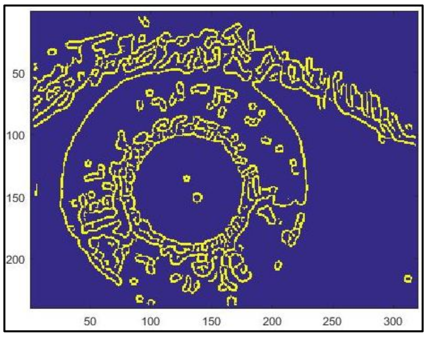

Figure.6. Hysteresis Thresholding

5) Step5: The feature extraction of the image is done with the circular Hough transform in which a novel approach is used to reduce the time taken in transform.
6) Step6: Now we detect the concentric pupil and iris boundary and draw the circle.
7) Step7: The feature encoding is done with the use of Gabour filter or 2D Gabor filters.
8) Step8: Now the pattern matching is done with the help of Hamming distance. If the distance is greater than a specified value then the output mismatched is shown.

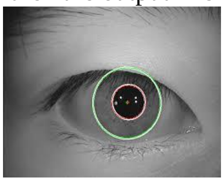

Fig. 7: Concentric pupil and iris boundary

V. EXPERIMENTAL SETUPS

The system was trained with the voice data of two persons with some specific command words like "OPEN" and "CLOSE" and the iris data captured in different environmental conditions due to which system can be trained in different intensity levels. The voice feature extracted data was tested in the neural network. The diagram of neural network is shown in the following figure:

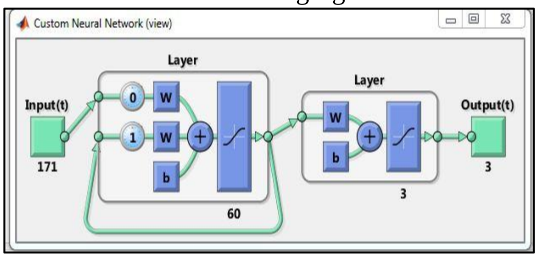

Fig. 8: Neural network model

The network gives a good performance when it is trained with limited number of person's data set. The performance plot of the voice subsystem is shown below:

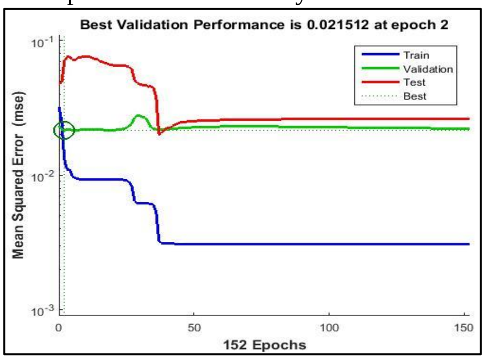

Fig. 8: Performance of network

VI. CONCLUSION

Due to the use of iris subsystem the voice biometric security [11] system has improved performance. As a novel approach was developed to reduce the time taken in Hough transform the system processing time is improved and it gives better results in less amount of time. The system can be used for real time biometric security systems with a specific number of users. The performance of the system can be further improved by increasing the number of training data set of the users.

REFERENCES

[1] "Biometrics: Overview". Biometrics.cse.msu.edu. 6 September 2007. Archived from the original on 7 January 2012. Retrieved 2012-06-10..
[2] Santosh Gaikwad, Bharti Gawali, and Pravin Yannawar, "Performance Analysis of MFCC & DTW for Isolated Arabic Digit", International Journal of Advanced Research in Computer Science, Vol. 2 (1), pp. 513-518 Jan. –Feb, 2011.
[3] S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 28, no. 4, pp. 357–366, August 1980.
[4] Young-Giu Jung, Mun-Sung Han, and Sang Jo Lee, "Development of an Optimized Feature Extraction Algorithm for Throat Signal Analysis", Electonics and Telecommunication research Institute Journal, Volume 29, Number 3, pp.292-299, June 2007
[5] Ram Singh, and Preeti Rao, "Spectral Subtraction Speech Enhancement with RASTA Filtering", Proceeding of National Conference on Communications (NCC), Kanpur, India, 2007.
[6] Lawrence R. Rabiner,"A tutorial on Hidden markov models and selected applications in speech recognition"
[7] D. Yu, K. Yao, H. Su, G. Li, and F. Seide, "KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition," In Proc. ICASSP'13, pp. 7893– 7897, 2013.
[8] Z. András, K. Daniil, S. Ralf and N. Hermann, "Using Multiple Acoustic Feature Sets for Speech Recognition", Speech Communication, Vol. 49, No. 6, 2007, pp.514-525.
[9] H. Doi, K. Nakamura, T. Toda, H. Saruwatari, K. Shikano, "Statistical approach to enhancing esophageal speech based on Gaussian mixture models," Proc. ICASSP2010, pp. 4250–4253 (2010)
[10] J. G. Daughman, "How iris recognition works," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, 2004.
[11] M. Pradhan, "Next Generation Secure Computing: Biometric in Secure E-transaction," International Journal of Advance Research in Computer Science and Management Studies, vol. 3, no. 4, pp. 473-489, 2015.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.