MULTIMODAL HUMAN VERIFICATION FOR FACE AND SPEECH

Madhushri Hiremath[1], Mr.Dinesh Kumar D.S[2]
[1]M.Tech Scholar, [2]Associate.Prof, Dept of ECE
East Point Institute of Technology, Bangalore, India

Abstract: Human biometric characteristics are unique, so it can hardly be duplicated (Kong et al.2005). Such information includes; facial, speech, hands, body, fingerprints, and gesture to name a few. Face detection and recognition techniques are proven to be more popular than other biometric features based on efficiency and convenience (Kriegman et al. 2002; Liu et al. 2002). It can also use a low-cost personal computer (PC) camera instead of expensive equipments, and require minimal user interface. Face authentication has become a potential a research field related to face recognition. Face recognition differs from face authentication because the former has to determine the identity of an object, while the latter needs to verify the claimed identity of a user. Speech (Gu and Thomas 1999) is one of the basic communications, which is better than other methods in the sense of efficiency and convenience. Each a single biometric information, however, has its own limitation. For this reason, we present a multimodal biometric verification method to reduce false acceptance rate (FAR) and false rejection rate (FRR) in real-time.

I. INTRODUCTION

Multimodal biometric framework, brushing two matchers face and voice, is proposed. With the proposed approach some biometric monomials verification frameworks confinements has been and the others, so the utilized combination system intended to based on entirety and item govern combinations. Further, we change the two fold sigmoid standardization strategy to enhance the multimodal framework execution. The created approach can be embraced with general multimodal validation frameworks including distinctive biometric highlights. Biometric frameworks consequently decide or check a man's character taking into account his anatomical and behavioural qualities, for example, unique mark, palm print, vein example, face and iris. A technique for distinguishing or checking the personality of a unique individual or subject in light of the physiological and behavioural attributes is biometric acknowledgment. Multimodal biometrics build exactness by considering other very particular organic attributes to restrain the number of candidate for a character. Multimodal biometric frameworks use more than one physiological or behavioural trademark for enlistment, check and recognizable proof. The motivation to consolidate distinctive modalities is to move forward acknowledgment rate. The multimodal biometrics aim is to reduce the following.

- False Accept Rate [FAR].
- False Reject Rate [FRR].
- Failure to enroll rate [FTE].

II. RELATED WORKS

A literature review discusses published information in a particular subject area. A literature review can be just a simple summary of the sources, but it usually has an organizational pattern and combines both summary and synthesis. The focus of a literature review, however, is to summarize and synthesize and the argument and ideas of others without adding new contributions.

In [2]S.R.Soruba Sree, Dr. N.Radha, "A Survey on Fusion Techniques for Multimodal Biometric Identification". In this paper various image fusion techniques are studied and their performances are evaluated on three criteria: Score, Decision and feature fusion.

In[3]Mohamed Soltane, Noureddine Doghmane, Noureddine Guersi, "Face and Speech Based Multi-Modal Biometric Authentication". The paper has presented a human authentication method combined face and speech information in order to improve the problem of single biometric authentication, since single biometric authentication has the fundamental problems of high FAR and FRR.

In[4]Y. M. Fouda, "Fusion of Face and Voice: An improvement". This paper examines the performance of multimodal biometric system using face and voice matchers on 140 individuals. The performance of sum rule, product rule, minimum rule, and maximum rule fusion on multimodal system (face and voice) has been evaluated.

In [5] Rahul Kala, Harsh Vazirani, Anupam Shukla, Ritu Tiwari, "Fusion of Speech and Face by Enhanced Modular Neural Networks". In this paper we took the problem of fusion of face and speech. Using this problem as a means we studied a good method of reducing dimensionality in the problem that was causing effects to the performance of the system

In[6]Samir Akrouf, Belayadi Yahia, Mostefai Messaoud and Youssef chahir, "A Multi-Modal Recognition System Using Face and Speech". This paper provides results obtained on a multi-modal biometric system that uses face and voice features for recognition purposes. We used fusion at the decision level with OR and AND operators.

In[7]Changhan Park and Joonki Paik, "Multi-Modal Human Verification Using Face and Speech". In this paper, a human verification method using combined face and speech information in order to improve the problem of single biometric verification has been presented. Single biometric verification has the fundamental problems of high FAR and FRR. Proposed a PCA for face recognition and HMM for speech recognition for real-time personal verification. As a result the proposed verification method can provides stable verification rate, and it overcomes the limitation of a single

mode system.

III. PROPOSED APPROACH

Proposed block diagram is show in the fig.1. Considering the template images and template speech. Extracting the face feature, unique variation of intensity or pixel will be unique. Here we have taken Histogram Base feature extraction technology for identifying the unique behavior of single person and also we are extracting the speech feature of a single person in different vibrations (Pitch) by using Singular Spectrum analysis. Both face and speech features will be fused in fusion algorithm then store it in the data base. Take input image and speech from the real time, extract the features then compare with data base. If it is matching person is identified alert system get triggered otherwise rejected

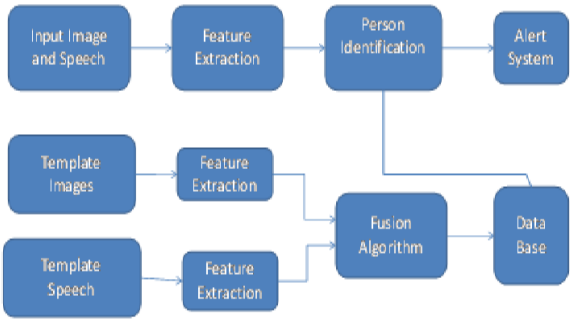

Figure 1: Block diagram of proposed system

A. FACE DETECTION

Face recognition is such an integral part of our lives and performed with such ease that rarely stop to consider the complexity of what is being done. It is the primary means by which people identify each other and so it isnatural to attempt to 'teach' computers to do the same. The applications of automated face recognition arenumerous: from biometric authentication; surveillance to video database indexing and searching. Face recognition systems are becoming increasingly popular in biometric authentication as they are non-intrusive and do not really require the users' cooperation. However, the recognition accuracy is still not high enough for large scale applications and is about 20 times worse than fingerprint based systems.

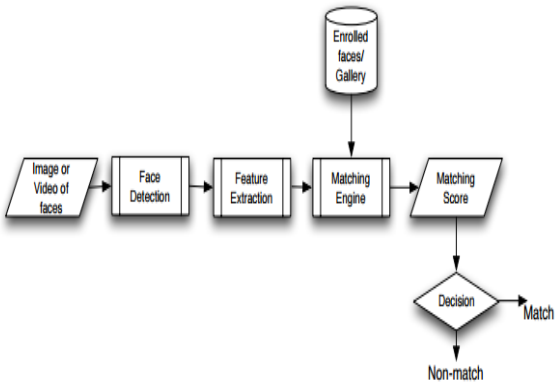

Figure 2: The basic flow of a recognition system.

Face recognition systems are examples of the general class of pattern recognition systems, and require similar components to locate and normalize the face; extract a set of features and match these to a gallery of stored examples, figure 1. An essential aspect is that the extracted facial features must appear on all faces and should be robustly detected despite any variation in the presentation: changes in pose, illumination, expression etc. Since faces may not be the only objects in the images presented to the system, all face recognition systems perform face detection which typically places a rectangular bounding box around the face or faces in the images. This can be achieved robustly and in real-time.

FLOW CHART:

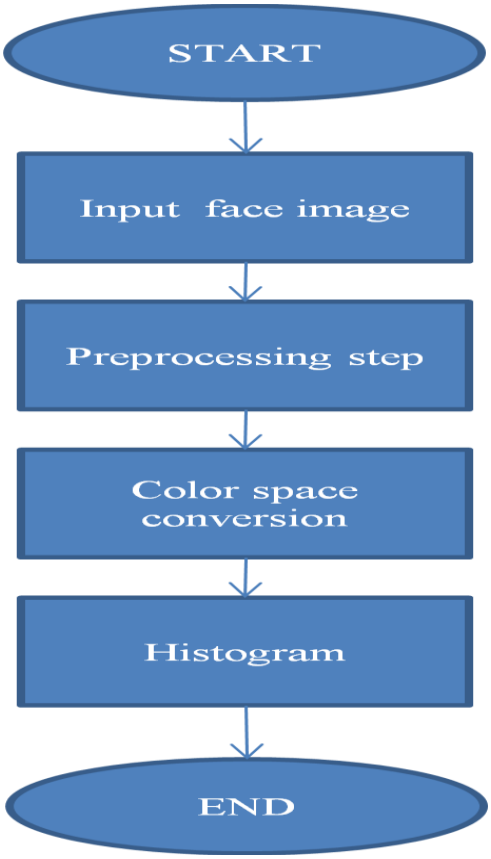

Figure .3: Flow chart

Face recognition flow chart shown in the figure 2. 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 obvious and common way to change the size of an image is to resize or scale an image. The content of the image is then enlarged or more commonly shrunk to fit the desired size. But while the actual image pixels and colors are modified, the content represented by the image is essentially left unchanged. The input image is resized to a new vale i.e. 250x250, 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. A histogram is a

graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable.

B. SPEECH DETECTION.

The speech is primary mode of communication among human being and also the most natural and efficient form of exchanging information among human in speech. So, it is only logical that the next technological development to be natural language speech recognition for HCI. Speech Recognition can be defined as the process of converting speech signal to a sequence of words by means Algorithm implemented as a computer program. Speech processing is one of the exciting areas of signal processing. The goal of speech recognition area is to developed technique and system to develop for speech input to machine. Based on major advanced in statically modeling of speech ,automatic speech recognition today find widespread application in task that require human machine interface such as automatic call processing. Since the 1960s computer scientists have been researching ways and means to make computers able to record interpret and understand human speech. Throughout the decades this has been a daunting task. Even the most rudimentary problem such as digitalizing (sampling) voice was a huge challenge in the early years. It took until the 1980s before the first systems arrived which could actually decipher speech. Off course these early systems were very limited in scope and power. Communication among the human being is dominated by spoken language, therefore it is natural for people to expect speech interfaces with computer .computer which can speak and recognize speech in native language. Machine reorganisation of speech involves generating a sequence of words best matches the given speech signal. Some of known applications include virtual reality, Multimedia searches, auto-attendants, travel Information and reservation, translators, natural language understanding and many more Applications.

FLOW CHART

A flowchart is a type of diagram that represents an algorithm, workflow or process, of the project. Fig.3 shows the flow chart of the algorithm, input audio signal is feed to the algorithm. Algorithm verifies the input signal with database where the different audio signal are present, if the input matches with any signal it will display input signal is matched else not matched. The input signal is applied to adaptive filter to remove the noise present in the audio signal. Initially input audio signal is in time domain 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.

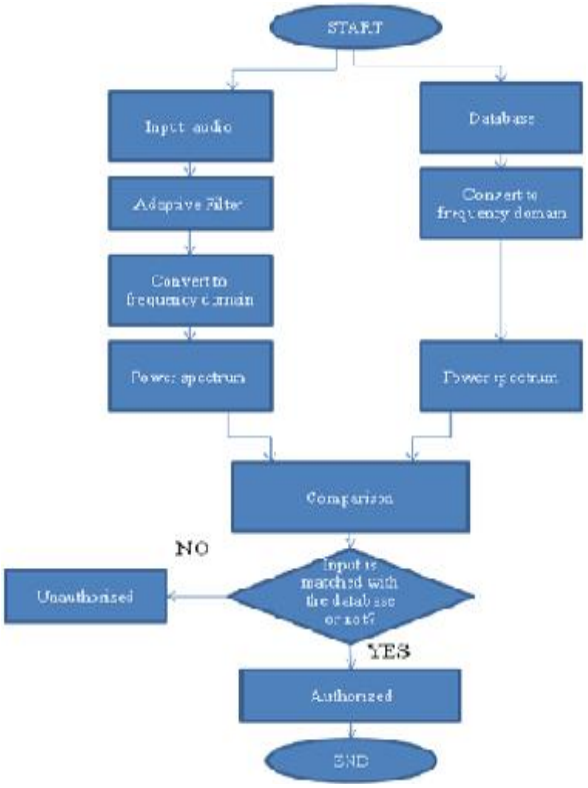

Fig .4: Flow chart

IV. EXPERIMENTAL RESULTS

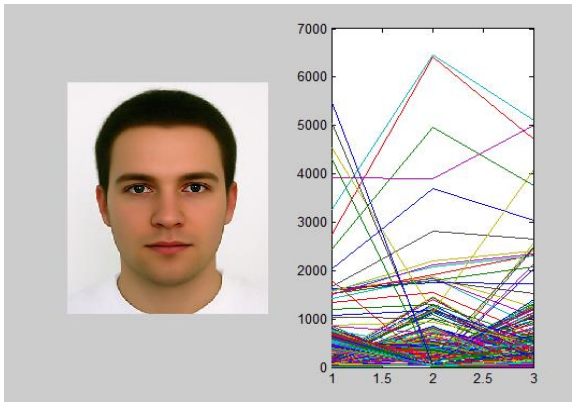

Figure 5: Input image and its Histogram plot

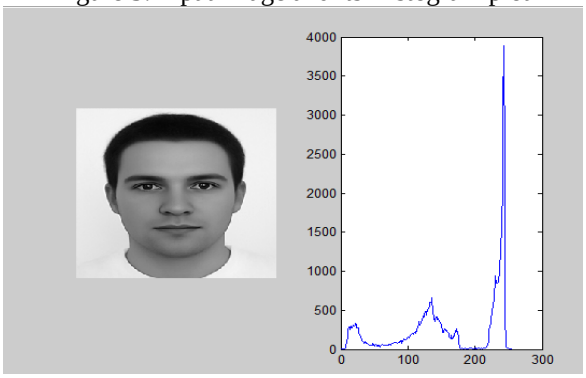

Figure 6: Gray converted image and its Histogram plot

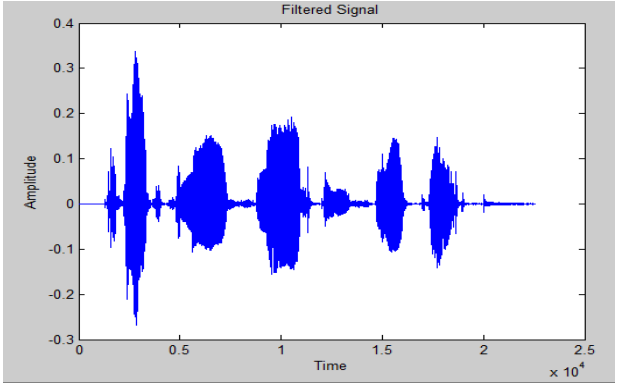

Figure 6: Adaptive Filter output

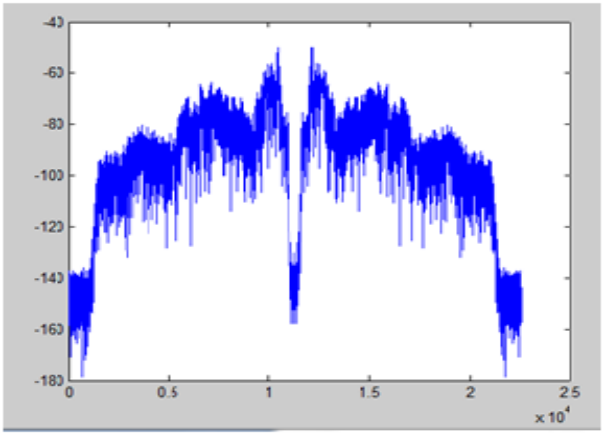

Figure 7: Power Spectrum of filtered signal

V. CONCLUSION

Human verification method using combined face and speech information in order to improve the problem of single biometric verification. Singlebiometric verification has the fundamental problems of high FAR and FRR. So we present a multimodal, biometric human verification method to improve the verification rate and reliability in real-time. We use PCA for face recognition and HMM for speech recognition for real-time personal verification. As a result the proposed verification method can provides stable verification rate, and it overcomes the limitation of a single mode system. Based on the experimental results, we show that FRR can be reduced down to 0.0001% in the human multimodal interface method using both face and speech information

REFERENCES

[1] S.R.Soruba Sree, Dr. N.Radha, "A Survey on Fusion Techniques for Multimodal Biometric Identification". International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 12, December 2014.

[2] Mohamed Soltane, Noureddine Doghmane, Noureddine Guersi, "Face and Speech Based Multi-Modal Biometric Authentication". International Journal of Advanced Science and Technology Vol. 21, August, 2010.

[3] Y. M. Fouda, "Fusion of Face and Voice: An improvement". IJCSNS International Journal of Computer Science and Network Security, VOL.12 No.4, April 2012.

[4] Rahul Kala, Harsh Vazirani, Anupam Shukla, Ritu Tiwari, "Fusion of Speech and Face by Enhanced Modular Neural Networks". Proceedings of the Springer International Conference on Information Systems, Technology and Management, Bankok, Thailand, pp 363-372.

[5] Samir Akrouf, Belayadi Yahia, Mostefai Messaoud and Youssef chahir, "A Multi-Modal Recognition System Using Face and Speech". IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 1, May 2011 ISSN (Online): 1694-0814 www.IJCSI.org

[6] Belhumeur, P.; Hispania, J. and Kriegman, D. (1997) Eigenfaces vs fisherfaces: recognition using class specification linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, (July 1997) Page(s) :711-720, 0162-8828.

[7] Kim, Y. ; Park, C. and Paik, J. (2004). A new 3D active camera system for robust face recognition by correcting pose variation. Proceedings of International Conference Circuits and Systems, pp. 1482-1487, August 2004

