Performance evaluation of unimodal, bimodal and trimodal biometric identification system

Binsu C. Kovoor
Assistant Professor, Information Technology Department, School of Engineering, Cochin University of Science and Technology, CUSAT (P.O.) Kochi, India, PIN 682022
Mobile no. 0091 9847788551

ABSTRACT

Multimodal biometric systems that integrate the evidences presented by multiple sources of information addressed the limitation of the unimodal biometric systems with a variety of problems, namely, noisy data, intra-class variations, restricted degrees of freedom, non-universality and spoof attacks. The likelihood ratio based score fusion framework was developed specifically for the verification scenario where the goal is to decide whether an input sample belongs to the genuine or impostor class. Twenty users were randomly selected from the population. The various biometric traits, namely face, voice and signature of ten users were used to train the respective modules of the proposed identification system. As a general observation for three scenarios, the p-value in the ANOVA analysis conclude that the fusion of three modalities brings systematically a clear improvement of the results in comparison to either unimodal and bimodal.

Key words: Multi modal, security, Signature, Biometric fusion

INTRODUCTION

The verification of identity of users, using human body characteristics or traits is more accurate and reliable than regular non biometric methods. Unlike in the case of passwords or identity cards, the biometric traits cannot be forgotten or lost. The factors affecting the selection of a particular trait in a biometric authentication are universality, uniqueness, permanence, measurability, performance, acceptability and circumvention [1]. A single biometric trait may not be sufficient to meet the requirements of protection from hacking. Multimodal biometric combines multiple traits from a user for authentication and thereby overcome the limitations of unimodal biometric systems.

A biometric system is essentially a pattern recognition system that operates by acquiring biometric data from a user [2]. In this paper, the performance of a multimodal biometric system using face, speech and signature traits is described. The face and signature traits are captured using a dedicated image capturing camera and the speech trait is captured using a microphone. The main advantage of using face, speech and signature traits is that they can
be collected easily or with minimum resistance from the users. The traits are collected sequentially, processed and authenticated based on Euclidean score.

The details of research work reported in the literature on multimodal biometric system using face, speech or signature traits are reviewed. The combinations of several biometric traits are used to build up multimodal identification or authentication systems. Bruneli (1995)[3] suggested that face and speech are the most popular traits used in biometric authentication systems. Hong (1998)[4] proposed a bimodal system using face and fingerprint traits. The recognition accuracy of this system was found to result in better recognition accuracy when compared to the corresponding data of the respective unimodal systems. Ghattis et al (1999)[5] explored the effectiveness of a biomodal system using speech and face traits. Frischholz et.al (2000)[6] developed a trimodal biometric system with face, speech and lip movement. Face and speech are static traits while the lip movement is a dynamic trait. Jain (2004)[1] discussed the strength and weakness of the various physiological and behavioural traits used in a computer aided authentication system. B.Due (1997)[7] cited that popularity and acceptability of the audio visual multibiometric systems are remarkable.

**BIOMETRIC TRAITS AND METHODS**

A multimodal biometric identification system using face, voice and signature traits is proposed and shown in Fig 1. The use of multiple traits results in greater protection against fraudulent attacks. The different input traits are captured using separate sensors. The sensor module, feature extraction module, matching module, database and normalization module corresponding to each of the traits are different. The score of the individual trait is combined in the fusion module and passed to the decision module. The identity of the user is accepted or rejected based on the decision criteria on the combined score in the decision module. The different sensor-different trait multimodal identification system is expected to be more robust, and accurate.

**Sensor module**

The face and signature images were captured using web camera (WC) with still image resolution 8.0 Mega Pixel. The microphone (MP) having a capturing frequency range of 100 to 16,000 Hz with sensitivity of 4dB is used to capture the speech signal.

**Feature extraction module**

In the feature extraction module, the measurable components are quantified from input traits. The features of the face image, speech signal and signature image are extracted using Principal Component Analysis (PCA), Mel Frequency Cepstrum Coefficients (MFCC) and Gabor Wavelet Transform (GWT) technique respectively.

Principal Component Analysis (PCA) is a mathematical procedure for computing the principal components of a set of observations as given in Fig 2.
Fig. 1 Component modules in multimodal biometric system with face, voice and signature

![Diagram](image)

Fig. 2 Face Feature Extraction Blocks

Mel Frequency Cepstrum Coefficient (MFCC) of speech signal is useful for identifying a person. The cepstral representation of speech signal provides unique characteristics of the spectrum. The MFCC processor blocks shown in Fig 3 consists of pre-emphasising, windowing, fast fourier transform, Mel-scale filter bank and MFCC vector computation.
The captured image of the signature is pre-processed prior to the feature extraction process. Image binarization, normalization and image enhancement are done in the pre-processing stage and is given in Fig 4.

**Matching module**

The feature vector of the user is compared with the feature vector enrolled in the database. Mahalanobis distance [8] accounts for differences in variance as well as correlation between vectors. The similarity between the feature vector and enrolled vector and is quantified by Mahalanobis distance.

**Normalization module**

In Min-Max (MM) normalization, the original distribution of the scores is mapped into the range [0,1] without distortion. Min-Max normalization is a straightforward approach when compared to Z-score and tanh [9] normalization technique. Min-Max (MM) normalization is used in the present study for face, speech and signature traits, and is given by

\[
S' = \frac{S_i - \min(S_i)}{\max(S_i) - \min(S_i)}
\]

where \{S_i\} is the score of the trait.

**Fusion Module**

Rank level fusion is adopted in this study. The rank for the normalized score is computed for face, speech and signature trait and is given by
where $S'_{\text{total}}$ is the average success ratio of a trait when used in a unimodal system. The total score $S'_{\text{total}}$ is computed by

$$S'_{\text{total}} = \sum r_i S'_i$$

\section*{Decision module}

The magnitude of trimodal fusion matching score is the separating hyper plane in the decision rule. The decision rule is given by

$$\text{Decision} = \begin{cases} \text{Accept, if } S'_{\text{total}} > 0.99 \\ \text{Reject, otherwise} \end{cases}$$

\section*{Characterization of biometric identification system and Experimental program}

The performance of the multimodal biometric identification system is evaluated using the parameters such as Success Ratio (SR), False Rejection Ratio (FRR) and False Acceptance Ratio (FAR). The success is defined as the approval of known person or rejection of unknown person. False rejection is defined as the rejection of known person, and false acceptance is defined as the approval of unknown person. The high magnitude of the SR or low magnitude of FRR and FAR indicate improved performance of the biometric identification system. The low magnitude of FRR or FAR can be obtained by arbitrarily adjusting the decision threshold, with a drawback of increasing the other value[10].

The biometric traits are submitted to the unimodal, bimodal and trimodal identification system and the details are given in Table.1. The magnitude of SR, FAR and FRR corresponding to unimodal, bimodal and trimodal system is analyzed statistically using ANOVA. The variance between the parameter vectors is determined and compared.

| Biometric traits | Face   | Speech | Signature | Face + Speech |
|------------------|--------|--------|-----------|--------------|
| Face             | Unimodal | --     | --        | --           |
| Speech           | Bimodal | Unimodal | --       | --           |
| Signature        | Bimodal | Bimodal | Unimodal  | Trimal       |

Twenty users were randomly selected from the population. The various biometric traits, namely face, voice and signature of ten users were used to train the respective modules of the proposed identification system. The characterization of performance of the
identification system is possible by permitting the predefined number of attempts of ‘known’ users and ‘unknown’ users. The ten users were ‘known’ and the remaining ten users were ‘unknown’ to the identification system. All the twenty users were allowed to submit the biometric traits in random order once in a day for twenty days. Every day, the users were allowed to use different costumes, facial pack, ornaments, hair style, pen for signing etc. Thereby the possible variability was accounted for in the proposed model. In the first five days, the biometric traits of the selected ten users were submitted to the training modules of the proposed identification. The face traits and signature were captured using dedicated web cameras and the speech trait was captured using microphone. The identity traits submitted by all the twenty users during the subsequent fifteen days were used for evaluating the performance of the system.

RESULTS AND DISCUSSIONS

The success, false acceptance and false rejection of the user identity were recorded for unimodal, bimodal and trimodal biometric systems. The statistical details of the Success Ratio (SR), False Acceptance Ratio (FAR) and False Rejection Ratio (FRR) of unimodal, bimodal and trimodal biometric systems are given in Table 2. The mean value success ratio of trimodal is greater than the bimodal. Similarly the success ratio of bimodal is greater than unimodal system. The mean value of false acceptance ratio and false rejection ratio of trimodal is less than bimodal. Similarly, the false acceptance ratio and false rejection ratio of bimodal is less than that of the unimodal system. The magnitude of false acceptance ratio and false rejection ratio depends upon the proportion of trait submission attempts made by ‘known’ users and ‘unknown’ users. A higher magnitude of the success ratio and lower magnitude of rejection ratio is an indication of better performance.

Table 2. The statistical details of SR, FAR and FRR of unimodal and multimodal system

| Biometric system parameter | Statistical parameter | Biometric identification system | Unimodal system | Bimodal system | Trimodal system |
|----------------------------|-----------------------|--------------------------------|----------------|---------------|----------------|
| Success ratio (SR)         | Mean                  | Face                          | 0.54           | 0.68          | 0.81           |
|                            |                       | Speech                        | 0.56           | 0.67          | 0.77           |
|                            |                       | Signature                     | 0.54           | 0.67          | 0.77           |
| False acceptance ratio (FAR)| Mean                | Face + Speech                 | 0.45           | 0.30          | 0.33           |
|                            |                       | Speech                        | 0.47           | 0.35          | 0.33           |
|                            |                       | Signature                     | 0.49           | 0.35          | 0.33           |
| False rejection ratio (FRR)| Mean                  | Face + Speech + Signature     | 0.47           | 0.34          | 0.18           |
|                            |                       | Speech + Signature            | 0.41           | 0.31          | 0.18           |
|                            |                       | Signature                     | 0.43           | 0.34          | 0.21           |
|                            | SD                    | Face                          | 0.05           | 0.07          | 0.05           |
|                            |                       | Speech                        | 0.09           | 0.05          | 0.07           |
|                            |                       | Signature                     | 0.08           | 0.05          | 0.07           |
The influence of various treatments namely, unimodal, bimodal and trimodal system was determined statistically by carrying out the ANOVA test. The box plot indicating the minimum, maximum, mean and the standard deviation of Success Ratio, False Acceptance Ratio and False Rejection Ratio of the biometric identification system is presented in Fig 5. The standard deviation and the range of minimum to maximum are greater for biometric system using speech trait. This may be attributed to the distortion of speech signal in presence of greater noise from the surrounding.

![Box plot for SR, FAR, FRR of biometric systems](image)

**Fig 5**: Box plot for the SR, FAR, FRR of the biometric identification systems

**CONCLUSION**

A multimodal biometric technique that can be devised to access control equipment is developed. The technique combines multiple biometrics in making a personal identification, can be used to overcome the limitations of individual biometrics. In this paper a multimodal biometric system, which integrates decision made by speaker
recognition, face recognition, and handwritten signature recognition to make a personal identification, is developed. Future works could go in the direction of using more robust modeling techniques against forgeries and fusion at the feature extraction level can be used. More than three modalities can be used together to make forgeries more difficult. In addition, the system should be tested on a larger database to validate the robustness of the model.

REFERENCE

[1] AK Jain, Ross A, Prabhakar S., An introduction to biometric recognition, *IEEE transactions on circuits and systems for video technology*, Vol.14(1), pp.167-172, 2004

[2] A.K Jain, Bolle, R., Pankanti, S., Biometrics: Personal Identification in Networked Society, *Kluwer Academic Publications*, 1999.

[3] Bruneli R and Falavigna D., Person Identification using multiple cues, *IEEE Trans PAMI*, Vol.17(10), pp.955-966, 1995.

[4] Hong L. and Jain A.K., Integrating faces and fingerprints for person identification, *IEEE PAMI*, Vol. 20(12), pp.1295-1307, 1998

[5] Ghattis V., Bors A.G and Pitas I, Multimodal decision level fusion for person authentication, *IEEE Trans Systems, Man and Cybernetics*, Vol.29(6), pp.674-680, 1999.

[6] Frischholz R.W. and Dieckmann U., Bioid: A multimodal biometric identification System, *IEEE Computer Society*, pp 64-68, 2000. Van Trees H. L., Detection, Estimation and Modulation Theory, *John Wiley & Sons*, New York, 1968.

[7] Due B, Fusion of audio and video information for multimodal person authentication, *Elsevier Pattern Recognition Letters*, Vol.18, pp.835-845, 1997.

[8] Mahalanobis, Prasanta Chandra, On the generalised distance in statistics, *Proceedings of the National Institute of Sciences of India*, Vol. 2 (1), pp.49–55, 1936.

[9] Jain A K, Nandakumar K., Ross A., A score normalization in multimodal biometric system, *Pattern Recognition Lett*, Vol.38, pp.2270-2285, 2005.

[10] Van Trees H. L., Detection, Estimation and Modulation Theory, *John Wiley & Sons, New York*, 1968.