ISSN 2229-5518

Score Level Fusion Based Personal Authentication Using Fingerprint and Speech

Praveen N, Tessamma Thomas

Abstract— In this paper development of a multimodal based biometric fusion system is discussed. A fingerprint recognition system is developed using global singularity features. Mel-frequency Cepstral Coefficients are used to recognise a speaker using the backpropagation artificial neural network. A score level fusion based recognition system is developed using fingerprint and speech match scores and the equal error rate (EER) measured shows a good improvement with 100% recognition rate is obtained for over a large span of match score threshold.

Index Terms— Biometrics, Fingerprint Recognition, Orientation Estimation, Mel-frequency Cepstral Coefficient, Artificial Neural Networks, Backpropagation, Score level Fusion, Sum Rule, Product Rule

1 Introduction

Biometric systems that use a single modality are usually affected by problems like noisy sensor data, non-universality and/or lack of distinctiveness of the biometric trait, unacceptable error rates, and spoof attacks [1]. A multibiometric system improves the performance of the system by considering several traits from different scores [2]. Multibiometric systems deals with two or more evidences are taken from different sources like multiple fingers of the same person, multiple samples of the same instances, multiple sensors for the same biometric, multiple algorithms for representation and matching or multiple traits. A multibiometric system that uses different biometric traits is expected to be more robust to noise, address the problem of non-universality, improve the matching accuracy and provide reasonable protection against spoof attacks and thus multibiometric system has received considerable attention among researchers [2].

Various levels of fusion are possible in a multibiometric system that uses different biometric traits: fusion at the features extraction level, matching score level or decision level. Feature level fusion is quite difficult to consolidate as the feature sets used by different modalities may either be inaccessible or incompatible. Fusion at the decision level is too rigid since only a limited amount of information is available at this level. Therefore integration of the matching score level is generally preferred due to the ease of in accessing and combining matching scores.

In the case of verification fusion at the matching score level can be approached in two distinct ways: in the first approach the fusion is viewed as a classification problem while in the second approach the fusion is viewed as a combination problem. In the classification approach, a feature vector is constructed using the matching scores output by the individual matchers; this feature is then classified into one of the two classes: "Accept" for a genuine user or "Reject" for an imposter [2]. In the other case the individual matching score are combined to generate a single scalar score which is then used to make the final decision. The combination approaches have shown to be the better approach by Ross and Jain [3]. Combination approach has been implemented in this work and modalities taken are fingerprint and speech.

2 Information Fusion in Multimodal Biometrics

According to Sanderson and Paliwal [4], information fusion can be classified into three main categories: Pre-mapping fusion, midst-mapping fusion and post-mapping fusion (Fig. 1). In pre-mapping fusion, information is combined before any use of classifiers or experts; in midst-mapping fusion, information is combined during mapping from sensor-data/feature space into opinion/decision space, while in post-mapping fusion, information is combined after mapping from sensor-data/feature space into opinion/decision space.

Pre-mapping fusion is categorized as sensor data level fusion and feature level fusion. In post-mapping fusion, there are two main sub-categories: decision fusion and opinion fusion. Ross and Jain [3] refer to opinion fusion as score fusion.

- *Dr. Praveen N is currently working in N S S College, Rajakumari, Idukki, Kerala, India as Associate Professor of Electronics E-mail: praveen.naniyat@gmail.com*
- *Prof. (Dr) Tessamma Thomas is currently working in Department of Electronics, Cochin University of Science and Technology, Kochi, Kerala, India as Professor. E-mail: tess@cusat.ac.in*

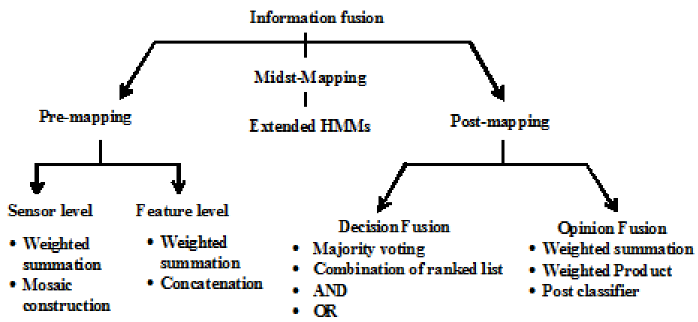

Fig 1 Information fusion- Various Levels

2.1 Post mapping fusion

This type of fusion is also referred to as score level fusion in which, an ensemble of experts provides an opinion on each possible decision [3][5][6]. Scores from each modality can be combined using weighted summation or weighted product approach. The main advantage of summation and product fusion is that the opinion from each expert can be weighted. Thus the weights can be set to reflect the reliability and discrimination ability of each expert [4]

2.1.1 Weighted summation fusion

In this method, the opinions regarding class *j* from N_E experts are combined using the expression:

$$f_j = \sum_{i=1}^{N_E} w_i o_{i,j} \quad (1)$$

where $o_{i,j}$ is the opinion from the i^{th} expert and w_i is the corresponding weight in the [0, 1] interval, with the constraints $\sum_{i=1}^{N_E} w_i = 1$ and $\forall i: w_i \geq 0$. This approach is also known as linear opinion pool [7] and sum rule [8][9].

2.1.2 Weighted Product fusion

Assuming the experts are independent, the opinions regarding class *j* from N_E experts can be combined using a product rule

$$f_j = \prod_{i=1}^{N_E} o_{i,j} \quad (2)$$

By introducing weighting to account for varying discrimination ability and reliability of each expert,

$$f_j = \prod_{i=1}^{N_E} (o_{i,j})^{w_i} \quad (3)$$

The weighted product approach is also known as logarithmic opinion pool [7] and product rule [8] [9].

In this work, a multimodal recognition system is developed based on fingerprint and speech modalities. The fingerprint recognition score is combined with speech based recognition score using sum and product rule. The FAR and FRR are plotted for various thresholds and EER is measured. The results are compared with the unimodal systems. Fig. 2 shows a bimodal biometric system showing the three levels of fusion: fusion at the feature extraction level, fusion at the matching score level and fusion at the decision level. In the figure, FU represents the fusion module, MM represents the matching module and DM represents the decision module.

3. IMPLEMENTATION

3.1 Fingerprint verification

Fingerprint verification involves extracting a feature set from a fingerprint image acquired by means of a good quality optical based fingerprint scanner with a resolution of 500 dpi and matching it with the template stored in the data.

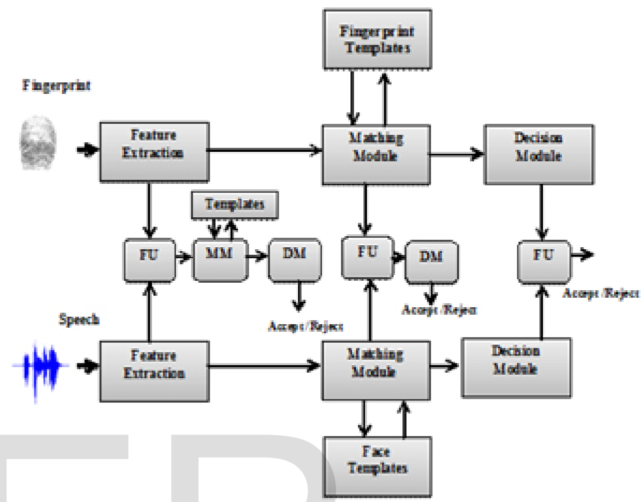

Fig 2. Bimodal biometric system showing three levels of fusion.

In this research a global singularity based fingerprint identification method has been used [10]. The feature set consists of 16 polygonal feature metrics and the methods used for feature extraction is given in fig. 3. The matching process involves the computation of a similarity measure using the distance between corresponding polygonal features in the input and feature template. A score of 1 gives a 100% matching and a score of 0 gives no matching. The threshold for a match can be fixed at 0.79 to achieve a 100% recognition rate and 0% EER as per the FAR-FRR curve shown in Fig. 4

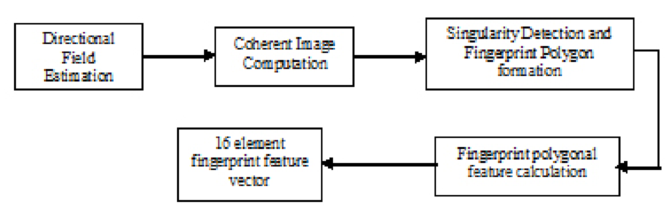

Fig 3. Fingerprint Feature Extraction

3.2 Text-Dependant Speaker Recognition

ISSN :

In text dependent speaker verification the speech samples of a known sentences or words were recorded using a good quality microphone in the laboratory environment. Feature extraction involves the computation of 13 MFC coefficients of frames which were clustered to give about a 13 × 5 element matrix [11]. These coefficients were input to a trained backpropagation based artificial neural network for classification/identification. The entire process is given in fig. 5. The neural network output ranges in between 0 to 1 in which a 0 output indicates the absence of speaker template in database

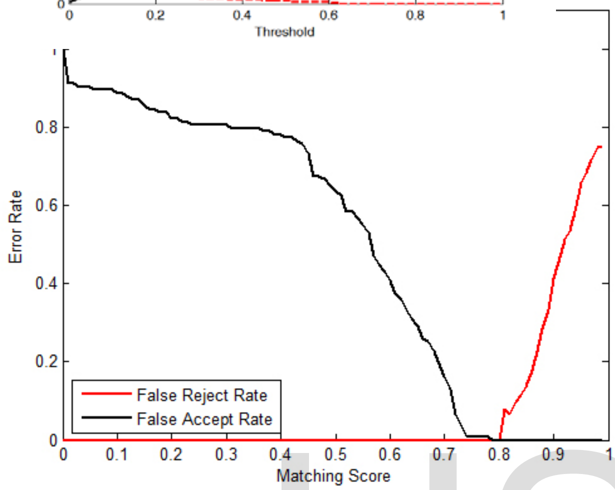

Fig 4. FAR-FRR Characteristics

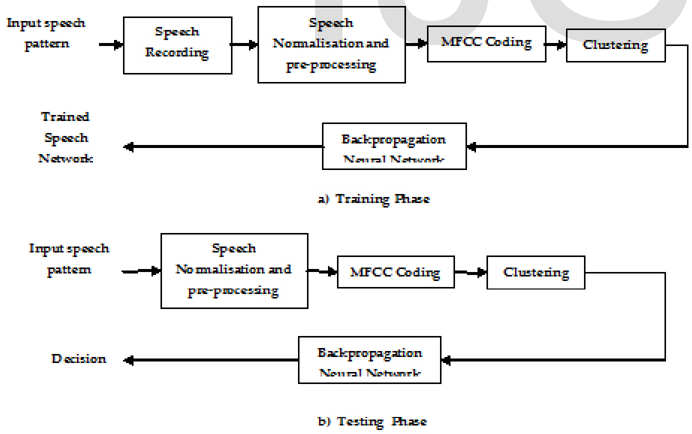

Fig 5. Speaker Recognition steps

and a score of 1 indicates the presence of template in the speaker database.

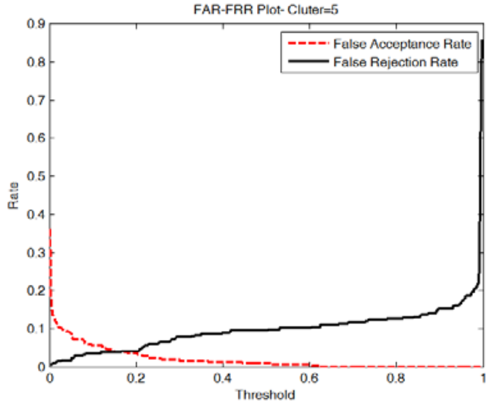

Fig 6. FAR FRR Characteristics EER=0.0382 at a Threshold of 0.152-0.167

Threshold can be fixed in between 0.152 and 0.167 as per the experiments described in [11]. The speaker is identified by the output neuron number which is triggered to a high output. The FAR-FRR characteristic curve for a cluster size of 5 is shown in Fig. 6, with the EER is approximately equal to 0.0382 within a matching threshold of 0.152-0.167. Thus the maximum genuine acceptance rate or recognition rate achieved for this system is 96.18%**.**

3.3 Combining the two modalities

The database for combining the two modalities consisted of matching scores of two different modalities- fingerprint and speech. The fingerprint and speech data were obtained from a user set consisting of 40 users. For each user, 6 fingerprints were acquired and 10 voice prints were recorded. Out of the 6 fingerprints, 3 fingerprint image features were extracted and the features were averaged to form the feature templates. Other 3 images of 40 users were used to generate the match scores using the distance based match score generation as per equation hence a 120 genuine scores were obtained for fingerprint modality and 5880 (120 × 49) imposter scores were obtained. Similarly, 10 speech samples of 40 users were recorded, out of which 7 were used for ANN training and 3 were used for generating the neural network output. Here also 120 genuine scores were obtained for speech modality and 5880 imposter scores were obtained.

3.3.1 Combining using Sum rule

In this combining method, weighted averages of the scores from the two modalities were taken. In this research the effect of weights on each modality combined together to give a score were investigated. Starting from an initial weight, w_i=0.1 the scores were calculated according to the sum rule:

$$M = m_f w + m_s(1 - w) \qquad (4)$$

ISSN 2229-5518

Where M is the combined match score, m_f is the match score for the fingerprint based recognition for the database taken as explained in the previous section, m_s is the match score for the speech based identification system and w is the weight taken which ranges from 0.1 to 0.9. For each weight, the False Accept Rate and False Reject Rate are calculated for various matching threshold. The FAR and FRR were calculated for each weights and plotted (Fig. 7). Table 1 gives the Equal Error Rates (EER) correspond to each of the above plots. The EER is zero for weights in the range 0.1-0.8 and for weights ≥ 0.9 the EER is greater than 0. At lower weights, the threshold range difference is large. The discriminative power of neural network classifier is predominant here even at lower weights.

3.3.2 Combining using Product rule

The match scores were combined using the product rule as per the rule[Altiçay, 2000] [Alexandre, 2001; Kittler, 1998]:

$$M = (m_f)^w . (m_s)^{1-w} \qquad (5)$$

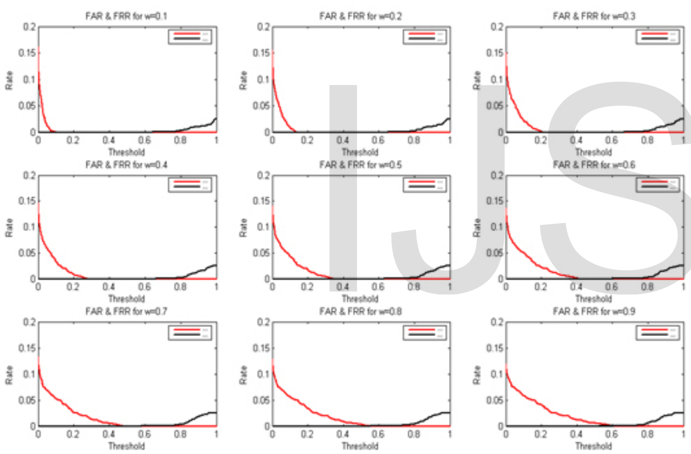

Fig 7 FAR and FRR Plots correspond to various weights taken with sum rule

Table 1 Weights and Threshold range for EER. For w =0.9, EER=6.24 × 10^{-4} and EER=0 for w<0.9

Weight	Threshold range	Difference
0.1	0.1-0.639	0.539
0.2	0.139-0.651	0.512
0.3	0.205-0.664	0.459
0.4	0.274-0.656	0.382
0.5	0.342-0.629	0.287
0.6	0.41-0.603	0.193
0.7	0.478-0.577	0.099
0.8	0.546-0.551	0.005
0.9	0.596-0.612	0.016

where M is the fused match score, m_f is the match score for the fingerprint based recognition, m_s is the match score for the speech based identification system and w is the weight taken which ranges from 0.1 to 0.9. Fig.8 give the ROC curves obtained using product rule and Table 2 show the corresponding weights and threshold range for EER.

From the two experiments on fusion using sum and product rule, EER was shown to be zero for a considerable range of threshold using sum rule with weight, w=0.1. The discriminative power of neural network classifier was quite evident from the plot as even for lower weights the EER was shown to be zero.

3.4 Discussion on Result

The experiments described above conclude that the sum rule performs better that the product rule. Even though in the sum and product rule the EER falls to zero for a reasonable range of thresholds and weights, the maximum range of thresholds for which the EER was zero was observed in the sum rule with weight, w=0.1 and the threshold range difference was about 0.539. Here a key limitation of the product rule was its sensitivity to errors in the estimation of the posterior probabilities. Even if one of the classifiers outputs a

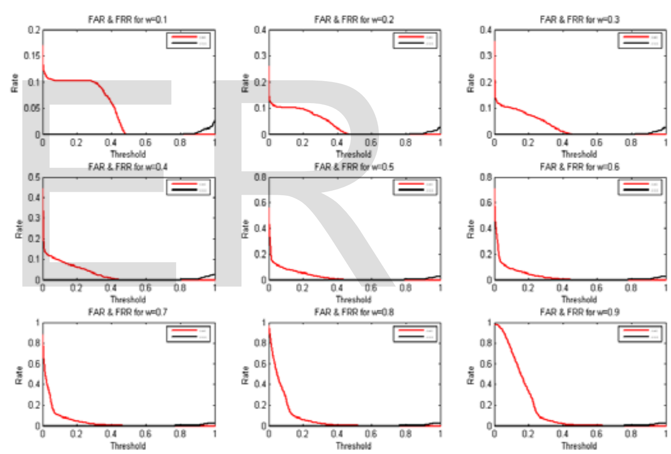

Fig. 8. FAR and FRR Plots correspond to various weights taken with product rule

Table 2 Weights and Threshold range for EER. EER=0 for all the shown weights

Weight	Threshold range	Difference
0.1	0.484-0.815	0.331
0.2	0.468-0.816	0.348
0.3	0.454-0.818	0.364
0.4	0.444-0.803	0.399
0.5	0.439-0.789	0.350
0.6	0.444-0.777	0.333
0.7	0.467-0.768	0.301
0.8	0.521-0.760	0.239
0.9	0.631-0.754	0.123

ISSN 2229-5518
probability close to zero, the product of the *M* posterior probabilities was rather small and this often leads to an incorrect classification decision. The sum rule was generally more effective than the product rule and by taking the weight, *w*=0.1-0.4 and by fixing the match score threshold around 0.4 the person can be effectively recognized as a genuine or imposter.

4. CONCLUSION

In this paper basic fusion techniques were discussed. Fingerprint recognition, which is one of the traditional Score level fusion based authentication system was developed using the sum and product rule. The sum rule was found to be more effective for the fusion of the scores in this system.

REFERENCES

[1] A.K. Jain, A. Ross, "Multibiometric systems", Commun. ACM 47 (1) (2004) 34–40 (special issue on multimodal interfaces)

[2] Jain, Anil, KarthikNandakumar, and Arun Ross. "Score normalization in multimodal biometric systems", Pattern Recognition, 38, 2270–2285, 2005

[3] A Ross, A Jain, "Information fusion in biometrics", Pattern Recogn.Lett. 24, 2003, 2115-2125.

[4] C. Sanderson, K.K. Paliwal, "Identity Verification Using Speech and Face Information", Digital Signal Processing, Vol. 14, No. 5, 2004, pp. 449-480.

[5] D.L. Hall, J. Llinas," Multisensor data fusion", D. L. Hall and J. Llinas (Eds.), Handbook of Multisensor Data Fusion, CRC Press, USA, 2001, pp. 1-1 - 1-10.

[6] S.S. Iyengar, L. Prasad, H. Min, "Advances in Distributed Sensor Technology", Prentice Hall PTR, New Jersey, 1995.

[7] H. Altiçay,M. Demirekler, "An information theoretic framework for weight estimation in the combination of probabilistic classifiers for speaker identification", Speech Commun. 30, 2000, 255–272.

[8] L.A. Alexandre, A.C. Campilho, M. Kamel, "On combining classifiers using sum and product rules", Pattern Recogn. Lett. 22, 2001, 1283–1289.

[9] J. Kittler, M. Hatef, R.P.W.Duin, J.Matas, "On combining classifiers", IEEE Trans. Pattern Analysis and Machine Intelligence 20 (1998) 226-239.

[10] Praveen N and Tessamma Thomas, "Singularity Based Fingerprint Identification", International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 2, No. 5, October 2011, ISSN: 2079-2557 Science Academy Publisher, United Kingdom, pp. 1155-1159

[11] Praveen N and Tessamma Thomas."A Text Dependent Speaker Recognition using MFCC features and BPANN", International Journal of Computer Applications 74(5):31-39, July 2013.