

Optimal Score Level Fusion using Modalities Reliability and Separability Measures

Anzar S.M.

Electronics and Communication Engineering
National Institute of Technology Calicut
India, Pin: 673601

Sathidevi P.S.

Electronics and Communication Engineering
National Institute of Technology Calicut
India, Pin: 673601

ABSTRACT

This paper presents two optimal fusion techniques using reliability and separability measures for a multibiometric system, employing fingerprints and voice. In the first method, reliabilities of fingerprint and voice modalities are measured and the integration weights are computed as the ratio of these two reliabilities. The computed reliability ratio is then optimized against the recognition accuracy. The optimizing parameter is estimated in the training/validating phase. Latter is a multi-normalization based fusion scheme, where the separability measures are used as the integration weights to improve the performance of the biometric system. The inter/intra class separability measures and the d-prime separability measures, under various noise conditions are estimated in the training/validation stage. Performance of the proposed methods are compared with that of the baseline techniques using score level fusion. Experimental studies show that, the proposed methods improve global recognition rate and reduce the False Acceptance Rate (FAR) and False Rejection Rate (FRR) over the baseline systems. The advantage of the proposed biometric techniques is that they can be easily integrated into any multibiometric system with score level fusion and find extremely useful in applications especially with less number of available training samples.

Keywords:

Reliability measures, Separability measures, Score level fusion, score normalization, noise robustness

1. INTRODUCTION

Data variations are considered as one of the main problems in multimodal fusion. In this paper, the focus is on improving the score level fusion of fingerprint and voice biometrics, under varying noise conditions [2]. Even though the recognition accuracy of the voice biometric is high in clean conditions, its weakness is mainly due to their inability to cope robustly with audio corruptions. Audio corruptions can arise from various sources such as environmental noises or voice interference, reverberation effects or transmission channel distortion etc. Voice features show large intra-class differences, inter-class similarities and may vary with environmental and physiological factors as well as time [10]. When a biometric measure obtained from one modality is corrupted by noise, the evidence presented by a comparatively reliable trait is used for the more accurate determination of identity. Noisy Biometric may not be successfully matched with the respective stored template in the database, resulting in genuine user being incorrectly rejected or an impostor accepted as a genuine [20]. A reliability ratio based integration weight optimization scheme is proposed here to determine the optimal weight factor for the complementary modalities, under

different noise conditions. As the feature vector varies with the amount of noise, the quality of score level fusion heavily depends on the reliability of the unimodal sources. Reliability measure gives the degree of trust in the recognition result drawn from individual subsystems [12]. The performance of biometric fusion can be improved if we could use the score-based reliability measures to find the best integration weight. Determining proper weights for the individual modalities is crucial for robust recognition as it determines the contribution of each modality towards the final decision. At high SNR (Signal to Noise Ratio), the voice matcher outperforms the fingerprint matcher and the final decision heavily relies on the score values of the voice matcher. When the voice biometric is contaminated by noise, the fingerprint matcher outperforms the voice matcher. In this case, the score values of the fingerprint matcher contribute more to the final decision. Therefore, it is crucial to estimate the best weight factor dynamically, to combine both the modalities for better performance, otherwise the system will show attenuating fusion [17]. Estimation of reliability measures and the fusion strategy are detailed in section 4. We estimated the optimizing parameter in the training/validation stage using Leave-One-Out Cross Validation technique (LOOCV). The optimal weight estimated in the training/validation stage is subsequently used in the fusion module for testing with the unknown input samples.

A weighting strategy, combining the separability measures from the feature space and the match score space is also proposed in this paper. In a biometric system the smaller the overlap between the impostor scores and the genuine scores, the better the recognition rate. As the amount of noise increases, the overlap between the genuine and the impostor score distribution will also increase. Thus the class separability and the score separability measures from the feature space and the matching score space respectively at different noise conditions give an indication of the quality of the biometric samples and the matcher. An efficient preprocessing on the raw vector of scores using multi-normalization (different score normalization techniques for the complementary modalities) is employed to improve the robustness and efficiency of the bimodal system under various noise conditions. The central idea behind score normalization is to reduce the data variations that are reflected in the matching scores [8]. The inter/intra class separability measures from the feature space and the d-prime separability measures from the match score space are estimated in the training/validating phase using LOOCV technique. The proposed fusion strategies are promising and it gives robust and improved recognition accuracy even at low SNR conditions.

2. RELATED WORK

Jain et al. in 1999 introduced a multimodal biometric system that integrates face, fingerprints and voice [9]. Kar-Ann Toh in 2003 proposed a generalized reduced multivariate polynomial model for combining fingerprint and speaker verification decisions [23]. Wang et al. in 2004 combined fingerprint and voiceprint biometrics [26]. Lewis et al. in 2004 shed some light on audio-visual speech recognition systems using dispersion measures as the integration weights [13]. Toh et al. in 2005 combined fingerprint and speaker verification decisions in the match score level using functional link network [24]. Poh et al. in 2005 proposed a margin derived confidence measure while fusing two system opinions [16]. Jain et al. in 2005 examined the effect of different score normalization techniques on the performance of multimodal biometric system [8]. Kryszczuk et al. in 2007 proposed a method of performing multimodal fusion using face and speech data combining signal quality measures and reliability estimates [11]. Sarkar et al. in 2008 fused fingerprint, face, voice and gait [21]. Bendris et al. in 2009 introduced quality measures in audio-visual identity verification [5]. Alsaade et al. in 2009 showed that score normalization and quality-based fusion improves the accuracy of multimodal biometrics [3]. Optimal integration weight estimation using least squares technique was reported in [23]. Reliability based optimal integration weight estimation for audio-visual decision fusion was reported in [17, 18]. We were motivated by [8, 10, 17] to develop a bimodal system, with fingerprint and voice biometrics, that is more robust to environmental and sensor noise. The focus here is to determine the *best integration weight* β using reliability and separability measures. The proposed integration weight proves to be simpler and gives robust and improved recognition accuracy at varying noise conditions. To the best of our knowledge, the proposed score level fusion of fingerprints and voice using reliability and separability measures has not been attempted until now.

3. INDIVIDUAL CLASSIFIERS

3.1 Fingerprint Classifier

We considered the minutiae-based fingerprint matching using ridge counting, as this approach is more robust against fingerprint distortions [27]. Given two sets of minutiae from the template (T) and the input fingerprint (I) images, the matching algorithm compares the minutiae points in the two images and returns a degree of similarity [15]. Each minutiae is represented as a triplet $m = \{x, y, \theta\}$ that indicates the x, y minutiae location coordinates and the minutiae angle θ . A minutiae m_i in T and a minutiae m'_j in I are considered matching, if the spatial distance (sd) between them is lesser than a given tolerance r_0 and the direction difference (dd) between them is lesser than an angular tolerance θ_0 [15].

$$sd(m'_j, m_i) = \sqrt{(x'_j - x_i)^2 + (y'_j - y_i)^2} \leq r_0 \quad (1)$$

$$dd(m'_j, m_i) = \min(|\theta'_j - \theta_i|, 360^\circ - |\theta'_j - \theta_i|) \leq \theta_0 \quad (2)$$

Elastic matching algorithm is used to perform matching between the two fingerprints [27]. Match score formula for the reference and the test print is given by [28],

$$Matchingscore = \frac{100N_{pair}}{\max\{M, N\}} \quad (3)$$

where N_{pair} is the number of matched minutiae, M is the number of minutiae in the template set, and N is the number of minutiae in the test set. Maximum similarity criterion is used for fingerprint pattern classification.

3.2 Voice Classifier

Short-time spectral analysis is used to characterize the quasi-stationary voice samples. To represent the voice samples in a parametric way, we considered the MFCC representation as they were proved to be efficient and compact [19]. The number of mel cepstrum coefficients, is chosen as 16 (here). Gaussian mixture model (GMM) is considered here for representing the acoustic feature vectors. The complete Gaussian mixture model is parametrized by the mean vectors, covariance and the mixture weights. These parameters are collectively represented by [19],

$$\lambda = \{a_i, \mu_i, \Sigma_i\}, \quad i = 1, \dots, M \quad (4)$$

So, by using the MFCC feature vectors and the statistical GMM, each enrolled speaker is uniquely represented by a specific λ . In the training stage itself, each enrolled speakers in g , where $bf_g = \{\hat{g}_1, \hat{g}_2, \dots, \hat{g}_G\}$ is represented by a unique GMM's (λ' 's). In the testing stage, the features from the unknown speaker's utterances are compared with statistical models of the voices of speakers known to the system. The Bayes rule suggests to allocate the test samples to the class \hat{g}_k , having the highest posterior probability, that is [19],

$$\hat{g}_k = \arg \max_{1 \leq k \leq G} p(X|\lambda_k) \quad (5)$$

where $p(X|\lambda_k)$ is the *a posteriori* probability for a given observation sequence. In this case, maximum likelihood classifier is used.

4. OPTIMAL FUSION USING RELIABILITY MEASURES

One of the popular approaches among various integration weight computation schemes is the reliability ratio based method. Reliability estimation has shown to be an efficient and accurate way for reducing the classification errors in both unimodal and multimodal systems [11]. Reliability measure can be used as an auxiliary quality information for the score level fusion. In this approach, the integration weight is determined from the relative reliability of the two modalities. The reliability parameters of each modality are captured from the matching score matrix. When the voice samples does not contain any noise, there are large differences in the matching score values. As the voice samples become noisy, these differences tend to become small. Considering this observation, the reliability of a modality is defined in several ways as mentioned in [13]. The modalities reliability parameters are estimated based upon the variances of the matching scores. The usual measure is to calculate the variance around the best or the least score rather than the mean or median. The reliability ratio can be calculated as follows [12, 17].

$$\gamma = \frac{S_V}{S_V + S_F} \quad (6)$$

where S_V and S_F are the reliability measures of the voice and fingerprint modality, respectively. This can be obtained from the respective matching score matrices using

$$S_m = \frac{1}{N} \sum_{n=1}^N (s_n^m - \min(s_n^m))^2 \quad (7)$$

where 'N' is the number of test samples considered from all the classes and 'm' stands for the reliability of either fingerprint(F) or voice (V) modality. This quantity measures the dispersion of the score values to the least score rather than the mean. Even though the integration weight using equation 6 can improve noise robustness under certain noise conditions, it is not always the optimal. Hence, a modified integration weight β given by equation 8 is employed to obtain better performance under low SNR

Table 1. Training/ Validation Accuracy of Individual Classifiers

Modality		Voice							Fingerprints
SNR in dB	Set.No.	-10	-5	0	5	10	15	20	Clean
Accuracy	Val.1	9.3478	16.087	33.9130	50.6522	65.6522	87.1738	94.5652	95.6522
	Val.2	4.3478	6.7391	25.8696	44.7826	70.2174	86.5217	96.5218	91.3043
	Val.3	6.7391	20.0000	27.6087	42.1739	66.9565	84.5652	99.1304	91.3043
	Val.4	8.2609	12.8261	18.4783	33.9130	58.0435	87.8260	100.0000	86.9565
	Val.5	7.8261	20.2174	21.9565	30.2174	59.3478	76.0869	95.2174	60.8696
	Val.6	4.3478	4.5652	23.9131	51.5217	81.0870	95.2174	100.0000	86.9565
	Val.7	9.7825	13.2609	16.3044	45.6522	74.7826	95.4348	98.2609	95.6522
Average		7.2360	13.3851	24.0062	42.7019	68.0124	87.5465	97.6708	86.9565

Table 2. Optimal Integration Weights estimated using LOOCV

No	SNR in dB	Baseline System		Proposed reliability based method						
		GS β	GA β	γ	GS x_{opt}	GA x_{opt}	PSO x_{opt}	GS β	GA β	PSO β
1	20	0.7071	0.8087	0.7421	0.8174	0.9612	0.9664	0.6066	0.7133	0.7172
2	15	0.6928	0.7387	0.7086	0.7446	0.9410	0.9800	0.5276	0.6668	0.6944
3	10	0.6500	0.6605	0.7035	0.7444	0.8722	0.9868	0.5237	0.6136	0.6942
4	5	0.1255	0.1255	0.6875	0.600	0.7796	0.8662	0.4125	0.5360	0.5955
5	0	0.0000	0.0000	0.6720	0.4165	0.2674	0.2722	0.2799	0.1797	0.1829
6	-5	0.0000	0.0000	0.6518	0.0677	0.1915	0.0500	0.0441	0.1248	0.0326
7	-10	0.0000	0.0000	0.6365	0.0528	0.1568	0.0148	0.0336	0.0998	0.0094

conditions [17].

$$\beta = x_{opt} \times \left[\frac{S_V}{S_V + S_F} \right] \quad (8)$$

where x_{opt} is the scaling factor which needs to be optimized. In order to emphasize or deemphasize the scores obtained from the unimodal systems, the integration weight factor must be adaptive and optimal. That is, the weights must be very appropriate and self adapted to the fluctuating inputs. So, we propose linear programming and stochastic optimization techniques to obtain the appropriate integration weights for fusion. The optimal integration weight was obtained in the training/ validation stage using LOOCV. The min-max normalized match scores from the two modalities were combined by the weighted sum rule to produce the final decision. Given the speaker scores $S^{(sc)}$ and the finger scores $S^{(fc)}$, the fused scores can be obtained by linearly combining the two scores.

$$S^{(fus)} = \beta S^{(sc)} + (1 - \beta) S^{(fc)} \quad (9)$$

The weighting factor $\beta (0 \leq \beta \leq 1)$ determines how much each modality contributes to the final decision. The proposed method systematically chooses the best scaling factor x_{opt} from a defined domain ($0 \leq x_{opt} \leq 1$) so as to maximize the objective function (recognition accuracy). The objective function is given by:

$$RecognitionAccuracy = - \frac{\sum diag(C_{Mat})}{\sum \sum (C_{Mat})} \times 100 \quad (10)$$

where C_{Mat} is the confusion matrix. The performance of the proposed scheme is compared with that of equal weight bimodal biometric system and the method proposed by [4]. To show the effectiveness of the proposed technique, we evaluated the performance of the system with a direct search optimization method like Grid Search (GS) and random search optimization methods such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO).

5. FUSION USING MULTI-NORMALIZATION AND SEPARABILITY MEASURES

The global recognition rate of the multibiometric system could be improved by incorporating the separability measures as integration weights in the fusion module. Here, we considered inter/intra class distance measure from the feature space and the d-prime separability measure from the matching score space.

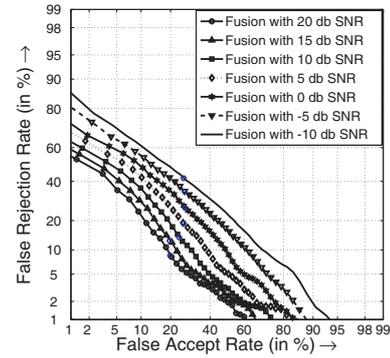


Fig. 1. System With Equal Weighting (Baseline)

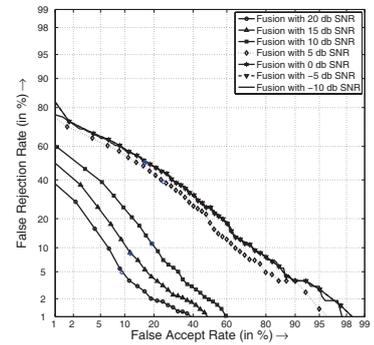


Fig. 2. GS Based Optimization (Baseline)

5.1 Estimation of Inter/Intra Class Distance

The inter/intra class distance based on Euclidean norms gives an indication of how well the classes in the feature space can be discriminated. Here, the basic assumption is that the class dependent distributions are such that the expectation vectors of different classes are discriminating [25]. Let T_S is a labelled training set with N_S feature vectors. The classes ω_k are represented by subsets $N_k \subset N_S$, each class having N_k features ($\sum N_k = N_S$). Feature vectors in T_S without reference to their classes are denoted by z_n . Feature vectors in T_k (i.e. vectors coming from the class ω_k) are denoted by $z_{k,n}$. In order to quantify the scattering of feature vectors in the space, we consider the

Table 3. Recognition Accuracy with Baseline and Reliability based methods

No	SNR in dB	Individual Classifiers		Baseline Systems			Proposed Method		
		Fingerprint	Voice	Equal Weightage	Grid Search	Genetic	Grid Search	Genetic	PSO
1	20	95.6522	98.6956	97.8261	100.0000	100.0000	100.0000	100.0000	100.0000
2	15	95.6522	91.4493	97.6812	97.8261	97.8261	100.0000	100.0000	100.0000
3	10	95.6522	69.4203	95.6522	97.2464	97.1015	97.8261	97.8261	98.0507
4	5	95.6522	33.1884	95.6522	95.6522	95.6522	96.8116	96.8116	96.5290
5	0	95.6522	23.1884	85.6522	95.6522	95.6522	93.4783	93.4783	93.2826
6	-5	95.6522	9.1304	68.6956	95.6522	95.6522	93.4783	93.4783	92.2464
7	-10	95.6522	5.6521	57.3913	95.6522	95.6522	93.4783	93.4783	92.0942

Table 4. Combining the Weight Factors from the Feature and Score space

No.	SNR in dB	d'_F	d'_V	W_1	ρ_F	ρ_V	W_2	$(W_1 + W_2)/2$	$W_1 \times W_2$	Adaptive weight
1	20	0.8823	2.1176	0.7059	0.0166	0.0314	0.6358	0.6799	0.4615	0.6799
2	15	0.8823	1.8633	0.6786	0.0166	0.0305	0.6479	0.6633	0.4397	0.6633
3	10	0.8823	1.5593	0.6386	0.0166	0.0263	0.6128	0.6257	0.3913	0.6257
4	5	0.8823	1.2692	0.5899	0.0166	0.0252	0.6033	0.5966	0.3559	0.5966
5	0	0.8823	0.9403	0.5159	0.0166	0.0159	0.4886	0.5023	0.2521	0.2521
6	-5	0.8823	0.6222	0.4136	0.0166	0.0079	0.3236	0.3686	0.1338	0.1338
7	-10	0.8823	0.2503	0.2210	0.0166	0.0038	0.1855	0.2033	0.0410	0.0410

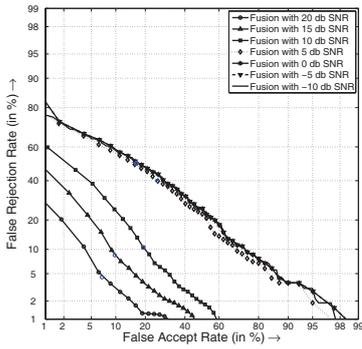


Fig. 3. GA Based Optimization (Baseline)

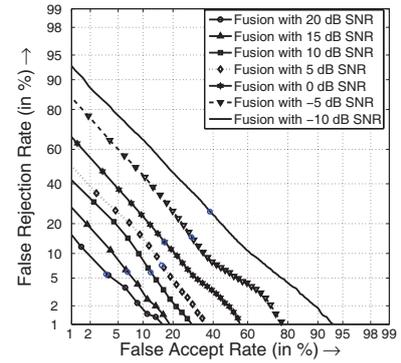


Fig. 4. Reliability Based GS (Proposed)

scatter matrices. Scatter matrices gives some information about the dispersion of the feature vectors around their mean. The sample mean of class ω_k is given by

$$\hat{\mu}_k = \frac{1}{S_k} \sum_{n=1}^{S_k} z_{k,n} \quad (11)$$

The sample mean of the entire training set is given by

$$\hat{\mu} = \frac{1}{S_N} \sum_{n=1}^{S_N} z_n \quad (12)$$

Matrix that describes the scattering of vectors from class ω_k is

$$S_k = \frac{1}{N_k} \sum_{n=1}^{N_k} (z_{k,n} - \hat{\mu}_k)(z_{k,n} - \hat{\mu}_k)^T \quad (13)$$

S_k provides information about the average distance of the scattering for class ω_k . Let S_w and S_b are the within class scatter matrix and between class scatter matrix respectively. S_w gives the average scattering within the classes. In order to obtain S_w , we find the scattering of all the classes and take the average.

$$S_w = \frac{1}{N_S} \sum_{k=1}^K N_k S_k \quad (14)$$

S_b gives the scattering of the class dependent sample means around the overall average.

$$S_b = \frac{1}{N_S} \sum_{k=1}^K N_k (\hat{\mu}_k - \hat{\mu})(\hat{\mu}_k - \hat{\mu})^T \quad (15)$$

The performance measure well suited to express the separability of classes is the ratio between interclass and intraclass distance [25]. That is,

$$\frac{J_{INTER}}{J_{INTRA}} = \frac{\text{trace}(S_b)}{\text{trace}(S_w)} \quad (16)$$

J_{INTER} denotes the fluctuations of the conditional expectations around the overall expectation, i.e. the fluctuations of the signal, while J_{INTRA} measures the fluctuations due to noise. Hence $\frac{J_{INTER}}{J_{INTRA}}$ can be considered as 'signal-to-noise-ratio' [25]. This measure is estimated from the feature space of both the fingerprints and voice modalities. The separability measures thus obtained in the training stage are used as the integration weights for fusion in the testing stage.

5.2 Score Normalization

Score normalization is essentially a transformation technique, that effectively normalizes for any unwanted peculiarities involved in the raw similarity computations. Various score normalization techniques were proposed in the literature [1, 8, 22]. For a good normalization scheme, the estimates of the location and scale parameters of the matching score distribution must be robust and efficient. All the normalization techniques are not equally suited for the different match score distributions. Here we use cohort and tanh normalization techniques for the fingerprint and voice similarity scores, to enhance the efficiency and the robustness of the system, under varying noise conditions. Tanh (TH) normalization is one of the robust and efficient normalization methods. This normalization method transforms the genuine scores to a distribution with a mean of 0.5 and standard deviation of approximately α . This transformation maps the raw

Table 5. Recognition Accuracy with Separability Measures

No.	SNR in dB	Accuracy of Classifiers		SMIW			MNSMIW		
		Fingerprint	Voice	Average	Product	Adaptive	Average	Product	Adaptive
1	20	95.6522	98.6956	100.0000	97.8261	100.0000	100.0000	100.0000	100.0000
2	15	95.6522	91.4493	97.8261	95.6522	97.8261	100.0000	100.0000	100.0000
3	10	95.6522	69.4203	96.5217	95.6522	96.5217	100.0000	100.0000	100.0000
4	5	95.6522	33.1884	95.2174	95.6522	95.6522	100.0000	100.0000	100.0000
5	0	95.6522	23.1884	85.9420	93.4783	93.4783	100.0000	100.0000	100.0000
6	-5	95.6522	9.1304	82.1739	93.4783	93.4783	100.0000	100.0000	100.0000
7	-10	95.6522	5.6521	87.5362	92.3188	92.3188	100.0000	100.0000	100.0000

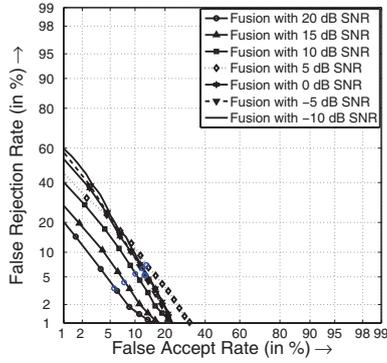


Fig. 5. Reliability Based GA (Proposed)

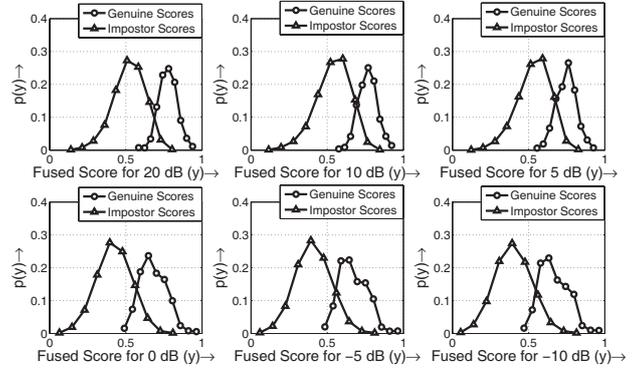


Fig. 8. Density Plot: Reliability GA (Proposed)

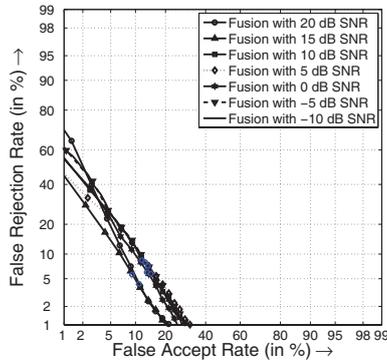


Fig. 6. Reliability Based PSO (Proposed)

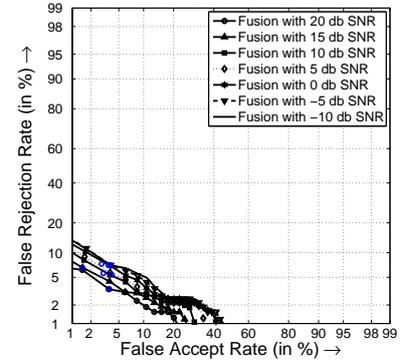


Fig. 9. Fusion With Separability Measures (Proposed)

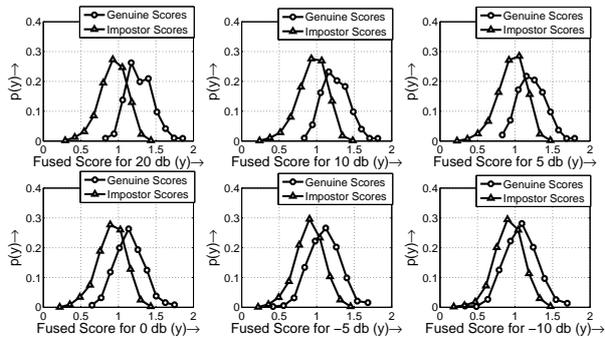


Fig. 7. Density Plot: Equal Weighting (Baseline)

scores to a range (0,1). Given a set of matching scores $s = \{s_i\}$, $i = 1, 2, \dots, n$, the normalized score s'_i is obtained by,

$$s'_i = \frac{1}{2} \left[\tanh \left(\alpha \frac{(s_i - \text{mean}(s))}{\text{std}(s)} \right) \right] \quad (17)$$

In Cohort (C) normalization method a given query/test-sample is compared with the claimed as well as the cohort (neighbours) of the claimed identity[1]. Let $s(x, \lambda)$ is the similarity score of the query with the claimed identity, $s(x, \bar{\lambda})$ is the similarity score of the query with the cohort, where x is the query template, λ is the claimant class and $\bar{\lambda}$ represents the background class ($U - \lambda$), where U represents the universal set. The normalized scores $c(x, \lambda)$ can be calculated by,

$$c(x, \lambda) = \frac{s(x, \lambda)}{s(x, \bar{\lambda})} \quad (18)$$

Assuming the cohort to a size of k , $s(x, \bar{\lambda})$ can be determined using the max-rule [1].

$$s(x, \bar{\lambda}) = \max \{s(x, \lambda^1), s(x, \lambda^2), \dots, s(x, \lambda^k)\} \quad (19)$$

$s(x, \lambda^1), s(x, \lambda^2), \dots, s(x, \lambda^k)$ is the set of similarity score of the query with the cohort λ^j 's for the enrolled identity λ . $j = 1, 2, \dots, k$.

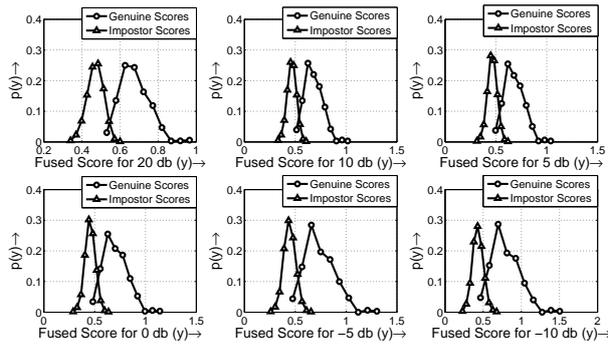


Fig. 10. Density Plot: MNSMIW (Proposed)

5.3 Estimation of d-prime Separability Measure

The d-prime separability measures gives the ratio of separation to spread (*Separation/spread*) of the genuine and the impostor score distributions. Separation gives the indication that how much mean of the distributions are separated and the spread gives an indication of the overlap. d-prime gives a measure how well the non-match score probability density and the match score probability density are separated. the d' is defined as [6],

$$d' = \frac{\mu_m - \mu_n}{\sqrt{(\sigma_m^2 + \sigma_n^2)}} \quad (20)$$

where μ_m = mean of genuine scores; σ_m^2 = variance of genuine scores; μ_n = mean of impostor scores; σ_n^2 = variance of impostor scores. A higher d-prime indicates that the Genuine scores can be more readily detected. Thus, the discriminability of a class depends both on the separation and the spread of the Genuine and impostor score distribution curves.

5.4 Fusion Using Separability Measures

We combined the separability measures from the feature space and the score level phase using sum rule and product rule. Let ρ_F and ρ_V denote the inter/intra class distance measures obtained from the fingerprint and voice modality respectively. d'_F and d'_V denote the d-prime separability measures obtained from preprocessed similarity scores of fingerprints and voice modality respectively. The results are shown in Table 4 and it is evident from the table that the separability measure decreases with the increase in noise. Following parameters are defined to obtain the fused scores and Table 3.2 depicts their numerical values.

$$\rho = \frac{\rho_V}{\rho_F + \rho_V} \quad (\text{Feature space}) \quad (21)$$

$$d' = \frac{d'_V}{d'_F + d'_V} \quad (\text{Score space}) \quad (22)$$

$$\beta = \frac{1}{2} \frac{\rho_V}{\rho_F + \rho_V} + \frac{1}{2} \frac{d'_V}{d'_F + d'_V} \quad (\text{Average}) \quad (23)$$

$$\beta = \left(\frac{\rho_V}{\rho_F + \rho_V} \right) \times \left(\frac{d'_V}{d'_F + d'_V} \right) \quad (\text{Product}) \quad (24)$$

We performed an adaptive score level fusion using the separability measures. It is experimentally found that for normal conditions the weighted average of the separability measures $\frac{1}{2}\rho + \frac{1}{2}d'$ provides optimal weighting (improves the recognition accuracy) while for the extreme noise conditions the product of the separability measures $\rho d'$ improves the recognition accuracy. In order to derive the advantages of combining the separability measures using sum and product rule, under normal and extreme condition, we use an adaptive approach for the score level fusion. Under normal noise conditions from 20 dB SNR to 5 dB SNR we

combined the separability measures using sum rule and for adverse conditions from 0 dB SNR to -10 dB SNR the separability measures are combined using the product rule. This strategy can improve the overall recognition accuracy as well as the robustness of the system. The results are shown in Table 5.

6. SIMULATION RESULTS AND DISCUSSIONS

Finger images from the FVC2002 fingerprint database [14] and voice samples from ELSDSR database [7] have been employed for the experimentation. We took nine samples per person from the fingerprint and the speaker database. Out of these nine samples from each biometric, seven samples were used for training the individual classifiers and two samples were used for testing. As the fingerprint biometric is more robust, the performance of the system under varying noise conditions was not considered. The performance of the weak, voice biometric system under varying noise conditions was investigated by artificially degrading the test samples with additive white Gaussian noise. The output of the two classifiers were consolidated into a single vector of scores using the sum rule of fusion.

6.1 Fusion With Baseline Systems

6.1.1 Fusion with equal weights. In this case a constant value of $\beta = 0.5$ was assigned as an integration weight at all SNR levels as discussed in [4]. This technique will not favour one modality over another. More over the combined recognition accuracy may not be maximum always. The training accuracy of the baseline system is shown in Table 3. Further insight could be obtained from Fig. 1 and Fig. 7.

6.1.2 Fusion With Optimal Integration Weight Without Reliability Information. We compared the proposed reliability based integration weight optimization technique with a non-reliability based integration weight estimation scheme as discussed in [4]. The overall testing accuracy of the non-reliability based integration weight estimation scheme is presented in Table 3. This method shows improved accuracy than any of the unimodal systems in the normal operating conditions and maintained the accuracy of the better unimodal ones for all the extreme noise conditions. Further insight could be obtained from the DET plots (Fig. 2 and Fig. 3). The disadvantage of the method is at the extreme noise conditions the fusion module contributed zero weighting to the voice modality. In these conditions, the overall performance depends solely on the fingerprint matcher.

6.2 Fusion With The Proposed Methods

6.2.1 Fusion with reliability based optimal integration weight. From the training/validation stage (using LOOCV), we obtained optimal integration weights (β 's) for different noise conditions from -10dB to 20 dB. We have applied one dimensional Grid Search method and the stochastic optimization techniques; GA and PSO for optimizing the reliability-based integration weight. The overall validation accuracy of the individual classifiers for various SNR conditions is shown in the Table 1. The relative reliability ratio estimates of the two modalities and the reliability-based optimal integration weight β estimated for the various SNR conditions are shown in the Table 2. The β values thus estimated during the training/validation stage is used for testing. The overall testing accuracy of the proposed method is depicted in Table 3. Even though the recognition accuracy of the proposed method shows attenuation at very extreme noise conditions (0db, -5db and -10db), the proposed method shows better performance in terms of recognition accuracy, FAR and FRR, than the baseline systems. This is evident from the Table 3, respective DET plots (Fig. 4, Fig. 5 and Fig. 6) and from the score density plots (Fig. 8). Experimental evaluations reveals that the classification errors show a more pronounced reduction when we use the qual-

ity measure for finding the optimal integration weight. This is because a higher reliability measure positively correlates with the chance of making a correct classification decision. The score density plot in Fig. 8 indicates that the effective overlap between the genuine and impostor score distributions get considerably reduced with the proposed method, making the system more robust to the fluctuating inputs. So we could demonstrate the advantages of using unimodal reliability information to find the best integration weights for score level fusion.

6.2.2 Fusion With Separability Measures as Integration Weights. The inter/intra class separability measures are derived from the feature space of the two modalities. The inter/intra class measures for voice modality (ρ_V) at different noise conditions are obtained by artificially degrading the training samples with AWGN from 20 dB SNR to -10 dB SNR. The separability measures thus obtained are used as the integration weights for fusion in the testing stage. The d-prime separability measures for various noise conditions (20 dB SNR to -10 dB SNR) are obtained from the genuine and impostor scores during the training/validation stage using LOOCV strategy. Table 4 shows the d-prime statistic for the finger print (d'_F) and voice (d'_V) biometrics. We combined the separability measures from the feature space and the matching score space using sum rule, product rule and an adaptive approach as described in section 5.4. Table 5 shows the recognition accuracy of the method, with Separability Measures as Integration Weights (SMIW). To improve the global recognition rate of the multibiometric system, the scores obtained from the degraded modality are pre-processed using score normalization. Fingerprint similarity scores are preprocessed using cohort normalization. To perform cohort normalization we divided the genuine scores with the maximum of the impostor scores. This transformation, increases the separability between the genuine and impostor scores, which in turn improves the overall performance of the biometric system. In order to improve the robustness of the noisy, voice modality the similarity scores are transformed using tanh normalization technique. We observed that better performance could be obtained by this multi-normalization technique. Table 5 shows the recognition accuracy of the proposed method. (MNSMIW - Multi-normalization with Separability Measures as Integration Weights). Separability measures combined with sum and product rule shows similar performance in terms of accuracy, FAR and FRR. Further insight could be obtained from the DET plots (Fig. 1 and Fig. 9) and the score density plots (Fig. 7 and Fig. 10). The DET plot of Fig. 9 show that the proposed method considerably reduces the FAR and FRR and the score density plot of Fig. 10 show that the overlap between the genuine and the impostor score distributions reduces with the multi-normalization and separability measures. Hence, the proposed score level fusion using separability measures and score preprocessing techniques outperforms the baseline fusion technique under various noise conditions.

7. CONCLUSION

In the first part of this paper, an optimum reliability ratio based integration weight optimization scheme for fingerprint and voice modalities is proposed and implemented. The performance of the multibiometric system is evaluated under various noise conditions. By estimating the optimal integration weight using stochastic optimization strategies and LOOCV techniques, we could automate the process and make the system more robust to fluctuating inputs. This method effectively reduces FAR and hence it will be highly suitable for applications like sharing networked computer resources, granting access to nuclear facilities, performing remote financial transactions or boarding a commercial flight. One drawback of this method is that it gives attenuating fusion under extreme noise conditions. A new weighting strategy is proposed in the second part of this paper by com-

binning the separability measures from the feature space and the match score space. An efficient preprocessing on the raw vector of scores using multi-normalization is employed, to improve the performance of the system. The proposed method is robust and could successfully eliminate the attenuating fusion even at low SNR conditions while reducing the FRR considerably. As the FRR is low, this method can be very effectively implemented in a multibiometric system for forensic applications like criminal investigation, parenthood determination etc.

8. REFERENCES

- [1] G. Aggarwal, N.K. Ratha, R.M. Bolle, and R. Chellappa. Multi-biometric cohort analysis for biometric fusion. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 5224–5227. IEEE, 2008.
- [2] Harbi Almahafzah and Maen Zaid Alrwashdeh. Article: A survey of multibiometric systems. *International Journal of Computer Applications*, 43(15):36–43, April 2012. Published by Foundation of Computer Science, New York, USA.
- [3] F. Alsaade, A. Ariyaeinia, A. Malegaonkar, and S. Pillay. Qualitative fusion of normalised scores in multimodal biometrics. *Pattern Recognition Letters*, 30(5):564–569, 2009.
- [4] S.M. Anzar and P.S. Sathidevi. Optimization of integration weights for a multibiometric system with score level fusion. In *Advances in Computing and Inform. Technology, AISC 177*, pages 833–842. Springer-Verlag Berlin Heidelberg, 2012.
- [5] M. Bendris, D. Charlet, and G. Chollet. Introduction of quality measures in audio-visual identity verification. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1913–1916. IEEE, 2009.
- [6] R. Bolle. *Guide to biometrics*. Springer-Verlag New York Inc, 2004.
- [7] L. Feng and L.K. Hansen. *A new database for speaker recognition*. IMM, Informatik og Matematisk Modellering, DTU, 2005.
- [8] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern recognition*, 38(12):2270–2285, 2005.
- [9] AK Jain, L. Hong, and Y. Kulkarni. A multimodal biometric system using fingerprint, face and speech. In *Proceedings of 2nd Int'l Conference on Audio-and Video-based Biometric Person Authentication, Washington DC*, pages 182–187, 1999.
- [10] HB Kekre and VA Bharadi. Ageing adaptation for multimodal biometrics using adaptive feature set update algorithm. In *Advance Computing Conference, 2009. IACC 2009. IEEE International*, pages 535–540. IEEE, 2009.
- [11] K. Kryszczuk, J. Richiardi, P. Prodanov, and A. Drygajlo. Reliability-based decision fusion in multimodal biometric verification systems. *EURASIP Journal on Applied Signal Processing*, 2007(1):74–74, 2007.
- [12] J.S. Lee and C.H. Park. Robust audio-visual speech recognition based on late integration. *Multimedia, IEEE Transactions on*, 10(5):767–779, 2008.
- [13] T.W. Lewis and D.M.W. Powers. Sensor fusion weighting measures in audio-visual speech recognition. In *Proceedings of the 27th Australasian conference on Computer science-Volume 26*, pages 305–314. Australian Computer Society, Inc., 2004.
- [14] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, and A.K. Jain. Fvc2002: Second fingerprint verification competition.

- In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 3, pages 811–814. IEEE, 2002.
- [15] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of fingerprint recognition*. Springer-Verlag New York Inc, 2009.
- [16] N. Poh and S. Bengio. Improving fusion with margin-derived confidence in biometric authentication tasks. In *Audio-and Video-Based Biometric Person Authentication*, pages 347–356. Springer, 2005.
- [17] Rajavel R. and Sathidevi P.S. A new GA optimised Reliability Ratio based integration weight estimation scheme for decision fusion Audio-Visual Speech Recognition. *International Journal of Signal and Imaging Systems Engineering*, 4(2):123–131, 2011.
- [18] Rajavel R. and Sathidevi P.S. Adaptive reliability measure and optimum integration weight for decision fusion audio-visual speech recognition. *Journal of Signal Processing Systems*, pages 1–11, 2011.
- [19] D. Reynolds. Gaussian mixture models. *Encyclopedia of Biometric Recognition*, 2008.
- [20] A.A. Ross, K. Nandakumar, and A.K. Jain. *Handbook of multibiometrics*, volume 6. Springer-Verlag New York Inc, 2006.
- [21] S. Sarkar and P. Mohanty. Globally linear embedding of biometric scores: An empirical study. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1–6. IEEE, 2008.
- [22] S.Thangasamy and L.Latha. Article:efficient approach to normalization of multimodal biometric scores. *International Journal of Computer Applications*, 32(10):57–64, October 2011. Published by Foundation of Computer Science, New York, USA.
- [23] K.A. Toh. Fingerprint and speaker verification decisions fusion. In *Image Analysis and Processing, 2003. Proceedings. 12th International Conference on*, pages 626–631. IEEE, 2003.
- [24] K.A. Toh and W.Y. Yau. Fingerprint and speaker verification decisions fusion using a functional link network. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 35(3):357–370, 2005.
- [25] F. Van der Heijden, R.P.W. Duin, D. De Ridder, and DMJ Tax. *Classification, parameter estimation and state estimation*. Wiley Online Library, 2004.
- [26] Y. Wang, Y. Wang, and T. Tan. Combining fingerprint and voiceprint biometrics for identity verification: an experimental comparison. *Biometric Authentication*, pages 289–294, 2004.
- [27] Wuzhili. *Fingerprint Recognition*. Honors Thesis, 2002.
- [28] N. Yager and A. Amin. Fingerprint verification based on minutiae features: a review. *Pattern Analysis & Applications*, 7(1):94–113, 2004.