AN EFFICIENT PSO OPTIMIZED INTEGRATION WEIGHT ESTIMATION USING D-PRIME STATISTICS FOR A MULTIBIOMETRIC SYSTEM

S. M. Anzar\textsuperscript{1} and P. S. Sathidevi\textsuperscript{2}

\textsuperscript{1}Department of Electronics & Communication Engineering, National Institute of Technology, Calicut, India, 673601
p090004ec@nitc.ac.in

\textsuperscript{2}Department of Electronics & Communication Engineering, National Institute of Technology, Calicut, India, 673601
sathi@nitc.ac.in

ABSTRACT

The efficiency of a multibiometric system can be improved by weighting the scores obtained from the degraded modalities in an appropriate manner. In this paper, we propose an efficient PSO (Particle Swarm Optimization) based integration weight optimization scheme using d-prime statistics to determine the optimal weight factors for the complementary modalities, under different noise conditions. Instead of treating the weight estimation process from an algebraic point of view, an attempt is made to consider the same from the principles of linear programming techniques. The performance of the proposed method is analysed in the context of fingerprint and voice biometrics using sum rule of fusion. The d-prime statistics of fingerprint and voice modalities are measured and the integration weights are computed as the ratio of these two statistics. The computed d-prime ratio is then optimized against the recognition accuracy. The optimizing parameter is estimated in the training/validation phase using Leave-One-Out Cross Validation (LOOCV) technique. Experimental studies show that the proposed method improves recognition rate under normal operating conditions and reduces the FAR (False Acceptance Rate) considerably even at extreme low SNR (Signal to Noise Ratio) conditions. The proposed biometric solution can be easily integrated into any multibiometric system with score level fusion. Moreover, it finds extremely useful in applications where there are less number of available training samples.

KEYWORDS

Multibiometric System, Integration Weight, d-prime Statistics, Particle Swarm Optimization, Robustness

1. INTRODUCTION

In this paper, the primary focus is on the determination of the optimal integration weights using d-prime statistics, for the score level fusion of fingerprints and voice biometrics. Although the recognition accuracy of the voice biometrics is high in clean conditions, its performance tends to be significantly degraded under the presence of background noise. This is not desirable in real world applications. At high SNR, 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 optimal weight factor dynamically, to combine both the modalities for the maximum recognition accuracy \cite{1}. Environmental noise and the quality in acquisition do affect the matching score distributions. Two examples of this are scanner quality in fingerprints and microphone quality for voice recognition \cite{2}. Integration weights depend on the matching score.
distributions of the voice biometrics since the matching score vectors are heavily affected by the amount of acoustic noise.

Figure 1: Block diagram representing Score Level Fusion

The smaller the overlap between the genuine and the impostor scores the better the recognition rate. As the amount of noise increases, the overlap between the genuine and the impostor score distributions will also increase. Thus, the score separability measures from the matching score space at different noise conditions give an indication of the quality of the matcher. It is well known that the d-prime statistics gives the ratio of separation to spread (separation/spread) of the genuine and the impostor score distributions. Here, we use this auxiliary information to find the best integration weights for the optimal fusion of the two modalities. The estimation of optimal integration weight is important because it determines the amount of contribution of each modality towards the final decision. Otherwise, the system will show attenuating fusion. We are motivated by [3, 4] to develop a bimodal biometric system that is more robust to environmental and sensor noise. Figure 1 shows the overall block diagram of the score level fusion strategy.

2. RELATED WORKS

As the inputs tend to be noisy, the sum rule of fusion is proved more effective [5]. However, this fusion strategy results in attenuating fusion, when the integration weights are not optimal. One can manually determine the integration weight, but this is a non-optimal approach, which often needs much iteration and lot of expertise [6]. Previously, the integration weight estimation was treated from an algebraic point of view. Optimal integration weights estimation using least squares technique was reported in [7]. Reliability based optimal integration weights estimation for audio-visual decision fusion was reported in [4, 6]. Optimal integration weights for a multibiometric system with fingerprints and voice were reported in [1] by the author’s of this paper. We compared the proposed PSO based integration weight optimization scheme with the integration weights optimization scheme reported in [1]. In order to automate the estimation of the “best available weight factor”, we formulated the integration weight estimation problem using mathematical programming techniques. The 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 were obtained by linearly combining the two scores as follows,

\[
S^{(fus)} = \beta S^{(sc)} + (1 - \beta) S^{(fc)}
\]  

(1)
The weight factor $\beta (0 \leq \beta \leq 1)$ determined the amount of contribution of each modality to the final decision. The objective was to find the optimal $\beta$ that maximise the recognition accuracy. LOOCV strategy was employed for the estimation of the optimal integration weight. The performances of the proposed method were evaluated with a direct search optimization method (Grid Search (GS) [8]) and a random search optimization method (Genetic Algorithm (GA) [9]). Even though the recognition accuracy of the proposed methods were pretty better than the unimodal systems under normal operating conditions, the demerits of the method was that, at the extreme noise conditions the fusion module contributed zero weighting to the voice modality. In these conditions, the overall performances of the multimodal system solely depend on the unimodal fingerprint matcher. Moreover the FAR and FRR hardly show any substantial improvements at low SNR conditions. In order to improve the global recognition accuracy and reduce the FAR under various noise conditions, we proposed PSO optimized integration weight estimation scheme using the d-prime statistics. To the best of our knowledge, the proposed optimal integration weight estimation using d-prime statistics has not been attempted until now.

3. PROPOSED METHOD

In order to improve the performance of score level fusion at various noise conditions, we present a new weighting strategy using the d-prime statistics [10]. For judging the quality of a matcher, it is sufficient to measure the separation between non-match score probability density and the match score probability density. The sensitivity index or the d-prime statistic could effectively characterise these distance measures. The d-prime statistics provides the separation between the means of the genuine and the impostor score distributions, in units of the standard deviation of the impostor score distribution [10]. The separation/spread $d$ is defined as,

$$d = \frac{\mu_m - \mu_n}{\sqrt{\sigma_m^2 + \sigma_n^2}}$$  \hspace{1cm} (2)

where, $\mu_m$ = mean of genuine scores; $\sigma_m$ = variance of genuine scores; $\mu_n$ = mean of impostor scores; $\sigma_n$ = variance of impostor scores. Numerator (separation) gives the indication that how much the mean of the distributions are separated and the denominator (spread) gives an indication of their overlap. A high d-prime value indicates that the genuine scores can be readily detected. Thus, the discriminability of a class depends on both the separation and the spread of the genuine and impostor score distribution curves. The sensitivity index captures both the separation and the spread of the genuine and impostor score distributions effectively. The global recognition rate of a multibiometric system can be improved by incorporating these separability measures as integration weights in the fusion module. The d-prime statistics of each modality were captured from the genuine and impostor scores during the training/validation stage. When the voice samples do not contain any noise, the d-prime statistic exhibits high values. As the voice samples become noisy, these values tend to become small. The d-prime ratio can be calculated as follows,

$$\alpha = \frac{d_v'}{d_v' + d_f'}$$  \hspace{1cm} (3)

where $d_v'$ and $d_f'$ are the d-prime indices of the voice and fingerprint modality respectively. Even though the integration weight obtained using equation (3) can improve the noise robustness under certain noise conditions, it is not always the optimal. Hence, a modified integration weight $\beta$ as given by equation (4) is employed to obtain better performance under low SNR conditions [6].
Here $k_{opt}$ is the scaling factor that needs to be optimized. We can automate the estimation of the “best available weight factor”, by formulating the problem using stochastic optimization techniques. The proposed PSO based optimization method systematically chooses the best ‘scale factor’ $k_{opt}$ from a defined domain $0 \leq k_{opt} \leq 1$ to maximize the objective function (Recognition Accuracy). The performances of the proposed method were compared with that of the baseline systems [1]. Robust feature vectors were extracted from the two biometric traits. We considered minutiae based system that used elastic matching algorithm for fingerprint matching [11]. MFCC features were extracted from the voice samples and Gaussian mixture model (GMM) was considered for representing the acoustic feature vectors [12]. Model with 16 MFCC feature vectors and 12 Gaussian mixtures were used. Score normalization was done to transform the match scores of the two matchers into a comparable (common) domain.

\[
\text{normalised \_score} = \frac{\text{unnormalised \_score} - \text{min\_score}}{\text{max\_score} - \text{min\_score}}
\]

The optimal integration weights were obtained in the training/validation stage. The normalized match scores from the two modalities were combined by the equation (6) to produce the final decision.

\[
S^{(fus)} = k_{opt} \left[ \frac{d'_v}{d'_v + d'_f} \right] S^{(vc)} + \left( 1 - k_{opt} \left[ \frac{d'_v}{d'_v + d'_f} \right] \right) S^{(fc)}
\]

The objective function is given by:

\[
\text{Recognition \_Accuracy} = -\frac{\sum \text{diag}(C_{Mat})}{\sum \sum C_{Mat}} \times 100
\]

$C_{Mat}$ is the confusion matrix. LOOCV strategy was employed for the estimation of the optimal integration weights.

### 3.1. Particle Swarm Optimization (PSO) Algorithm

The PSO algorithm is a stochastic optimization strategy, inspired by the social behaviour of the flock of swarms. Here the underlying concept is that, for every time instant, the velocity of each particle (potential solution), changes between its $pbest$ and $lbest$ locations [13]. The particle associated with the best solution (fitness value) seems to be the leader and each particle keeps track of its coordinates in the solution space. This fitness value is stored which is referred to as $pbest$. Another “best” value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbours of the particle. This location is called $lbest$. When a particle takes all the population as its topological neighbours, the best value is a global best and is called $gbest$. The algorithm is as follows:

1. Randomly generate initial candidate solutions.
2. Assign the position and velocity of the associated particles randomly.
3. Evaluate the fitness (objective function) of each particle.
4. Compare each particle's objective function value with this particle's personal best value. If better, update \( pbest \) and take record current position as the particles personal best position.

5. Find the lowest objective function value of the whole particles. If the value is better than \( gbest \), replace \( gbest \) with this objective function value, and take record the global best position.

6. Change velocities and positions.

Velocity and the position update is given by the following equations,

\[
X = c_1 * \text{rand1} * (pbest[old] - \text{present}[old])
\]

\[
\text{v}[\text{new}] = w * \text{v}[\text{old}] + X + c_2 * \text{rand2} * (gbest[old] - \text{present}[old])
\]  

(8)

\[
\text{present}[\text{new}] = \text{present}[\text{old}] + \text{v}[\text{new}]
\]  

(9)

where ‘\( w \)’ is the inertia weight, \( \text{v}[] \) is the particle velocity, \( \text{present}[] \) is the current particle (solution). \( \text{rand1} \) and \( \text{rand2} \) are the random number between \([0,1]\). \( c_1, c_2 \) are learning factors. Usually \( c_1 = c_2 = 2 \).

7. Repeat 3 to 7 until stop criteria are satisfied.

![Figure 2: Performance of the system with Equal Weighting (Baseline)](image)

**Table 1: Training/Validation accuracy of the individual classifiers**

<table>
<thead>
<tr>
<th>Modality</th>
<th>SNR in dB</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>Clean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val.1</td>
<td>9.3478</td>
<td>16.087</td>
<td>33.9130</td>
<td>50.6522</td>
<td>65.6522</td>
<td>87.1738</td>
<td>94.5652</td>
<td>95.6522</td>
<td></td>
</tr>
<tr>
<td>Val.2</td>
<td>4.3478</td>
<td>6.7391</td>
<td>25.8696</td>
<td>44.7826</td>
<td>70.2174</td>
<td>86.5217</td>
<td>96.5218</td>
<td>91.3043</td>
<td></td>
</tr>
<tr>
<td>Val.3</td>
<td>6.7391</td>
<td>20.0000</td>
<td>27.6087</td>
<td>42.1739</td>
<td>66.9565</td>
<td>84.5652</td>
<td>99.1304</td>
<td>91.3043</td>
<td></td>
</tr>
<tr>
<td>Val.4</td>
<td>8.2609</td>
<td>12.8261</td>
<td>18.4783</td>
<td>33.9130</td>
<td>58.0435</td>
<td>87.8250</td>
<td>100.000</td>
<td>86.9565</td>
<td></td>
</tr>
<tr>
<td>Val.5</td>
<td>7.8256</td>
<td>20.2174</td>
<td>21.9565</td>
<td>30.2174</td>
<td>59.3478</td>
<td>76.0859</td>
<td>95.2174</td>
<td>60.8566</td>
<td></td>
</tr>
<tr>
<td>Val.6</td>
<td>4.3478</td>
<td>4.5625</td>
<td>23.9131</td>
<td>51.5217</td>
<td>81.0870</td>
<td>95.2174</td>
<td>100.000</td>
<td>86.9565</td>
<td></td>
</tr>
<tr>
<td>Val.7</td>
<td>9.7825</td>
<td>13.2609</td>
<td>16.8044</td>
<td>43.8522</td>
<td>74.7826</td>
<td>95.4348</td>
<td>98.2609</td>
<td>95.6522</td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>7.2360</td>
<td>13.3851</td>
<td>24.0062</td>
<td>42.7019</td>
<td>68.0124</td>
<td>87.5465</td>
<td>97.6708</td>
<td>86.9565</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2: Fusion with Manually determined Integration Weight

<table>
<thead>
<tr>
<th>No</th>
<th>SNR in dB</th>
<th>Fingerprints</th>
<th>Voice</th>
<th>$w_f$</th>
<th>$w_s$</th>
<th>Combined Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>95.6522</td>
<td>98.6956</td>
<td>0.30</td>
<td>0.70</td>
<td>100.0000</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>95.6522</td>
<td>91.4493</td>
<td>0.35</td>
<td>0.65</td>
<td>97.8261</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>95.6522</td>
<td>69.4203</td>
<td>0.35</td>
<td>0.65</td>
<td>97.2464</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>95.6522</td>
<td>33.1884</td>
<td>0.65</td>
<td>0.35</td>
<td>95.6522</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>95.6522</td>
<td>23.1884</td>
<td>1.00</td>
<td>0.00</td>
<td>95.6522</td>
</tr>
<tr>
<td>6</td>
<td>-5</td>
<td>95.6522</td>
<td>9.1304</td>
<td>1.00</td>
<td>0.00</td>
<td>95.6522</td>
</tr>
<tr>
<td>7</td>
<td>-10</td>
<td>95.6522</td>
<td>5.6521</td>
<td>1.00</td>
<td>0.00</td>
<td>95.6522</td>
</tr>
</tbody>
</table>

### Table 3: Optimal Integration Weights Estimated using LOOCV

<table>
<thead>
<tr>
<th>No</th>
<th>SNR in dB</th>
<th>Baseline</th>
<th>Proposed Method</th>
<th>$d_f$</th>
<th>$d_v$</th>
<th>$\alpha$</th>
<th>$\text{PSO } k_{\text{opt}}$</th>
<th>$\text{PSO } \beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>0.7071</td>
<td>0.8087</td>
<td>2.1176</td>
<td>0.7059</td>
<td>0.9795</td>
<td>0.6914</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0.6928</td>
<td>0.7387</td>
<td>1.8633</td>
<td>0.6786</td>
<td>0.9699</td>
<td>0.6582</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0.6500</td>
<td>0.6605</td>
<td>1.5593</td>
<td>0.6386</td>
<td>0.8449</td>
<td>0.5396</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.1255</td>
<td>0.1255</td>
<td>1.2692</td>
<td>0.5899</td>
<td>0.7790</td>
<td>0.4595</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.8823</td>
<td>0.9403</td>
<td>0.5159</td>
<td>0.3556</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.8823</td>
<td>0.6222</td>
<td>0.4136</td>
<td>0.3452</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-10</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.8823</td>
<td>0.2503</td>
<td>0.2210</td>
<td>0.1751</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Recognition Accuracy with the Baseline and the Proposed methods

<table>
<thead>
<tr>
<th>No</th>
<th>SNR in dB</th>
<th>Fingerprints</th>
<th>Voice</th>
<th>Equal weighting</th>
<th>GS</th>
<th>GA</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>95.6522</td>
<td>98.6956</td>
<td>97.8261</td>
<td>100.0000</td>
<td>100.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>95.6522</td>
<td>91.4493</td>
<td>97.8261</td>
<td>97.8261</td>
<td>100.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>95.6522</td>
<td>69.4203</td>
<td>97.2464</td>
<td>97.1015</td>
<td>97.8261</td>
<td>97.7174</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>95.6522</td>
<td>33.1884</td>
<td>95.6522</td>
<td>95.6522</td>
<td>95.6522</td>
<td>93.4783</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>95.6522</td>
<td>23.1884</td>
<td>85.6522</td>
<td>95.6522</td>
<td>95.6522</td>
<td>93.4783</td>
</tr>
<tr>
<td>6</td>
<td>-5</td>
<td>95.6522</td>
<td>9.1304</td>
<td>68.6956</td>
<td>95.6522</td>
<td>95.6522</td>
<td>93.4783</td>
</tr>
<tr>
<td>7</td>
<td>-10</td>
<td>95.6522</td>
<td>5.6521</td>
<td>57.3913</td>
<td>95.6522</td>
<td>95.6522</td>
<td>93.4783</td>
</tr>
</tbody>
</table>

### 3.2. Leave-One-Out Cross Validation (LOOCV)

As the number of available biometric samples was limited, LOOCV strategy was employed to fine-tune the training/validation phase and estimate the best optimal weights. LOOCV involves partitioning the data samples into complementary subsets and using a single observation from the data sample as validation data, and the remaining observations as training data. To reduce variability, multiple rounds of cross-validation were performed using different partitions, and the validation results were averaged over the rounds. Here, seven training samples from each modality were used for cross validation. So, seven cross validation sets were considered from
both modalities (Val.1 to Val.7). We choose each sample from the training set, for cross validation testing and the remaining six samples for cross validation training one at a time.

The average training/validation accuracy of the individual classifiers is depicted in Table 1. The d-prime statistics and the scale factor $k_{opt}$ for various noise conditions were obtained with LOOCV technique. After performing the validation (testing), match score outputs from each validation set were stored separately. The d-prime values for each noise conditions were obtained from the genuine and impostor scores of the match score outputs of the two modalities.
(validation sets) using equation (2). The d-prime values for each noise conditions were used to find the best scaling factor $k_{\text{opt}}$, using equation (6) to maximize the objective function. A more detailed description of the LOOCV procedure we adopted is given in [1].

4. RESULTS AND DISCUSSIONS

Finger images from the FVC2002 fingerprint database [14] and voice samples from ELSDSR database [15] have been employed for the experimentation. ELSDSR database contains nine text independent speech samples of twenty-three persons. So, finger images of twenty-three different persons with nine impressions per finger were considered. Out of these nine samples per 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 were investigated by artificially degrading the test samples with additive white Gaussian noise. The performance of the system with varying SNR conditions were considered systematically from the feature extraction and model building

![Figure 5: Score density plot for GS based optimization (Baseline)](image)

![Figure 6: Score density plot for GA based optimization (Baseline)](image)
stage to the testing stage. MFCC feature vectors of the order 12, 16 and 20 and the GMM with 12 and 16 mixtures were considered for the simulation studies, as they are widely used. Different model combinations were considered to select the best model that gives better recognition accuracy. The voice model with 16 MFCC feature vectors with 12 Gaussian mixtures gave improved recognition accuracy under normal operating conditions (10 dB SNR to 20 dB SNR). So, this model combination was considered for the subsequent analysis.

4.1. Fusion with Baseline Systems

The outputs of the different classifiers were consolidated into a single vector of scores using the sum rule of fusion. We explored the sum rule of fusion with equal and different weights to the two modalities.

4.1.1. Fusion with Equal Weightings

In this case a constant value of \( \beta = 0.5 \) was assigned as an integration weight at all SNR levels. The score transformation could be achieved by the following equation.

\[
S^{(\text{fus})} = \frac{1}{n} \sum_{j=1}^{n} S_j \quad ; n = 2
\]

\[
S^{(\text{fus})} = \frac{1}{2} [S^{(\text{sc})} + S^{(\text{fc})}]
\]

where \( S^{(\text{sc})} \) - Speaker Scores; \( S^{(\text{fc})} \) - Finger Scores. This technique will not favour one modality over another. More over the combined recognition accuracy may not be the maximum always. The testing accuracy of the bimodal system is shown in Table 4. Figure 2 depicts the DET performance curve.

4.1.2. Manually Determined Integration Weights

This is a heuristic approach and often needs a lot of expertise. Weighted averages of the individual scores were considered here. The integration weights were randomly chosen on a trial and error basis for different SNR conditions and the weights that give higher accuracy were considered as the best choice. Sum rule of fusion with integration weights can be represented as follows.

\[
S^{(\text{fus})} = \frac{1}{n} \sum_{j=1}^{n} w_j S_j \quad ; n = 2
\]

where \( \sum_{j=1}^{n} w_j = 1 \). This process required lot of time and effort and the results are not guaranteed to be optimal. The testing accuracy of the manually determined integration weight is shown in Table 2. The parameters \( w_f \) and \( w_v \) are the individual weights given to the fingerprint and voice modality respectively. The results show that for different noise conditions we need different weighting factors for both the modalities.

4.1.3. Fusion with Optimal Integration weights without Separability Information

Blind optimization of the parameter \( \beta \) using equation (1) was performed in [1] to maximise the recognition accuracy. Grid Search (GS) and Genetic Algorithm (GA) based optimization were performed to obtain the best integration weights. The testing accuracy of this method is shown in the Table 4. The method does show improved performance than the sum rule of fusion method with equal weighting and it also helps to automate the integration weights estimation.
process. Moreover, it showed improved recognition accuracy than any of the unimodal systems in the normal operating conditions and maintained the accuracy of the better unimodal ones for all adverse conditions. Further insight could be obtained from the DET plots [16] (Figure 2- Figure 4) and the score density plots (Figure 5- Figure 6). The disadvantage of the method is that at the extreme noise conditions the fusion module contributed zero weighting to the voice modality and there is no substantial reduction in FAR and FRR.

4.2. Performance of the Proposed Method

From the validation stage, we obtained optimal integration weights \((\beta = k_{opt}\alpha)\) for different noise conditions from -10dB to 20 dB. We have applied Particle Swarm Optimization technique for optimizing the integration weight factor. The overall validation accuracy of the individual classifiers for various SNR conditions is shown in the Table 1. The relative d-prime ratio estimates of the two modalities and the optimal integration weight \(\beta\) estimated for various noise conditions are shown in Table 3. The \(\beta\) values thus estimated during the training/validation stage is used for testing. The overall testing accuracy is depicted in Table 4. Though the recognition accuracy of the proposed method shows attenuating fusion at very extreme noise conditions (0dB, -5dB and -10dB), the proposed method gave better recognition accuracy above
0dB SNR. The added advantage of the proposed technique is that FAR remains more robust and is reduced considerably even for low SNR conditions. The DET performance plot (Figure 7) and the score density plot (Figure 8) depict this observation. Hence, these experimental studies reveal that, the FAR is substantially reduced when the d-prime statistic is employed as an auxiliary quality measure for finding the optimal integration weight. The score density plot indicates that the proposed method reduces the effective overlap between the genuine and the impostor score distributions.

5. CONCLUSION

PSO based integration weight optimization technique is proposed here for improving the performance of a multibiometric system under various noise conditions. We studied the performance of the proposed technique in the context of fingerprint and voice biometrics using score level fusion. The proposed method could successfully reduce the FAR under varying noise conditions. Moreover, by estimating the optimal integration weight using stochastic optimization technique and LOOCV techniques, we could automate the process and make the system robust to fluctuating inputs. Thus, the proposed method could save lot of overheads in terms of time and effort involved in the determination of the optimal integration weight. This method finds extremely useful in applications such as 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 (for SNR's < 0dB).

REFERENCES


12. Douglas Reynolds, Gaussian Mixture Models* MIT Lincoln Laboratory, 244 Wood St., Lexington, MA 02140, USA.


Authors

S. M. Anzar obtained his Bachelor’s Degree in Electronics and Instrumentation Engineering from Cochin University of Science and Technology, Cochin, India in 2003. He received the Masters of Engineering Degree in Applied Electronics from Anna University, Chennai, India in 2008. His general research interests include signal processing, image processing, pattern recognition, biometrics, and neural networks. Currently he is a research scholar in National Institute of Technology Calicut, India since July 2009.

P. S. Sathidevi is currently serving as Professor and Head in the Department of Electronics and Communication Engineering at National Institute of Technology Calicut, India. She received B.Tech. Degree in Electronics Engineering from Regional Engineering College Calicut, India in 1985, M.Tech. in Electronics from Cochin University of Science and Technology, Cochin, India in 1987 and Ph.D. from Regional Engineering College, Calicut, India in 2003, in the field of Speech and Audio Processing. Her current research interests include speech processing, perceptual audio coding, image processing, Computational Auditory Scene Analysis (CASA), cryptography, and biometrics. She has over 50 publications to her credit in various international journals and conferences.