Analytical Study of Biometrics Normalization and Fusion Techniques For Designing a Multimodal System

Divya Singhal\(^1\), Ajay Kumar Yadav\(^2\)

\(^1\) M.Tech, EC Department, Mewar University Chittorgarh, Rajasthan, INDIA
\(^2\) Asst. Professor, College of Engineering and Rural Technology Meerut INDIA

Abstract
Unimodal biometric systems are reliable; still the popularity of these systems is degrading due to the problems such as noisy sensor data, non-universality and spoof attacks. These limitations can be overcome if more than one trait is fused to make a system called multi-biometric system. In this paper, I am presenting different techniques for the fusion of different traits to make a multimodal system after normalization of all features in one domain. Also, a comparative study of different normalization methods has been done. Performance is analyzed on basis of two parameters: Genuine acceptance rate (GAR) and false acceptance rate (FAR). These results obtained through MATLAB programs yields curves for GAR and FAR.

Keywords: Normalization techniques, multimodal biometric system, fusion methods.

2. INTRODUCTION

2.1 Multimodal Systems
A multi-modal biometric system conceives more than one biometric modality in the process of identification. In one of the research work develop by Ko [1] proposes that multi-modal fusion welfares the most when the biometric cues are extraneous. These biometric evidences can be considered extraneous to each other when the match performance of one cue does not anticipate the others operation. It is supposed that ideally all of the biometric evidences would be extraneous to each other, at the same instant captured with the same sensor, and at superiority. However, it is beneficial to have few algorithm of categorizing multi-biometric systems, and realizing the rewards and hinders linked with each approach is essential to effective system invention.

2.2 Score Level Fusion
It is obvious for every biometric system that they possess the information at sensors level in raw form only. Due to the incompatibility among the samples of different trait, in raw form, it is very unusual and impractical to fuse the acquired information at sensor level. However, at feature extraction level, if a salient piece of information is looked for and gathered, the feature vectors out of raw data can be created. Although, feature vectors contain a sufficient amount of information, achieving fusion is a bit complicated due to incompatibility in feature sets [2, 3].

At decision level, user claims an identity to be verified or identified; depending upon the screening is either 1:1 or 1:N. In addition to this, the possibility of ties among the classifier results, make it impractical for fusion and the use of more number of classifier makes the system complex. When various feature vectors are compared with matchers, matching scores are generated. The main advantage of matching score level fusion is that either the problem of incompatibility of scores is eliminated or the system can be trained. There are several methodologies have been developed for score level fusion which shows the importance of this level, such as:
• Transformation-based score level fusion
• Classifier-based score level fusion
• Density-based score level fusion

In this thesis, on transformation based fusion work is done; the matching score are needed to be normalized or transform to convert them in common domain.

2.3 Normalization
Normalization is a method to convert the matching scores obtained from the different matchers in a common domain [4]. In other words, normalization is used to unionize the database and to eliminate the inconsistency in the data. An effective transformation scheme not only estimates the location but also the scale parameters of the database. The matching scores obtained after normalization must be robust and efficient over the entire distribution. Robustness is necessary in case if outliers are present in the distribution and efficiency is required as to check the proximity of the estimated distribution [5]. But, the main issue is to select a technique which is robust and efficient in nature. Many normalization techniques have been discussed in the literature for the transformation of the data such as, Min-Max normalization, Z-Score normalization, Tanh-estimators normalization, Reduction of high- scores effect (RHE) normalization, Adaptive Normalization, Mathematical Function Normalization, Double Sigmoid Function Normalization technique, Decimal scaling etc. In this thesis, Adaptive Normalization technique, Mathematical Function Normalization technique, Double Sigmoid Function Normalization technique and Z-score normalization technique [6] have been used to evaluate the performance of a biometric system.

2.3.1 Z-Score Normalization: z-score normalization technique estimates mean and standard deviation of matching scores to normalize the entire distribution. If μ and σ are the mean and standard deviation of the given database then normalized scores are given as [4]:

\[ s'_k = \frac{s_k - \mu}{\sigma} \]

2.3.2 Mathematical Functional Normalization: This approach employs a mathematical function which has two different forms; one is used for dissimilarity matching scores and other for the similarity matching scores. After normalization, the whole distribution spreads in the range of 0 and 1, i.e. the minimum values approaches toward 0 and maximum toward 1. This method is both efficient and robust in nature because it does not estimate the distribution and reduces the effect of outliers too. If \( s_k \) is the original matching score then normalized scores \( s'_k \) are given by [6].

\[ s'_k = \frac{1}{2} \left( 1 - \frac{s_k}{\sqrt{s_k^2 + \alpha}} \right) \]  

and

\[ s'_k = \frac{s_k}{\sqrt{s_k^2 + \alpha}} \]  

Equation (1) and (2) give the transformation of dissimilarity and similarity scores respectively in the range of [0, 1]. Where, \( \alpha \) is a constant which is known as smoothing parameter, i.e. it removes the irregularities from the distribution when assigned the higher value (e.g. \( \alpha = 50, 100 \ldots \)).

2.3.3 Adaptive Normalization: The errors of individual biometric matchers stem from the overlap of the genuine and impostor distributions. This region is characterized with its center \( c \) and its width \( w \). To decrease the effect of this overlap on the fusion algorithm, an adaptive normalization procedure is proposed that aims to increase the separation of the genuine and impostor distributions. [7]

\[ n_{AD} = \frac{1}{(c-w)^2} n_{MM}^2 \leq \frac{(c-w)}{2} \]

\[ n_{MM} \leq \frac{(c-w)}{2} < n_{MM} \leq \frac{(c+w)}{2} \]

\[ (c+w)^2 + \sqrt{(1-(c-w))(n_{MM}-c-w)^2}, otherwise \]

Where \( n_{MM} \) = normalized score by min max normalization.

2.3.4 Double Sigmoid Function Normalization: Cappelli et al. have used a double sigmoid function for score normalization in a multimodal biometric system that combines different fingerprint classifiers. The normalized
The input pattern delegated to the class \( c \) is given by \([50]\):

\[
P(w_j \mid \mathbf{x}_i) = \arg \max_{c} \left( \prod_{j=1}^{R} PW_j \right) |\mathbf{X}_i|
\]

These rules for \( R \) number of matchers are given below.

3.1.1 Sum Rule: This is one of the productive rules because it eliminates the problem of equivocalness during classification. In sum rule, transformed scores of every class are added together to get the final score. Here, input pattern is delegated to the class \( c \) such that \([4]\):

\[
c = \arg \max_{j} \sum_{i=1}^{R} PW_j |\mathbf{X}_i|
\]

3.1.2 Product Rule: The product rule provides a less intended results than sum rule because it is based on the statistical independence of the feature vectors. The input pattern delegated to the class \( c \) is given by \([50]\):

\[
c = \arg \max_{j} \prod_{i=1}^{R} PW_j |\mathbf{X}_i|
\]

3.1.3 Min Rule: In this rule, a minimum posterior probability is collected out of all classes. Hence, the input pattern delegated to the class \( c \) such that \([4]\):

\[
c = \arg \max_{j} \min_{i} PW_j |\mathbf{X}_i|
\]

3.1.4 Max Rule: In max rule, the posterior probability is approximated by the maximum value of the input pattern. The input pattern delegated to the class \( c \) is given by \([4]\):

\[
c = \arg \max_{j} \max_{i} PW_j |\mathbf{X}_i|
\]

3.2 Supervised Methods: The supervised methods of fusion make up a major factor in integration and decision making. Especially, a healthy number of studies in the field of pattern classification can be followed freely without any ambiguity for fusion and decision. The basis of the fusion work in this paper is unsupervised methods so the discussion to unsupervised methods of fusion only is done.

4. PROCEDURE FOR FUSION APPROACH

For the procedural examination of fusion approach, various set of matching scores are selected. Every time these data sets are normalized for the distribution of different scores in a common domain. These data sets are normalized using above discussed normalization techniques. This is to mention here while transforming the data in a common domain the removal of the outliers in from the matching score is not considered. Since the presence of outliers affect the ultimate result of the normalization so a comparative study is taken in to account to check the robustness and efficiency of every
method. A novel approach which utilizes a mathematical function also used in the model development and has shown an appreciable performance over other methods. The evaluation of the unsupervised rule (sum, product, min and max) based score level fusion, is demonstrated in the form of FAR, FRR and GAR. In order to calculate FAR, FRR and GAR from the matching scores, the setting up a minimum value after the fusion in the form of threshold is required in order to identify a person as an imposter or as genuine. The FAR and FRR value are the percentage of falsely accepted impostor scores divided by the total number of impostor scores and falsely rejected genuine scores divided by total number of rejection, respectively. The GAR value is the percentage of the number of genuinely given an entry divided by the total number of genuine scores. Hence, performance evaluation of unsupervised rules based score level fusion is dependent on FAR, FRR and GAR values obtained for different data set.

The purpose of analytical study is to investigate how multiple biometric modalities can be fuse for the creation of an effective authentication system. In this thesis, the fusion of different biometric traits using transformation schemes and fusion rules is examined, and it is evaluated that every rule has its different advantages and drawbacks. The significant distinction between these methods has been made on the basis of recognition rates. It is clear from the performance tables that on an appropriate database, a good percentage of GARs over FARs & FRRs can be achieved with ease.

5. RESULTS

5.1 Performance Analysis on The Basis of Genuine Acceptance Rate (GAR) and False Acceptance Rate (FAR)

The operational consequences of unsupervised rules based fusion strategy have been evaluated in terms of Genuine Acceptance Rate (GAR) and False Acceptance Rate (FAR) for various thresholds. The GARs and FARs for all normalizations and fusion rules are computed as follows.

5.1.1 GAR and FRR Calculation for four Normalizations with sum rule: The Genuine Acceptance Rate and False Acceptance Rate for four normalizations with sum fusion rule have been evaluated and are shown in table 5.1.

<table>
<thead>
<tr>
<th>Sum Rule</th>
<th>Threshold 6 users</th>
<th>Threshold 10 users</th>
<th>Threshold 100 users</th>
<th>6 users</th>
<th>10 users</th>
<th>100 users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z Score</td>
<td>12.678</td>
<td>4.4508</td>
<td>5.8439</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Mathematical</td>
<td>0.7889</td>
<td>0.7888</td>
<td>0.7992</td>
<td>10</td>
<td>10</td>
<td>95</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.9908</td>
<td>0.7129</td>
<td>0.8271</td>
<td>10</td>
<td>10</td>
<td>99.99</td>
</tr>
</tbody>
</table>

5.1.2 GAR and FRR Calculation for four Normalizations with product rule: The Genuine Acceptance Rate and False Acceptance Rate for four normalizations with product fusion rule have been evaluated and are shown in table 5.2.

<table>
<thead>
<tr>
<th>Product Rule</th>
<th>Threshold 6 users</th>
<th>Threshold 10 users</th>
<th>Threshold 100 users</th>
<th>6 users</th>
<th>10 users</th>
<th>100 users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z Score</td>
<td>24.844</td>
<td>1.881</td>
<td>0.8047</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Mathematical</td>
<td>0.0049</td>
<td>0.0049</td>
<td>0.0039</td>
<td>100</td>
<td>10</td>
<td>98</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.0032</td>
<td>0.0024</td>
<td>0.0846</td>
<td>100</td>
<td>10</td>
<td>98</td>
</tr>
</tbody>
</table>

5.1.3 GAR and FRR Calculation for four Normalizations with min rule: The Genuine Acceptance Rate and False Acceptance Rate for four normalizations with min fusion rule have been evaluated and are shown in table 5.3.

| Double Sigmoid | 0.0231 | 0.0207 | 0.0439 | 90     | 10     | 60     |

www.ijaert.org
Table 5.3: GAR and FRR for four normalizations with min fusion rule.

<table>
<thead>
<tr>
<th>Min Rule</th>
<th>Threshold for 6 users</th>
<th>Threshold for 10 users</th>
<th>Threshold for 100 users</th>
<th>6 users</th>
<th>10 users</th>
<th>100 users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z Score</td>
<td>1.2306</td>
<td>0.6549</td>
<td>0.1197</td>
<td>100</td>
<td>0</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>0.6549</td>
<td>0.1197</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mathematical</td>
<td>0.0287</td>
<td>0.0287</td>
<td>0.0291</td>
<td>100</td>
<td>95</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.0174</td>
<td>0.0174</td>
<td>0.0438</td>
<td>83.3</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.6</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Double Sigmoid</td>
<td>0.0078</td>
<td>0.1794</td>
<td>0.1636</td>
<td>80</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>0.1794</td>
<td>0.1636</td>
<td></td>
<td>20</td>
<td>30</td>
<td>25</td>
</tr>
</tbody>
</table>

The results shown above are derived for 6 and 10 user data. The same procedure for the database of 100 users have also conducted, and have shown the results in graphs.

5.2 MATLAB Simulation results

(a) (b) (c) (d) (e)

5.1.4 GAR and FRR Calculation for four Normalizations with max rule: The Genuine Acceptance Rate and False Acceptance Rate for four normalizations with max fusion rule have been evaluated and are shown in table 5.4.

Table 5.4: GAR and FRR for four normalizations with max fusion rule.

<table>
<thead>
<tr>
<th>Max Rule</th>
<th>Threshold for 6 users</th>
<th>Threshold for 10 users</th>
<th>Threshold for 100 users</th>
<th>6 users</th>
<th>10 users</th>
<th>100 users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z Score</td>
<td>9.3541</td>
<td>2.6687</td>
<td>4.1582</td>
<td>10</td>
<td>0</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>2.6687</td>
<td>4.1582</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mathematical</td>
<td>0.4642</td>
<td>0.4310</td>
<td>0.4300</td>
<td>10</td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0.5194</td>
<td>0.4064</td>
<td>0.7675</td>
<td>99</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>0.4064</td>
<td>0.7675</td>
<td></td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Double Sigmoid</td>
<td>0.0143</td>
<td>0.0611</td>
<td>0.0187</td>
<td>95</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>0.0611</td>
<td>0.0187</td>
<td></td>
<td>5</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>
Fig 5.1 (a), (b), (c) and (d) depicts the GARs and FARs for sum rule, product rule, min rule and max rule fusion respectively through adaptive normalization. Fig 4.1 (e), (f), (g) and (h) depicts the GARs and FARs for sum rule, product rule, min rule and max rule fusion respectively through Z score normalization.

6. CONCLUSION

The operational performance of unsupervised rule-based fusion devolves on the alternative of standardization technique.

REFERENCES