Score Normalization in Multimodal Biometric Systems

Karthik Nandakumar and Anil K. Jain

Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan 48824

nandakum@cse.msu.edu, jain@cse.msu.edu

Arun A. Ross

Lane Department of Computer Science and Engineering, West Virginia University, Morgantown, West Virginia 26506

ross@csee.wvu.edu

1. Introduction

Unimodal biometric systems are often affected by several practical problems like noisy sensor data, non-universality and/or lack of distinctiveness of the biometric trait, unacceptable error rates, and spoof attacks. Multimodal biometric systems overcome some of these problems by consolidating the evidence obtained from different sources [1]. In a multimodal biometric system, various levels of fusion are possible: fusion at the feature extraction level, matching score level or decision level. Integration at the matching score level is generally preferred due to the ease in accessing and combining matching scores.

2. Fusion at the Matching Score Level

In the context of verification, there are two approaches for consolidating the scores obtained from different matchers. One approach is to formulate it as a classification problem, where a feature vector is constructed using the matching scores output by the individual matchers; this feature vector is then classified into one of the two classes: "Genuine user" or "Impostor". In the combination approach, the individual matching scores are combined to generate a single scalar score, which is then used to make the final decision. Ross and Jain [2] have shown that the combination approach performs better than some classification methods like decision tree and linear discriminant analysis. In this paper, we use the combination approach and address some of the issues involved in computing a single score from the scores of the different modalities.

3. Score Normalization

The matching scores at the output of the individual matchers may not be homogeneous. For example, one matcher may output a dissimilarity measure while another may output a similarity measure. Further, the scores of the individual matchers need not be on the same numerical scale and may follow different statistical distributions. Due to these reasons, score normalization is essential to transform the scores of the individual matchers into a common domain prior to combining them. Score normalization is a critical part in the design of a combination scheme for matching score level fusion.

Score normalization refers to changing the location and scale parameters of the matching score distributions at the output of the individual matchers, so that the scores of different matchers are transformed into a common domain. In a good normalization scheme, the estimates of the location and scale parameters must be *robust* and *efficient*. ***Robustness*** refers to insensitivity to the presence of outliers. ***Efficiency*** refers to the proximity of the obtained estimate to the optimal estimate when the distribution of the data is known. Huber [3] explains the concepts of robustness and efficiency of statistical procedures and emphasizes the need to have both these desirable characteristics. Although many techniques can be used for score normalization, the challenge lies in identifying a technique that is both robust and efficient.

Snelick et al. [4] evaluated the effects of normalization techniques like min-max, z-score, median and MAD, and tanh estimators and fusion methods like simple sum of scores, max rule, min rule, sum rule, and product rule on the performance of a multimodal biometric system using face and fingerprint modalities. Their experiments showed that the min-max normalization followed by the sum of scores fusion method outperform other schemes. However, they do not offer any reasons for such a behavior. In this paper, we have systematically studied the different normalization techniques to ascertain their role in the performance of a multimodal system consisting of face, fingerprint and hand-geometry modalities. In addition to the normalization techniques employed in [4], we have also analyzed the double sigmoid normalization. We have used the Parzen window density estimation method for converting the scores into probabilities. This non-parametric method does not require the scores to have a normal distribution and can also make use of the prior probabilities of the genuine and impostor users that may be available to the system.

4. Experimental Results

Experiments conducted on a database of 100 users (each user provided 5 instances of their fingerprint, face, and hand-geometry) show that a multimodal system employing the sum of scores method provides better performance than the best unimodal system (fingerprint in this case) for all normalization techniques except median-MAD normalization (see Figure 1). The distribution of the fingerprint scores deviates wildly from the Gaussian assumption and hence, median and MAD are poor estimates of location and scale, respectively. In this case, the combined score is approximately equal to the fingerprint score and the performance of the multimodal system is close to that of the fingerprint module. Among the other normalization techniques, we observe that the tanh and minmax normalization techniques outperform other techniques at low FARs. At higher FARs, z-score normalization provides slightly better performance than tanh and min-max normalization. We have also demonstrated the sensitivity of the min-max and z-score normalization techniques to the presence of outliers, by artificially introducing outliers in the fingerprint scores (see Figures 2 and 3). On the other hand, Figure 4 shows that the tanh normalization technique is not affected by the introduction of outliers and hence, it is a more robust normalization technique.

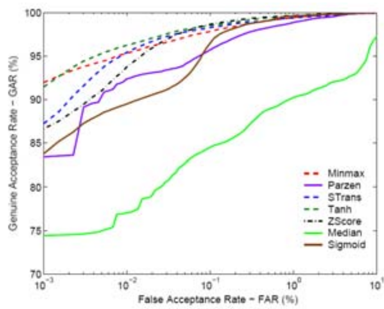

Fig. 1. ROC curves for sum of scores fusion method

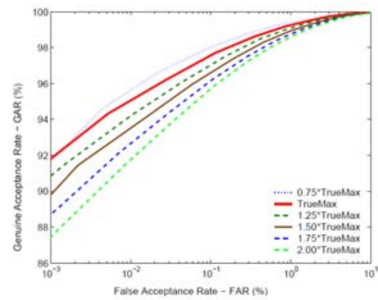

Fig. 2. Robustness analysis of min-max normalization

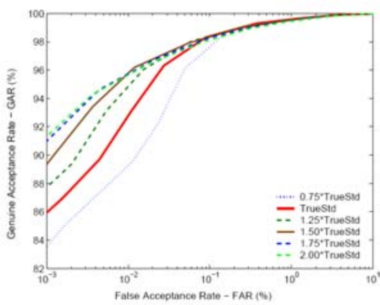

Fig. 3. Robustness analysis of z-score normalization

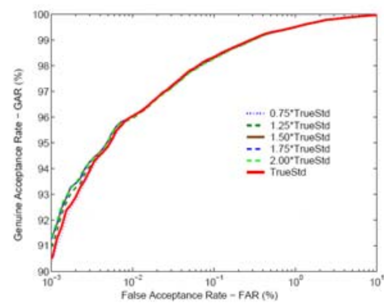

Fig. 4. Robustness analysis of tanh normalization

5. Conclusion and Future Work

We have studied the effect of different score normalization techniques in a multimodal biometric system. Min-max, z-score, and tanh normalization techniques followed by a simple sum of scores fusion method result in a higher GAR than all the other normalization and fusion techniques. We have shown that both min-max and z-score methods are sensitive to outliers. On the other hand, tanh normalization method is both robust and efficient. We need to develop rules that allow a practitioner to choose a normalization scheme after analyzing the genuine and impostor score distributions of the individual matchers. Alternative normalization and fusion techniques such as those based on Bayesian statistics also need to be explored.

6. References

[1] A. K. Jain and A. Ross, "Multibiometric Systems," Communications of the ACM 47(1), 34-40 (2004).

[2] A. Ross and A. K. Jain, "Information Fusion in Biometrics," Pattern Recognition Letters 24 (13), 2115–2125 (2003).

[3] P. J. Huber, *Robust Statistics* (John Wiley & Sons, 1981).

[4] "Multimodal Biometrics: Issues in Design and Testing," in Proceedings of Fifth International Conference on Multimodal Interfaces, R. Snelick, M. Indovina, J. Yen, and A. Mink, (Vancouver, 2003), pp. 68–72.

[5] "Combining COTS Finger and Face Biometrics for Identity Verification," in Proceedings of Workshop on MultiModal User Authentication, M. Indovina, U. Uludag, R. Snelick, A. Mink, A. K. Jain, (Santa Barbara, 2003), pp. 99-106.