MULTI-BIOMETRIC COHORT ANALYSIS FOR BIOMETRIC FUSION

Gaurav Aggarwal[††]*, Nalini K. Ratha*[†]*, Ruud M. Bolle*[†] *and Rama Chellappa*[††]

[††]Center for Automation Research
University of Maryland, College Park
{gaurav, rama}@cfar.umd.edu

[†]IBM T. J. Watson Research Center
Hawthorne, NY, USA
{ratha, bolle}@us.ibm.com

ABSTRACT

Biometric matching decisions have traditionally been made based solely on a score that represents the similarity of the query biometric to the enrolled biometric(s) of the claimed identity. Fusion schemes have been proposed to benefit from the availability of multiple biometric samples (e.g., multiple samples of the same fingerprint) or multiple different biometrics (e.g., face and fingerprint). These commonly adopted fusion approaches rarely make use of the large number of non-matching biometric samples available in the database in the form of other enrolled identities or training data. In this paper, we study the impact of combining this information with the existing fusion methodologies in a cohort analysis framework. Experimental results are provided to show the usefulness of such a cohort-based fusion of face and fingerprint biometrics.

Index Terms— cohort analysis, multi-modal biometrics, classifier fusion

1. INTRODUCTION

Biometrics refers to the measurement and analysis of physical and behavioral traits of humans [1]. Such an analysis is often directed towards the goal of verifying or determining the identity of humans. Biometrics provide a more authentic alternative to establish identity as compared to passwords, ID cards, etc. which can be stolen or passed on to others fairly easily. A biometric characteristic should have the following characteristics for it to be truly useful in real scenarios:

- Universality (every person should have the biometric),
- Uniqueness (every person's biometric signature should be different from others),
- Permanence (the biometric should be invariant over time),
- Collectibility (an easy, quick, inexpensive, non-intrusive way to acquire the biometric),
- Acceptability (acceptable to people),
- Difficult to circumvent (spoof-proof), and
- low underlying system errors (FAR, FRR, etc).

No matter how good a matching algorithm is, it may not be possible for a single biometric to have all the mentioned desirable properties. This has led to the rise of research in multi-biometric systems that rely on fusing information from multiple biometric evidences. Fusion of multiple biometric characteristics has been shown to increase accuracy while decreasing the vulnerability to spoofing [2]. In addition, use of multiple biometrics provides a better coverage of population to deal with situations like indistinguishable unimodal biometric characteristic (like illegible/indistinguishable fingerprint).

Biometric fusion research so far has concentrated on fusing multiple (independent) biometrics, multiple samples of the same biometric and fusing matching scores for a single biometric sample obtained using multiple matching algorithms. Not much has been done to fuse vast number of non-matching biometric samples available in the database in the form of other enrolled biometrics and training data. The earlier works that make use of non-match templates have mostly been restricted to unimodal biometrics like speech [3] and fingerprint [4].

In this paper, we extract and combine information from the non-matching templates in a multi-biometrics framework. In cohort analysis [4], neighbors for each enrolled identity are identified as the pre-processing step. In the absence of a suitable statistical model for biometrics like face and fingerprints, the neighbor (cohort) selection is performed based on the raw similarity scores as provided by the available matcher. No assumption whatsoever is made on the nature of the biometric or matching algorithm. Given a query, it is compared not only with the claimed identity but also with the neighbors (cohort) of the claimed identity. The final similarity score is determined by combining the similarity score of the query with the claimed identity and its cohort. For multi-biometrics scenarios, the fusion of biometric characteristics can be performed either before cohort determination or late fusion after computing cohort-normalized scores independently for the multiple biometrics. Experimental evaluation shows that both the fusion strategies perform significantly better than direct fusion of raw similarity scores from the matcher.

1.1. Organization of the paper

The rest of the paper is organized as follows. The following section describes related literature on biometric fusion. Section 3 provides an insight into the cohort formulation for biometric fusion. Section 4 describes the proposed cohort-based approach to combine multiple biometrics for improved matching performance. Results of experimental evaluation performed to test the approach are shown in Section 5. The paper concludes with a brief summary in Section 6.

2. PREVIOUS WORK

There exists a large number of techniques for fusing biometric characteristics. Fusion has been performed at feature level, matching score level and decision level [5]. Kittler *et al.* [6] evaluated several fusion rules on frontal face, face profile and voice biometrics. Jain *et al.* [7] examine the effect of different score normalization techniques on the performance of a multimodal biometric system. Snelick *et al.* [8] compare combinations of several different normalization and fusion rules for matching level fusion of face and fingerprint biometrics. It was observed that though fusing biometrics perform bet-

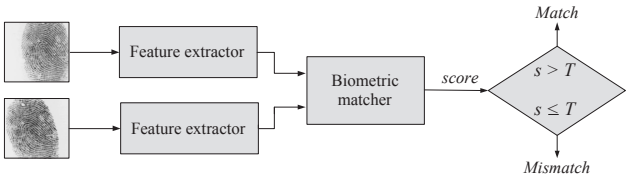

Fig. 1. A typical verification system. A matcher determines the similarity score s between two biometrics. The decision is made by comparing the similarity score with a suitable pre-set threshold T.

ter than unimodal systems, the performance gain is limited by the high performance of the unimodal systems. Ross and Jain [9] fuse face, fingerprint and hand geometry features for biometric verification. Their experimental results indicate that the sum rule performs better than the decision tree and linear discriminant classifiers.

3. COHORT ANALYSIS

Most biometric matching approaches make verification or identification decisions based purely on the similarity of the query with the enrolled biometric samples of the claimed identity (Fig. 1). The similarity is usually determined based on the distance of the query from the enrolled biometrics as determined by matching algorithm. To perform well, such approaches expect the biometric classes to be reasonably compact (around the available sample for each enrolled identity) with respect to the inter-class distances, and similarly distributed. When the class distributions vary across identities, the verification threshold (Fig. 1) may turn out to be too stringent for a few classes while too lenient for others. Additionally, biometric classes may not be isotropically distributed around the available sample(s) in feature space, making it difficult to even set a good threshold separately for each class. The performance of biometric systems gets particularly affected in situations when there are significant peculiarities that are not modeled by the matching algorithm. For example, illumination or pose variations in face, scanner quality in fingerprint, phone/microphone quality for speaker verification, etc. If the matching algorithm is unable to factor out these peculiarities effectively, the raw similarity scores obtained are dependent on these factors. This increases inter-class similarity scores while decreasing the intra-class ones.

Potentially these situations can be dealt with if the knowledge of class distributions is available. In most practical scenarios, learning these distributions is infeasible with just a few (often just one) samples per enrolled identity. It is in these situations that one can make use of large number of non-match biometric samples already present in the database. Normalizing the raw similarity score of the query with the claimed identity using its similarity with the neighbors of the claimed identity provides a sense of class distributions and normalizes for any unwanted peculiarities involved in raw similarity computation. Such a score normalization using neighbors of the claimed identity is termed as cohort analysis and has been shown to improve biometric matching performance significantly in unimodal scenarios [4]. In this paper, we build on our previous work [4] and perform biometric fusion in a cohort analysis framework to reap the benefits of the availability of multiple evidences and the non-match templates in the database for improved matching performance.

It is worthwhile to note that the cohort analysis differs from other popular techniques like Z-norm (zero normalization). These techniques are often based on the hypothesis that the output scores of each biometric class follow normal distribution which is not the case with cohort analysis [4]. Z-norm scheme normalizes each query using same impostor mean and variance (can be different for each enrolled identity but does not depend on the query). Unlike Z-norm, cohort analysis normalizes each query differently based on its similarity with the impostors of the target identity.

4. COHORT-BASED BIOMETRIC FUSION

In this section, we describe the proposed cohort analysis framework for biometric fusion using example of face and fingerprint fusion. Therefore, each biometric identity is characterized by a fingerprint and a face. It is assumed that we have access to matcher(s) that can provide fingerprint and face similarity scores. No assumption whatsoever is made on the nature of features or representation used or the goodness of the matching algorithm(s). All the similarity scores are normalized to lie in the range $[0, 1]$ by subtracting the minimum score in the database and dividing by the maximum one for each modality.

4.1. Cohort-based normalization

In cohort analysis framework for biometric matching [4], the similarity of a query with the claimed identity is computed as the ratio of its raw similarity with the claimed identity divided by the raw similarity with the cohort of the claimed identity $\bar{w}$, i.e.,

$$S(x, w) = \frac{s(x, w)}{s(x, \bar{w})}. \tag{1}$$

Here $s(x, \bar{w})$ is the similarity score of the query with the cohort. The raw similarity with the claimed identity can directly be determined using the available matcher. Assuming the cohort set to be of size k, $s(x, \bar{w})$ is determined using the following max-rule

$$s(x, \bar{w}) = \max\{s(x, w^1), s(x, w^2), \ldots, s(x, w^k)\}, \tag{2}$$

where $\{s(x, w^1), s(x, w^2), \ldots, s(x, w^k)\}$ is the set of similarity scores of the query with the cohort w^j's for the enrolled identity w.

Biometric fusion in this framework can be performed in the following two different ways.

- **Late fusion:** In this scheme, the two (can be more) biometrics are treated independently for cohort normalization and the combined score is obtained by fusing the final cohort-normalized scores of individual biometrics as follows

  $$S_f(x, w) = f(S_1(x, w), S_2(x, w)) \tag{3}$$

  Here, $S_f(x, w)$ denotes the final combined score of the two biometrics and f is a fusion function like simple sum rule or product rule. $S_1(x, w)$ and $S_2(x, w)$ denotes the cohort-normalized scores of the two individual biometrics as determined using (1). Quite clearly, one can use such a scheme for more than two biometrics.

- **Early fusion:** An alternative approach is to perform early fusion by combining the raw similarity scores before cohort-normalization. In early fusion, the raw similarity score $s(x, w)$ in (1) is replaced by the corresponding fused raw similarity score as follows

  $$s_f(x, w) = f(s_1(x, w), s_2(x, w)) \tag{4}$$

Here, $s_f(x,w)$ is the fused raw similarity score and $s_1(x,w)$ and $s_2(x,w)$ are the raw similarity scores for the two biometrics. Scores $\{s(x,w^1), s(x,w^2), \ldots, s(x,w^k)\}$ in (2) are also replaced by the corresponding fused scores in a similar fashion. Using these combined raw similarity scores, the final cohort-normalized score is obtained using

$$S_f(x,w) = \frac{s_f(x,w)}{s_f(x,\bar{w})}. \quad (5)$$

Experimental results indicate that both these schemes significantly outperform simple fusion of raw similarity scores.

4.2. Cohort selection

The described normalization scheme assumes knowledge of cohort (neighbors) for each enrolled identity. Though one can potentially use all available samples in the database as the cohort set, a large cohort set will make the proposed normalization scheme extremely inefficient. Therefore, as a pre-processing step, we need to determine cohort for each enrolled identity. Cohort can be chosen from a separate training data or from the enrolled identities themselves as done in this paper. For each enrolled identity, we compute its neighbors based on a combination of its face and fingerprint biometrics. The scores from the two modalities are combined using the simple sum rule, which is a simple addition of the scores obtained from the two modalities. Other combination rules we used included the product rule and L_2-norm of the two dimensional similarity vector (one dimension for face similarity and other for fingerprint). A preselected suitable number (10 in our experiments) of closest neighbors for each enrolled identity are selected based on the combined similarity measure. This corresponds to early fusion of biometrics for cohort selection. For late fusion scheme, one can select separate cohort sets for each biometric independently.

5. EXPERIMENTAL EVALUATION

We present the results of experimental evaluation performed to show the efficacy of the proposed cohort-based biometric fusion approach. We present biometric fusion results using a combination of early and late fusion alternatives described in the previous section. Fingerprint and face modalities are combined in our fusion experiments. We combine face and fingerprints from different datasets and associate them with each other to create a virtual multimodal face-fingerprint biometric dataset. The PIE dataset [10] is used for facial data while FVC 2002 [11] DB1 (Set A) database is used for fingerprints.

5.1. Database and matching algorithms

The FVC 2002 fingerprint dataset consists of eight fingerprints each of 100 different subjects. There is a significant variation in the quality of the eight copies of the same print. The raw similarity scores are computed using the NIST Fingerprint Image Software 2 [12]. The Bozorth 3 matcher included in the software is a minutiae-based matcher. The PIE face database consists of 68 subjects with variations in illumination, pose and expression. We use only the illumination part of the PIE dataset in our experiments. There are 21 images of each subject in 21 different illumination conditions. The face recognition approach proposed in [13] is used to generate the facial similarity scores. The combined bimodal dataset we create consists of 68 subjects with 8 copies of face-fingerprint signatures for each. For each verification experiment, one face-fingerprint sample per identity is enrolled in the gallery while the remaining 7 samples per identity are used as queries. Therefore, there are 68×7 genuine pairs while $67 \times 68 \times 7$ impostor pairs in each experiment.

5.2. Performance evaluation

Fig. 2 compares performance of the proposed cohort-based fusion approach with the simple raw similarity score based fusion. The fusion is performed using the simple sum rule for both raw and cohort-normalized scores. In the early fusion scheme, cohort selection is performed on the two biometrics jointly. In the late fusion scheme, cohort selection and cohort-based normalization for the two modalities is performed independently followed by fusion of cohort-normalized scores using the simple sum rule. As shown in Fig. 2, though fusion of raw scores performs worse than fingerprint by itself, both early and late cohort-based fusion schemes show significant performance improvement. The bad performance of the raw fusion is probably due to the poor performance of the face signatures. Interestingly, cohort-based approaches are able to account for this, thereby improving the overall fusion performance. We also perform fusion of raw scores using weighted sum rule giving higher weight to fingerprint similarity scores. As shown in the figure, though weighted sum rule improves the performance of raw fusion, the performance is still much worse than the cohort-based fusion schemes.

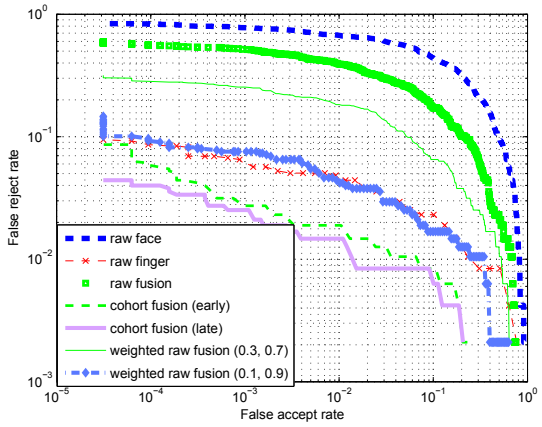

Fig. 2. Receiver Operator Characteristic (ROC) curves showing the verification performance of the proposed cohort-based biometric fusion approach. Simple sum rule is used to fuse both raw and cohort-normalized scores. (Best viewed in color)

Fig. 3 and Fig. 4 shows the genuine-impostor score distributions for the two modalities and fusion approaches. There is a big overlap in genuine and impostor score distributions for the face modality (Fig. 3) leading to bad verification performance which negatively affects the raw similarity score based biometric fusion (Fig. 4). This leads to a verification performance which is worse than using fingerprint alone for matching. On the other hand, the proposed cohort-based fusion scheme nicely separates the genuine and impostor distributions leading to good matching performance.

We also evaluate the biometric fusion performance using the product rule. In the product rule, the $[0,1]$-normalized scores from the two modalities are multiplied to obtain a combined score. For cohort-based fusion, this amounts to using product rule-based com-

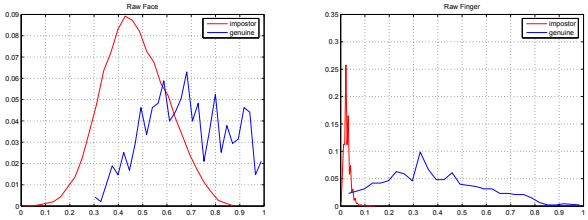

Fig. 3. Genuine and impostor score distributions of the raw similarity scores for face (left) and fingerprint (right) modalities

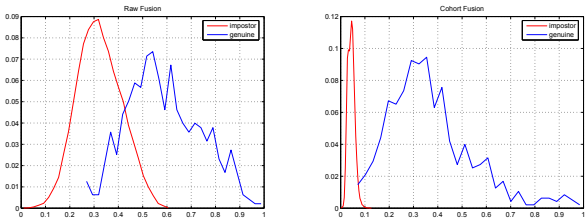

Fig. 4. Genuine and impostor score distributions of the fused similarity scores using raw similarity scores (left) and proposed cohort-based fusion approach.

bined similarity for cohort selection and cohort-based normalization. Fig. 5 shows the verification performance obtained in this experiment. As shown in the plot, the raw score-based fusion performs much better than it did using the simple sum rule in the previous experiment. The proposed cohort-based fusion again performs significantly better than raw-similarity based fusion and only fingerprint based similarity measure.

6. SUMMARY

In this paper, we combined popular biometric fusion techniques with cohort analysis to improve biometric matching performance. Unlike existing multi-biometrics approaches, the proposed approach utilizes the information present in the form of large number of non-match biometric samples present in the database, resulting in matching performances significantly better than those of the existing techniques. Both early and late fusion schemes in the cohort framework show much better performance than direct fusion of the similarity scores.

7. REFERENCES

[1] R. Bolle, J. H. Connell, S. Pankati, N. K. Ratha, and A. W. Senior, *Guide to Biometrics*, Springer verlag, 2003.

[2] A. K. Jain, R. Bolle, and S. Pankati, *Biometrics: Personal Identification in Networked Society*, Springer verlag, 2003.

[3] A. Higgins, L. Bahler, and J. Porter, "Speaker verification using randomized phrase prompting," *Digital Signal Processing*, vol. 1, no. 2, pp. 89–106, 1991.

[4] G. Aggarwal, N. K. Ratha, and R. M. Bolle, "Biometric verification: Looking beyond raw similarity scores," in *Proceedings of IEEE Computer Society Workshop on Biometrics (CVPR)*, june 2006.

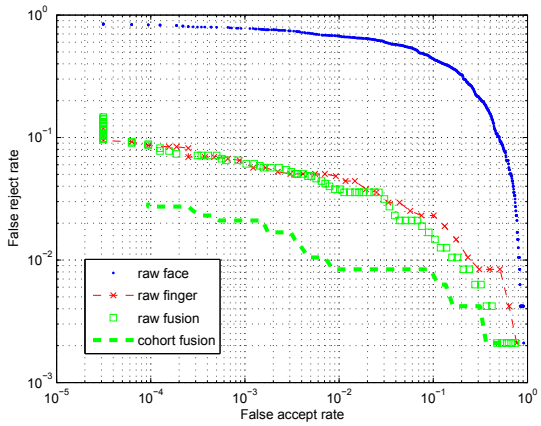

Fig. 5. ROC curves showing the verification performance of the proposed cohort-based biometric fusion approach. Product rule is used to fuse both raw and cohort-normalized scores. Similar to the sum-rule, both early and late cohort-based fusion schemes give similar performance in this experiment. Therefore, only one is shown for clarity.

[5] D. Maltoni, D. Maio, and A. K. Jain, *Handbook of Fingerprint Recognition*, Springer verlag, 2003.

[6] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, March 1998.

[7] A. K. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vol. 38, pp. 2270–2285, 2005.

[8] R. Snelick, U. Uludag, A. Mink, M. Indovina, and A. Jain, "Large scale evaluation of multimodal biometric authentication using state-of-the-art systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 450–455, March 2005.

[9] A. Ross and A. K. Jain, "Information fusion in biometrics," *Pattern Recognition Letters*, vol. 24, pp. 2115–2125, 2003.

[10] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," vol. 25, no. 12, pp. 1615–1618, 2003.

[11] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "Fvc2002: Second fingerprint verification competition.," in *Proceedings of the 16th International Conference on Pattern Recognition (3)*, 2002, pp. 811–814.

[12] C. Watson and M. Garris, "Nist fingerprint image software 2 (nfis2) http://fingerprint.nist.gov/nfis/index.html," .

[13] G. Aggarwal and R. Chellappa, "Face recognition in the presence of multiple illumination sources," in *International Conference on Computer Vision (ICCV)*, 2005, pp. 1169–1176.