Quality-based Score Level Fusion in Multibiometric Systems*

Karthik Nandakumar, Yi Chen, Anil K. Jain
Department of Computer Science and Engineering
Michigan State University

Sarat C. Dass
Department of Statistics and Probability
Michigan State University

Abstract

The quality of biometric samples has a significant impact on the accuracy of a matcher. Poor quality biometric samples often lead to incorrect matching results because the features extracted from these samples are not reliable. Therefore, dynamically assigning weights to the outputs of individual matchers based on the quality of the samples presented at the input of the matchers can improve the overall recognition performance of a multibiometric system. We propose a likelihood ratio-based fusion scheme that takes into account the quality of the biometric samples while combining the match scores provided by the matchers. Instead of estimating the quality of the template and query images individually, we estimate a single quality metric for each template-query pair based on the local image quality measures. Experiments on a database of 320 *users with iris and fingerprint modalities demonstrate the advantages of utilizing the quality information in multibiometric systems.*

1. Introduction

Several studies have shown that the poor quality of biometric samples leads to a significant reduction in the accuracy of a unimodal biometric system [13, 3, 4]. Multibiometric systems integrate the evidence presented by multiple biometric sources. Such systems are more robust to variations in the sample quality than unimodal systems due to presence of multiple (and usually independent) pieces of evidence. However, the accuracy of a multibiometric system can be further improved by estimating the quality of the biometric samples and adaptively weighting the individual matchers based on the quality values.

Several schemes have been proposed to exploit the signal quality in fusion at the match score level. Bigun et al. [2] developed a model based on Bayes theory to normalize and combine the match scores provided by multiple matchers. In their model, the quality of the match score is used to estimate the variance of the score distribution. Toh et al. [12] use multivariate polynomials that take into account the quality information for fusing fingerprint and voice. Fierrez-Aguilar et al. [8] employ a SVM classifier to classify a 2-D score vector (with fingerprint and signature match scores) into genuine and impostor classes. Samples with good quality are assigned higher cost of misclassification during the training of the SVM classifier. Quality-weighted sum rule has been used for fusion of multiple fingerprint matchers in [7] and for fusion of face and voice matchers in [11]. Baker et al. [1] combine match scores corresponding to all ten fingers of a person using a Bayesian belief network.

We present a likelihood ratio based approach to perform quality-based fusion of match scores in a multibiometric system (see Figure 1). We employ the generalized likelihood ratio-based (GLRF) scheme for fusion [5] since it circumvents the need for score normalization and selection of optimal weights for fusion on a case-by-case basis. We extend the GLRF scheme to account for the quality of the samples presented to each matcher. The effectiveness of the proposed fusion scheme is demonstrated on a multimodal database with fingerprint and iris modalities. Rather than estimating the quality of the template and query images individually, we compute a single quality metric for each template-query pair based on the local quality measures. Our quality metric represents the quality of the matching process, given the template-query pair.

2. Automatic quality assessment

We determine the quality of local regions in fingerprint and iris images and utilize the local quality to derive an overall quality of the match between each pair of template and query images.

2.1. Fingerprint quality

We estimate the local quality in a fingerprint image using the coherence measure described in [3]. Let T_f and Q_f represent the template and the query fingerprint images, respectively. We partition T_f and Q_f into blocks of size 12×12 pixels and estimate the coherence γ and γ' for each

*This research was supported by NSF ITR grant # CNS-0325640

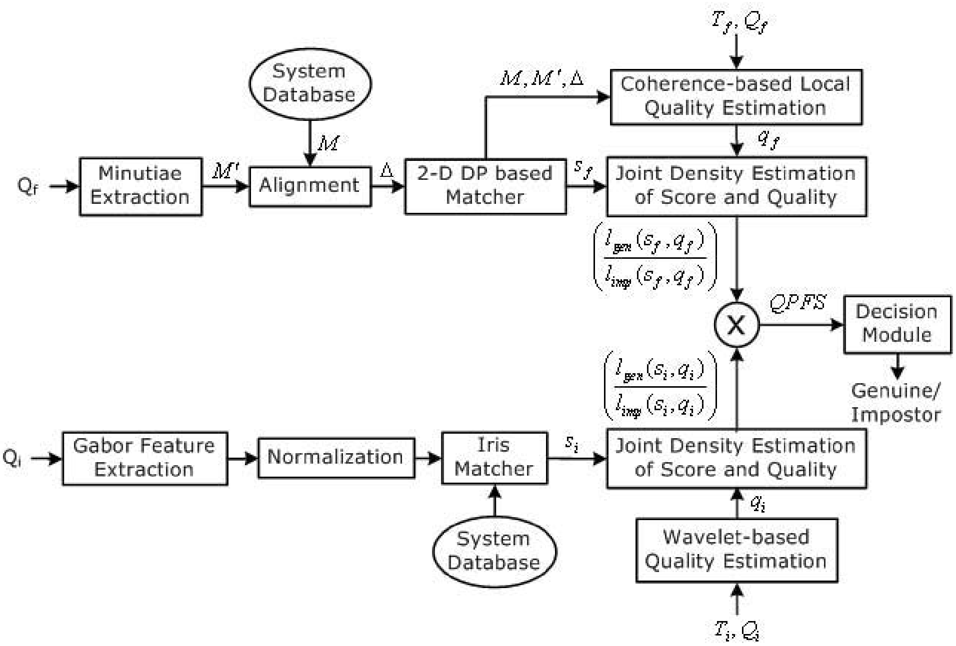

Figure 1. Framework for quality-based fusion. Here, T_f and Q_f (T_i and Q_i) represent the template and query fingerprint (iris) images, s_f and q_f (s_i and q_i) represent the fingerprint (iris) match score and quality, M (M') represents the template (query) fingerprint minutiae, Δ represents the rigid transformation between T_f and Q_f and QPFS represents the quality-based product fusion score.

block in T_f and Q_f, respectively. Let $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_m$ be the m minutiae in T_f, where $\boldsymbol{M}_i = \{x_i, y_i, \theta_i\}$, $i = 1, \ldots, m$. Let $\boldsymbol{M}'_1, \ldots, \boldsymbol{M}'_n$ be the n minutiae in Q_f, where $\boldsymbol{M}'_j = \{x'_j, y'_j, \theta'_j\}$, $j = 1, \ldots, n$. Let $\gamma(x, y)$ and $\gamma'(x, y)$ be the quality (coherence) of the block which contains the location (x, y) in T_f and Q_f, respectively. Let $t(x, y, \Delta)$ be the rigid transformation function that transforms a point (x, y) in T_f to a point (x', y') in Q_f. Here, $\Delta = [\Delta x, \Delta y, \Delta\theta]$ represents the translation and rotation parameters which are estimated using the 2-D dynamic programming based minutiae matcher described in [9]. Let A and A' be the area of the fingerprint regions in the template and the query. The area of overlap, A_o, between the fingerprint regions of T_f and Q_f can be computed using Δ. The overall quality of the matching between the template and query fingerprint images, $q_f(T_f, Q_f)$, is then defined as follows.

$$q_f(T_f, Q_f) = \left(\frac{r_1 + r_2}{m + n}\right)\left(\frac{2A_o}{A + A'}\right), \text{where} \quad (1)$$

$$r_1 = \sum_{i=1}^{m} \gamma(x_i, y_i)\gamma'(t(x_i, y_i, \Delta)) \text{ and}$$

$$r_2 = \sum_{j=1}^{n} \gamma(t(x'_j, y'_j, -\Delta))\gamma'(x'_j, y'_j).$$

Here, $0 \leq q_f(T_f, Q_f) \leq 1$. Note that if a minutia point in the template (query) falls outside the fingerprint region of the query (template) image, then the quality of that minutia is set to zero. Given good quality template and query fingerprint images with large overlap, $q_f(T_f, Q_f) \approx 1$.

2.2. Iris quality

We estimate the quality of match between the template and query iris images using a modified version of the wavelet-based iris quality assessment scheme proposed in [4]. The template (T_i) and query (Q_i) iris images are segmented into iris and non-iris regions [4]. A 2-D isotropic Mexican hat wavelet filter is applied to the iris regions of T_i and Q_i at three different scales $(0.5, 1.0, 2.0)$ and the product of the responses at the three scales is obtained. In order to account for the variations in the pupil dilation, iris size and rotation, the rubber sheet model proposed by Daugman [6] is used to normalize the wavelet responses. Let $w_{r,s}$ be the product of the wavelet responses at the r^{th} radius $(r = 1, \ldots, R)$ and s^{th} angle $(s = 1, \ldots, S)$ in T_i and let $w'_{r,s}$ be the corresponding wavelet response in Q_i. The average wavelet response at each radius r is computed as w_r $(= \frac{1}{S}\sum_{s=1}^{S} w_{r,s})$ and w'_r $(= \frac{1}{S}\sum_{s=1}^{S} w'_{r,s})$ in T_i and Q_i, respectively. The quality of match between the template and query iris images, $q_i(T_i, Q_i)$, is defined as the correlation coefficient between the vectors $\boldsymbol{w} = [w_1, \ldots, w_R]$ and $\boldsymbol{w'} = [w'_1, \ldots, w'_R]$. Here, $-1 \leq q_i(T_i, Q_i) \leq 1$.

3. Quality-based likelihood ratio fusion

Let S_{gen} (S_{imp}) denote a genuine (impostor) match score with distribution function $L_{gen}(s)$ ($L_{imp}(s)$) and density $l_{gen}(s)$ ($l_{imp}(s)$). The Neyman-Pearson theorem [10] states that the *optimal* test for deciding between the genuine and impostor class is the likelihood ratio (LR) test. Dass et al. [5] model the genuine and impostor score densities as generalized densities (mixture of discrete and continuous components). Given a generic set of match scores, the generalized density can be estimated as follows. For a fixed threshold T, the discrete match scores are identified as those scores s_0 with $P(S = s_0) > T$, where $0 \leq T \leq 1$. In our experiments, the value of T is set to 0.02. The probability $P(S = s_0)$ is estimated as $N(s_0)/N$, where $N(s_0)$ is the number of observations in the given score set that equals s_0 and N is the total number of observations. After removing the discrete components, the continuous component of the density is estimated using a non-parametric kernel density estimation technique with a Gaussian kernel. In a multibiometric system with R matchers, let $l_{j,gen}(s_j)$ (($l_{j,imp}(s_j)$) be the generalized density estimated from the genuine (impostor) scores of the j^{th} matcher, for $j = 1, \ldots, R$. If the R biometric matchers are independent, the joint density of the R match scores, $\boldsymbol{s} = [s_1, \ldots, s_R]$ is the product of the R marginal densities. Hence, the combined likelihood ratio (also known as the product fusion score), $PFS(\boldsymbol{s})$, is

$$PFS(\boldsymbol{s}) = \prod_{j=1}^{R} \frac{l_{j,gen}(s_j)}{l_{j,imp}(s_j)}. \qquad (2)$$

An observed set of match scores $\boldsymbol{s}$ is assigned to the genuine class if $PFS(\boldsymbol{s})$ is greater than a fixed threshold η, $\eta > 0$.

The proposed quality-based fusion technique is based on the following observation. When the samples presented to a matcher are of poor quality, it cannot reliably distinguish between genuine and impostor users and the likelihood ratio will be closer to 1. On the other hand, for good quality samples, the likelihood ratio will be greater than 1 for genuine users and less than 1 for impostors. Hence, if we estimate the joint density of the match score and the quality of the match for each matcher, the resulting likelihood ratios of the individual matchers will be implicity weighted. Since the match score and the quality metric are correlated, we use the copula-based multivariate density estimation technique [5] for the joint density estimation. Let q_j be the quality of the match provided by the j^{th} matcher, for $j = 1, \ldots, R$. Let $l_{j,gen}(s_j, q_j)$ (($l_{j,imp}(s_j, q_j)$) be the joint density of the match score and the quality estimated from the genuine (impostor) template-query pairs of the j^{th} matcher. The quality-based product fusion score, $QPFS(\boldsymbol{s})$, is given by

$$QPFS(\boldsymbol{s}) = \prod_{j=1}^{R} \frac{l_{j,gen}(s_j, q_j)}{l_{j,imp}(s_j, q_j)}. \qquad (3)$$

4. Experimental results

We have used the multimodal database collected at West Virginia University consisting of fingerprint and iris modalities for 320 users with 5 samples per modality. The fingerprint match scores (s_f) are obtained using the 2-D dynamic programming-based minutiae matcher [9]. The iris match score (s_i) is the Hamming distance between the template and query Iriscodes [6]. The quality of the fingerprint (q_f) and iris (q_i) matches are estimated using the procedure described in Section 2. 50% of the genuine matches (match scores and the corresponding quality measures) and 25% of the impostor matches are used for density estimation, while the remaining data is used for evaluating the performance. This training-test partitioning is repeated 20 times and we report the average performance over the 20 trials.

Figure 2 shows the performance of the product and the quality-based product fusion rules. Fusion of fingerprint and iris using the product rule gives a large improvement in the GAR compared to the best single modality (here, iris). The quality-based product fusion rule further improves the GAR. For example, at a FAR of 0.01%, the GAR of the iris modality is 75.2%, while the GAR of the product and quality-based product fusion rules are 89.5% and 94.8%, respectively. From figure 2, we also observe that the performance of the quality-based product fusion rule is comparable to the quality-weighted sum of scores fusion method[1] proposed in [11]. However, it must be emphasized that the weighted sum rule requires normalization of match scores and tuning of matcher and quality weights to achieve higher accuracy. On the other hand, the proposed likelihood ratio-based fusion scheme is a principled approach that does not involve any ad-hoc normalization and weighting.

Figure 3 shows the biometric samples of a user whose iris images are of good quality ($q_i = 0.7$), but fingerprint images are of poor quality ($q_f = 0.4$). The product fusion score (PFS) is low ($\log(PFS) = 1.4$), resulting in a false reject. However, the quality-based fusion rule implicitly assigns a higher weight to the modality with better quality (iris). Hence, the QPFS is high ($\log(QPFS) = 10.6$) and the user is accepted by the multibiometric system.

5. Summary

We have proposed a likelihood ratio-based approach to achieve quality-dependent match score fusion. The proposed method does not use any ad-hoc weighting scheme to

[1] In quality-weighted sum of scores fusion method, the fused score is obtained as $QSS(\boldsymbol{s}) = \sum_{j=1}^{R} \alpha_j s_j$, where $\alpha_j = w_j + \beta_j q_j$. Here, w_j and β_j are known as matcher and quality weights, respectively. In our experiments, weights of 0.4 and 0.6 are assigned to the iris and fingerprint matchers, respectively. The quality weight β_j is set to 0.1.

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

Figure 3. Illustration of quality-based fusion. This user's iris images ((a) template and (b) query) are of good quality and the fingerprints ((c) template and (d) query) are of poor quality. Quality-based fusion rule implicitly assigns a higher weight to the iris modality, resulting in a correct acceptance of a genuine user who was falsely rejected by the simple product fusion rule.

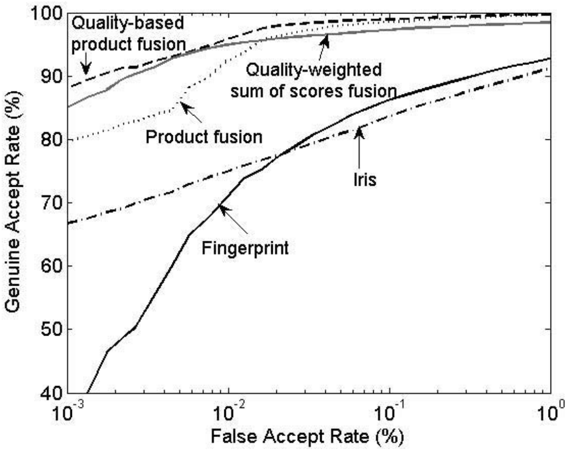

Figure 2. ROC curves for product fusion and quality-based product fusion rules.

combine the match scores. Instead, by estimating the joint densities of the match score and the quality of the genuine and impostor classes, the likelihood ratios of the individual matchers are implicity weighted. We have also modified existing quality assessment algorithms for fingerprint and iris images to determine the joint quality of a template-query pair. The proposed quality-based fusion scheme provides significant improvement in the performance of a multimodal (iris + fingerprint) biometric system.

References

[1] J. P. Baker and D. E. Maurer. Fusion of Biometric Data with Quality Estimates via a Bayesian Belief Network. In *The Biometric Consortium Conference (BC2005)*, pages 21–22, Arlington, USA, September 2005.

[2] E. S. Bigun, J. Bigun, B. Duc, and S. Fischer. Expert Conciliation for Multimodal Person Authentication Systems using Bayesian Statistics. In *First AVBPA*, pages 291–300, Crans-Montana, Switzerland, March 1997.

[3] Y. Chen, S. C. Dass, and A. K. Jain. Fingerprint Quality Indices for Predicting Authentication Performance. In *Fifth AVBPA*, pages 160–170, Rye Brook, July 2005.

[4] Y. Chen, S. C. Dass, and A. K. Jain. Localized Iris Image Quality Using 2-D Wavelets. In *Intl. Conf. on Biometrics (ICB)*, Hong Kong, January 2006.

[5] S. C. Dass, K. Nandakumar, and A. K. Jain. A Principled Approach to Score Level Fusion in Multimodal Biometric Systems. In *Fifth AVBPA*, pages 1049–1058, Rye Brook, July 2005.

[6] J. Daugman. How Iris Recognition Works? *IEEE Trans. on CSVT*, 14(1):21–30, 2004.

[7] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia, and A. K. Jain. Incorporating Image Quality in Multi-algorithm Fingerprint Verification. In *Intl. Conf. on Biometrics (ICB)*, Hong Kong, January 2006.

[8] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun. Discriminative Multimodal Biometric Authentication based on Quality Measures. *Pattern Recognition*, 38(5):777–779, May 2005.

[9] A. K. Jain, S. Prabhakar, and S. Chen. Combining Multiple Matchers for a High Security Fingerprint Verification System. *Pattern Recognition Letters*, 20(11-13):1371–1379, November 1999.

[10] E. L. Lehmann and J. P. Romano. *Testing Statistical Hypotheses*. Springer, 2005.

[11] N. Poh and S. Bengio. Improving Fusion with Margin-Derived Confidence in Biometric Authentication Tasks. In *Fifth AVBPA*, pages 474–483, Rye Brook, July 2005.

[12] K.-A. Toh, W.-Y. Yau, E. Lim, L. Chen, and C.-H. Ng. Fusion of Auxiliary Information for Multi-modal Biometrics Authentication. In *First Intl. Conf. on Biometric Authentication (ICBA)*, pages 678–685, Hong Kong, July 2004.

[13] C. Wilson, A. R. Hicklin, M. Bone, H. Korves, P. Grother, B. Ulery, R. Micheals, M. Zoepfl, S. Otto, and C. Watson. Fingerprint Vendor Technology Evaluation 2003: Summary of Results. Technical Report 7123, NIST, June 2004.