

Score Fusion in Multibiometric Identification Based on Fuzzy Set Theory

Khalid Fakhar¹, Mohammed El Aroussi^{1,2},
Mohamed Nabil Saidi^{1,3}, and Driss Aboutajdine¹

¹ GSCM-LRIT Research Laboratory (associated to CNRST, URAC 29),
Mohammed V University – Agdal, Rabat, Morocco

kld.fakhar@gmail.com, aboutaj@fsr.ac.ma

² LETI, EHTP, Casablanca, Morocco

moha387@yahoo.fr

³ INSEA, Rabat, Morocco

msaidi@insea.ac.ma

Abstract. Multimodal biometric systems consolidate or fuse information from multiple biometric sources. They have been developed to overcome several limitations of each individual biometric system, such as sensitivity to noise, intra class invariability, data quality, non-universality and other factors. In this paper, we propose a general framework of multibiometric identification system based on fusion at matching score level using fuzzy set theory. The motivation for using fuzzy set theory is that it offers methods suited to treat (modeling, fusion,...) and take into account the information inherently uncertain and ambiguous. We note that our fusion system is based on face and iris modalities. Experimental results exhibit that the proposed method performance bring obvious improvement compared to unimodal biometric identification methods and classical combination approaches at score level fusion.

Keywords: Multimodal biometrics, identification, score level fusion, fuzzy set theory.

1 Introduction

Biometrics is the science of establishing human identity based on their physical or behavioral characteristics [1]. Biometric systems can operate in two modes [2]: verification and identification. In the verification mode, the system validates a query biometric by comparing only the captured biometric data with the biometric template of a specific identity stored in the database. There is one-to-one comparison in this case. In the identification mode, the user's biometric input is compared with the templates of all the persons enrolled in the database. Therefore, the system conducts a one-to-many comparisons to establish individual's identity (closed-set) or fails if the input is not enrolled in the database (open-set). The identification task is significantly more challenging than verification [3].

Biometric systems based on a single biometric trait suffer from limitations such as sensitivity to noise, data quality, non-universality and spoof attacks [4]. Multimodal biometric systems which combine multiple sources of biometric information have been developed in order to overcome those problems and to achieve better recognition performance [5]. The technique implementation of this kind of systems requires the use of information fusion theory. There are three main fusion strategies [6]: fusion at the feature extraction level, matching score level and decision level.

In our work, we focused on biometric fusion at score level because it offers the best trade-off in terms of the information content and the ease in fusion [5]. Unlike most studies that have been used in this fusion level and examined for the verification mode [7], we aim, in this paper, to introduce a general framework for the closed-set identification based on fuzzy set theory [8] where each biometric matcher output will be modeled as a fuzzy set.

The rest of this paper is organized as follow. In section 2, we introduce the proposed system, then we describe the multimodal biometric system used in our work in section 3. The experimental results are discussed in section 4. Finally some conclusions end this paper.

2 Fuzzy Set Theory and Multibiometric Identification System

2.1 Fuzzy Set Theory

Fuzzy set theory was introduced by Zadeh in 1965 as a means of representing and manipulating imprecise or uncertain information [8]. It plays a significant role in many complex systems because of their capability to model the vagueness and ambiguity data. A fuzzy set F is a subset of the universe of discourse U , represented as:

$$F = \{ (x, \mu_F(x)), x \in U \},$$

where $\mu_F(\cdot)$ is a membership function which gives to each $x \in U$ a degree of belongingness from $[0, 1]$. When $\mu_F(\cdot)$ takes a value only in $\{0, 1\}$, F reduces to a crisp set and $\mu_F(\cdot)$ represents the characteristic function of the set F .

Measure of fuzziness is an intrinsic property that estimates the average ambiguity in a fuzzy set. In the literature, several definitions have been proposed. In general, a measure of fuzziness is a function

$$h : \mathfrak{F}(U) \longrightarrow \mathbb{R}^+,$$

where $\mathfrak{F}(U)$ denotes the set of all fuzzy subsets of U . This function is required to satisfy the following conditions [9]:

1. $\forall F \subset U$, if $h(\mu_F) = 0$ then F is a crisp set in U .
2. $h(\mu_F)$ assumes a unique maximum if $\forall x \subset U, \mu_F(x) = 0.5$.

3. $\forall (\mu_F, \mu_{F'}) \in U^2, f(\mu_F) \geq f(\mu_{F'})$ if F' is "crisper" than F ,
 i.e., $\forall x \in U \begin{cases} \mu_G(x) \geq \mu_F(x) & \text{if } \mu_F(x) \geq 0.5 \\ \mu_G(x) \leq \mu_F(x) & \text{if } \mu_F(x) \leq 0.5 \end{cases}$
4. $\forall F \subset U, h(\mu_F) = h(\mu_{\bar{F}})$ where \bar{F} is the complement of F .

For each fuzzy set F , this function assigns a nonnegative real number $h(F)$ that expresses the degree to which the boundary of F is not sharp. De Luca and Termini [9] borrowed the concept of Shannon's [10] information theoretic entropy to define a fuzzy entropy satisfying the above properties. This fuzzy entropy is given by:

$$H_{DTE}(\mu_F) = K \sum_{i=1}^n S(\mu_F(x_i)), \tag{1}$$

where $S(x) = -x \log(x) - (1 - x) \log(1 - x)$ and K is a normalizing constant. Many other measures of fuzziness have been suggested, with similar properties, but fuzzy entropy is still one of the most interesting and important measures of fuzziness in a fuzzy set.

2.2 Score Fusion in Multibiometric Identification System

When a biometric system operates in the closed-set identification, the output of the system is always a non-empty candidate list. Therefore, our goal is to determine the true identity of the given query.

Let M denote the number of biometric modalities in the multibiometric system and N be the number of users enrolled in the system. Each of these users has his own reference model. To simplify the notation, we assume that there is only a single reference model associated with each user for each biometric modality. A way to work at the score level is the use a score matrix which containing the match scores output by each biometric matchers. Let s_m^n denote the generic match score output by the m^{th} biometric matcher for n^{th} user in the database, $m = 1, 2, \dots, M; n = 1, 2, \dots, N$. For a given query, we can get a $M \times N$ score matrix \mathbf{S} defined as:

$$\mathbf{S} = \begin{pmatrix} s_1^1 & \cdots & s_1^N \\ \vdots & \ddots & \vdots \\ s_M^1 & \cdots & s_M^N \end{pmatrix} = \begin{pmatrix} S_1 \\ \vdots \\ S_M \end{pmatrix},$$

where $S_m = \{s_m^1, s_m^2, \dots, s_m^N\}$, for $m = 1, 2, \dots, M$.

At the score level fusion, a normalization step is generally necessary before the match scores from different matchers can be combined, because the matching scores output by various modalities are heterogeneous. In general, the normalized score is obtained by using a normalization function that maps S_m to \bar{S}_m defined as:

$$\begin{aligned} \mu : S_m &\longrightarrow \bar{S}_m \subset [0, 1] \\ s_m^n &\longmapsto \mu(s_m^n) \end{aligned},$$

where S_m is the set of all the raw output values of the corresponding matcher m and \bar{S}_m is the set of normalized matching scores of the matcher m . Therefore, the normalized score matrix $\bar{\mathbf{S}}$ is defined as:

$$\bar{\mathbf{S}} = (\mu(s_m^n))_{M \times N} = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_M)^\top.$$

If we replace the normalization function by the membership function using the fuzzy set theory, we can consider the set $\bar{S}_m = \{\mu(s_m^1), \mu(s_m^2), \dots, \mu(s_m^N)\}$ as a fuzzy set in S_m , and we can write it in the following way:

$$\bar{S}_m = \{(s_m^n, \mu_{S_m}(s_m^n)), s_m^n \in S_m\},$$

where $\mu_{S_m}(s_m^n) \in [0, 1]$ is the grade of the membership function of s_m^n in \bar{S}_m . Therefore, for a given query, the output of each biometric matcher is then modeled as a fuzzy set. In other words, the outcome of the system can be viewed as the fusion of the different membership functions provided by the different biometric matchers. In this study each biometric matcher m is represented by piecewise-linear membership function defined as:

$$\mu_{S_m}(s_m^n) = \begin{cases} 1 & \text{if } s_m^n < \alpha_1, \\ (\alpha_2 - s_m^n)/(\alpha_2 - \alpha_1) & \text{if } \alpha_1 < s_m^n < \alpha_2, \\ 0 & \text{if } s_m^n > \alpha_2, \end{cases}$$

where α_1 and α_2 are the minimum value of the impostor and the maximum value of the genuine score distributions, respectively.

2.3 Proposed Fusion Approach

All matching scores provided by a biometric matcher is represented as a fuzzy set. We assume that when a biometric matcher provides a reliable result, this set will be close to a binary set. On the contrary, when the biometric matcher is unreliable no opportunity should be significantly higher than the others. In term of fuzziness degree, the set constructed with a reliable biometric matcher will have a low degree contrary to a set constructed with an unreliable biometric matcher. Based on measure of fuzziness we define the weights

$$w_i = \frac{\sum_{k=0, k \neq i}^m H_{DTE}(\mu_{S_k})}{(m - 1) \sum_{k=0}^m H_{DTE}(\mu_{S_k})} \tag{2}$$

where $H_{DTE}(\mu_{S_k})$ is the fuzziness degree of biometric matcher k defined in (1), and m is the number of biometric modalities. When a source has a low fuzziness degree, w_i is close to 1 and it only slightly affects corresponding fuzzy set.

For every person (enrolled in the database) submits a biometric query samples, m fuzzy sets are computed, one for each modality. Then each fuzzy set will be

weighted by w_i defined in (2). This set of fuzzy sets constitutes the input for the fusion process:

$$\mathbf{D} = \{w_1\bar{S}_1, w_2\bar{S}_2, \dots, w_M\bar{S}_M\}.$$

Based on fuzzy set theory, the query should be assigned to the identity I_{n_0} if

$$n_0 = \arg \max_{n=1}^N \bigcup_{m=1}^M \{w_m \mu_{S_m}\},$$

where $\bigcup_{m=1}^M \{w_m \mu_{S_m}\} = \max_{m=1}^M \{w_m \mu_{S_m}\}.$

3 Multimodal Biometric System

In this paper, the face and iris biometric traits are selected to construct multimodal biometric system, because face recognition is most natural and acceptable in identity authentication whereas iris recognition is one of the most accurate biometrics.

3.1 Face Recognition

Among various face recognition algorithms, the most popular are appearance based approaches. Here we use the Steerable Pyramid (S-P) wavelet transform [11]. The first task in face recognition is an initial alignment, all images are aligned with respect to the manually detected eye coordinates, scaled and histogram equalized. After this stage, each sample face image preprocessed is described by a subset of band filtered images containing steerable pyramid coefficients which characterize the face textures. We divide the S-P sub-bands into small sub-blocks, from which we extract compact and meaningful feature vectors using simple statistical measures. For recognition, the city-block distance is used to measure the dissimilarity of different feature vectors.

3.2 Iris Recognition

Iris recognition is receiving increased attention due to its high reliability [12]. The iris recognition system employed in this paper is mainly based on the open source code provided by Libor Masek [13]. The human iris is an annular region between the black pupil and the white sclera. The first stage of iris recognition is to isolate the boundaries between the pupil and iris region and between sclera and iris region using a method that based on canny edge detection and circular Hough Transform [14]. The iris region not only contains the region of interest, but also some unuseful parts such as eyelid, eyelash and pupil. The eyelashes and eyelids occlude the upper and lower parts of the iris region are marked so that they will not be taken into account in the comparison stage. A noise mask

was generated to record the position of the occluded region. After localizing the region of interest, the Rubber Sheet Model [12] is used for unwrapped the iris to a rectangular region with prefixed size where the iris texture is analyzed. Then, a template representing iris pattern information is created using a 1D log-Gabor wavelets. The created templates were compared using the Hamming distance [12].

4 Experimental Results

To evaluate the robustness and the performance of our approach, we used a multimodal database containing face and iris samples. Considering the independence between two biometric traits, our database is constructed by concatenating FERET face database [15] and CASIA iris database (version 1) [16]. Finally, The obtained virtual multimodal database contains 756 records corresponding to 108 subjects (7 records for each subject), and each record represents a face sample and an iris sample.

In the following experiments, the whole database is divided randomly into two data sets, 128 records to construct the training data and the remaining records are used for testing. The training dataset was used to get the distributions of the unimodal matching scores in order to estimate the parameters of the fuzzy membership function. For the testing dataset, we use it to evaluate the performance of multimodal system. In the closed-set identification system the cumulative match characteristic (CMC) is widely used as an accuracy performance measure. The CMC shows the ratio of correct identification versus the matching rank.

The goal of the multimodal biometric is to increase precision and enhance reliability than unimodal biometrics. In order to demonstrate the effectiveness of our proposed method, figure 1 shows the CMC curves of the individual face and iris matchers and for the proposed method. These results clearly show that our multimodal system provides a significantly better rate than the individual monomodal system.

The second experiment consist to evaluate our fusion system with different score fusion methods such as sum, product, min and max rules. The recognition accuracy with min-max normalization are presented in table 1. According to table 1, we note that the proposed method outperforms the sum rule, which is considered to be one of the best combination approaches at score level fusion [5].

Table 1. Recognition accuracy: proposed method vs. combination approaches

Sum rule	Product rule	Max rule	Min rule	Proposed method
96.76%	96.30%	92.13%	91.67%	97.69%

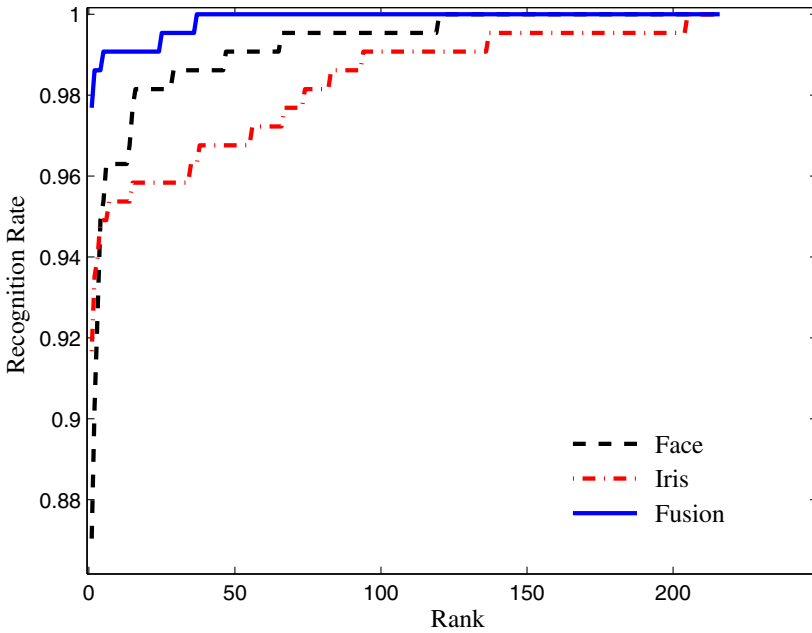


Fig. 1. Cumulative Match Characteristic (CMC) curves: proposed method vs. unimodal systems

5 Conclusion

In this paper, we proposed a general framework of multibiometric identification system based on fusion at matching score level using fuzzy set theory. In this approach, the output of each biometric matcher is modeled as a fuzzy set, and the corresponding degree of fuzziness estimates the reliability of the information provided by each biometric matcher. Then, the results are aggregated with a fuzzy combination rule. In order to prove the efficiency of our fusion approach, we have used a virtual multimodal database which integrates face and iris biometric modalities. Experimental results exhibit that our fusion approach achieves better performance than the best unimodal system and the classical combination approaches at score level fusion. However, the system needs to be tested on a large database.

References

1. Jain, A.K., Flynn, P., Ross, A.A.: Hand book of Biometrics. Springer (2008)
2. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology 14(1), 4–20 (2004)

3. Nandakumar, K., Jain, A.K., Ross, A.: Fusion in Multibiometric Identification Systems: What about the Missing Data? In: Tistarelli, M., Nixon, M.S. (eds.) ICB 2009. LNCS, vol. 5558, pp. 743–752. Springer, Heidelberg (2009)
4. Jain, A., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. *Pattern Recognition* 38(12), 2270–2285 (2005)
5. Ross, A.A., Nandakumar, K., Jain, A.K.: *Handbook of Multibiometrics* (International Series on Biometrics). Springer-Verlag New York, Inc., Secaucus (2006)
6. Ross, A., Jain, A.: Information fusion in biometrics. *Pattern Recognition Letters* 24, 2115–2125 (2003)
7. Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.K.: Likelihood Ratio Based Biometric Score Fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(2), 342–347 (2008)
8. Zadeh, L.A.: Fuzzy sets. *Information and Control* 8(3), 338–353 (1965)
9. De Luca, A., Termini, S.: A definition of a nonprobabilistic entropy in the setting of fuzzy entropy. *Inform. Control* 20, 301–312 (1972)
10. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. Jr.* 379–423, 621–659 (1946)
11. El Aroussi, M., El Hassouni, M., Ghouzali, S., Rziza, M., Aboutajdine, D.: Local appearance based face recognition method using block based steerable pyramid transform. *Signal Processing* (2010)
12. Daugman, J.G.: How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology* 14(1), 21–30 (2002)
13. Masek, L.: Recognition of human iris patterns for biometric identification. Bachelor of Engineering Degree Thesis, The University of Western Australia, Australia (2003)
14. Wildes, R.P.: Iris recognition: an emerging biometrics technology. *Proc. IEEE* 85, 1348–1363 (1997)
15. Phillips, P.J., nad Moon, H., Rauss, P.J., Rizvi, S.: The FERET evaluation methodology for face recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence* 22, 891–906 (2000)
16. Institute of Automation, Chinese academy of Science, CASIA Iris Image Database, <http://www.sinobiometrics.com/chinese/chinese.htm> (retrieved on April 2010)