Biometric Recognition: When Is Evidence Fusion Advantageous?

Hugo Proença

Department of Computer Science
IT - Instituto de Telecomunicações
SOCIA - Soft Computing and Image Analysis Group
University of Beira Interior, 6200-Covilhã, Portugal
hugomcp@di.ubi.pt

Abstract. Having assessed the performance gains due to evidence fusion, previous works reported contradictory conclusions. For some, a consistent improvement is achieved, while others state that the fusion of a stronger and a weaker biometric expert tends to produce worst results than if the best expert was used individually. The main contribution of this paper is to assess when improvements in performance are actually achieved, regarding the individual performance of each expert. Starting from readily satisfied assumptions about the score distributions generated by a biometric system, we predict the performance of each of the individual experts and of the fused system. Then, we conclude about the performance gains in fusing evidence from multiple sources. Also, we parameterize an empirically obtained relationship between the individual performance of the fused experts that contributes to decide whether evidence fusion techniques are advantageous or not.

1 Introduction

Private and governmental entities are paying growing attention to biometrics and nationwide systems are starting to be deployed. Pattern recognition (PR) systems have never dealt with such sensitive information at these high scales, which motivated significant efforts to increase accuracy, comfort, scale and performance. Currently deployed systems achieve remarkable low error rates (e.g., the Daugman's iris recognition system [1]) at the expenses of constrained data acquisition setups and protocols, which is a major constraint regarding their dissemination.

Most biometric systems use a single trait for recognition (e.g., fingerprints, face, voice, iris, retina, ear or palm-print) and are called *unimodal*. These systems have high probability of being affected by noisy data, non-universality, lack of distinctiveness and spoof attacks [2]. *Multimodal* systems make use of more than one source to perform recognition and are an attempt to alleviate these problems. Used sources may be different recognition strategies from the same data or from different sensors and from a unique or multiple traits. Here, fusion can occur at any stage of the PR process: (1) At the data acquisition level; (2) At the match score level, if the scores generated by each feature comparison strategy are used; (3) At the decision level, if the output of each PR system is used to generate the final response. Fusion at the early stages is believed to

G. Bebis et al. (Eds.): ISVC 2009, Part II, LNCS 5876, pp. 698–708, 2009.

be more effective [3], essentially due to the amount of available information. However, it is more difficult to achieve in practice, due to usual incompatibilities between feature sets. At the other extreme, fusion at the decision level is considered too rigid, due to the limited amount of available information. Fusion at the match score level is seen as a trade-off: it is relatively easy to perform and combines appropriately the scores generated by different modalities.

The idea of fusing scores to perform biometric recognition is largely described in the literature. Ross and Jain [4] reported a significant improvement in performance when using the sum rule. Wang *et* al. [5] used the similarity scores of a face and an iris recognition module to generate 2D feature vectors that are redirected as inputs of a neural network classifier. Duca *et* al. [6] framed the problem according to the Bayes theory and estimated the biases of individual expert opinions. These were used to calibrate and fuse scores into a final decision. Brunelli and Falavigna [7] used the face and voice for identification and Hong and Jain [8] associated different confidence measures with the individual matchers, when integrating the face and fingerprint traits. These works reported significant improvements in performance due to evidence fusion and did not pointed any constraint about the individual performance of each fused expert. However, as stated by Daugman [9], fusing different scores may not be good for all situations. Although the combination of tests enables the decision based in more information, on the other hand, if a stronger test is combined with a weaker one, the resulting decision environment is in some sense averaged, and the combined performance will lie somewhere between that of the two tests conducted individually. Accordingly, Poh and Bengio [10] analyzed four typical scenarios encountered in biometric recognition, mainly concerned about the issues of score correlation and variance, having concluded that fusing in not always beneficial.

The main purpose of this study is to contribute to the decision about when evidence fusion should actually be used, by predicting the performance of the fused classifier and comparing it with the corresponding value of the best expert used in fusion. To do so, we simulate the scores generated by each expert, assuming that they are unimodal, independent and identically-distributed (i.i.d.) for the intra-class and inter-class comparisons and can be approximated by normal distributions. These assumptions may be readily satisfied and are a reasonable practice in multimodal biometrics research. Everywhere in this paper, the term “performance” refers to the accuracy performance of the biometric experts. We will restrict our study to fusion of two biometric experts, although the results could be extended to fusion of multiple biometrics by induction, as pointed by Hong *et* al. [11].

The remainder of this paper is organized as follows: Section 2 briefly summarizes the most usual information fusion techniques. Section 3 describes our empirical framework and presents and discusses the results. Finally, Section 4 concludes.

2 Evidence Fusion

Kittler *et* al. [17] developed a theoretical framework for combing multiple experts and derived the most usual classifiers combinations schemes, such as the $product$, sum, min, max and $median$ rules.

Let R be the number of biometric experts B_i operating in the environment, such that $i = \{1, \ldots, R\}$. Let Z be an input pattern that is to be assigned to one of m classes $w_1, \ldots, w_m$. For our purposes, the value of m was set to 2, which corresponds to a trivial verification system (w_0=intra-class comparison, w_1=inter-class comparison). Let $\overrightarrow{x_i}$ be the biometric signature (encoded from Z) that is presented to the i^{th} biometric expert and generates a corresponding dissimilarity score d_i for each enrolled template. Before fusing scores it is necessary to perform normalization, so that none of the base experts dominates the decision. Without any assumption about the priori probabilities, the approximation of the posterior probability that $\overrightarrow{x_i}$ belongs to class w_j is given by:

$$P(w_j|\overrightarrow{x_i}) = \frac{P(\overrightarrow{x_i}|w_j)}{\sum_{s=1}^{R} P(\overrightarrow{x_i}|w_s)} \tag{1}$$

where $P(\overrightarrow{x_i}|w_j)$ denotes the probability density function of the j^{th} class, estimated from the d_i values observed in a training set.
Product Rule: Assuming statistical independence of the $\overrightarrow{x_i}$ values, the input pattern is assigned to class w_0 iff:

$$\prod_{i=1}^{R} P(w_0|\overrightarrow{x_i}) > \prod_{i=1}^{R} P(w_1|\overrightarrow{x_i})$$

Sum Rule: Apart from the assumption of statistical independence of the $\overrightarrow{x_i}$ values, this rule also assumes that posteriori probabilities from each system and corresponding priori probabilities are similar. The input pattern is assigned to class w_0 iff:

$$\sum_{i=1}^{R} P(w_0|\overrightarrow{x_i}) > \sum_{i=1}^{R} P(w_1|\overrightarrow{x_i})$$

Max Rule: This rule approximates the sum of the posterior probabilities by the maximum value. As in the previous rules, statistical independence of the $\overrightarrow{x_i}$ values is assumed. The input pattern is assigned to class w_0 iff:

$$\max_i P(w_0|\overrightarrow{x_i}) > \max_i P(w_1|\overrightarrow{x_i})$$

Min Rule: Similarly to the previous rule, statistical independence of the $\overrightarrow{x_i}$ values is assumed. The input pattern is assigned to class w_0 iff:

$$\min_i P(w_0|\overrightarrow{x_i}) > \min_i P(w_1|\overrightarrow{x_i})$$

Expert Weighting: As proposed by Snelick *et* al. [18], an intuitive idea is to assign different weights to each individual expert, hoping to increase the role played by the stronger ones. In this work, the weight e_i associated to each expert was assigned according to its d-prime value d'_i, proportionally with the values of the remaining experts. The input pattern is assigned to class w_0 iff:

$$\sum_{i=1}^{R} e_i \; P(w_0|\overrightarrow{x_i}) > \sum_{i=1}^{R} e_i \; P(w_1|\overrightarrow{x_i}), \; s.t. \; e_i = \frac{d'_i}{\sum_{j=1}^{R} d'_j}$$

Dempster-Shafer Theory: It is based on belief functions [19] and combines different pieces of evidences into a single value that approximates the probability of an event. Let X denote our frame of discernment, composed uniquely by two states: the assignment of the input pattern to classes w_0 or w_1. The power set $\mathbb{P}(X)$ contains all possible subsets of X: $\{\emptyset, \{w_0\}, \{w_1\}, \{w_0, w_1\}\}$. Assigning null beliefs to the $\emptyset$ and $\{w_0, w_1\}$ states, the mass of w_0 is given by $m(w_0) = \max\{0, 1 - F_0 - F_1\}$, where F_0 and F_1 are the cumulative distribution functions of classes w_0 and w_1. $m(w_1) = 1 - m(w_0)$, so that $\sum_{A \in \mathbb{P}(X)} m(A) = 1$. The combination of two masses is given by:

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B) \; m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) \; m_2(C)} \qquad (2)$$

where m_1 and m_2 are the masses of individual experts.

3 Experiments and Discussion

Our empirical framework comprises v virtual subjects. Let these be denoted by $P = \{p_1, \ldots, p_v\}$. For simplicity purposes, we assume that (1) all subjects appear with identical frequency; (2) no other subjects attempt to be recognized and (3) data is properly acquired by all the biometric devices operating in the environment. Operating in the identification mode, each of the R samples acquired in a recognition attempt is matched against all the enrolled templates, performing a total of 1 intra-class and $v-1$ inter-class comparisons for each expert. Thus, a recognition attempts give a total of a (intra-class) and $(v-1) \times a$ (inter-class) dissimilarity scores for each expert. We denote these sets respectively by $X = \{X_1, \ldots, X_a\}$ and $Y = \{Y_1, \ldots, Y_{(v-1) \times a}\}$. We consider that X and Y are drawn from populations with distributions F_I and F_E, such that, $F_I(x) = Prob(X \leq x)$ and $F_E(x) = Prob(Y \leq x)$. Also, multiple comparisons provide unimodal i.i.d. dissimilarity scores that follow the normal distribution. An estimate $\hat{F_I}(x)$ of $F_I(x)$ at some $x > 0$ is given by:

$$\hat{F_I}(x) = \frac{1}{a} \sum_{i=i}^{a} \mathbb{I}_{\{X_i \leq x\}} \qquad (3)$$

where $\mathbb{I}_{\{.\}}$ denotes the characteristic function. As suggested by Bolle *et* al. [20], the law of large numbers guarantees that $\hat{F_I}(x)$ is distributed according to a normal distribution $N(\hat{F_I}(x), \sigma(x))$. An estimate of the standard deviation is given by $\hat{\sigma}(x) = \sqrt{\frac{\hat{F_I}(x)(1-\hat{F_I}(x))}{s_i}}$ and confidence intervals can be found with percentiles of the normal distribution. For all our results, 99% confidence intervals were chosen and given by $-2.326\,\hat{\sigma}(x) < \hat{F}_I(x) < 2.326\,\hat{\sigma}(x)$. The procedure is similar for $\hat{F_E}(x)$.

In our experiments, we used $v = 10\ 000$ and $a = 20\ 000$. The dissimilarity scores generated by each biometric expert were simulated through a pseudo-random generator of normally distributed numbers (Zigurat method [21]), according to the corresponding parameters of the expert and type of comparison (intra-class and inter-class). Using 32-bit integers to store data, this method guarantees a period for the overall generator of about 2^{64}, which is more than enough for the purposes of this work. Also, we simulated different levels of correlation between the scores generated by experts and analyzed the corresponding effect in fusion. The Pearson product moment correlation $\rho(X, Y)$ measures the linear dependence between variables, yielding a value between 1 (maximal correlation) and 0 (independence):

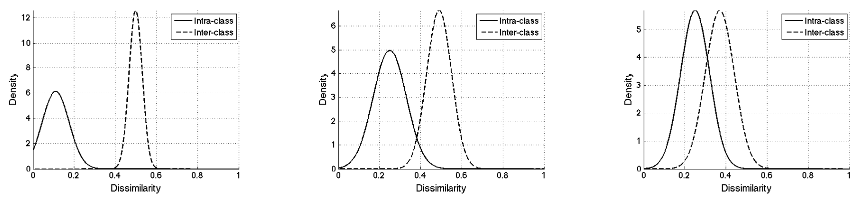
$$\rho(X, Y) = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{X_i - \mu_X}{\sigma_X}\right)\left(\frac{Y_i - \mu_Y}{\sigma_Y}\right) \tag{4}$$

As suggested by Daugman [12], the d-prime value (d') appropriately quantifies the decidability of a biometric system, informing about the typical separation between dissimilarity scores generated for intra-class and inter-class comparisons:

$$d' = \frac{|\mu_E - \mu_I|}{\sqrt{\frac{1}{2}(\sigma_I^2 + \sigma_E^2)}} \tag{5}$$

where μ_I and σ_I are the mean and standard deviation of the intra-class comparisons and μ_E and σ_E are similar values of the inter-class comparisons. As figure 1 illustrates, d' has inverse correspondence with the overlap area of the two distributions and, hence, acts as a measure of the error expected for the biometric system.

Figure 2 illustrates the typical distributions of the $P(w_0|\overrightarrow{x_i})$ values for each evidence fusion variant, having as base experts the ones illustrated in figures 1b and 1c.

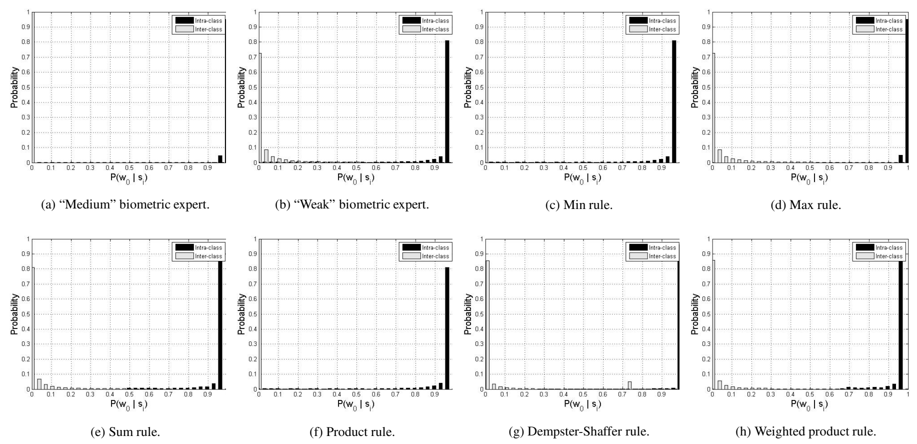


(a) "Strong" biometric expert ($\mu_I = 0.11$, $\sigma_I = 0.065$, $\mu_E = 0.49$, $\sigma_E = 0.031$, $d' = 7.60$).

(b) "Medium" biometric expert ($\mu_I = 0.25$, $\sigma_I = 0.08$, $\mu_E = 0.49$, $\sigma_E = 0.06$, $d' = 3.39$).

(c) "Weak" biometric expert ($\mu_I = 0.25$, $\sigma_I = 0.07$, $\mu_E = 0.37$, $\sigma_E = 0.07$, $d' = 1.75$).

Fig. 1. Typical intra-class (continuous lines) and inter-class distributions (dashed lines) of the dissimilarity scores generated by biometric experts with heterogenous performance: good (figure 1a, that refers to iris recognition in constrained imaging setups [1]), medium (figure 1b, based in results of ear [13], face [14] and palm-print [15] recognition) and poor (figure 1c, based in results of a gait classifier [16])



(a) "Medium" biometric expert. (b) "Weak" biometric expert. (c) Min rule. (d) Max rule.

(e) Sum rule. (f) Product rule. (g) Dempster-Shaffer rule. (h) Weighted product rule.

Fig. 2. Intra-class (dark bars) and inter-class (white bars) distributions of the posterior probabilities for class w_0 given a dissimilarity score s_i, of the biometric experts of figures 1b and 1c and of the different evidence fusion variants

3.1 Scenario 1: Fusing Independent Scores

In the first scenario we assume that the base experts analyze different traits and, thus, the scores generated by them are statistically independent. Figure 3 illustrates the independence between similarity scores generated by a "Strong" and a "Weak" biometric expert.

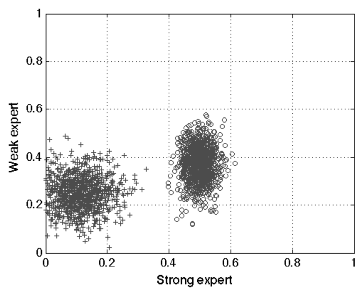


Fig. 3. Independence between scores ($\rho_I(X,Y) \approx 0,\ \rho_E(X,Y) \approx 0$) generated by a "Strong" and a "Weak" biometric expert. Cross points denote the intra-class and circular points the inter-class comparisons.

Table 1 compares the results obtained by evidence fusion, where S, M and W stand for the "Strong", "Medium" and "Weak" biometric experts illustrated in figure 1. Note that the "Individual" column gives the values obtained individually by the best biometric expert. d' and EER denote the decidability and approximated equal error rate. It is interesting to note that the weighted rule outperformed all the other combination rules for most of the times.

Table 1. Comparison between the performance obtained by evidence fusion techniques, according to the individual performance of independent experts. S, M and W denote the "Strong", "Medium" and "Weak" biometric experts of figure 1. $\{A, B\}$ denotes de fusion of the A and B experts.

	d'							**EER**						
Experts	Individual	Min	Max	Prod.	Sum	DS	Weig.	Individual	Min	Max	Prod.	Sum	DS	Weig.
$\{S,\ S\}$	360	8690	250	8591	499	18.9	499	0.01	0	0.01	0	0.01	0.04	0.01
$\{S,\ M\}$	360	6.98	8.69	6.98	10.89	7.23	18.12	0.01	0.04	0.03	0.04	0.03	0.04	0.03
$\{S,\ W\}$	360	4.41	4.13	4.41	6.03	2.75	20.11	0.01	0.05	0.04	0.05	0.04	0.05	0.04
$\{M,\ M\}$	5.26	4.97	6.20	4.98	7.44	7.73	7.44	4.15	4.60	3.37	4.47	3.61	3.5	3.61
$\{M,\ W\}$	5.29	3.41	3.48	3.47	4.46	3.98	5.31	4.15	6.80	6.74	6.61	5.74	6.07	4.12
$\{W,\ W\}$	1.88	2.22	2.18	2.26	2.60	1.87	2.60	18.82	14.21	11.19	14.03	14.01	16.12	14.81

In order to assess in more detail when evidence fusion techniques do indeed improve the recognition performance, we compared the results obtained by the evidence fusion variant that was observed to be the best (for each case) with the performance of the best individual expert. We carried out all combinations between experts with individual d' in $[1,\ 7.75]$, using 0.25 steps. Figure 4 exhibits the results. The vertical axis gives the proportion between the results of the best individual expert and the ones obtained by evidence fusion. The remaining axes give the individual d' values of the fused experts. It can be observed that the higher improvements are achieved when the base experts have close performance. If their performance is notoriously different, evidence fusion techniques lead to the deterioration of the performance, when compared to the best individual expert. Also, a diagonal structure can be seen in the plot, which suggests that improvements tend to be in direct correspondence with the individual performance of the base experts. The figure at the upper-right corner shows the intersection of the 3D performance surface with the plane $z = 1$, which is to say that reveals the region where evidence fusion was observed to be advantageous (Adv). Here, a power relationship between the stronger (d'_s) and the weaker (d'_w) expert appears to be evident, as the R^2 value of the fitted function (given in (6)) confirms.

3.2 Scenario 2: Fusing Correlated Scores

It was also found pertinent to evaluate the improvements in performance when the scores generated by the fused biometric experts are correlated, perhaps because they resulted from the same data or from different data extracted from the same trait. In our experiments, the Pearson product moment correlation ρ of the generated scores was approximately of 0.75, both for the intra-class $(\rho_I(X, Y))$ and inter-class $(\rho_E(X, Y))$ comparisons, as illustrated in figure 5.

As in the previous scenario, we varied the d' values of the fused experts and compared the results obtained by fusion and individually by the best expert. Table 2 lists the obtained results. It is evident that the gains break down, suggesting that data correlation fully constraints the improvements due to fusion. This was confirmed for all types of fused experts ("Strong", "Medium" or "Weak") and for all evidence fusion variants, which is in agreement with a previous conclusion made by Kitter *et* al. [17].

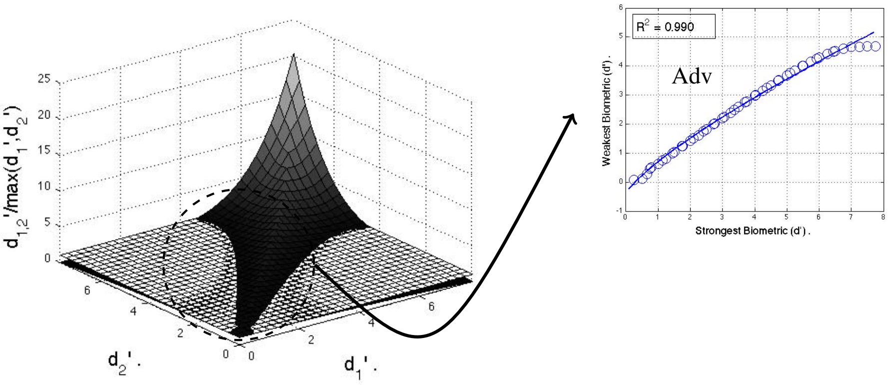


Fig. 4. Comparison between the performance obtained by evidence fusion and individually by the best expert, according to the d' values of the fused experts and assuming the independence between scores. A power function was fitted to the intersection of the 3D structure with the plane $Z = 1$, revealing the region (Adv) where evidence fusion is actually advantageous.

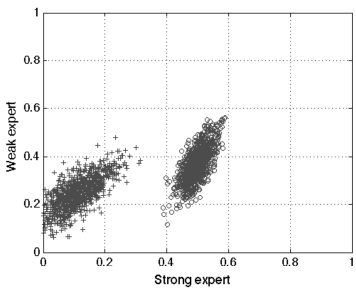


Fig. 5. Correlated scores ($\rho_I(X,Y),\ \rho_E(X,Y)\ \approx\ 0.75$) generated by a "Strong" and a "Weak" biometric expert. Cross points denote the intra-class and circular points the inter-class comparisons.

As in the previous scenario, figure 6 compares the results obtained by the best individual expert with the ones obtained by the evidence fusion variant that was observed to be the best. We confirmed that the higher improvements occur when the fused experts have close individual performance. Again, a power function defines the region where evidence fusion is advantageous (Adv), as illustrated in the 2D plot at the upper right-corner.

The above given results suggest that *evidence fusion improves the relative performance mostly when the fused experts have similar performance*. On the contrary, when one of the fused biometrics is considerably weaker, the overall performance of the system tends to decrease, confirming the essential of the Daugman's note: "*A strong biometric is better used alone than in combination with a weaker one*" [9]. Also, it should be noted that *data correlation is an important feature for the overall improvements*. This can be confirmed in (6), that defines the regions where evidence fusion should

Table 2. Improvements in performance obtained by evidence fusion techniques, according to the performance of correlated experts. S, M and W denote the "Strong", "Medium" and "Weak" biometric experts of figure 1. $\{A, B\}$ denotes de fusion of the A and B experts.

	d'							EER						
Classifs.	Individual	Min	Max	Prod.	Sum	DS	Weig.	Individual	Min	Max	Prod.	Sum	DS	Weig.
$\{S, S\}$	360	970	29	909	61	3.1	61	0.01	0.01	0.01	0	0.01	0.04	0.01
$\{S, M\}$	360	1.98	2.02	2.13	2.00	1.61	4.00	0.01	0.05	0.04	0.05	0.04	0.05	0.04
$\{S, W\}$	360	1.39	1.15	1.40	1.61	1.12	5.14	0.01	0.07	0.05	0.06	0.04	0.06	0.05
$\{M, M\}$	5.26	3.61	4.03	3.91	4.55	4.69	4.55	4.15	4.80	4.01	4.03	3.99	4.52	3.99
$\{M, W\}$	5.29	2.28	2.32	2.27	2.86	2.26	2.79	4.15	7.09	8.10	7.14	7.00	9.07	7.03
$\{W, W\}$	1.88	1.95	1.93	1.99	2.02	1.80	2.02	18.82	16.09	13.53	15.66	15.69	18.12	15.69

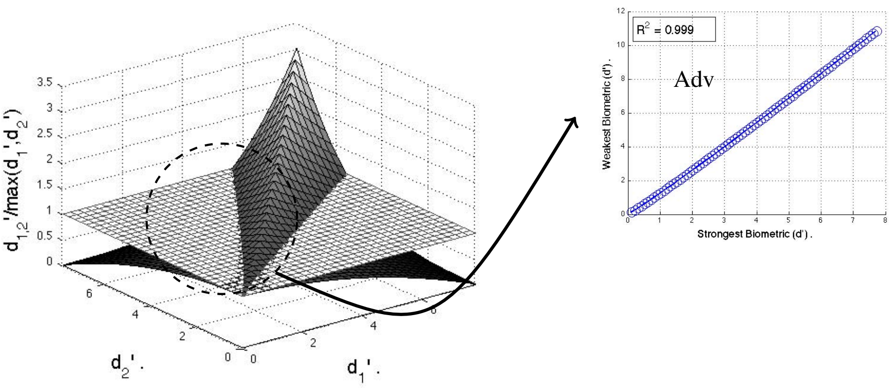


Fig. 6. Comparison between the performance obtained by the best evidence fusion variant and by the best individual expert used in fusion, according to the decidability (d') of the fused experts (left) and assuming scores correlation $(\rho(X_{i,j}, Y_{i,j}) \approx 0.75)$. A fitted power function approximates the relationship between the performance of the fused experts in order to improve performance by fusion (Adv region of the figure at the upper-right corner).

actually improve performance, either when the scores generated by experts are independent $(\rho(X_{i,j}, Y_{i,j}) \approx 0)$ or not. This equation relates the decidability value of the weaker biometric expert d'_w with the corresponding value of the stronger one (d'_s) in order to improve results by fusion.

$$\begin{cases} d'_w > 1.114\, d_s^{'0.7884} - 0.4133\;,\; \rho(X_{i,j}, Y_{i,j}) \approx 0 \\ d'_w > 1.254\, d_s^{'-1.056} + 0.002\;,\; \text{otherwise} \end{cases} \tag{6}$$

4 Conclusions

Previous research works reported substantial improvements in biometrics performance by fusing the evidence from multiple sources, with emphasis to the fusion at the score level. However, others authors claim that these strategies are not particularly useful and even tend to deteriorate the recognition performance.

Starting from readily satisfied assumptions about the dissimilarity scores generated by each biometric expert (i.i.d. unimodal values for the intra-class and inter-class comparisons that can be modeled by normal distributions), we simulated the outputs generated by different biometric experts and analyzed the performance gains obtained by the most usual evidence fusion techniques. We concluded that effectiveness is maximized when the fused biometrics have similar performance. Oppositely, if their performance is notoriously different, the overall performance tends to decrease, when compared to the best expert.

Also, we confirmed that the independence between the fused similarity scores is an important requirement for the effectiveness of score fusion techniques. If the fused data is strongly correlated, the performance achieved by evidence fusion is likely to be worst than the one obtained individually by the strongest expert. Finally, we fitted two boundary decision curves by power functions that define the regions where evidence fusion should actually be advantageous, either if the fused scores are independent or not. This was made according to the individual performance of each fused expert.

Acknowledgments

We acknowledge the financial support given by "FCT-Fundação para a Ciência e Tecnologia" and "FEDER" in the scope of the PTDC/EIA/69106/2006 research project "BIOREC: Non-Cooperative Biometric Recognition".

References

1. Daugman, J.G.: Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. Proceedings of the IEEE 94(11), 1927–1935 (2006)
2. Jain, A.K., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognition 38, 2270–2285 (2005)
3. Ross, A., Jain, A.K.: Multimodal biometrics: An overview. In: Proceedings of the 12^{th} European Signal Processing Conference (EUSIPCO), pp. 1221–1224 (2004)
4. Ross, A., Jain, A.K.: Information fusion in biometrics. Pattern Recognition Letters 24, 2115–2125 (2003)
5. Wang, Y., Tan, T., Jain, A.K.: Combining face and iris biometrics for identity verification. In: Kittler, J., Nixon, M.S. (eds.) AVBPA 2003. LNCS, vol. 2688, pp. 805–813. Springer, Heidelberg (2003)
6. Duca, B., Bigiin, E.S., Bigiin, J., Maitre, G., Fischer, G.: Fusion of audio and video information for multi modal person authentication. Pattern Recognition Letters 18(9), 835–843 (1997)
7. Brunelli, R., Falavigna, D.: Person identification using multiple cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 955–966 (1995)
8. Hong, L., Jain, A.K.: Integrating faces and fingerprints for personal identification. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, 1295–1307 (1998)
9. Daugman, J.G.: Biometric decision landscapes (2000), `http://www.cl.cam.ac.uk/TechReports/`

10. Poh, N., Bengio, S.: How Do Correlation and Variance of Base-Experts Affect Fusion in Biometric Authentication Tasks? IEEE Transactions on Signal Processing 53(11), 4384–4396 (2005)
11. Hong, L., Jain, A.K., Pankanti, S.: Can multibiometrics improve performance? In: Proceedings of the IEEE Workshop on Automatic Identification Advanced Technologies, pp. 59–64 (1999)
12. Daugman, J.G., Williams, G.O.: A proposed standard for biometric decidability. In: Proceedings of the CardTech/SecureTech Conference, pp. 223–234 (1996)
13. Arbab-Zavar, B., Nixon, M., Hurley, D.: On Model-Based Analysis of Ear Biometrics. In: IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, pp. 1–5 (2007)
14. Chen, X., Flynn, P., Bowyer, K.: IR and visible light face recognition. Computer Vision and Image Understanding 99(3), 332–358 (2005)
15. Han, Y., Sun, Z., Tan, T.: Combine hierarchical appearance statistics for accurate palmprint recognition. In: 19^{th} International Conference on Pattern Recognition, pp. 1–4 (2008)
16. Boyd, J., Little, J.: Biometric Gait Recognition. In: Tistarelli, M., Bigun, J., Grosso, E. (eds.) Advanced Studies in Biometrics. LNCS, vol. 3161, pp. 19–42. Springer, Heidelberg (2005)
17. Kittler, J., Hatef, M., Duin, R., Matas, J.: On combing classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(3), 226–239 (1998)
18. Snelick, R., Uludag, U., Mink, A., Indovina, M., Jain, A.K.: Large scale evaluation of multimodal biometric authentication using state-of-the-art systems. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(3), 450–455 (2005)
19. Yager, R.R., Liu, L.: Classic works of the Dempster-Shafer theory of belief functions. In: Studies in fuzziness and soft computing, vol. 219. Springer, Heidelberg (2008)
20. Bolle, R., Pankanti, S., Ratha, N.: Evaluation techniques for biometrics-based authentication systems(frr). In: Proceedings of the 15^{th} International Conference on Pattern Recognition, June 2000, vol. 2, pp. 831–837 (2000)
21. Marsaglia, G., Tsang, W.W.: The ziggurat method for generating random variables. Journal of Statistical Software 5(8) (2000)