Multiple classifiers in biometrics. Part 2: Trends and challenges

Julian Fierrez, Aythami Morales, Ruben Vera-Rodriguez, David Camacho

School of Engineering, Universidad Autonoma de Madrid, Madrid, Spain

Abstract

The present paper is Part 2 in this series of two papers. In Part 1 we provided an introduction to Multiple Classifier Systems (MCS) with a focus into the fundamentals: basic nomenclature, key elements, architecture, main methods, and prevalent theory and framework. Part 1 then overviewed the application of MCS to the particular field of multimodal biometric person authentication in the last 25 years, as a prototypical area in which MCS has resulted in important achievements.

Here in Part 2 we present in more technical detail recent trends and developments in MCS coming from multimodal biometrics that incorporate context information in an adaptive way. These new MCS architectures exploit input quality measures and pattern-specific particularities that move apart from general population statistics, resulting in robust multimodal biometric systems. Similarly as in Part 1, methods here are described in a general way so they can be applied to other information fusion problems as well. Finally, we also discuss here open challenges in biometrics in which MCS can play a key role.

Keywords:
classifier, fusion, biometrics, multimodal, adaptive, context

1. Introduction

The present paper is Part 2 in this series of two papers. In Part 1 we provided an introduction to Multiple Classifier Systems (MCS) with a focus into the fundamentals [1]: basic nomenclature, key elements, architecture, main methods, and prevalent theory and framework. Part 1 then overviewed the application of MCS to the particular field of multimodal biometric person

authentication in the last 25 years [2], as a prototypical area in which MCS has resulted in important achievements. That overview of MCS applied to biometrics was developed using a generic MCS framework and mathematical notation, with the purpose of facilitating the transfer of MCS achievements from here to other pattern recognition applications like video surveillance [3], speech technologies [4], human-computer interaction [5], data analytics applications [6], or recommender systems [7].

Here in Part 2 we present in more technical detail recent trends and developments in MCS coming from multimodal biometrics that incorporate context information in an adaptive way. These new MCS architectures exploit input quality measures [8] and pattern-specific particularities that move apart from general population statistics [9], resulting in robust multimodal biometric systems. Similarly as in Part 1, methods here are described in a general way so they can be applied to other information fusion problems as well.

We end this series of two papers with a discussion of open challenges in biometrics. The challenges exposed largely follow the excellent survey and outlook of the field of biometric person recognition by Jain et al. [2], which we complement with our personal view, and augment with the way MCS developments can advance those key challenges in biometrics. With that, we also hope to provide some light about the future of other pattern recognition and information fusion areas as well.

2. Trends in biometrics: Context-based MCS

This section is focused on MCS for multimodal biometric authentication, adapted both to user-specificities and to the input biometric quality. In the following sections we summarize key related works in these areas.

The adaptive MCS schemes for multimodal biometrics are divided into three classes: 1) user-dependent, 2) quality-based, and 3) user-dependent and quality-based. Although the last class includes the first two classes as particular cases, the three classes are introduced sequentially in order to facilitate the description.

For each class of methods, we first sketch the system model and then we derive particular implementations by using standard pattern recognition methods, either based on generative assumptions following Bayesian theory, or discriminative criteria using Support Vector Machines. These two classes

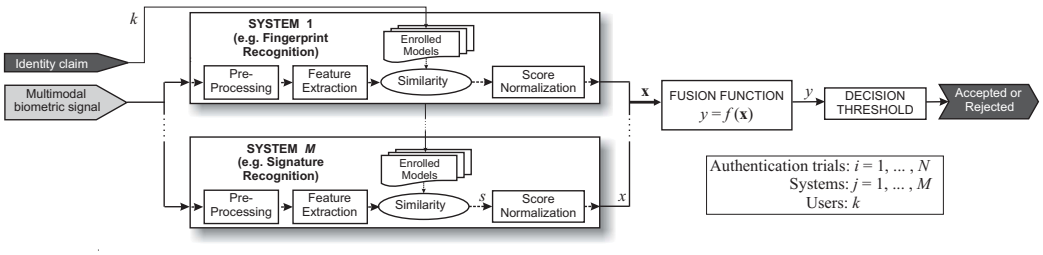

Figure 1: General system model of multimodal biometric authentication using score level fusion including name conventions.

of implementations aim at minimizing the Bayesian error and the Structural Risk of the verification task, respectively.

In the rest of the paper we use the following nomenclature and conventions. Given a multimodal biometric verification system consisting of M different unimodal systems $j = 1, \ldots, M$, each one computes a similarity score s between an input biometric pattern and the enrolled pattern or model of the given claimant k. The similarity scores s are normalized to x. Let the normalized similarity scores provided by the different unimodal systems be combined into a multimodal score $\mathbf{x} = [x_1, \ldots, x_M]^T$. The design of a fusion scheme consists in the definition of a function $f : \mathbb{R}^M \rightarrow \mathbb{R}$, so as to maximize the separability of client $\{f(\mathbf{x})|\text{client attempt}\}$ and impostor $\{f(\mathbf{x})|\text{impostor attempt}\}$ fused score distributions. This function may be trained by using labelled training scores $(\mathbf{x}_i, z_i)$, where $z_i = \{0 = \text{impostor attempt}, 1 = \text{client attempt}\}$.

In Fig. 1 we depict the general system model including all the notations defined above.

2.1. User-dependent multimodal biometrics

The idea of exploiting user-specific parameters at the score level in multimodal biometrics was introduced, to the best of our knowledge, by [10]. In this work, user-independent weighted linear combination of similarity scores was demonstrated to be improved by using either user-dependent weights or user-dependent decision thresholds, both of them computed by exhaustive search on the testing data. The idea of user-dependent fusion parameters was also explored by [11] using non-biased error estimation procedures. Other attempts to personalized multimodal biometrics include the use of the claimed identity index as a feature for a global trained fusion scheme based on Neu-

ral Networks [12], computing user-dependent weights using lambness metrics [13], and using personalized Fisher ratios [14].

Toh et al. [15] proposed a taxonomy of score-level fusion approaches for multi-biometrics. Multimodal fusion approaches were classified as global or local depending firstly on the fusion function (i.e., user-independent or user-dependent fusion strategies) and secondly depending on the decision making process (i.e., user-independent or user-dependent decision thresholds): global-learning and global-decision (GG), local-learning and global-decision LG, and similarly GL and LL. Some example works of each group are:

GG: [16, 17, 18, 19, 20, 21, 22].

LG: [10, 23, 12, 24, 11, 15, 13].

GL: [10, 15, 9]

LL: [15, 9]

These local methods (user-dependent fusion or decision) are confronted with a big challenge: training data scarcity, as the amount of available training data in localized learning is usually not sufficient and representative enough to guarantee good MCS parameter estimation and generalization capabilities. To cope with this lack of robustness derived from partial knowledge, the use of robust adaptive learning strategies based on background information was proposed in related research areas [25]. The idea of exploiting background information and adapt from there the fusion functions of MCS based on context information was introduced in biometrics by Fierrez et al. [9, 26], and was soon followed by others [27]. In brief, in these context-based MCS methods, the relative balance between the background information (from a pool of background users) and the local data (a given user) is performed as a tradeoff between both kinds of information.

The system model of user-dependent score fusion including the mentioned adaptation from background information is shown in Fig. 2.

Two selected algorithms implementing the discussed adapted user-dependent fusion are summarized in the following sections.

2.1.1. User-dependent MCS: combination approach

Here we outline this algorithm, representative of context-based MCS by adapting the score fusion function to each user from general background

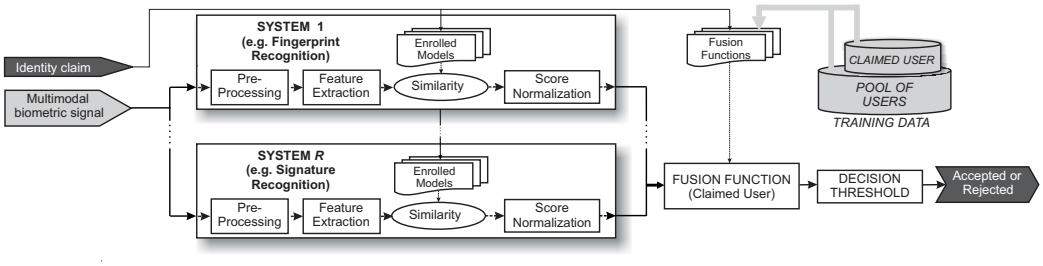

Figure 2: System model of multimodal biometric authentication with adapted user-dependent score fusion.

information. For a more detailed description and experimental evaluation see [26].

Impostor and client score distributions are modelled as multivariate Gaussians $p(\mathbf{x}|\omega_0) = N(\mathbf{x}|\boldsymbol{\mu}_0, \boldsymbol{\sigma}_0^2)$ and $p(\mathbf{x}|\omega_1) = N(\mathbf{x}|\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1^2)$, respectively[1]. The fused score y_T of a multimodal test score $\mathbf{x}_T$ is defined then as follows

$$y_T = f(\mathbf{x}_T) = \log p(\mathbf{x}_T|\omega_1) - \log p(\mathbf{x}_T|\omega_0), \tag{1}$$

which is known to be a Quadratic Discriminant (QD) function consistent with Bayes estimate in case of equal impostor and client prior probabilities [28]. The score distributions are estimated using the available training data as follows:

Global. The training set $X_{\mathrm{G}} = (\mathbf{x}_i, z_i)_{i=1}^{N_{\mathrm{G}}}$ includes multimodal scores from a number of different clients, and $(\{\boldsymbol{\mu}_{\mathrm{G},0}, \boldsymbol{\sigma}^2_{\mathrm{G},0}\}, \{\boldsymbol{\mu}_{\mathrm{G},1}, \boldsymbol{\sigma}^2_{\mathrm{G},1}\})$ are estimated by using the standard Maximum Likelihood criterion [29]. The resulting fusion rule $f_{\mathrm{G}}(\mathbf{x})$ is applied globally at the operational stage regardless of the claimed identity.

Local. A different fusion rule $f_{k,\mathrm{L}}(\mathbf{x})$ is obtained for each client k enrolled in the system by using Maximum Likelihood density estimates $(\{\boldsymbol{\mu}_{k,\mathrm{L},0}, \boldsymbol{\sigma}^2_{k,\mathrm{L},0}\}, \{\boldsymbol{\mu}_{k,\mathrm{L},1}, \boldsymbol{\sigma}^2_{k,\mathrm{L},1}\})$ computed from a set of development scores X_k of the specific client k.

Adapted. The adapted fusion rule $f_{k,\mathrm{A}}(\mathbf{x})$ of client k trades off the general knowledge provided by the user-independent development data X_{G},

[1]We use diagonal covariance matrixes, so $\boldsymbol{\sigma}^2$ is shorthand for $\mathrm{diag}(\Sigma)$. Similarly, $\boldsymbol{\mu}^2$ is shorthand for $\mathrm{diag}(\boldsymbol{\mu}\boldsymbol{\mu}')$.

and the user specificities provided by the user-dependent training set X_k, through Maximum a Posteriori density estimation [29]. This is done by adapting the sufficient statistics as follows

$$\begin{aligned} \boldsymbol{\mu}_{k,\mathrm{A},l} &= \alpha_l \boldsymbol{\mu}_{k,\mathrm{L},l} + (1-\alpha_l)\boldsymbol{\mu}_{\mathrm{G},l}, \\ \boldsymbol{\sigma}^2_{k,\mathrm{A},l} &= \alpha_l(\boldsymbol{\sigma}^2_{k,\mathrm{L},l} + \boldsymbol{\mu}^2_{k,\mathrm{L},l}) + (1-\alpha_l)(\boldsymbol{\sigma}^2_{\mathrm{G},l} + \boldsymbol{\mu}^2_{\mathrm{G},l}) - \boldsymbol{\mu}^2_{j,\mathrm{A},l}. \end{aligned} \tag{2}$$

For each class $l = \{0 = \text{impostor}, 1 = \text{client}\}$, a data-dependent adaptation coefficient

$$\alpha_l = N_l/(N_l + r) \tag{3}$$

is used, where N_l is the number of local training scores in class l, and r is a fixed relevance factor.

Note that other statistical models or other techniques for trading-off the general and local knowledge can be used in a similar way.

2.1.2. User-dependent MCS: classification approach

Similarly as before, we only outline here the main aspects of this context-based MCS approach, which also adapts the score fusion function to each user from general background information. This particular implementation is based on SVM, but the approach is easily extensible to any other binary classifier. For a detailed description and experimental evaluation see [9].

Without loss of generality, suppose we train a SVM classifier with the following training set: $X = (\mathbf{x}_i, z_i)_{i=1}^N$ where N is the number of multimodal scores in the training set, and $z_i \in \{-1, 1\} = \{\text{Impostor}, \text{Client}\}$. We train the SVM classifier by solving the following quadratic programming problem [30]:

$$\min_{\mathbf{w}, w_0, \xi_1, \ldots, \xi_N} \left(\tfrac{1}{2} \|\mathbf{w}\|^2 + \sum_{i=1}^{N} C_i \xi_i \right) \tag{4}$$

subject to

$$\begin{aligned} z_i(\langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle_{\mathbb{H}} + w_0) &\geq 1 - \xi_i, \qquad i = 1, \ldots, N, \\ \xi_i &\geq 0, \qquad i = 1, \ldots, N, \end{aligned} \tag{5}$$

where slack variables ξ_i are introduced to take into account the eventual non-separability of $\Phi(X)$ and parameter $C_i = C$ is a positive constant that controls the relative influence of the two competing terms.

The optimization problem in Eqs. (4) and (5) is solved with the Wolfe dual representation by using the kernel trick [31]:

$$\max_{\alpha_1,\ldots,\alpha_N} \left(\sum_{i=1}^{N} \alpha_i - \tfrac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j z_i z_j K(\mathbf{x}_i, \mathbf{x}_j) \right) \tag{6}$$

subject to

$$\begin{aligned} & 0 \leq \alpha_i \leq C_i, \quad i = 1, \ldots, N \\ & \sum_{i=1}^{N} \alpha_i z_i = 0 \end{aligned} \tag{7}$$

where the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle_{\mathbb{H}}$ is introduced to avoid direct manipulation of the elements of $\mathbb{H}$. Typical kernel functions include radial basis functions

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2 \right), \tag{8}$$

and linear kernels

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j. \tag{9}$$

resulting in complex and linear separating surfaces between client and impostor distributions, respectively.

The fused score y_T of a multimodal test pattern $\mathbf{x}_T$ is defined as follows:

$$y_T = f(\mathbf{x}_T) = \langle \mathbf{w}^*, \Phi(\mathbf{x}_T) \rangle_{\mathbb{H}} + w_0^*, \tag{10}$$

which is a signed distance measure from $\mathbf{x}_T$ to the separating surface given by the solution of the SVM problem. Applying the Karush-Kuhn-Tucker (KKT) conditions to the problem in Eqs. (4) and (5), y_T can be shown to be equivalent to the following sparse expression

$$y_T = f(\mathbf{x}_T) = \sum_{i \in \mathrm{SV}} \alpha_i^* y_i K(\mathbf{x}_i, \mathbf{x}_T) + w_0^*, \tag{11}$$

where $(\mathbf{w}^*, w_0^*)$ is the optimal hyperplane, $(\alpha_1^*, \ldots, \alpha_N^*)$ is the solution to the problem in Eqs. (6) and (7), and $\mathrm{SV} = \{i | \alpha_i^* > 0\}$ indexes the set of support

vectors. The bias parameter w_0^* is obtained from the solution to the problem in Eqs. (6) and (7) by using the KKT conditions [31].

As a result, the training procedure in Eqs. (6) and (7) and the testing strategy in Eq. (11) are obtained for the problem of multimodal fusion.

Global. The training set $X_{\mathrm{G}} = (\mathbf{x}_i, z_i)_{i=1}^{N_{\mathrm{G}}}$ includes multimodal scores from a number of different clients and the resulting fusion rule $f_{\mathrm{G}}(\mathbf{x})$ is applied globally at the operational stage regardless of the claimed identity.

Local. A different fusion rule $f_{k,\mathrm{L}}(\mathbf{x})$ is obtained for each client enrolled in the system k by using development scores X_k of the specific client k. At the operational stage, the fusion rule $f_{k,\mathrm{L}}(\mathbf{x})$ of the claimed identity k is applied.

Adapted. This scheme trades off the general knowledge provided by a user-independent training set X_{G}, and the user specificities provided by a user-dependent training set X_k. To obtain the adapted fusion rule, $f_{k,\mathrm{A}}(\mathbf{x})$, for user k, we compute both the global fusion rule, $f_{\mathrm{G}}(\mathbf{x})$, and the local fusion rule, $f_{k,\mathrm{L}}(\mathbf{x})$, as described above, and finally combine them as follows:

$$f_{k,\mathrm{A}}(\mathbf{x}) = \alpha f_{k,\mathrm{L}}(\mathbf{x}) + (1-\alpha) f_{\mathrm{G}}(\mathbf{x}), \tag{12}$$

where α is a trade-off parameter. This can be seen as a user-dependent fusion scheme adapted from user-independent information. The idea can also be extended easily to trained fusion schemes based on other classifiers. Worth noting, sequential algorithms to solve the SVM optimization problem in Eqs. (4) and (5) have been already proposed [32], and can be used to extend the proposed idea, first constructing the user-independent solution and then refining it by incorporating the local data.

2.1.3. User-dependent decision

The system model of user-dependent decision is shown in Fig. 3. Once a fused similarity score has been obtained by using either a global, local or an adapted fusion method, the score is compared to a decision threshold in order to accept or reject the identity claim. This decision making process, also subject to training, can also be made globally, locally, or can be adapted from global to local information. For this purpose, the methods presented

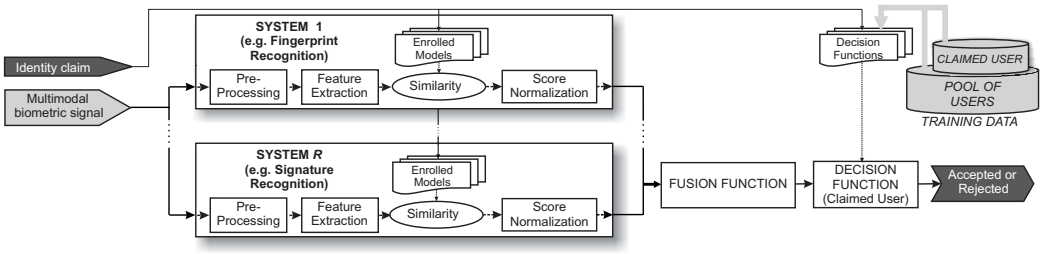

Figure 3: System model of multimodal biometric authentication with adapted user-dependent decision.

in Sects. 2.1.1 and 2.1.2 can be directly applied exchanging the input multimodal scores $\mathbf{x}$ for fused scores y.

2.2. Quality-based multimodal biometrics

The 21st century began with a growing interest in studying the effects of signal quality on the performance of biometric systems [33, 34, 35]. As a result, it was shown in several works that the performance of an unimodal system can drop significantly under noisy conditions [36]. Multimodal systems have been demonstrated to overcome this challenge to some extent by combining the evidences provided by a number of different traits. This idea can be extended by explicitly considering quality measures of the input biometric signals and weighting the various pieces of evidence based on this quality information. Following this idea, various quality-based multimodal authentication schemes were proposed and studied since mid 2000s [8].

Quality measures of the input biometric signals can be used for adapting the different modules of a multimodal authentication system [36]. Here we concentrate in quality-based score fusion. The system model of quality-based score fusion is shown in Fig. 4.

Bigun et al. [17] studied the problem of multimodal biometric authentication by using Bayesian statistics. The result was an Expert Conciliation scheme including weighting factors not only for the accuracy of the experts but also for the confidence of the experts on the particular input samples. Experiments were provided by combining face and voice modalities. The idea of relating the confidence value to quality measures of the input biometric signals was nevertheless not developed.

The concept of confidence measure of matching scores was also studied by [37]. In that work Bengio et al. demonstrated that the confidence of matching

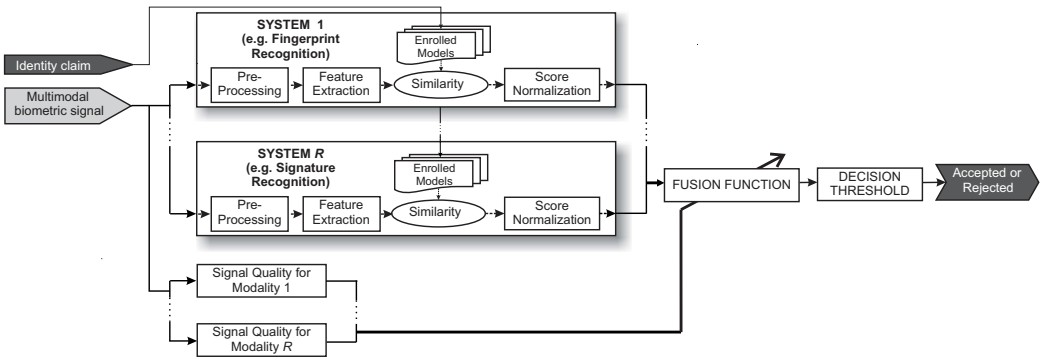

Figure 4: System model of multimodal biometric authentication with quality-based score fusion.

scores can help in the fusion process. In particular, they tested confidence measures based on: 1) Gaussian assumptions on the score distributions, 2) the adequacy of the trained biometric models to explain the input data, and 3) resampling techniques on the set of test scores. This research line was further developed by Poh and Bengio [38], who devised confidence measures based on the margin between impostor and client score distributions.

Chatzis et al. [21] evaluated a number of fusion schemes based on clustering strategies. In this case quality measures obtained directly from the input biometric signals were used to fuzzify the scores provided by the different systems. They demonstrated that fuzzy versions of k-means and Vector Quantization including the quality measures outperformed slightly, and not in all cases, the standard non-fuzzy clustering methods. This work is, to the best of our knowledge, the first one reporting results of quality-based fusion. One limitation in the experimental setup of this work was the reduced number of individuals used, only 37.

Another work in quality-based fusion without the success of previous methods was reported by Toh et al. [39], who developed a score fusion scheme based on polynomial functions. Quality measures were introduced in the optimization problem for training the polynomials as weights in the regularization term. Unexpectedly, no performance improvements were obtained by including the quality measures. One inconvenience of this work was the use of a chimeric multimodal database combining the data from 3 different face, voice and fingerprint databases.

2.2.1. Quality-based MCS: combination approach

One straightforward way to incorporate the input biometric quality to the score fusion approach is by including weights in simple combination approaches. In the case of the weighted average presented in Part 1 Eq. (10), this can be achieved by using $w_j = q_j$ in order to obtain the following quality-based score fusion function

$$y = \sum_{j=1}^{M} q_j x_j, \tag{13}$$

where q_j is a quality measure of the score x_j. This score quality should be ideally related to the confidence of the system j in providing a reliable matching score for the particular biometric signal being tested [40, 41]. The score quality proposed and used in [8] is as follows:

$$q = \sqrt{Q \cdot Q_{\text{claim}}}, \tag{14}$$

where Q and Q_{claim} are the input biometric quality and the average quality of the biometric signals used for enrollment, respectively. The two quality measures Q and Q_{claim} are supposed to be in the range $[0, 1]$ where 0 corresponds to the poorest quality, and 1 corresponds to the highest quality.

Other definitions of score quality found in the literature include [36]: $q = (Q + Q_{\text{claim}})/2$, $q = \min\{Q, Q_{\text{claim}}\}$, etc.

Preliminaries.. The nomenclature and conventions summarized in Fig. 1 are extended here:

x_{ij} Similarity score i delivered by system j

v_{ij} Variance of x_{ij} as estimated by system j

z_i The true label corresponding to score i

ζ_{ij} The error score $\zeta_{ij} = z_i - x_{ij}$

With respect to the previous cases developed in this paper, note that here we introduce the variance v_{ij} of the input scores x_{ij}. The true labels z_i can take only two numerical values corresponding to "Impostor" and "Client". If x_{ij} is between 0 and 1 then these values are chosen to be 0 and 1, respectively. The fusion function is trained on shots $i \in 1 \dots N$ (i.e. x_{ij} and z_i are known for $i \in 1 \dots N$) and we consider the trial $N + 1$ as a test shot on the working multimodal system (i.e. $x_{(N+1)j}$ is known, but z_{N+1} is not known).

Statistical Model. The model for combining the different systems (here also called machine experts) is based on Bayesian statistics and the assumption of normal distributed expert errors, i.e. ζ_{ij} is considered to be a sample of a normally distributed random variable. It has been shown experimentally [17] that this assumption does not strictly hold for common audio- and video-based biometric machine experts, but it is shown that it holds reasonably well when client and impostor distributions are considered separately. Taking this result into account, two different fusion functions are constructed, one of them based on genuine scores

$$\mathcal{C} = \{x_{ij}, v_{ij} | 1 \leq i \leq N \text{ and } z_i = 1, 1 \leq j \leq M\}, \tag{15}$$

while the other is based on impostor scores

$$\mathcal{I} = \{x_{ij}, v_{ij} | 1 \leq i \leq N \text{ and } z_i = 0, 1 \leq j \leq M\}. \tag{16}$$

The two fusion functions will be referred to as *client function* and *impostor function* respectively.

The client function estimates the expected true label of an input claim based on its expertise on recognizing client data. More formally, it computes $M''_{\mathcal{C}} = E\left[Z_{N+1} | \mathcal{C}, x_{N+1,j}\right]$. Similarly, the impostor function computes $M''_{\mathcal{I}} = E\left[Z_{N+1} | \mathcal{I}, x_{N+1,j}\right]$. The conciliated overall score M'' takes into account the different expertise of the two fusion functions and chooses the one which came closest to its goal, i.e. 0 for the impostor function and 1 for the client function:

$$M'' = \begin{cases} M''_{\mathcal{C}} & \text{if } |1 - M''_{\mathcal{C}}| - |0 - M''_{\mathcal{I}}| < 0 \\ M''_{\mathcal{I}} & \text{otherwise} \end{cases}. \tag{17}$$

Based on the normality assumption of the errors, the fusion training and testing algorithm described in [17] is obtained, see [42] for further background and details. In the following paragraphs we summarize the resulting algorithm in the two cases where it can be applied.

Bayesian simplified quality-based score fusion. When only the similarity scores x_{ij} are available, the following simplified fusion function is obtained by using $v_{ij} = 1$:

Training. Estimate the bias parameters of each system. The bias parameters for the client function are

$$M_{\mathcal{C}j} = \frac{1}{n_{\mathcal{C}}} \sum_i \zeta_{ij} \quad \text{and} \quad V_{\mathcal{C}j} = \frac{\alpha_{\mathcal{C}j}}{n_{\mathcal{C}}}, \tag{18}$$

where i indexes the training set $\mathcal{C}$, $n_{\mathcal{C}}$ is the number of training samples in $\mathcal{C}$ and

$$\alpha_{\mathcal{C}j} = \frac{1}{n_{\mathcal{C}} - 3} \left(\sum_i \zeta_{ij}^2 - \frac{1}{n_{\mathcal{C}}} \left(\sum_i \zeta_{ij} \right)^2 \right). \tag{19}$$

Similarly $M_{\mathcal{I}j}$ and $V_{\mathcal{I}j}$ are obtained for the impostor function.

Authentication. At this step, both fusion functions are operational, so that the time instant is $N + 1$ and the fusion functions have access to the similarity scores $x_{N+1,j}$ but not to the true label z_{N+1}. First the client and impostor functions are calibrated according to their past performance, yielding (for the client function)

$$M'_{\mathcal{C}j} = x_{n+1,j} + M_{\mathcal{C}j} \text{ and } V'_{\mathcal{C}j} = (n_{\mathcal{C}} + 1)V_{\mathcal{C}j}, \tag{20}$$

and then the different calibrated systems are combined according to

$$M''_{\mathcal{C}} = \frac{\sum_{j=1}^{M} \frac{M'_{\mathcal{C}j}}{V'_{\mathcal{C}j}}}{\sum_{j=1}^{M} \frac{1}{V'_{\mathcal{C}j}}}. \tag{21}$$

Similarly, $M'_{\mathcal{I}}$, $V'_{\mathcal{I}}$ and $M''_{\mathcal{I}}$ are obtained. The final fused output is obtained according to Eq. (17).

The algorithm described above has been successfully applied in [43] in a multimodal authentication system combining face and speech data. Verification performance improvements of almost an order magnitude were reported as compared to the best modality.

Bayesian quality-based score fusion. When not only the scores but also the score variances are available, the following algorithm is obtained:

Training. Estimate the bias parameters. For the client function

$$M_{\mathcal{C}j} = \frac{\sum_i \frac{\zeta_{ij}}{\sigma_{ij}^2}}{\sum_i \frac{1}{\sigma_{ij}^2}} \quad \text{and} \quad V_{\mathcal{C}i} = \frac{1}{\sum_i \frac{1}{\sigma_{ij}^2}}, \tag{22}$$

where the training set $\mathcal{C}$ is used. The variances σ_{ij}^2 are estimated through $\bar{\sigma}_{ij}^2 = v_{ij} \cdot \alpha_{\mathcal{C}j}$, where

$$\alpha_{\mathcal{C}j} = \frac{1}{n_{\mathcal{C}} - 3} \left(\sum_i \frac{\zeta_{ij}^2}{v_{ij}} - \left(\sum_i \frac{\zeta_{ij}}{v_{ij}} \right)^2 \left(\sum_i \frac{1}{v_{ij}} \right)^{-1} \right). \tag{23}$$

Similarly $M_{\mathcal{I}j}$ and $V_{\mathcal{I}j}$ are obtained for the impostor function.

Authentication. First we calibrate the systems according to their past performance, for the client function

$$M'_{\mathcal{C}j} = x_{N+1,j} + M_{\mathcal{C}j} \text{ and } V'_{\mathcal{C}j} = v_{N+1,j}\alpha_{\mathcal{C}j} + V_{\mathcal{C}j}, \tag{24}$$

and then the different calibrated systems are combined according to Eq. (21). Similarly, $M'_{\mathcal{I}}$, $V'_{\mathcal{I}}$ and $M''_{\mathcal{I}}$ are obtained. The final fused score is obtained according to Eq. (17). This combined output can be expressed in the form of Eq. (11) from Part 1.

The algorithm described above has been successfully applied not only in biometrics where it was originated [44], but also in other unrelated fields like risk assessment of aircraft accidents [42].

The variance v_{ij} of the score x_{ij} concerns a particular authentication assessment. It is not a general reliability measure for the system itself, but a certainty measure based on the performance of the system and the data being assessed. Typically the variance of the score is chosen as the width of the range in which one can place the score when considering human opinions. Because such intervals can be conveniently provided by a human expert, the algorithm presented here constitutes a systematic way of combining human and machine expertise in MCS applications. An example of such an application is forensic reporting using biometric evidences, where machine expert approaches are increasingly being used [45] and human opinions must be taken into consideration.

The context-based MCS approach summarized here calculates v_{ij} as a function of quality measures computed on the input biometric signals (see Fig. 4). This implies taking into account Eq. (24) right, that the trained fusion function adapts the weights of the experts using the input signal quality. For that purpose the quality q_{ij} of the score x_{ij} is defined as:

$$q_{ij} = \sqrt{Q_{ij} \cdot Q_{\text{claim},j}}, \tag{25}$$

where Q_{ij} and $Q_{\text{claim},j}$ are the quality label of the biometric trait j in trial i and the average quality of the biometric signals used by the system j for modelling the claimed identity respectively. The two quality labels Q_{ij} and $Q_{\text{claim},j}$ are supposed to be in the range $[0, Q_{\max}]$ with $Q_{\max} > 1$, where 0 corresponds to the poorest quality, 1 corresponds to normal quality and $Q_{\max}$ corresponds to the highest quality. Finally, the variance parameter is calculated according to

$$v_{ij} = \frac{1}{q_{ij}^2}. \tag{26}$$

Experimental evaluation of this quality-based fusion approach can be found in [44, 42].

2.2.2. Quality-based MCS: classification approach

Instead of assuming particular statical models on the genuine and impostor score distributions like in previous section, here we exemplify a quality-based score fusion approach based on any binary classifier. Without loss of generality, we sketch the approach considering SVM classifiers [8].

Let $\mathbf{q} = [q_1, \ldots, q_M]^T$ denote the quality vector of the multimodal similarity score $\mathbf{x} = [x_1, \ldots, x_M]^T$, where q_j is a scalar quality measure corresponding to the similarity score x_j with $j = 1, \ldots, M$ being M the number of modalities. As in the case of the Bayesian quality-based fusion algorithm, the quality values q_j are computed as follows:

$$q_j = \sqrt{Q_j \cdot Q_{\text{claim},j}}, \tag{27}$$

where Q_j and $Q_{\text{claim},j}$ are the quality measure of the sensed signal for biometric trait j, and the average signal quality of the biometric signals used by unimodal system j for modelling the claimed identity, respectively. The two quality labels Q_j and $Q_{\text{claim},j}$ are supposed to be in the range $[0, Q_{\max}]$ with $Q_{\max} > 1$, where 0 corresponds to the poorest quality, 1 corresponds to standard quality, and $Q_{\max}$ corresponds to the highest quality.

The score-level fusion scheme based on SVM classifiers and quality measures proposed in [8] is as follows:

Training. An initial fusion function:

$$f_{\text{SVM}} : \mathbb{R}^M \to \mathbb{R}, f_{\text{SVM}}(\mathbf{x}_T) = \langle \mathbf{w}, \Phi(\mathbf{x}_T) \rangle + w_0 \tag{28}$$

is trained by solving the problem:

$$\min_{\mathbf{w}, w_0, \xi_1, \ldots, \xi_N} \left(\tfrac{1}{2} \|\mathbf{w}\|^2 + \sum_{i=1}^{N} C_i \xi_i \right) \tag{29}$$

subject to

$$y_i(\langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle_{\mathbb{H}} + w_0) \geq 1 - \xi_i, \quad i = 1, \ldots, N, \tag{30}$$

$$\xi_i \geq 0, \qquad i = 1, \ldots, N, \tag{31}$$

as described in Sect. 2.1.2, but using as cost weights

$$C_i = C \left(\frac{\prod_{j=1}^{M} q_{i,j}}{Q_{\max}^{M}} \right)^{\alpha_1}, \tag{32}$$

where $q_{i,j}$, $j = 1, \ldots, M$ are the components of the quality vector $\mathbf{q}_i$ associated with training sample $(\mathbf{x}_i, z_i)$, $z_i \in \{-1, 1\} = \{\text{Impostor}, \text{Client}\}$, and C is a positive constant. As a result, the higher the overall quality of a multimodal training score the higher its contribution to the computation of the initial fusion function. Additionally, M SVMs of dimension $M - 1$ (SVM_1 to SVM_M) are trained leaving out traits 1 to M respectively. Similarly to Eq. (32)

$$C_i = C \left(\frac{\prod_{r \neq j} q_{i,r}}{Q_{\max}^{(M-1)}} \right)^{\alpha_1}, \tag{33}$$

for SVM_j with $j = 1, \ldots, M$.

Authentication. Let the sensed multimodal biometric sample generate a quality vector $\mathbf{q}_T = [q_{T,1}, \ldots, q_{T,M}]^T$. Re-index the individual traits in order to have $q_{T,1} \leq q_{T,2} \leq \ldots \leq q_{T,M}$. A multimodal similarity score $\mathbf{x}_T = [x_{T,1}, \ldots, x_{T,M}]'$ is then generated. The combined quality-based similarity score is computed as follows:

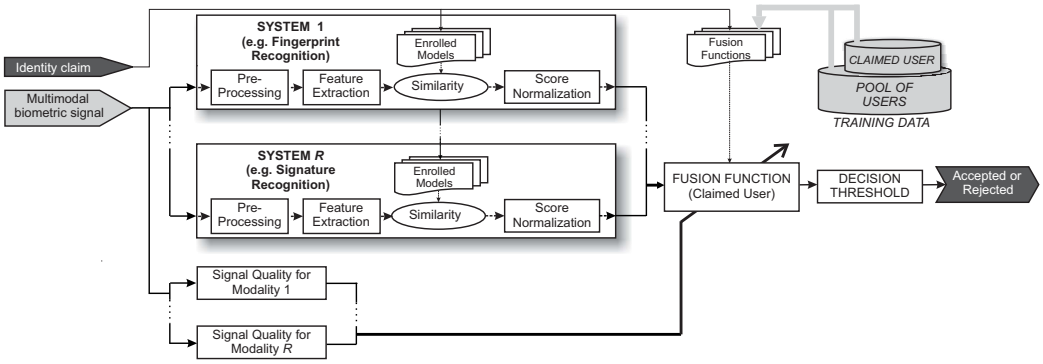

Figure 5: System model of multimodal biometric authentication with user-dependent and quality-based score fusion.

$$f_{\mathrm{SVM}_Q}(\mathbf{x}_T) = \beta_1 \sum_{j=1}^{M-1} \frac{\beta_j}{\sum_{r=1}^{M-1} \beta_r} f_{\mathrm{SVM}_j}(\mathbf{x}_T^{(j)}) + (1-\beta_1) f_{\mathrm{SVM}}(\mathbf{x}_T), \quad (34)$$

where $\mathbf{x}_T^{(j)} = [x_{T,1}, \ldots, x_{T,j-1}, x_{T,j+1}, \ldots, x_{T,M}]^T$ and

$$\beta_j = \left(\frac{q_{T,M} - q_{T,j}}{Q_{\max}} \right)^{\alpha_2}, \quad j = 1, \ldots, M-1. \quad (35)$$

As a result, the adapted fusion function in Eq. (34) is a quality-based trade-off between not using and using low quality traits.

2.3. User-dependent and quality-based multimodal biometrics

Finally, we may combine previous strategies to derive fusion systems adapted both to the user at hand and to the input biometric quality, as shown in Fig. 5.

Practical implementations of this scheme can be obtained by combining some of the procedures described previously in the present paper. One possibility is to use Bayesian user-dependent score fusion plus discriminative quality-based adaptation.

3. Challenges in biometrics: Role of MCS

In the present section, similarly as in the excellent exposition by Jain et al. [2], we discuss main challenges in biometrics, adapting their discussion based on our personal view, and commenting how new MCS developments may play a role in overcoming those challenges.

Note that biometrics person recognition shares architectures, methods, issues, and challenges with almost any other pattern recognition application. Therefore, the challenges exposed here have a parallel in other research areas, and may provide some light on the future of other pattern recognition applications as well.

Challenge 0: Better understanding about the nature of biometrics (distinctiveness and permanence). Current knowledge about the nature of the variety of biometric modalities useful for person recognition is quite limited [2]. Although practical systems based on fingerprint or face recognition can satisfy certain applications, a better understanding of factors like their intrinsic distinctive capacity [46, 47], or their permanence [48, 49], will open the way to new improved recognition, and will rationalize the application of such technologies depending on the scenario of application and potential population of use [50].

There have been some advances in these areas, but still much work is necessary to fully understand the nature of biometrics for person authentication. Towards this objective, MCS approaches can be instrumental for analyzing the increasing amount of multimodal biometric data available nowadays [51, 52]. MCS methods can be quite helpful to analyze those data as they permit to simultaneously analyse and model complex yet structured relations on heterogenous data [53], which is the case in biometrics, e.g.: the different representation levels existing in fingerprint [54, 55], or speech [56, 57].

Challenge 1: Design of robust algorithms (representation and matching) from uncooperative users in unconstrained and varying scenarios. This challenge has been the main focus of research in biometrics during the last 50 years [2], and still the desirable performance level for many biometric applications in realistic scenarios is not yet satisfactory. There are a myriad of pattern representation schemes and matching procedures depending on the biometric modality (e.g., face image vs speech time-sequences) and acquisition scenario (e.g., controlled vs latent fingerprints), and one can find in the vast and growing literature representation and matching methods specifically adjusted for

many practical applications. Most of these approaches are variants of successful representation and matching techniques coming from other research areas like image and signal processing, speech analysis, or computer vision, e.g., LBP or SIFT features [58].

As developed in Part 1 in our review of MCS applied to multimodal biometrics, combining various of such representation-matching schemes provide significant benefits, not only when one has multiple evidences to combine [59], but also when one has only one biometric evidence but wants to be robust against degraded or varying conditions by combining various representation schemes [60]. The success of such MCS schemes is related to the diversity of classifiers being combined, a topic attracting much attention in the MCS community [61, 7].

MCS strategies in previous paragraph supposed that there are various classifiers available to be combined, but one can also generate multiple base classifiers, e.g., the highly successful AdaBoost approach in the Viola-Jones cascade MCS [62]. These MCS approaches are specially useful when patterns to be recognized are difficult to be represented, or vary in time due to its intrinsic nature or environmental changes. An adaptive generation of multiple base classifiers, and adaptive fusion schemes, like AdaBoost, may track and adapt well under those unconstrained and varying conditions. This topic of adaptive pattern recognition is also source of interesting research in MCS under multiple names like concept drift [63, 64]. Advances in adaptive MCS can be instrumental for the future of this Challenge 1. In addition to such adaptive schemes, a better understanding of such unconstrained scenarios through benchmarks and public databases is also of outmost importance [65, 66].

On the other hand, in the last 5 years or so we have witnesses the triumph of data-agnostic (i.e., without any explicit representation) end-to-end machine learning approaches such as deep neural networks that, given enough representative training data, can generate very robust classifiers for many problems in unconstrained scenarios with highly varying conditions, e.g., face [67] or speaker recognition [68].

MCS methods exploiting deep learning [69], and new deep learning strategies exploiting and considering both existing classifiers (a common case in biometric applications) and contextual information [70] are also very promising lines for advancing in this area.

Challenge 2: Integration with end applications. Most traditional and widely deployed biometric solutions for person recognition are designed for access control or forensic scenarios. One important challenge in biometrics is how to properly integrate biometrics technologies in other application scenarios like mobile authentication [71, 72], video surveillance [3], forensics [73], large-scale ID [74], cloud biometrics or ubiquitous biometrics [75].

Depending on the scenario at hand, the traditional biometric technologies will need to be adapted, or perhaps designed again in order to satisfy new application requirements. In this case adaptive MCS techniques incorporating context information, like the ones described here in Section 2, can be quite useful.

Challenge 3: Understanding and improving the usability. As mentioned in Challenge 2, the number and variety of biometric applications for person recognition is ever growing, and some of them are strongly dependent on an adequate interaction between the user and the biometric sensor, e.g., in mobile authentication [71].

We currently lack a good understanding of how the people naturally interact with some biometric sensors, and in which conditions the authentication mechanisms generated with biometric technology perform best. There has been some research in the past to analyze those factors between the user and the biometric sensor in general [76], including specific models to analyze and exploit the interaction between the user and the biometric sensor [77]. More recently, we can see some targeted studies towards understanding the interaction between users and technology for key biometric end applications like border control [78], or smartphone unlock [79].

Similar to Challenge 0, MCS approaches can be exploited here as a tool for analyzing multiple sources of heterogenous data [53], complex yet structured, as is the case of human-biometric sensor interaction data [77].

Challenge 4: Understanding and improving the security. Pattern recognition applications based on biometrics are usually intended for securing information or control the access to services or places [2]. Note this is not the only usage possible, as biometric technologies may be also used to analyze personal data towards other objectives, like behaviour analysis [80] or medical diagnosis [81].

When biometrics are used for security applications, one may want to know the level of security provided by the application at hand, given a set of operational conditions. This question has been already addressed in the general

information security community, where various international standards have been generated under the umbrella of Common Criteria (ISO/IEC 15408) since 1990 [82]. That standardization effort includes some specific developments for biometric systems [83]. The basic idea behind those standards is to measure quantitatively the effort required for potential attackers to bypass the protection provided by biometrics, and the impact of such attacks.

These ideas have generated much research in biometrics towards understanding possible attacks [84], and the generation of protection methods against attacks [85]. When MCS approaches are applied to biometrics, specific vulnerabilities appear [86], and protection methods can be generated by exploiting specific MCS fusion strategies [87].

The topic of security against attackers seeking illicit access is related to the privacy protection of users, and in particular their biometric templates. Securing such templates against potential identity theft has also generated much research activity in the last decade [88]. There are some recent developments in this area exploiting advances in cryptography like homomorphic encryption [89], but still there are no general satisfactory solutions for generating secure biometric templates at the same time 1) non-invertible, 2) non-linkable, and 3) with high discrimination [2]. Current trends for better protecting templates containing multiple biometric data are usually based on advanced cryptographic constructions and the principles of MCS described in Part 1 [90].

4. Conclusions

The present paper is the Part 2 in a series of two papers. In Part 1 we first provided a brief introduction to Multiple Classifier Systems (MCS) including basic nomenclature, architecture, and key elements [1]. Our main focus there was into the fundamentals of MCS, providing pointers for detailed descriptions of MCS algorithms.

Part 1 then overviewed the application of MCS to the particular field of multimodal biometric person authentication in the last 25 years [2], including general descriptions of main MCS elements, methods, and algorithms generated in the biometrics field. The presentation there was general with a generic mathematical formulation, in order to facilitate the export of experiences and methods to other information fusion problems, e.g.: video surveillance [3], speech technologies [4], biomedical applications [91], human-computer interaction [5], data analytics [6], or recommender systems [7].

Part 1 was intended for the non-expert in MCS, or any other reader interested in overviewing the field of multimodal biometrics. Here in Part 2 we provide more advanced material intended for researchers knowledgeable already in MCS and multimodal biometrics, readers that completed Part 1, and any other researcher seeking ideas and prospects about the future of biometrics that can be parallel to other pattern recognition areas as well.

We began this Part 2 describing in technical detail recent trends and developments in MCS from multimodal biometrics that incorporate context information in an adaptive way, using the framework and mathematical tools introduced in Part 1. These new MCS architectures exploit input quality measures [8] and pattern-specific particularities that move apart from general population statistics [9], resulting in robust multimodal biometric systems.

Similarly as in Part 1, methods here in Part 2 were introduced in a general way so they can be applied to other information fusion problems as well. In related works such as [92], one can find an excellent treatment of general context-based information fusion, in which there are indications on how to apply the methods and specific algorithms developed here to other information fusion architectures.

Finally, we have discussed open challenges in biometrics in which MCS may play a key role: 0) limited knowledge about the nature of biometrics (in terms of distinctiveness and permanence for different populations), 1) design of robust algorithms (representation and matching) from uncooperative users in unconstrained and varying scenarios, 2) integration with end applications, 3) understanding and improving the usability, and 4) understanding and improving the security.

5. Acknowledgements

This work was funded by projects CogniMetrics (TEC2015-70627-R) from MINECO/FEDER and RiskTrakc (JUST-2015-JCOO-AG-1). Part of this work was conducted during a research visit of J.F. to Prof. Ludmila Kuncheva at Bangor University (UK) with STSM funding from COST CA16101 (MULTI-FORESEE). Author J.F. want to thank Prof. Kuncheva for fruitful discussions during his visit.

References

[1] L. I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, Wiley, 2014.

[2] A. K. Jain, K. Nandakumar, A. Ross, 50 years of biometric research, Pattern Recogn. Lett. 79 (2016) 80–105.

[3] A. Garcia-Martin, J. M. Martinez, People detection in surveillance: classification and evaluation, IET Computer Vision 9 (2015) 779–788(9).

[4] I. Lopez-Moreno, J. Gonzalez-Dominguez, D. Martinez, O. Plchot, J. Gonzalez-Rodriguez, P. J. Moreno, On the use of deep feedforward neural networks for automatic language identification, Computer Speech and Language 40 (2016) 46 – 59.

[5] D. Rozado, T. Moreno, J. S. Agustin, F. B. Rodriguez, P. Varona, Controlling a smartphone using gaze gestures as the input mechanism, Human-Computer Interaction 30 (1) (2015) 34–63.

[6] G. Bello-Orgaz, J. J. Jung, D. Camacho, Social big data: Recent achievements and new challenges, Information Fusion 28 (2016) 45 – 59.

[7] P. Castells, N. J. Hurley, S. Vargas, Recommender Systems Handbook, Springer US, 2015, Ch. Novelty and Diversity in Recommender Systems, pp. 881–918.

[8] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, J. Bigun, Discriminative multimodal biometric authentication based on quality measures, Pattern Recognition 38 (5) (2005) 777–779.

[9] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Adapted user-dependent multimodal biometric authentication exploiting general information, Pattern Recognition Letters 26 (16) (2005) 2628–2639.

[10] A. K. Jain, A. Ross, Learning user-specific parameters in a multibiometric system, in: Proc. of IEEE Intl. Conf. on Image Processing, ICIP, Vol. 1, 2002, pp. 57–60.

[11] Y. Wang, Y. Wang, T. Tan, Combining fingerprint and voice biometrics for identity verification: An experimental comparison, in: D. Zhang, A. K. Jain (Eds.), Proc. of Intl. Conf. on Biometric Authentication, ICBA, Springer LNCS-3072, 2004, pp. 663–670.

[12] A. Kumar, D. Zhang, Integrating palmprint with face for user authentication, in: Proc. of Workshop on Multimodal User Authentication, MMUA, 2003, pp. 107–112.

[13] R. Snelick, U. Uludag, A. Mink, M. Indovina, A. K. Jain, Large scale evaluation of multimodal biometric authentication using state-of-the-art systems, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (3) (2005) 450–455.

[14] N. Poh, S. Bengio, An investigation of f-ratio client-dependent normalisation on biometric authentication tasks, in: Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, ICASSP, Vol. 1, 2005, pp. 721–724.

[15] K. A. Toh, X. Jiang, W. Y. Yau, Exploiting local and global decisions for multimodal biometrics verification, IEEE Trans. on Signal Processing 52 (2004) 3059–3072.

[16] R. Brunelli, D. Falavigna, Person identification using multiple cues, IEEE Trans. on Pattern Anal. and Machine Intell. 17 (10) (1995) 955–966.

[17] E. S. Bigun, J. Bigun, B. Duc, S. Fischer, Expert conciliation for multi modal person authentication systems by Bayesian statistics, in: J. Bigun, G. Chollet, G. Borgefors (Eds.), Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA, Springer LNCS-1206, 1997, pp. 291–300.

[18] J. Kittler, M. Hatef, R. Duin, J. Matas, On combining classifiers, IEEE Trans. on Pattern Anal. and Machine Intell. 20 (3) (1998) 226–239.

[19] L. Hong, A. K. Jain, Integrating faces and fingerprints for personal identification, IEEE Trans. on Pattern Anal. and Machine Intell. 20 (12) (1998) 1295–1307.

[20] S. Ben-Yacoub, Y. Abdeljaoued, E. Mayoraz, Fusion of face and speech data for person identity verification, IEEE Trans. on Neural Networks 10 (5) (1999) 1065–1074.

[21] V. Chatzis, A. G. Bors, I. Pitas, Multimodal decision-level fusion for person authentication, IEEE Trans. on System, Man, and Cybernetics, part A 29 (6) (1999) 674–680.

[22] P. Verlinde, G. Chollet, M. Acheroy, Multi-modal identity verification using expert fusion, Information Fusion 1 (1) (2000) 17–33.

[23] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero, J. Gonzalez-Rodriguez, A comparative evaluation of fusion strategies for multimodal biometric verification, in: J. Kittler, M. S. Nixon (Eds.), Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA, Springer LNCS-2688, 2003, pp. 830–837.

[24] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Exploiting general knowledge in user-dependent fusion strategies for multimodal biometric verification, in: Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, ICASSP, Vol. 5, 2004, pp. 617–620.

[25] C. H. Lee, Q. Huo, On adaptive decision rules and decision parameter adaptation for automatic speech recognition, Proceedings of the IEEE 88 (8) (2000) 1241–1269.

[26] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Bayesian adaptation for user-dependent multimodal biometric authentication, Pattern Recognition 38 (8) (2005) 1317–1319.

[27] N. Poh, J. Kittler, T. Bourlai, Quality-based score normalization with device qualitative information for multimodal biometric fusion, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 40 (3) (2010) 539–554.

[28] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, Wiley, 2001.

[29] D. A. Reynolds, T. F. Quatieri, R. B. Dunn, Speaker verification using adapted Gaussian Mixture Models, Digital Signal Processing 10 (2000) 19–41.

[30] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, 2000.

[31] S. Theodoridis, K. Koutroumbas, Pattern Recognition, Academic Press, 2003.

[32] A. Navia-Vazquez, F. Perez-Cruz, A. Artes-Rodriguez, A. R. Figueiras-Vidal, Weighted least squares training of support vector classifiers leading to compact and adaptive schemes, IEEE Trans. on Neural Networks 12 (5) (2001) 1047–1059.

[33] J. C. Junqua, G. V. Noord (Eds.), Robustness in Language and Speech Technology, Kluwer Academic Publishers, 2001.

[34] D. Simon-Zorita, J. Ortega-Garcia, J. Fierrez-Aguilar, J. Gonzalez-Rodriguez, Image quality and position variability assessment in minutiae-based fingerprint verification, IEE Proceedings Vision, Image and Signal Processing 150 (6) (2003) 402–408.

[35] C. Wilson, et al., FpVTE2003: Fingerprint Vendor Technology Evaluation 2003, nIST Research Report NISTIR 7123 (http://fpvte.nist.gov/) (June 2004).

[36] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, Quality measures in biometric systems, IEEE Security and Privacy 10 (9) (2012) 52–62. doi:http://dx.doi.org/10.1109/MSP.2011.178.

[37] S. Bengio, C. Marcel, S. Marcel, J. Mariethoz, Confidence measures for multimodal identity verification, Information Fusion 3 (4) (2002) 267–276.

[38] N. Poh, S. Bengio, Improving fusion with margin-derived confidence in biometric authentication tasks, in: Proc. of Intl. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA, Vol. Springer LNCS-3546, 2005, pp. 474–483.

[39] K. A. Toh, W. Y. Yau, E. Lim, L. C. a C. H. Ng, Fusion of auxiliary information for multi-modal biometrics authentication, in: D. Zhang, A. K. Jain (Eds.), Proc. of Intl. Conf. on Biometric Authentication, ICBA, Springer LNCS-3072, 2004, pp. 678–685.

[40] P. Grother, E. Tabassi, Performance of biometric quality measures, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (4) (2007) 531–543.

[41] F. Alonso-Fernandez, J. Fierrez, D. Ramos, J. Gonzalez-Rodriguez, Quality-based conditional processing in multi-biometrics: application to sensor interoperability, IEEE Tansactions on Systems, Man and Cybernetics Part A 40 (6) (2010) 1168–1179.

[42] E. S. Bigun, Risk analysis of catastrophes using experts' judgments: An empirical study on risk analysis of major civil aircraft accidents in Europe, European J. Operational Research 87 (1995) 599–612.

[43] J. Bigun, B. Duc, S. Fischer, A. Makarov, F. Smeraldi, Multi modal person authentication, in: H. Wechsler, et al. (Eds.), NATO-ASI Advanced Study on Face Recogniton, Vol. F-163, Springer, 1997, pp. 26–50.

[44] J. Bigun, J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Multimodal biometric authentication using quality signals in mobile communications, in: Proc. of Intl. Conf. on Image Analysis and Processing, ICIAP, IEEE CS Press, 2003, pp. 2–13.

[45] J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, D. Ramos-Castro, J. Ortega-Garcia, Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems, Forensic Science International 155 (2-3) (2005) 126–140.

[46] J. Daugman, Information theory and the iriscode, IEEE Transactions on Information Forensics and Security 11 (2) (2016) 400–409.

[47] S. Gong, V. N. Boddeti, A. K. Jain, On the capacity of face representation, CoRR abs/1709.10433 (2017) 1–9.
URL `http://arxiv.org/abs/1709.10433`

[48] J. Galbally, M. Martinez-Diaz, J. Fierrez, Aging in biometrics: An experimental analysis on on-line signature, PLOS ONE 8 (7) (2013) e69897.

[49] S. Yoon, A. K. Jain, Longitudinal study of fingerprint recognition, Proceedings of the National Academy of Sciences 112 (28) (2015) 8555–8560.

[50] N. Yager, T. Dunstone, The biometric menagerie, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2) (2010) 220–230.

[51] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M. Freire, J. Gonzalez-Rodriguez, C.Garcia-Mateo, J.-L.Alba-Castro, E.Gonzalez-Agulla, E.Otero-Muras, S.Garcia-Salicetti, L.Allano, B.Ly-Van, B.Dorizzi, J.Kittler, T.Bourlai, N.Poh, F.Deravi, M.Ng, M.Fairhurst, J.Hennebert, A.Humm, M.Tistarelli, L.Brodo, J.Richiardi, A.Drygajlo, H.Ganster, F.M.Sukno, S.-K.Pavani, A.Frangi, L.Akarun, A.Savran, The multi-scenario multi-environment biosecure multimodal database (bmdb), IEEE Trans. on Pattern Analysis and Machine Intelligence 32 (6) (2010) 1097–1111.

[52] B. Rios-Sanchez, M. F. Arriaga-Gomez, J. Guerra-Casanova, D. de Santos-Sierra, I. de Mendizabal-Vazquez, G. Bailador, C. Sanchez-Avila, gb2sumod: A multimodal biometric video database using visible and ir light, Information Fusion 32 (2016) 64 – 79.

[53] L. Sorber, M. V. Barel, L. D. Lathauwer, Structured data fusion, IEEE Journal of Selected Topics in Signal Processing 9 (4) (2015) 586–600.

[54] H. Fronthaler, K. Kollreider, J. Bigun, J. Fierrez, F. Alonso-Fernandez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Fingerprint image quality estimation and its application to multi-algorithm verification, IEEE Trans. on Information Forensics and Security 3 (2) (2008) 331–338.

[55] M. Vatsa, R. Singh, A. Noore, Unification of evidence-theoretic fusion algorithms: A case study in level-2 and level-3 fingerprint features, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 39 (1) (2009) 47–56.

[56] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Speaker verification using adapted user-dependent multilevel fusion, in: Proc. 6th IAPR Intl. Workshop on Multiple Classifier Systems, MCS, Vol. 3541 of LNCS, Springer, 2005, pp. 356–365.

[57] H. Quene, Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo, The Journal of the Acoustical Society of America 123 (2) (2008) 1104–1113.

[58] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, Exploring facial regions in unconstrained scenarios: Experience on icb-rw, IEEE Intelligent Systems (2018) 1–3.

[59] N. Poh, T. Bourlai, J. Kittler, L. Allano, F. Alonso-Fernandez, O. Ambekar, J. Baker, B. Dorizzi, O. Fatukasi, J. Fierrez, H. Ganster, J. Ortega-Garcia, D. Maurer, A. A. Salah, T. Scheidat, C. Vielhauer, Benchmarking quality-dependent and cost-sensitive score-level multimodal biometric fusion algorithms, IEEE Trans on Information Forensics and Security 4 (4) (2009) 849–866.

[60] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia, A. Jain, Incorporating image quality in multi-algorithm fingerprint verification, in: D. Zhang, A. K. Jain (Eds.), Proc. of IAPR Intl. Conf. on Biometrics, ICB, Springer LNCS-3832, 2006, pp. 213–220.

[61] L. I. Kuncheva, C. J. Whitaker, Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy, Machine Learning 51 (2) (2003) 181–207.

[62] P. Viola, M. J. Jones, Robust real-time face detection, Int. J. Comput. Vision 57 (2) (2004) 137–154.

[63] R. Elwell, R. Polikar, Incremental learning of concept drift in nonstationary environments, IEEE Transactions on Neural Networks 22 (10) (2011) 1517–1531.

[64] L. I. Kuncheva, Classifier ensembles for changing environments, in: 5th International Workshop on Multiple Classifier Systems, MCS 04, Vol. 3077 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 1–15.

[65] J. Neves, J. C. Moreno, H. Proenca, Quis-campi: An annotated multibiometrics data feed from surveillance scenarios, IET Biometrics (2018) 1–20.

[66] E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez, F. Alonso-Fernandez, Facial soft biometrics for recognition in the wild: Recent works, annotation and cots evaluation, IEEE Trans. on Information Forensics and Security (2018) 1–12.

[67] O. M. Parkhi, A. Vedaldi, A. Zisserman, Deep face recognition, in: British Machine Vision Conference, 2015.

[68] E. Variani, X. Lei, E. McDermott, I. L. Moreno, J. Gonzalez-Dominguez, Deep neural networks for small footprint text-dependent speaker verification, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 4052–4056.

[69] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, A. Y. Ng, Multimodal deep learning, in: ICML, 2011, pp. 689–696.

[70] P. S. Aleksic, M. Ghodsi, A. H. Michaely, C. Allauzen, K. B. Hall, B. Roark, D. Rybach, P. J. Moreno, Bringing contextual information to google speech recognition, in: INTERSPEECH 2015, 16th Annual Conference of the International Speech Communication Association, Dresden, Germany, September 6-10, 2015, 2015, pp. 468–472.

[71] V. M. Patel, R. Chellappa, D. Chandra, B. Barbello, Continuous user authentication on mobile devices: Recent progress and remaining challenges, IEEE Signal Processing Magazine 33 (4) (2016) 49–61.

[72] J. Fierrez, A. Pozo, M. Martinez-Diaz, J. Galbally, A. Morales, Benchmarking swipe biometrics for mobile authentication, IEEE Trans. on Information Forensics and Security (2018) 1–12.

[73] M. Tistarelli, C. Champod (Eds.), Handbook of Biometrics for Forensic Science, Springer, 2017.

[74] D. Wang, C. Otto, A. K. Jain, Face search at scale, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (6) (2017) 1122–1136.

[75] R. He, B. Lovell, R. Chellappa, A. Jain, Z. Sun, Editorial: Special issue on ubiquitous biometrics, Pattern Recognition 66 (2017) 1–3.

[76] R. Blanco-Gonzalo, R. Sanchez-Reillo, J. Liu-Jimenez, C. Sanchez-Redondo, How to assess user interaction effects in biometric performance, in: 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), 2017.

[77] M. Brockly, S. Elliott, R. Guest, R. Blanco-Gonzalo, Encyclopedia of Biometrics, Springer, 2015, Ch. Human-Biometric Sensor Interaction, pp. 887–893.

[78] J. J. Robertson, R. M. Guest, S. J. Elliott, K. O'Connor, A framework for biometric and interaction performance assessment of automated border control processes, IEEE Transactions on Human-Machine Systems 47 (6) (2017) 983–993.

[79] M. Harbach, A. De Luca, S. Egelman, The anatomy of smartphone unlocking: A field study of android lock screens, in: Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI, 2016, pp. 4806–4817.

[80] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, S. Zafeiriou, End-to-end multimodal emotion recognition using deep neural networks, CoRR abs/1704.08619 (2017) 1–9.
URL http://arxiv.org/abs/1704.08619

[81] J. Garre-Olmo, M. Faundez-Zanuy, K. Lopez-de Ipina, L. Calvo-Perxas, O. Turro-Garriga, Kinematic and pressure features of handwriting and drawing: Preliminary results between patients with mild cognitive impairment, alzheimer disease and healthy controls, Current Alzheimer Research 14 (9) (2017) 960–968.

[82] D. Mellado, E. Fernandez-Medina, M. Piattini, A common criteria based security requirements engineering process for the development of secure information systems, Computer Standards and Interfaces 29 (2) (2007) 244 – 253.

[83] A. Merle, J. Bringer, J. Fierrez, N. Tekampe, Beat: A methodology for common criteria evaluations of biometrics systems, in: Intl. Common Criteria Conf., London, UK, 2015.

[84] A. Hadid, N. Evans, S. Marcel, J. Fierrez, Biometrics systems under spoofing attack: An evaluation methodology and lessons learned, IEEE Signal Processing Magazine 32 (5) (2015) 20–30.

[85] J. Galbally, S. Marcel, J. Fierrez, Image quality assessment for fake biometric detection: Application to iris, fingerprint and face recognition, IEEE Trans. on Image Processing 23 (2) (2014) 710–724.

[86] M. Gomez-Barrero, J. Galbally, J. Fierrez, Efficient software attack to multimodal biometric systems and its application to face and iris fusion, Pattern Recognition Letters 36 (2014) 243–253.

[87] B. Biggio, G. Fumera, G. L. Marcialis, F. Roli, Statistical meta-analysis of presentation attacks for secure multibiometric systems, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (3) (2017) 561–575.

[88] K. Nandakumar, A. K. Jain, Biometric template protection: Bridging the performance gap between theory and practice, IEEE Signal Processing Magazine 32 (5) (2015) 88–100.

[89] M. Gomez-Barrero, J. Galbally, A. Morales, J. Fierrez, Privacy-preserving comparison of variable-length data with application to biometric template protection, IEEE Access 5 (2017) 8606–8619.

[90] M. Gomez-Barrero, E. Maiorana, J. Galbally, P. Campisi, J. Fierrez, Multi-biometric template protection based on homomorphic encryption, Pattern Recognition 67 (2017) 149–163.

[91] L. Nanni, C. Salvatore, A. Cerasa, I. Castiglioni, Combining multiple approaches for the early diagnosis of alzheimer's disease, Pattern Recognition Letters 84 (2016) 259 – 266.

[92] L. Snidaro, J. Garca, J. Llinas, Context-based information fusion: A survey and discussion, Information Fusion 25 (2015) 16 – 31.