HAL
open science

An adaptive multi-biometric incremental fusion strategy in the context of BMEC 2007

Lorène Allano, Sonia Garcia-Salicetti, Bernadette Dorizzi

To cite this version:

Lorène Allano, Sonia Garcia-Salicetti, Bernadette Dorizzi. An adaptive multi-biometric incremental fusion strategy in the context of BMEC 2007. ICARCV 2008 : 10th International Conference on Control, Automation, Robotics and Vision, Dec 2008, Hanoi, Vietnam. pp.1144 - 1149, 10.1109/ICARCV.2008.4795682 . hal-01391747

HAL Id: hal-01391747

https://hal.science/hal-01391747

Submitted on 3 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Adaptive Multi-biometric Incremental Fusion Strategy in the Context of BMEC 2007

Lorene Allano, Sonia Garcia-Salicetti and Bernadette Dorizzi
Institut TELECOM; TELECOM & Management SudParis; Intermedia Team
Evry; FRANCE
{Lorene.Allano; Sonia.Salicetti; Bernadette.Dorizzi}@it-sudparis.eu

***Abstract*—We present the general principles of a dynamical sequential fusion strategy for multibiometric systems allowing reducing the cost associated to the use of different modalities while preserving a good performance compared to that of a global fusion system. We have implemented this strategy in the BMEC (BioSecure Multimodal Evaluation Campaign) 2007 cost-sensitive evaluation. We show that this approach was very robust to the occurrence of missing data thanks to its adaptive implementation.**

***Keywords*—Multi-biometrics, fusion, cost optimization, evaluation, incremental fusion, missing values.**

I. INTRODUCTION

Multi-biometrics has already been the subject of important studies and surveys [1,2]. Fusion of multimodal scores, if performed carefully, provides in all cases an improvement in terms of performance compared to the use of a single modality [3]. Multi-biometrics is also claimed to bring extra advantages, such as robustness to forgeries and to missing values when for some reasons we do not have at disposal the score value of one of the modality at hand. On the other side, multimodality has a cost as additional sensors are needed, both in enrollment and test phases; it also reduces user convenience and is time and complexity consuming. It is therefore interesting to evaluate the merits and limits of multi-biometrics in terms of performance as well as induced cost. This was one of the objectives of the BioSecure Multimodal Evaluation Campaign (BMEC 2007) [4] organized in 2007 by the BioSecure [5] Network of Excellence. This multimodal biometric evaluation campaign was composed of two pre-defined scenarios: access control and mobile scenario. More details on the corresponding Multimodal Database as well as the experimental results of BMEC 2007 can be found at [4].

We will focus in this paper on the access control scenario and more particularly on one of the two proposed evaluation schemes: cost-sensitive score fusion. The aim of this cost-sensitive evaluation scheme was to consider the fusion task as an optimization problem whose goal is to achieve the highest performance with a minimal cost. We refer to "cost" as the price paid for acquiring and processing more information, e.g., requesting more samples from the same device or using more biometric devices (which entails longer processing time).

In this paper we present the system that we submitted to the cost-sensitive evaluation of the access control scenario of BMEC 2007. It is based on a dynamical sequential fusion architecture. The aim of the proposed fusion architecture is to reduce the cost of a multi-biometric system in using the least number of systems as possible. The architecture is sequential because systems are used one after another, and dynamical because the number of systems used depends on each access. In Section 2 of this paper we will describe the methodology we introduced in order to build such architecture. In Section 3, we will present in more details the cost-sensitive evaluation and the implementation of our methodology to this context while our results are exposed in Section 4.

II. DYNAMICAL SEQUENTIAL FUSION STRATEGY TO REDUCE THE COST

In presence of several systems, the most frequent strategy in the literature is to fuse the scores provided by all the systems simultaneously to take the final decision, without making any distinctions on the values of the different scores. At the opposite, the general idea that underlies our methodology is to use the different modalities sequentially. By multi-biometrics we mean using different biometrics (like face and iris for instance) or using different realizations of the same modality (like the scores obtained on the 5 or 10 fingerprints of the same person). Similarly to the multistage approach introduced in [6,7] in the general case of multiple classifiers, we propose to introduce a hierarchy in the different systems leading therefore to classifiers of increasing complexity (a single modality system, then a fusion of two modalities system etc…) as shown in Fig. 1. More precisely, the idea is to use a single system (at the beginning reduced to one single modality, this system will correspond at further steps to the fusion of modalities) as long as this system allows taking a decision that will induce no errors (or very few errors).

As presented in [6], we do not consider the classical "2-classes" decision that is accepting or rejecting a user identity thanks to a unique threshold on the scores of the system. In our strategy, at each stage of our architecture, three decisions are possible: acceptance, rejection or no-decision. This is possible thanks to the use of two different thresholds as shown in Fig. 2 and explained in more details in the following section. In the case of "no-decision" we add an extra modality (and then extra information) to try to take the decision. Our strategy consists in choosing this additional modality in such a way that it is the most complementary to the existing system (i.e. it corresponds to the lowest value of the resulting fusion error).

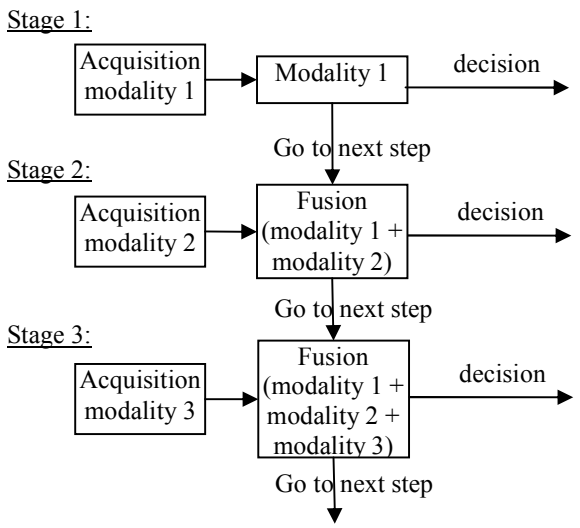

Figure 1. Incremental architecture scheme.

As several modalities are available (in BMEC 2007, there are 8 possible scores for each client access), we must order the different modalities and their fusion in order to optimize the above described methodology. This is done on a development set on which we have at disposal the client and impostor scores for each modality. This way, our sequential fusion strategy is able to dynamically adapt and choose the systems depending on the context and the difficulty to take the decision. It thus allows reducing the cost while still keeping the advantages of multi-biometric systems in terms of performance for difficult cases and also in terms of forgeries detection.

The sequential architecture, described in Fig. 1, is defined by two aspects: a given ordering of monomodal systems and the decision rules at each step.

The incremental system is dynamical because for each access, the number of systems used can be different depending on the step at which the decision is taken. The only thing that is fixed in advance thanks to the development database is the ordering of modalities.

III. TRAINING METHODOLOGY TO BUILD THE SEQUENTIAL FUSION STRATEGY

We will now describe in more detail the two steps of the training methodology we have settled in order to build the system. The available data are scores from each biometric system. In some cases, quality measures can also be available and used in the decision rules.

Step 1: Determining monomodal systems' ordering

The ordering of the monomodal systems in the incremental architecture consists in the choice of the sequence of modalities to be considered.

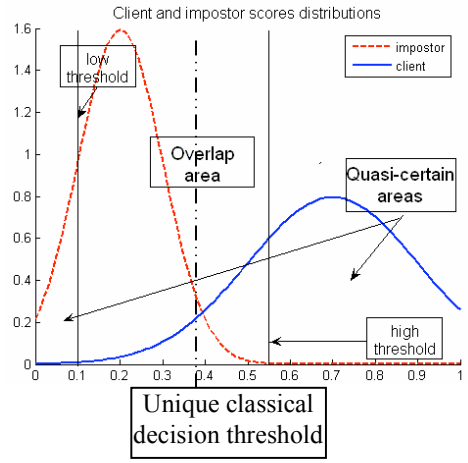

Figure 2. Client and impostor scores distributions.

We select the first system using the following criteria:

- good individual performance
- good complementarity with other systems

To study the complementarity between individual biometric systems, we consider the performance of 2 or several monomodal systems using a simple fusion method: fusion by average of independently normalized scores using MinMax [8]. The less the EER (Equal Error Rate), the more the systems are complementary.

At each step, we choose the next system among the remaining systems using the same two criteria.

Step 2: Determining rules of decision at each step

Once the ordering of monomodal systems is fixed, we determine the rules of "decision" at a given step and the rules of "transition" to the next step.

The rules of decision are defined using high and low thresholds corresponding to quasi-certain areas (see Fig. 2) on:

- the score of the modality acquired at this step and the associated quality measures if available
- the fusion score of all monomodal systems available at this step and the previous steps.

At the first stage (see Fig. 1), in order to define the high and low thresholds associated to the chosen monomodal system according to Step 1 we observe the shape of its client and impostor scores distributions (see Fig. 2) on the development set.

We then determine areas of quasi-certain decision using high (t_h) and low (t_l) thresholds represented by vertical lines on Fig. 2. The 2 quasi-certain areas $]-\infty; t_l]$ and $[t_h;+\infty[$ are determined so that only one class (client, impostor) is present. Those high and low thresholds allow to take a decision by

using only the considered monomodal system (that is without performing fusion) when a score belongs to the 2 previously mentioned intervals.

At stage 2, the chosen monomodal system according to step 1, is analyzed in term of client and impostor distributions to see whether it is possible to take a decision. In that case no fusion is necessary; otherwise we fuse the score of modality and the score of modality 2 in Fig. 1 and study again the client and impostor distributions to determine the two thresholds.

All the rules of the architecture are defined using this kind of double thresholding but in some cases, as explained above, being applied to a monomodal system score and in other cases to a score from a fusion system.

Both the ordering of the monomodal systems and rules are fixed empirically on the development set that should be representative of the population of the application. Rules should not be over fitted to the development set in order to allow a good generalization on an unknown evaluation set.

When evaluating our architecture, we will compare our sequential architecture to a "classical" fusion system using all available monomodal biometric systems. This so-called "reference" fusion system is a fusion method based on Gaussian Mixture Models (GMM) [9]. Client and impostor scores likelihoods are modeled by 2 different GMMs as described in [10]. The posterior probability of the client class is then computed using Bayes rule and used as the final score.

IV. EXPERIMENTS AT BMEC 2007 IN THE COST-SENSITIVE EVALUATION

A. *The cost-sensitive protocol*

As already said, the cost-sensitive evaluation is a score-level fusion evaluation. The objective is to optimize at the same time the multimodal system's performance and its cost. Such cost can be seen as the financial cost of devices used for biometric data capture as well as the processing time to capture and process such data. For each multimodal access, 8 scores are available:

- 1 face score
- 1 iris score
- 6 fingerprint scores for 6 different fingers obtained from the same fingerprint recognition algorithm.

For each score, quality measures are given. Quality measures are available for both the reference image and the test image:

- For face images, 14 quality algorithms give 14 quality measures for the reference image and 14 quality measures for the test image.
- For iris images, 3 quality algorithms give 3 quality measures for the reference image and 3 quality measures for the test image.
- For each of the 6 fingerprint images, 1 quality algorithm gives 1 quality measure for the reference image and 1 quality measure for the test image of each finger.

For this evaluation, the value of the costs has been fixed by the organizers. The global cost is defined by a weight coefficient of 1 for each new biometric modality used (face, iris, fingerprint), a weight of 0.3 for each additional finger in the fingerprint modality. For example, using only one modality corresponds to a cost of 1, using face and iris corresponds to a cost of 2 (1+1), using 2 fingerprint scores corresponds to a cost of 1.3 (1+0.3) and using all scores corresponds to a cost of 4.5 (1+1+1+5*0.3).

For evaluation, it is important to distinguish between two data sets, i.e., the development and the evaluation sets. The development set was released to the participants and is used for algorithm development, e.g., finding the optimal parameters of an algorithm, including setting the global decision threshold. In this evaluation, the development set consists of 51 persons with 2 client scores and 412 impostor scores each. The evaluation set consists of 156 persons with 2 client scores and 504 impostor scores each. For unbiased performance assessment, the population of users in these two data sets is disjoint and the evaluation set remained sequestered by the organizers. More details on the precise protocol can be found on BMEC 2007 web site [4].

In Table 1, the performance of the 8 monomodal systems on the development set and on the evaluation set are presented. On the development set we consider the EER criterion (Equal Error Rate) in order to fix a threshold which will be used on the evaluation set. Then, Performance on the evaluation set is measured by HTER (Half Total Error Rate).

We can observe in Table 1 that the performance of the different systems are surprising compared to habits. Indeed fingerprint and iris have poor performance whereas they are considered as very reliable modalities. Those results are due to the bad quality of the database and particularly to some sensing characteristics. Indeed, fingerprints have been acquired with a thermal sensor that gives bad quality images with distortions compared to those obtained through commonly used optical sensors. Iris images in this database are also of very bad quality due to some defect of the sensors.

B. *Our proposed incremental fusion architecture submitted at BMEC 2007*

Step 1.Ordering of the system

The architecture has been designed on the development set. The ordering of 8 systems in the incremental architecture is: Face, Fingerprint2 (right index), Iris, Fingerprint1 (right thumb), Fingerprint5 (left index), Fingerprint3 (right middle), Fingerprint4 (left thumb) and Fingerprint6 (left middle). The ordering has been defined using each system's performance separately (See Table 1) and also its complementarity with others by considering fusion of 2 or more systems as shown in Table 2 and Table 3.

Face is chosen to be the first biometrics because:

- The face system's performance is among the best on the development set
- Fusion of face with the other modalities gives good performance as shown in Table 2.

TABLE I. PERFORMANCE OF THE 8 MONOMODAL SYSTEMS ON DEVELOPMENT AND EVALUATION SETS

	Face	Iris	Fing1	Fing2	Fing3	Fing4	Fing5	Fing6
EER (%) on the development set	8.98	14.46	17.41	9.13	16.12	19.49	12.5	13.37
HTER (%) on the evaluation set for a priori threshold corresponding to EER on the development set	10.23	18.90	19.18	10.98	18.18	19.25	13.68	17.56

TABLE II. PERFORMANCE OF FUSION OF FACE WITH OTHER SYSTEMS. THE FUSION METHOD USED IS MINMAX NORMALISATION AND AVERAGE OF SCORES.

Fusion of face with	Iris	Fing1	Fing2	Fing3	Fing4	Fing5	Fing6
EER (%) on the development set	4.91	6.83	1.98	3.94	4.69	3.05	2.95

TABLE III. PERFORMANCE OF FUSION OF FACE AND FINGERPRINT 2 WITH OTHER SYSTEMS. THE FUSION METHOD USED IS MINMAX NORMALISATION AND AVERAGE OF SCORES.

Fusion of face and fingerprint 2 with	Iris	Fing1	Fing3	Fing4	Fing5	Fing6
EER (%) on the development set	1.79	1.98	2.71	1.93	2.88	1.97

The second biometrics used is Fingerprint2 (right index) because it has good performance (see Table 1) and good complementarity with face (see Table 2).

When comparing Table 2 and Table 3 we can see that fusing 3 systems is not always better than fusing only 2 systems (face and fingerprint2). In order to have a better performance for fusion, we could have used a complex fusion method as GMM presented in the previous part, or we could have weighted the 3 systems. However, the use of this simple fusion method allows seeing that iris is the most complementary modality with face and fingerprint 2. This is the reason why iris is chosen to be the third biometrics used in the sequential architecture.

Following this methodology, we chose the final ordering of systems used in the sequential architecture submitted to BMEC 2007 cost sensitive evaluation.

Step 2. Definition of decision rules at each step of the incremental architecture

Ordering of modalities has been chosen in order to use first the best systems. We now need to define the decision rules at each step.

At the first step, only one modality is available: face. Quality measures (14 measures) are also available for face system, more details about quality measures can be found in [4]. Fig. 3 presents the distribution of face scores versus one of the face quality measure (quality measure 2) which gives indication on the brightness. Decision on Face consists in high and low decision thresholds for which the face system allows a quasi-certain decision as shown in Fig. 3 with the two vertical lines. The lines correspond to high (t_h=75) and low (t_l=10) thresholds that determine 2 quasi-certain areas $]-\infty; t_l]$ and $[t_h;+\infty[$: scores below 10 and scores higher than 75. In these two areas, there is no overlap between clients and impostors distributions. As explained in Section 2, those two thresholds allow classifying part of the scores using only face.

Also, dashed lines in Fig. 3, illustrate the use of quality measures in the definition of decision rules. For example, for this measure, we can observe that the high threshold (t_h) can be rather set to 55 if the quality measure has a value below 50. This additional rule allows classifying more data using only the face modality.

An example of rule of decision for the first step of our sequential architecture is:

- If face score is $\geq t_h$=75 , then accept
- If face score is $\leq t_l$=10, then reject
- If face score is $\geq t_{hq}$=55 and quality is $\leq$ 50, then accept (quality adapted rule)
- Otherwise no decision is taken and we go to the next step.

At each of the other next steps, two kinds of decision are made: decision on the new modality alone using high and low threshold as for face, and decision on the fusion of all the previous systems. For example at the third step, the fusion of Face, Fingerprint 2 and Iris is also used to take a decision using high and low "quasi-certain" thresholds on the fused score.

At each step, if one or several biometric scores are missing (due to acquisition errors or in score computation by biometric systems), the corresponding rules are not used. If no rules can be used (the conditions are not verified), we directly move to the next step. At the last step, if all scores are available but none of the rules allows taking a decision, we use a "classical" fusion method as follows. If one or several scores are missing, we use the simple MinMax [8] method that independently normalizes each score before taking their average. If all scores are available, we use the "reference" method using GMM-based fusion as described in section III.

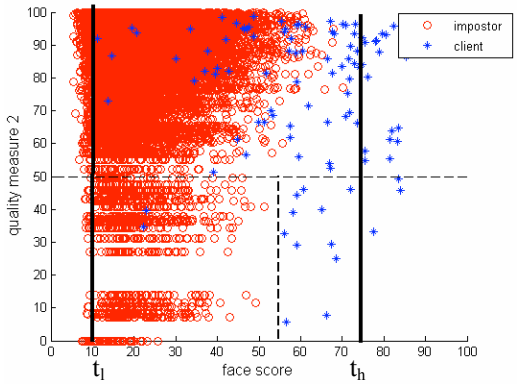

Figure 3. Face score distribution vs. quality measure.

V. RESULTS OF OUR SEQUENTIAL FUSION ARCHITECTURE AT BMEC 2007

In this section we present the results of our sequential fusion strategy at the cost-sensitive evaluation of BMEC 2007. We also compare it to the "classical" fusion method based on GMM and using all the available scores that we also submitted as a point of comparison. We do not compare our system to all the other participant systems. For information on the results and comparison of all the submitted systems the reader should refer to [4].

The cost-sensitive evaluation allows studying two different aspects of multibiometric systems:

- cost minimization of a multibiometric system (reduction of the number of modalities acquired and used)
- estimation of biometric systems robustness to missing values (missing values correspond to impossibility to acquire or to process a biometric sample)

As described above, cost is estimated depending on the number of biometric systems used. Robustness to missing values is studied using 5 different evaluation datasets. Indeed, the organizers have provided 5 different evaluation datasets corresponding to different percentages of missing values. In each case, the organizers have randomly "hidden", respectively 0%, 10%, 20%, 30% and 40% of the scores of each system.

In Fig. 4, each curve corresponds to 5 points associated to the 5 different percentages of missing values (0%, 10%, 20%, 30% and 40%). Only points of each curve corresponding to original data with no artificially missing values (0%) are indicated by a symbol (□, ▲, ►, ▼, ◄, ◊, o, * …). It thus allows reporting the evolution of performance when the percentage of missing data increases (starting from the extremity with the previously mentioned symbols).

Performance on evaluation set corresponds to HTER (Half Total Error Rate) with an a priori threshold computed with the EER (Equal Error Rate) criterion on the development set.

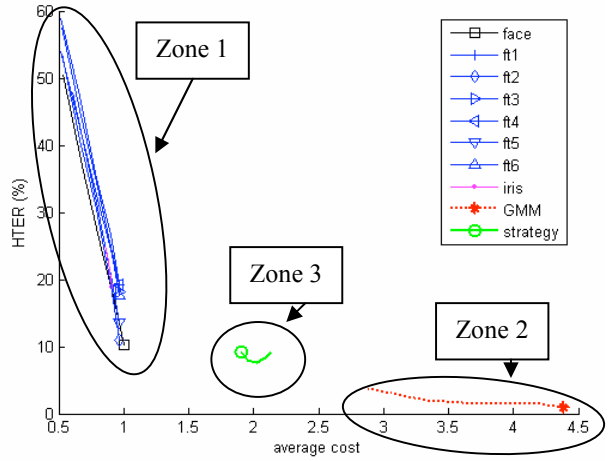

Figure 4. Performance of systems versus cost depending on missing values on evaluation set.

In Fig. 4, we can identify 3 zones:

- Zone 1 corresponds to monomodal systems when being considered individually.
- Zone 2 corresponds to fusion systems using all the monomodal biometric systems.
- Zone 3 corresponds to fusion systems that tend to minimize dynamically the cost as proposed by our architecture.

On the upper left side of Fig. 4, systems in Zone 1 correspond to monomodal systems. As already mentioned in Table 1, monomodal systems' performance varies approximately from 10% to 20% on the evaluation set as shown in Fig. 4. All those systems have a cost equal to 1 when no data are missing and the cost decreases as data is missing. We can also see in Fig. 4, that monomodal systems of Zone 1 are extremely sensitive to missing values as expected. The HTER of the 8 monomodal systems is strongly affected by missing values, going from 10% or 20% to 50% or 60% for the maximum of missing values, when the average cost on all the accesses reaches 0.5.

On the lower right side of Fig. 4, in Zone 2 at the opposite of what we observe in Zone 1, the system has a very good performance but a high cost. The system in this zone is a classical fusion system based on GMM using all the available scores, therefore the cost is maximal but the performance is very good. We can also see that, at the opposite of monomodal systems in Zone 1, system performance in Zone 2 is very slightly degraded by missing values. The HTER of GMM-based fusion system goes from 0.9% to 3.81% when 40% of values are missing. At the same time, the cost decreases (from 4.4 to 2.9) due to missing values.

The objective of the evaluation is to reach the lower left corner of Fig. 4 by having at the same time a good performance and a low cost. Then, we can see that Zone 1 is "better" in case

of no missing values but Zone 2 is “better” in case of missing values.

Zone 3 corresponds to our fusion system using dynamical selection of monomodal systems for fusion. Our system is comparable to monomodal systems when no values are missing with a slightly better performance but a slightly higher cost. We notice that the real interest of those systems is that they are very stable when values are missing, both in terms of performance and cost. The cost is always around 2 and the HTER remains between 8% and 9%. We notice that the global cost increases slightly when the percentage of missing values increases. It is due to the fact that for some accesses, the “strong” modalities (the best modalities in terms of our criteria in Step 1) are missing and then the strategy needs to use “weaker” modalities and therefore more systems.

Depending on the application, the needs in terms of performance and cost are different. Fig. 4 allows choosing the best type of systems for an application considering the three “zones” that appear. Indeed:

Zone 1 is interesting for applications needing a low cost even with higher error rates. Zone 1 can only be used when no or very few values are missing and there are no advantages of using a multimodal system, with regards to enrolment problems or problems of robustness to forgeries.

Zone 2 is interesting for applications requiring a very high performance even at the price of a high cost. It can also be used when data are missing and it keeps the advantages of multimodality (robustness to forgeries, robustness to enrolment problems, …)

Zone 3 is very interesting in applications requiring simultaneously to reduce the cost and to keep a good level of performance, and particularly when data are missing. The sequential architecture, in Zone 3, also has other advantages compared to individual monomodal systems (Zone 1) due to the multibiometric aspect:

- depending on the application, we can change the rules in order to adapt the cost to the context
- multibiometrics also leads to a more robust system to forgeries
- it is particularly suited for situations in which one modality cannot be used (in case of missing values), since it allows considering other modalities in replacement.

It is thus a very flexible architecture, reactive to situations of actual acquisition scenarios.

VI. CONCLUSION

In this article we presented a new incremental and dynamical fusion strategy that we submitted to the BMEC 2007 cost-sensitive evaluation. We have shown, on this rather unrealistic situation, that this strategy allows to get a good level of performance while reducing significantly the cost over methods performing global fusion (fusion of all the 8 available scores). Moreover, this method can be used when data are missing, as it adapts to the context for each client access and it keeps the advantages of multimodality (robustness to forgeries, to enrolment problems, increase of performance). Of course this evaluation is not representative of a possible application (8 modalities for each client access is unrealistic) but this theoretical framework was interesting to stress the interest of our strategy which can be used with profit in more practical situations where reducing the cost in the fusion scheme is an important issue.

REFERENCES

[1] A. Ross, A.K. Jain, “Information Fusion in Biometrics”, Pattern Recognition Letters, vol. 24, N° 1, pp. 2115–2125, 2003.

[2] J. Kittler, M. Hatef, R.P.W. Duin, J. Matas, “On Combining Classifiers”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, N°3, pp. 226-239, March 1998.

[3] M. Indovina, U. Uludag, R. Snelick, A. Mink, A. Jain, “Multimodal Biometric Authentication Methods : A COTS Approach”, in Proc. MMUA 2003, Santa Barbara, California, USA, Dec. 2003, pp. 99-106.

[4] BioSecure Multimodal Evaluation Campaign (BMEC) 2007: http://www.int-evry.fr/biometrics/BMEC2007

[5] BioSecure Network of excellence: http://www.biosecure.info

[6] P. Pudil, J. Novovicoca, S. Blaha and J. Kittler, “Multistage Pattern Recognition with Reject Option”, In: Proc. 11th IAPR Internat. Conf. on Pattern Recognition, The Hague. IEEE Computer Society Press, Los Alamos, pp. 92–95 , 1992.

[7] E Alpaydin, C Kaynak, "Cascading Classifiers," Kybernetika, 1998, 34(4), 369-374.

[8] Arun A. Ross, Karthik Nandakumar and Anil K. Jain, “Handbook of multibiometrics”, Springer, 2006.

[9] D.A. Reynolds and R.C. Rose, "Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models", IEEE Transactions on Speech and Audio Processing, Vol. 3, No. 1, Jan 1995.

[10] L. Allano, A. C. Morris, H. Sellahewa, S. Garcia-Salicetti, J. Koreman, S. Jassim, B. Ly-Van, D. Wu, B. Dorizzi, Non intrusive multi-biometrics on a mobile device: a comparison of fusion techniques. In Proc of SPIE 2006, conference on Biometric Techniques for Human Identification III, Orlando, Florida, USA, 17-21 April, 2006.