Data Normalization and Fusion in Multibiometric Systems

Andrea F. Abate, Michele Nappi, Daniel Riccio
Dipartimento di Matematica e Informatica
Università di Salerno
{abate, mnappi, driccio}@unisa.it

Maria De Marsico
Dipartimento di Informatica
Università degli Studi di Roma "La Sapienza"
demarsico@di.uniroma1.it

Abstract— **Nowadays, the present trend in biometrics is towards multimodal systems. In particular, this work analyzes the combination of the two different biometries face and ear. The twofold contribution is given by the introduction of a new normalization function, namely the *mapping function*, able to overcome the limitations of commonly used functions, and by the definition of the *SRR* (System Response Reliability) index, whose purpose is to estimate the reliability degree of each system response.**

Experimental results in the final part of our work provide a positive feedback about theoretical statements within the body of the paper.

I. Introduction

Biometrics allows recognizing an individual based on physical or behavioural features and its potential is extremely high [2]. Most present systems rely on a single biometry; this makes them vulnerable to possible attacks, as well as little robust with respect to a number of issues. An example is the possible non universality of the chosen biometric feature, as in the case of subjects lacking in a limb for fingerprints recognition or of deaf-mute subjects for voice recognition. A multimodal system provides an effective solution, as flaws of an individual system can be compensated by the availability of a higher number of alternative biometries [6]. Nevertheless, using more biometries in a single recognition system requires to consider some critical aspects. First, each subsystem labels each subject with a numeric value (score) in the range [0, ∞]. However, such values generally come from measurements performed on different features and by using different scales and procedures: a direct combination of such values would give an incorrect result. According to this consideration it is clear the need to normalize the *scores* assigned by each biometry to individual subjects, before combining them using any fusion rule. The normalization module is a very important component in the context of multi-biometric systems [3]. On the other hand, due to the possibly different quality of data inputted from time to time to each subsystem, and to the possibly different accuracy of exploited recognition procedures, it could happen that not all responses are equally reliable. The definition of a measure for the response reliability of the single subsystems would be significant for fusing the single results in an overall final response; such aspect is too often neglected, though especially crucial. The main reason can be found in the considerable difficulty implicit in binding result reliability to input data quality. This would require an absolute metric to estimate the difficulty encountered by the system in identifying a subject, given an acquisition of his/her biometric features. Such metric could not be independent of the adopted classification. Moreover it should consider the relation between data used for enrolment and data used for testing. Finally, it is not possible to directly use an existing performance measure such as the system Recognition Rate, because it provides an estimate of the global ability of the system to recognize a subject, and not an evaluation for the reliability of a single testing operation. Even less it is possible to rely entirely on the single value of the distance between the probe biometric key and the biometric key of the first retrieved subject, as this information alone does not take into account the relation between the latter and all the remaining subjects in the gallery. In other words, similarity among gallery subjects should influence the results evaluation.

In this work we propose a solution to the above two problems by defining both a new *mapping function* normalizing scores from different subsystems, and a *system dependent metric*, namely *System Response Reliability (SRR)*, based on the ability by a recognition procedure of separating genuine subjects from impostors in the most sharp and unequivocal way possible.

II. The Normalization Function

A number of different solutions have been proposed in literature to solve this problem. However, each of them presents some limitation. For example, the Min-max normalization technique performs a “shifting” of the minimum and maximum values in the interval between 0 and 1. Given a set of matching scores $\{S_k\}$, k=1,2,...,n , the corresponding normalized values are given by the (1) in Table 1. Such technique assumes that the minimum and maximum ever generated by a matching module are known.

The Z-score technique is the most widespread and uses arithmetic average and standard deviation of scores returned by the single subsystem. Normalized values are given by (2) in Table 1, where μ represents the arithmetic average of scores and σ is the standard deviation. The problem with Z-score is that it does not guarantee a common interval for normalized values coming from different subsystems.

The Median/MAD technique uses the median and the MAD (median of absolute values); it is quite robust, and the normalized value from the matching module is given by (3) in Table 1, where MAD = median(|s_k – median|). As compared with Z-score normalization, Median/MAD is less effective, most of all when values have a non-Gaussian distribution; in such cases both median and MAD are non-significant estimations and the normalization technique neither preserves the original value distribution nor transforms the values in a common numeric interval.

In this section we propose a normalization function derived from the family of sigmoidal functions. A *sigmoidal* function is defined by the (4) in Table 1. Such function has the open interval (0,1) as codomain, so that it represents a possible mapping function from score values from different biometric subsystems into a single interval; as noticed before, such function would be applied before passing the score values to a fusion module. However we have two drawbacks: a) the distortion introduced by the function when *x* tends to the extremes of the interval is excessive; b) the shape of the function depends on the two parameters *c* and *k* that in turn strongly depend on the domain of *x* parameter. It is possible to reduce the distortion by deriving a new function *F(x)* from *f(x)*, with a pseudo-linear behaviour in the whole codomain though preserving the property such that $F(x) \in [0,1), \forall x$.

Table 1 Some of the normalization functions from the literature.

NORMALIZATION FUNCTIONS		
MIN/MAX	$s'_k = \dfrac{s_k - \min}{\max - \min}$	(1)
Z-SCORE	$s'_k = \dfrac{s_k - \mu}{\sigma}$	(2)
MEDIAN/MAD	$s'_k = \dfrac{s_k - median}{MAD}$	(3)
SIGMOIDAL	$s'_k = \dfrac{1}{1 + ce^{-ks_k}}$	(4)

Given *f(x)* we can find that it has 0 as horizontal asymptote when $x \to -\infty$ and 1 when $x \to \infty$. As noticed above, though *f(x)* shows a pseudo-linear behaviour in its central part, it introduces a non-linear distortion at the extremes, as can be seen in Fig. 1 (blue line). The first step is to compute the extremes of the exploitable region, i.e. the region where the distortion degree introduced by *f(x)* is still sufficiently small. Examining the third derivative of *f(x)*, and in particular the points where it becomes null, it is possible to find out such extremes, which can be used for mapping the scores from the single biometries into the same interval [0,1). The third derivative of *f(x)* is given by:

$$f^3(x) = 6\frac{c^3k^3e^{(-kx)^3}}{(1+ce^{-kx})^4} - 6\frac{c^2k^3e^{(-kx)^2}}{(1+ce^{-kx})^3} + \frac{ck^3e^{(-kx)}}{(1+ce^{-kx})^2}. \qquad (5)$$

It becomes null in two points:

$$xmin = -\frac{1}{k}\log\left(\frac{2+\sqrt{3}}{c}\right) \text{ and } xmax = -\frac{1}{k}\log\left(\frac{2-\sqrt{3}}{c}\right). \qquad (6)$$

Let us assume that, in the general case, the maximum distance value returned by the biometric system is $xmax$, while the minimum value is $xmin$. Let us set up a system of equations using the two solutions and let us solve it with respect to the two variables *c* and *k*; the solutions obtained are:

$$c = \frac{2+\sqrt{3}}{e^{-\frac{xmin \cdot \log(7-4\sqrt{3})}{xmin - xmax}}} \qquad (7)$$

and

$$k = \frac{\left[(xmax - xmin)\log\left(\frac{2-\sqrt{3}}{2+\sqrt{3}}\right) + xmin \cdot \log(7-4\sqrt{3})\right]}{xmax(xmin - xmax)} \qquad (8)$$

In the specific case at hand the values in the domain of $f(x)$ represent the distances between feature vectors belonging to the subjects compared by the different biometries; they then fall in the interval [0, ∞[, so that we can fix $xmin = 0$, so simplifying the obtained solutions:

$$c = 2+\sqrt{3} \qquad (9)$$

and

$$k = -\frac{1}{xmax}\log\left(\frac{2-\sqrt{3}}{2+\sqrt{3}}\right) \qquad (10)$$

We are left with a last problem. The function we are looking for must have the interval [0, 1) as codomain. Let us then define a new function $g(x) = f(x) - f(0)$ and compute its limit when $x \to \infty$:

$$L = \lim_{x\to\infty} g(x) = \frac{2+\sqrt{3}}{3+\sqrt{3}} \qquad (11)$$

Our mapping function is then given by:

$$F(x) = \frac{1}{L}g(x) = \frac{1-b^{\frac{x}{xmax}}}{ab^{\frac{x}{xmax}}+1} \qquad (12)$$

with $a = (2+\sqrt{3})$ and $b = (7-4\sqrt{3})$.

The function *F(x)* assures a pseudo-linear mapping for all values of *x* included in the interval [0, *xmax*]; if we admit some distortion, it also yields to normalized values for x greater than $xmax$, still guaranteeing the constraint $F(x)<1$. This is essential in those biometric systems where the value of $xmax$ is not known in advance and we use an estimate of it. The graph of such function when $xmax$=400 is shown in Fig. 1 (black line).

When the value of $xmax$ is fixed in advance and $x \le xmax$ $\forall x$ it is possible to obtain a better result using the function

$$\overline{F}(x) = \frac{ab+1}{1-b} F(x), \tag{13}$$

as confirmed by the corresponding graph in Fig. 1 (red line).

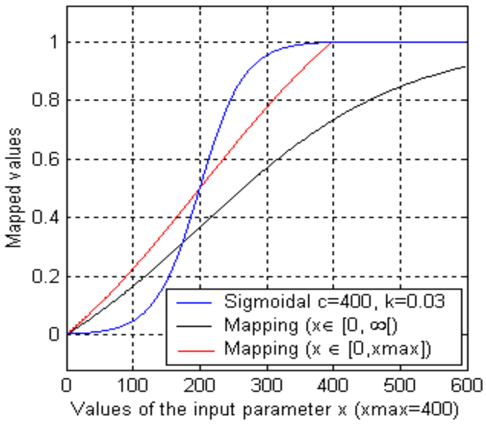

Fig. 1 Graphs of the sigmoidal function, of the mapping function when 0≤x<∞, of the mapping function when 0≤x≤xmax.

III. System Response Reliability

Let A be an identification system and G its gallery of genuine subjects who were correctly enrolled. Assume there are at least n>0 acquisitions for each subject. Moreover, let p be a person to be identified. First of all, the system computes distances between p and all the subjects in the gallery: $d(p, g_i)$ where $i = 1, \ldots, |G|$. Such distances are then ordered increasingly, so that $d(p, g_{i_1}) \le d(p, g_{i_2}) \le \ldots \le d(p, g_{i_{|G|}})$.

We analysed two different measures: the relative distance between the scores of the two first retrieved distinct identities, namely SRR I, and the number of subjects near to the retrieved identity which are present in the gallery, namely SRR II.

We define the relative distance as:

$$\varphi(p) = \frac{\left(F\left(d\left(p, g_{i_2}\right)\right) - F\left(d\left(p, g_{i_1}\right)\right)\right)}{F\left(d\left(p, g_{i_{|G|}}\right)\right)}, \tag{14}$$

where F is a suited data normalization function and g_{i_2} is the second distinct identity in the ordered list of distances. As experimentally observed, the relative difference tends to be small for impostor subjects and high for genuine ones, independently from the biometry and from the classification method. The second measure is computed in the following way. Given the normalized distance $F\left(d\left(p, g_{i_1}\right)\right)$ between the probe and the first identity retrieved from the gallery, the function $\varphi(p)$ is computed using the ratio between the number of subjects in the gallery having a distance from the probe lower than twice $F\left(d\left(p, g_{i_1}\right)\right)$ and the cardinality |G| of the gallery:

$$\varphi(p) = 1 - |N_b| / |G|, \text{ where} \tag{15}$$

$$N_b = \left\{ g_{i_k} \in G \mid F\left(d\left(p, g_{i_k}\right)\right) < 2 \cdot F\left(d\left(p, g_{i_1}\right)\right) \right\}.$$

In both cases, the role of the data normalization function F is to normalize distances in the interval $[0,1[$ such that the function $\varphi(p)$ is independent from the specific function d used for the distance. We need to establish a value φ_k for the reliability index separating genuine subjects from impostor ones. Such value represents the point of maximum uncertainty. It varies depending on the biometric feature and on the classifier, so that it must be estimated from time to time for the single subsystem. The optimal φ_k is given by that value able to minimize the wrong estimates of function $\varphi(p)$, i.e. impostors with $\varphi(p)$ higher than φ_k or genuine subjects with $\varphi(p)$ lower than φ_k. SRR index can then be defined as:

$$SRR = |\varphi(p) - \varphi_k|. \tag{16}$$

Notice that SRR gets high values both for $\varphi(p)$ much higher than φ_k (genuine subjects) and $\varphi(p)$ much lower than φ_k (impostors).

A. How to integrate SRR index into the fusion protocol

A multimodal system is composed of a number of subsystems each working with a specific biometry; as a consequence single subsystems will simultaneously produce a response and a reliability measure for it. The fusion module has then the additional task of carrying out an integration policy able to weight the single responses based on the corresponding reliability degree.

Different choices exist to perform integration, all equally sound in theory, yet possibly taking to very different results in practice. We find it worth then to analyse some of these hypotheses, considering their specific features and postponing to Section 4 their experimental evaluation. Let us assume to have a system S composed by N subsystems $T_1, \ldots, T_N$, each able to produce a list $T_i(1,\ldots,|G|)$ of $|G|$ subjects (where G is the gallery) ordered according to the distance $d_i(p,g_j)$ with $j=1, \ldots, |G|$ from the probe image p, and a numeric value srr_i representing an estimate of the reliability degree of its own response. In order to perform a consistent fusion of data from the single subsystems it is necessary to guarantee the constraint $\sum_i srr_i = 1$ by normalizing the individual indexes using the formula:

$$w_i = \frac{srr_i}{\sum_{j=1}^{N} srr_j}, \ \forall \, i \,. \tag{17}$$

Moreover a consistent threshold th_i is estimated for each subsystem T_i above which we can consider its reliability satisfactory enough. The main integration policies that have been used are OR and AND. For the former, the system considers the combined response as valid only if at least one of its subsystems guarantees its own response with a reliability degree above the corresponding threshold, while for the latter the system considers the combined response as valid only if all of its subsystems guarantee their own response with a reliability degree above the corresponding threshold.

The problem that remains to be solved is to determine the optimal thresholds th_i for the single subsystems. In order to be able to automatically estimate a significant threshold for a given subsystem it is possible to rely on a number of observations. Let us assume that T_i subsystem has executed M times producing as many responses $\{T_i^1(1),\ldots,T_i^M(1)\}$ with their reliability measures $\{srr_i^1,\ldots,srr_i^M\}$ attached; the latter can be considered as the components of a vector $\bar{S}_i = \{srr_i^1,\ldots,srr_i^M\}$ containing the *history* of what happened till a certain moment. We observe that the value to assign to th_i threshold is strictly correlated to the features of vector $\bar{S}_i$, in particular to its average and to its variance. As a matter of fact, if $\mathrm{E}[\bar{S}_i]$ is an high value we can expect that system responses are reliable on the average and that the corresponding threshold is then proportionally high. However, an important role is also played from the variance $\sigma[\bar{S}_i]$ of $\bar{S}_i$ vector, that measures the stability of T_i subsystem; in particular, if the variance has a low value the system can be considered very stable in always giving reliable (unreliable) responses, while an high variance highlights an anomalous behaviour of T_i subsystem. The attractive characteristic searched for a given T_i subsystem is then that its vector $\bar{S}_i$ has a high value for the average and a low value for the variance; in such case it is possible to fix a high value for threshold th_i. We can formally summarize the above observations in the following formula:

$$th_i = \left| \frac{\mathrm{E}[\bar{S}_i]^2 - \sigma[\bar{S}_i]}{\mathrm{E}[\bar{S}_i]} \right|, \tag{18}$$

As a matter of fact such formula represents an indicative criterion to choose the thresholds for the single T_i subsystems, but also a rule for their dynamic update. In the next section we will experimentally show its validity.

IV. Experimental Results

The most used databases were FERET [5] and AR-Faces [4] for the face, and Notre-Dame [7] for the ear. Both for FERET and AR-Faces a subset of the whole database was selected; this is mainly due to the need of building a multimodal database where each subject had corresponding images for face and ear. For this reason FERET subset contains 100 subjects from each of the FAFB, FAFC and DUP I, probe/gallery, i.e. the first 100 subjects (with respect to the original label). In each test only onesubset was used. Such value was imposed because Notre-Dame contains more or less 100 subjects (114 to be precise). The choice of the first 100 labels, i.e. of a well identifiable subset, aims at facilitating the comparison of this study with possible future ones. The subsystems Face and Ear work in parallel and independently [1], while system performances are analysed based on Recognition Rate and Equal Error Rate (EER).

We start from describing experimental results related to the *mapping function*. The main goal of a normalization function is to map values from its domain to its co-domain [0,1], preserving the original input data distribution as much as possible. In order to evaluate the contribution provided by the new proposed normalization function, we studied the distortion degree introduced by the functions in Section 2, by comparing their performances when input data features are known in advance. In more detail, we chose the two following test functions: $f_1(x) = 2 \cdot (\cos(x) + 1)$ and $f_2(x) = 2 \cdot \log(x+1) \cdot (\cos(x) + 1)$ in [0, 2π] interval and applied the given normalization functions to both of them. In Fig. 2 a graphical representation is shown of how the different normalization functions discussed in Section 2 act upon the values of our two test functions, while in Table 2 numerical performances of the whole systems are reported. When minimum and maximum values are known, the normalization process is trivial; for this reason, we assumed to miss an exact estimate of the maximum value for the two functions $f_1(x)$ and $f_2(x)$ and we chose the average value in its place, in order to stress normalization functions even more. From Fig. 2 it comes out that *Min-Max*, *Z-scores* e *Median/MAD* do not comply with the constraint of having normalized valued fall within the [0,1] interval, moreover the last two also produce negative values besides introducing a sharp distortion in the original distribution. The *Sigmoidal* function guarantees a limited co-domain]0,1[. However it introduces an excessive distortion that is particularly evident for input values near to zero. Finally, our Mapping Function provides performances which are better when compared to all the preceding ones. It assures to have normalized minimum and maximum values within the desired range and also preserves input data distribution in a satisfying way, though relying on a maximum estimate which are far from the real value.

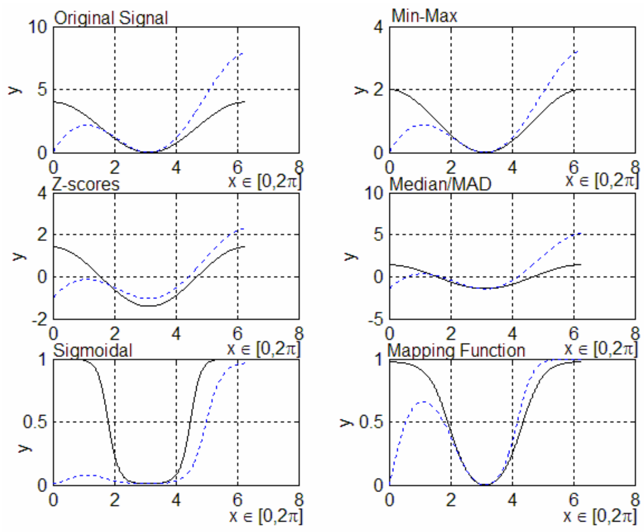

Fig. 2 The distortion introduced by the different normalization functions.

Table 2 Comparison between the performance of biometric systems for different normalization functions.

SYSTEM		PERFORMANCES				
		MIN MAX	Z SCORES	MEDIAN MAD	SIGM.	**MAPPING FUNCTION**
FACE	RR	93%	93%	93%	93%	**93%**
	EER	0.03	0.23	0.12	0.04	**0.03**
EAR	RR	72%	72%	72%	72%	**72%**
	EER	0.14	0.25	0.17	0.16	**0.14**
Face ⊕ *Ear*	RR	95%	93%	93%	94%	**98%**
	EER	0.018	0.23	0.11	0.02	**0.015**
	NRR	100	100	100	97	**90**

For the multibiometric architecture Face⊕Ear in Table 2, the subsystems have been combined using OR policy and SRR II reliability index. Results in Table 2 underline that the Face/Ear Recognition Rate remains unchanged across different methods, while the Equal Error Rate undergoes to significant variations, this is likely due to an increment of the FAR. The worst performance is given by the Z-scores approach that provides an EER above 0.22 for both single subsystems and the multi-biometric architecture. On the contrary, the proposed mapping function equals the optimal performances of the Min/Max approach, yet providing a high robustness with respect to the estimation of the maximum score. Indeed, while the performance of the Min/Max method degrades with inaccurate estimations of the maximum score, the mapping function still provides similar RR/EER values for large variations of the *xmax* parameter (see Section II). An example of how the effectiveness of SRR II index for face varies together with the value chosen for φ_k is given in the first graphic in Fig. 3; an asterisk underlines the point where the index precision is maximum. The second graphic in Fig. 3 shows which probes among the full set of 100 were correctly recognized (circle) and which were incorrectly identified (square). We can observe that squares are concentrated on the lower part of the graphic, underlining that $\varphi(p)$ function gets a low value, while circles are mostly located in the upper part, indicating a higher value of such function.

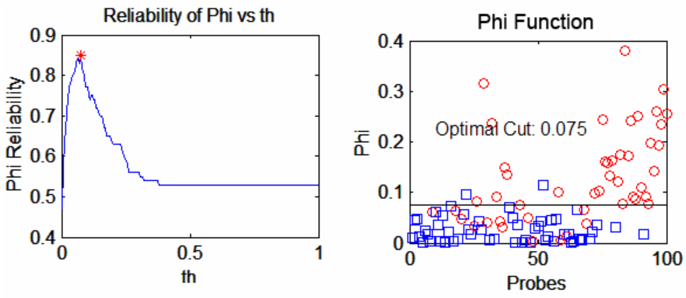

Fig. 3 Experimental evaluation of SRR index.

It is also important to verify the validity of the criterion for choosing the single th_i thresholds. To do this a comparison was performed between estimated thresholds (formula (18)) and optimal ones, the latter being computed by maximizing the number of system responses consistent with the value of SRR index. In this case the selected fusion rule is the AND, and the reliability index used is SRR II. For both face and ear, estimated thresholds are very close to the optimal ones, for all test sets (e.g.: EAR: 0.06/0.07, FAFB 0.6/0.5, DUP I 0.4/0.5, AR-LEFT LIGHT 0.5/0.7, SUNGLASSES 0.6/0.7). We tested both SRR indicators introduced in Section III, in combination with two different integration policies based on them, for data coming from single subsystems. Experiments were performed for the three different datasets extracted from FERET, combined with Notre-Dame ears database. The single facial recognition subsystem assures, on three different databases (FAFB, FAFC, DUP I), a Recognition rate of respectively 93%, 16% and 47%, and an Equal Error Rate of 0.03, 0.29, 0.19; the ear recognition module on Notre-Dame database provides a Recognition Rate of 72% and an Equal Error Rate of 0.15. Table 3 reports results obtained from the multimodal system Face/Ear using each of the integration policies (OR, AND) and varying the face database; the results are presented in terms of Recognition Rate (RR), Equal Error Rate (EER) and Number of Reliable Responses (NRR). Table 3 shows that on the whole all the integration policies guarantee an higher Recognition Rate than the one coming from face or ear (87% for DUP I with OR/SRR II instead of 47% e 72%), and at the same time considerably reduce the Equal Error Rate (0.015 for FAFB with OR/SRR II instead of 0.030 and 0.150), that in some cases even cancels out. The counterbalance of such recognition performance increment is the number of responses that are invalidated because considered as unreliable. However, it is worth underline that an unreliable response has not to be necessarily discarded or considered as wrong, but could represent a valid reason to perform a further check. This could even be more expensive, like for example fingerprints; as such check would be performed only on a limited number of cases, the average computational cost of the multimodal system would be reduced anyway, with the advantage of an higher security degree. The first column shows that all responses are considered as reliable when fusion is performed without any SRR index. However, both RR and EER are significantly worse, showing its ability to discard little reliable responses.

Table 3 Comparison between the performance of the fusion rules.

Database		Statistics				
		None	SRR I		SRR II	
		Or	Or	And	Or	And
FAFB	RR	98%	99%	100%	96%	100%
	EER	0.028	0.016	0.003	0.015	0.000
	NRR	100	75	63	94	38
FAFC	RR	55%	76%	100%	84%	-
	EER	0.167	0.153	0.002	0.117	-
	NRR	100	85	2	74	0
Dup I	RR	75%	81%	100%	87%	100%
	EER	0.238	0.228	0.001	0.177	0.000
	NRR	100	91	18	84	22

From a quantitative point of view, we can observe from that OR always assures a higher number of responses considered as reliable, as expected. Finally, the two reliability indexes SRR I and SRR II have been tested on subsets extracted from AR-Faces database, combining face and ear using AND policy. The results of such experiment are shown in Table 4.

Table 4 Comparison between the performance of SRR I and SRR II.

Distortions on the Face		Statistics				
		Face	Ear		Face ⊕ Ear	
					SRR I	SRR II
Left Light	RR	83%	72%	RR	100%	100%
	EER	0.11	0.12	EER	0.001	0.000
	REL	93%	90%	NRR	37	30
Sad	RR	95%	72%	RR	100%	100%
	EER	0.07	0.12	EER	0.009	0.010
	REL	100%	90%	NRR	86	86
Scarf	RR	80%	72%	RR	100%	100%
	EER	0.17	0.12	EER	0.000	0.004
	REL	83%	90%	NRR	20	47
Scream	RR	47%	72%	RR	100%	100%
	EER	0.18	0.12	EER	0.002	0.000
	REL	83%	90%	NRR	19	10
Sun glasses	RR	87%	72%	RR	100%	100%
	EER	0.10	0.12	EER	0.020	0.020
	REL	87%	90%	NRR	87	77

The single subsystems performances in terms of Recognition Rate, Equal Error Rate and Reliability are reported in the first two columns of the table. As the database used for the ear is always the same, data are repeated on each row. The results obtained using the combined system with AND policy are reported in the following two columns; the reliability index which was used is on top of the column. Results in Table 4 further enforce our considerations. As a matter of fact, we can observe that the combined system guarantees in all cases better performances than single subsystems (in some cases like *scarf* we pass from 47% and 72% to 100% as for Recognition Rate). However, when the single subsystems are particularly stressed by input quality, the number of responses considered as reliable significantly decreases. This confirms the need for more biometries to further strengthen the system.

V. Conclusions

The paper faces crucial problems for a multimodal system, namely data normalization and system response reliability. In particular, we defined a new data normalization function, the *mapping function*, able to overcome the limits of the presently used ones. We also proposed two different reliability indices, providing the conditions for their actual exploitability within the single subsystems, such as the *a priori* identification of thresholds.

All the theoretical concepts introduced in the first part of this work are then validated in the section showing experimental data, through a number of tests performed on different databases chosen among those commonly used.

Future developments can certainly regard the extension to a higher number of either biometries or fusion models, so that a higher number of subsystems could be involved in our analysis.

References

[1] A. F. Abate, M. Nappi, D. Riccio, Face and Ear: A Bimodal Identification System, Proceedings of the Int. Conf. On Image Analysis and Recognition ICIAR 2006, Póvoa de Varzim, Portugal, September 2006.

[2] A. K. Jain, A. Ross, and S. Prabhakar, An introduction to biometric recognition, IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, pp. 4–20, Jan 2004.

[3] A. K. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, in Pattern recognition, Vol. 38, No. 2, pp. 2270-2285, 2005

[4] A. M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class", in *IEEE Transans. on Pattern Analysis and Machine Intelligence*, Vol.24, No.6, pp. 748–763, June 2002.

[5] J. P. Phillips, H. Moon, A. S. Rizvi and P. J. Rauss, The FERET Evaluation Methodology for Face-Recognition Algorithms, in IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.22, No.10, pp.1090–1104, 2000.

[6] A.Ross, A. K. Jain, Multimodal Biometrics: An Overview, in Proceedings. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.

[7] Notre Dame Ear Database, http://www.nd.edu/~cvrl/UNDBiometricsDatabase.html, 2007-01-07.