A Biometric Authentication System Based on Face Recognition and RFID tags

F. Battaglia, G. Iannizzotto, L. Lo Bello

Abstract. *Authentication systems usually adopt either the conventional identifier-password paradigm or different kinds of tokens (e.g., badges, keys). However, passwords can be disclosed while being input and tokens can be stolen and used by impostors. As a result, in the last decades biometric techniques were developed to identify a user through physiological features that cannot be stolen or counterfeited. However, even those techniques have their flaws, and for this reason recent research addressed the combination of multiple identification factors. In this context, this work proposes VisilabFaceRec, a multi factor authentication system based on the combination of a dual-stage cascading classifier for biometric identification (face recognition) with an encrypted RFID tag for token-based authentication. Unlike other approaches in the literature that propose a centralized database for storing biometric data, with serious risks regarding user privacy, our work avoids a centralized database and stores sensitive data in the RFID, thus also making the system performance independent of the total number of subjects enrolled. The proposed architecture is able to simultaneously minimize the False Acceptance Rate and the False Rejection Rate, thanks to an innovative approach for the calculation of the decision thresholds for the two discriminators. VisilabFaceRec has been realized on a commercial board for embedded computing and proven to be able to run in near real-time. The paper describes the system architecture and the algorithm used to jointly determine the couple of decision thresholds for the cascading classifiers, and proposes a performance evaluation, in terms of both accuracy and speed, on a well-known and publicly available face database.*

Keywords: Biometrics, security, authentication, RFID tags.

1. Introduction

Access control (or authentication) mechanisms aim to guarantee only authorized users the access to a given resource or service at any time, while blocking impostors.

The simplest authentication mechanism is password protection [Jain et al, 2006]. However, the need for systematically typing a password to obtain access to resources or restricted areas (i.e., buildings, offices, devices) is a tedious and inefficient operation, that is also prone to serious security flaws, due to the possibility that the password be either read by someone else (e.g., while it is being typed) or involuntarily disclosed.

Although tokens (such as, badges or keys) offer effective security features (e.g., encryption, public/private keys, etc.), their use is not the ultimate security solution either, as a token is not exclusively bound to its owner and therefore could be used by an impostor.

For this reason, a significant effort has been made over the years to develop authentication systems based on biometric data, which offer the advantage of being always available at the very place where the user to be authorized is while being quite difficult to counterfeit.

Although several features can be used for biometric recognition (e.g., digital footprint, iris, hand shape, voice), face recognition offers multiple advantages. First, face recognition does not require expensive sensors. Second, there is no need for physical contact between user and sensor. Finally, it can be used also for video-surveillance and in the non-decisive phases of judicial investigation.

However, when authentication techniques are not combined with other approaches, they suffer from the limitation of requiring a database, either centralized or distributed, to maintain the biometric data of all the users to be authenticated, as such a database allows performing authentication through the comparison between the biometric characteristics of the user and those that are stored in the database. Moreover, it has not been proven yet that automated face recognition is able to achieve 100% accuracy for an arbitrarily large database and this fact raises doubts about the adoption of this technology.

An approach that is being broadly investigated nowadays is Multi Factor Authentication, which combines multiple techniques to obtain more reliable results.

For instance, in the past some solutions that integrate RFID tags (i.e., radio-frequency tokens) with face recognition systems were proposed [Min et al, 2011] [Jing et al, 2009] [Nguyen et al, 2012]. However, these solutions maintain the set of poses of the authorized users in a centralized database and store in the RFID tags only very few data, such as, the user identifier (declared identity).

Furthermore, these methods require, for a single claimant, a large number of poses to be acquired at high resolution under different lighting conditions and stored in the database [Nguyen et al, 2012].

These approaches raise problems, not only due to the database size, but also due to legal issues: The authorities, both at a national and supranational level,

have stated several times that biometric data are personal data, therefore it is in general not advisable to maintain them in a centralized database run by the service provider, unless there exist important and proven security reasons for doing so (principle of finality, necessity, proportionality). In addition, a clear preference towards *serverless* solutions, in which the biometric data are stored on a chip that the authorized user is able to take with them, was given [Art. 29 WP, 2003].

Authentication systems like the one proposed in this paper, i.e., combining RFID tags with face recognition and based on a serverless architecture, are quite innovative compared to the current state of the art. For instance, in [Meng et al, 2010], the authors propose an embedded system that stores in the RFID tag only the set of the *n* principal decomposition components (PCA) [Pentland et al., 1991] associated to the owner's face. However, *for both the enrolment and authentication stages* the PCA representation requires the availability, local to the authentication system, of a set of images (*Eigenfaces*) which depend on the whole of the images of the enrolled subjects. Moreover, every time a new subject is enrolled, the set of Eigenfaces changes and therefore must be recalculated. Unfortunately, when the set of Eigenfaces changes, also the set of principal components of each enrolled user changes, therefore all RFID tags must be redistributed. The paper does not clearly state where the Eigenfaces are stored (either in the tag or in the local memory of the authentication device) and what happens when a new subject is enrolled.

Furthermore, none of previously cited works was tested for robustness against the intrusion of impostors by simulating a significant number of illegal access attempts. The systems were optimized for recognition accuracy or processing speed (for example, in [Jing et al, 2009] the system was tested on three subjects only and an average false acceptance rate of 3.89% was obtained).

The system presented in this paper, called *VisilabFaceRec,* integrates RFID recognition technology with a face recognition system, in such a way that the resulting system provides the following properties:

- The user biometric data are encrypted and stored in the RFID badge; therefore, there is no need for centralized databases or for arbitrarily large databases.
- The algorithms adopted allow for high accuracy, although the amount of stored data is compliant with the capacity of a commercial RFID. This significantly reduces costs and architectural complexity compared to other solutions based on centralized databases.
- The system is fully scalable, as the operations of adding or deleting a subject in the list of the enrolled users do not require any recalculation or the redistribution of the tags.

The main contribution of this work is therefore an authentication system that, combining biometric and RFID authentication, does not need a centralized database, thus avoiding any problems related to privacy issues. The proposed system obtains good results, and its performance in terms of accuracy is independent of the number of subjects to be authenticated. This solves a

significant problem found in other approaches in which the reliability of the results decreases when the number of subjects grows. Finally, the system works on very low resolution face images (e.g. 40x30 pixels), so the image acquisition can be performed by simple, cheap, off-the-shelf cameras and the communication with the RFID tag is reasonably fast.

As will be shown in the following, the proposed authentication system is based on a *two-cascaded discriminator* stage that is able to realize simultaneously the minimization of the False Acceptance Rate (proportion of impostors accepted) and of the False Rejection Rate (proportion of genuine claimants rejected), thanks to an innovative approach for the calculation of the decision thresholds for the two discriminators.

The proposed system also offers, in addition to the already listed advantages, the appealing possibility of being implemented on embedded devices at relatively low cost.

The paper is organized as follows. Sect. 2 describes the VisilabFaceRec architecture, discussing the rationale behind the design choices and providing details on the mathematical formulation of the approach devised for determining the optimal thresholds for the two cascaded discriminators. Sect. 3 presents experimental results obtained by testing the performance of VisilabFaceRec using a well-known publicly available face database and evidences the accuracy and speed of the system. Finally, Sect. 4 provides conclusions and directions for future work.

2. VisilabFaceRec

Before a subject can be authenticated, they must be enrolled, i.e. a set of images of their face must be recorded together with their identity information. In VisilabFaceRec the images of the enrolled subject are saved in an RFID tag and the enrollment step is a one-time process.

In the authentication step, when a subject to be authenticated approaches the camera, a sequence of images of their face is acquired and, at the same time, the content of the RFID tag is read. The acquired images are processed one at a time. If the similarity level between at least one of the acquired faces and the images retrieved from the RFID is sufficient, the subject is authenticated.

The adopted approach (see Fig. 1) is composed of two cascaded authentication stages, optimized for low-resolution images which can be saved in the small memory available in commercially available RFID tags (4-32 KB). The first stage is based on the 2D Principal Components Analysis (2DPCA) algorithm [Yang et al, 2004] and operates a first discrimination of the input images. The second stage operates only on the images which pass the first stage and exploits a *Scalar Image Feature Transform* (SIFT) [Lowe, 2004] to produce a final selection.

2DPCA is a template-matching algorithm originally proposed for face recognition. Given a collection B of *m* face images (poses) belonging to *p* different *known subjects* (classes) with *k* poses for each class, 2DPCA receives as an input a pose belonging to an *unknown subject* and selects from the collection the pose most similar to it, thus identifying the claimant.

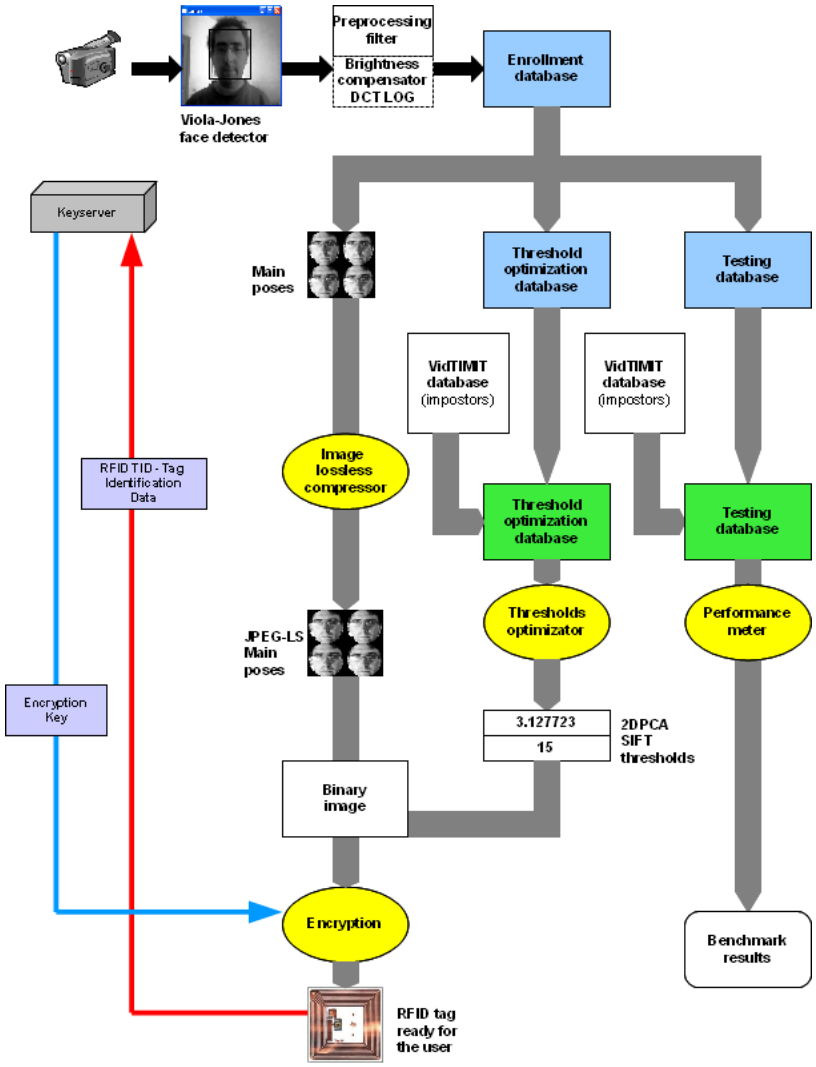

Figure 1
Schematization of the VisilabFaceRec enrolling process

Within 2DPCA an image I of size *w*h* pixels can be decomposed into a set of z_{2DPCA} <=w vector components (called *decomposition vectors*):

$$(\mathbf{w}_0, \mathbf{w}_1, ..., \mathbf{w}_{z_{2DPCA}-1}) \qquad z_{2DPCA} \leq w \quad \mathbf{w}_s \in R^h \tag{1}$$

where the number of components z_{2DPCA} is chosen according to the fraction of the original information associated to the image, which we want to retain.

We therefore can decompose each pose in the collection B of known subjects and the image I of an unknown subject, and then select the pose in B which is most similar to I by minimizing the *distance in the feature space* (DIFS) defined in (2):

$$DIFS(\mathbf{I},B) = d(\mathbf{I},\mathbf{W_i}) = \sum_{s=0}^{z_{2DPCA}-1} \left\| \mathbf{w}_s - (\mathbf{w}_i)_s \right\|_2 \qquad i \in \{0...(m-1)\} \qquad (2)$$

where $\mathbf{W}_i$ is the i-th pose from B, $\mathbf{w}_s$ is the s-th component vector of the unknown pose I, $(\mathbf{w}_i)_s$ is the s-th component vector of $\mathbf{W}_i$ and *m=p*k*.

The 2DPCA decomposition of the image I in respect to the collection B is obtained by multiplying the matrix A containing the pixel intensities of I by each one of the z_{2DPCA}<=w main eigenvectors of the covariance matrix $\mathbf{G}_t$ (of size w*w**)** defined in (3):

$$\mathbf{G}_t = \frac{1}{m}\sum_{j=0}^{m-1}(\mathbf{A}_j - \overline{\mathbf{A}})^T(\mathbf{A}_j - \overline{\mathbf{A}}) \qquad (3)$$

where $\mathbf{A}_j$ is the matrix containing the pixel intensities of the j-th pose from the collection B.

In our approach B contains only the poses of the enrolled subject (p=1 and therefore m=k), B is named *main poses database* and it is possible to verify if a newly acquired image I belongs to the same subject by simply applying a decision threshold ρ to the distance DIFS defined by (2). However, in single threshold systems, False Acceptance Rate (FAR) and False Rejection Rate (FRR) cannot be minimized at the same time [Flach, 2003]. The decision threshold ρ must therefore be determined as a trade-off between the two minimization objectives. Moreover, due to (3) depends on the poses of the collection B and therefore is specific for each collection. In Section 2.1 we describe a solution for the two problems introduced above.

The second stage of our authentication algorithm is based on the SIFT algorithm, which belongs to the class of feature-matching algorithms. When applied to a couple of images, it determines a set of *keypoints* in both images and applies a robust and reliable criterion to establish a correspondence between couples of matching keypoints from the two images [Lowe, 2004]. For each acquired image which passes the 2DPCA stage, and the corresponding most similar pose from the collection B, the SIFT stage determines the number *s* of matching SIFT features and, if this is higher than a decision threshold $\mathbf{o}_{SIFT}$, the acquired image passes also the second stage and the unknown subject is authenticated. Otherwise, the subject is rejected.

2.1. The Enrollment phase

A new subject is enrolled by acquiring $n_{enrolldb}$ images of their face (poses). Each pose is then decomposed according to 2DPCA and the resulting representations are clusterized by applying a K-means algorithm using the distance (2). For each one of the $n_{main.poses}$ clusters obtained, the pose closest to the centroid of the cluster is selected, thus creating a set of *main representative*

poses of the subject. Those poses compose the collection B for a specific subject and are compressed with a lossless algorithm (LS-JPEG) and saved in the RFID of the subject.

Unlike other solutions proposed in literature, our approach uses several poses for each subject, coded at 256 gray levels. We save in the tag the images and not their 2DPCA representations (1) because saving a number z_{2DPCA} of floating point components suitable for our authentication purposes would take more space than the original 8 bit gray level images. We therefore prefer to recalculate on-the-fly the covariance matrix (3) and the corresponding eigenvectors every time the RFID tag is read. As shown in Section 3, the time required by such recalculation does not affect severely the execution time of the whole authentication process.

As stated earlier, VisilabFaceRec uses two cascaded authentication stages which each require a custom threshold for each collection of representative poses. We therefore need a couple of thresholds (ρ_{2DPCA}, σ_{SIFT}) for each subject. Such thresholds are saved together with the main representative poses of the subject in the RFID tag. Determining such thresholds is not a trivial task because the outcome of the 2DPCA stage affects the behavior of the SIFT stage. In the following we describe how we determine the two thresholds and in Section 3 we show some relevant experimental results which support the suitability of the described solution.

We define the *threshold optimization database* (TOdb) composed of p_{TOdb} rows and k_{TOdb} columns. The first row is composed of the $k_{TOdb}= n_{enrolldb} - n_{main.poses}$ poses of the enrolled user. The other rows contain k_{TOdb} poses of (p_{TOdb}-1) *impostors*, i.e. people not belonging to the set of enrolled users. In our case we took the impostors from the VidTIMIT public domain database [Sanderson, 2002]. We calculate the covariance matrix (3) from the set of main representative poses of the subject being enrolled and apply the 2DPCA authentication stage, *using the calculated matrix*, to the poses of the TOdb. For each pose of TOdb we calculate the DIFS distance (2) from the set of main representative poses of the subject. Finally, we assign to d_{max} the maximum distance value among all the calculated distances. We can then draw the *Primary Receiver Operating Characteristic* curve (ROC). On a (FRR,FAR) plane, the primary ROC shows the behavior of False Rejection Rate (FRR) and False Acceptance Rate (FAR) for the case where we select a given value of ρ(ρ in [0..d_{max}]) as a threshold for the first authentication stage (note that we always have FRR(ρ=d_{max})=0 and FAR(ρ=0)=0). The primary ROC is a parametric curve where each value of ρ corresponds to a point P and vice-versa.

We therefore choose as a threshold for the first stage the value ρ_{2DPCA} which minimizes the following cost function (4):

$$C(\rho) = C_{FRR} FRR(\rho) + C_{FAR} FAR(\rho) \qquad (4)$$

where C_{FAR} and C_{FRR} are the costs heuristically assigned, respectively, to the case of false acceptance and false rejection (see Section 3 for an example of how they can be assigned). In Fig. 2, on the right, the calculated ρ corresponds to the point $P(\rho_{2DPCA})$.

So far, we have determined the threshold for the first stage and still need to determine the corresponding threshold for the second authentication stage.

For each pose of the TOdb which passes the first stage (i.e. for which $DIFS(I_n,B)<\rho_{2DPCA}$) we consider the corresponding most similar pose (according to the previous stage) in the set of main representative poses of the subject. We then calculate the number *s* of matching SIFT features for each couple so determined and assign to s_{max} the maximum value of *s* over the whole set of poses analyzed. Now we can draw a second graph, named *secondary ROC*, which shows the behavior of FRR and FAR of the aggregated system (i.e. the system composed of the two cascaded authentication stages) when the threshold σ of the second stage varies in the interval $[0..s_{max}]$. In other words, the secondary ROC is a parametric curve describing the performance of the aggregated system when σ varies from 0 to the maximum number of features calculated according to threshold optimization database (note that a pose is accepted by the second stage if the number of detected matching features is *higher* than σ).

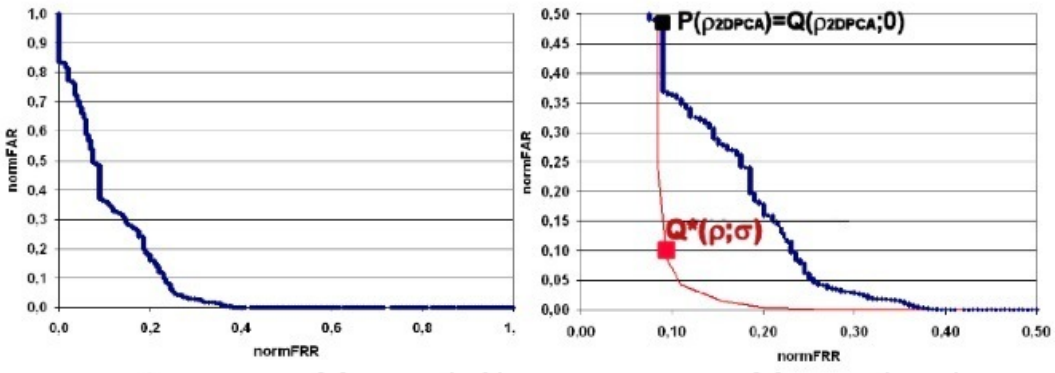

Figure 2
Primary ROC curve (left) and secondary ROC curve (right)

Note that when σ=0, the SIFT stage is "*turned off*" and all images pass through it, so FAR and FRR for this case are the same as those of a single stage 2DPCA system. Therefore, the secondary ROC always intersects the primary ROC in the point $Q(\rho_{2DPCA};0)=P(\rho_{2DPCA})$.

Finally, in order to minimize at the same time both FAR and FRR we choose for s the value which minimizes the function (5):

$$\sqrt{[FRR^*(\rho_{2DPCA};\sigma)]^2+[FAR^*(\rho_{2DPCA};\sigma)]^2} \qquad \sigma \in [0..s_{\max}] \qquad (5)$$

where FRR* and FAR* are FRR and FAR normalized in the interval [0..1], respectively. The corresponding point in Fig. 2 is $Q^*(\rho_{2DPCA}; \sigma_{SIFT})$. The couple $(\rho_{2DPCA}, \sigma_{SIFT})$ so determined is saved in the RFID tag together with the collection of main representative poses of the subject. The data is AES-128 encrypted for security reasons and the encryption key is different for each tag. Indeed, should we use the same key for all the tags, an impostor might find the correct value of the key and create a false RFID tag with their poses and arbitrary identity data (*spoofing attack*). Instead, in VisilabFaceRec every tag is encrypted with a different key $K(x_{EPC}, x_{ID})$ which is generated during the enrolment phase and associated unambiguously to the two codes x_{EPC} and x_{ID}. The first code, x_{EPC}, is determined by the Tag Identification Data, which is unique to each RFID tag, and the second code is the identification code assigned to the enrolled subject by the issuing agency. The association between the encryption key K and the couple (x_{EPC}, x_{ID}) is maintained by a centralized keyserver which does not introduce privacy issues because it does not contain biometric data.

During the authentication phase, the TID data and the x_{ID} are read from the tag and the encryption key is retrieved from the key server. Note that an impostor would not be able to retrieve the correct encryption key from the server because their (x_{EPC}, x_{ID}) couple would not correspond to any entry in the keys database.

2.2. The authentication phase

The authentication algorithm is depicted in Fig. 3.

During the authentication phase a sequence of images of the unknown subject is acquired. In each image, the region containing the face of the subject is detected through a face detection algorithm [Viola and Jones, 2001], cropped and scaled in order to normalize its size. The normalized face region is then compensated for brightness [Chen et al, 2006]. The sequence of images is then passed to the two cascaded authentication stages. At the same time, the content of the RFID has been read and the covariance matrix (3) calculated. If at least one of the images passes both stages, the subject is authenticated.

3. Performance Evaluation

In order to evaluate the FAR and FRR performance of our approach we adopted the “A priori performance type A” procedure based on the VidTIMIT database and proposed in [Sanderson, 2002]. The VidTIMIT database contains a set of poses belonging to 43 different subjects from two different classes, 35 *true claimants* and 8 *impostors*. Tests have been performed with two different configurations, using 2 and 4 main representative poses for each user, respectively.

For each configuration, the poses of each subject are grouped in 3 sessions. The first session is used to create a model of the subject, by determining a set of 2 or 4 main representative poses (according to the configuration of the test) through the clustering procedure described in Section 2.1. The second and third sessions are used for a two-phase testing procedure as follows.

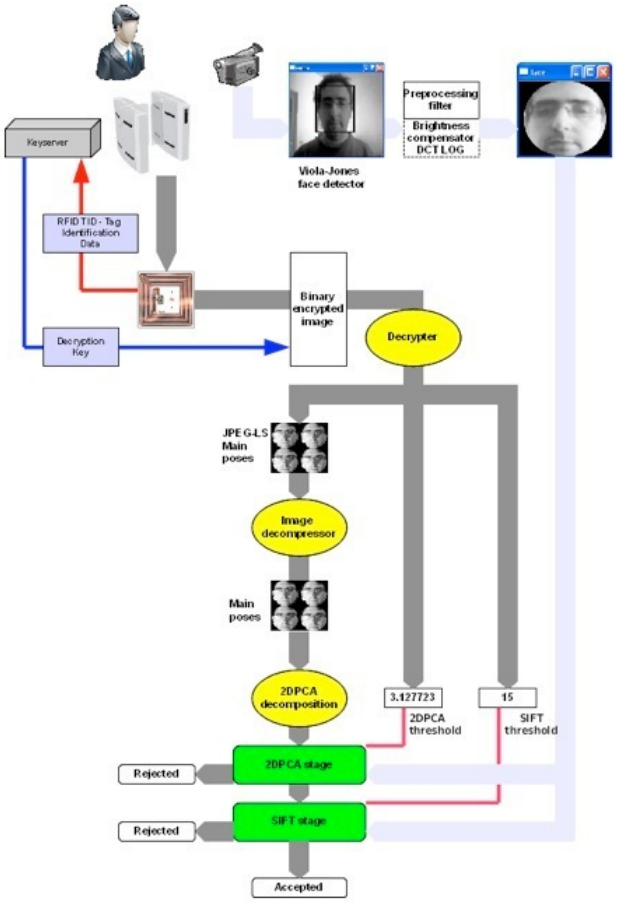

Figure 3
Schematization of the authentication process

During the first phase (phase A) the second session acts as a threshold optimization database (TOdb) for the procedure described in Section 2.1 and the third session is used to test the authentication performance and to measure the FAR and FRR with the thresholds (ρ_{2DPCA}, σ_{SIFT}) just calculated.

During the second phase (phase B) the roles of the second and third sessions are swapped and the procedure of phase A is repeated. Then the FAR and FRR measurements thus obtained are averaged in order to produce a single couple (FAR,FRR) for each subject.

Finally, the values avgFAR and avgFRR are calculated as the averages of FAR and FRR over all the subjects. The couple (avgFAR, avgFRR) constitutes an estimate of the system performance at the given resolution and configuration (2 or 4 main representative poses). For each one of the 35 true claimants we thus simulate 204 authorized authentication trials and 1632 unauthorized (i.e. impostors) authentication trials. The tests were performed at 4 different resolutions (20x15, 40x30, 80x60 and 160x120 pixels) and the number of vector

components for the 2DPCA stage was set as $z_{2DPCA}=w$, i.e. to its theoretical maximum. We now need to assign a value to the ratio C_{FAR}/C_{FRR} (see (4)), which also affects the performance of the system as shown in Fig. 4, drawn for a resolution of 80x60 pixels.

In order to find the value to the C_{FAR}/C_{FRR}, which maximizes the performance of the system, we report the values of normalized FAR and normalized FRR (FAR* and FRR*, respectively) on the plane [0,1]x[0,1] of Fig. 5, thus obtaining a point in the plane for each value of the ratio C_{FAR}/C_{FRR}, representing the performance obtained with that value.

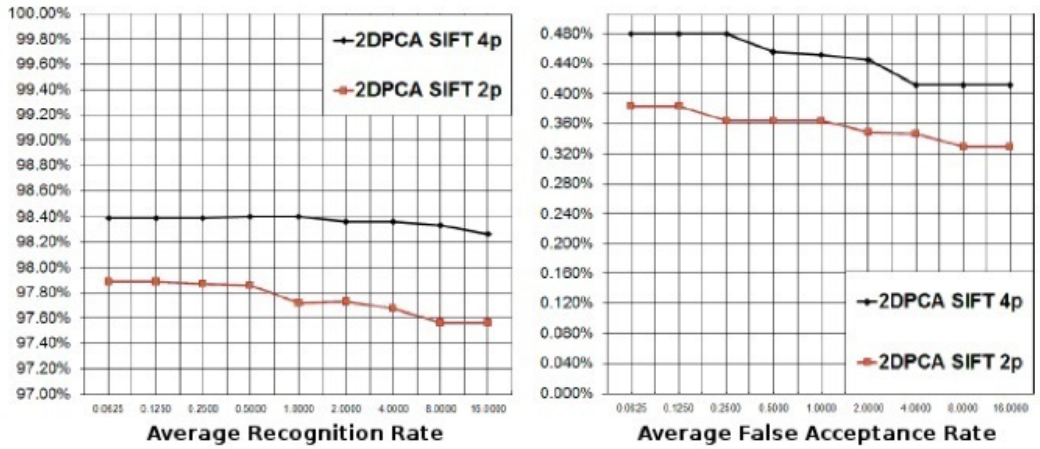

Figure 4
Average recognition rate(left) and average FAR (right) for different values of the C_{FAR}/C_{FRR} ratio. Resolution 80x60 pixel

Our target is to minimize both FAR and FRR, therefore we choose the value for the ratio corresponding to the point closer to the origin of the axes in the plane.

As shown in Fig. 5, the optimal ratio changes according to the number of main representative poses (actually, 2 and 4 for the two configurations). We choose to assign as a common value the average of the optimal values (i.e. 3), assuming that the performance would not be sensibly affected. Similarly, we assumed that the value chosen for the resolution 80x60 would also be acceptable for lower resolutions.

Fig. 6 shows the performance of VisilabFaceRec with changing resolution and two different configurations (2 and 4 main representative poses). The proposed system produces a recognition rate RR>97.69% at all resolutions higher than 20x15 pixels and a false acceptance rate FAR<0.45% for resolutions equal to or higher than 80x60 pixels.

In our two-stage architecture the SIFT stage allows a reduction of FAR while keeping a high value of the recognition rate RR (RR=1-FRR). In order to assess the actual improvement produced in terms of FAR while keeping constant RR, compared to the single-stage architecture, we designed the following experiment.

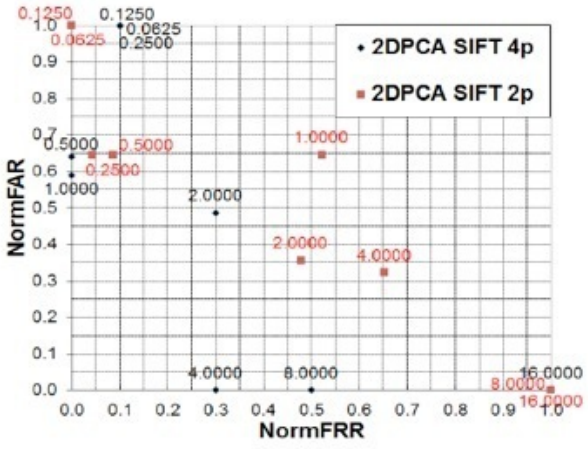

Figure 5
FAR and FRR* for different values of the C_{FAR}/C_{FRR} ratio at a resolution of 80x60 pixels and for 2 and 4 main representative poses for each subject*

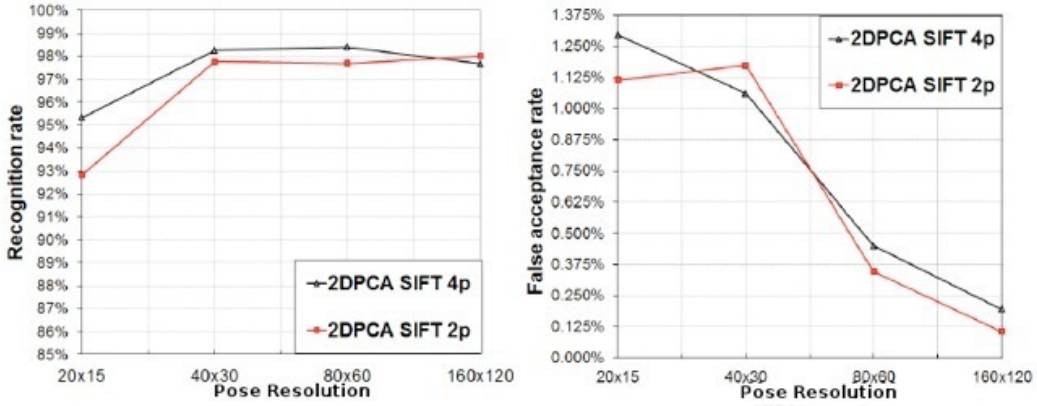

Figure 6
Average recognition rate (left) and FAR (right) vs. resolution

For each subject in the VidTIMIT database we consider the couple ($exFRR^{dblstage}$, $exFAR^{dblstage}$) (expected FRR, expected FAR) produced by the two-stage system according to the procedure in Section 2.1 and also the primary ROC corresponding to the single-stage 2DPCA system. Let us choose a new threshold $\mathbf{p}_{2DPCA}{}^{onestage}$ in such a way that:

$$exFRR^{onestage}(\rho_{2DPCA}^{onestage}) = exFRR^{dblstage}(\rho_{2DPCA}^{twostage}; \sigma_{SIFT}^{twostage}) \qquad (6)$$

In other words, we configure a 2DPCA single stage system so that we obtain the same *expected recognition rate* (exRR = 1-exFRR) than the two-stage system.

Then we run the single-stage system on the testing database in order to produce the actual (i.e., not *expected*) values of the single-stage RR and FAR. The result is shown in Fig. 7.

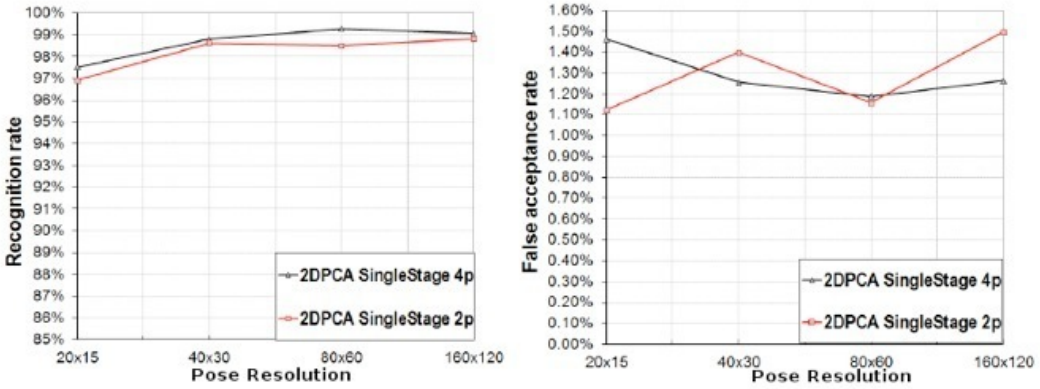

Figure 7
Average RR (left) and average FAR (right) versus pose resolution for the single stage system, while imposing the same expected FRR than the corresponding two-stage 2DPCA-SIFT system

Owing to (6), the recognition rate (Fig. 7, on the left) for the single stage system appears very similar to that of the corresponding two-stage system. Indeed, the single stage performs slightly better, because the second stage, besides rejecting a number of impostors (thus reducing the FAR), also rejects a small minority of true claimants (thus reducing the RR).

The actual improvement in the two-stage over the single-stage is shown by the FAR values (Fig. 7 on the right). The single stage produces a FAR sensibly higher than the two-stage for all resolutions higher than 20x15 pixels (below this resolution the SIFT stage does not operate correctly): Such results demonstrate the correctness of our approach.

Fig. 8 shows the minimum authentication time t_{min} of VisilabFaceRec and its components. The reported values were measured on an ASRock E350M1 board equipped with an AMD Fusion dual core processor running at 1.6GHz and an Inpinji Speedway RFID reader. The system uses high capacity RFID tags produced by Tego Inc.

The authentication time is minimum if the user is authorized immediately after the RFID is read, i.e. when the authentication of the very first image of the claimant succeeds. In this case we have:

$$t_{min}=t_{rfid}+t_{fdet}+t_{fspace}+t_{frec}+t_{sift} \qquad (7)$$

where:

t_{rfid} is the time needed for RFID tag reading, binary decryption, and decompression of the user main poses.

t_{fdet} is the time needed by the Viola-Jones cascade classifier for the detection of the user face.

t_{fspace} is the time needed to calculate the covariance matrix G_t according to (3), its eigenvalues and eigenvectors, and the 2DPCA decomposition vectors of the main poses according to (1).

t_{frec} is the time needed for the 2DPCA decomposition of the acquired image and its comparison with the decomposition vectors of the main representative poses retrieved from the RFID (according to (2)).

t_{sift} is the time needed for the calculation of the SIFT features of the image acquired and the application of the Lowe criterion.

At the resolution of 80x60 pixels, using the configurations with 4 or 2 main representative poses, the authentication requires 13.09 seconds and 6.72 seconds respectively. The dominant component of the authentication time, however, is the time t_{frid} needed to transfer the data from the RFID tag to the board and decode (i.e. decrypt and decompress) it. After the first read, if the authentication of the first image of the claimant fails, the next trials require much shorter times, $t_{min^*} = t_{min} - t_{rfid}$ (0.36 s and 0.29 s respectively), because the RFID data is temporarily cached. As, usually, the first trial is performed when the claimant is still approaching the camera, the first, slower trial in most cases does not significantly affect the perceived speed of the authentication system.

		4p 20x15	4p 40x30	4p 80x60	4p 160x120	2p 20x15	2p 40x30	2p 80x60	2p 160x120
RFID tag reading time t_{rfid} (s)	Avg	1.0072	3.6983	12.7322	50.1272	0.5517	1.9477	6.4282	25.1802
	StdDev	0.0787	0.2021	0.6568	2.5363	0.0395	0.1156	0.3333	1.2846
Face detection time t_{fdet} (s)	Avg	0.0605	0.0650	0.0638	0.0640	0.0605	0.0650	0.0638	0.0640
	StdDev	0.0062	0.0067	0.0071	0.0067	0.0062	0.0067	0.0071	0.0067
Face space rebuilding time t_{fspace} (s)	Avg	0.1208	0.1593	0.2549	0.5812	0.1129	0.1320	0.1887	0.3940
	StdDev	0.0066	0.0093	0.0172	0.0540	0.0048	0.0064	0.0178	0.0787
Face recognition time t_{frec} (s)	Avg	0.0002	0.0007	0.0031	0.0133	0.0002	0.0007	0.0028	0.0122
	StdDev	0.0000	0.0000	0.0001	0.0004	0.0000	0.0000	0.0001	0.0004
SIFT feature computation time t_{sift} (s)	Avg	0.0060	0.0147	0.0393	0.1088	0.0057	0.0143	0.0397	0.1097
	StdDev	0.0014	0.0019	0.0059	0.0218	0.0013	0.0018	0.0059	0.0214
$t_{min} = t_{rfid} + t_{fdet} + t_{fspace} + t_{frec} + t_{sift}$		1.1948	3.9382	13.0933	50.8946	0.7309	2.1597	6.7232	25.7601
$t_{min}^* = t_{fdet} + t_{fspace} + t_{frec} + t_{sift}$		0.1876	0.2398	0.3611	0.7674	0.1793	0.2120	0.2950	0.5799

Figure 8
Operating times of VisilabFaceRec (in seconds)

For a given configuration (number of main representation poses and resolution), the data transfer and decoding time t_{rfid} can be considered as almost constant, as it mainly depends on the transfer rate between the RFID tag and the processing board. As it is shown in Fig. 8, the other time components are characterized by small average values and very small standard deviations in all the tests we performed, therefore we suggest that our approach can be successfully used also for near real-time applications.

4. Conclusions

This paper presented VisilabFaceRec, a multi factor authentication system for controlling the access to services and restricted areas, which combines RFID tags and biometric recognition (face recognition) for the sake of improved accuracy, reliability and privacy. The system is specifically devised to work with very low resolution images, thus allowing for the storing of sensitive user data (e.g., face images) directly into the RFID tag, without the need for a centralized database. To the best of our knowledge, there are no other systems that obtain similar results, in terms of FAR/FRR balance, when working with the same resolution and operating constraints (i.e., using RFIDs to store sensitive data). The proposed system was realized and tested on a commercial board for embedded systems. The obtained execution times are short enough to be suitable for adoption in applications, such as access control of restricted areas, which cannot tolerate long authentication times or afford expensive hardware. Future work will deal with further improvements in the calculation of the two decision thresholds for the cascading stages and with the assessment of other algorithms for the implementation of the two stages with the aim to further increase the reliability of the authentication while reducing the execution time.

References

[Art. 29 WP, 2003]. Data Protection Working Party, Working document on biometrics. 12168/02/EN WP 80. Available online at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2003/wp80_en.pdf.

[Chen et al, 2006] Chen W., Er M.J., Wu S., Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain, IEEE Transactions on Systems, Man, and Cybernetics, 36, 2, 2006, 458-466.

[Flach, 2003] Flach P.A., The Geometry of ROC Space: Understanding Machine Learning Metrics through ROC Isometrics, Twentieth Intnl. Conf. on Machine Learning, 2003, 194–201.

[Jain et al, 2006] Jain A.K., Ross A., Pankanti S., Biometrics: A Tool for Information Security, IEEE Transactions on Information Forensics and Security, 1, 2, 2006.

[Jing et al, 2009] Jing B.Z., Yeung D.S., Ng W.W.Y, Ding H.L., Wu D.L., Wang Q.C., Li J.C., RFID Access authorization by face recognition,

Eighth Intnl. Conf. on Machine Learning and Cybernetics, Baoding,2009, 302-307.

[Lowe, 2004] Lowe D., Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60, 2004, 91-110.

[Meng et al, 2010] Meng X.L., Song Z.W., Li X.Y., RFID-Based Security Authentication System Based on a Novel Face-Recognition Structure, WASE International Conference on Information Engineering, 1, 2010, 97-100.

[Min et al, 2011] Min D.G., Kim J.W., Jun J.S., The Entrance Authentication System in Real-Time using Face Extraction and the RFID Tag, Intnl. Conf. on Ubiquitous Computing and Multimedia Applications, 2011, 20-24.

[Nguyen et al, 2012] Nguyen T.D., Quang L.D., Van N.C., Thanh L.T., Hoang T. M., De Souza-Daw T., An efficient and reliable human resource management system based on a hybrid of face authentication and RFID technology, Fourth Intnl. Conf. on Communications and Electronics (ICCE) 2012, 333-338.

[Pentland, 1991] Turk M., Pentland A., Eigenfaces for Recognition, Journal of Cognitive Neuroscience, 3, 1, 1991, 71-86.

[Sanderson, 2002] Sanderson C., The VidTIMIT Database, Available online at: http://publications.idiap.ch/index.php/publications/show/710.

[Viola and Jones, 2001] Viola P., Jones M., Rapid Object Detection Using a Boosted Cascade of Simple Features, IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, 2001, 511-518.

[Yang et al, 2004] Yang J., Zhang D., Frangi A.F., Yang J.Y., Two-dimensional PCA: a new approach to appearance-based face representation and recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, 26, 1, 2004, 131-137.

Biographies

Filippo Battaglia received the M.S. degree in Electronic Engineering from the University of Messina, Italy, in July, 2008 and the Ph.D. degree in Information Technology from the University "Mediterranea" of Reggio Calabria, Italy, in February, 2013. His research interests include artificial vision, voice synthesis, operating systems, micro-optoelectronics and communication systems.
email: filbattaglia@libero.it

Giancarlo Iannizzotto received the M.D. degree in Electronic Engineering from the University of Catania, Italy, in 1994 and the Ph.D. in Computer Science from the same University in February, 1998. From 1996 to 2006 he was Assistant Professor at the Faculty of Engineering, University of Messina, Italy. From 2006 to 2012 he was Associate Professor at the same Faculty. Currently he is Associate Professor at the Department of Cognitive Sciences, Education and Cultural Studies at the University of Messina. His research activity is in the fields of Computer Vision, Artificial Intelligence, Human-Computer Interaction and Human Factors in ICT.
email: ianni@unime.it

Lucia Lo Bello received the M.S. degree in Electronic Engineering and the Ph.D. degree in Computer Engineering from the University of Catania, Italy, in 1994 and 1998, respectively. She was a Visiting Researcher with the Department of Computer Engineering, Seoul National University, Korea (2000-01). She is currently an Associate Professor with tenure with the Department of Electrical, Electronic and Computer Engineering, University of Catania. Her research interests include real-time systems, wireless networks and sensor networks, factory communications, and embedded systems. She published more than 120 technical papers on peer-reviewed international conferences, books, and journals.
email: lucia.lobello@dieei.unict.it