A person authentication system based on RFID tags and a cascade of face recognition algorithms

Filippo Battaglia, Giancarlo Iannizzotto, Lucia Lo Bello

Abstract—Face recognition represents an appealing solution for biometrics-based unobtrusive and flexible person authentication. However, most solutions proposed in the literature suffer from some significant drawbacks, such as, high computational complexity, the need for a centralized biometrics database (which is not desirable, due to widespread international provisions discouraging collections of sensitive personal data) and limited scalability on a large number of enrolled subjects. We propose a novel person authentication solution based on a cascade of face recognition and pattern matching algorithms, that not only provides for high reliability and robustness against impostors, but also stores in a personal RFID tag all the needed individual biometrics information of the user, who therefore always remains in control, and has the exclusive availability, of such sensitive data. The paper describes the proposed approach, called RFaceID, and discusses its performance in terms of the ratio between false acceptance rate and false rejection rate and in terms of authentication time, when applied to the VidTIMIT, Extended Yale B and MOBIO widely adopted face databases.

Index Terms—Biometrics, face recognition, RFID tags

EDICS Category: Face Recognition, Biometrics, Authentication

1 Introduction

The interest towards Biometric authentication for high security protection systems has been consistently increasing over the last decade. Several realistic applications have been introduced, under the strict constraint that the selected physiological or behavioral characteristics of the subject cannot be stolen or imitated [1]. Face characteristics are among the most frequently adopted features, as they can be easily "sensed" from a distance of a few meters without requiring direct contact or close-up image acquisition [2]. Moreover, local as well as transnational regulations often explicitly enforce that people expose their face while approaching guarded gates or otherwise accessing controlled areas. However, face characteristics are considered as sensitive personal data in several countries and therefore their collection is strongly discouraged or even prohibited. Several provisions enacted by the responsible authorities (e.g., art. 29 Working Party of the European Commission for the protection of biometric data), advise against the use of a centralized server in these kind of systems, *except for few specific purposes*. The EC authority considers advisable that *"biometric systems are based on the reading of biometric data stored as encrypted templates on media that are held exclusively by the relevant data subjects (e.g. smart cards or similar devices)"* [3]. Therefore the approaches storing the biometric data in the authentication system (as most works in the literature do) should be avoided.

In order to overcome the issues described above, in this work we introduce RFaceID, a novel two-factor authentication architecture based on an RFID tag and a cascade of two face recognition stages. The first stage uses the *Two-Dimensional Principal Component Analysis* (2DPCA) [4], while the second is based on the *Speeded-up Robust Features Detector* (SURF) [5] pattern matching algorithm. The main value-added is threefold. First, a technique based on the novel *BestPoint model* to jointly calculate the optimal parameters for the two stages, which allows a very high recognition rate together with an extremely low false acceptance rate. Second, the proposed approach does not require a centralized database storing the biometric data of all the authorized subjects. It only relies on the personal biometric information stored into the RFID tag, which always remains available exclusively to the user. Third, RFaceID is devised to work on images captured at a very low resolution, compatible with the storing capacity of the small memories (8-32 Kbytes) of the passive RFID tags currently on the market, with improved accuracy over the existing state of the art (both in terms of false acceptance rate and of false rejection rate) even in presence of largely varying illumination.

This paper is organized as follows. Section II provides a description of related works and addresses their limitations. Section III proposes a relation between face recognition and face authentication algorithms, thus paving the way for the proposed authentication approach, which is introduced in Section IV. Sections V

Giancarlo Iannizzotto is with Department of Cognitive Sciences, Education and Cultural Studies, University of Messina, Italy, e-mail: {ianni@unime.it}

Filippo Battaglia and Lucia Lo Bello are with DIEEI, Department of Electrical, Electronic and Computer Engineering, University of Catania, Italy, e-mail: {filippo.battaglia@dieei.unict.it}, {lucia.lobello@unict.it}

and VI describe the enrollment and the authentication phases, respectively. Section VII presents a testing protocol developed in order to assess the performance of the proposed approach and compares its results with those produced by a recent state-of-the-art algorithm based on Gabor Disparity [6]. Section VIII demonstrates the superiority of RFaceID with respect to the VisilabFaceRec algorithm, which shares some structural similarities with RFaceID and was recently presented in [7]. Finally, Section IX concludes the paper and gives hints for future work.

2 Related work

A significant effort has been made over the years to develop several template matching algorithms for face recognition e.g., *Eigenfaces*, based on Principal Components Analysis (PCA) [8], *Fisherface*, based on Linear Discriminant Analysis (LDA) [9], and *Two-Dimensional Principal Component Analysis* (2DPCA) [4]. However, a unified database was always needed to keep the biometric data of all the enrolled users. Moreover, currently available biometric recognition algorithms have not proven, yet, to be able to achieve 100% accuracy for an arbitrarily large database. Consequently, the interest towards authentication systems that combine token-based and biometrics-based techniques in order to obtain more reliable architectures, is recently increasing [10]. Thanks to their wide availability and versatility, in the last decade RFID tags became the leading candidates for the role of physical authentication tokens. In [11] Seo and Baek proposed an interesting authentication system based on face recognition and 64-bit capable RFID tags. In their architecture, each face is transformed into a set of 8 PCA [8] coefficients. Due to the RFID capacity constraints, the typical 32-bit size (float) of each of the 8 PCA components used is compressed into an 8-bit size (byte). During the user authentication phase, the system restores the user facial information from the data stored in the RFID by referring to a *backward database*, i.e., a centralized database managed by the service provider with the face information of the authorized users. Similarly, integration of RFID tags (i.e., radio-frequency tokens) with face recognition systems were proposed by Min et al. in [12], by Jing et al. in [13], by Nguyen et al. in [14], by Jong et al. in [15] and by Affandi et al. in [16]. However, those 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). Moreover, for each single claimant, a large number of poses at high resolution must be acquired under different lighting conditions and stored in the database [14].

Meng et al. in [17] proposed an embedded system that stores in the RFID tag only the set of the n principal decomposition components (PCA) [8] associated to the owner's face. However, for both the enrollment and authentication stages the PCA representation requires the availability, local to the authentication system, of a set of images (Eigenfaces) which depend on all 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 cited 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.

A two-stage authentication system, named *VisilabFaceRec*, was introduced in [7] but it was not devised to work in strongly variable illumination conditions and some working parameters have to be manually tuned by the operator before using the system. RFaceID, instead, does not need to manually set any critical parameter and introduces an adaptive algorithm that provides better performance.

In our knowledge, all the other solutions currently available in the literature share one or more of the following drawbacks:

1) They need to recalculate *all* PCA coefficients, every time a new authorized user is added;
2) They need a remote centralized database for biometric data and, therefore, an *always-on* connection for each checkpoint.

Furthermore, the solutions cited above were mainly tested on large images (e.g. 320x240p in [12], 200x200p in [14]) which makes them unsuitable for the small memories of the RFID tags, acquired in strongly controlled conditions or by using domestic databases ([13][14]). In [13] a false rejection rate (FRR) of 12.22% and a false acceptance rate (FAR) of 3.89% were reported over a domestic dataset consisting of only 3 users, whereas in other works a measure of the FAR is not provided [11][12][15], so their applicability to authentication cannot be assessed. Conversely, RFaceID provides a good recognition rate (RR) while being robust against hundreds of intrusion attempts even in presence of strongly variable conditions of illumination and pose. Finally, the RFaceID architecture is designed to resist against both *biometric data stealing* and *tag counterfeiting*, while the works cited above offer no protection against such attacks.

3 Face Recognition and Face Authentication

A face recognition algorithm, given a face database B of people whose identity is known *"a priori"* and a face image I of unknown identity, determines the pose in B that is most similar to I. In this paper we call B *definition database*, because it defines the face space in which the algorithm works. A definition database of $m = r*k$ poses belonging to r subjects (k poses per subject) can be seen as a *face table* of r rows containing the k poses. In other words, each row corresponds to a subject and contains

her k poses. A recognition algorithm can be modeled as in (1):

$$\{i_{subject}, i_{pose}, d\} \rightarrow \{g_{i\_s}(I,B), g_{i\_p}(I,B), g_d(I,B)\} \quad (1)$$

where I is the face image of the subject to be recognized, B is the definition database, $i_{subject}$ and i_{pose} are the row (i.e., the subject index) and column (i.e., the specific pose of that subject) indexes of the image $I^* \in B$ that is most similar to I, d is the distance between the images I and I^*, while $g_{i\_s}(I,B)$, $g_{i\_p}(I,B)$ and $g_d(I,B)$ represent their dependence upon I and B.

An authentication algorithm, instead, acquires as an input the image I of the claimant face and the declared identity i_{dec_id} of the claimant, and compares I with one or more face images I^* associated with i_{dec_id} in a database B that is known *"a priori"*. If I and I^* match, the algorithm authenticates the claimant. The behavior of such an algorithm can be represented as a function $f(I, i_{dec\_id}, B) \in \{0,1\}$ that returns 1 when the claimant can be authenticated and 0 otherwise.

Face recognition and face authentication are, in general, very different tasks. However, by introducing an *ideal database* D made of infinite face images (poses) of the whole world population, we can write the following general equation, linking face authentication and face recognition:

$$f(I, i_{dec\_id}, D) = \begin{cases} 1, \text{if } g_{i\_s}(I,D) = i_{dec\_id} \\ 0, \text{if } g_{i\_s}(I,D) \neq i_{dec\_id} \end{cases} \quad (2)$$

A real database B can only include the poses of a limited number of subjects, therefore, when an impostor tries to authenticate, two mutually exclusive events can happen:

1) The impostor is recognized as a subject in the database B, associated to an identity different from the declared identity i_{dec_id}. As a result, $g_{i\_s}(I,B) \neq i_{dec\_id}$ and the impostor is correctly rejected.
2) The impostor is very similar to the claimant whose identity is the declared identity i_{dec_id}, and, being B of limited size, there is no other subject that is more similar to the impostor. In this case, $g_{i\_s}(I,B) = i_{dec\_id}$ and the impostor is authenticated (*false acceptance*).

The probability that the last event occurs can be reduced by adding a further condition to Eq. (2), that allows the authentication only if the similarity level $g_d(I,B)$ is higher than a threshold ρ. Moreover, without loss of generality, we can assume that, reordering the indexes, the poses of the authorized subject can be moved to row 0, while the other rows will be dedicated to a set of impostors (named *extrasubjects*), chosen according to some criterion. As a result, condition 2) can be written as in (3):

$$f(I,B) = \begin{cases} 1, \text{if } g_{i\_s}(I,B) = 0 \wedge g_d(I,B) < \rho \\ 0, \text{otherwise} \end{cases} \quad (3)$$

A face recognition algorithm, represented as in (1), can be used for face authentication, provided that (3) is adopted as the discrimination condition. As a corner case, if the database B contains only the poses of the authorized subject and does not contain any pose of impostors, Eq. (3) becomes:

$$f(I,B) = \begin{cases} 1, \text{if } g_d(I,B) < \rho \\ 0, \text{otherwise} \end{cases} \quad (4)$$

According to Eq. (4), the system authenticates a claimant only based on the similarity between the pose I, grabbed by the camera, and the pose I^*, chosen as the most similar to I among those belonging to the authorized user in B.

4 THE RFACEID ARCHITECTURE

Before a subject can be authenticated, she must be *enrolled*, i.e., a set of images of her face (named *main poses*) must be recorded and associated to her identity information. In RFaceID, this task is performed by storing together the images and the identity information of the new user in the RFID tag (see Section V-D). The enrollment is a one-time process.

The authentication phase begins when a subject approaches an authentication gate, claiming for an identity. A sequence of images of the face of the subject (poses) is acquired and the content of the RFID tag is read. The acquired poses are processed one by one, and as soon as one of them is authenticated, the subject is granted the claimed identity.

RFaceID exploits two cascaded authentication stages, optimized for low-resolution images which can be stored in the small memory embedded in commercially available RFID tags (usually 4-32 KB). The first stage is based on the 2DPCA recognition algorithm [4]. An image I of size $w * h$ pixels can be decomposed into the set $(\mathbf{W_0}, \mathbf{W_1}, ..., \mathbf{W_{z^{2dpca}-1}})$ of *decomposition vectors*, calculated by multiplying the matrix of the pixel intensities of I by each one of the z^{2dpca} principal eigenvectors $\mathbf{\Psi_i}$ of the 2DPCA covariance matrix $\mathbf{G}_t$ (of size $w * w$):

$$\mathbf{G}_t = \frac{1}{m}\sum_{j=1}^{m}(\mathbf{B}_j - \bar{\mathbf{B}})^T(\mathbf{B}_j - \bar{\mathbf{B}}) \qquad \mathbf{W}_{j,I} = \mathbf{I}\mathbf{\Psi}_j \quad (5)$$

The number $z^{2dpca} \leq w$ of decomposition vectors is chosen according to the fraction of the original information that we want to retain. The behavior of 2DPCA as an authentication system can in general be modelled by Eq. (3) or, if only the poses of a single authorized subject are stored in the database B, by Eq. (4). In order to evaluate the similarity of the pose I belonging to an unknown subject, with each pose B_i in B, we decompose both I and B_i and calculate their *distance in the 2DPCA feature space* (DIFS), defined as in (6):

$$DIFS(I, B_i) = \sum_{j=1}^{z^{2dpca}} \parallel \mathbf{W}_{j,I} - \mathbf{W}_{j,B_i} \parallel_2 \quad (6)$$

where $\mathbf{W}_{j;I}$ and $\mathbf{W}_{j,B_i}$ are the j^{th} decomposition vectors of I and B_i. As a result, in Eq. (3), $g_{i\_s}(I,B)$ is the pose I^* in B that minimizes (6) and $g_d(I,B) = DIFS(I,I^*)$.

Usually, the performance of a recognition algorithm is evaluated by analyzing its false acceptance rate (FAR) and false rejection rate (FRR) measures. Both measures are supposed to be minimal for an optimal recognition algorithm, but, unfortunately, in single-threshold systems, FAR and FRR cannot be minimized at the same time [18]. The decision threshold ρ^{2dpca} must therefore be determined as a trade-off between the two minimization objectives. Moreover, it can be shown that ρ^{2dpca} depends on the poses of the collection B and therefore is specific for each collection. In Section V-B we propose a solution for the two above mentioned problems.

The poses that pass the threshold of the 2DPCA stage are sent as an input to the second stage. This is based on the *Speeded-Up Robust Features Detector* (SURF) algorithm, proposed by Bay in [5] and aimed at extracting distinctive invariant features from an image. For an input image I, SURF returns a set of *SURF features* (nearly 30 at 80x60p size) that are very resistant to changes of illumination, rotations, translations, and noise. Each feature consists of the coordinates of a keypoint (x_{kp_i}, y_{kp_i}) and a keypoint descriptor $\mathbf{v_i} \in R^{128}$.

The system compares the features extracted from the image I of the unknown claimant and from the most similar pose I^* in B determined by the first stage. If the number of the matching features (that can be considered as the *similarity score* between I and I^*) is higher than a threshold σ^{surf}, the image I also passes this second stage and the unknown claimant is finally authenticated.

5 The Enrollment phase

The enrollment phase is performed off-line by the issuing agency and is aimed to determine the main poses for the new authorized user and the decision thresholds for the two authentication stages. As a first step, $n_{enrolldb}$ poses of the user are captured and stored in the *enrollment database*.

For every pose I, the face is located through the Haar-like face detector proposed by Viola and Jones in [19] [1]. The area in the image I where a face is present is extracted and cropped. The cropping size is then normalized to $w * h$ and the area is made rotation-invariant [21]. The resulting image is recropped so that the nose is at the center of the image and the size of the resulting sample is $w * h$. Next, the average brightness and the standard deviation of the pixels inside a centered elliptic area of rays $(r_x = 0.9w, r_y = 1.0h)$ is calculated. The average value is subtracted from the pixel intensities of the area and the resulting values are then divided by the standard deviation. Conversely, the intensities of the pixels outside the considered area are set to zero, thus realizing a *background subtraction*. The enrollment database is then divided in two subsets, i.e., the *main poses subset*, ($n_{mainposes}$ poses), and the *threshold optimization database*, ($n_{throptdb}$ poses).

The main poses are extracted through clusterization in the 2DPCA space. First the 2DPCA decomposition vectors of the $n_{enrolldb}$ poses are calculated, and, afterwards, a clustering algorithm is applied to obtain $n_{mainposes}$ clusters. For each cluster, the pose closest to the cluster centroid is chosen to be one of the *main poses*, while the remaining poses are used for the threshold optimization process (see Section V-B). Therefore, if the enrollment set contains samples that were grabbed under different conditions of lightning or orientation, the main poses represent $n_{mainposes}$ main changes in the aspect of the authorized user [2].

Since the number of clusters is determined a priori, k-means [22] would be a leading candidate as the clustering algorithm. However, as in our case each pose is represented by a set $\{\mathbf{W}_j\}$ of z^{2dpca} decomposition vectors (i.e. a *vector set*) instead of a single vector (as it would be in PCA/LDA), the traditional k-means algorithm based on Euclidean distance between points is not applicable. As a consequence, we exploit a modified version named *mm-kmeans*, that processes the vector sets by analogy with how the traditional k-means algorithm processes the points. The mean vector set between two vector sets $\{\mathbf{W}_{I_A}\}$ and $\{\mathbf{W}_{I_B}\}$ is defined as in (7):

$$\{\bar{\mathbf{W}}_j\} = \left\{ \frac{\mathbf{W}_{j,I_A} + \mathbf{W}_{j,I_B}}{2} \right\} \qquad j \in [0..z^{2dpca}[\quad (7)$$

The distance between two vector sets is the *pair-wise euclidean* distance defined in (8):

$$d(I_A, I_B) = \sum_{j=1}^{z^{2dpca}} \parallel \mathbf{W}_{j,I_A} - \mathbf{W}_{j,I_B} \parallel_2 \quad (8)$$

The i^{th} *mm-centroid* $\{\mathbf{W}_{C_i}\}$ is determined by applying Eq.(7) to all the vector sets in the cluster C_i ($i \in [1..n_{mainposes}]$). The clustering process iterates a number Γ of times and, finally, the $n_{mainposes}$ poses corresponding to the vector sets that are nearest to the $n_{mainposes}$ *mm-centroids* are chosen as the main poses. In our experiments we used $\Gamma = 10000$ and $z^{2dpca} = w$, i.e., the maximum number of available 2DPCA decomposition vectors. Our approach was tested with different values of $n_{mainposes}$ ($\{2, 4, 8\}$) and the results are shown in Section VII. The main poses are used to build a single-row definition database B, that does not contain any pose of impostors (see Section V-A for the motivations of this choice).

Unlike the solutions proposed in [11] and [17], in RFaceID the decision thresholds associated to the two

1. In order to improve the effectiveness of our approach on images affected by strong brightness variations, the Haar-like detector is applied to a histogram-normalized [20] copy of the input image I, while the recognition stage is applied directly to the original input image.

2. Some examples of the main poses before and after preprocessing and a representation of the related 2DPCA *vectorsets* for all databases used in this work are provided in separate document, uploaded while submitting this paper, that will also be made available to the interested Readers.

stages are optimized for each single subject and stored into the RFID tag together with the biometric information. Such subject-specific thresholds are calculated as described in Section V-B and finally stored in the RFID tag as described in Section V-D.

5.1 The size of the definition database

Holistic algorithms like PCA, LDA or 2DPCA attempt to retain information along the z *principal directions* with the largest between-class variance [4; 23] over all samples in B. The image variations orthogonal to all those principal directions are unavailable for comparison. By increasing z, more and more directions are considered, usually achieving better recognition rates [4; 9; 24]. As a result, including the highest available number of different directions would be appealing, in order to take into consideration also changes in the smallest details in the image. Unfortunately, in PCA and LDA the maximum number of different directions (associated to non-null eigenvalues) is upper-bounded by the numbers of training classes and samples ($z^{pca} < r*k$ and $z^{lda} < r$, respectively) [8; 9]. As a consequence, z might be increased only by increasing the number k of samples per class, or the number r of classes (i.e. adding the poses of $r-1$ impostors to those belonging to the authorized user). Conversely, in 2DPCA the upper bound on z does not depend on the size of B ($z^{2dpca} \leq w$) [4], therefore, in this work a simplified model is used (described by Eq. (4)) where the database B contains only the poses of the authorized subject, without impostors.

5.2 Determination of the thresholds pair

For a *single-stage matcher*, the expected performance is usually represented by the *receiver operating curve* (ROC), which plots the percentage of false acceptances (FAR) and of false rejections (FRR) as a function of a threshold value ρ. Such a value is commonly determined by choosing a predefined FAR/FRR ratio (for example FAR=FRR=ERR that is named *Equal Error Rate*) or by minimizing a linear cost function $W_{ER}(\rho, c) = cFRR(\rho) + (1-c)FAR(\rho)$.

For a serial system made up of two matchers the problem becomes very hard. Some methods were recently proposed in the literature, but they show some limitations:

1) Using the Neyman-Pearson lemma requires an analytic approximation of the *Score Density curves* (derived from the ROC), through *Gaussian Mixture Model* (GMM) [25] or *Kernel Density Estimator* (KDE) that sensibly increases the computational complexity of the training algorithm [26]. The lemma can be easily applied only if the features used by the matchers are independent of each other (e.g. face and fingerprint);
2) The Marcialis-Roli model [27] exploits a pair of thresholds (s_1^l, s_1^u) on the *similarity score* s_1 achieved by the first matcher (such as $FRR_1(s_1^l) = 0$ and $FAR_1(s_1^u) = 0$) and a single threshold s_2^* for the second one. The subject is immediately authorized if $s_1 > s_1^u$, and rejected if $s_1 < s_1^l$. Only if $s_1^l < s_1 < s_1^u$ the second matcher is used. The method states no rule for the threshold s_2^* of the second matcher and is devised to minimize the verification time rather than the overall performance [3].

The method proposed here (*BestPoint model*) is based on searching, for each enrolled user, the pair $(\rho^{2dpca}, \sigma^{surf})$ that minimizes both FAR and FRR. The idea is to progressively increase ρ^{2dpca} and decrease σ^{surf} in order to make more (less) permissive the first (second) matcher, by allowing that more and more genuine poses can go through the first stage, thus improving the RR while not increasing the FAR. The threshold ρ^{2dpca} is increased as long as the second stage is able to compensate the higher rate of impostors that pass the first stage. This method provides some advantages:

1) Unlike the Marcialis-Roli method, the BestPoint method exploits *both* matchers in synergy, in order to improve the overall recognition rate without affecting the FAR;
2) It can be easily applied also when the features used by the matchers are not independent of each other;
3) It can be generalized to a large class of serial face authentication systems, consisting of matchers based on a *template-matching* and on a *feature-matching* algorithm.

The first step is creating a database of poses which includes both a set of poses of the user being enrolled and a set of poses of impostors. Such a database $B_{throptsdb}$ is called "*threshold optimization super-database*", as it is an extension of the *threshold optimization database* $B_{throptdb}$ created during the enrolling phase. It is composed of $r_{throptsdb}$ rows and $k_{throptsdb}$ columns. The $n_{throptdb}$ main poses of the subject are saved in the row 0, so $k_{throptsdb} = n_{throptdb}$. The other $r_{throptsdb} - 1$ rows contain the poses of different impostors, obtained from a freely available database (such as VidTIMIT [28]). Each image in the threshold optimization super-database is then tested for authentication with respect to the definition database B_{defdb}. As a result, the system simulates $n_{throptsdb} = r_{throptsdb}k_{throptsdb}$ authentication attempts, $n_{throptdb}$ of which are taken from the genuine user and $(r_{throptsdb}-1)k_{throptsdb}$ from impostors. During the simulation, the behavior of the 2DPCA stage can be modeled, according to Eq. (1) by the functions in Eq. (9):

$$i_{pose} = g_n(I, B_{defdb}) \qquad d = g_d(I, B_{defdb}) \tag{9}$$

where I is an image of the super-database $B_{throptsdb}$ to be authenticated, i_{pose} is the closest pose in B_{defdb} belonging to the authorized user, and d is their distance.

3. As the pair (s_1^l, s_1^u) depends only on the ROC of the first matcher, there is no guarantee that the load of impostors that pass the first stage can be blocked by the second one without increasing the overall FAR.

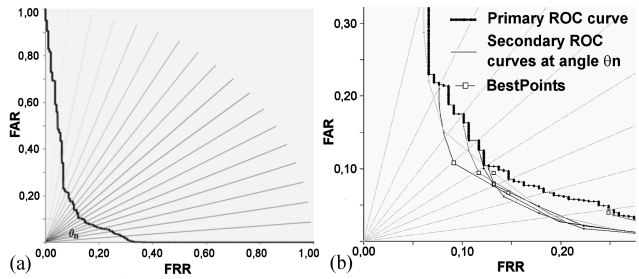

Fig. 1. (a) The primary ROC for a single user with some of the r_n lines intersecting the curve. (b) A magnified detail of the normalized ROC plane. The primary ROC and the secondary ROCs were drawn only for some angles for the sake of clarity.

If ρ_{max} is the maximum value of the distance d, determined by comparing each image in the super-database with the main poses in the definition database, then all the possible values of the first-stage threshold ρ can be considered between 0 and ρ_{max} and, for each of them, it is possible to evaluate how many poses, of either the authorized subject or the impostors, would be accepted and how many would be rejected. In this way, the first stage ROC can be drawn on a plane (FRR, FAR) scaled to $([0,1] \times [0,1])$ (see Fig. 1(a)).

Let us, now, consider the sequence of straight lines passing through the origin of the normalized plane (FRR, FAR):

$$r_n : FAR - tg\theta_n FRR = 0 \tag{10}$$

where $\theta_n = n\theta_0$ with $n = [1..\varsigma]$ and $\theta_0 = 90° \cdot (1+\varsigma)^{-1}$. For our tests, the parameter ς was set to 35 as a heuristic trade-off between accuracy in threshold determination and required computational power, and therefore $\theta_0 = 2.5°$. By varying n in its definition interval, the whole quadrant can be probed and, for each value of θ_n, the intersection point p_n between the straight line r_n and the first stage ROC curve, as well as the corresponding value of ρ_n, can be determined (see Fig. 1(a)). The final result of this step is a set of 35 values for the first stage threshold $\{\rho_n\}$, corresponding to a set of points $\{p_n\}$ on the first stage ROC curve.

The "impostors load" that the second stage can tolerate can thus be estimated by executing 35 simulations, and assuming for each simulation a different threshold $\rho^{2dpca} = \rho_n$ (n=1..35) for the first stage. For each simulation, the system considers each pose I in $B_{throptsdb}$ and, if it goes through the first stage (with threshold $\rho^{2dpca} = \rho_n$), the corresponding nearest pose in the definition database ($I^* = g_{i\_p}(I, B_{defdb})$) is retrieved. Then the second stage calculates the SURF features of I and I^* and applies the SURF matching criterion to determine, and record, the *matching score* (i.e., the number of matching features) for the image I. Finally, the maximum matching score σ_n^{max} over the whole super-database is determined.

By varying the SURF threshold σ within the integer interval $[0..\sigma_n^{max}]$ and applying for each value of σ the two cascaded stages to each image of the super-database (with fixed ρ_n threshold for the first stage), a *secondary ROC curve* $c_n(\rho_n, \sigma)$ on the plane (FRR, FAR) can be drawn. As this analysis can be repeated for each ρ_n, 35 secondary ROCs can be plotted (see Fig.1(b)). These curves have the following properties:

1) The secondary ROC curve always starts from the corresponding point p_n on the first stage ROC curve. In fact, for $\sigma = 0$ the second stage is *disabled* and accepts all the poses which pass through the first stage;
2) The secondary ROC curve can only reduce the FAR or increase the FRR respect to the values corresponding to p_n. In fact, the second stage can only discard poses and cannot accept poses that were discarded by the first stage.

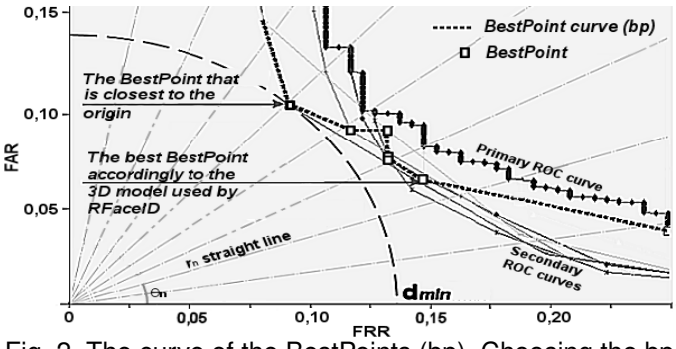

Fig. 2. The curve of the BestPoints (bp). Choosing the bp closest to the origin in the 2D model could entail a high FAR. As a consequence, the bp closest to the origin in $(FAR, FRR, FAR/FRR)$ 3D model is chosen (it is the so-called "best BestPoint"). d_{min} is the distance between the origin and the closest BestPoint.

For each secondary ROC curve, the software determines the point that is closest to the origin of the plane. This point b_n, associated to a threshold pair $(\rho_n^{2dpca}, \sigma_n^{surf})$, is hereon called *BestPoint*. More precisely, this is the point on the n^{th} secondary ROC curve that minimizes the distance:

$$d(FRR; FAR) = \sqrt{FRR^2 + FAR^2} \tag{11}$$

There are ς BestPoints available (35 in our tests), each one representing a suboptimal solution for the problem dealt with here. The optimal configuration $(\rho^{2dpca}, \sigma^{surf})$ for the user being enrolled corresponds to one of them.

A possible choice could be the BestPoint that minimizes the distance in Eq. (11) but, unfortunately, there would no guarantee that the FAR is minimized (see Fig.2). For this reason, a new three-dimensional model is proposed, where the minimization of the FAR is enforced. For each BestPoint $\{bp_n\}$, three values are considered:

$$\begin{aligned} x(bp_n) &= FRR(bp_n) \\ y(bp_n) &= FAR(bp_n) \\ \eta(bp_n) &= \frac{FAR(bp_n)}{FRR(bp_n)} \end{aligned} \tag{12}$$

The proposed algorithm calculates the extrema (x^{min}, y^{min}, η^{min}, x^{max},y^{max},η^{max}) over all the available BestPoints and normalizes their values according to Eq. (13):

$$x^\star(bp_n) = \begin{cases} \frac{x(bp_n)-x^{min}}{x^{max}-x^{min}} & \text{if } x^{max} \neq x^{min} \\ 0 & \text{otherwise} \end{cases}$$
$$y^\star(bp_n) = \begin{cases} \frac{y(bp_n)-y^{min}}{y^{max}-y^{min}} & \text{if } y^{max} \neq y^{min} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$
$$\eta^\star(bp_n) = \begin{cases} \frac{\eta(bp_n)-\eta^{min}}{\eta^{max}-\eta^{min}} & \text{if } \eta^{max} \neq \eta^{min} \wedge y^{max} \neq y^{min} \\ 0 & \text{otherwise} \end{cases}$$

Now the ς BestPoints can be represented in a tridimensional space ($x^\star(bp_n)$, $y^\star(bp_n)$, $\eta^\star(bp_n)$) (see Fig. 3(a)). The chosen BestPoint (called *best BestPoint*) is the one which minimizes the distance:

$$d(x^\star, y^\star, \eta^\star) = \sqrt{(x^\star)^2 + (y^\star)^2 + (\eta^\star)^2} \quad (14)$$

This *best BestPoint* represents a trade-off between the opposite requirements to minimize FRR, FAR, and their ratio. The corresponding threshold pair is the optimal configuration ($\rho^{2dpca}, \sigma^{surf}$) for the user being enrolled.

5.3 An alternative model for critical databases

The model described in V-B works well on databases where the genuine poses are uniformly illuminated. Conversely, in a non-uniform illumination dataset it can fail, because a large part of the genuine subset may contain very few SURF matching features. The system tries to improve RR reducing the fraction of these samples discarded by the second stage, but this fatally leads to choose $\sigma^{surf} = 0$.

In order to overcome this issue two changes to the proposed model are needed. The first is the deployment of a *homomorphic filter* [21; 29] before the SURF stage. In this work, a high-pass Gaussian filter is applied on the frequency image $Z(f_x, f_y)$ in the logarithmic domain related to the input image of the SURF stage. The filter is defined as in (15):

$$H(f_x, f_y) = 1 - e^{-\frac{1}{\mu^2}\frac{{f_x}^2+{f_y}^2}{(w^\star)^2+(h^\star)^2}} \quad (15)$$

where $w^\star = min(w, 80)$ and $h^\star = min(h, 60)$ and μ is proportional to the cut-off frequency of the filter. A too high value for μ attenuates the effect of illumination changes but it may cut some details that are distinctive of the subject. An assessment of the value chosen for μ is in Sec. VII-E.

The second change is modifying the FRR formula, in order to reduce the importance of those genuine samples that, being acquired under poor illumination, are characterized by few SURF matching features (i.e. less than a threshold σ_{KT}). This workaround is used when the value of σ^{surf} achieved by the standard method is less than a preset *security value* σ_T.

Usually $FRR(\rho, \sigma) = n_{FN}(\rho, \sigma)/k_{throptdb}$ is used, where $n_{FN}(\rho, \sigma)$ is the number of rejected genuine samples composing the set of *false negatives* $S_{FN}(\rho, \sigma) \subset B_{throptdb}$. Assigning a *weight* $w(I_i) = 1$ to the poses I_i, FRR becomes:

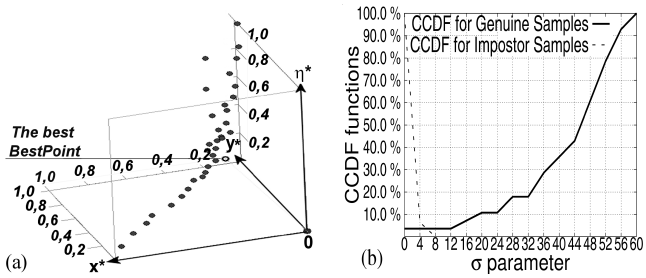

Fig. 3. (a) Best points in the three-dimensional space; (b) The CCDF functions $\vartheta_{gen}(\sigma)$ and $\vartheta_{imp}(\sigma)$ drawn for a claimant of the Extended Yale Database B ($\beta^{surf} = 0.5$, $\mu = 0.25$, res. 80x60p).

$$FRR^\star(\rho, \sigma) = \frac{\sum_{I_i \in S_{FN}(\rho,\sigma)} w(I_i)}{\sum_{I_j \in B_{throptdb}} w(I_j)} \quad (16)$$

Now consider the complementary cumulative distribution functions (CCDF) $\vartheta_{gen}(\sigma)$ and $\vartheta_{imp}(\sigma)$ related to the genuine and impostor samples $\mathbf{x} \in B_{throptsdb}$, defined as:

$$\vartheta_{gen}(\sigma) = \frac{\text{card}\{\mathbf{x} : \textit{surfmt}(\mathbf{x}) \geq (\sigma_{max} - \sigma)\}}{k_{throptdb}} \quad (17)$$
$$\vartheta_{imp}(\sigma) = \frac{\text{card}\{\mathbf{x} : \textit{surfmt}(\mathbf{x}) \geq \sigma\}}{(r_{throptsdb} - 1)k_{throptsdb}} \quad (18)$$

where $\text{card}\{\mathbf{S}\}$ is the number of elements (cardinality) of the subset $\mathbf{S}$, the function *surfmt*($\mathbf{x}$) returns the number of matching SURF features of the sample $\mathbf{x}$ and σ^{max} is the maximum number of matching features detected over the whole $B_{throptsdb}$ database.

$\vartheta_{gen}(\sigma)$ and $\vartheta_{imp}(\sigma)$ are respectively monotone increasing and decreasing curves (see Fig.3(b)). Assuming that, for an *ideal* genuine pose, $\sigma^*_{gen} = \sigma_{max}$ and for an *ideal* impostor pose, $\sigma^*_{imp} = 0$, they express the fraction of samples accepted when setting the threshold σ^{surf} to $|\sigma^* - \sigma|$.

The threshold σ_{KT} can be determined through the condition $\vartheta_{gen}(\sigma_{KT}) = \vartheta_{imp}(\sigma_{KT})$, that defines the intervals $[0, \sigma_{KT}]$ for the impostor samples and $[\sigma_{max} - \sigma_{KT}, \sigma_{max}]$ for the genuine ones characterized by the same probability of acceptance. By assuming that the genuine samples in $[0, \sigma_{KT}]$ were acquired under poor illumination conditions, thus causing a reduction of the number of SURF features being detected, such genuine samples can be ignored in the FRR estimation. Accordingly, the *weighted FRR* is defined by (16) and (19):

$$w(I, \rho, \sigma) = \begin{cases} 0, & \text{if } (\sigma < \sigma_{KT} \wedge \rho < \rho^{2dpca}) \\ \vartheta_{gen}(\sigma), & \text{if } (\sigma \geq \sigma_{KT} \wedge \rho < \rho^{2dpca}) \\ 1, & \text{if } \rho \geq \rho^{2dpca} \end{cases} \quad (19)$$

and the Bestpoints can be found through (11), (16) and (19).

TABLE 1
Memory requirements for compressed main poses (in Bytes)

pose size	2poses	4poses	8poses
80x60p	7680	15360	30720
40x30p	2160	4320	8640
20x15p	540	1080	2160

The advantage of the proposed model is that it is not affected by the distribution of the genuine samples in the neighborhood of $\sigma = 0$. Instead, it depends on the distribution of the genuine samples in the neighborhood of σ_{max}. Furthermore, it converges to the model described in V-B, when it is used on a database consisting of uniformly illuminated poses. In fact, in such a database, almost all the genuine poses are distributed in a narrow neighborhood of $\sigma = \sigma_{max}$. As a consequence, the CCDF function $\vartheta_{gen}(\sigma)$ grows very quickly: $\vartheta_{gen}(\sigma) \approx 1$ for $\sigma \geq 0$ and thus $\sigma_{KT} \approx 0$. If those conditions are replaced in Eq. (19), $w(I) \approx 1$ and $FRR^{\star}(\rho, \sigma)$ in Eq. (16) converges to the usual $FRR(\rho, \sigma)$ definition.

5.4 Saving data to the RFID tag

After the enrollment phase is completed, the main poses, the 2DPCA, and the SURF thresholds are determined and tailored to the enrolled subject, and can be stored into the RFID tag. The main poses of the subject are merged in a single superpose (8 bits depth) compressed using a lossless algorithm such as JPEG-LS. By using a single superpose, the compression algorithm can take advantage of similarities among the poses of the same subject. Table I reports storage requirements for the compressed main poses, assuming that the compression algorithm can reduce the size of the data by 20% at least when applied to images of 80x60p, and by 10% for smaller size images. The "2-poses" configurations can be used with common 64 Kbits RFID tags, while the other configurations require the new large-capacity tags such as TegoChip [30] or Xerafy-XL [31].

Instead of storing the main poses, their 2DPCA decomposition vectors might be stored as well. However, as the components in 2DPCA vectors are floating point numbers (32-bit depth), this would force to reduce to 25% the number of vector components, thus decreasing the effectiveness of the 2DPCA stage. Moreover, the SURF algorithm needs the original images. As a consequence, the compressed 8-bit depth images are directly stored into the tag and the 2DPCA face space is rebuilt *on-the-fly* during the authentication phase.

In order to prevent *spoofing attacks*, RFaceID exploits a system based on both symmetric and asymmetric keys, which are specific for the user and for the tag (see Fig.4). It encrypts the face samples (FACEBLK) and a 16-Byte (16B) data block named CFGBLK. The latter contains some configuration data: a 32-bit service code indicating the algorithms used during the authentication phase, a checksum and the two thresholds (represented as floating point numbers) for the 2DPCA and SURF stages. During the enrollment phase, the issuing agency assigns

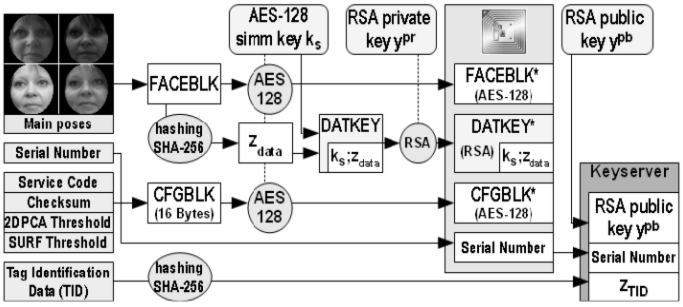

Fig. 4. The RFID encryption scheme used during the enrollment phase.

a serial number s to the user and randomly generates an AES-128 bit simmetric key k_s, a pair of public/private RSA-1024 bit keys $(y_s^{pb}; y_s^{pr})$ and two SHA-256 hash codes. The former, named z_{data}, is built on the biometric content in FACEBLK, and the latter, named z_{TID}, on the serialized *Tag Identification Data* (TID) of the chip. TID is a set of data that uniquely identifies the RFID chip and the manufacturer.

As the encryption of the whole dataset through an asymmetric algorithm (such as RSA) would be too computational expensive [32], in RFaceID FACEBLK and CFGBLK are encrypted using the AES algorithm, thus obtaining the protected data block FACEBLK* and CFGBLK*, and, next, only the 128-bit AES key k_s and the 256-bit hash z_{data} are encrypted through RSA. More in detail, RFaceID creates a DATKEY data block (48B), consisting of the 16B AES key k_s and of the 32B z_{data} hash. Next, the DATKEY block is padded using the OAEP algorithm [33] and encrypted using the RSA private key y_s^{pr}. The result is a protected pair $(rsa(k_s, z_{data}))$ that is contained in a data block named DATKEY* (128B).

Next, the serial number s, the data block DATKEY* and the encrypted biometric data FACEBLK* and CFGBLK* are written into the tag, which is finally delivered to the user, so no centralized database of biometric features is needed.

The private key y_s^{pr} for the user is no longer necessary, so it is erased. The tuple $\{s, z_{TID}, y_s^{pb}\}$ is stored in a *keyserver*, a copy of which is distributed to each authentication gate.

During the authentication phase, RFaceID retrieves y_s^{pb} from the keyserver and uses it for decrypting DATKEY*, thus achieving k_s and z_{data}. The AES key k_s is used for decoding FACEBLK* and CFGBLK* and the authenticity of the resulting FACEBLK is verified by calculating its SHA-256 hash z_{data}^{*} and by comparing it with the z_{data} registered by the agency in DATKEY*. If they do not match, the tag is considered counterfeit and rejected.

The RSA key y_s^{pb} is never transmitted on the RF channel and it cannot be stolen if the keyserver is secure, thus preserving confidentiality of the data belonging to the already registered tags. Furthermore, any hypothetical attacker, who could read the data contained in the keyserver, would never be able to forge a tag in order to be authenticated. In fact, even by attempting

to emulate the serial s and the TID code z_{TID} of a legally registered chip, the impostor would not be able to correctly encode the key k_s and the hash z_{data} of the owned biometric data, because the stolen RSA key y_s^{pb} would be exploitable only for decryption. Moreover, counterfeiting the TID of a RFID tag would not be a trivial task because, to the best of our knowledge, in all Electronic Product Code (EPC) chips [34] currently available the TID memory is locked [35].

6 THE AUTHENTICATION PHASE

When a claimant approaches an authentication gate, the tag is detected and the binary image stored in it is read, together with the user serial number s and with the TID data of the chip. RFaceID generates the hash z_{TID} and looks for the pair $\{s; z_{TID}\}$ in the keyserver. If the claimant is not registered or the tag is counterfeit, the authentication fails at this step.

If the claimant is registered the public key y_s^{pb} is retrieved and used for decryption. The datablock DATKEY* is RSA-decrypted. The AES-128 key k_s contained in the block is used to decrypt FACEBLK* and CFGBLK*, containing the configuration data and the main poses (see Fig.5).

Next the main poses are decompressed, the thresholds ρ^{2dpca} and σ^{surf} are extracted from the configuration data and the 2DPCA face space associated with the custom definition database B_{defdb} containing the poses of the authorized user is rebuilt, using Eq. (5) with $z^{2dpca} = w$. We used 2DPCA because of the small size of its covariance matrix [4], which allows to quickly rebuild its face space. Other algorithms such as 2DLDA [36] or ERE [37] assume that multiple classes of samples, representing different users, are available at the authentication time, and would not be applicable to our case, as only the poses belonging to the authorized user are available[4].

In the meantime, a set of images of the approaching claimant has been acquired by the camera. The area containing the face is detected in each image using the Viola-Jones detector [19], the image is cropped to the area of the face, scaled and preprocessed through the same steps already described in V. Finally, the sample is submitted to the two-stage cascaded authentication module. Up to $l = 10$ images of the face of the claimant are acquired and each image is submitted to the authentication system. If at least one of the images is recognized, the claimant is authenticated, otherwise she is rejected. If the authentication process fails, the tag is put into a black list for $240s$, thus avoiding further immediate attempts[5]. As the second stage may be executed several times (once for each sample acquired) before a user is finally authenticated or rejected, its execution time heavily affects the execution time of the whole system. For this reason, SURF was preferred to the slower SIFT [38] and to the even heavier ROLG [39] for the architecture proposed.

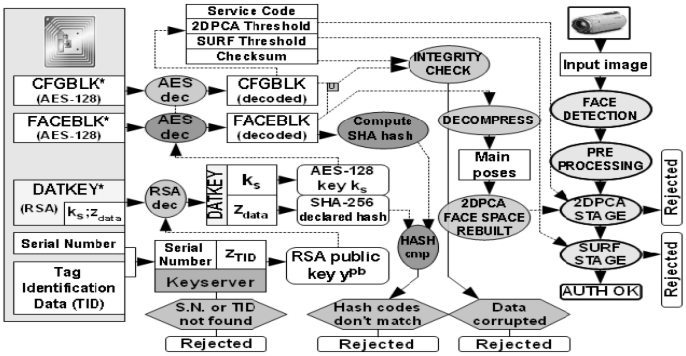

Fig. 5. The workflow scheme for the authentication phase.

6.1 The 2DPCA stage

Each image I that is submitted to the first stage is 2DPCA decomposed using Eq. (5). Then the pose I^* that, among those present in the definition database, is closest to I is determined and the distance $d(I, I^*)$ in the face space is calculated. In this phase, the behavior of the first stage can be modeled by Eq. (9). The first stage accepts I if:

$$g_d(I, B_{defdb}) < \rho^{2dpca} \tag{20}$$

6.2 The SURF stage

An image I that is accepted by the first stage is passed to the second stage for response verification. The system extracts the SURF features of I and of its closest pose I^* in the definition database that has been determined by the previous stage. Then, a matching criterion is applied to the SURF features of I and I^*. A modified version of the second-closest neighbor criterion proposed for SIFT by Lowe in [38], and also used for SURF in [5], is adopted. Given two images I and I^*, we call $S(I)$ the set of SURF features of the image I and $S(I^*)$ the set of SURF features of the image I^*. Each SURF feature is composed of a keypoint $\mathbf{p_I}$ and a keypoint descriptor $\mathbf{v_I}$:

$$s_I = \{\mathbf{p_I} = (x_I, y_I) \in R^2; \mathbf{v_I} \in R^{128}\} \in S(I) \tag{21}$$

$$s_{I^*} = \{\mathbf{p_{I^*}} = (x_{I^*}, y_{I^*}) \in R^2; \mathbf{v_{I^*}} \in R^{128}\} \in S(I^*) \tag{22}$$

4. 2DLDA or ERE would be exploitable by using the model described by Eq. 3. However, this would require a criterion to select an opportune set of *extrasubjects* as a function of the main poses of the user. Such a problem is beyond the scope of this work and is not dealt with in this paper.

5. In theory, the overall probability of a false positive can be calculated through the formula: $P_{FP} = 1 - P_{TN} = 1 - p_{TN}^l = 1 - (1 - p_{FP})^l$ where $p_{FP} = n_{FP}/N_{attempts}$ is the probability of accepting an impostor for a single frame. Unfortunately, p_{FP} is hard to estimate, due to the changing environmental conditions in which the experiments are performed. As a consequence, the system is usually tested on some standard dataset acquired under several different illumination conditions. The fraction (or *rate*) of samples that are recognized over the whole dataset (in a single trial), is assumed as the score for the algorithm (see Sec.VII-E).

For a feature s_I and its keypoint descriptor $\mathbf{v_I}$ of the image I, the software determines the features $s^1_{I^*}$ and $s^2_{I^*}$ of the image I^* that are associated to the closest keypoint descriptor $\mathbf{v^1_{I^*}}$ and the second closest keypoint descriptor $\mathbf{v^2_{I^*}}$ in R^{128}. The features s_I and $s^1_{I^*}$ match if their descriptors comply with the following conditions:

$$|\mathbf{v_I} - \mathbf{v^1_{I^*}}| < \beta^{surf} |\mathbf{v_I} - \mathbf{v^2_{I^*}}| \quad (23)$$

$$|x_I - x^1_{I^*}| < 0.05w \quad (24)$$

$$|y_I - y^1_{I^*}| < 0.05h \quad (25)$$

where β^{surf} is a parameter affecting the discrimination performance of SURF. In [5] a default value of 0.7 is advised, but in Sec.VII-B we deal with the determination of the value that optimizes the performance of the algorithm. If the number of matching features is equal to, or bigger than, the σ^{surf} threshold, the pose is authenticated, otherwise it is rejected.

7 Experimental results

In this work, in order to assess the performance of the proposed face authentication algorithm, a twofold *cross-validation procedure*, based on the *Sanderson's "a priori performance type A" protocol* [28], was developed. It is composed of three testing configurations, using $n_{mainposes} = 8, 4, 2$ respectively, for each authorized user. The tests were performed on gray-scale images (poses) of size 320x240p, 160x120p, 80x60p, 40x30p, 20x15p, and 14x10p.

7.1 The testing super-database

In the testing protocol the poses of 43 people (24 men and 19 women) were considered, taken from the VidTIMIT database (VDT) [40]. For each subject, the VidTIMIT database provides $n_{totgrab}$ poses divided in three sessions, registered with a few days delay:

$$n_{totgrab} = n_{session1} + n_{session2} + n_{session3}. \quad (26)$$

The first session is used to produce the $n_{mainposes}$ poses representative of the subject. The $n_{session1}$ poses are 2DPCA-decomposed and $n_{mainposes} < n_{session1}$ items are selected following the procedure described in Section V.

According to the Sanderson protocol, two phases are needed. In the first phase the poses of the second session are used to build the *threshold optimization database* (see V-B) and the poses of the subject belonging to the session 3 are used to build a "single-row archive" named *testing database* with cardinality $k_{testingdb} = n_{session3}$.

The *testing database* is then extended with poses belonging to impostors, thus building a "face table" composed of $r_{testingsdb}$ rows and $k_{testingsdb}$ columns, named *testing super-database*. The $n_{session3}$ poses of the testing database are saved in the row 0, so that:

$$k_{testingdb} = k_{testingsdb} = n_{session3} \quad (27)$$

The other rows of the super-database are reserved to the poses of $r_{testingsdb} - 1$ impostors. During this first phase, the *threshold optimization super-database* is used to determine the optimal threshold pair $(\rho^{2dpca}, \sigma^{surf})_1$ and the authentication algorithm is then applied to the *testing super-database*, in order to obtain a set of values $(FRR, FAR)_1$ for the subject.

In the second phase, the role of sessions 2 and 3 are exchanged. A new *threshold optimization super-database*, containing poses coming from session 3 of the VidTIMIT database, is defined and used to determine a second optimal threshold pair $(\rho^{2dpca}, \sigma^{surf})_2$. Then the authentication algorithm is applied to a second *testing superdatabase*, containing poses coming from session 2 of the VidTIMIT database, and a second set of values $(FRR, FAR)_2$ is thus obtained.

The performance values, for the subject, in both phases, are finally averaged and then those user-specific values are again averaged over all the subjects, in order to obtain an overall performance assessment for each resolution and for each value assigned to $n_{mainposes}$. In the tests, the threshold optimization and the testing superdatabases have the same number of rows and columns. The poses of 35 true claimants and 8 impostors (chosen as stated in [28]) are used. For each subject, 102 poses from session 2 and 102 poses from session 3, randomly chosen before starting the experiment, are used. For each true claimant the system simulates $2 * 102 = 204$ legal access attempts and $2 * 8 * 102 = 1632$ violation attempts made by impostors.

7.2 The optimal β^{surf} parameter

As stated in Section VI-B, the parameter β^{surf} affects the discrimination performance of the SURF algorithm. In [5] the value 0.7 for β is recommended, however, in this work a series of tests with 8,4 and 2 poses for each tag were performed, in order to assess to what extent a change in the value of β significantly affects the performance of the second stage.

Fig. 6 shows the results of the tests. As the aim here is minimizing both FAR and FRR, the FAR and FRR (in percentage) as a function of β^{surf} are plotted. The optimal value of β^{surf} corresponds to the point of the graph that is closest to the origin. By applying the described criterion, $\{0.6, 0.4, 0.5\}$ were selected as the optimal values of β^{surf} for the configurations with 8,4, and 2 poses for tag, respectively.

7.3 Effects of the image size on the performance

In this work we also evaluate how the performance of the proposed approach depends on the size of the poses. The tests show that the average recognition rate (RR) is only slightly affected by changes in the size of the poses (see Fig. 7).

In particular, RR is always over 97% for all pose sizes larger then 20x15p (with a variation lower than 2.5% for the same sizes), whereas FAR is always below 0.08% for

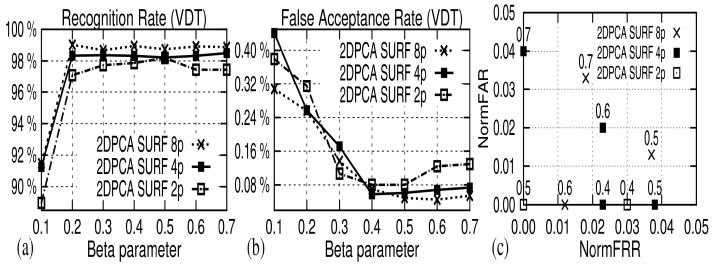

Fig. 6. Measured performance of RFaceID varying β^{surf} parameter. (a) Recognition rate. (b) False acceptance rate. (c) Normalized values of (FRR, FAR) pairs (for clearness, only the points in the zone [0.05x0.05] are shown).

TABLE 2
Measured performance for VidTIMIT, ExYaleB and MOBIO databases

	algo	8pRR	8pFAR	4pRR	4pFAR	2pRR	2pFAR
VDT 80x60p	RFaceID	98.9%	0.04%	98.3%	0.05%	98.1%	0.08%
VDT 40x30p	RFaceID	98.2%	0.19%	97.9%	0.41%	97.8%	0.38%
EXY 80x60p	RFaceID	57.0%	2.30%	47.9%	2.58%	31.5%	1.65%
EXY 40x30p	RFaceID	55.2%	5.22%	42.6%	3.98%	31.5%	4.24%
MOB 80x60p	RFaceID	86.4%	0.22%	73.0%	0.28%	59.4%	0.33%
MOB 40x30p	RFaceID	85.9%	1.27%	74.6%	2.26%	64.4%	3.74%
VDT 80x60p	Disparity	100.0%	0.87%	99.9%	0.86%	99.7%	0.89%
VDT 40x30p	Disparity	99.7%	1.11%	99.8%	1.10%	99.0%	1.27%
EXY 80x60p	VSFR	54.8%	2.65%	39.6%	1.75%	25.8%	1.77%
EXY 40x30p	VSFR	29.4%	1.53%	24.9%	1.68%	19.1%	1.86%

all sizes larger than 40x30p. Furthermore, RR increases and FAR decreases when the number of poses increases. As a result, the proposed approach is very reliable for pose sizes equal to or larger than 80x60p, and can still be satisfactorily applied to poses of 40x30p when the security constraints are not very restrictive.

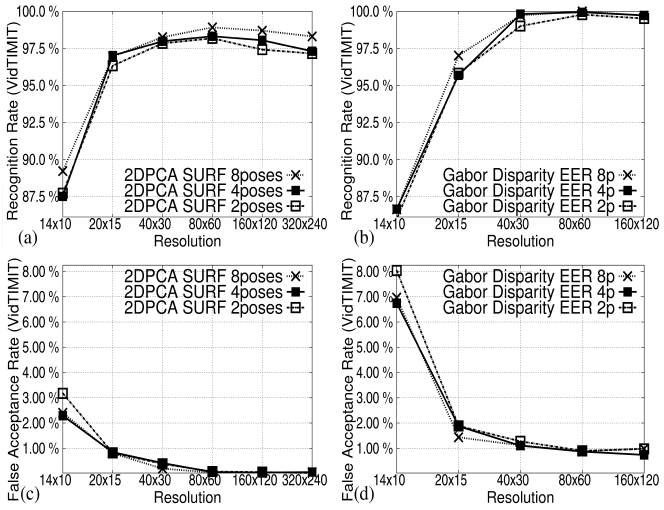

Fig. 7. Measured performance for RFaceID and for Gabor Disparity on VidTIMIT database with varying sample size. (a,b) Recognition rate. (c,d) False acceptance rate.

7.4 Comparison with other authentication algorithms

In the next experiment RFaceID was compared with *Gabor Disparity* [6], a face authentication system that gained one of the highest scores in the *Face Recognition Evaluation in Mobile Environment* contest[6] [42]. The Sanderson protocol was applied to the VidTIMIT database and the Gabor Disparity recognition algorithm. According to this algorithm, a grid of 23x16 points for the resolutions of 160x120p, 80x60p, 40x30p and a grid of 7x5 points for the resolutions of 20x15p and 14x10p is applied to the image, using the position of the eyes (detected through a Viola-Jones detector [19]) as a reference. For each point $\bar{x}$ of the grid, a vector of 40 complex numbers (named *GaborJet*) is created [6]. The j-th component $\mathbf{J}(\bar{x}))_j$ of the jet in $\bar{x}$ is calculated through the convolution product $(\mathbf{J}(\bar{x}))_j = (\mathbf{I} \star \mathbf{\Psi}_{\bar{k}_j})(\bar{x})$, where $\mathbf{\Psi}_{\bar{k}_j}$ is the Gabor wavelet calculated for a vector $\bar{k}_j$ which contains information about scale and orientation. Finally, the components of the GaborJet are L_2 normalized [43]. The distance between two images I and I^* can be calculated according to (28):

$$d(I, I^*) = \left[1 + \frac{1}{g_x g_y} \sum_{n_x=1}^{g_x} \sum_{n_y=1}^{g_y} S(n_x, n_y)\right]^{-1} \quad (28)$$

where $S(n_x, n_y)$ is the similarity between the GaborJets related to the points in I and I^* at the position (n_x, n_y) of the grid. In the performed tests, the S_{n+C} similarity function was used, as it achieved the best results in [6]:

$$S_{n+C}(n_x, n_y) = S_n(n_x, n_y) + S_C(n_x, n_y) \quad (29)$$

where S_C is the Canberra similarity [43] and S_n is the *new similarity based on the Disparity vector* [6] between two Jets in the position (n_x, n_y) of the grid.

The Sanderson protocol was applied to all $k_{throptdb}$ poses. The distances were calculated through (28), thus generating the ROC curve. As Disparity is a single-stage algorithm, the threshold ρ^{Gabor} was chosen according to the criterion $FRR(\rho^{Gabor}) = FAR(\rho^{Gabor})$. Next, ρ^{Gabor} was used for the testing protocol. Fig. 7 shows that Disparity obtained a slightly better RR than RFaceID, but at the cost of a very higher FAR at the same sizes. For example, at 80x60p (8 poses), Disparity achieved a RR that is 1.01 times the one measured for RFaceID, whereas the FAR is 21.75 times the one measured for RFaceID. As the priority for an authentication system is the rejection of impostors, RFaceID outperforms Disparity for the proposed task.

Moreover, it can be shown that RFaceID outperforms a single-stage 2DPCA system when both work on the same dataset under the same operational constrains [7].

7.5 Effects of illumination changes on performance

The VDT database contains poses acquired under controlled conditions. However, the real working conditions

6. Despite other algorithms showed better scores during the competition, they are not comparable directly to RFaceID being either commercial (such as GRADIANT, http://www.gradiant.org/en/research-lines/human-sensing/facial-processing.html) or based on color images (such as MR-PCA[41]).

7. Due to the limitation of the page number, those further results are reported in a separate document, uploaded while submitting this paper, that will also be made available to the interested Readers.

for an authentication system are very different. Variations in illumination or face orientation are frequent and can considerably affect the recognition effectiveness. A second issue is the generalizability of the β^{surf} optimal values in VII-B. This should be verified by applying to a second database the settings found for the first one [44] and checking that the performance still remains acceptable. In order to assess the behavior of RFaceID under variable illumination, some tests on the *Extended Yale B* (ExYaleB) [45; 46] and on the *MOBIO* [47] databases were run. The ExYaleB database contains the poses of 28 subjects under 9 orientations and 64 illumination conditions (576 poses for subject). The MOBIO database contains the poses of 150 subjects acquired under different conditions of orientation and illumination by a mobile phone camera. Both were reorganized for the Sanderson protocol.

The first experiment (FR) measured the performance, under different illumination conditions and frontal orientation, on the P00 subset of ExYaleB, containing $n_{enrolldb} = 64$ poses for subject. From this set, three different numbers $n_{mainposes}$ of sample poses representative of each subject were chosen and the experiment was repeated for each number (8,4,2). At each iteration of the experiment, the remaining $n_{enrolldb} - n_{mainposes}$ items were randomly reordered and then divided in two equal-length sets, thus selecting the $n_{throptdb}$ and $n_{testingdb}$ samples for the threshold optimization and testing database (see Sec. V-B and VII-A). Thus, for each subject, $n_{session2} = n_{session3} = \{28, 30, 31\}$ poses (8p, 4p, 2p main poses configurations respectively). For each one of the 28 claimants, the other 27 were considered as impostors, thus performing at least $2 * 28 = 56$ legal access attempts and $27 * 2 * 28 = 1512$ violation attempts per subject.

The second experiment (OR) dealt with both changes in illumination and orientation, thus considering all the $n_{enrolldb} = 576$ poses per subject. From this set, the $n_{mainposes} = \{8, 4\}$ main poses were extracted, reserving respectively $n^{frontal}_{mainposes} = \{3, 2\}$ items for frontally illuminated samples. The remaining $n_{enrolldb} - n_{mainposes}$ ones were reordered and split in such a way as to obtain $n_{session2} = n_{session3} = \{284, 286\}$ (i.e., $8p^{OR}, 4p^{OR}$ configurations respectively). For each subject, at least $2*284 = 568$ authorized access attempts, and $27*2*284 = 15336$ violation attempts were simulated.

All tests were performed at the resolutions of 14x10, 20x15, 40x30, 80x60, 160x120 and 320x240 pixels. For the SURF stage we used the average of the optimal values previously found for the VidTIMIT database for the 8, 4 and 2 poses configurations ($\beta^{surf} = 0.5$). The model described in V-C was adopted for threshold optimization, with $\sigma_T = 4$.

Some preliminary tests were performed on ExYaleB at a resolution of 80x60p in order to optimize the *cut-off frequency* of the homomorphic filter defined by Eq.(15). Fig.8 shows that for the FR configurations $(8p, 4p, 2p)$ the RR is maximized choosing $\mu = 0.25$. The next step

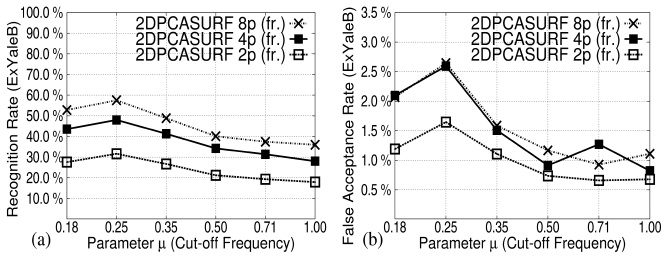

Fig. 8. Measured performance of RFaceID varying cut-off frequency of the homomorphic prefilter for SURF (Extended Yale B Database, res. 80x60p.) (a) Recognition rate. (b) False acceptance rate.

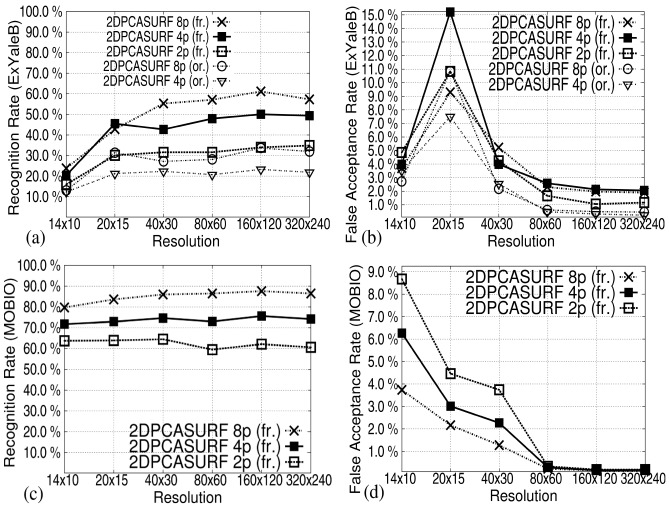

Fig. 9. Measured performance for 2DPCA-SURF on the ExYaleB and MOBIO databases obtained by varying the size of the poses. (a,c) Recognition rate. (b,d) False acceptance rate.

was to measure the performance of the system using the optimal parameter μ just determined and changing the resolution. The results are shown in Fig. 9 and show that in these tests RFaceID optimized RR and FAR differently. If the ratio $q = n_{defdb}/n_{throptsdb}$ is favourable (several samples and a small dataset, as in 8poses/4poses FR tests), the algorithm mostly optimizes RR (that was between 43% and 61% for all sizes larger than 40x30p), and this entails a higher FAR (between 1.5% and 2.5% for all sizes larger than 80x60p). Conversely, if q is unfavourable (few samples and a large training set, as in 2 poses FR test or in OR tests), RFaceID partially gives up optimizing the RR, thus concentrating on FAR reduction. In this case, for all sizes larger than 40x30p, the obtained RR was between 20.6% and 33.6%, whereas FAR was less than 1.65%. These RR values are comparable to those reported, for the same FAR, by other recent works applied to the ExYale B database [48][49]. Moreover, RFaceID, by exploiting a second stage based on SURF, outperforms at the lowest resolutions the analogous system exploiting SIFT, when both applied to the same database with the same parameters.[8]

8. Due to the limitation of the page number, those further results are reported in a separate document, uploaded while submitting this paper, that will also be made available to the interested Readers.

Finally, a third experiment on the MOBIO database was performed, in order to test the generalizability of the optimal value of μ on a different dataset. For each one of the 150 subjects, $n_{enrolldb} = 190$ poses were randomly chosen. Next, the set of $n_{mainposes}$={8,4,2} poses was extracted, thus generating the two subsets of $n_{session2} = n_{session3} = \{91, 93, 94\}$ samples for threshold optimization and testing. For each subject, the other 149 were considered as impostors, thus performing at least $2*91 = 182$ legal access attempts and $149*2*91 = 27118$ violation attempts per subject. Fig. 9 shows that on the MOBIO dataset RFaceID performed even better than on ExYaleB, thus achieving a RR between 87.57% and 59.47% and a FAR between 0.33% and 0.13% for all sizes larger than 40x30p. Moreover, at a resolution of 80x60p, the *Half Total Error Rate* $HTER = (FRR + FAR)/2$ measured for 8p, 4p and 2p configurations was respectively 6.89%, 13.63% and 20.43%. These values are better than, or comparable with, those reported for the same dataset in [42][50].

8 RFaceID vs VisilabFaceRec

A two-stage authentication system, named *VisilabFaceRec* (VSFR) and based on a 2DPCA-SIFT combination and RFID tags, was presented in [7]. Though RFaceID shares some important architectural elements with VisilabFaceRec, it also features a number of fundamental differences:

- *The algorithm for threshold optimization*. VSFR chooses a single point $p(\rho)$ on the primary ROC curve minimizing the cost function:

$$c(\rho) = FRR(\rho) + \frac{C_{FAR}}{C_{FRR}} FAR(\rho) = FRR(\rho) + \upsilon FAR(\rho) \quad (30)$$

  Starting from $p(\rho)$, the secondary ROC curve (based on SIFT) is drawn and the point nearest the origin of the axes (FRR, FAR) is chosen (thus determining the pair $(\rho^{2dpca}, \sigma^{sift})$). As a consequence, VSFR performance strongly depends on the choice of the cost ratio $\upsilon = C_{FAR}/C_{FRR}$, which has to be manually selected "a priori" by the operator. A better optimization choice would have been tailoring the parameter υ for each user on the basis of its sample set acquired during the enrollment phase. Unfortunately, this is not possible in VSFR. The best optimization consists in selecting a value for υ that is heuristically and uniquely determined *for all users* through a simulation. Conversely, RFaceID exploits the *BestPoint model* (see Sec. V-B) that optimizes automatically the cost ratio for each user, in order to obtain the best compromise between FRR and FAR.
- *The algorithm for illumination compensation*. VSFR uses a pair of Chen's filters [51], placed before the 2DPCA and SIFT stages. Chen's compensator firstly provides for a sample transformation into logarithmic DCT domain (LDCT) and, next, the LDCT coefficient $C(0;0)$ is set in such a way as to normalize the average brightness of the image, and low frequency LDCT coefficients $C(u,v)$ such as $|u+v| < D_{dis}$ are set to zero (hence, D_{dis} is the main parameter of the filter). Finally, the inverse DCT is applied to the image, and the latter is then back-transformed into the logarithmic domain (Chen's compensator does not provide for a back-transformation in linear domain). As a consequence, VSFR performance strongly depends on the choice of two parameters (D_{dis}^{2dpca} and D_{dis}^{sift}) that have to be manually preset before using the system. Such an architecture may misbehave in presence of samples affected by non-uniform illumination changes, as the Chen's filter can lead the 2DPCA stage to choose the *wrong* sample on the first row of B_{defdb}, thus affecting the effectiveness of the SIFT stage. Conversely, RFaceID does not exploit any filter before the first stage (leveraging on samples grabbed under different illumination conditions and stored into the tag), and applies only a homomorphic filter (see Eq.15) before the SURF stage.
- *The algorithm used by the 2nd stage*. VSFR uses a SIFT-based classifier [38], whereas RFaceID exploits SURF.

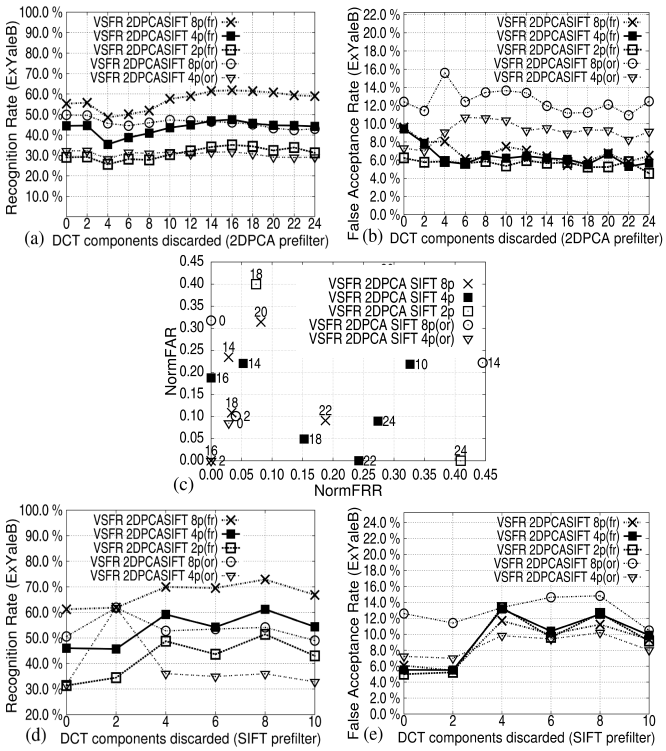

Fig. 10. Performance for VisilabFaceRec on the ExYaleB db (res. 80x60p). (a-b) RR and FAR measured varying D_{dis}^{2dpca}. (c) Normalized values of (FRR, FAR) pairs measured varying D_{dis}^{2dpca} (only the points in the area [0.45x0.45] are shown). (d-e) RR and FAR measured varying D_{dis}^{sift}.

This section is aimed at showing that RFaceID outperforms VSFR on Extended Yale B database, even when all

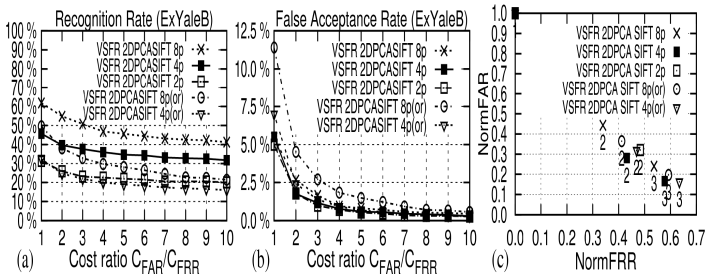

Fig. 11. Measured performance for VisilabFaceRec on the Extended Yale B database obtained varying the cost ratio υ (res. 80x60p.). (a) Recognition rate. (b) False acceptance rate. (c) Normalized values of (FRR, FAR) pairs.

three working parameters $(D_{dis}^{2dpca}, D_{dis}^{sift}, \upsilon)$ of the latter are optimized. As a first experiment, some tests were done varying the VSFR parameter D_{dis}^{2dpca} and setting up $D_{dis}^{sift} = 2$, $\upsilon = C_{FAR}/C_{FRR} = 1$ and $\beta^{sift} = 0.5$. The results are shown in Fig.10. The optimal values for D_{dis}^{2dpca} can be found normalizing both FRR and FAR in the plane [0,1]x[0,1] (see Fig.10(c)) and choosing the points nearest the origin of the axes. In this experiment, the values found through this method were $D_{dis}^{2dpca} = \{16, 18, 18, 2, 2\}$ respectively for the VSFR configurations $(8p, 4p, 2p, 8p^{OR}, 4p^{OR})$.

As a second experiment, a number of tests were performed by varying D_{dis}^{sift}, setting up $\upsilon = 1$, $\beta^{sift} = 0.5$ and using the optimized values for D_{dis}^{2dpca} just found. Fig.10(d)10(e) show that the SIFT stage is negatively affected by the presence of Chen's compensator. For $D_{dis}^{sift} > 2$ the system shows a better RR, but at the cost of high FAR. In this case we decided to minimize only the FAR, and the optimal values chosen for the VSFR configurations $(8p, 4p, 2p, 8p^{OR}, 4p^{OR})$ were respectively $D_{dis}^{sift} = \{2, 2, 0, 2, 2\}$.

The next experiment dealt with the optimization of the cost ratio υ. The simulations were carried out using the values just found for D_{dis}^{2dpca} and D_{dis}^{sift}. Fig. 11 shows that FRR and FAR can be minimized, for all configurations, choosing $\upsilon = 2$.

The previous tests were all performed at a resolution of 80x60p. The final step was to measure the performance of VSFR by varying the resolution and using the optimized parameters just found. Fig.12(a)12(b) show the measured performance, and Fig.12(c)12(d) show a comparison with RFaceID (positive values indicate a better behavior of RFaceID).

In the FR tests, RFaceID achieved better RR values at all resolutions larger than 20x15p. The new system was very effective at the intermediate resolutions (40x30p and 80x60p) with a peak difference of 25.8% at 40x30p. It achieved also better FAR, but only at the resolutions of 160x120p and 320x240p, with gains up to 0.9%. Conversely, at the intermediate resolutions the FAR were slightly worse, with differences up to 2.3% at 40x30p (a result still acceptable being large the RR gain at the same size). In the OR tests the two systems performed differently. At intermediate and largest resolutions, RFaceID obtained worse RR values than VSFR (differences up to 9.8%), but better FAR values at (differences up to 4.31%). In conclusion, when used at resolutions larger than 20x15p, RFaceID always outperforms VSFR, at least in one of the two score indexes (RR or FAR).

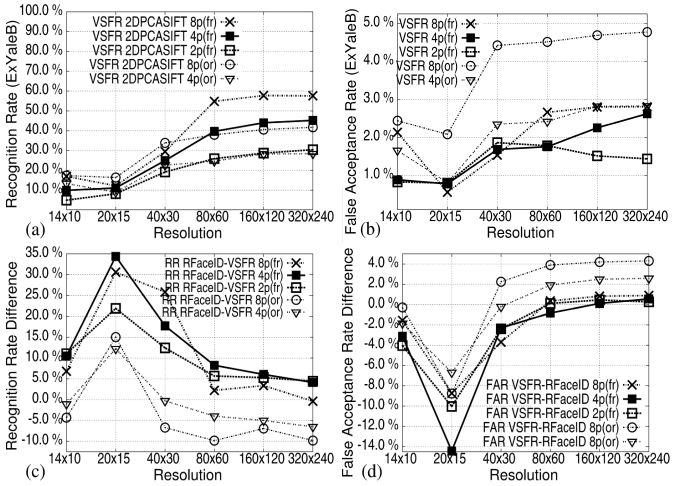

Fig. 12. Measured performance for VisilabFaceRec on the Extended Yale B database obtained varying the sizes of the poses. (a) Recognition rate. (b) False acceptance rate. (c) RR difference between RFaceID and VisilabFaceRec. (d) FAR difference between VisilabFaceRec and RFaceID.

8.1 Authentication times

Tab. III shows the minimum authentication times t_{min} measured for RFaceID, VSFR and Disparity using an ASRock E350M1 board (cpu AMD Fusion dual core 1.6GHz) and an Inpinji Speedway RFID reader. This time is minimum when the user is immediately authorized after the RFID is read. In this case, $t_{min} = t_{rfid} + t_{frec}$, where t_{rfid} is the time required for tag reading (at a distance of 50 cm) and decompression, and t_{frec} is the time required for 2DPCA face space rebuilding, 2DPCA matching, and SURF (or SIFT for VSFR) matching with homomorphic filtering. For Disparity, t_{frec} is the time required for GaborJet calculation and comparison through Canberra similarity. The tests show that RFaceID experienced the lowest recognition times.

9 Conclusion and Future Works

This paper presented RFaceID, a multifactor authentication system for access control of services and restricted areas, which combines face recognition and token-based authentication for the sake of improved accuracy, reliability, and privacy. The system was specifically devised to work with very low resolution images, thus allowing the storage of the sensitive biometrics user data (e.g., the face images) directly into the RFID tag, without the need for a centralized biometric database. To the best of our knowledge, RFaceID performs better, and with lower resolution face images, than the other approaches in the

TABLE 3
Authentication times $[s]$

		8poses 40x30p	8poses 80x60p	4poses 40x30p	4poses 80x60p	2poses 40x30p	2poses 80x60p
avg(t_{rfid})	-	7.296	25.434	3.698	12.732	1.947	6.428
stddev(t_{rfid})	-	0.403	1.312	0.202	0.656	0.115	0.333
avg(t_{frec})	RFaceID	0.545	0.849	0.405	0.597	0.316	0.367
stddev(t_{frec})	RFaceID	0.022	0.146	0.012	0.036	0.011	0.036
avg(t_{frec})	VSFR	0.734	1.720	0.477	1.032	0.343	0.537
stddev(t_{frec})	VSFR	0.050	0.128	0.028	0.071	0.013	0.028
avg(t_{frec})	Disparity	9.781	21.140	4.878	11.431	2.634	5.417
stddev(t_{frec})	Disparity	0.113	0.803	0.216	0.392	0.040	0.142
t_{min}	RFaceID	7.841	26.283	4.103	13.329	2.263	6.795
t_{min}	VSFR	8.031	27.154	4.175	13.764	2.291	6.966
t_{min}	Disparity	17.078	46.574	8.577	24.163	4.582	11.845

literature integrating RFID tags and biometric authentication [12–17]. Moreover, thanks to its novel BestPoint model, RFaceID achieved better FAR/FRR ratio than other state-of-the-art algorithms, such as Gabor Disparity [42]. Furthermore, despite the low execution times and the small amount of data available, RFaceID is able to ensure an acceptable recognition rate together with a low false acceptance rate even in presence of strong variations in the aspect of the authorized user due to illumination changes. Future work will be aimed to improve the recognition rate on ExYaleB database without increasing the FAR, by using the preprocessor proposed in [52] before the 2DPCA-stage or by using a novel 2DLDA-based [36] first matcher devised to attenuate only the *intra-class* component of the large-scale band variations [53] in the LDCT feature space.

REFERENCES

[1] A. K. Jain, A. Ross, S. Prabhakar, *et al.*, "An Introduction to Biometric Recognition," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 14, Jan. 2004.

[2] K. Nasrollahi *et al.*, "Extracting a Good Quality Frontal Face Image from a Low-Resolution Video Sequence," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 21, Oct. 2011.

[3] European Commission: Art. 29 Working Party, "Working document on biometrics 00720/12/EN WP 193," 2012.

[4] J. Yang, "Two dimensional PCA: a new approach to appearance based face representation and recognition," *IEEE Trans. on Pattern Anal. and Machine Intell.*, Jan 2004.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, no. 110, pp. 346–359, 2008.

[6] M. Gunther, D. Haufe, and R. Wurtz, "Face Recognition with Disparity Corrected Gabor Phase Differences," pp. 411–418, Springer-Verlag, 2012.

[7] F. Battaglia, G. Iannizzotto, and L. Lo Bello, "A biometric authentication system based on face recognition and RFID tags," *Mondo Digitale*, vol. 13, Feb 2014.

[8] A. Pentland and M. Turk, "Face recognition using Eigenfaces," *IEEE Conf. CVPR '91*, pp. 586–591, 1991.

[9] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. on PAMI*, vol. 19, pp. 711–720, July 1997.

[10] A. T. B. Jin *et al.*, "Biohashing: two factor authentication featuring fingerprint data and tokenised random number," *Pattern Recognition*, vol. 37, no. 11, pp. 2245–2255, 2004.

[11] D.-H. Seo, J.-M. Baek, M.-Y. Yoon, J.-W. Choi, *et al.*, "Improved authentication scheme using facial recognition scheme based on RFID tag," *Int. Conf. on Multimedia Information Networking and Security*, 2010.

[12] D.-G. Min, J.-W. Kim, and M.-S. Jun, "The entrance authentication system in real-time using face extraction and the rfid tag," in *Ubiquitous Computing and Multimedia Applications (UCMA), 2011 Int. Conf. on*, pp. 20–24, 2011.

[13] B.-Z. Jing, D. Yeung, W. Ng, H.-L. Ding, D.-L. Wu, Q.-C. Wang, and J.-C. Li, "Rfid access authorization by face recognition," in *Machine Learning and Cybernetics, 2009 Int. Conf. on*, vol. 1, pp. 302–307, 2009.

[14] T. Nguyen, L. D. Quang, N. C. Van, L. T. Thanh, T. M. Hoang, and T. de Souza-Daw, "An efficient and reliable human resource management system based on a hybrid of face authentication and RFID technology," in *Communications and Electronics (ICCE), 2012 Fourth Int. Conf. on*, pp. 333–338, 2012.

[15] G. Jong, P. Peng, *et al.*, "Multi-recognition combined security system for intelligent car electronics," *Int. Journal of Innovative Computing, Information and Control*, vol. 8, Apr. 2012.

[16] A. Affandi, M. Awedh, *et al.*, "RFID and Face Recognition Based Security and Access Control System," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 2, Nov 2013.

[17] X.-L. Meng, Z.-W. Song, and X.-Y. Li, "RFID-based security authentication system based on a novel face-recognition structure," *WASE Int. Conf. on Information Engineering*, 2010.

[18] P. A. Flach, "The geometry of ROC space: Understanding machine learning metrics through ROC isometrics.," in *Int. Conf. on Machine Learning*, pp. 194–201, AAAI Press, 2003.

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518, 2001.

[20] P. Laytner, C. Ling, and Q. Xiao, "Robust face detection from still images," in *2014 IEEE Symp. on Comp. Intell. in Biometrics and Identity Management (CIBIM)*, pp. 76–80, Dec 2014.

[21] Moura, E.S., Gomes, H.M., and De Carvalho, J.M., "An Improved Face Verification Approach Based on Speedup Robust Features and Pairwise Matching," *26th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 362–369, Aug. 2013.

[22] S. Lloyd, "Least squares quantization in pcm," *IEEE Trans. Inf. Theor.*, vol. 28, pp. 129–137, Sept. 2006.

[23] N. Vaswani and R. Chellappa, "Principal Components Null Space Analysis for Image and Video Classification," *IEEE Trans. on Image Proc.*, vol. 15, Jun 2006.

[24] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 23, no. 2, 2001.

[25] N. Poh and J. Kittler, "A Unified Framework For Biometric Expert Fusion Incorporating Quality Measures," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 34, Jan 2012.

[26] B. Ulery, A. Hicklin, C. Watson, *et al.*, "Studies of Biometric Fusion," Tech. Rep. IR 7346, Sept. 2006.

[27] G. L. Marcialis, F. Roli, and L. Didaci, "Personal identity verification by serial fusion of fingerprint and face matchers," *Pattern Recognition*, no. 42, pp. 2807–2817, 2009.

[28] C. Sanderson, "The VidTIMIT Database," tech. rep., IDIAP Communication, Aug 2002.

[29] J. Short, J. Kittler, *et al.*, "A Comparison of Photometric Normalisation Algorithms for Face Verification," *Sixth IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2004.

[30] "Tegochip." Available: http://www.tegoinc.com/products.

[31] "Xerafy XL." Available: http://www.xerafy.com/.

[32] B. Padmavathi and S. Ranjitha Kumari, "A Survey on performance analysis of DES, AES and RSA algorithm along with LSB substitution technique," *Intnl. Journ. of Science and Research*, vol. 2, Apr. 2013.
[33] H. Agrawal, "MATLAB implementation, analysis & comparison of some RSA family cryptosystems," *2010 IEEE International Conference on Computational Intelligence and Computing Research*, pp. 1–3, Dec 2010.
[34] "EPCGlobal." Available: http://www.gs1.org/epcglobal.
[35] M. Lehtonen, A. Ruhanen, F. Michahelles, and E. Fleisch, "Serialized TID numbers: A headache or a blessing for RFID crackers ?," *IEEE Int. Conf. on RFID*, pp. 233–240, 2009.
[36] M. Li and B. Yuan, "2D-LDA: A statistical linear discriminant analysis for image matrix," *Pattern Recognition Letters*, vol. 26, no. 5, pp. 527–532, 2005.
[37] X. Jiang, B. Mandal, and A. Kot, "EigenFeatures Regularization and Extraction in Face Recognition," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 30, March 2008.
[38] D.Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, Nov 2004.
[39] Z. Miao and X. Jiang, "Interest point detection using rank order LoG filter," *Pattern Recognition*, vol. 46, Nov 2013.
[40] C. Sanderson, *VidTIMIT Database*, vol. Biometric Person Recognition: Face, Speech and Fusion. VDM-Verlag, 2008.
[41] V. Struc, J. Z. Gros, *et al.*, "Exploiting representation plurality for robust and efficient face recognition," *22nd International Electrotechnical and Computer Science Conference*, 2013.
[42] M. Gunther, A. C. Pazo, C. Ding, *et al.*, "The 2013 Face Recognition Evaluation in Mobile Environment," *Biometrics (ICB), Int. Conf. on Biometrics Compendium*, pp. 1–7, Jun 2013.
[43] D. Gonzalez-Jimenez, M. Bicego, J. W. H. Tangelder, B. A. M. Schouten, *et al.*, "Distance Measures for Gabor Jets Based Face Authentication: a Comparative Evaluation," *Intnl. Conf. on Advances in Biometrics*, Gen 2007.
[44] Y.-L. Tian, T. Kanade, *et al.*, "Recognizing action units for facial expression analysis," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 23, pp. 97–115, Feb 2001.
[45] "The Extended Yale Database B." Available online: http://vision.ucsd.edu/ leekc/ExtYale-Database/ExtYaleB.html.
[46] K. Lee, J. Ho, and D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 27-5, pp. 684–698, 2005.
[47] C. McCool, S. Marcel, *et al.*, "Bi-Modal Person Recognition on a Mobile Phone: using mobile phone data," *IEEE ICME Workshop on Hot Topics in Mobile Multimedia*, 2012.
[48] X. Xie, "Illumination preprocessing for face images based on empirical mode decomposition," *Signal Processing, Elsevier*, vol. 103, pp. 250–257, Oct 2014.
[49] X. Xie, W.-S. Zheng, *et al.*, "Normalization of Face Illumination Based on Large-and Small-Scale Features," *IEEE Trans. on Image Proc.*, vol. 20, pp. 1807–1821, Jun 2011.
[50] S. Mau, S. Chen, *et al.*, "Video face matching using subset selection and clustering of probabilistic Multi-Region Histograms," in *25th Int. Conf. of Image and Vision Computing New Zealand (IVCNZ), 2010*, pp. 1–8, Nov 2010.
[51] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 2, 2006.
[52] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. on Image Processing*, vol. 19, pp. 1635–1650, June 2010.
[53] H. Han, S. Shan, X. Chen, and W. Gao, "A comparative study on illumination preprocessing in face recognition," *Pattern Recognition*, vol. 46, no. 6, pp. 1691–1699, 2013.

Filippo Battaglia received the M.S. degree in electronics engineering from University of Messina, in 2008 and the PhD degree in information engineering from University of Reggio Calabria, Italy, in 2013. His research interests include computer vision, IoT/M2M communication and sensor networks.

Giancarlo Iannizzotto is Associate professor with tenure at the Department of Cognitive Science, Education and Cultural Studies of the University of Messina, Italy. His research interests include Computer Vision and its application to face recognition and Human-Computer Interaction.

Lucia Lo Bello is Associate Professor at the University of Catania, Italy. She received the M.S. degree in electronic engineering in 1994 and the Ph.D. degree in computer engineering in 1998. She authored more than 140 papers in the area of real-time systems, industrial automation and sensor networks.