Computer Vision and Image Understanding 117 (2013) 1512–1525

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms ☆

Javier Galbally [a,*], Arun Ross [b], Marta Gomez-Barrero [a], Julian Fierrez [a], Javier Ortega-Garcia [a]

[a] Biometric Recognition Group – ATVS, EPS, Universidad Autonoma de Madrid, C/Francisco Tomas y Valiente 11, 28049 Madrid, Spain
[b] Integrated Pattern Recognition and Biometrics Lab (i-PRoBe), Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 26 June 2012
Accepted 4 June 2013
Available online 18 June 2013

Keywords:
Image reconstruction
Biometric systems
Iris recognition
Binary iriscodes
Security
Privacy

ABSTRACT

A binary iriscode is a very compact representation of an iris image. For a long time it was assumed that the iriscode did not contain enough information to allow for the reconstruction of the original iris. The present work proposes a novel probabilistic approach based on genetic algorithms to reconstruct iris images from binary templates and analyzes the similarity between the reconstructed synthetic iris image and the original one. The performance of the reconstruction technique is assessed by empirically estimating the probability of successfully matching the synthesized iris image against its true counterpart using a commercial matcher. The experimental results indicate that the reconstructed images look reasonably realistic. While a human expert may not be easily deceived by them, they can successfully deceive a commercial matcher. Furthermore, since the proposed methodology is able to synthesize multiple iris images from a single iriscode, it has other potential applications including privacy enhancement of iris-based systems.

1. Introduction

Biometrics is the science of establishing human identity based on the physical and behavioral attributes of an individual such as fingerprints, face, iris or voice. Since biometric traits are inherently associated with a person, they offer substantial advantages over traditional human authentication schemes based on passwords and ID cards [1,2]. Consequently, the deployment of biometric systems in identity management and access control applications is on the increase.

A classical biometric system acquires the biometric trait of an individual, extracts salient features from the trait, and compares the extracted features against those in a database in order to *verify* a claimed identity or to *identify* an individual. For security and privacy reasons, biometric systems typically do not store the raw biometric data that may disclose sensitive information about the subjects (i.e., race, diseases, etc.). Rather, they store the extracted template (feature set) containing the most discriminative information about the individual and relevant for recognition purposes. However, recent research has looked into the possibility of recovering the original biometric data from the reduced template [3,4]. Such studies, which are also relevant from an information theory perspective (i.e., what is the amount of information necessary to reverse engineer a biometric template?), have set a new research trend in the biometrics field known as *inverse biometrics*. The present work falls under this category.

Among the various biometric traits that have been researched and used, iris is traditionally regarded as one of the most reliable and accurate [5]. After some preprocessing steps in which the iris is localized, segmented and normalized, the vast majority of iris recognition systems perform some type of filtering operation in order to generate the iris template (e.g., using 2-D Gabor wavelets). The phase information of the filtered normalized image is quantized to produce the final binary template (i.e., iriscode) which is stored in the database during enrollment. Then, in the authentication or recognition phase, iriscodes are compared using bit-based metrics like the Hamming distance [6–8]. This way iris recognition is accomplished based only on phase-related information, while the amplitude data is discarded due to its sensitivity to external factors such as imaging contrast, illumination or camera gain.

The iriscode has been adopted as a *de facto* standard by most iris-based systems, as it is a very efficient and compact representation of the discriminative characteristics contained within a person's iris pattern. It has been a common belief in the biometric community that binary templates do not have sufficient information to reconstruct the original iris image from them [9]. Furthermore, iriscodes from *real* iris images have been demonstrated to be significantly unique across individuals [10].

☆ This paper has been recommended for acceptance by Yingli Tian, PhD.

* Corresponding author. Fax: +34 91 497 2114.
E-mail addresses: javier.galbally@uam.es (J. Galbally), arun.ross@mail.wvu.edu (A. Ross), marta.barrero@uam.es (M. Gomez-Barrero), julian.fierrez@uam.es (J. Fierrez), javier.ortega@uam.es (J. Ortega-Garcia).

1077-3142/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2013.06.003

Are iriscodes really resilient to being reverse-engineered in order to recover the original iris pattern from them? Is it possible to generate different *synthetic* iris-like patterns which yield iriscodes very similar to the one given? In summary, can we generate *synthetic* images that match a specific binary template thereby potentially deceiving an iris recognition system?.

In the present work we address these questions by proposing a novel probabilistic approach based on genetic algorithms for the generation of iris-like synthetic patterns whose corresponding iriscodes match that of a genuine user. Two main goals are pursued:

- On the one hand, explore whether the phase information embedded in the iriscode is sufficient to reconstruct an iris image that can be successfully matched to the real one from which the template was generated. As a validation of the proposed reconstruction approach, we will investigate if the synthetically produced images may be used to deceive state-of-the-art commercial matchers.
- On the other hand, determine if it is possible to generate not just one, but a class of synthetic patterns with very similar iriscodes to that of a real one (i.e., exhibiting similarities in phase but differences in magnitude with respect to the original genuine pattern). As a validation of this second objective, we will determine if producing more than one reconstructed sample results in a better chance of deceiving iris recognition systems.

If the aforementioned goals are realized, it would imply that it is possible to generate synthetic iris images that are visually different from the original iris sample but which produce iriscodes that fall within the intra-class tolerance of a genuine user.

The work has been carried out from a computer-based perspective. This means that our goal is not to generate iris images that could fool a human expert; rather, the goal is to successfully match the synthesized iris images with their true counterparts using an automated iris matcher. Even so, different strategies to make the synthetic patterns look as realistic as possible are also explored in the experimental part of the article, where statistical results regarding the visual perception that experts and non-experts have of the reconstructed image are presented.

In order to provide a fully reproducible experimental protocol, which permits the comparison of the results with future studies, experiments are carried out on two publicly available databases. Furthermore, the iris recognition systems used for development and testing are well known matchers that can be easily obtained by any interested party.

The rest of the article is structured as follows. Related work is discussed in Section 2. The concept of automated iris recognition is briefly summarized in Section 3. The proposed iris reconstruction algorithm is presented in Section 4. The databases and iris matchers used in the experimental protocol are described in Section 5. The performance of the proposed approach is reported and analyzed in Section 6, while results evaluating the visual realism of the reconstructed images are given in Section 7. A preliminary quality assessment of the synthesized iris images is presented in Section 8. Conclusions are drawn in Section 9.

2. Related work and contributions

Recently, several researchers have addressed the problem of generating different synthetic biometric traits such as iris [11–15], fingerprints [16], signature [17,18], face [19], handwriting [20], and voice [21]. All these efforts have been mainly focused on the generation of *new* synthetic data, intended in general to overcome the limitation of assembling large biometric databases for performance assessment purposes. However, none of these very valuable efforts directly addresses the main objective of the present work, that is, the reconstruction of a synthetic biometric sample from a raw genuine template and an evaluation of the ensuing security implications.

In one of the earliest works on image reconstruction from templates, Hill [22] reported an experiment that challenged the notion of non-reversibility of fingerprint minutiae templates. His study suggested that the information contained in minutiae templates might allow for the reconstruction of images that are somewhat similar to the original fingerprints. Since that pioneer study, various researchers have successfully undertaken the challenge of generating a fingerprint image from minutiae points alone [3,23]. More recently, researchers have succeeded in manufacturing a gummy finger from minutiae templates [24].

In the context of face recognition, different approaches have been used by researchers to reconstruct face images from raw templates [25,26,4,27].

In the context of iris recognition, a very recent study has been the first to address the problem of generating iris images from binary iriscodes [28]. In this work, the authors take advantage of the prior knowledge of the feature extraction scheme used by the recognition system (i.e., functions defining the filters used during feature extraction) in order to reverse engineer the iriscode. Then, real images are used to impart a more realistic appearance to the synthetic iris patterns generated.

The differences between the deterministic technique described in [28] and the probabilistic method proposed in the present study will be pointed out throughout the article. However, the most important differences are as follows:

- *Type of approach*. In [28], given an iriscode and a fixed set of parameter values, the resulting reconstructed synthetic pattern is always the same (i.e., deterministic approach). Our methodology permits the reconstruction of potentially a large number of synthetic iris patterns with very similar iriscodes (i.e., probabilistic approach). As will be shown in the experimental part of the work, having more than one synthetic iris significantly increases the chances of matching against the true counterparts. Furthermore, apart from security-related studies such as the one carried out in the present paper, the proposed probabilistic method presents other potential applications that will be discussed in Section 9.
- *Knowledge required*. In the development stage, the method proposed in [28] requires knowledge of the feature extraction scheme being used by the recognition system. On the other hand, our technique only requires the output score of an iris matcher to reconstruct the image and does not need any prior information about how the recognition system obtains that score.
- *Images required*. In order to generate somewhat realistic iris-like patterns, the algorithm described in [28] relies on information from real iris images. No original samples are needed in the present study to obtain realistic-looking synthetic images.
- *Experimental protocol*. Although consistent, the experimental protocol followed in [28] does not allow for the comparison of its results with other methods, as the iris matchers used for development and validation are proprietary implementations and not publicly available. In the present work, the experimental protocol has been designed to be fully reproducible (i.e., publicly available databases and matchers are used) so that an objective comparison may be carried out with other reconstruction approaches proposed in the future.

A great deal of attention has been given recently to another key area directly related to the present work: evaluating the security of biometric systems. Many different researchers have addressed the

vulnerabilities of biometric systems to spoofing attacks (those carried out at the sensor level using, for instance, a gummy finger or a printed iris image) [29–32], and to software-based attacks (carried out against some of the internal modules of the system) [33,27,34]. Furthermore, the interest in the analysis of system vulnerabilities has permeated the scientific field and different standardization initiatives at the international level have emerged in order to deal with the problem of security evaluation in biometric systems, such as the Common Criteria [35] and its associated Biometric Evaluation Methodology BEM [36] or the ISO/IEC-19792:2009 for biometric security evaluation [37].

This new concern which has arisen in the biometric community regarding the security of biometric systems has also led to the initiation of several international projects, like the European Tabula Rasa [38], which involves the joint effort of researchers, developers and evaluators to improve this technology using the security-through-transparency principle: in order to make biometric systems more secure and reliable, their vulnerabilities need to be detected and analyzed so that useful countermeasures may be developed.

Following the same transparency principle which is beginning to prevail in the biometric community, the most significant contributions of the present work compared to those previously mentioned in this section are: (i) proposal of a novel genetic-based probabilistic approach for the reconstruction of iris patterns from their iriscodes, (ii) evaluation of the vulnerabilities of a commercial iris matcher to different attacks carried out with the reconstructed iris images using a systematic and reproducible experimental protocol, and (iii) demonstration of the feasibility of generating multiple synthetic iris patterns with similar iriscodes to that of a real iris image.

3. Summary of iris recognition

The objective of this section is to briefly summarize those aspects of an iris recognition system which are directly related to the present study and which are essential for the correct understanding of the work. For a more comprehensive, descriptive and self-contained review on automatic iris recognition the reader is referred to [10,7,39–42].

Common iris recognition systems comprise of five different stages: image acquisition, iris location and segmentation, normalization, encoding and matching. As has been mentioned before, the main objective of this work is to reconstruct an iris pattern from its encoded binary template. Thus, although the acquisition and segmentation tasks may be very challenging under certain scenarios (e.g., long distance acquisition, uncontrolled lighting conditions, eye deviation, etc.) they are not relevant to this study and will not be treated here.

- *Normalization*. Once the iris has been segmented, the vast majority of iris recognition systems transform the annular-like iris pattern in cartesian coordinates to a normalized rectangular image of fixed dimensions in pseudo-polar coordinates. These are the type of images that will be reconstructed using the algorithm described in this work. The normalization process may be reversed and the normalized iris patterns can be incorporated again into the original eye images (of the same or of a different user).
- *Encoding*. Although a number of methods have been reported in this stage, most of them use some type of filtering strategy (typically based on Gabor Wavelet filters) prior to quantizing the phasor response of the filtered output resulting in a binary representation of the iris image (i.e., the iriscode).

Finally, two iriscodes are compared using a bitwise operator such as the Hamming distance. In most cases, in the segmentation stage, a mask showing the occluded areas of the iris (e.g., due to the eyelids or eyelashes) is also generated. Thus, the matching score is only computed using the "non-masked" bits of the iriscode.

In Fig. 1 an example of the normalization and encoding stages is shown. The original iris image appears on top (a) with the two white circles denoting the outer and inner boundaries of the segmented iris. The corresponding normalized image along with the mask indicating the occluded areas (b) and the final iriscode (c) are also shown.

4. The reconstruction method

To extend formality to the problem being addressed, some mathematical notations are introduced in this section. Let $\mathbf{B}$ represent the iriscode of the user whose iris image is being reconstructed, $\mathbf{I_R}$ represent the reconstructed *normalized* iris image which is a solution to the problem, $\mathbf{B_R}$ be its associated iriscode and $\delta$the matching threshold that determines if two iris images are of the same eye.

Problem statement. Consider a $R \times C$ dimensional matrix $\mathbf{I_R}$ of real values, which is divided into $H \times L$ square blocks of dimension $R/H \times C/L$, with $H \leqslant R$ and $L \leqslant C$. This matrix is mapped by some unknown function $\mathcal{F}$ to a binary matrix $\mathbf{B_R}$ (i.e., $\mathbf{B_R} = \mathcal{F}(\mathbf{I_R})$) of dimensions $K \times W$ (K is a multiple of R and W is a multiple of C).

Consider the problem of finding an $\mathbf{I_R}$ matrix such that, its associated $\mathbf{B_R}$ matrix (unknown), produces a similarity score (s) greater than a certain threshold δ, when it is compared to a *known* binary matrix $\mathbf{B}$ according to some unknown matching function $\mathcal{J}$, i.e., $\mathcal{J}(\mathbf{B}, \mathbf{B_R}) > \delta$.

For clarity, we will define a new function $\mathcal{V}$ as: $\mathcal{V}(\mathbf{B}, \mathbf{I_R}) = \mathcal{J}(\mathbf{B}, \mathcal{F}(\mathbf{I_R})) = \mathcal{J}(\mathbf{B}, \mathbf{B_R}) = s$.

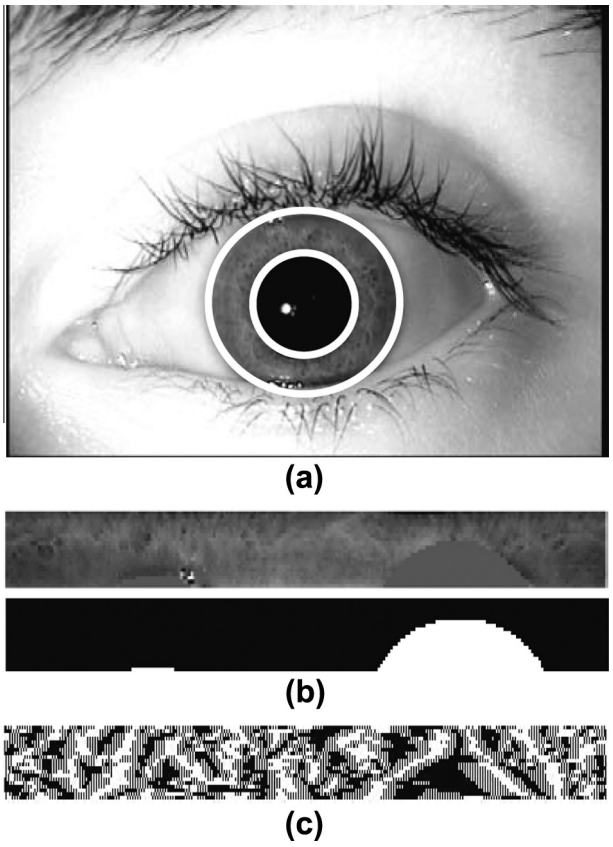

Fig. 1. Example illustrating the segmentation (a), the normalization and occlusion mask (b), and the encoding (c) stages used by most iris recognition systems.

Assumptions. Let us assume that we have access to the evaluation of the function $\mathcal{V}(\mathbf{B}, \mathbf{I_R})$ for several trials of $\mathbf{I_R}$.

Algorithm. The problem stated above may be solved using a genetic algorithm to optimize the similarity score given by the system, according to the general diagram shown in Fig. 2. Genetic algorithms, which have shown remarkable performance in optimization problems [43], are search methods that iteratively apply certain rules inspired by biological evolution to a population of individuals (possible solutions) according to a given fitness function. During each iteration the algorithm moves towards better solutions in terms of the fitness function which has to be optimized. In our particular problem, the following observations ought to be made.

- The fitness value associated with each individual (normalized iris image) is the matching score, $s = \mathcal{V}(\mathbf{B}, \mathbf{I_R})$.
- Usually genetic algorithms operate with individuals that are binary vectors. In this problem, the genetic algorithm has been modified to work with matrices of real values (i.e., $\mathbf{I_R}$) where each of the $H \times L$ blocks represents a gene of the individual.

As can be seen in Fig. 3, the steps followed by the reconstruction algorithm are:

1. Generate an initial population P_0 with N individuals of size $R \times C$ (i.e., dimensions of the normalized iris images), and tessellate each individual into $H \times L$ rectangular blocks.
2. Compute the similarity scores s^i of the individuals ($\mathbf{I}_\mathbf{R}^i$) of the population P_0, $s^i = \mathcal{V}(\mathbf{B}, \mathbf{I}_\mathbf{R}^i)$, with $i = 1, \ldots, N$.
3. Four rules are used at each iteration to create the next generation P_n of individuals from the current population:
   (a) *Elite*: The two individuals with the maximum similarity scores are retained unaltered for the next generation.
   (b) *Selection*: Certain individuals, the *parents*, are chosen by stochastic universal sampling [44]. Therefore, the individuals with the highest fitness values (similarity scores) are more likely to be selected as parents for the next generation: one subject can be selected 0 or many times. From the original N individuals, only $N-2$ are eligible (as the best two are retained as elite) from which $N/2-1$ *fathers* and $N/2-1$ *mothers* are chosen.
   (c) *Crossover*: Parents are combined to form $N-2$ *children* for the next generation by employing a scattered crossover method: a random binary matrix of size $H \times L$ is created and the genes (blocks) for the first child are selected from the first parent if the value of an entry is 1, and from the second when it is 0 (vice versa for the second child).
   (d) *Mutation*: Random changes are applied to the blocks of the new children with a mutation probability p_m. When a certain block is selected for mutation, the equivalent block in the individual of the population with the highest fitness value is changed.
4. Redefine $P_0 = P_n$ and return to step 2.

Stopping criteria. The algorithm stops when: (i) the best fitness score of the individuals in the population is higher than the threshold δ (i.e., the image has been successfully reconstructed), (ii) the variation of the similarity scores obtained in successive generations is lower than a previously fixed value, or (iii) when the maximum number of generations (iterations) is exceeded.

Criteria ii and iii are set in order to prevent an infinite loop in case condition i is not reached. In the particular case of the experiments described in Section 6 the first criterion was always met (i.e., all images were successfully reconstructed).

Additional note. There are some important characteristics of the reconstruction method presented above that should be highlighted as they differentiate it from other previously published iris reconstruction techniques [28]:

- Due to the probabilistic nature of the four rules being applied, the algorithm produces different solutions at each execution, even when the initialization and parameter values are the same. This facilitates the reconstruction of multiple normalized iris images ($\mathbf{I_R}$) whose iriscodes ($\mathbf{B_R}$) are very similar to the target ($\mathbf{B}$).
- The algorithm does not require prior knowledge of the mapping function $\mathcal{F}$ between the normalized iris images ($\mathbf{I_R}$) and their corresponding iriscodes ($\mathbf{B_R}$).
- The algorithm does not require knowledge of the matching function $\mathcal{J}$.
- The algorithm does not require knowledge of the function $\mathcal{V}$, but just its output to the given inputs.
- No *real* iris images are involved in the reconstruction process. As will be explained in Section 5, the initial population P_0 is taken from a database of fully *synthetic* iris images.

A genetic search algorithm was used in this work, since the nature of the search space is unknown to us. Specifically, it is not clear if the objective function results in a smooth or even a continuous search space. Consequently, the efficiency of classical stochastic gradient descent methods would be at least unclear. Although previous work in [45] partially supports the assumption of smoothness/continuity, this could not be easily substantiated in our case. Therefore, by simultaneously searching for multiple

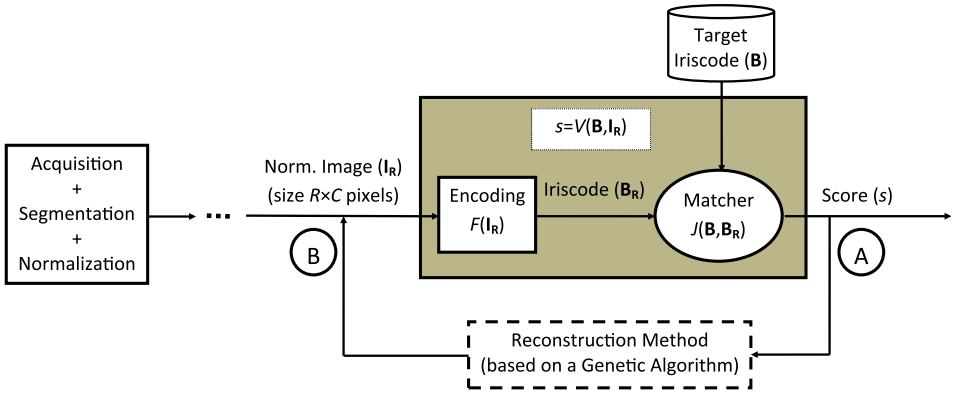

Fig. 2. General diagram of the scheme followed in the present work. A detailed diagram of the reconstruction approach (dashed rectangle) is given in Fig. 3 where points A and B show, respectively, the input and output of the algorithm.

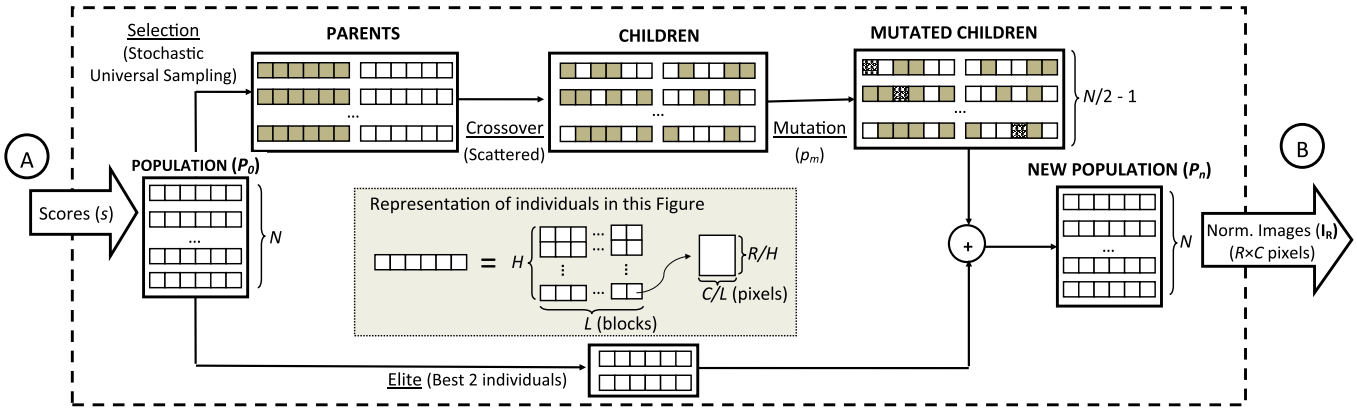

Fig. 3. Diagram of the probabilistic method proposed in the present work for the reconstruction of iris images from their iriscode. Points A and B (input and output of the reconstruction algorithm respectively) may be seen for reference in Fig. 2. As is shown in the shaded chart in the center of the figure, although individuals are represented as vectors for simplicity, strictly they are matrices of size $R \times C$ pixels divided into $H \times L$ blocks.

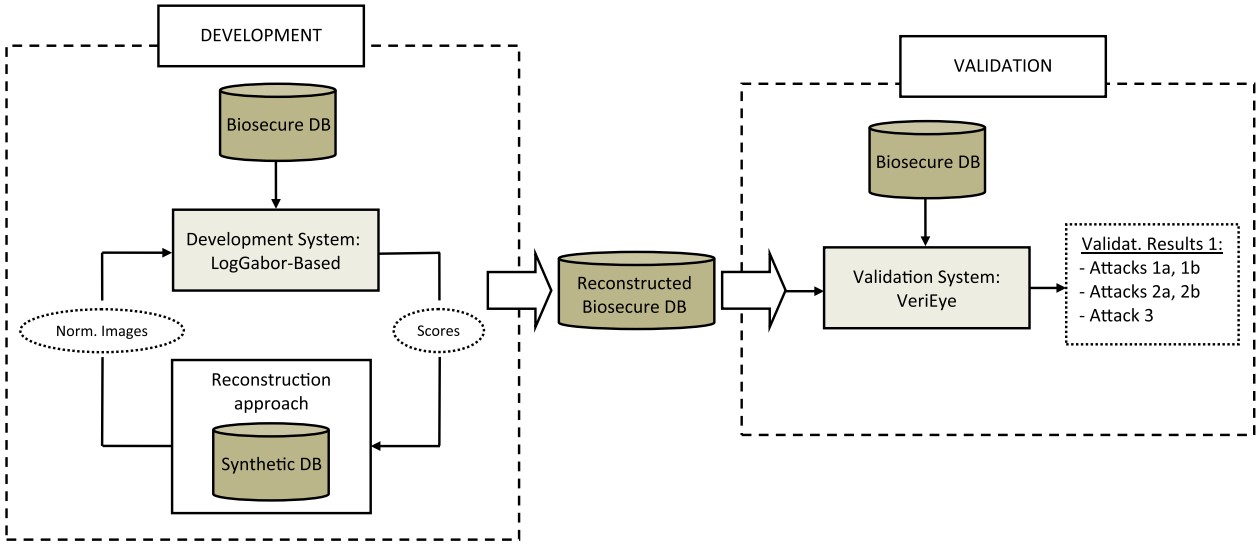

Fig. 4. Diagram of the experimental protocol followed in the present work. The databases and systems used are highlighted with a darker shade. The protocol is described in Sections 5 and 6.

solutions in the solution space, genetic algorithms are more likely to avoid potential minima or even plateaus in the search space (much like simulated annealing schemes).

5. Experimental protocol: databases and systems

As shown in Fig. 4 the experimental protocol is divided into a development stage and a validation stage, where two different databases and two different iris matchers have been used in order to ensure unbiased results. The databases and iris matchers are publicly available and so the results obtained in this study are fully reproducible and may be compared with other methods in the future.

5.1. Databases

Two databases, one containing real iris samples and another containing synthetic samples, are used in the experiments. The iris images to be reconstructed are taken from the real database (Biosecure DB), while the synthetic dataset (SDB) is used for the initialization of the reconstruction algorithm (see Fig. 4).

As was described in Section 4, the reconstruction method proposed in the present work needs a set of iris images for its initialization. This pool of initial samples is taken from a database of fully synthetic iris images for two main reasons: on the one hand, this avoids any possible overlap between the reconstructed images and those used in the reconstruction process (which could lead to overoptimistic results), and, on the other hand, it avoids the need for using real iris images in the reconstruction method.

- *The real database: Biosecure DB.* The real images to be reconstructed in the experiments are taken from the iris corpus included in the Desktop Dataset of the multimodal BioSecure Database [46] that contains voice, fingerprints, face, iris, signature and hand data of 210 subjects, captured in two time-separated acquisition sessions. This database was acquired by the joint effort of 11 European institutions and has become one of the standard benchmarks for biometric performance and security evaluations [47]. It is publicly available through the BioSecure Association.[1]
  The database consists of three datasets captured under different

[1] http://biosecure.it-sudparis.eu/AB.

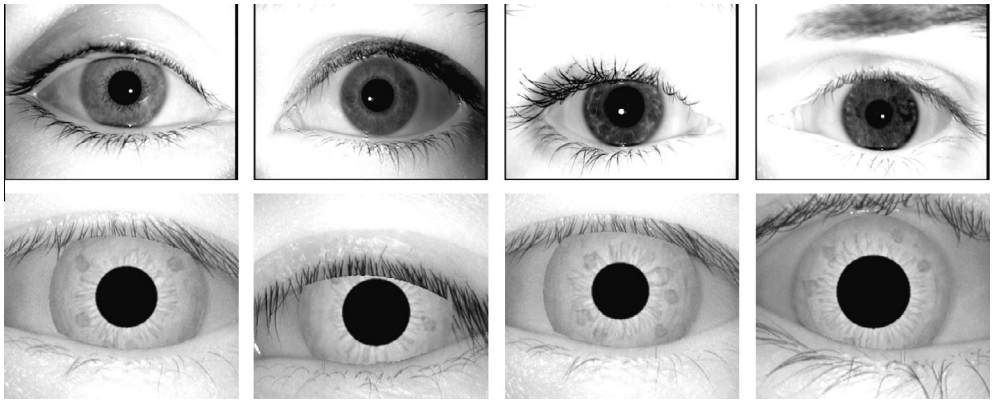

Fig. 5. Examples of iris images from the two databases used in the experiments: real (top) and synthetic (bottom).

acquisition scenarios: (i) Internet Dataset (DS1, captured through the Internet in an unsupervised setup), (ii) Desktop Dataset (DS2, captured in an office-like environment with human supervision), and (iii) the Mobile Dataset (DS3, acquired using mobile devices withinuncontrolled conditions). The iris subset used in this work includes four gray-scale images (two per session) per eye, all captured using the Iris Access EOU3000 sensor from LG. In the experiments the two eyes of each subject have been considered as separate users, resulting in a total of $210 \times 2 \times 4 = 1680$ iris samples.

- *The Synthetic Database: SDB*. Being a database that contains only fully synthetic data, it is not subjected to any legal constraints and is publicly available through the CITeR research center.[2] The synthetic irises are generated following the method described in [13], which has two stages. In the first stage, a Markov Random Field model is used to generate a background texture representing the global iris appearance [12]. In the next stage, a variety of iris features such as radial and concentric furrows, collarette and crypts, are generated and embedded in the texture field. The database includes seven gray-scale images of 1000 different subjects.

Typical examples of the eye images that can be found in Biosecure DS2 (top) and SDB (bottom) are shown in Fig. 5. We can observe that the samples in both datasets are totally different: this underscores our intention to avoid biased results.

5.2. Iris recognition systems

Two different iris matchers are used in the experiments (see Fig. 4). The first one, consisting of fully accessible software modules, is used as the development system for the reconstruction of the iris images. The second one, completely independent from the previous one, is used in the validation stage in order to match the reconstructed images against the real ones.

- *Development: LogGabor filter-based* [48]. In the development stage, where iris images are reconstructed from real iriscodes, a modified version of the iris matcher developed by Masek and Kovesi [48] is used. This system was selected for several reasons: (i) it is publicly available and its source code may be freely downloaded[3], (ii) although its performance is certainly lower than that of current state-of-the-art iris recognition systems, it is widely used in many iris-related publications to give baseline results, and (iii) it is partitioned into independent software modules that permit access to the matching score (needed by the proposed reconstruction method).
  In Masek's matcher, the different stages involved in iris recognition (described in Section 3) are implemented following a *classical* approach: (i) *segmentation*, the method proposed in [49] is followed, modeling the iris and pupil boundaries as circles; (ii) *normalization*, a technique based on Daugman's rubber sheet model that maps the segmented iris region into a 2D array is used [7]; (iii) *feature encoding*, produces a binary template of $20 \times 480 = 9600$ bits by filtering the normalized iris pattern with 1D Log-Gabor wavelets and quantizing the filtered output to four levels (i.e., two bits) according to [7]; and (iv) *matching*, a modified Hamming distance that takes into account the noise mask bits is used.
- *Validation: VeriEye* [50]. For the validation experiments, the VeriEye commercial matcher marketed by Neurotechnology[4] is used to determine the matching potential of the reconstructed iris images. The motivation for its selection is twofold: (i) it was ranked among the top performing matchers in the NIST Iris Exchange (IREX) independent evaluation in 2009 [51], and, (ii) being a commercial matcher it works as a black-box for the user, who has no knowledge of the algorithms used in any of the stages of the iris recognition process (being a commercial matchers its implementation details are proprietary). Therefore, the results of our proposed method are ensured to be unbiased and not due to a specific adaptation of the reconstruction algorithm to a given validation system.

6. Results: performance

Besides avoiding biased results, the experimental framework has been designed to evaluate the performance of the reconstruction algorithm and its compliance with the main objectives set in this work: (i) can the iris images reconstructed using the proposed method be successfully matched with the original iris counterpart? (main goal of the present work), (ii) can the reconstruction scheme synthesize different iris-like patterns whose iriscodes are very similar to the original real iris? (secondary goal of the present work).

6.1. Development experiments: LogGabor filter-based system

The objectives of this first set of experiments are: (i) to fix the values of the different parameters involved in the reconstruction algorithm and, (ii) (once the parameters have been set) to reconstruct the real iris images in Biosecure DB starting from their iriscodes.

[2] http://www.citer.wvu.edu/.
[3] www.csse.uwa.edu.au/pk/studentprojects/libor/sourcecode.html.
[4] http://www.neurotechnology.com/verieye.html.

In order to achieve these two goals, one sample of each of the 420 users present in the Biosecure DB (right and left irises of 210 subjects) were randomly selected and their iriscode computed according to the publicly available iris recognition system developed by Masek and Kovesi [48]. The dimensions of the normalized iris images produced by this system are $R \times C = 20 \times 240$ and the size of their corresponding binary templates $K \times W = 20 \times 480$ (i.e., each pixel is coded with two bits).

In order to determine the parameter values of the genetic algorithm effectively, certain general guidelines should be taken into account. Probably, the key factor is to determine the population size. On the one hand, if it is too small the risk of converging prematurely to a local minima is increased since the population does not have enough genetic material to sufficiently cover the problem space (i.e., the diversity is too low). On the other hand, a larger population has a greater chance of finding the global optimum at the expense of drastically increasing the computation load (i.e., CPU time) as the number of iterations needed for convergence is greater.

In most GA-related solutions, the individual's size (i.e., number of blocks $H \times L$) is determined by the problem at hand. However, in our specific case, the same reasoning used for the population size applies to the individual's dimensions as well. Therefore, for this particular problem, a good balance must be obtained between both parameters. As a general rule of thumb, in this specific case, good results are usually obtained when $N \simeq L$.

Besides the aforementioned trade-off, in most GA-related problems the mutation probability is usually kept below 1% in order to avoid losing diversity.

With these general principles in mind, extensive experiments were undertaken to determine a good set of parameter values for the reconstruction algorithm, resulting in the following efficient operating point: population size $N = 80$, mutation probability $p_m = 0.003$, and block size $R/H \times C/L = 2 \times 2$ pixels (i.e., each normalized image is divided into $H \times L = 10 \times 120$ blocks).

It must be emphasized that these parameter values could be further optimized. Furthermore, different strategies than those used here may be adopted in order to implement each of the four rules described in Section 4 (i.e., elite, selection, crossover and mutation). However, the above (or other) improvements related to genetic algorithms are outside the scope of the present work, which is not focused on the study and optimization of this search tool, but rather on the reversibility of binary iris templates. For a more detailed description of different architectures for genetic algorithms the reader is referred to [43,52].

Once the parameter values of the reconstruction method were determined and fixed, Masek's matcher [48] was then used to compute the matching scores needed by the optimization algorithm, in order to generate 5 different reconstructed images of each binary template (i.e., the algorithm was applied 5 times to reconstruct and image from each iriscode), thus leading to a database of $5 \times 420 = 2,100$ reconstructed iris images (referred to as Reconstructed Biosecure DB in Fig. 4).

In order to determine the positive matching threshold δ at which an iriscode is considered to have been successfully reconstructed, the iris recognition system performance was evaluated on the Biosecure DB. Genuine scores were computed by matching the first sample of each user to the other 3 images of that same user (i.e., $420 \times 3 = 1260$ genuine scores), while impostor scores were generated by comparing the first iris of each user to the first sample of the remaining users in the database (i.e., $420 \times 419 = 175{,}980$ impostor scores). The two sets of similarity scores are depicted in Fig. 6, where the selected positive matching threshold has been highlighted with a vertical dotted line. We can observe that, below that value, $\delta = 0.3$, the probability of having an impostor score is almost zero. Thus, two iris images producing such a similarity score may be considered to come from the same user.

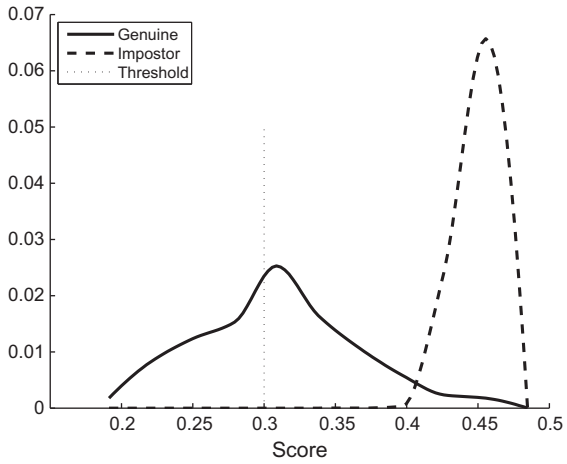

Fig. 6. Genuine and impostor score distributions of the iris matcher used in the development experiments. The selected positive matching threshold is marked with a vertical dotted line, $\delta = 0.3$.

In Fig. 7 three different reconstruction outcomes corresponding to a single real iriscode are shown. Although the reconstructed patterns do not visually resemble the original one and block artifacts are discernible, their corresponding iriscodes are all very similar to each other and exhibit a high degree of resemblance with the original. The visual dissimilarity between the original and the reconstructed patterns may be explained by the absence of amplitude-related information in the iriscodes. This leads to arbitrary amplitude values in the synthetically generated samples which, nevertheless, present comparable phase information, resulting in accurate iriscode reproductions. Above each reconstructed image in this figure, the evolution of the score across iterations is shown. Marked with a horizontal dashed line is the positive matching threshold $\delta = 0.3$.

6.2. Validation experiments: VeriEye

The iris images reconstructed in the development stage are used to test the vulnerabilities of the VeriEye iris matcher (see the validation chart in Fig. 4). As mentioned in Section 5.2, this system operates as a black-box, i.e., given an input, it returns an output with no information about the internal algorithms used to get that final result. Several remarks have to be made regarding the inputs and outputs of VeriEye:

- *Inputs*. Normalized iris samples in polar coordinates are not accepted by VeriEye. The input to the system has to be an image containing a *circular* iris in cartesian coordinates. For this reason, in order to attack the system, all the reconstructed irides were reconverted into Cartesian coordinates as shown in Fig. 8.
- *Outputs*. The system outputs a non-zero similarity score in case of a positive match. When the matching threshold is not reached, a 0 is returned, thereby making it difficult to launch a hill-climbing attack [27]. In case an error occurs during the recognition process (most likely during the segmentation stage), a negative score value is returned.

The performance of the attack is measured in terms of its Success Rate (SR), which is defined as the percentage of successful attacks (A_s) out of the total carried out (A_T), i.e., $\mathrm{SR} = A_s/A_T \times 100$. The key factors to compute the SR are to define (a) what constitutes an attack, and (b) when an attack is considered to be successful. In

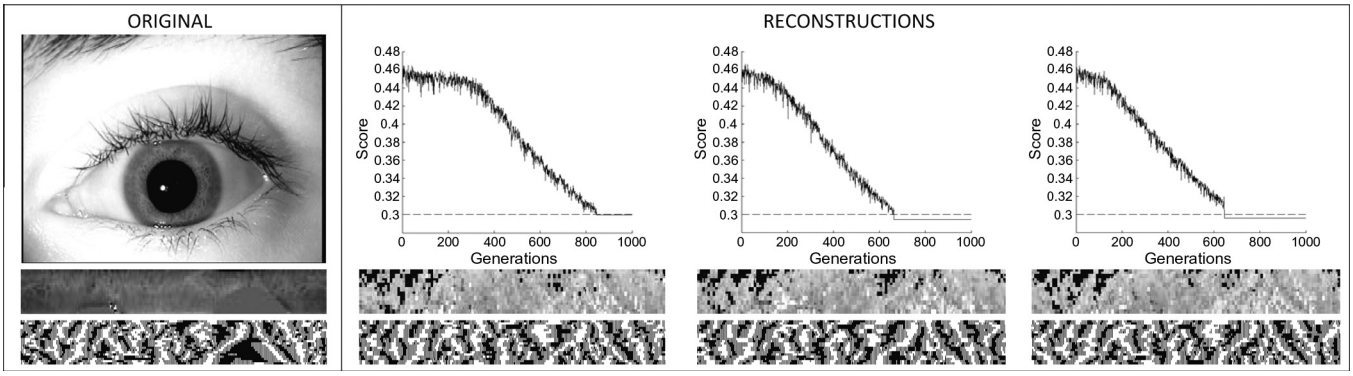

Fig. 7. Three example executions (right) of the reconstruction algorithm for the same original image (left). For the reconstruction samples, the evolution of the score through the generations is shown on top (positive matching threshold marked with a horizontal dashed line), with the final reconstructed normalized image and its corresponding iriscode shown below.

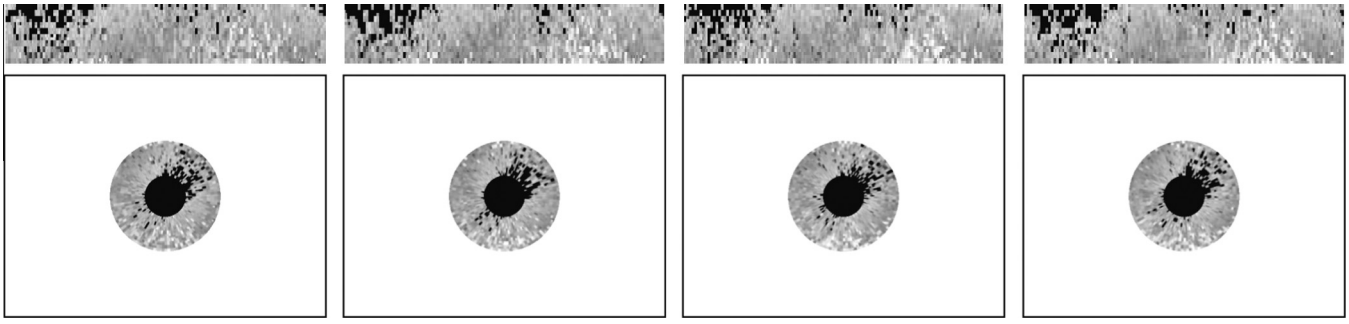

Fig. 8. Four reconstructed iris images in pseudo-polar coordinates (top) all recovered from the same original iris, and their corresponding denormalized images in cartesian coordinates used to attack the VeriEye commercial matcher (bottom).

the experiments, three representative attacks will be taken into account in order to estimate the performance of the proposed reconstruction method:

1. *Attack 1*: 1 reconstructed image vs 1 real image. In this case the attack is carried out on a 1-on-1 basis. That is, one reconstructed image is matched against one real image and, if the resulting score exceeds the fixed matching threshold, the attack is deemed to be successful. Two possible scenarios may be distinguished in this case depending on the real image being attacked:
   (a) The real image being attacked is the original sample from which the synthetic image was reconstructed. In this scenario the total number of attacks performed which will be used to compute SR_{1a} is $A_{T1a} = 420 \times 5 = 2,100$.
   (b) The real image being attacked is one of the other three samples of the same user present in the Biosecure DB. For this experiment the total number of attacks performed which will be used to compute SR_{1b} is $A_{T1b} = 420 \times 3 \times 5 = 6300$.
2. *Attack 2*: 5 reconstructed images vs 1 real image. In this case all five reconstructed images are matched against the real sample. The attack is successful if at least one of the synthetic images matches against the real image. This represents the most likely attack scenario analyzed in other related vulnerability studies [23]; here, the iriscode of a legitimate user in the database is compromised and the intruder reconstructs multiple images of the iris to try and break the system. The attacker will gain access if any one of the reconstructed images results in a positive score.
   The same two scenarios as in attack 1 can be considered here and, so, the total number of attacks carried out in each scenario will be $A_{T2a} = 420$ and $A_{T2b} = 420 \times 3 = 1260$. The resulting success rates will be denoted as SR_{2a} and SR_{2b}, respectively.
3. *Attack 3*: 5 reconstructed images vs average (4 real images). It is a common practice in many biometric recognition systems to match the test sample against several stored templates and return the average score. To emulate this scenario, each reconstructed iris image is matched against the four samples of the real user available in the Biosecure DB. The attack is successful if the average score due to *of any of the five reconstructed images* is higher than the given operating threshold. Thus, in this case, the total number of attacks performed in order to compute SR_3 is $A_{T3} = 420$.

In general, the success of an attack is highly dependent on the False Acceptance Rate (FAR) of the system. Thus, the vulnerability of the system to the attacks with the reconstructed images is evaluated at three operating points corresponding to FAR = 0.1%, FAR = 0.05%, and FAR = 0.01%, which, according to [53], correspond to a low, medium and high security application, respectively. For completeness, the system is also tested at a very high security operating point corresponding to FAR $\ll$ 0.01%.

As was mentioned before, this commercial matcher does not return impostor scores (i.e., they are always 0) which means that its FAR may not be statistically computed on a given database. In order to fix the threshold for the different operating points, a deterministic equation is given in the documentation enclosed with the system.

In the experiments, the system was unable to segment (i.e., reported an error) 1.4% of the real images in the Biosecure DB. This implies that, for these cases, a sample from a legitimate user would have not been able to access the system. Thus, the highest SR that can be reached by the attacks is 98.6%. Moreover, 0.5% of the reconstructed images were not correctly segmented (these are regarded as unsuccessful attacks).

Several observations can be made from the results of the validation experiments carried out on VeriEye as shown in Table 1:

Table 1
SR of the different attacking scenarios considered for the VeriEye matcher at the four operating points tested.

FAR (%)	SR (%) – VeriEye					
	SR_{1a}	SR_{1b}	SR_{2a}	SR_{2b}	SR_3	Average
0.1	81.2	66.7	96.2	92.8	96.7	86.7
0.05	79.2	63.4	96.2	91.4	95.2	85.1
0.01	77.3	60.9	95.2	90.9	93.8	83.6
0.0001	69.0	49.1	92.8	82.8	82.9	75.3

- The high performance of the proposed reconstruction algorithm is confirmed, reaching an average SR of around 85% for the three usual operating points considered and over 95% for the most likely attacking scenario (i.e., SR_{2a}).
- Even for an unrealistically high security point (i.e., FAR = 0.0001%), the reconstructed images would have, on average, almost 75% chances of breaking the system.
- As expected, it is more probable that the synthetic samples are positively matched to the original image from which they were reconstructed than to other real images of the same user (see the decrease in the SR between SR_{1a} vs SR_{1b} and between SR_{2a} vs SR_{2b}).
- Even so, the reconstructed images still present a high probability of breaking the system even when the stored templates are not the one from which they were recovered (average SR of SR_{1b} and SR_{2b} around 75%).
- Furthermore, in the case of using several real samples of the user for verification (SR_3), the reconstructed images are still able to access the system ~94% of the time at the usual operating points, and for 80% of the attempts in the extremely high operating point tested.
- Besides, a new possible vulnerability of iris recognition applications has been raised, as the tested system positively matches images with a black circle in the middle and a white background (such as the ones shown in Fig. 8) that should by no means be recognized as an eye image.

The last observation emphasizes the need for incorporating some type of pre-checking stage, prior to the localization and segmentation of the iris, in order to confirm that the sample presented to the system is really that of an eye, and not some simple iris-like image.

The results presented in Table 1 confirm the first objective set in the present work: iris patterns may be recovered from their iriscodes, and the reconstructed images represent a threat to the integrity of automatic recognition systems.

Recall that the second goal of the work is to determine the feasibility of generating multiple synthetic iris patterns with iriscodes very similar to a real one. In order to address this point, results from experiment 2.a (i.e., all 5 synthetic images are compared against the original image) are presented in Table 2 from a different perspective. In this case we report in each column the percentage of attacks in which only n out of the 5 reconstructed images (with $n = 1, \ldots, 5$) were positively matched to the original real image. In each case, the total number of attacks performed is $A_{Tn} = 420$ and the success rate is denoted as SR_n.

Table 2
Percentage of successful attacks where n out of the total 5 reconstructed images were positively matched against the original iris image from whose iriscode they were reconstructed. Results are given for the four operating points tested on VeriEye.

FAR (%)	SR_n (%) – VeriEye				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
0.1	1.9	5.3	13.3	24.8	50.9
0.05	2.4	6.7	13.8	27.6	45.7
0.01	3.8	6.2	13.8	28.1	43.3
0.0001	7.6	6.7	21.9	24.7	31.9
Average	3.9	6.2	15.7	26.3	42.9

Averaging over the four operating points, all five reconstructed images were positively matched to the original image in 42.9% of the cases. This increases to 69.2% if we consider $n = \{4, 5\}$, and to 84.9% when taking into account $n = \{3, 4, 5\}$. These results confirm the ability of the proposed probabilistic reconstruction method to generate multiple iris patterns that match successfully against one specific iriscode. As can be seen in Table 1, this ability gives the proposed method a much higher attacking potential than deterministic algorithms that can only generate one image from each iriscode: the success rate increases by around 27% on an average when several reconstructions of the iris image are available (i.e., attack 2: 1 vs 5) compared to the case in which only one reconstructed sample is used to access the system (i.e., attack 1: 1 vs 1).

6.3. Validation experiments: variation of the reconstruction threshold

As explained in Section 6.1, in the previous validation experiments, the reconstruction threshold was set to $\delta = 0.3$ since, as seen in Fig. 6, it is very unlikely that an impostor score falls within the range of this value (impostor scores for the development system are typically below 0.4.) In the present set of experiments, this reconstruction threshold is modified in order to gain a deeper understanding of the proposed reconstruction approach. In particular, the objective of these experiments is to: (i) determine the impact of the reconstruction threshold selection on the success rate of the studied attacks, and (ii) establish if it is possible to obtain a perfect reconstruction following the proposed scheme (i.e., generate an iris image with the exact same iriscode as the original sample), and analyze the increase in the computational cost for perfect reconstruction.

In order to accomplish the first objective, the reconstruction threshold was set to $\delta = 0.4$, at the lower bound of the impostor scores range as shown in Fig. 6. This means that the reconstructed images will be less similar to the original counterparts compared to the ones generated in Section 6.2 where the threshold was fixed at $\delta = 0.3$. Therefore, the SR of the attacks carried out with these new synthetic images should be lower than that achieved in the previous experiments.

In Table 3 we show the comparative performance of the attack for the two considered reconstruction thresholds, where we observe that, as expected, the success rate decreases from over 90% in the best reconstruction case ($\delta = 0.3$), to less than 10% when the threshold is increased. For clarity, the comparison has just been established for the operating point corresponding to FAR = 0.5%.

These results show that selecting a better reconstruction threshold results in a higher success rate for the attack at the cost of a higher computational cost (i.e., time) to generate each reconstructed sample.

In order to determine this increase in the computational cost (i.e., objective ii for this set of experiments), three "best possible" reconstructions were generated from five randomly selected irises.

Table 3
SR of the different attacking scenarios considered for the VeriEye matcher at the operating point FAR = 0.01% for two different reconstruction thresholds.

FAR = 0.01%	SR (%) – VeriEye					
	SR_{1a}	SR_{1b}	SR_{2a}	SR_{2b}	SR_3	Average
$\delta = 0.3$	77.3	60.9	95.2	90.9	93.8	83.6
$\delta = 0.4$	5.1	0.0	9.4	8.6	8.8	7.9

To reach these "best possible" synthetic images, the parameters of the algorithm were set as follows: population size $N = 150$, mutation probability $p_m = 0.003$, and block size $R/H \times C/L = 1 \times 1$ pixels (i.e., each normalized image is divided into $H \times L = 20 \times 240$ blocks, which means that each pixel is considered as a single block). No reconstruction threshold was fixed and the algorithm iterated 10,000 times to determine the lowest threshold that can be reached by the proposed method (i.e., best possible reconstruction).

In Fig. 9, we show three optimal reconstructions of one original iris, together with their corresponding iriscodes and the evolution of the matching score through the iterations. It can be seen that the best images generated by the algorithm produce a score close to 0.1. This means that, following this approach, although we can generate extremely good reconstructions, it is unlikely that *the exact* iriscode can be recovered, mainly due to two factors: on the one hand, GAs do not guarantee finding the global minimum of the function being optimized and, on the other hand, finding an *identical* iriscode would require a larger population with sufficient diversity (which in this case is limited by the amount of synthetic initialization data).

We note that each of these best possible reconstructions took around 12 execution hours, compared to the few minutes taken when the threshold was set to $\delta = 0.3$ (also with a smaller population and smaller individuals). These results reinforce the need to reach a trade-off between computational cost and quality of the reconstructed samples.

7. Results: appearance

Although the primary objective of this work is to determine if *automatic* iris recognition systems may be deceived by the reconstructed iris images (see results in Section 6), in this section we statistically evaluate the *visual realism* of the reconstructed iris images from a human point of view.

Firstly, the reconstructed samples are given a more realistic appearance by suppressing the block artifacts. To this end, the synthetic iris images are smoothed using standard image processing tools (Gaussian filtering); transformed into the cartesian space; and then embedded in a real eye image. Some examples of the outcome of this process are shown in Fig. 10. It has to be noted here that real eye images are being introduced for the first time in the experimental protocol with the only purpose of *hosting* the synthetic samples, in order to obtain a fair comparison between the two classes (real and synthetic). The smoothing operation carried out on the reconstructed images barely affects matching performance, which remains as shown in Section 6 (with negligible variations).

In order to compute statistically significant results, a set of 100 randomly chosen real and reconstructed samples (50 of each class) was given to two groups of people: (i) *non-experts*, comprising 25 subjects with naive knowledge on iris recognition and (ii) *experts*, consisting of 15 researchers working in the iris field. Both groups were asked to mark each specimen from 0 (fully synthetic) to 4 (somewhat synthetic) and from 6 (somewhat real) to 10 (fully real) according to their impression after a quick inspection of the iris. The maximum time permitted to complete the experiment was 15 min.

Two types of errors can be committed in the classification task: (i) a real iris is marked as synthetic (0–4), measured by the False Synthetic Rate (FSR), and (ii) a synthetic iris is marked as a real sample (ranked 6–10), measured by the False Real Rate (FRR). The final Average Classification Error (ACE) is defined as ACE = (FSR + FRR)/2. These error rates are presented in the first three columns of Tables 4 and 5. In the next two columns we give the average score given by all subjects to the 50 real and synthetic samples. Finally the average time taken to complete the experiment is shown.

From the results presented in Table 4 we can see that over one third of the irides (38%) were misclassified by non-expert participants, proving the real-like appearance of synthetic samples (a random classifier would present an ACE of 50%). It should also be noticed that both error rates FSR and FRR are very close (36.2% and 39.3%, respectively) which means that the number of mistaken real and synthetic samples is very similar and that it is not easier to distinguish one class over the other. Furthermore, the average score given by the participants to real (5.61) and synthetic specimens (4.23) is quite close, reinforcing the idea that human subjects have a very similar perception of both types of irides.

As expected, the error rates reported in Table 5 for the experts group (close to 10%) are much lower than that of the non-expert group. These results show that, although it would be very difficult for synthetic irides to deceive an expert after close inspection, he can still make some errors when a non-detailed examination is done. This confirms their relatively high level of similarity to real samples.

Finally, we remark that the average time taken by the participants to carry out the task was a little less than 10 min (around 6 s per iris), which is consistent with the overall objective of the experiment of not making a detailed analysis of each iris, but rather estimating the general visual appearance of these samples after a brief inspection.

8. Results: quality assessment

In order to complement the human-aided results presented in Section 7, a coarse estimation of the quality of the genuine and

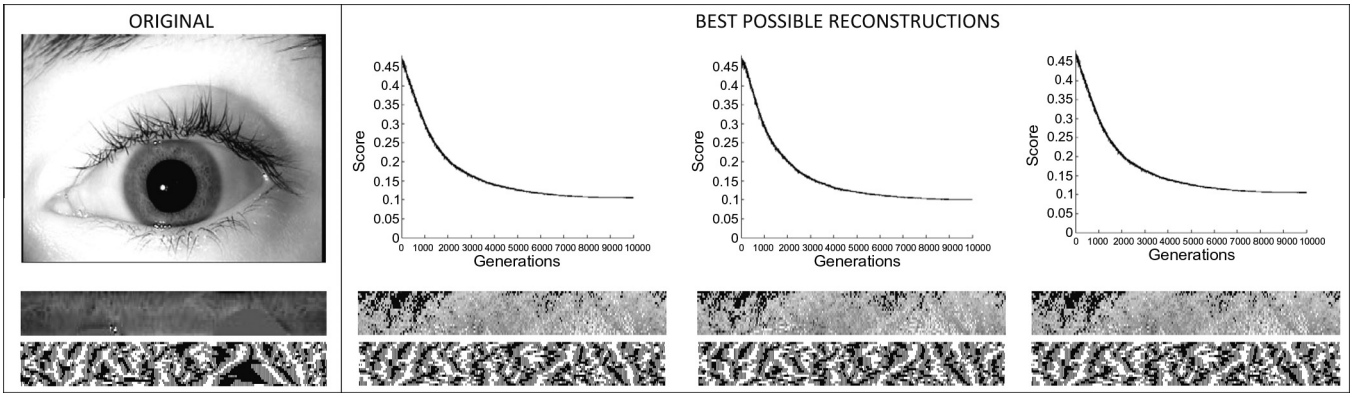

Fig. 9. Three "best possible" reconstructions reached with the proposed algorithm (right) for the same original image (left). For the reconstruction samples, the evolution of the score through the generations is shown on top, with the final reconstructed normalized image and its corresponding iriscode shown below.

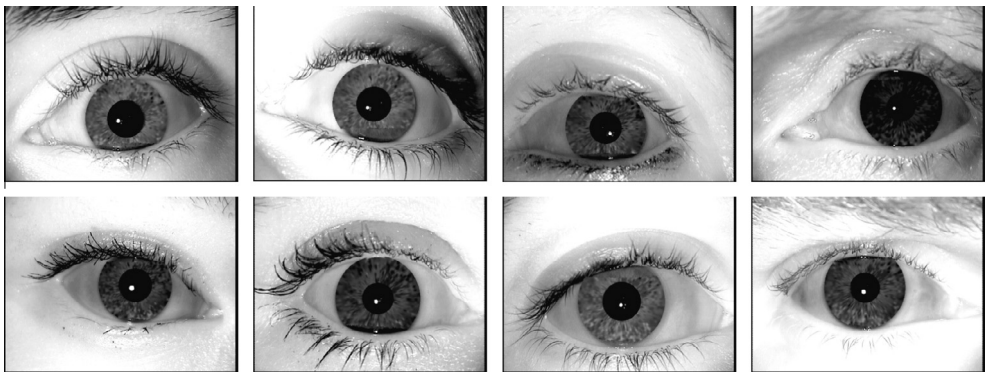

Fig. 10. Examples of reconstructed images used in the appearance evaluation experiments.

Table 4
Error rates, average score and average time of the 25 non-expert participants in the appearance evaluation experiment. FSR stands for False Synthetic Rate, FRR for False Real Rate, and ACE for Average Classification Error.

Non-expert participants (25)					
Error rates (%)			Average score		Average time (min)
FSR	FRR	ACE	Real	Synthetic	
36.2	39.3	37.7	5.61	4.23	9.7

Table 5
Error rates, average score and average time of the 15 expert participants in the appearance evaluation experiment. FSR stands for False Synthetic Rate, FRR for False Real Rate, and ACE for Average Classification Error.

Expert participants (15)					
Error rates (%)			Average scoring		Average time (min)
FSR	FRR	ACE	Real	Synthetic	
9.0	7.6	8.3	7.5	1.9	8.6

reconstructed samples from the perspective of automatic recognition systems was performed. For this purpose, both real and synthetic images were compared in terms of their quality.

From a biometric point of view, the quality of iris images can be assessed by measuring one of the following properties: (i) motion blur, (ii) occlusion, (iii) contrast, and (iv) other factors, including the image focus or the dilation of the pupil. A number of sources of information can be used to measure these properties including the high frequency power spectrum, angle information provided by directional filters, pixel intensity of certain eye regions, or different ratios comparing the iris area to that of the image, and the diameters of the iris and pupil. Iris quality can be assessed by either analyzing the image in a holistic manner, or combining the quality from local blocks in the image.

The four parameters used in this study are:

- *Motion: Frequency Distribution Rates (FDR1 and FDR2)* [54]. These are different combinations of three different parameters which consider, respectively, the power of the low (F_1), medium (F_2), and high (F_3) frequencies (computed according to the 2D Fourier Spectrum) of two local regions situated on the sides of the pupil. For the present work, two different features have been considered: FDR1 $= F2/(F1 + F3)$ and FDR2 $= F3$.
- *Occlusion: Region of Interest (RoI)* [55]. It analyzes the average value of the pixels in the region of interest, located 50 pixels above the pupil center.
- *Contrast: Local Contrast (LC)* [56]. This quality feature is adapted from the technique presented in [56] for occlusion estimation. A square region covering the iris and pupil is divided into a 10×10 cell grid. Each cell is assigned a value which corresponds to the power of its medium frequencies. The final quality measure is obtained by taking a ratio between the number of cells whose values fall between 20 and 60, and the total number of cells.

The four quality-related features described above, each measuring a different image property, were computed for all the samples of (i) the Biosecure Database, (ii) the Synthetic Database (SDB) comprising of simple iris-like images with a white background (see Fig. 8), and (iii) the synthetic iris images embedded in a real eye used in the appearance experiments described in Section 7.

The results for the four quality features applied to the three databases are depicted in Fig. 11. As was expected, for all four measures, the quality of the embedded synthetic irides is closer to that of the genuine samples than the raw reconstructed images. Nevertheless, there is a fairly good separation between these distributions, which suggests that quality assessment may be a suitable approach for developing a biometric-based countermeasure against the disclosed threat, as has already been achieved for other modalities [57]. However, due to the detailed experimentation and analysis that would be required, such a study falls out of the scope of the present contribution and will be the topic of future research.

9. Conclusions

This work has shown that the phase information summarized in iriscodes is sufficient to generate synthetic iris-like images with very similar binary templates to that of the original iris pattern. The experimental findings indicate that an eventual attack against iris matchers using such reconstructed images would have a very high chance of success. Such an attack presupposes that (a) the system stores unencrypted templates (or the attacker is able to override this protection) and that (b) synthetic iris samples can be input to the matcher. Since iriscodes only encode phase-related data of the original iris image and discard the amplitude information [8], there are visual differences between the reconstructed iris and the original iris. However, results indicate that it is quite likely to deceive a non-expert human observer with the reconstructed samples even though the synthetic grayscale iris patterns are not a fully accurate reproduction of the original patterns.

The experimental findings have also shown the ability of the proposed probabilistic approach to reconstruct not just one, but multiple synthetic samples from a given iriscode. This not only significantly increases the success rate of the attack compared to methods that can generate only one synthetic sample from an iriscode, but it also opens up the possibility of other applications besides inverse biometrics such us its use for privacy preserving purposes.

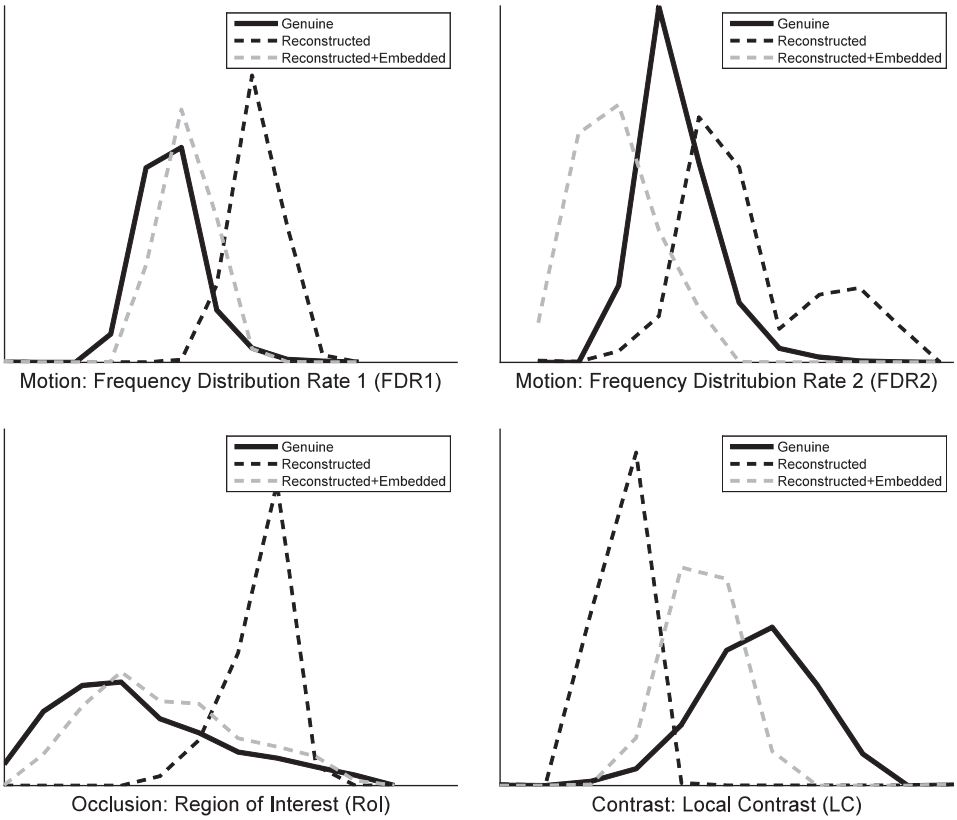

Fig. 11. Distributions for the four quality features considered in the quality assessment study for the three databases used in the experiments.

Biometric samples are personal data and different privacy concerns have arisen regarding their distribution and protection [58]. The proposed reconstruction method is able to generate synthetic iris patterns visually different to the original (see Fig. 7) which are, nevertheless, positively matched to the user's identity. This means that the synthetic samples may be considered as an alternative representation of the user's identity and, as such, may be stored in the database thereby avoiding possible privacy issues (e.g., deducing gender, age or ethnicity from the original iris images).

Furthermore, the work has reinforced the need for including template protection schemes in commercial iris systems as well as for adopting a verification strategy that confirms if the biometric samples presented to the system are those of a genuine eye and not that of a digital or physical artifact of the iris.

It may be argued that attacks such as the one considered in this work can be successful only when the template stored in the database is compromised. This may be difficult (although possible) in classical biometric systems where the enrolled templates are kept in a centralized database. In this case, the attacker would have to access the database and extract the information, or intercept the communication channel when the stored template is released for matching. But the threat is heightened in Match-on-Card (MoC) applications where an individual's biometric template is stored in a smartcard possessed by the person. Such applications are rapidly growing due to several appealing characteristics such as scalability and privacy [59]. Similarly, biometric data is being stored in many official documents such as the new biometric passport [60], some national ID cards [61], the US FIPS-201 Personal Identity Verification initiatives (PIV) [62] and the ILO Seafarers Identity Card Program [63]. In spite of the clear advantages that these type of applications offer, templates are more likely to be compromised as it is easier for the attacker to have physical access to the storage device and, as has already been demonstrated [64], fraudulently obtain the information contained inside. This makes MoC systems potentially more vulnerable to the type of threat described in this article especially when the biometric data is stored without any type of encryption [62], or printed in the clear on plastic cards as 2D barcodes [63].

Thus, there is an acute need to deflect the type of attack outlined in this article. This can be accomplished using two complementary approaches:

- *Prevention*. Here the goal is to avoid the users' templates from being compromised, for example by securely storing biometric data using encrypted templates [39,65] or protecting the communication channels through encryption [66].
- *Protection*. Here the goal is to minimize the probability of a successful attack even when a template is compromised. This could be accomplished by using biometric-based countermeasures to distinguish synthetic images from real iris images or to employ liveness-detection techniques [67].

Research work, such as the one presented in this article, or previous studies dealing with other modalities like fingerprint [23,3] or face [4], bring to the fore the difficulty in estimating the amount of information present within a biometric trait and the issue of biometric template generation.

Furthermore, from a security perspective, we believe that these examples may serve as a wake-up call for vendors and developers to be aware of the potential risks of not securing biometric templates, as is the case in some operational systems already installed in sensitive areas. There is an urgent need to design effective countermeasures that minimize the effects of these threats and increase the confidence of the end users in this rapidly emerging technology.

Acknowledgments

This work has been partially supported by projects Contexts (S2009/TIC-1485) from CAM, Bio-Challenge (TEC2009-11186) and Bio-Shield (TEC2012-34881) from Spanish MECD, TABULA RASA (FP7-ICT-257289) and BEAT (FP7-SEC-284989) from EU, and *Cátedra UAM-Telefónica*.

Javier Galbally was awarded with a fellowship "José Castillejo" from the Spanish MECD in order to carry out this work.

Arun Ross was supported by a grant from the National Science Foundation (NSF CAREER Award IIS 0642554).

Marta Gomez-Barrero is supported by a FPU Fellowship from Spanish MECD.

Javier Galbally would also like to thank Arun Ross of the Integrated Pattern Recognition and Biometrics Laboratory (iPRoBe) at West Virginia University for hosting him and collaborating with him during the development of this research work.

References

[1] A.K. Jain, A. Ross, S. Pankanti, Biometrics: a tool for information security, IEEE Trans. Inform. Foren. Secur. 1 (2) (2006) 125–143.
[2] J. Wayman, A. Jain, D. Maltoni, D. Maio, Biometric Systems. Technology, Design and Performance Evaluation, Springer, 2005.
[3] A. Ross, J. Shah, A.K. Jain, From template to image: reconstructing fingerprints from minutiae points, IEEE Trans. Pattern Anal. Mach. Intel. 29 (2007) 544–560.
[4] P. Mohanty, S. Sarkar, R. Kasturi, From scores to face templates: a model-based approach, IEEE Trans. Pattern Anal. Mach. Intel. 29 (2007) 2065–2078.
[5] A. Jain, P. Flynn, A. Ross (Eds.), Handbook of Biometrics, Springer, 2008.
[6] J. Daugman, How iris recognition works, in: Proc. IEEE Int. Conf. on Image Processing (ICIP), 2002, pp. I.33–I.36.
[7] J. Daugman, How iris recognition works, IEEE Trans. Circ. Syst. Video Technol. 14 (2004) 21–30.
[8] J. Daugman, Encyclopedia of Biometrics, Springer, 2009. Ch. Iris Encoding and Recognition Using Gabor Wavelets, pp. 787–797.
[9] International Biometric Group, Generating Images from Templates, White paper, 2002.
[10] J. Daugman, Probing the uniqueness and randomness of iris codes: results from 200 billion iris pair comparisons, Proceedings of the IEEE 94 (2006) 1927–1935.
[11] J. Cui, Y. Wang, J. Huang, T. Tan, Z. Sun, An iris image synthesis method based on pca and super-resolution, in: Proc. IAPR Int. Conf. on Pattern Recognition (ICPR), 2004, pp. 471–474.
[12] S. Makthal, A. Ross, Synthesis of iris images using markov random fields, in: Proc. of 13th European Signal Processing Conference (EUSIPCO), 2005.
[13] S. Shah, A. Ross, Generating synthetic irises by feature agglomeration, in: Proc. IEEE Int. Conf. on Image Processing (ICIP), 2006, pp. 317–320.
[14] J. Zuo, N.A. Schmid, X. Chen, On generation and analysis of synthetic iris images, IEEE Trans. Inform. Forens. Secur. 2 (2007) 77–90.
[15] Z. Wei, T. Tan, Z. Sun, Synthesis of large realistic iris databases using patch-based sampling, in: Proc. IAPR Int. Conf. of Pattern Recognition (ICPR), 2008, pp. 1–4.
[16] R. Cappelli, Handbook of Fingerprint Recognition, Springer, 2003. Ch. Synthetic Fingerprint Generation, pp. 203–231.
[17] J. Galbally, R. Plamondon, J. Fierrez, J. Ortega-Garcia, Synthetic on-line signature generation. Part I: Methodology, algorithms, Pattern Recognition 45 (2012) 2610–2621.
[18] J. Galbally, J. Fierrez, J. Ortega-Garcia, R. Plamondon, Synthetic on-line signature generation. Part II: Experimental validation, Pattern Recognition 45 (2012) 2622–2632.
[19] N. Poh, S. Marcel, S. Bengio, Improving face authentication using virtual samples, in: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2003.
[20] A. Lin, L. Wang, Style-preserving english handwriting synthesis, Pattern Recogn. 40 (2007) 2097–2109.
[21] T. Dutoit, An Introduction to Text-to-Speech Synthesis, Kluwer Academic Publishers (2001).
[22] C.J. Hill, Risk of Masquerade Arising from the Storage of Biometrics, Master's thesis, Australian National University, 2001.
[23] R. Cappelli, D. Maio, A. Lumini, D. Maltoni, Fingerprint image reconstruction from standard templates, IEEE Trans. Pattern Anal. Mach. Intel. 29 (2007) 1489–1503.
[24] J. Galbally, R. Cappelli, A. Lumini, G.G. de Rivera, D. Maltoni, J. Fierrez, J. Ortega-Garcia, D. Maio, An evaluation of direct and indirect attacks using fake fingers generated from ISO templates, Pattern Recogn. Lett. 31 (2010) 725–732.
[25] A. Adler, Sample images can be independently restored from face recognition templates, in: Proc. Canadian Conference on Electrical and Computer Engineering (CCECE), vol. 2, 2003, pp. 1163–1166.
[26] A. Adler, Images can be regenerated from quantized biometric match score data, in: Proc. Canadian Conference on Electrical and Computer Engineering (CCECE), 2004, pp. 469–472.
[27] J. Galbally, C. McCool, J. Fierrez, S. Marcel, On the vulnerability of face verification systems to hill-climbing attacks, Pattern Recogn. 43 (2010) 1027–1038.
[28] S. Venugopalan, M. Savvides, How to generate spoofed irises from an iris code template, IEEE Trans. Inform. Forens. Secur. 6 (2011) 385–394.
[29] L. Thalheim, J. Krissler, Body check: biometric access protection devices and their programs put to the test, ct Magazine (2002) 114–121.
[30] T. Matsumoto, Artificial irises: importance of vulnerability analysis, in: Proc. Asian Biometrics Workshop (AWB), vol. 45, 2004.
[31] J. Galbally, J. Fierrez, F. Alonso-Fernandez, M. Martinez-Diaz, Evaluation of direct attacks to fingerprint verification systems, J. Telecommun. Syst. Special Issue Biomet. Syst. Appl. 47 (2011) 243–254.
[32] R.N. Rodrigues, L.L. Ling, V. Govindaraju, Robustness of multimodal biometric fusion methods against spoof attacks, J. Visual Lang. Comput. 20 (2009) 169–179.
[33] U. Uludag, A. Jain, Attacks on biometric systems: a case study in fingerprints, in: Proc. SPIE Seganography and Watermarking of Multimedia Contents VI, vol. 5306, 2004, pp. 622–633.
[34] M. Martinez-Diaz, J. Fierrez, J. Galbally, J. Ortega-Garcia, An evaluation of indirect attacks and countermeasures in fingerprint verification systems, Pattern Recogn. Lett. 32 (2011) 1643–1651.
[35] CC, Common Criteria for Information Technology Security Evaluation, v3.1, 2006. <http://www.commoncriteriaportal.org/>.
[36] BEM, Biometric Evaluation Methodology. v1.0, 2002.
[37] ISO/IEC 19792, ISO/IEC 19792:2009, Information Technology – Security Techniques – Security Evaluation of Biometrics, 2009.
[38] Tabula Rasa, Trusted biometrics under spoofing attacks (tabula rasa), 2010, (http://www.tabularasa-euproject.org/). <http://www.tabularasa-euproject.org/>.
[39] J. Daugman, The importance of being random: statistical principles of iris recognition, Pattern Recogn. 36 (2003) 279–291.
[40] J. Daugman, New methods in iris recognition, IEEE Trans. Syst. Man Cybernet. – Part B: Cybernet. 37 (2007) 1167–1175.
[41] J. Daugman, Iris Recognition, Springer, 2008 (Chapter 4, pp. 71–90).
[42] K. Bowyer, K. Hollingsworth, P. Flynn, Image understanding for iris biometrics: a survey, Comput. Vis. Image Understand. 110 (2008) 281–307.
[43] D.E. Goldberg, Genetic Algorithms in Search Optimization and Machine Learning, Addison Wesley (1989).
[44] J.E. Baker, Reducing bias and inefficiency in the selection algorithm, in: Proc. Int. Conf. on Genetic Algorithms and their Application (ICGAA), L. Erlbaum Associates Inc., 1987, pp. 14–21.
[45] C. Rathgeb, A. Uhl, Attacking iris recognition: an efficient hill-climbing technique, in: Proc. Int. Conf. on Pattern Recognition (ICPR), 2010, pp. 1217–1220.
[46] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M.R. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J.-L. Alba-Castro, E. Gonzalez-Agulla, E. Otero-Muras, S. Garcia-Salicetti, L. Allano, B. Ly-Van, B. Dorizzi, J. Kittler, T. Bourlai, N. Poh, F. Deravi, M.W.R. Ng, M. Fairhurst, J. Hennebert, A. Humm, M. Tistarelli, L. Brodo, J. Richiardi, A. Drygajlo, H. Ganster, F.M. Sukno, S.-K. Pavani, A. Frangi, L. Akarun, A. Savran, The multi-scenario multi-environment Bio Secure multimodal database (BMDB), IEEE Trans. on Pattern Analysis and Machine Intelligence 32 (2010) 1097–1111.
[47] A. Mayoue, B. Dorizzi, L. Allano, G. Chollet, J. Hennebert, D. Petrovska-Delacretaz, F. Verdet, Guide to biometric reference systems and performance evaluation, Springer, 2009. Ch. BioSecure multimodal evaluation campaign 2007 (BMEC 2007), pp. 327–372.
[48] L. Masek, P. Kovesi, Matlab source code for a biometric identification system based on iris patterns, Master's thesis, School of Computer Science and Software Engineering, University of Western Australia, 2003.
[49] V. Ruiz-Albacete, P. Tome-Gonzalez, F. Alonso-Fernandez, J. Galbally, J. Fierrez, J. Ortega-Garcia, Direct attacks using fake images in iris verification, in: Proc. COST 2101 Workshop on Biometrics and Identity Management (BioID), LNCS, vol. 5372, Springer, 2008, pp. 181–190.
[50] Neurotechnology. <http://www.neurotechnology.com/verieye.html>.
[51] P. Grother, E. Tabassi, G. W. Quinn, W. Salamon, IREX I: performance of iris recognition algorithms on standard images, Tech. rep., National Institute of Standards and Technology, 2009.
[52] D.E. Goldberg, The Design of Innovation: Lessons from and for Competent Genetic Algorithms, Kluwer Academic Publishers (2002).
[53] ANSI-NIST, ANSI x9.84-2001, Biometric Information Management and Security, 2001.
[54] L. Ma, T. Tan, Y. Wang, D. Zhang, Personal identification based on iris texture analysis, IEEE Trans. Pattern Anal. Mach. Intel. 25 (2003) 1519–1533.
[55] Z. Wei, T. Tan, Z. Sun, J. Cui, Robust and fast assessment of iris image quality, in: Proc. IAPR Int. Conf. on Biometrics (ICB), LNCS, vol. 3832, Springer, 2006, pp. 464–471.
[56] A. Abhyankar, S. Schukers, Iris quality assessment and bi-orthogonal wavelet based encoding for recognition, Pattern Recogn. 42 (2009) 1878–1894.
[57] J. Galbally, F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, A high performance fingerprint liveness detection method based on quality related features, Future Gener. Comput. Syst. 28 (2012) 311–321.
[58] S. Prabhakar, S. Pankanti, A.K. Jain, Biometric recognition: security and privacy concerns, IEEE Secur. Privacy 1 (2003) 33–42.

[59] C. Bergman, Advances in Biometrics: Sensors, Algorithms and Systems, Springer, 2008. Ch. Match-on-card for Secure and Scalable Biometric Authentication, pp. 407–422.
[60] ICAO, ICAO document 9303, Part 1, Volume 2: Machine Readable Passports – Specifications for Electronically Enabled Passports with Biometric Identification Capability, 2006.
[61] Government of Spain. <http://www.dnielectronico.es/>.
[62] NIST, NIST Special Publication 800-76, Biometric Data Specification for Personal Identity Verification, 2005.
[63] ILO, ILO SID-0002, Finger Minutiae-Based Biometric Profile for Seafarers Identity Documents, Intl Labour Organization, 2006.
[64] J. van Beek, ePassports reloaded, in: Black Hat USA Briefings, 2008.
[65] A. Ross, A. Othman, Visual cryptography for biometric privacy, IEEE Trans. Inform. Foren. Secur. 6 (2011) 70–81.
[66] U. Uludag, S. Pankanti, S. Prabhakar, A.K. Jain, Biometric cryptosystems: issues and challenges, Proc. IEEE 92 (2004) 948–960.
[67] Z. Wei, X. Qiu, Z. Sun, T. Tan, Counterfeit iris detection based on texture analysis, in: Proc. IAPR Int. Conf. on Pattern Recognition (ICPR), 2008.