

# SYNTHETIC IRIS IMAGES FROM IRIS PATTERNS BY MEANS OF EVOLUTIONARY STRATEGIES

## *How to Deceive a Biometric System based on Iris Recognition*

Alberto de Santos Sierra, Javier Guerra Casanova

*Centro de Domótica Integral, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Madrid, Spain*

Carmen Sánchez Ávila, Vicente Jara Vera

*Applied Mathematics Department, ETSIT, Polytechnical University of Madrid, Madrid, Spain*

**Keywords:** Biometric iris recognition, Synthetic iris images, Iris falsification, Security, Evolutionary strategies.

**Abstract:** Synthetic Biometric is emerging nowadays as a new research field in biometrics. An artificial iris tissue or a synthetic fingerprint could compromise the security, allowing a non-registered individual to enter the system. However, inverse biometric can also improve current identification systems, enhancing not only its strength against fake-based attacks, but also by replicating unavailable or corrupted data, due to a bad acquisition, for instance. The methods proposed in this document aim to provide a procedure to create a synthetic iris tissue from a stored biometric template, so that a non-registered user could access the system under a registered identity. These algorithms will come out with the result that synthetic sample could be so similar to original as desired.

## 1 INTRODUCTION

Biometric systems based on Iris Recognition are known to be one of the most reliable systems in terms of security (Daugman, 1993; Daugman, 2004). Considering the fact that Iris is already stable in very early stages of life and unalterable since that moment, it is obviously impossible to modify the iris itself. Thus, any method attempting to deceive a system based on iris recognition should focus on the acquired image, which is already stored in the system when captured.

Despite of being suggested that there exists no previous development of an automated iris reconstruction, (Yanushkevich, 2006), there is indeed a previous work on iris synthesis, (Cui et al., 2004), but focused on creating synthetic biometric databases for testing purpose, far beyond the scope provided in this paper, where the aim consists on accessing the system by means of synthetic iris images. On the other hand, several attempts already exist in other biometric techniques such as fingerprint (Capelli et al., 2006), face recognition (Yanushkevich et al., 2007), handwritten signature (Guyon, 1996; Yanushkevich, 2006), voice (Cook, 2002), and so forth.

Overall biometric systems relies their strength on the fact that once the templates are extracted, there is

no function able to find an acquisition, whatever its nature (fingerprint, iris, hand veins), whose template coincides with the required pattern.

This paper provides two approaches attempting to implement such a function, i.e., a function able to come out with an image iris, whose pattern may be so similar to the desired template as required. Based on evolutionary strategies, the procedure will provide the best solution to meet the required goal. Therefore, this algorithm may be extended to other biometric techniques, and all biometric systems will benefit of it, since their main applications consider not only enhancing the system against fakes, but also improving its performance, by replicating corrupted data or testing algorithms with synthetic databases.

## 2 CREATING A SYNTHETIC IRIS

The main problem this approach tackles with is about creating an artificial iris tissue. In order to meet this goal, two approaches have been implemented with a twofold purpose: first, providing two possible solutions, and second, comparing them in terms of visual aspect and overall performance.

Each approach is based on evolutionary strategies

(Eiben and Smith, 2003; Schwefel, 1995), but the schemes on which they are based are completely different. Nonetheless, before describing the algorithms in detail, the problem to be solved must be stated.

## 2.1 Problem Statement

The problem these two approaches attempt to tackle with can be stated mathematically as a function  $\mathcal{A}$ , whose input is the template of a user already registered in the database,  $T_{RU}$ . More in detail,  $\mathcal{A}$  can be expressed as:

$$\begin{aligned} \mathcal{A} : \mathcal{T} &\mapsto \mathcal{D} \\ T \in \mathcal{T} &\mapsto \mathcal{A}(T) \in \mathcal{D} \end{aligned} \quad (1)$$

where  $\mathcal{T}$  represents the set of all possible templates in a Biometric System, in other words,  $\mathcal{T}$  is the information stored in the biometric database, (Daugman, 2004); and  $\mathcal{D}$  stands for the set of all possible images of a given database, related to previous  $\mathcal{T}$ . Both sets verify to have the same number of elements, same cardinality, and each element in  $\mathcal{T}$  corresponds to only one element in  $\mathcal{D}$ , and vice-versa.

This paper will suppose that the information of the database is available, but in most current systems, the stored information is protected and encrypted, so unveiling patterns in  $\mathcal{T}$  is an arduous, but possible, task.

As an overview,  $\mathcal{A}$  is a function which provides an image  $\mathcal{A}(T) \in \mathcal{D}$  when being applied to a template  $T \in \mathcal{T}$ .

Continuing on mathematical representation,  $\mathcal{Z}$  is defined as the algorithm responsible for extracting the template from an image. In other words,

$$\begin{aligned} \mathcal{Z} : \mathcal{D} &\mapsto \mathcal{T} \\ I \in \mathcal{D} &\mapsto \mathcal{Z}(I) \in \mathcal{T} \end{aligned} \quad (2)$$

In fact,  $\mathcal{Z}(\mathcal{D}) = \mathcal{T}$ , because each element in  $\mathcal{T}$  corresponds to only one element in  $\mathcal{D}$ .

Furthermore, a measure must be defined in order to assess how similar two templates are. Let  $\eta \in \mathbb{R}$  be a function able to measure to what extend two templates coincides or not. This measurement can be defined in terms of  $\mathcal{T}$  as follows,

$$\begin{aligned} \eta : \mathcal{T} \times \mathcal{T} &\mapsto \mathbb{R} \\ T_1, T_2 \in \mathcal{T} &\mapsto \eta(T_1, T_2) \end{aligned} \quad (3)$$

For the sake of simplicity,  $\eta$  is implemented by Euclidean distance in this document, but other measures can be deployed, (González et al., 2004). In any case,  $\eta(T_1, T_2)$  will tend to zero, when  $T_1$  tends to  $T_2$ .

Finally, considering previous equations Eq. (1)-(3), the main problem can be stated as:

Given  $T_{RU}$  (a registered user template) and  $\eta_0$  (a threshold), implement  $\mathcal{A}$  in order to verify that

$\eta(\tilde{T}, T_{RU}) \leq \eta_0$ , where  $\tilde{T} = \mathcal{Z}(\mathcal{A}(T_{RU}))$ . In other words, the problem can be defined in terms of functions as:

Attempt to find  $\mathcal{A}$  so that  $\mathcal{A} = \mathcal{Z}^{-1}$ .

Before dealing with the problem of finding  $\mathcal{A}$ , several constraints concerning  $\mathcal{Z}$  should be made. First of all, and for the sake of simplicity,  $\mathcal{Z}$  will be defined as in previous approaches (Sánchez-Ávila and Sánchez-Reillo, 2002; de Santos-Sierra et al., 2007), losing some generality. This approach will be explained in next section 2.2. However, the main ideas are suitable not only for other feature extractors (Boles and Boashash, 1998; Chun and Chung, 2004; Daugman, 2004), but also for other biometric techniques. Secondly,  $\mathcal{A}$  has been tested in CASIA v3 database (<http://www.sinobiometrics.com>), a low quality database, so that,  $\mathcal{Z}$  will provide low resolution templates. Testing these algorithms in different databases, remains as a future work.

## 2.2 Template Description

Evidently, a falsification must consider how the template is extracted, otherwise the problem could be unattainable. The extraction step begins after pre-processing the image, once the iris has been isolated from pupil, eyelids, and so forth, (Daugman, 1993; de Santos-Sierra et al., 2007). Furthermore, a first 'raw' template is extracted from the isolated iris, being afterwards processed in order to obtain a better result in identification.

This approach regards only these previous 'raw' data, for the sake of simplicity. Note these considerations do not lack of generality, since posterior processing is inherent to each Biometric System, and furthermore, such post-processing is easily reversible due to the algorithms involved in that step, (Sánchez-Ávila and Sánchez-Reillo, 2002; de Santos-Sierra et al., 2007).

The template is extracted by  $\mathcal{Z}$ , using a circular section (Fig. 1), averaging the intensities values along the radius of the crown,  $\rho$ , for each angle of the previous section,  $\alpha_i$ ,  $\forall i \in \mathbb{N} \cap [0, 255]$ , (de Santos-Sierra et al., 2007).

The circular section provides a template of 256 points, where each point is represented by a double precision value in  $[0, 255]$ , due to the grayscale representation of the iris image. By the use of a  $256^\circ$  circular section, eyelids are avoided, although a  $360^\circ$  section can be considered when suitable, regarding eye aperture, (de Santos-Sierra et al., 2007).

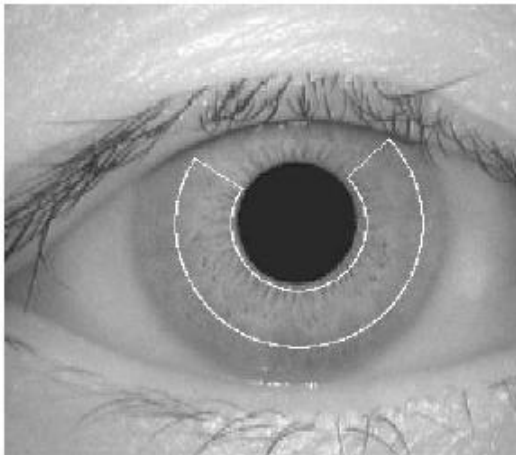


Figure 1: Region of Interest in Feature Extraction.

## 2.3 From Original to Synthetic

Once the template extraction has been presented, the reader could realize that not the whole image is considered as a part of the template, but only many pixels. In fact, those pixels are included in the previous circular section. Furthermore, a certain image  $I$  is considered as a base image, from which the algorithms will be able to evolve a fake image. This must be done, if a close-human fake is required. It is far easier to fake several points in an image than the whole image.

According to these previous statements, several definitions must be considered, before starting with the explanation of the approaches. First of all, the image  $I$  is randomly selected. From now,  $I$  will be considered as a matrix, where  $X$  and  $Y$  represent the sets for horizontal and vertical coordinates respectively, assuming that a pixel of  $I$  is represented by its position within the image,  $(x, y) \in X \times Y$ , and its color intensity,  $I(x, y) \in \mathbb{N} \cap [0, 255]$ , (González et al., 2004).

Let  $V$  be the matrix containing all the intensity values within previous circular section. Matrix  $V$  has 256 columns and  $\rho$  rows, where  $\rho$  is the previous circular section width. Therefore,  $\Xi \subset X$  and  $\Upsilon \subset Y$  are defined as those positions, in terms of  $X$  and  $Y$ , whose pixels are contained in  $V$ . In other words,  $\Xi$  and  $\Upsilon$  are the matricial representation of previous circular section.

Matrix  $V$  is calculated once for image  $I$ , since  $V$  gathers the unique values to be modified to fake image  $I$ . When image  $I$  is presented to the system, feature extraction step will come up with a template read from those pixels located in  $\Xi$  and  $\Upsilon$ , the rest of  $I$  is neglected for template extraction.

In fact, template  $T$  associated with image  $I$  can be obtained as:

$$T(j) = \frac{\sum_i V(i, j)}{\rho} \quad (4)$$

Once the nomenclature is stated, the two approaches are presented. First approach regards evolutionary algorithms, (Michalewicz, 1996), finding optimal solution by evolving those pixels contained in  $V$ . On the other hand, second approach contemplates the possibility of changing the order in  $\Xi$  and  $\Upsilon$ , in other words, changing the position of each pixel, but not altering its intensity. Both approaches entail evolutionary strategies, but with different approaches. The performance of both will be presented in section 3.

### 2.3.1 Evolutionary Strategies on $V$

An evolutionary algorithm is a suitable manner of tackling minimization and maximization problems, (Michalewicz, 1996; Eiben and Smith, 2003). However, the problem must be adapted to make it suitable for evolutionary strategies. Since each template  $T \in \mathcal{T}$  verifies to be a 256 components vector, the strategy involves dividing the problem into 256 independent problems. The  $j$ -component of  $T$  corresponds to the average of the components in column  $j$  from matrix  $V$ , being the representation for the subproblem  $j$  a vector called  $v_j = \langle a_{1,j}, a_{2,j}, \dots, a_{\rho,j} \rangle$  with  $\rho$  components and  $a_{i,j} \in V$  where  $i \in \{1, \dots, \rho\}$  and  $j \in \{1, \dots, 256\}$ .

The procedure deployed under this section contemplates several steps before meeting the goal of obtaining an optimal solution:

- *Fitness Function:* This function provides the algorithm with a *Termination Condition*. The evolutionary algorithm will seek the solution until a certain value of fitness function is achieved. An appropriate fitness function is as follows in Eq. 5.

$$|T(j) - \bar{v}_j| = \left| T(j) - \frac{\sum_{i=1}^{\rho} a_i}{\rho} \right| \quad (5)$$

This fitness function must be minimized until a certain extent  $\eta_0$  according to previous problem statement. This function exists for each subproblem.

- *Population and Initialization:* The population of each subproblem consists of 50 members, each of them different to each other, in order to ensure *diversity* property, (Eiben and Smith, 2003). Each member of the population is created based on a uniform distribution. Such a distribution is uniform in the set formed by minimum and maximum value of column  $j$  for each  $j$ -subproblem.

For each subproblem, 50 members of the population (vectors) are created. The reason why this initialization is performed is due to the fact that they present a previous solution quite close to the final result which still remains unknown.

- *Parent Selection:* Parent Selection will provide those individuals in the population most suitable to mate, without neglecting the fact that less suitable individuals should be also considered as possible parents to create next offspring.

Thus, every individual is selected based on:

$$\xi_j = |T(j) - \bar{v}_j|$$

which indicates to what extent the results differs from aim. According to  $\xi_j$ , those individuals more capable (with  $\xi_j \rightarrow 0$ ) are more likely to be selected than those whose  $\xi_j$  is far from 0. In the literature, (Eiben and Smith, 2003), the Parent Selection described within this document is similar to Fitness Proportional Selection.

- *Recombination:* Having parents selected, a new offspring emerges as a result of mating previous parents. However, a recombination must be done in order to provide more richness to such offspring. The algorithm proposed in this approach is known as Whole Arithmetic Recombination, (Eiben and Smith, 2003), with its parameters set to  $\alpha = 0.3$ , accordingly to (Eiben and Smith, 2003; Back et al., 2000a; Back et al., 2000b), where  $0 < \alpha < 1$  is suggested. On the other hand, a Uniform Crossover operator is proposed to carry out such an operation with a  $p_c = 0.7$  probability. Lower values may make the algorithm not to converge in a reasonable time, (Back et al., 2000a; Back et al., 2000b).
- *Mutation:* Mutation is an operator capable of modifying the offspring by changing several elements in the alleles (individuals). For the sake of an optimal solution, mutation must always be carried out, but with lower probability than Recombination. In this work, mutation probability was experimentally set to  $p_m = 0.25$ . Higher values of  $p_m$  make the algorithm not to converge, (Eiben and Smith, 2003; Back et al., 2000a; Back et al., 2000b).
- *Survivor Selection Mechanism:* This mechanism is responsible for managing the process whereby the population of parents ( $\mu$ ) and the new offspring ( $\lambda$ ) is reduced to the size of population. Concretely, this approach considers Elitism, (Eiben and Smith, 2003), as the most suitable mechanism for survivor selection, since it combines Age-based replacement and Fitness-based

Replacement, in the same way parent selection is carried out.

Finally, the algorithm under this approach will be indicated as  $\mathcal{A}_V$ , as it is based on  $V$  matrix. The performance of this algorithm will be considered in section 3.

### 2.3.2 Evolutionary Strategies on $\Xi$ and $\Upsilon$

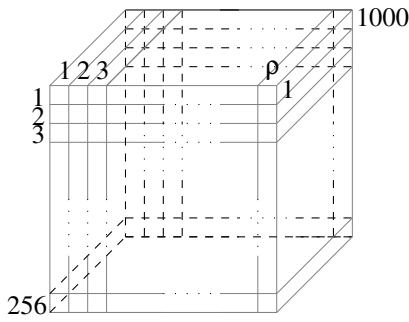
This approach, namely  $\mathcal{A}_{\Xi, \Upsilon}$ , despite of being different from previous algorithm, attempts the same aim. Nonetheless, this approach intends to fake the iris, by rotating rows in matrix  $V$ . In other words, this procedure attempts to find the right combination for what a solution is found in terms of problem statement.

Matrix  $V$  has  $\rho$  columns and 256 rows and is modified in column directions. In fact, the algorithm  $\mathcal{A}$  is conceived in this approach as an operator responsible for shifting the columns of  $V$ , so that the template  $T$  changes as the evolutionary strategies attempts to find the optimal solution. The operator able to carry out every shift is called  $\sigma_i^m$  with  $i = \{1, \dots, nPop\}$  and  $m = \{1, \dots, \rho\}$ , where nPop is the number of individuals in the population of the evolutionary algorithm. This value will be set to nPop = 1000, according to (Eiben and Smith, 2003; Back et al., 2000a; Back et al., 2000b).

Despite of considering only shifts in row direction, there exist the possibility of also shifting in column directions. The operator responsible for this strategy is defined as  $\alpha_j^m$  with  $j = \{1, \dots, 256\}$  and  $m = \{1, \dots, \rho\}$ . Although possible, the implementation of an evolutionary algorithm considering both  $\sigma_i^m$  and  $\alpha_j^m$  involves a non-acceptable processing time and the results compared to an approach using only  $\sigma$  operator, does not differ significantly. Thus,  $\sigma_i^m$  operator is considered, but  $\alpha_j^m$  operator is ignored, although its integration remains as future work.

The description of this approach follows in the same manner as previous algorithm was described. Most of the concepts coincide with those above in section 2.3.1, since most representation has been conserved, for the sake of simplicity regarding implementation.

- *Fitness Function:* Same function as in previous section is selected. However, several considerations must be taken into account. For instance, threshold  $\eta_0$  cannot be assured to be achieved, but minimum of fitness function can be obtained, although the performance of the whole algorithm could not be so promising as previous approach. Note the constraints of this algorithm are more restrictive than before, being on the contrary less complex in terms of processing time. Thus, same


 Figure 2: Population structure for  $\mathcal{A}_{\Xi, \Upsilon}$ .

function is considered as in Eq. (5) but searching the minimum, and not a defined threshold  $\eta_0$ .

- **Population and Initialization:** In this case, 1000 individuals for each population is selected, since a wider population is require due to the simplicity of the operator  $\sigma_i^m$ , (Eiben and Smith, 2003). Each individual represent a  $V$  matrix, where the operator  $\sigma_i^m$  has been carried out. Fig. 2 intends to clarify the structure of the population and provides a visual idea about the population. However,  $\sigma_i^m$  is the actual operator to be evolved, although it is easier to think on the problem in terms of matrices  $\Xi$  and  $\Upsilon$ .
- **Parent Selection:** Since this approach slightly differs from previous approach in the fact that what is evolved is not the data, but the operators  $\sigma_i^m$ , parents must be selected based on those displacements which provide a  $\xi_j$  (as defined in Eq. (4)) close to zero.
- **Recombination and Mutation:** Both steps have been gathered under this section. The operators evolved are  $\sigma_i^m$  and thus, the operators used before lack of any meaning in this section. Instead, permutation operators are involved. More in detail, Partially Mapped Crossover was carried out for recombination with a probability of  $p_c = .7$ , and Swap Mutation was considered as a suitable operator for Mutation, carried out with a probability of  $p_m = .25$ , (Eiben and Smith, 2003; Back et al., 2000a; Back et al., 2000b). These previous values were set according to experimental results, lower values for  $p_c$  and higher values for  $p_m$  made the algorithm not converge.
- **Survivor Selection Mechanism:** As previously stated, Elitism is selected as the most suitable solution, based on experimental results.

### 3 RESULTS

Considering the aim of these algorithms, a function is required to measure how similar an original and a fake are. Many different proposals exist in literature, (González et al., 2004), in order to assess how different two signals are. The selected function, namely  $\varepsilon$ , provide with a fast and accurate manner of meeting this goal, and coincides with *fitness function* in its definition. Given two functions,  $f$  and  $g$ ,  $\varepsilon$  is defined as follows:

$$\varepsilon = \int |f - g| \quad (6)$$

Notice that this function is the simplest measurement function. Main reason to use this function regards the fact that if with the simplest function the results are promising and acceptable, (see section 3), more complicated measurement functions will provide more thorough and precise results. Nonetheless, an implementation with more complex functions remains as future work.

Once evaluation function is defined, an experiment comparing deviation in terms of  $\varepsilon$  was carried out among different samples of same individuals, and different samples of their corresponding synthetic copies. Next picture highlights the results obtained in these experiments: horizontal lines indicate the average ( $\mu_{\text{Orig}}$ ,  $\text{---}\bullet\text{---}$ ) and standard deviation range ( $\sigma_{\text{Orig}}$ ,  $\text{---}\bullet\text{---}$ ) of  $\varepsilon$  among real samples of a same individual. On the other hand, dotted lines ( $\text{---}\diamond\text{---}$  and  $\text{---}\blacksquare\text{---}$ ) indicate the performance of the algorithms  $\mathcal{A}_V$  and  $\mathcal{A}_{\Xi, \Upsilon}$  respectively. Notice how horizontal lines split the space in three different regions: a region where falsifications are more similar to original than real samples of the same user from which original was taken ('Best'); a region where fake is similar to real samples ('Good') and a region where falsifications could not deceive the system ('Bad').

Since both implementations of the algorithm,  $\mathcal{A}_V$  and  $\mathcal{A}_{\Xi, \Upsilon}$ , depends on  $\eta$ , several performance of both implementations were carried out for each value of  $\eta$ , (Fig. 3). However, notice that although  $\mathcal{A}_{\Xi, \Upsilon}$  depends on  $\eta$ , its aim attempts to achieve the closest minimum to  $\eta_0$ , which might not coincide with  $\eta_0$ .

Several conclusions can be extracted from Fig. 3. First of all,  $\mathcal{A}_V$  and  $\mathcal{A}_{\Xi, \Upsilon}$  come out with solutions more similar to their respective originals (from which they were copied) than even those samples taken from the same individual from which original images were taken. In other words, depending on  $\eta$ , falsifications can be 'Good' if  $\varepsilon$  is in the deviation interval among real samples, better if  $\varepsilon$  is under that range, or 'Bad' otherwise. In this former case, the result provided by  $\mathcal{A}_V$  and  $\mathcal{A}_{\Xi, \Upsilon}$  is not considered as a valid solution.

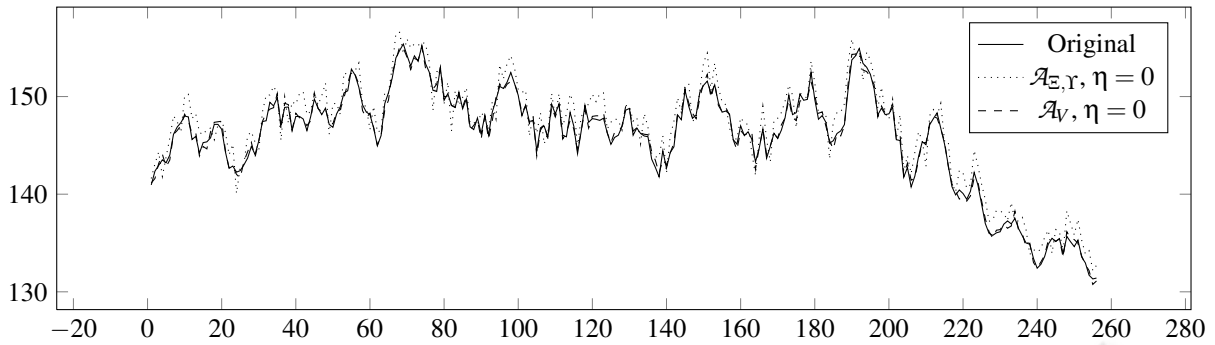


Figure 4: Result of both approaches,  $\mathcal{A}_V$  and  $\mathcal{A}_{E,\gamma}$ , compared to the original pattern.

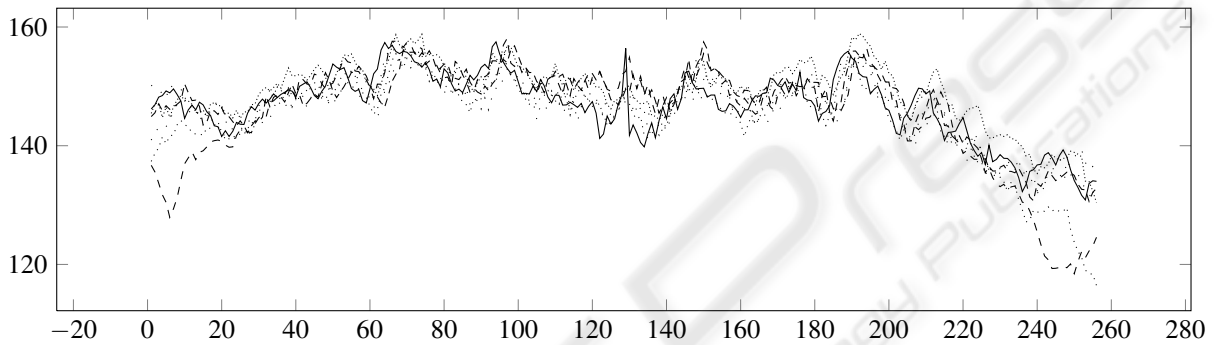


Figure 5: Patterns from different samples belonging to the same user.

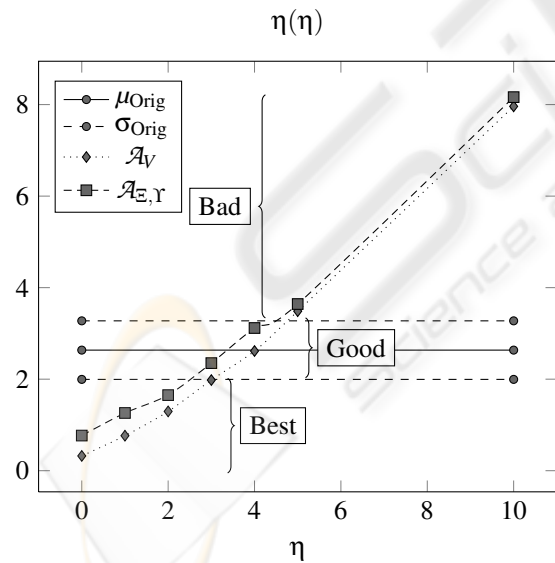


Figure 3: Comparison between  $\mathcal{A}_V$ ,  $\mathcal{A}_{E,\gamma}$  and original samples in terms of  $\eta$ .

The performance of  $\mathcal{A}_V$  and  $\mathcal{A}_{E,\gamma}$  is presented in Fig. 4. Notice that  $\mathcal{A}_V$  is almost equal to the original, and therefore, some difficulties can arise when trying to distinguish Original from  $\mathcal{A}_V$  in Fig. 4. A vi-

sual evaluation of the performance of both algorithms can be seen in Fig. 8 and in Fig. 9, where (—) represents an original pattern, (.....) represents a performance of  $\mathcal{A}_{E,\gamma}$  with  $\eta = 0$  and (---) represents a performance of  $\mathcal{A}_V$  with  $\eta = 0$ . The performance of  $\mathcal{A}_V$  is in general better than  $\mathcal{A}_{E,\gamma}$  due to its own definition, however both provide optimal solutions when compared to real data from same individuals.

This statement can be assessed in Fig. 5, where patterns of the same user are depicted. Notice how different are two samples from the same user. In fact, that difference is far more noticeable among original samples belonging to the same user, than among falsification of a certain original sample. Furthermore, the fake provided by  $\mathcal{A}_V$  (---) is extremely similar to original.

Together with these previous visual patterns, and important question comes up: How does a fake iris image look like? Obviously, reader must notice that a considerable amount of pixel within the image have been altered, so that the image may look strange to a human. However, that appearance lacks of meaning for a computer, since the system extracts only that region suitable for a posterior pattern extraction, ignoring how 'human' the iris looks like.

Thus, although an altered image by these former

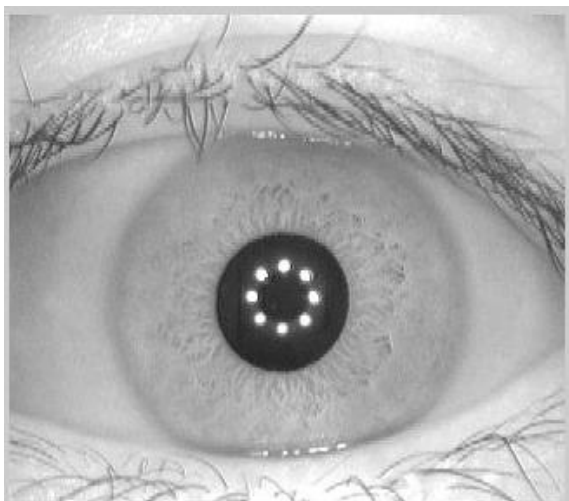


Figure 6: Image where falsification will take place.

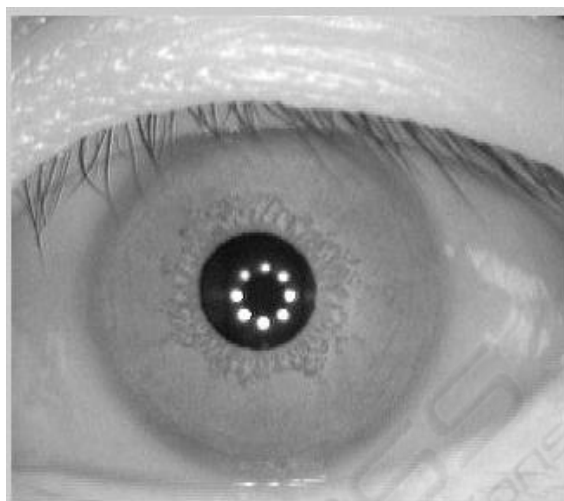


Figure 7: Image whose template is  $T_{RU}$ .

methods could seem weird at first sight, it is totally accepted by the computer. In the end, the computer will be responsible of processing the image, not the human. Actually, providing the synthetic image with a more human-like appearance remains as future work.

However, the main difference between  $\mathcal{A}_V$  and  $\mathcal{A}_{\varepsilon, \gamma}$  consists of the fact that  $\mathcal{A}_V$  achieves more precise results and can obtain a solution closer to the target, but on the contrary, the visual aspect of a fake image is far from being real, in other words, it is obvious that the image has been altered. On the other hand,  $\mathcal{A}_{\varepsilon, \gamma}$  achieves worst results and is more time-consuming. However, the visual aspect provided by this latter approach ( $\mathcal{A}_{\varepsilon, \gamma}$ ) does not look very different from a human iris, and in fact, this difference is negligible.

Fig. 6 shows the image used to make falsification. This image is chosen randomly, and the algorithm does not depend on this previous image selection. Moreover, Fig. 7 provides the image to fake. In other words, the idea is to enter the system where Fig. 7 is registered, using Fig. 6.

The results of the algorithms are shown in Fig. 8 and in Fig. 9, which represent the performance of  $\mathcal{A}_V$  and  $\mathcal{A}_{\varepsilon, \gamma}$  respectively. Notice how the appearance of Fig. 8 is slightly different from a human iris, and how Fig. 9 almost does not differ from Fig. 6. It must be emphasized the fact that although visually they do not look ‘standard’ human iris, the system makes no difference. Furthermore, achieving a more human appearance in both performance remains as future work.

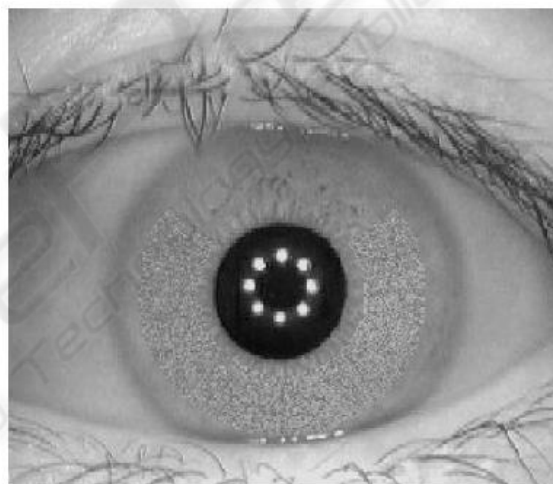


Figure 8: Visual result of  $\mathcal{A}_V$ .

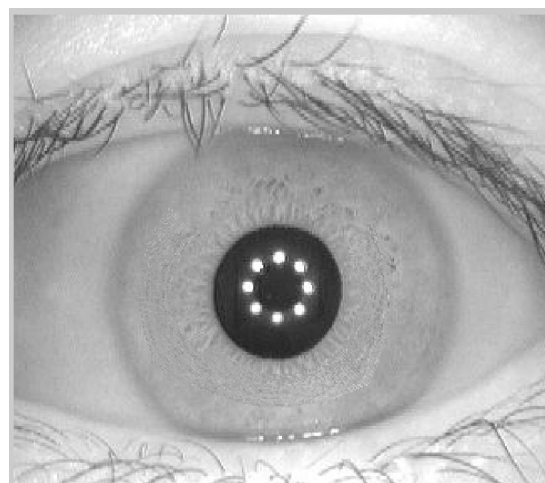


Figure 9: Visual result of  $\mathcal{A}_{\varepsilon, \gamma}$ .

## 4 CONCLUSIONS

This article presents two approaches for biometric synthesis. Both provide an accurate performance when creating a fake copy from an original pattern, although both compromise accuracy, time and human appearance. Main conclusion highlights the fact that copies are better than original samples.

Use cases inferred from synthetic biometric data regard biometric data reconstruction, synthetic data for testing biometric systems, or tightening up biometric security.

Regarding future work, stricter constraints will lead to a more human-like appearance in final synthetic image, focusing on the texture. Furthermore, tests will be carried out in more databases apart from CASIA v3, once the algorithm is able to deal with more complex textures. Despite of being far beyond the scope of this article, time performance must be considered for a real time application. However, reader must notice that time is not a constraint when doing 'hacking' activities.

One more aspect regards performance of the whole biometric system when identifying/authenticating a certain individual. No efforts have been done focusing on this aspect, since synthetic copies are so similar to originals, that the performance of the biometric system will be obviously decreased if these images are included within the database. However, to what extent a biometric system loses accuracy with these images remains as future work.

Finally, a more complex algorithm, merged as a result of both approaches, will be considered as a more precise and promising procedure, gathering best characteristics of each algorithm, including also  $\alpha_j^m$  operator as described before, (González et al., 2004).

## REFERENCES

- Back, T., Fogel, D. B., and Michalewicz, Z. (2000a). *Evolutionary Computation 1: Basic Algorithms and Operators*. Taylor & Francis, Bristol, 1<sup>st</sup> edition.
- Back, T., Fogel, D. B., and Michalewicz, Z. (2000b). *Evolutionary Computation 2: Advanced Algorithms and Operators*. Taylor & Francis, Bristol, 1<sup>st</sup> edition.
- Boles, W. and Boashash, B. (1998). A human identification technique using images of the iris and wavelet transform. In *IEEE Trans. Signal Processing*, volume 46, page 1185-1188.
- Capelli, R., Lumini, A., Maio, D., and Maltoni, D. (2006). Can fingerprints be reconstructed from iso templates? In *Proc. International Conference on Control, Automation, Robotics and Vision (ICARCV2006)*.
- Chun, C. N. and Chung, R. (2004). Iris recognition for palm-top application. In *International Conference on Biometric Authentication (ICBA 2004)*, volume 3072, pages 426-433. Springer-Verlag.
- Cook, P. R. (2002). *Real Sound Synthesis for Interactive Applications*. AK. Peters, 1<sup>st</sup> edition.
- Cui, J., Wang, Y., Huang, J., Tan, T., Sun, Z., and Ma, L. (2004). An iris image synthesis method based on pca and super-resolution. In *Proc. Int. Conf. on Pattern Recognition*.
- Daugman, J. (1993). High confidence visual recognition of persons by a test of statistical independence. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 15.
- Daugman, J. (2004). How iris recognition works. In *IEEE Transactions on Circuits and Systems For Video Technology*, volume 14.
- de Santos-Sierra, A., Sánchez-Ávila, C., and Sánchez-Reillo, R. (2007). Sistema de identificación biométrica mediante patrón de iris utilizando operadores morfológicos y representación. In *Congreso Iberoamericano de Seguridad Informática (CIBSI2007)*, pages 427-434.
- Eiben, A. E. and Smith, J. E. (2003). *Introduction to Evolutionary Computing*. Springer, Berlin.
- González, R. C., Woods, R. E., and Eddins, S. L. (2004). *Digital Image Processing*. Prentice Hall, 2<sup>nd</sup> edition.
- Guyon, I. (1996). Handwriting synthesis from handwritten glyphs. In *Proc. 5th Int. Workshop on Frontiers of Handwriting Recognition*, page 309-312.
- Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolutionary Programs*. Springer, Berlin, 3<sup>rd</sup> edition.
- Sánchez-Ávila, C. and Sánchez-Reillo, R. (2002). Iris-based biometric recognition using dyadic wavelet transform. In *IEEE Aerospace and Electronic Systems Magazine*, pages 3-6.
- Schwefel, H. P. (1995). *Evolution and Optimum Seeking*. Wiley, New York.
- Yanushkevich, S. N. (2006). Synthetic biometrics: A survey. In *International Joint Conference on Neural Networks*, pages 676-683.
- Yanushkevich, S. N., Wang, P. S. P., and Gavrilova, M. L. (2007). *Image Pattern Recognition: Synthesis and Analysis in Biometrics*. Imperial College Press.