

Multimodal Algorithm for Iris Recognition with Local Topological Descriptors*

Sergio Campos¹, Rodrigo Salas², Hector Allende¹, and Carlos Castro¹

¹ Universidad Técnica Federico Santa María, Dept. de Informática, Valparaíso-Chile
scampos@inf.utfsm.cl, hallende@inf.utfsm.cl, ccastro@inf.utfsm.cl

² Universidad de Valparaíso, Departamento de Ingeniería Biomédica,
Valparaíso-Chile
rodrigo.salas@uv.cl

Abstract. This work presents a new method for feature extraction of iris images to improve the identification process. The valuable information of the iris is intrinsically located in its natural texture, and preserving and extracting the most relevant features is of paramount importance. The technique consists in several steps from acquisition up to the person identification. Our contribution consists in a multimodal algorithm where a fragmentation of the normalized iris image is performed and, afterwards, regional statistical descriptors with Self-Organizing-Maps are extracted. By means of a biometric fusion of the resulting descriptors, the features of the iris are compared and classified. The results with the iris data set obtained from the Bath University repository show an excellent accuracy reaching up to 99.867%.

Keywords: Iris recognition, SOM, Voronoi polygons, regions descriptors.

1 Introduction

In recent years, the use of biometric systems, mainly for reasons of security, has grown almost exponentially. The authentication of the person's identity in an univocal and automatic way is a requirement by nowadays standards, even more due to large-scale applications that work with hundreds or thousands of users. For these reasons, it is necessary to improve the algorithms to expand the possible scenarios where these systems can be applied.

Biometric systems are commonly employed in verification mode, i.e., they work verifying the identity of the user by fitting it with the identity previously stored by any device used for accurate identification (card ID, password, etc.). This functionality is indispensable by banks, hospitals, government institutions, airports, to name a few possible applications. Moreover, it is also necessary to consider scenarios where we have no a priori information about the subject. Under this scene, the biometric system can work in identification mode, which without any information, and only with an input, searches within the database

* This work was supported by the Fondecyt 1070220 and 1070268 research grants and DGIP-UTFSM grant.

and tries to fit that input with the most similar user. This way of working is complicated when the system has more users, since it has to deal with the problem of reducing intra-class variance and increase the inter-class variance.

In this paper we propose a new method for iris recognition, which works the iris image as a multimodal system by fragmenting the image in two sectors. With an unsupervised method we obtain regions descriptors for each fragment, where this features preserves and compresses the iris texture information. These features are used to classify the images as identification mode.

The structure of our work is the following: section 2 shows a state of art of iris recognition; in section 3 we state our proposal of a new method, starting from the pre-processing up to the classification process; section 4 shows the results of the experiments, and section 5 gives some concluding remarks about this work.

2 Related Works on Iris Recognition

John Daugman[1] proposed the first known algorithm for iris recognition. The process basically consists in the following steps. First, the pre-processing stage detects the edges of the pupil and iris to locate the position of the iris within the image. This stage is known as pre-processing. Then the process continues with the extraction of a pattern of the iris image by methods of feature extraction and finally performs the classification process.

The literature focuses mainly on creating new methods for these 3 major processes (pre-processing, feature extraction and classification). Daugman uses an integrated operator differential, two dimensions Wavelets and matching using XOR function for this processes respectively.

There are many other jobs that are based on Daugman's job, but they used different techniques, for example, in [8] they used independent component analysis (ICA) for feature extraction. In this paper the authors state that their method creates a more compact *iriscode*[1] than the more classical approaches and therefore the matching is faster. In addition, this process use only a portion of the iris, as in [5], but these use Gabor Filters.

Since 2005 approximately, various learning algorithms have been employed for this purpose, in [10] a proposal based on HSOM (a variant of SOM (*Self-Organizing Maps*) to hold the Hamming distance) is presented to calculate the adjustment of the patterns obtained in the process of feature extraction, in [9] an algorithm based in LVQ (*Learning Vector Quantization*) is presented to classify the patterns extracted previously through two dimensions Wavelets transform.

In recent years researchs oriented to biometric recognition have been focused on constructing algorithms for multimodal systems. Multimodal systems are those who work with multiple sources of information to make a decision, for example, several samples of the fingerprint, both iris of a person, different algorithms that work independently and deliver different outcomes, etc[6]. The fusion process is highly considered because the how and where the fusion takes place has a direct influence on the systems performance. Daugman notes the importance of standardizing the process of fusion and some methods to make the fusion at decision level, as AND and OR rule[4].

Generally to work with more sources of information should improve the classification rates, which is shown in some studies as [11] y [12], where the first involves iris and face and attacks the problem of score fusion as an optimization problem, where the task is to find the threshold which minimizes the total error, and the second which work with 3 traits: iris, face and palms, which performs fusion at score level, through multiple SVMs in parallel, but this only works if good classifiers are used and/or the information source is not noisy.

3 The Proposed Method

3.1 Localization

The location is primarily to recover the portion of the iris in the image. For this purpose, we proceeded to find a preliminary center of the pupil (x_p, y_p) by checking the vicinity of each pixel, which should have a low intensity and be central to the image. After establishing the point (x_p, y_p) , we proceed to find the real center (x_c, y_c) through a square that circumscribe the pupil, detected by the difference of intensity between the pupil and the iris.

After obtaining the center of the pupil, we proceed to detect the inner boundary (border) of the iris and with it, the radius of the pupil r_p .

Finally, we use the procedure described in [3], which creates a circle of radius R ($r_p < R$) concentric to the pupil that is used to clean the image. Thus, the portion of the iris considered for the recognition is the area located in between the pupil and the circle of radius R . Everything outside this circle is eliminated by resetting the pixels to 0.

3.2 Normalization

It is necessary to extract the texture information of the iris of the original image once it is fully identified, so we can work with it more easily. We use the Polar transform [5], which is based on Daugman’s work [1]. The difference lies mainly in the sweep angle. In [5], the authors took a sweep angle of 90 for the left and right section of the iris, not considering the upper and lower sections of the iris, because of the possible occlusions produced by eyelashes.

Due to the characteristics of the images in the database[13], we modify the polar transform proposed in [5], written in the following way:

$$J(x, y) = IE(x_0 + r \cos(\theta), y_0 + r \sin(\theta)) \tag{1}$$

where:

$$r = r_p + (x - 1)\Delta_r \qquad \forall x \in N : x \leq \frac{r_i - r_p}{\Delta_r}$$

$$\theta = \begin{cases} \text{ang_1} + (y - 1)\Delta_\theta & , \text{ if } y \leq \frac{\text{ang_bar}}{2\Delta_\theta} \\ \text{ang_2} + (y - 1)\Delta_\theta & , \text{ if } y > \frac{\text{ang_bar}}{2\Delta_\theta} \end{cases} \qquad \forall y \in N : y \leq \frac{\text{ang_bar}}{\Delta_\theta}$$

where (x_0, y_0) are the coordinates of the center of the pupil, r_p and r_i are the values of the radius of the pupil and the iris, respectively, Δ_r is the interval of separation between pixels of the same radius (if $\Delta_r = 1$ means that there is no separation), Δ_θ is the interval of separation angle between radius and radius, ang_1, ang_2 are starting angles to make sweeping and ang_bar is the total angle covered by the full sweep of the process. The result of this process can be seen in Figure 1.

3.3 Enhancement of the Image

The normalized iris image has low contrast and could have a non-uniform brightness due to the light of the moment when the image was obtained. This makes the iris texture to be more uniform than it really is. This is why we performed an improvement to the image by means of a histogram equalization, which makes an expansion of the histogram of the image, ie, it makes it fill the full spectrum of shades of gray, hence increasing the contrast in such a way that the texture patterns are easily noted. Figure 1 shows the process.

3.4 Fragmentation of the Image

The process of fragmentation is applied to the normalized and enhanced image. The idea is to divide this image into N smaller fragments of the same size, to reduce the computational complexity of the process of feature extraction. The figure 1 shows an example of fragmentation with $N = 2$. Note that each fragment is considered as if they were different iris but of the same subject, which allows us to perform fusion.

3.5 Construction of the Topological Graph

The iris texture is extremely rich in detail, and as mentioned above, the shape and distribution of these items is unique in each iris. To rescue the pattern obtained from the texture of the iris, we proceeded to construct a graph which preserve the topology of the iris in the graph. This graph will be the template for each iris image, which has the sufficient information to make a direct fit with the images considered in-put, ie, carry out a direct fit between the topological graph and image.

For the construction of the topological graph a neural network SOM was used. Before handing the input to the network is performed a binarization of the image using the Otsu method[7], which chooses a threshold that minimizes the variation of the intra-class white and black pixels in the image.

With the application of the SOM to the binarized image, the most relevant patterns are identified by the resulting graph. The training vectors of the network SOM are the coordinates (x_i, y_i) of the black pixels of the image. The neurons of the grid are positioned such that they describe the distribution of black pixels. The figure 2 shows the process of feature extraction.

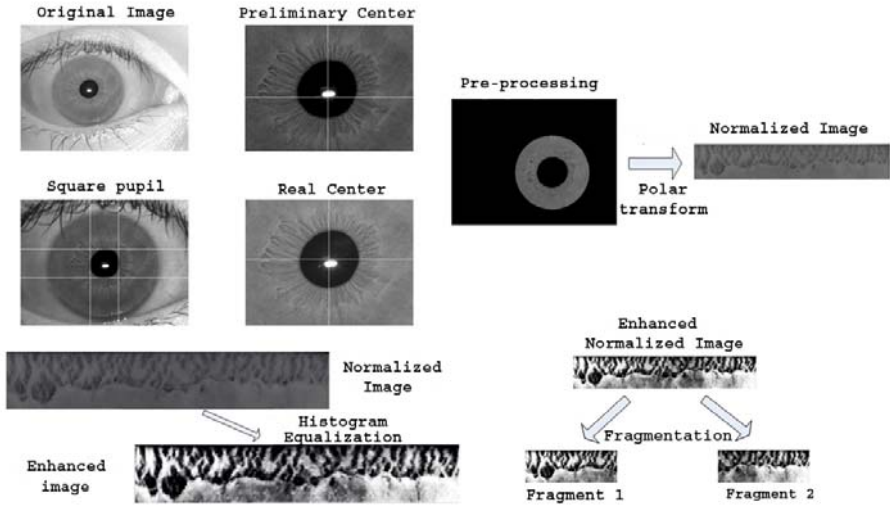


Fig. 1. Pre-processing process

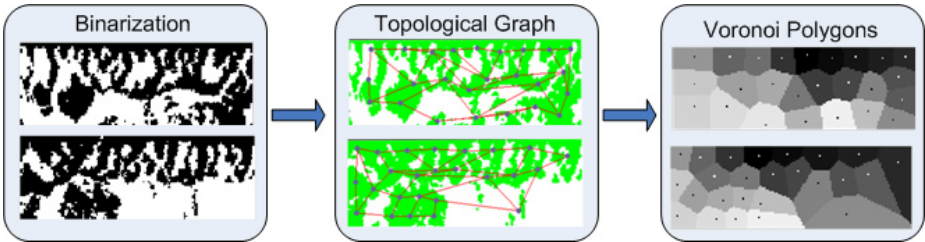


Fig. 2. Feature extraction process

3.6 Region Statistical Descriptors

Once the graph is constructed, the next step consists in obtaining the set of regional statistical descriptors. For this purpose, each node in the topological graph identify their areas of influence using Voronoi polygons.

Consider a finite number, n , of points in the Euclidean plane and assume that $2 \leq n < \infty$. The n points labeled by p_1, p_2, \dots, p_n with Cartesian coordinates $(x_{11}, x_{12}), (x_{21}, x_{22}), \dots, (x_{n1}, x_{n2})$. The n points are different in the sense that $p_i \neq p_j$ to $i \neq j, i, j \in I_n = [1, 2, \dots, n]$.

Let p be an arbitrary point in the Euclidean plane with coordinates (x_{p1}, x_{p2}) . Then the Euclidean distance between p and p_i is given by $d(p, p_i) = \|x_p - x_i\| = \sqrt{(x_{p1} - x_{i1})^2 + (x_{p2} - x_{i2})^2}$. If p_i is the point closest to p , then the relationship $\|x_p - x_i\| \leq \|x_p - x_j\|$ to $i \neq j, i, j \in I_n$ holds. In this case, p is assigned to p_i . So, mathematically we can define a Voronoi diagram as follows [2]:

Let $P = [p_1, p_2, \dots, p_n] \subset \mathbb{R}^2$, where $2 \leq n < \infty$ and $x_i \neq x_j$ to $i \neq j, i, j \in I_n$. Let the region given by: $V(p_i) = \{x \mid \|x - x_i\| \leq \|x - x_j\| \text{ for } j \neq i\}$

$i, j \in I_n$ planar Voronoi polygon associated with p_i , and the set given by: $\mathbf{V} = \{V(p_1, V(p_2), \dots, V(p_n))\}$ planar Voronoi diagram generated by P . After identifying the areas of influence of each node, we proceed to calculate statistical descriptors of the regions.

The statistical descriptors calculated were the mean, variance and skewness. The first 3 moments have a great descriptive power in terms of intensity distribution of pixels covered. In this way, for every region of influence the first 3 moments were calculated, normalized between 0 and 1 (only μ and σ) as follows: $\mu = \frac{1}{255} \frac{1}{n} \sum_{i=1}^n x_i$, $\sigma = \frac{1}{\mu \cdot 255^2} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, and $\gamma = \frac{E(x-\mu)^3}{\sigma^3}$ obtaining a vector of features as follows: $\mathbf{V} = (\mu_1, \sigma_1, \gamma_1, \dots, \mu_{25}, \sigma_{25}, \gamma_{25})$.

3.7 Classification

For the classification process we use the vectorial similarity function that measures how similar are 2 images:

$$F_{sim} = |\mathbf{V}_{in} - \mathbf{V}_t| \quad (2)$$

where \mathbf{V}_{in} is the vector of characteristics of the input image and \mathbf{V}_t is the template vector previously stored in the database.

4 Experimental Results

For our experiments, the database employed was obtained from the University of Bath repository[13]. The data set consists of 1000 images of 25 different subjects with 20 images for each eye. The images are in grayscale and an original size of 1280 by 960 pixels. The computer where the experiments were conducted was a Pentium 4 1.5 Ghz with 1 GB of RAM. The experiments were done in Identification mode, that is, given an input image, it was classified according to the closest class of the total database. This mode only generates one error rate, unlike the Verification mode that generates 2 types of error: False Accept rate (FAR) y False Reject Rate (FRR)[6].

Due to the difference in the patterns of the iris of the left and right eye of each person, we considered each eye as a different class, so the total number of classes that worked were 50. For each iris we got 2 fragments which we used to get the score level fusion.

The experiments were performed with 5-fold cross validation. The fusion was performed using a scoring rule that considers the following weights for each fragment:

$$Score_t = 0.5 \cdot score_{frag1} + 0.5 \cdot score_{frag2} \quad (3)$$

where $Score_t$ is the total score for the classification.

Table 1 shows the accuracy rate, variance and time obtained in the identification processes for with fragmentation and without fragmentation. Three different configuration of statistical descriptors were tested: the mean, the mean and the

Table 1. Summary of experimental results

Descriptors	Without Fragmentation			With Fragmentation		
	Acc. [%]	Var. [%]	Time [$\frac{sec}{match}$]	Acc. [%]	Var. [%]	Time [$\frac{sec}{match}$]
μ	99,655	0,066	0,095	99,867	0,026	0,077
μ and σ	99,517	0,104	0,113	99,203	0,102	0,083
μ , σ and γ	99,787	0,056	0,122	99,787	0,005	0,109

variance, and the mean, variance and skewness. The best outcome was obtained for the mean reaching up to 99.867%.

While it is logical to expect that including more features will describe in a better way the processed image, the addition can deteriorate the performance of the classification because, probably they do not supply any distinctive information for each class or combined with another occurrence of the classes overlap, that is, does not increase the inter-class variance.

With respect to time, it clearly wins by getting 0.077 seconds by each comparison, which apparently is not significant when compared with the third configuration's 0.109 seconds, but if we think in large scale, i.e. thousands of users, the improvement would be in the order of ten minutes.

The main goals of our proposed algorithm is twofold: first, it was to reduce the dimension of the images of the iris and in this way to reduce the complexity for the feature extraction process. The second was to have the opportunity to make a single image fusion without having the original pattern splitted, and therefore, work with only one part of the iris image.

5 Conclusions and Further Works

In this work, a new method for iris recognition were presented. This method allows us to recover the topology of the iris through the neural network SOM, technique which gives us important points within the iris image represented by the neurons. Thanks to these points, we can identify areas of influence of each neuron using Voronoi polygons, and also characterize these polygons using regions descriptors.

The techniques employed in our method describe very well the texture of the iris, which is defined as a sub-structure and each of these are quantified according to the distribution of intensities of the pixels.

The results of the experiments show that the fractionation of the image enhances the rate of accuracy, to almost 100 %. This tells us that it is not necessary to have more sources of information, in order to improve the rate of accuracy which involves more complex computational, but if each fragment separately allows a good classification, together will surely improve overall performance[4].

Our future work will focus primarily on 2 tasks to develop: be calculated more regions descriptors, such as a translations and rotations invariant moments, entropy, etc., and all these will develop a feature selection process to know which are the best descriptors in terms of it's descriptive quality and computational

complexity. As a second task, since we are interested in working in Identification mode, perform clustering of users by the topological graph, well to address environments to deal with so many users. For this work with more than one database.

References

1. Daugman, J.: How Iris Recognition Works. *IEEE Trans. on Circuits and Systems for Video Technology* 14, 21–30 (2004)
2. Okabe, A., Boots, B., Sugihara, K., Chiu, S.: *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, segunda edición. John Wiley & Sons Ltd, Chichester (2000)
3. Ganeshan, B., Theckedath, D., Young, R., Chatwin, C.: Biometric iris recognition system using a fast and robust iris localization and alignment procedure. *Optics and Lasers in Engineering* 44, 1–24 (2006)
4. Daugman, J.: Biometric Decision landscape. Technical report 482, University of Cambridge (2000)
5. Sanchez-Reillo, R., Sanchez-Avila, C.: Iris recognition with Low Template Size. In: Bigun, J., Smeraldi, F. (eds.) AVBPA 2001. LNCS, vol. 2091, pp. 324–329. Springer, Heidelberg (2001)
6. Jain, A., Ross, A., Nandakumar, K.: *Handbook of Multibiometrics*. Springer, Heidelberg (2006)
7. Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9(1), 62–66 (1979)
8. Noh, S., Bae, K., Park, K., Kim, J.: A new Iris recognition method using independent component analysis. *IEICE Trans. Inf. And Syst* E88-D (2005)
9. Cho, S., Kim, J.: Iris Recognition using LVQ Neural Network. In: Wang, J., Yi, Z., Žurada, J.M., Lu, B.-L., Yin, H. (eds.) ISNN 2006. LNCS, vol. 3972, pp. 26–33. Springer, Heidelberg (2006)
10. Neagoe, V.: New Self-Organizing Maps with Non-conventional metrics and their applications for Iris recognition and automatic translation. In: *Proceedings of the 11th WSEAS International Conference on Computers* (2007)
11. Toh, K., Kim, J., Lee, S.: Biometric scores fusion based on total error rate minimization. *Patter Recognition* 41, 1066–1082 (2008)
12. Wang, F., Han, J.: Robust Multimodal Biometric Authentication Integrating Iris, Face and Palmprint. *124X Information Technology and Control* 37(4) (2008)
13. Monro, D.: Bath University iris database (2008), <http://www.bath.ac.uk/elec-eng/research/sipg>