Combining Wavelets and 2D Gabor Descriptors for Iris Recognition in Noncooperative Environments

Sirlene Pio, Pedro Silva, Alvaro Guarda, Eduardo Luz, David Menotti
Departamento de Computação
Universidade Federal de Ouro Preto
Ouro Preto, MG, Brazil
Email: {sirlenepg,pedroh21.silva,alvaroguarda,eduluz,menottid}@gmail.com

Abstract—Nowadays, there are many open issues related to iris recognition in noncooperative environments. This work explores the fusion of Wavelet and 2D Gabor descriptors, two well successful feature descriptors used in the iris recognition literature. The main contribution of this work is how the fusion of Wavelet and 2D Gabor filter descriptors is made, resulting in a powerful iris descriptor. We conducted experiments on two widely used databases, CASIA-IrisV4, in which the iris images are acquired in a well constrained environment, and UBIRIS-v2, where the iris images were captured in unconstrained conditions, representing more realistic situations. The experiments show that our implementation of combined descriptors achieves Equal Error Rate (EER) of 1.46% for the CASIA-Iris.v4 database, which are very close to the state-of-the-art, and EER of about 15% for the UBIRIS-V2 database (decreasing the error in 25%).

Keywords—Iris recognition; Wavelet descriptor; 2D Gabor descriptor; CASIA and UBIRIS databases

I. INTRODUCTION

Among all biometrics source in study, the eye iris is taken as the most unique phenotype feature visible in a person’s face [1]. It is composed of particular and random textures for each individual, which are even different for each eye of the same person. This characteristic increases its uniqueness (possibility of 1 among 10^72 individuals of being equal), making it difficult even to fraud. In constrained environments, in short distances, and smaller databases, there are currently efficient and effective methods for correct (or almost perfect) iris identification [2], [3]. Nonetheless, the main question is how to identify it on adverse conditions in natural environments.

Thus, a major problem in using iris for person identification is in its recognition at an uncontrolled distance, moving people, and use of some accessories such as lenses and glasses, among others. Difficulties also increase when the system runs on a large dataset, with degraded images by light reflections and other noise, where a misidentification may cause great damage.

Currently there are several techniques applied to the iris recognition in order to overcome classification difficulties caused by these problems [4], [5], [6], [7]. In this work, we propose to explore the fusion of Wavelet and 2D Gabor descriptors, two well successful feature descriptors used in the iris recognition literature [8], [9], [10].

For evaluating these descriptors, we implement an iris recognition method focused on noncooperative environment that uses multiple signatures extracted from six overlapping iris sub-regions to improve the description of noisy iris images [8]. These same authors also use a particular and interesting classification scheme, employing an ordered set of six thresholds to minimize the false rejection and false acceptance rates based on the dissimilarities obtained from the sub-regions between the query and enrolled images. Our main contribution relies on the combination of the feature extraction method used in [8] (2D Gabor Representation) with another method which is based on Wavelets [9], achieving the best results in our experiments on CASIA and UBIRIS databases. The combination of descriptors for two images is made by simply adding the obtained Hamming distances of descriptors in each sub-region.

The remainder of this work is organized as follows. In Section II, some works related to iris recognition in noncooperative environments are reviewed. Section III presents the method developed here including our fusion strategy to combine Wavelet and 2D Gabor descriptors. Section IV details the experiments performed. In Section V, a brief conclusion and possible future work are drawn.

II. RELATED WORKS

In general, an iris recognition system is divided in six main steps such as: image acquisition; iris segmentation; normalization; feature extraction; representation of features (in a binary vector, for instance); and classification. However, further steps, such as pre-segmentation, pos-segmentation and pos-recognition, can be employed to improve the effectiveness of the system. In this study we are concerned with the feature extraction and representation, and classification steps. And more specifically, with the situations that deal with iris images acquired in noncooperative environments. In the seminal model proposed by Daugman in 1993 [1], the 2D Gabor filter is used to extract features from the original image and generate a binary vector, known as IrisCode. The classification is by means of the Hamming distance to calculate the dissimilarities between the query and the enrolled images. Similar processes are found in [9], [4]. Although this model [1] is still seen as a reference for the development of systems/methods of iris recognition, it is not robust when images are acquired in noncooperative environments. The problem of representing the information present in degraded iris images is often discussed and investigated in works related to iris recognition in noncooperative environments.

In the literature, there are works that extract features from
different regions of the eye [5], [11], [12], [13]. In [5], zigzag collarette is used to extract features. Other iris aspects, such as ordinal measures and color analysis, and eye features, texton and semantic information for instance, are present in [11]. The periocular region also provides promising descriptors, allowing combination of them with other iris features [12], [13].

Proença & Alexandre [8] divide the iris into two independent blocks, one composed of four sub-regions and the other composed of two sub-regions. Although the sub-regions of the same block are independent (non-overlapping), there is overlap among the sub-regions of distinct blocks. The rationale for this division is the robustness to noisy environments and loss of biometric signature. From these sub-regions, six dissimilarity values are obtained and then fused by means of a classification rule, which is also employed in our work and involves an ordered set of optimized thresholds obtained to minimize the false rejection (false negative) and false acceptance (false positive) rates. Similar strategies are employed in [14], [15]. In the work of Li et al. [6], the iris region is divided into small fragments without overlapping. In each fragment, it is used the strategy of Weighted Co-occurrence Phase Histogram (WCPH) to represent texture patterns.

Szewczyk et al. [9] initially segmented the iris image and then pre-processed, using the following steps: blue channel removal; conversion to monochrome images; histogram equalization; and removal of reflections, eyelashes, and occlusions caused by eyelid. They also analyze and choose the best Wavelet function for iris feature extraction.

Other works in the literature coping with iris images in noncooperative environment are focused on performance/effectiveness comparison of different strategies. In [16], the features extracted from the LoG-Gabor filters, Haar wavelet, Discrete Cosine Transform (DCT), and Fast Fourier Transform (FFT) are compared. In [9], [7], the use of Discrete Wavelet Transform (DWT) and an ideal subset of Gabor filters, respectively, are evaluated. Marsico et al. [17] combine two techniques, Linear Binary Patterns (LBP), which produces a local texture description, and Discriminable Textons (BLOBs), which highlights uniqueness of the texture (furrows, crypts and spots), and verify that the resulting method increases the final recognition performance.

In [5], [6], the focus is on the classification step. In [5], the Support Vector Machines (SVM) classifier is used with the Haar Wavelet features. If the classifier does not indicate a class with high confidence degree, the Laplacian of Gaussian (LoG) filter and 2D Gabor filters are used to extract descriptors from the images and then the Hamming distance is applied to generate a new classifier. In [6], the matching process used is based on Bhattacharyya Distance and Image Registration using Simple Image Patch Registration Method to find the distance between two images.

Despite the large number of techniques presented in the literature for iris recognition in noncooperative environments, here we propose the fusion of two well successful feature descriptors [8], [9]. In the following, besides describing the methodology used, we describe in details the Wavelet and 2D Gabor descriptors and how we combine them.

III. Method

In [8], Proença & Alexandre improved the recognition method described by Daugman [1] using a new iris partition strategy, allowing the extraction of more robust features for iris recognition in noncooperative environments. Figure 1 presents the development stages of this method.

That method introduces the segmentation/partition of the iris region into six sub-regions, allowing independent feature or descriptor extraction from each sub-region. Furthermore, by the comparison of extracted descriptors of the corresponding iris areas, six dissimilarity values are obtained that are fused by means of a classification rule. According to [8], the feature extraction process and independent comparison of each of these sub-regions prevents the biometric signature to be completely compromised due to some noise located in any of the sectors of the iris.

In the present work the same preliminary steps of the method proposed in [8] are used and then the reader is referred to original work. To improve the representation of the information contained in the iris, and consequently to obtain a higher effectiveness, Wavelets descriptors are also extracted. The combination of descriptors given two images is made by simply adding the obtained Hamming distances of descriptors in each sub-region. This is the main contribution to the literature of our work and brings promising results.

A. Feature extraction

In this work, we use the 2D Gabor representation [8] and Wavelets [9] as descriptors. We emphasize that the process of feature extraction, whatever the technique used (2D Gabor representation or Wavelet), consists in the creation of six independent biometric signatures, each one corresponding to a specific iris region. This feature partition is specially developed to deal with iris images acquired in noncooperative environments.

1) 2D Gabor representation: The feature extraction process using 2D Gabor filter consists in decomposing the input signal, a normalized image, using two Gabor filters in quadrature: real and imaginary parts, specified by cosine and sine functions modulated by a Gaussian. Basically, it is performed the projection of Gabor filter in fixed size blocks to obtain a set of centered filters in which they are used as parameters varying with the frequency inverse [18].

2) Wavelet: Wavelets are used for decomposition of (image) signals using special filters. The Wavelet transform can be viewed as a feature extraction technique based on windows with fixed area, but its width and length vary. There are several families of mother Wavelets, each having several variations, such as: Haar, Daubechies, Symlets, Coiflets, Discrete Meyer, Biorhogonal (bior1.3, bior2.2, bior6.8) and the Reverse Biorhogonal (rbio1.3, rbio2.2, rbio3.1, rbio3.3, rbio6.8).

Among these functions, given the best results achieved in [9], we choose to study here the rbio3.1 (Wavelet Reverse Biorhogonal 3.1) and the rbio2.2 (Wavelet Reverse Biorhogonal 2.2) mother Wavelets. The image signal can be represented by j-th scale and all k displacement components. This representation involves the detail coefficients
and scalar/approximations coefficients. For 2D signals (images), the details coefficients are composed of three types: horizontal, vertical, and diagonal coefficients. The signal can be sequentially decomposed, given as input the approximation coefficients for the next level, using many levels as necessary.

B. Encoding

The encoding step generates a binary vector, known as IrisCode, and the encoding strategies to the 2D Gabor representation and Wavelet are described below.

1) 2D Gabor representation: For each complex output value after the 2D Gabor filter on the sub-region of the normalized image, the phase value $\phi$ is decoded in one of four possible quadrants. That is, given a phase value, one verifies in what quadrant the point is located and then replaced it by two bits using the following representation \{00, 01, 10, 11\}. Thus, after feature extraction and encoding, we obtain six IrisCode one for each subregion of the iris.

2) Wavelet: For the Wavelet descriptor, the IrisCode is obtained by analyzing the vertical detail coefficients [9], i.e.,

$$\text{IrisCode}(x, y) = \begin{cases} 0, & \text{if } V_i^l < \text{median}(V^l) \\ 1, & \text{if } V_i^l \geq \text{median}(V^l) \end{cases}$$

in which $V_i^l$ is the $i$-th coefficient (referenced at $(x, y)$ position) at the $l$-th vertical level, and median$(V^l)$ stands for the median among all $V^l$ coefficients.

C. Classification

Initially, we calculated dissimilarity between two images in all six sub-regions $D_i = HD(I_i^1, I_i^2)$, in which $i = 1, \ldots, N$ ($N = 6$ is the number of sub-regions), $I_i^1$ the sub-region $i$ of the image 1, $I_i^2$ the sub-region $i$ of the image 2.

The Hamming Distance (HD) is used to calculate the dissimilarity between two images. Given two binary sets $A = \{a_1, \ldots, a_N\}$ and $B = \{b_1, \ldots, b_N\}$ the Hamming Distance is given by

$$HD(A, B) = \frac{1}{N} \sum_{i=1}^{N} a_i \otimes b_i$$

in which $a_i \otimes b_i$ is the logical XOR operation.

Besides separately evaluating the 2D Gabor filter and Wavelets descriptors, we propose their combination by summing up the Hamming distance obtained for each sub-region descriptors, i.e.,

$$D_i = HD(\text{IrisCode}W_i^1, \text{IrisCode}W_i^2) + HD(\text{IrisCode}G_i^1, \text{IrisCode}G_i^2)$$

in which $\text{IrisCode}W_i^1$ and $\text{IrisCode}W_i^2$ stand for the encoded Wavelet descriptors for images 1 and 2, respectively, and $\text{IrisCode}G_i^1$ and $\text{IrisCode}G_i^2$ for the encoded 2D Gabor descriptors for the same images 1 and 2, respectively, with $i = \{1, \ldots, 6\}$.

Given the sets of dissimilarity $D_i = [D_1, \ldots, D_N]$ and the thresholds $T_i = [T_1, \ldots, T_N]$, explained later, the next step is to count the number of $D_j \in D$ that are less or equal to $T_i$:

$$C(D, T_i) = \sum_{j=1}^{N} \Pi_{D_j \leq T_i}$$

where, $\Pi(\cdot)$ is the indicator function.

The images $I^1$ and $I^2$ are classified as corresponding to the same iris if:

$$\exists_i : C(D, T_i) \geq i, i = 1, \ldots, N.$$  (5)

The False Acceptance Rate (FAR) is the probability of error occurrence of the type I. The Error of type I, also known as $\alpha$-error or false positive, happens when it is accepted as genuine sample but it is false. The False Rejection Rate (FRR) is the probability of error type II. Type error II, also known as $\beta$-error or false negative, happens when we reject something that should be accepted.

FAR and FRR can be formally computed as

$$\text{FAR} = \sum_{i=1}^{6} P \left( C(D^E, T_i) \geq i, \bigcap_{j=1}^{i-1} C(D^E, T_j < j) \right),$$  (6)

$$\text{FRR} = \prod_{i=1}^{6} P \left( C(D^I, T_i) < i \right).$$  (7)

Figure 1. Development stages of the method proposed in [8]
in which $D^E$ and $D^I$ represent the interclass and intraclass dissimilarities sets, respectively.

Using Eq. 6 and Eq. 7, we conducted an exhaustive search in the interval $[0, 1]$ for some incrementing values (in our experiments, we used $[0.1; 0.05; 0.025; 0.01]$), in order to find the (sub)optimal set of thresholds $T_i = [T_1, ..., T_N]$, that minimizes the equal error rate (EER). Let $T_i = [T_1, ..., T_N]$, $T_i \in \mathbb{R}^N$, be a set of $N$ threshold values such that $T_i \leq T_j$, $\forall i < j$.

IV. Experiments

In this section, we describe the experiments performed in order to test the proposed method. Initially, we describe details of the iris images and their databases and then the results of the experiments performed are presented. Finally, a brief discussion of the results is made.

A. Databases

To the best of our knowledge, the CASIA-IrisV4 database [19] is the most used for evaluation of iris recognition systems. CASIA-Iris.v4 contains a total of 54,601 iris images from more than 1,800 natural individuals and 1,000 virtual individuals. However, the iris images of this database were acquired in a very constrained and controlled environment. Figure 2(a) shows an image example of CASIA database.

In the database UBIRIS.v2 [20], the images were captured in unconstrained conditions (at different distances, in motion and in the visible wavelength) with the corresponding more realistic noise factors. UBIRIS.v2 [20] contains 11,102 iris images from more than 261 individuals. Figure 2(b) shows an image example of UBIRIS.v2 database.

In our experiments, we use 800 images of 80 subjects (10 images per individual) of UBIRIS database and 800 images of CASIA database. Note that, for both databases, we randomly select the individuals as done in previous works [8], [9], [15], which makes difficult a fair comparison between works published in different papers, even when the same amount of samples and individuals are used. However, in order to surmount such issues, we propose to let available our selection of individuals and also all the source code used in a webpage once the paper is accepted to publication.

From these 800 images of each database, 400 images of 40 individuals were used in the training phase for estimating the classifying thresholds, leaving 400 images of the other 40 individuals for testing/evaluation.

B. Results

In the followings, we report the results of our experiments. Initially, we evaluate the results obtained by using the Wavelet descriptors and then the 2D Gabor representations ones, by varying some parameters of the descriptors and the increments used to obtain the classification thresholds. Finally, we took the parameters that obtained the best results for each descriptors and evaluated the combination of them.

It is important to note that we follows the same evaluation protocol used in previous works to report the results here [8], [9], [15]. That is, we estimate the classification thresholds in the training set, and in the evaluation/testing set all images of each individuals are taken as query image in a scheme similar to leave-one-out.

As several other works in the literature, we use the equal error rate (EER) as our effectiveness measure of analysis. The EER is defined as the point in which $FAR = FRR$, and its value is exactly the same of FAR and FRR.

1) Wavelet: As in [9], we use the vertical component (its coefficients) of the Wavelet transform [21], [22] to represent the information contained in the iris. In order to understand how the effectiveness of the iris recognition system would be affected by the parameters of the Wavelet transform, we performed experiments by varying the level of decomposition and the normalized image size and also as indicated in [9], we choose rbio3.1 and rbio2.2 as Wavelet bases. Instead of evaluating both the level of decomposition and the normalized image dimension, we choose to first analyze the level of decomposition to be used, since the evaluating process of two parameters simultaneously will require more experiments and space.

Initially, we perform experiments in the CASIA database, the one in which the images Table I presents the results obtained in the CASIA database using the rbio3.1 and rbio2.2 Wavelet bases by varying the level of decomposition, from 1 to 5 ($v_1$, $v_2$, $v_3$, $v_4$, and $v_5$), used to represent the iris image. For this experiment we keep the normalized input image with a fixed sized of $16 \times 128$ pixels, as an initial guess. From the EER values, we can observe that the results obtained when an incrementing of 0.1 is used to compute the classification thresholds, the EER values is almost 50%, while quite better results are obtained when steps of 0.05 are used to compute the classification thresholds. Note that the classification thresholds step is $O(\text{range}^6)$, in which $\text{range}$ is the range of values used. For instance, if increment is 0.1, then we have $\text{range} = \{0.0, 0.1, 0.2, 0.3, ..., 0.9, 1.0\}$. In other words, the lower the value of increment used, the higher the number of computations to obtain the classification thresholds is ($O(\text{range}^6)$). The best EER value among all presented in Table I is obtained using the vertical component of the second level of decomposition ($v2$) with the rbio2.2 Wavelet basis.

Having the level of decomposition parameter chosen, we vary the number of rows and columns of the normalized image, and the results are presented in Table II. Observing these figures, we note that higher numbers of columns (# cols) to the normalized image do not produce better EER. However, a slightly better EER was achieved (9.31) when the normalized image was setup to $16 \times 256$ pixels, when using an incrementing of 0.05 in the classification threshold step. Using

![Figure 2. Sample image: (a) CASIA and (b) UBIRIS](image.png)
this same configuration but now using increment values smaller
\([0.025; 0.01]\) and we obtained, respectively, the EERs 3.84% and 3.47%. Experiments with 0.025 and 0.01 increments are
performed only here because they are very time consuming,
taking 20 minutes and 20 hours, respectively, in an Intel core
i3 Notebook with 4 Gb RAM. These results when varying the
increment step for the classification threshold also show how
sensitive is the iris recognition system to this parameter.

Then we perform the same sequence of experiments in
the UBIRIS database, the one with iris images obtained in
noncooperative environments. Table III presents the results
obtained in the UBIRIS using respectively the Wavelet rbio3.1
and Wavelet rbio2.2 to the decomposition levels \(\{1, 2, 3, 4, 5\}\)
and normalized image size of \(16 \times 128\) pixels.

We selected the parameters that have achieved the lowest
EER, 22.3% in Table III (rbio3.1 Wavelet basis with
two decomposition levels) and then we vary the number of
rows and columns of normalized image. These results are
presented in Table IV. The smallest EER, using Wavelets
descriptors in the UBIRIS dataset, obtained was 22.3% to
rbio3.1 Wavelets basis with two decomposition levels and
the normalized image with dimension of \(128 \times 16\) pixels.
To these parameters, we performed experiments with smaller
increments, i.e., \([0, 0.025; 0.01]\), and obtained, respectively, EER
values of 20.76% and 20.63%.

2) 2D Gabor Representation: For the 2D Gabor Filter
descriptors, we do not have parameters to evaluate. Then, we
perform the experiments by simply varying the increment step
of the classification threshold phase in the CASIA and UBIRIS
databases.

Table V and VI show the results obtained using the 2D Gabor Filter as feature representation of iris. The smaller
EERs obtained in the CASIS and UBIRIS databases are
2.39 and 19.58, respectively, when using increment value
of 0.01. Smaller increments, such as 0.005 or even 0.001,
were not evaluated since our implementation would require
about a month to obtain the results using the computational
architecture available. However, note that, in [8], we can find
results reported using steps of 0.001.

3) Descriptor Combination: Once experiments were per-
formed by analyzing the effectiveness of the iris recognition
method using the descriptors separately, here we report results
by combining the descriptors. Tables VII and VIII show these
results obtained in the CASIA and UBIRIS respectively. As
naturally expected, in both, the CASIA and UBIRIS databases,
the lower the increment values used in the classification
threshold step, the smaller the EER is. Moreover, we can see
that the EER obtained with the combination of the descriptors
outperformed the EER results obtained previously when they
were used separately.

C. Discussion
In both databases, the best results is obtained for the feature
extraction method that combines the 2D Gabor Representa-
tion and Wavelet descriptors. In Figure 3, we can see that
the ROC (Receiver Operating Characteristic) curve generated when the 2D Gabor Representation + Wavelet is applied to the recognition (red curve) is closer to the origin than the ones generated when the recognition method uses 2D Gabor Representation or Wavelet descriptors separately (blue and green curves, respectively).

When comparing the best results achieved by 2D Gabor representation and Wavelet descriptors separately with the combined one, we observe that a relative improvement of about 25% is achieved, i.e., the 2D Gabor representation achieved an EER of about 20% while an EER of about 15% is obtained when the combined descriptor is applied.

V. CONCLUSION AND FUTURE WORK

In this work, we studied, implemented, and evaluated the extension of a iris recognition method for noncooperative environments using 2D Gabor multiply descriptors extracted from two blocks of iris sub-regions [8] by combining Wavelets descriptors [9] through summation of Hamming distances. We performed experiments of two databases (UBIRIS and CASIA), and in both were obtained a lower EERs when we combined the dissimilarities found when a 2D Gabor Representation and Wavelets were used to feature extraction. The combination of different feature descriptors showed able to exploit the strengths of each of them in images with different characteristics and improve the results obtained when the feature descriptors are applied individually.

As future work we plan to optimize, by using Evolutionary Algorithms, for instance, the stage of selection of classification thresholds making possible evaluations with smaller increment values.

AGRADECIMENTOS

The authors thank Pró-reitoria de Graduação (PROGRAD)/UFOP, CAPES, CNPq (#307010/2014-7), and FAPEMIG for the financial support to this work.

REFERENCES