

# Iris Recognition: What's Beyond Bit Fragility?

Hugo Proença, *Senior Member, IEEE*

**Abstract**—The concept of *fragility* of some bits in the iris codes regards exclusively their within-class variation, i.e., the probability that they take different values in templates computed from different images of the same iris. This paper extends that concept, by noticing that a similar phenomenon occurs for the between-classes comparisons, i.e., some bits have higher probability than others of assuming a predominant value, which was observed for near-infrared and (in a more evident way) for visible wavelength data. Accordingly, we propose a new measure (bit *discriminability*) that takes into account both the within-class and between-classes variabilities, and has roots in the Fisher discriminant. Based on the bit discriminability, we compare the usefulness of the different regions of the iris for biometric recognition, with respect to multi-spectral data and to different filters parameterizations. Finally, we measure the amount of information lost in codes quantization, which gives insight to further research on iris matching strategies that consider both phase and magnitude. Albeit augmenting the computational burden of recognition, such kind of strategies will consistently improve performance, particularly in poor-quality data.

**Index Terms**—Iris Recognition, Biometrics, Gabor Filtering, Multi-Lobe Differential Filtering.

## I. INTRODUCTION

The iris is undoubtedly among the most popular traits in the biometrics literature. This is a topic that dates back to 1893, when Bertillon suggested to use the iris in anthropometric identification. Then, in 1951, it was confirmed that “*the texture of the iris is so distinctive among different individuals that it could be used as mean of identification*” [20], and, in 1993, Daugman [5] proposed the pioneer automated iris recognition algorithm, that was subsequently enhanced [7]. Currently, the iris trait is the subject of intensive research efforts, being some of the most relevant summarised by Bowyer *et al.* [2].

Most of the previous works that studied the effectiveness of the iris as a biometric trait concentrated in the levels of false rejections and in the concept of bit *fragility*, observing uneven levels of within-class variation among bits, i.e., the probabilities that bits “*end up a 0 for some images of the iris and a 1 for other images of the same iris*” [13] are uneven, as firstly formalised by Bolle *et al.* [1]. This paper explores beyond the concept of fragility, by jointly considering the within-class and between-classes variabilities. The insight is that not only the probabilities of bits flipping among samples of one iris are uneven, but a similar phenomenon occurs for samples of different irises. i.e., some filters configurations used in particular regions of the irises augment the probability that bits

predominantly take a particular value. We propose the concept of bit *discriminability*, which shares the roots of the Fisher discriminant and has an intuitive (visual) representation. A discriminant bit should: 1) keep a constant value among samples of one iris; and 2) have a value that is independent of the values in codes from different irises. According to this formulation, the *discriminability* can be regarded as an extension to *fragility*: a bit is fragile due to the high probability of flipping its value in genuine comparisons. To classify a bit as discriminant, we regard not only its fragility but also the probability of obtaining agreeing values in impostor comparisons.

Based on this concept, we infer the suitability of each region of the iris for biometric recognition. Three additional novelties are given: 1) results are shown not only for the classical Gabor-based texture description, but also for Multi-Lobe Differential Filters (MLDF) [32], which were reported as a relevant advance to the field; 2) we consider different levels of image quality, corresponding to a broad range of data acquisition protocols; and 3) we compare results for multi-spectral data (near-infrared (NIR) and visible wavelength (VW)), enabling to perceive the potential of each spectrum for biometric recognition.

Our results are based in four well known data sets: 1) the University of Bath, representing good quality NIR data; 2) the CASIA-Iris-Distance, representing NIR data of moderate quality; the 3) UBIRIS.v2 and 4) FRGC datasets, both representing VW data acquired in uncontrolled setups. Based on disjoint learning / test datasets, more than 500,000 features were extracted from the normalized iris images, corresponding to every filter parameterization centered at a different position in the iris. Then, four feature selection algorithms obtained the best combinations of bits to be included in iris codes, from where the most suitable iris regions / filters parameterizations were inferred.

Finally, the concept of bit discriminability supports a change to the classical iris recognition processing chain: in the code quantization phase, if the sign-based function is replaced by a sigmoid, not only the phase of coefficients is considered, but also their weighted magnitude. This enables to match irises with additional amounts of information, which contributes for solid improvements in performance, albeit augmenting the computational burden of recognition.

The remainder of this paper is organized as follows: Section II summarizes the related work. Section III formalizes the concept of bit discriminability. Section IV describes the evaluation protocol and setup. Sections V and VI give the results respectively at the bits and classifiers levels. Finally, Section VII concludes the paper.

## II. RELATED WORK

### A. Iris Recognition

A long road has been travelled since the pioneering iris recognition algorithm, due to J. Daugman [5]. The iris is currently among the most popular traits in biometrics research and the recently published algorithms can be classified with respect to the images spectrum they aim to deal with: NIR or VW.

Regarding the first family, strides were given to improve the recognition performance against *hard* subjects, acquisition artifacts, and inter-sensor operability. Dong *et al.* [8] proposed an adaptive personalized matching scheme that highlights the discriminating features of each class (iris) and augments the robustness against several data degradation factors. Pillai *et al.* [24] used the sparse representation for classification algorithm in randomly projected iris patches, which was observed to improve the robustness against segmentation errors and acquisition artifacts. The possibility to perform recognition in data acquired from multiple kinds of sensors motivated the algorithm proposed by Pillai *et al.* [25], that learns transformations between data acquired by different sensors and avoid that users are re-enrolled each time a new sensor is deployed.

The VW family of algorithms is mostly focused in the recognition robustness against several degraded data, acquired in as much unconstrained as possible acquisition setups. The algorithm due to Tan *et al.* [35] got the best performance in the NICE [27] contest and is considered the state-of-the-art for visible light data. More recently, Kumar *et al.* [19] used the local Radon transform to feed the Sparse Representation for Classification (SRC) algorithm, and stressed the good performance that this kind of method attains in iris data. Kumar and Chan [18] approached the problem from the data representation perspective, exploring the use of hyper-complex numbers to store the orientation of texture elements in the iris, later classified by the SRC algorithm. Also, Tan and Kumar [34] weighted the bits in iris codes in a specific way, according to the predominating noisy regions, which was observed to increase the recognition robustness against hard environments.

### B. Study of Iris Codes

Several works studied the nature of the iris texture and the properties of the resulting codes. As relevant examples, Kong *et al.* [15] provided a deep understanding of the geometric structures of the codes, regarded as a clustering algorithm. These authors showed the relation between the Hamming distance used in matching and the bitwise phase distance, arguing that Gabor kernels are actually phase-steerable filters. Subsequently [16] [17], Kong focused on the geometrical relationships of bits in iris codes, regarded as convex polyhedral cones. The relationships detected imply that a property (central ray) is enough to reveal patterns among codes, which might be used to break systems without a liveness and quality checker.

The recognition performance respect to covariates were also previously studied: Bowyer *et al.* [3] tested three of these factors: 1) the effect of pupillary dilation in performance; 2) the iris stability over lifetime; and 3) the effect of contact

lenses. They concluded that these factors bias the genuine distribution toward the impostors', but also confirmed that the probability for false acceptances is practically invariant to these factors. More recently, Mehrotra *et al.* [21] claimed that the movement of the genuine match scores toward the imposters distribution was due to other covariates (such as blur, occlusions and pupillary dilation), perhaps even at a higher degree than the ageing effect.

Concerns about the fragility of some bits in the iris codes date back to the earliest implementations of the acknowledged Daugman's recognition algorithm, by disregarding the bits with responses near the axes. Then, Bolle *et al.* [1] formalized the term of *fragile* bit and observed that, due to imaging noise, not all bits have equal possibilities to flip among samples of one iris. Hollingsworth *et al.* [12] analyzed the fragility in iris codes and a similar idea had been reported in [26]. Subsequently, Hollingsworth *et al.* [13] found that the middle bands of the iris are better than the inner parts and that large filters provide more consistent bits than small filters, due to the attenuated effect of acquisition artefacts. Finally, the same authors used the notion of bit *fragility* to propose [14] a new matching distance based in the linear combination between the proportion of disagreeing bits and the *fragile bit distance* (FBD), that expresses the fraction of unoccluded bits masked for fragility in the comparison. They observed that the FBD carries complementary information to the traditional distance and that results obtained by fusing both measures are better than when using any of these alone. The discriminability of the bit coefficients due to the coarse quantization of the phase response was also studied [13], being suggested to ignore bits with amplitude in the lower quartile.

In terms of the selection of the most reliable bits of iris codes, Dozier *et al.* [9] based their work in the concept of bit fragility. When compared to the classical code of 2,048 elements, they were able to reduce the number of bits by 30%, without significantly increasing the error rates. They even reduced the number of bits in 90%, but in this case observed a significant increase in the error rates. Rathgeb *et al.* [30] obtained the bit-error occurrences and a corresponding global-rank of bit positions. Based on this information, the less reliable bits were discarded, which improved performance and simultaneously reduced the size of codes.

## III. BIT DISCRIMINABILITY

For comprehensibility, we adopt a notation similar to the used by Bolle *et al.* [1]. Let  $\mathcal{I}^{(p)}$  and  $\mathcal{I}^{(q)}$  be two real world irises, from where the binary iris codes  $\mathcal{C}^{(p)} = F(\mathcal{I}^{(p)})$  and  $\mathcal{C}^{(q)} = F(\mathcal{I}^{(q)})$  are extracted ( $F$  is a composition of an imager and a feature encoding system). There are two hypotheses:

- $H_0$  :  $\mathcal{C}^{(p)}$  and  $\mathcal{C}^{(q)}$  are from the same iris ( $p = q$ );
- $H_a$  :  $\mathcal{C}^{(p)}$  and  $\mathcal{C}^{(q)}$  are from different irises ( $p \neq q$ ).

Let  $C_k^{(p)}$  denote the  $k^{th}$  bit (out of  $t$ ) of an iris code, i.e.,  $\mathcal{C}^{(p)} = \{C_1^{(p)}, \dots, C_t^{(p)}\}$ . We are interested in defining a metric for each bit discriminability, in terms of its (a priori) effectiveness for biometric recognition.

**Definition III.1.** Let  $\oplus$  denote the exclusive-or logical operation. The  $k^{th}$  bit is considered *discriminant*

for biometric recognition if two conditions are met: 1)  $1 - P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0) < \delta$ ; and 2)  $|\frac{1}{2} - P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)| < \delta$ , for a small  $\delta \in \mathbb{R}^+$ .

The  $k^{\text{th}}$  bit *contributes* for a *Type I* classification error (false match) with probability  $P(H_a) P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ . Similarly, it contributes for a *Type II* classification error (false non-match) with probability  $P(H_0) P(C_k^{(p)} \oplus C_k^{(q)} = 1 | H_0)$ , i.e.,  $P(H_0) (1 - P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0))$ . Hence, the probability that the bit contributes for a classification error  $\epsilon$  is given by:

$$\epsilon(k) = P(H_0)(1 - P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0)) + P(H_a)P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a), \quad (1)$$

where  $P(H_0)$  and  $P(H_a)$  are the prior probabilities for genuine and impostors comparisons.

**Definition III.2.** Let  $x_k = P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0)$  and  $y_k = P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ . We define  $(x_k, y_k)$  as the *visual representation* of the bit discriminability.

For the discussion below, two (readily satisfied) assumptions are made:

- we assume that  $P(H_0) = P(H_a) = 0.5$ , i.e., for the purpose of our analysis we make no assumptions about the prior probabilities of genuine and impostor comparisons. However, it should be noted that in most practical scenarios  $P(H_a) \gg P(H_0)$ , i.e., for databases with a single template per eye, an identification process will require much more impostor than genuine comparisons (at most one).
- we assume that  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0) \geq P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ . In any practical biometric system, there is no reason for observing agreeing bits more likely in impostors than in genuine comparisons.

**Definition III.3.** A quantitative *measure* of the bit discriminability is given by the Euclidean distance between  $(x_k, y_k)$  and the straight line  $y = x$ ,  $\tau(k) = |y_k - x_k|$ .

**Theorem III.1.** Let  $(x_i, y_i)$  and  $(x_j, y_j)$  be the visual representations of the discriminability of the  $i^{\text{th}}$  and  $j^{\text{th}}$  bits of iris codes, such that  $\tau(i) > \tau(j)$ , and  $x_i \geq y_i \geq 0.5$ . Then,  $\epsilon(i) < \epsilon(j)$  and the  $i^{\text{th}}$  bit is less likely to contribute for a classification error than the  $j^{\text{th}}$  bit.

*Proof:* By hypothesis,  $\tau(i) > \tau(j)$ , i.e.,  $|y_i - x_i| > |y_j - x_j|$ . Then,  $y_i - x_i > y_j - x_j \vee x_i - y_i < x_j - y_j$ . As  $x_i \geq y_i$ , it follows that  $x_i - y_i \geq 0$ . Then,  $x_i - y_i > x_j - y_j$ . Multiplying both sides by -1, we know that  $y_i - x_i \leq y_j - x_j$ . Adding a constant in both sides and dividing everything by another constant, we have  $\frac{(1-x_i)+y_i}{2} \leq \frac{(1-x_j)+y_j}{2}$ , i.e.,  $\epsilon(i) < \epsilon(j)$ . ■

Fig. 1 illustrates the concept of bit discriminability: the central plot is the 2D histogram of the  $(x, y)$  visual representations for bits extracted from the University of Bath dataset. The left plot schematizes this histogram, and marks the non-discriminability line ( $y = x$ ). The "A" symbol denotes an *optimal* feature and "C" denotes features that keep the

TABLE I  
RESULTS OF THE KOLMOGOROV-SMIRNOV [11] NORMALITY TEST, USING THE *null* HYPOTHESIS THAT  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$  FOLLOW A BINOMIAL DISTRIBUTION  $\mathcal{B}(n, 0.5)$ , AT THE 5% CONFIDENCE LEVEL. RESULTS ARE GIVEN FOR ALL THE DATA SETS CONSIDERED IN THIS PAPER.

Dataset	n	null hypothesis	P-value
BATH	200,000	✗	$6.80e^{-4}$
CASIA	200,000	✗	$1.93e^{-6}$
UBIRIS.v2	200,000	✗	$6.71e^{-11}$
FRGC	200,000	✗	$5.27e^{-10}$

same value both for genuine and impostor comparisons. At the other extreme, the "R" region corresponds to features that behave randomly (like dynamic noise). Features close to "R" predominantly contribute for Type-II classification errors (Region FR), and features in the upper-right corner contribute more for Type-I classification errors (Region FA). The fact that most bits fall in the region FR is the root for the extraordinary small probability of false acceptances in current recognition systems, but also justifies their relatively high false rejection rates.

The bar plots at the right side of Fig. 1 are particularly important for the context of this work: the upper plot (vertical projection) gives evidence of the levels of bit *fragility*. Complementary, the bottom plot (horizontal projection) shows the varying levels for  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ , and supports the concept of bit *discriminability*.

However, the concept of discriminability depends of whether the relative frequency of  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$  is not simply function of a random effect. In that case, values should follow a Binomial distribution  $\mathcal{B}(n, p)$ , where  $n$  is the number of pairwise bit comparisons, each of which yields success with probability  $p$  and probability mass function given by:

$$P\left((C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a) = \frac{s}{n}\right) = \binom{n}{s} p^s (1-p)^{n-s}, \quad (2)$$

where  $s$  is the number of successes (bits agreement). Using a large set of  $n$  pairwise bit comparisons and assuming that  $p = 0.5$ , the Normal distribution  $\mathcal{N}(np, \sqrt{np(1-p)})$  may be used to closely approximate results from the afore binomial distribution. According to the Kolmogorov-Smirnov [11] normality test, the *null* hypothesis stating that values follow  $\mathcal{B}(n, p)$  was rejected with asymptotic P-values lower than  $1e^{-7}$  for all the datasets used in this paper:

According to the results given in Table I, it can be concluded that  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$  varies consistently with respect to some other factor apart randomness. Also, this phenomenon is more notorious for VW than for NIR data (substantially lower P-values for VW than for NIR data), which we believe to have roots in the corneal reflections determined by the ambient VW wavelengths that are not blocked in the camera. In this case, the images of the different subjects tend to display brighter intensities in similar positions of the iris, which has some influence in the  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$  value.

#### IV. DATASETS AND PREPROCESSING

Four freely available data sets were used in the experiments, each one representing a data acquisition scenario. Fig. 2

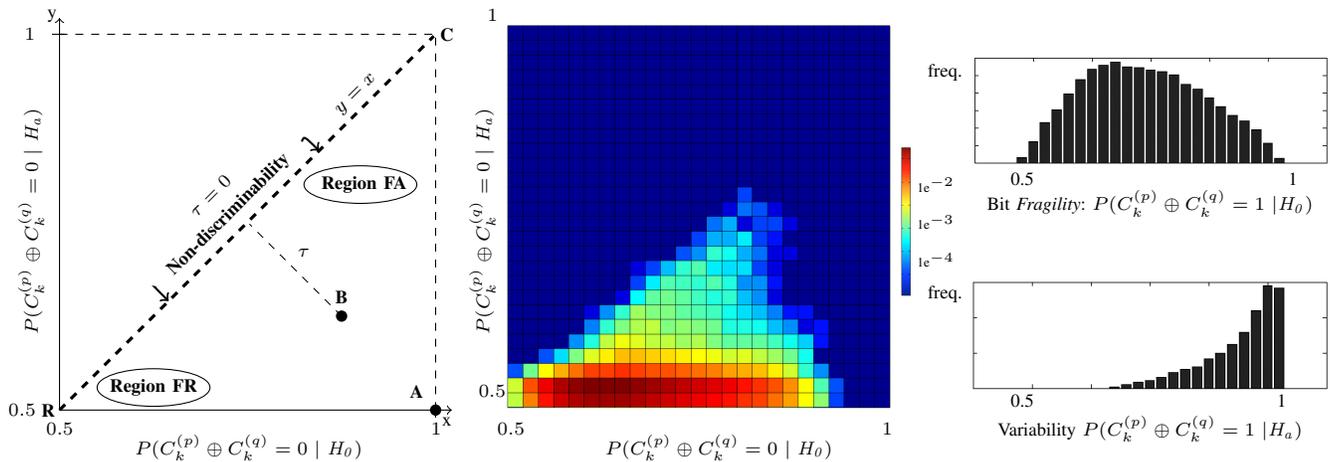


Fig. 1. The left figure is a schematic representation of bit *discriminability*. The central plot is a 2D histogram of the discriminability of bits extracted from the University of Bath data set. The histograms at the right side are the vertical and horizontal projections of the center plot: the upper histogram evidence different levels of bits fragility, whereas the bottom histogram points for additional non-random variation.

illustrates some of the images considered: the upper row regards the BATH data set and the subsequent rows represent the CASIA-Iris-Distance, UBIRIS.v2 and FRGC data sets.

#### A. DataSets

- The University of Bath data set<sup>1</sup> contains 32,000 NIR images from 800 subjects. From these, 6,000 images from 1,000 different classes (eyes) with very good quality were considered, to represent the optimal conditions where a recognition system work. All irises are sharp, without relevant occlusions and in frontal view.
- The CASIA-Iris-Distance set<sup>2</sup> was collected by the CASIA long-range device in a relatively unconstrained setup. Images feature blink, motion blur, off-axis gaze and other small anomalies, representing NIR data of moderate quality. A set of 9,521 images (127 subjects, 814 classes) was used, for which segmentation and noise detection was confirmed by visual inspection.
- The UBIRIS.v2 [28] dataset has 11,102 images from 261 subjects, acquired at visible wavelengths between three and eight meters away, under dynamic lighting conditions and unconstrained setups. Images are high heterogenous in terms of quality, with glossy reflections across the iris, significant occlusions due to eyelids and eyelashes, off-angle and blurred data. 5,340 images from 518 classes) were selected from this dataset, all of them accurately segmented. All these images were converted to grayscale.
- The FRGC [23] data set served initially for face recognition experiments and is a specially hard set for iris recognition, due to its limited resolution. The still images subset from both the controlled / uncontrolled setups was used. Images are typically frontal, with varying amounts of light, shadows and glossy reflections that occlude

portions of the irises. 4,360 from 868 classes images were selected from this data set. All these images were reasonably segmented, according to visual inspection, and were converted to grayscale.

- The UBI\_SPECTRAL is a set of iris data acquired in a synchronous way in the NIR and VW wavelengths, with a multispectral JAI AD080-GE camera, in a laboratorial controlled acquisition protocol. It contains data from 34 subjects, with 80 images per subject (20 NIR and 20 visible images, divided into 2 acquisition sessions). All images are frontal, sharp and practically noise-free. All the images were accurately segmented and VW images was converted to grayscale.

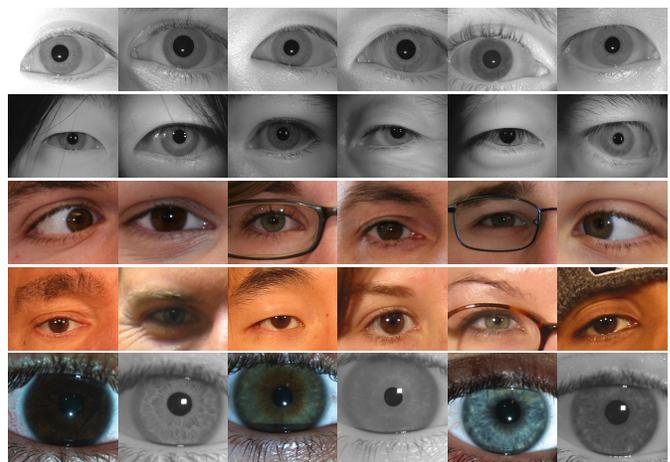


Fig. 2. Examples of the data sets used in the experimental evaluation. From top to bottom rows: BATH, CASIA-Iris-Distance, UBIRIS.v2, FRGC and UBI\_SPECTRAL datasets.

For the BATH, CASIA, UBIRIS and FRGC data sets, random samples composed by half of the within-class comparisons available and the same number of between-classes comparisons

<sup>1</sup><http://www.smartsensors.co.uk/products/iris-database/32-000-full-set/>

<sup>2</sup><http://biometrics.idealtest.org/>

were created. Next, in an iterative way, fully disjoint sets of the learning data were used to evaluate the recognition performance and estimate  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0)$  and  $P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ . Starting with  $t_w = 1,000$  within-class and  $t_b = 5,000$  between-classes comparisons, the recognition performance was obtained. At each iteration  $t$ , the number of comparisons was increased by a constant factor ( $t_w^{(t+1)} = 1.1 t_w^{(t)}$ ,  $t_b^{(t+1)} = 1.1 t_b^{(t)}$ ) until the performance values converged (after  $t = 49$  iterations).

The procedure described above can also be regarded as a way to mitigate the different number of degrees-of-freedom (DOF) in each sample and the way this factor might increase the correlation between bits and bias the subsequent results. Note that the data samples had varying number of classes, corresponding to different DOFs in the sets of pairwise comparisons. Even though, as the learning process was only stopped when performance was observed to converge, this implies that for large amounts of data, these changes in the number of DOF do not lead to substantial changes in the recognition performance. Finally, it should be stressed that none of these comparisons was used in feature selection, i.e., the learning and test sets were mutually exclusive. For reproducibility of the results, the set of all the pairwise comparisons used in the scope of this paper is available at<sup>3</sup>.

### B. Iris Segmentation, Noise-Free Texture Detection and Normalization

An overview of the preprocessing chain and of the resulting images is given in Fig. 3. For all the data sets considered in this paper, the unoccluded regions of the irises were obtained according to the algorithm of Tan *et al.* [33] (image at the upper-right corner). Next, an elliptical parameterization was chosen for both iris boundaries, using the random elliptic Hough transform. Based on the parameterization of the pupillary and scleric iris boundaries, the translation into the dimensionless pseudo-polar coordinate system was carried out according to the Daugman's *rubber sheet* model (images at the bottom-left corner).

## V. AMOUNT OF INFORMATION AND INDIVIDUAL BIT DISCRIMINABILITY

### A. Amount of Information in Iris Patches

The amount of information available in small iris patches was measured by the Shannon entropy criterion, quantifying (in terms of bits) the expected value for the amount of information in square regions  $p \times p$  of the normalized image  $I$ :

$$h(I_{p \times p}) = - \sum_i P(I_{p \times p} = i) \log_2 (P(I_{p \times p} = i)), \quad (3)$$

where  $P(I_{p \times p} = i)$  is the probability for the  $i^{\text{th}}$  intensity in the patch.

In order to fairly compare the local entropy between NIR and VW data, without concerns about the lighting conditions, the levels of iris pigmentation of the intrinsic features of

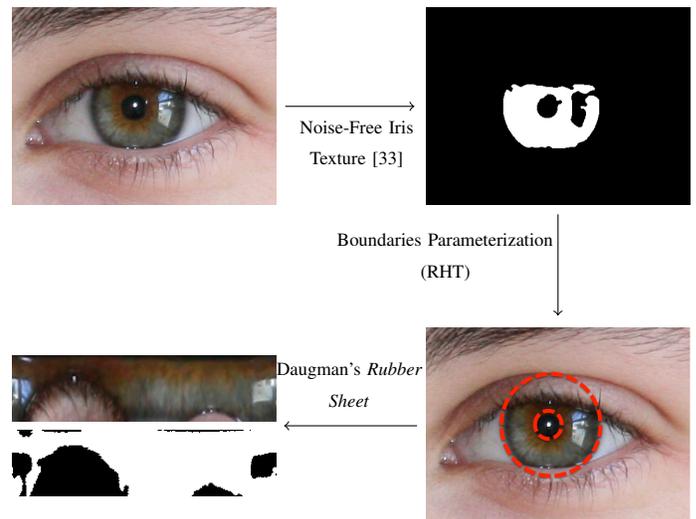


Fig. 3. Processing chain for segmenting the irises, detecting the noise-free iris regions, parameterizing the boundaries and converting them into the polar domain.

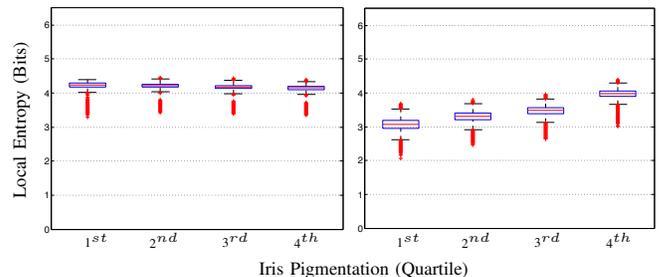


Fig. 4. Comparison between the average entropy values (3) observed in iris patches of NIR (left plot) and VW data (right plot). Results regard the UBI\_SPECTRAL set.

the subjects in each dataset, the UBI\_SPECTRAL data set was used. In this case, as all images were acquired in highly controlled lighting conditions and in synchronous way for the NIR and VW wavelengths, the effect of the above factors should be minimised. Also, we used the average intensity inside the iris of the grayscale version of the VW images as an estimator of the levels of iris pigmentation. This way, the highest values correspond to light pigmented irises (light blue), whereas the lowest intensities are from the heavily pigmented irises (dark brown / black). Fig. 4 compares the box plots of the local entropy values obtained for the NIR (left plot) and VW (right plot) data, with respect to the levels of iris pigmentation (horizontal axes). Four groups of pigmentation were considered, corresponding to dark brown / black (1<sup>st</sup> quartile of the average intensities), light brown (2<sup>nd</sup> quartile), green / dark blue (3<sup>rd</sup> quartile) and light blue (4<sup>th</sup> quartile) irises. We confirmed that values vary much more in VW than in NIR data, and, for the former wavelength, is notoriously higher for light pigmented than for dark irises. For NIR data, the heavy pigmented irises (3<sup>rd</sup> and 4<sup>th</sup> quartiles) have slightly higher local entropy than the remaining classes, which is in

<sup>3</sup><http://www.di.ubi.pt/~hugomcp/BitFragility>

exact opposition of the VW case. Also, the entropy in patches across the iris is more heterogeneous for VW images than for NIR, which is particularly evident for light pigmented irises. With regard to local variations, heavy pigmented irises acquired in VW have not only a relatively low amount of local information, but also display low variability between patches. i.e., in practice provide much flatter distributions for VW than for NIR data.

For the remaining datasets, Fig. 5 quantifies the amount of information in  $p = 9$  patches. Even noting that the comparison between data sets might be unfair (the original images have different resolution), the immediate conclusion is the higher homogeneity of values observed in NIR data than in the VW case. Note that the average values were also much higher in NIR than in VW data, which actually implies that the NIR images provide more heterogeneity in terms of intensities in iris patches than VW data.

Also, we observed that the pupillary regions are the most valuable in NIR images, which is not evident in VW. Regarding the FRGC dataset, there are two regions near the pupillary boundary with values notoriously higher than the remaining regions. We confirmed that they were due to frequent reflections not detected by the noise-free segmentation phase. Also, noted that in the FRGC set the bottom parts of the irises have evidently smaller amounts of information than the upper parts, probably due to the lighting sources from above that propitiate shadows in these regions.

### B. Filters Parameterizations

The discriminating power provided by each region of the iris was assessed with respect to two families of filters: 1) Gabor kernels, which faithfully model simple cells in the visual cortex of mammalian brains [4] and are used in the most acknowledged iris recognition algorithm; and 2) Multi-lobe differential filters (MLDF), which were recently reported as a relevant advance in the iris recognition field [32].

The impulse response of a Gabor kernel is defined by the multiplication of a harmonic and a Gaussian function:

$$\mathcal{G}[x, y, \omega, \varphi, \sigma] = \exp\left[\frac{-x^2 - y^2}{\sigma^2}\right] \exp[2\pi\omega i\Phi], \quad (4)$$

where  $\Phi = x \cos(\varphi) + y \sin(\varphi)$ ,  $\omega$  is the spatial frequency,  $\varphi$  is the orientation and  $\sigma$  the standard deviation of a Gaussian kernel (isotropic in our experiments,  $\sigma = 0.65\omega$ ). A more general form of Gabor filters can be found in the literature (e.g., [7]), allowing for different scales along the axes ( $\sigma_x$  and  $\sigma_y$ ). In this paper, to keep moderate the dimension of the parameterisation space, we decided to use exclusively filters with the same scale along the axes.

Regarding the MLDF filters, they can be parameterised in terms of the number of positive/negative lobes, location, scale, orientation and inter-lobe distance. To keep the number of possibilities moderately low, only Gaussian kernels with balanced number of positive / negative lobes (1/1, 2/2, ...) and equal scale for both types of lobes were considered. Hence, the MLDF filters are expressed by:

$$\mathbf{m}[\mathbf{x}_j, \mu_j, \sigma_j] = \sum_{j=1}^{t_l} (-1)^{j+1} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[\frac{-(\mathbf{x}_j - \mu_j)^2}{2\sigma_j}\right], \quad (5)$$

where  $\mathbf{x}_j = (x_j, y_j)$  is the center of each of the  $t_l$  lobes. Next,  $\mathbf{k} = \{\mathbf{m}, \mathbf{g}\}$  filters were convolved with each normalized iris image  $\mathbf{I}$ , providing a set of coefficients. The sign of the coefficients was obtained, i.e.,  $\mathbf{C}$  is the vector representation of  $\text{sgn}(\mathbf{I} * \mathbf{k})$ . In terms of parameterisations tested per filter, for Gabor kernels the wavelength (px.)  $\omega : \{1 : 1 : 14\}$ , the orientation  $\varphi : \{0, \pi/4, \pi/2, 3\pi/4\}$  and the Gaussian sigma  $\sigma : 0.65\omega$ . Regarding MLDFs, the number of lobes  $t_l : \{1/1, 2/2, 3/3, 4/4\}$  and the Gaussian sigma  $\sigma : \{1, 2, 3, 4, 5, 6\}$ .

Fig. 6 expresses the variations in discriminability with respect to each parameter of the filters. The continuous lines represent the BATH dataset, the dashed lines with the diamond marks regard the CASIA-Iris-Distance. The UBIRIS.v2 is given by the dotted lines with triangular marks and the FRGC dataset by the dashed lines with circular marks. Above each plot we illustrate a normalized iris image and represent the filters that correspond to the nearby positions in the plot. Generally, the discriminability was substantially higher for MLDF than for Gabor filters. In case of the latter filters, larger wavelengths consistently increased the discriminability, essentially because they have a reduced sensitivity to outlier values due to acquisition artifacts. Orientation is another relevant parameter for Gabor kernels, where filters that analyze features that spread radially in the normalized data provided much better results. Regarding MLDF filters, filters with more lobes got worse results, which might be due to the *cross-elimination* effect of differences between lobes. Surprisingly, the variation in results with respect to the sigma of the Gaussian kernel were not so evident as in the case of Gabor kernels.

### C. Bit Discriminability

The discriminability  $\tau$  of each bit extracted was obtained. Note that the iris patches evolved in the convolution for each bit contribute to the result in different degree, according to the magnitude of the kernel at each point, i.e., if a kernel has very small value at a specific position, the corresponding intensity on the patch almost does not affect the result. This way, the contribution of each location  $[x, y]$  in the iris to the bit value is given by:

$$\Psi[x, y] = \frac{\sum_i (|\mathbf{k}_i[x - r_i, y - c_i]| \tau(i))}{\sum_i |\mathbf{k}_i[x - r_i, y - c_i]|}, \quad (6)$$

where  $[r_i, c_i]$  is the central position of the  $i^{\text{th}}$  filter  $\mathbf{k}_i$  and  $\tau(i)$  is the discriminability of the  $i^{\text{th}}$  bit.

Fig. 7 gives the discriminability provided by each region of the iris in the Cartesian and polar coordinate systems: the maximal values were obtained for the NIR data sets, both for Gabor and MLDF filters. Interestingly, in all cases the lower parts of the iris were better than the upper parts, which are

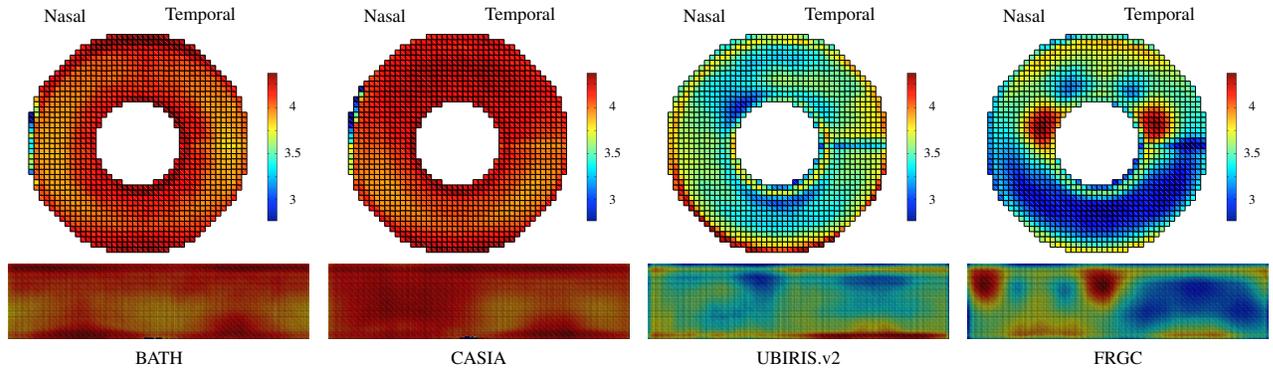


Fig. 5. Average amount of information (Shannon entropy in  $9 \times 9$  patches of the normalised images) across the different regions of the irises in the BATH, CASIA-Iris-Distance, UBIRIS.v2 and FRGC datasets. Values are expressed in bits, and enable to perceive the gap of information between NIR (BATH and CASIA) and VW (UBIRIS.v2 and FRGC) iris data.

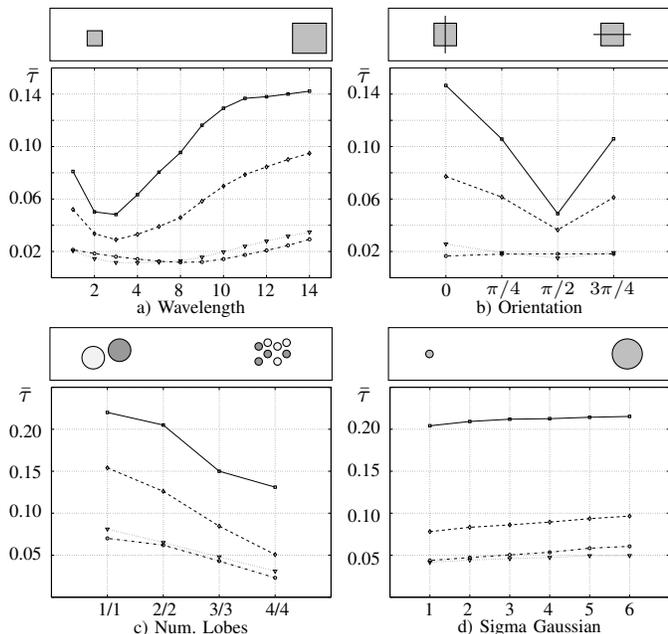


Fig. 6. Average discriminability  $\bar{\tau}$  of the bits in iris codes, regarding filters parameterization. The upper row regards the Gabor kernels (wavelength and orientation parameters) and the bottom row corresponds to the MLDF filters (number of lobes and sigma of the Gaussian kernel).

more frequently occluded by eyelids. Globally, MLDF filters provided more homogeneous values than Gabor filters. For VW data, regions nearby the pupillary boundary are worse than the middle and outer bands, probably due to the difficulty in obtaining reliable estimates of the pupillary boundary in VW images.

Regarding the radial bands in the iris, even though the maximal discriminability was observed for the middle bands, this might not be due to biological properties of the iris texture. Instead, the middle bands are the regions where the largest filters can be applied without surpassing the iris boundaries. As illustrated in Fig. 6, large filters tend to produce more

discriminant bits, which accords the results given in [13].

It is interesting to note the reduced correlation between the amounts of information in iris patches and the discriminability of each patch. For the BATH data set, the levels of linear correlation between variables  $h[x, y]$  and  $\Psi[x, y]$  were  $-0.12/-0.38$  (Gabor/MLDF filters), and  $-0.40/-0.22$  for the CASIA-Iris-Distance set. Regarding the VW data, values were  $0.16/-0.02$  for the UBIRIS.v2 and  $-0.34/-0.41$  for the FRGC datasets. These low correlation values in terms of magnitude and sign (negative in 7/8 of the cases) give space for additional research about iris feature extraction / matching strategies that profit in a better way from the amount of information that is locally available.

#### D. Discriminability vs. Fragility

This section illustrates the differences between the previously reported concept of *fragility* and the concept of *discriminability* discussed in this paper. In Fig. 8 we highlight the bits from one of the datasets used in the paper (BATH) where the largest differences in the fragility and discriminability z-scores  $z(\cdot)$  were observed (using Gabor filters), i.e.,  $z(\Psi[x, y]) - z(1-\text{fragility})$ . Here, red / orange regions are particularly discriminative but - even though- their bits have a relatively large fragility, whereas blue colours represent bits that are not fragile, but still have not particularly high discriminability. The less interesting cases (bits consistent and non-fragile or bits non-consistent and fragile) have values near to 0. This clearly distinguishes both concepts, i.e., apart fragility, there is a new family of bits (discriminability), which can be used to develop better iris recognition algorithms in the future.

## VI. CLASSIFIERS DISCRIMINABILITY AND PERFORMANCE ANALYSIS

The discriminability provided by each bit is the root of iris recognition performance. However, this value regards the performance of *individual players* (bits), and not of the *team* (iris code), which is the one that actually matters. Various highly discriminating bits do not necessarily constitute a good

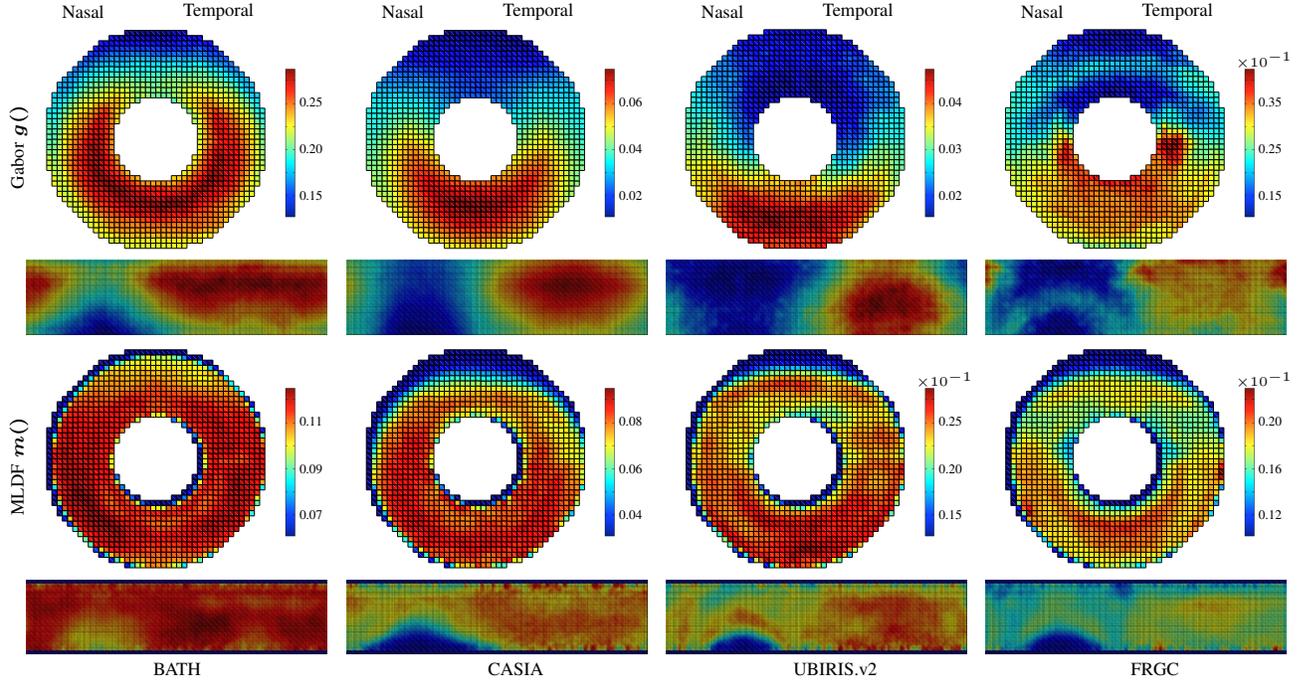


Fig. 7. Average bit discriminability  $\Psi[x, y]$  across the iris. Values are given for the Cartesian and polar coordinate systems, for the four data sets considered in this paper.

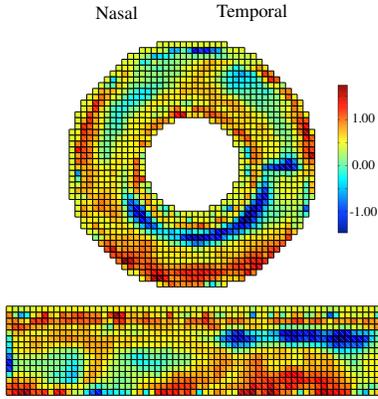


Fig. 8. Illustration of the differences between the *fragility* and *discriminability* concepts. Regions in red / orange illustrate discriminant bits that - even though - are relatively fragile, whereas blue regions denote the opposite case, i.e., bits that are not particularly fragile but still have not the highest discriminability (results obtained using Gabor  $g()$  kernels, in the BATH dataset).

recognition system, if they are strongly correlated. This kind of analysis requires to obtain the conditional probabilities  $P((\sum_{k^*=1}^t C_{k^*}^{(p)} \oplus C_{k^*}^{(q)}) = c | H_a)$  and  $P((\sum_{k^*=1}^t C_{k^*}^{(p)} \oplus C_{k^*}^{(q)}) = c | H_0)$ , for  $c \in \{0, \dots, t\}$ , where  $k^*$  denotes the index of the bits in the iris code.

#### A. Feature Selection and Performance

To obtain the best combination of bits for iris codes, three families of feature selection algorithms were considered. The

simplest was the Fisher-score [10]. The goal is to find a feature subset such that in the projected subspace the distances between points in different classes are maximal, while maintaining the distances between points in the same class as small as possible. For simplicity, let  $C_k$  denote the  $k^{th}$  bit of feature set. The objective function is given by:

$$F(C_k) = \frac{\sum_j t_j (\mu_j^k - \mu^k)^2}{\sum_j t_j (\sigma_j^k)^2}, \quad (7)$$

where  $\mu_j^k$  and  $\sigma_j^k$  are the mean and standard deviation of the  $k^{th}$  bit in the  $j^{th}$  ( $\in \{0, 1\}$ ) class.  $\mu^k$  is the overall mean of the  $k^{th}$  feature and  $t_j$  the number of elements in the  $j^{th}$  class. After obtaining the Fisher-scores, the top- $m$  ranked features were selected.

The Minimum Redundancy - Maximum Relevance (mRMR) algorithm [22] for discrete data was also considered. For every pair of features, the mutual information was obtained:

$$M(C_p, C_q) = \sum_{i=0}^1 \sum_{j=0}^1 P(C_p = i, C_q = j) \cdot \log_2 \frac{P(C_p = i, C_q = j)}{P(C_p = i)P(C_q = j)}. \quad (8)$$

The minimum redundancy condition of a feature set  $\mathcal{S}$  corresponds to  $\min \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} M(C_i, C_j)$ . Then, using the class variable  $h \in \{0, 1\}$ , the maximum relevance of  $\mathcal{S}$  is given by  $\max \sum_{j \in \mathcal{S}} M(h, C_j)$ . Both terms enable to select features iteratively:

$$\max_{i \in \Omega} \left[ M(i, h) - \frac{1}{S} \sum_{j \in S} M(i, j) \right]. \quad (9)$$

This process attains a near optimal solution in  $O(|S|N)$  time complexity, with  $|\cdot|$  denoting the set cardinality. Finally, the Sequential Floating Feature Selection (SFFS) algorithm [29] was used, with the objective function:

$$J(\{C_p, S\}) = \frac{|\mu_0^{\{C_p, S\}} - \mu_a^{\{C_p, S\}}|}{\sqrt{(\sigma_0^{\{C_p, S\}})^2 + (\sigma_a^{\{C_p, S\}})^2}}, \quad (10)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviations. The subscript denotes the class and the superscript  $\{\cdot, \cdot\}$  is the concatenation operator. Starting with the empty set  $S = \emptyset$ , at each iteration the best feature was taken:  $C_p^* = \arg_p \max J(C_p, S)$  and added to the selected set  $S = \{S, C_p^*\}$ . After each insertion, the exclusion of features previously selected was considered:  $C_p^* = \arg_p \max J(\{S \setminus C_p^*\})$ , where " $\setminus$ " denotes set complement. Whenever  $J(\{S \setminus C_p^*\}) > J(S)$ ,  $C_p^*$  was excluded from  $S$ .

Fig. 9 gives the ROC curves with respect to the feature selection algorithms. For all cases, 2,048 features were selected. The upper row expresses the results for Gabor filters and the bottom row regards the results for MLDF filters. In summary, the SFFS algorithm fed by MLDF features produced the best results for all datasets, in some cases with substantial differences in performance with respect to the remaining configurations. For Gabor kernels it was harder to perceive which was the best feature selection algorithm, but the mRMR algorithm got the best results in half of the cases. Table II summarizes the performance (in terms of the  $d'$  index [11] and of the Area Under Curve (AUC) value) for each data set, feature selection algorithm and type of kernel used in feature encoding.

It is evident that MLDF filters performed better than Gabor, both in NIR and VW data. This might be due to the interlacing structure of the muscles in the vascularised stroma of the iris, that augment the correlation of intensities between non-adjacent regions. As pointed out by Daugman [4], Gabor kernels optimally resemble the receptive field profiles of neurons in the visual cortex of the brains of mammals and - as such - exclusively analyze adjacent patterns. Also, MLDFs are actually a generalization of Gabor filters, meaning that any Gabor kernel can be roughly modeled by a MLDF parameterization.

Fig. 10 illustrates the iris regions from where the best bits are more frequently extracted, noting the stable patterns obtained among all the datasets: for Gabor kernels, the lateral and lower parts of the iris are the most important, almost ignoring the upper parts. Note that the difference between the values observed for the bottom / upper parts of the iris is more notorious in the Gabor than in MLDFs kernels. This is due to the fact that Gabor kernels analyse adjacent patterns, and to extract a pattern from the upper part of the iris, a large and continuous patch should be considered noise-free (which is not too frequent, due to eyelids and eyelashes). In opposition, as MLDFs analyse non-adjacent patterns, enable that even small patches of the upper regions of the iris can be used to extract

an useful bit.

Both for Gabor and MLDF kernels, the middle bands are the most frequently selected and can be considered the best for biometric recognition. In opposition, regions near the boundaries the worst, both in NIR / VW data and Gabor / MLDF filters. Finally, note the higher homogeneity of the values obtained for MLDF filters than for Gabor, where highly salient regions appear in small patches of the iris.

### B. Codes Quantization: How Much Discriminating Information Is Lost?

In the most acknowledged iris recognition algorithm, only phase information is used in recognition. Amplitude information is not considered reliable, as it depends of imaging contrast, illumination and camera gain. Accordingly, Hollingsworth *et al.* [13] observed that most inconsistencies in iris codes are due to the coarse quantization of the phase response, and disregarded bits from filter responses near the axes.

Even considering the afore arguments reasonable, we assessed the amounts of discriminating information contained in the filter responses near the axes. With respect to the traditional strategy of keeping only the sign of coefficients (function A) in Fig. 11), two other strategies were considered: a linear mapping of the magnitude of the responses, yielding real-valued coefficients matched by the  $\ell_2$  norm (function C) in Fig. 11); and a trade-off of both strategies, according to a sigmoid-based transform that maps large magnitude values to the 0/1 values, but weights values near the axes to real values in the [0,1] interval. In this case, the  $\ell_2$  norm was also used as matching function.

The ROC curves given at the right side of Fig. 11 compare the recognition performance with respect to each quantization strategy and Table III summarizes the results, giving the Area Under Curve (AUC) and the decidability index  $d'$  that, as suggested by Daugman [6], measures how well separated the genuine / impostor distributions are:

$$d' = \frac{|\mu_G - \mu_I|}{\sqrt{\frac{1}{2}(\sigma_I^2 + \sigma_G^2)}}, \quad (11)$$

where  $\mu_I = \frac{1}{k} \sum_i d_i^I$  and  $\mu_G = \frac{1}{m} \sum_i d_i^G$  are the means of the genuine (G) and impostor (I) scores and  $\sigma_I = \frac{1}{k-1} \sum_i (d_i^I - \mu_I)^2$  and  $\sigma_G = \frac{1}{m-1} \sum_i (d_i^G - \mu_G)^2$  their standard deviations.

Two opposite conclusions were drawn: for Gabor filters, the best results were observed when using the traditional  $\text{sign}()$  quantization function. In this case, using scalars instead of sign bits even decreased the recognition performance. Oppositely, for MLDF filters, the best results were observed when using the proposed sigmoid function, i.e., when the coefficients of small magnitude were also considered for the matching process. This points for the conclusion that there is actually reliable discriminating information in the coefficients near the origin. However, these coefficients are less reliable than those with large magnitude, as in no case the linear mapping strategy got results close to any of the remaining strategies.

Note that the above conclusions were drew based on the reported AUC and  $d'$  values, which in the large majority of

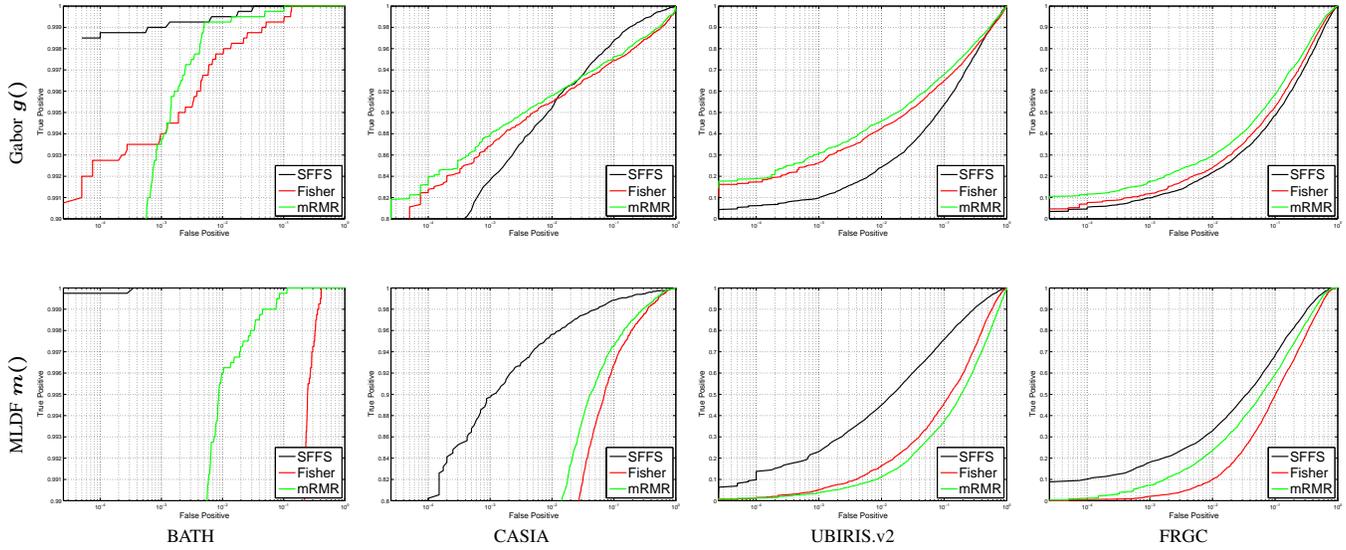


Fig. 9. Comparison between the recognition performance observed per data set and type of kernel, according to the best three feature selection strategies.

TABLE II

SUMMARY OF THE RECOGNITION PERFORMANCE OBSERVED FOR EACH DATA SET, WITH RESPECT TO THE TYPE OF KERNELS AND THE FEATURE SELECTION ALGORITHMS.

Dataset		BATH		CASIA-Iris-Distance		UBIRIS.v2		FRGC	
Feat.	Feat. Sel.	d'	AUC	d'	AUC	d'	AUC	d'	AUC
Gabor	Fisher-score	7.54 ± 0.01	0.991 ± 0.002	3.36 ± 0.01	0.970 ± 0.002	1.43 ± 0.01	0.8471 ± 0.006	1.28 ± 0.02	0.809 ± 0.003
MLDF	Fisher-score	7.05 ± 0.01	0.990 ± 0.002	2.79 ± 0.01	0.962 ± 0.002	0.74 ± 0.01	0.779 ± 0.007	1.29 ± 0.02	0.809 ± 0.006
Gabor	mRMR	6.35 ± 0.02	0.989 ± 0.001	3.44 ± 0.02	0.971 ± 0.002	1.52 ± 0.02	0.812 ± 0.006	1.41 ± 0.02	0.838 ± 0.004
MLDF	mRMR	5.81 ± 0.01	0.987 ± 0.001	3.02 ± 0.01	0.968 ± 0.003	0.79 ± 0.01	0.707 ± 0.004	1.53 ± 0.01	0.851 ± 0.004
Gabor	SFFS	8.79 ± 0.01	0.994 ± 0.001	3.20 ± 0.01	0.976 ± 0.001	1.23 ± 0.01	0.818 ± 0.006	1.12 ± 0.02	0.778 ± 0.006
MLDF	SFFS	9.15 ± 0.01	0.994 ± 0.001	3.89 ± 0.01	0.990 ± 0.001	1.88 ± 0.01	0.904 ± 0.003	1.74 ± 0.01	0.892 ± 0.006

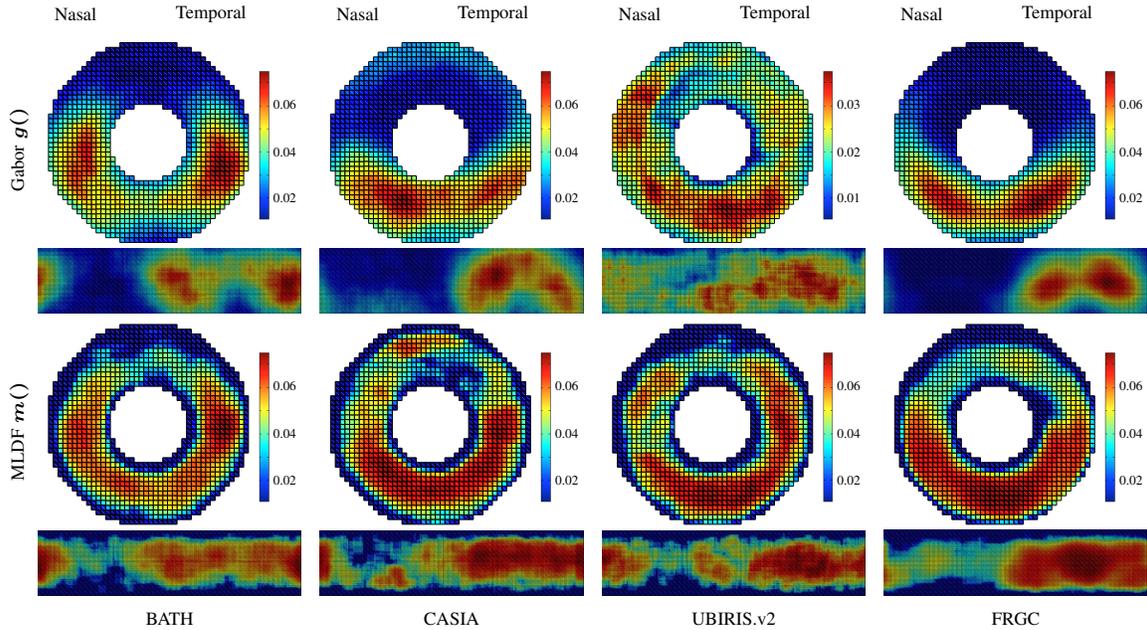


Fig. 10. Most important regions of the iris for biometric recognition. Results express how many times each position of the iris was considered to extract bits selected to be included in the final iris codes.

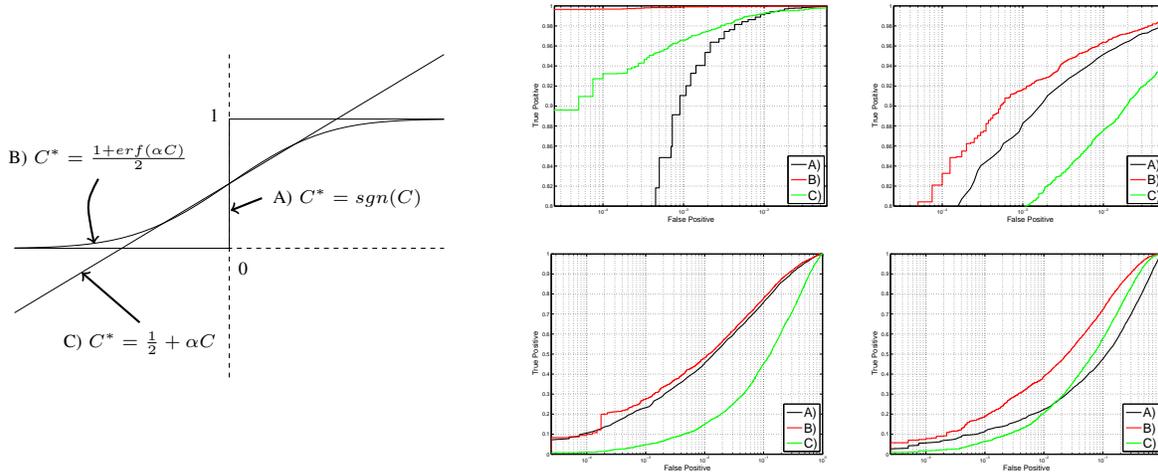


Fig. 11. At left: Three different strategies for code quantization: A) binary; B) sigmoid function; and C) linear mapping. At right: recognition performance with respect to A), B) and C) code quantization strategies for BATH (upper-left plot), CASIA-Iris-Distance (upper-right), UBIRIS.v2 (bottom-left) and FRGC data sets (bottom-right).

the cases were observed to be in agreement. The exceptions occurred mostly in cases where the shape of the genuine / impostor distributions were the farthest from Gaussian distributions. For these particular cases, we relied mostly in the AUC value, as it does not require a specific data distribution to report meaningful results.

## VII. CONCLUSIONS

This paper introduced the concept of bit *discriminability*, for iris recognition purposes. As a complement to the previously proposed concept of bit *fragility*, we noticed that not only bits have different probabilities of flipping their values among samples of an iris, but also the probability of observing agreeing bit values for different irises varies in a statistically significant way. Even though this phenomenon was observed both for NIR and VW data, it is much more evident for the latter wavelength, which should be due to the corneal reflections that are determined by the ambient light and appear (typically) in the same positions of the irises in a data set.

Based on the proposed concept of bit discriminability, we compared the effectiveness of each region of the iris for biometric recognition, with respect to multi-spectral data and to two widely used types of filters (Gabor and MLDFs). Our main conclusions were:

- Iris Anatomy: 1) there is a poor correlation between the amount of information in iris patches and their usability for biometric recognition. This gives space for future research in terms of novel iris feature encoding / matching algorithms, particularly for NIR data, where the highest amounts of local information were observed. However, note that by converting the VW data to grayscale we disregarded a substantial amount of the information in this kind of images. In the scope of color information, there is evident space for further analysis, as the effectiveness of different color spaces, or of different linear / non-linear strategies to fusing color channels; 2) the bottom parts

of the iris are less likely to be occluded by eyelids than the upper parts. Shadows are more frequent in the upper parts, particularly when illumination from above is used. However, note that commercial iris systems do not use illumination from above; 3) there is a direct correlation between the size of the filters and the discriminability of the resulting bits. This turns the middle radial bands of the iris as the most important, as they are those where the largest filters can be used there without surpassing the iris boundaries; 4) there is no evidence that either the temporal / nasal sides of the iris should be preferred over the other. However, in case of illumination from the side, shadows by the nose are likely to appear in the iris, which might also decrease performance;

- Filters: 1) MLDFs provide better performance than Gabor kernels due to their ability of exploiting non-adjacent patterns. This property is particularly interesting for tissues with interlacing fibers, such as the human iris; 2) there is a strong agreement between the best iris regions obtained for MLDF and Gabor filters, suggesting that the choice for the best regions to perform iris recognition is relatively independent of the kind of filters used;
- Data Spectrum and Discriminating Information: 1) Both in NIR and VW data, the signal magnitude carries valuable discriminating information, which should be particularly useful for hard acquisition environments; 2) NIR data provides more discriminating information than VW data, particularly in the pupillary bands; 3) the bit discriminability in VW data appears to spread in a more uneven way than in NIR data. This topic should be subject of further research, as we cannot determine whether this was due to the evidently *wilder* acquisition setups of the VW data sets than of the NIR sets used in this paper.

## ACKNOWLEDGEMENTS

Portions of the research in this paper use the CASIA-IrisV4 collected by the Chinese Academy of Sciences, Institute of

TABLE III  
VARIATIONS IN RECOGNITION PERFORMANCE WITH RESPECT TO DIFFERENT STRATEGIES FOR CODE QUANTIZATION.

Dataset	Features	A) sign()		B) sigmoid()		C) linear (no quantization)	
		d'	AUC	d'	AUC	d'	AUC
BATH	Gabor	8.79 ± 0.01	0.994 ± 0.001	7.08 ± 0.01	0.992 ± 0.001	6.52 ± 0.01	0.990 ± 0.001
BATH	MLDF	9.15 ± 0.01	0.994 ± 0.001	8.82 ± 0.01	0.993 ± 0.001	5.89 ± 0.01	0.988 ± 0.001
CASIA-Iris-Distance	Gabor	3.20 ± 0.01	0.982 ± 0.001	3.16 ± 0.01	0.982 ± 0.001	3.05 ± 0.02	0.971 ± 0.001
CASIA-Iris-Distance	MLDF	3.89 ± 0.01	0.990 ± 0.001	4.12 ± 0.01	0.984 ± 0.001	3.13 ± 0.01	0.982 ± 0.001
UBIRIS.v2	Gabor	1.23 ± 0.01	0.813 ± 0.006	1.16 ± 0.02	0.793 ± 0.007	0.82 ± 0.02	0.720 ± 0.006
UBIRIS.v2	MLDF	1.88 ± 0.01	0.904 ± 0.003	1.96 ± 0.01	0.917 ± 0.003	1.02 ± 0.01	0.766 ± 0.009
FRGC	Gabor	1.12 ± 0.02	0.792 ± 0.006	1.01 ± 0.02	0.770 ± 0.008	0.83 ± 0.01	0.731 ± 0.007
FRGC	MLDF	1.74 ± 0.01	0.892 ± 0.006	1.88 ± 0.02	0.908 ± 0.002	1.47 ± 0.02	0.849 ± 0.007

Automation (CASIA).

## REFERENCES

- [1] R.M. Bolle, S. Pankanti, J.H. Connell and N. Ratha. Iris Individuality: A Partial Iris Model. Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), vol. 2, pag. 927-930, 2004.
- [2] K.W. Bowyer, K.P. Hollingsworth and P.J. Flynn. A Survey of Iris Biometrics Research: 2008-2010. *Handbook of Iris Recognition*, M.J. Burge and K.W. Bowyer (eds.), Advances in Computer Vision and Pattern Recognition, Springer-Verlag, pag. 15-54, 2013.
- [3] K.W. Bowyer, S.E. Baker, A. Hentz, K.P. Hollingsworth, T. Peters and P.J. Flynn. Factors that degrade the match distribution in iris biometrics. *Identity in the Information Society*, vol. 2, Issue 3, pag. 327-343, 2009.
- [4] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America, A*, vol. 2, no. 7, pag. 1160-1169, 1985.
- [5] J. Daugman. High Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pag. 1148-1161, 1993.
- [6] J. Daugman. Biometric decision landscapes. *University of Cambridge Technical Report*, UCAM-CL-TR-482, ISSN: 1476-2986, 2000.
- [7] J. Daugman. Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, vol. 94, no. 11, pag. 1927-1935, 2006.
- [8] W. Dong, Z. Sun and T. Tan. Iris matching based on personalized weight map. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pag. 1744-1757, 2011.
- [9] G. Dozier, K. Frederiksen, R. Meeks, M. Savvides, K. Bryant, D. Hopes and T. Munemoto. Minimizing the Number of Bits Needed for Iris Recognition via Bit Inconsistency and GRIT. Proceedings of the *IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications* (CIB 2009), pag. 30-37, 2009.
- [10] R. Duda, P. Hart and D. Stork. Pattern Classification. Wiley-Interscience Publication, ISBN: 978-0-471-05669-0, 2001.
- [11] M. Hazewinkel (ed.). Kolmogorov-Smirnov test. *Encyclopedia of Mathematics*, Springer, 2001.
- [12] K.P. Hollingsworth, K.W. Bowyer and P.J. Flynn. All Iris Code Bits are Not Created Equal. First IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS 2007), pag. 1-6, 2007.
- [13] K.P. Hollingsworth, K.W. Bowyer and P.J. Flynn. The best bits in an iris code. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, pag. 964-973, 2009.
- [14] K.P. Hollingsworth, K.W. Bowyer and P.J. Flynn. Improved Iris Recognition Through Fusion of Hamming Distance and Fragile Bit Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pag. 2465-2476, 2011.
- [15] A. Kong, D. Zhang and M. Kamel. An Analysis of IrisCode. *IEEE Transactions on Image Processing*, vol. 19, no. 2, pag. 522-532, 2010.
- [16] A. Kong. IrisCode Decompression Based on the Dependence between Its Bit Pairs. *IEEE Transactions on Image Processing*, vol. 34, no. 3, pag. 506-520, 2012.
- [17] A. Kong. Modeling IrisCode and Its Variants as Convex Polyhedral Cones and Its Security Implications. *IEEE Transactions on Image Processing*, vol. 22, no. 3, pag. 1148-1160, 2013.
- [18] A. Kumar and T. Chan. Iris recognition using quaternionic sparse orientation code (QSOC). In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pag. 59-64, 2012.
- [19] A. Kumar, T. Chan, and C. Tan. Human identification from at a distance face images using sparse representation of local iris features. In *Proceedings of the 5th IAPR International conference on biometrics*, pag. 303-309, 2012.
- [20] L. Levin, S. Nilsson, J. Hoeve, S. Wu, P. Kaufman and A. Alm. Adler's Physiology of the Eye. *Saunders*, 11th edition, ISBN: 978-0-323-05714-1, 2011.
- [21] H. Mehrotra, M. Vatsa, R. Singh and B. Majhi. Does Iris Change Over Time? *PLoS ONE*, vol. 8, no. 11, e78333, 2013.
- [22] H. Peng, F. Long and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pag. 1226-1238, 2005.
- [23] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pag. 947-954, 2005.
- [24] J. Pillai, V. Patel, R. Chellappa and N. Ratha. Secure and robust iris recognition using random projections and sparse representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pag. 1877-1893, 2011.
- [25] J. Pillai, M. Puertas and R. Chellappa. Cross-sensor iris recognition through kernel learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pag. 73-85, 2014.
- [26] H. Proença and L.A. Alexandre. Iris Recognition: Measuring Feature's Quality for the Feature Selection in Unconstrained Image Capture Environments. Proceedings of the 2006 International Conference on Computational Intelligence for Homeland Security and Personal Safety (CIHSPS 2006), pag. 35-40, 2006.
- [27] H. Proença and L. A. Alexandre. Toward Covert Iris Biometric Recognition: Experimental Results From the NICE Contests. *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pages 798-808, 2012.
- [28] H. Proença, S. Filipe, R. Santos, J. Oliveira and L. A. Alexandre. The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-The-Move and At-A-Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pag. 1502-1516, 2010.
- [29] P. Pudil, J. Ferri, J. Novovičová. Floating search methods for feature selection with nonmonotonic criterion function. Proceedings of the 12th International Conference on Pattern Recognition, vol. 2, pag. 279-283, 1994.
- [30] C. Rathgeb, A. Uhl and P. Wild. On Combining Selective Best Bits of Iris-Codes. Proceedings of the *Biometrics and ID Management Workshop* (BioID'11), LNCS 6583, pag. 227-237, 2011.
- [31] A. Ross and A.K. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, vol. 24, pag. 2115-2125, 2003.
- [32] Z. Sun and T. Tan. Ordinal Measures for Iris Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pag. 2211-2226, 2009.
- [33] T. Tan, Z. He, and Z. Sun. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image and Vision Computing*, vol. 28, no. 2, pag. 223-230, 2010.
- [34] C. Tan and A. Kumar. Adaptive and localized iris weight map for accurate iris recognition under less constrained environments. In *Proceedings of the IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pag. 1-7, 2013.
- [35] T. Tan, X. Zhang, Z. Sun and H. Zhang. Noisy iris image matching by using multiple cues. *Pattern Recognition Letters*, vol. 33, pag. 970-977, 2012.



**Hugo Proença** received the B.Sc. degree from the University of Beira Interior, Portugal, in 2001, the M.Sc. degree from the Faculty of Engineering, University of Oporto, in 2004, and the Ph.D. degree from the University of Beira Interior, in 2007. His research interests are focused in the artificial intelligence, pattern recognition and biometrics. Currently, he serves as Assistant Professor in the Department of Computer Science, University of Beira Interior. He is the area editor (ocular biometrics) of the IEEE Biometrics Compendium Journal and member of the

Editorial Board of the International Journal of Biometrics. Also, he served as Guest Editor of special issues of the Pattern Recognition Letters, Image and Vision Computing and Signal, Image and Video Processing journals.