Iris Recognition in Visible Wavelengths and Unconstrained Conditions

Hugo Proença

Abstract One of the most challenging goals in biometrics research is the development of recognition systems to work in unconstrained environments and without assuming the subjects’ willingness to be recognised. This has led to the concept of non-cooperative recognition, which broaden the application of biometrics to forensics / criminal seek domains. In this scope, one active research topic seeks to use as main trait the ocular region acquired at visible wavelengths, from moving targets and large distances. Under these conditions, performing reliable recognition is extremely difficult, because such real-world data have features that are notoriously different from those obtained in the classical constrained setups of currently deployed recognition systems. This chapter discusses the feasibility of iris / ocular biometric recognition: it starts by comparing the main properties of near-infrared and visible wavelength ocular data, and stresses the main difficulties behind the accurate segmentation of all components in the eye vicinity. Next, it summarises the most relevant research conducted in the scope of visible wavelength iris recognition and relates it to the concept of periocular recognition, which is an attempt to augment classes separability by using - apart from the iris - information from the surroundings of the eye. Finally, the current challenges in this topic and some directions for further research are discussed.

1 Introduction

The iris is one of the most valuable traits for human identification and growing efforts have been concentrated in the development of this technology [7]. Fundamentally, three reasons justify this interest: (1) it is a naturally protected internal organ that is visible from the exterior; (2) it has a near circular and planar shape that
turns easier its segmentation and parameterization and (3) its texture has a predominantly phenotypic or chaotic appearance that is stable over lifetime. The accuracy of the deployed iris recognition systems is remarkable: a study of 200 billion cross-comparisons conducted by Daugman [15] reported false acceptance rates of order $10^{-6}$ with false rejections of 1%. Other independent evaluations ([30] and [47]) confirmed these results.

Current systems require high illumination levels, sufficient to maximize the signal-to-noise ratio in the sensor and to capture images of the discriminating iris features with sufficient contrast. However, if similar processes were used to acquire iris images from large distances, acceptable depth-of-field values would demand significantly higher f-numbers for the optical system, corresponding directly (squared) with the amount of light required for the process. Similarly, the motion factor will demand very short exposure times, which again will require too high levels of light. The American and European standards councils ([2] and [10]) proposed safe irradiance limits for near-infrared (NIR) illumination of near 10 mW/cm$^2$. In addition to other factors that determine imaging system safety (blue light, non-reciprocity and wavelength dependence), these limits should be taken into account, as excessively strong illumination can cause permanent eye damage. The NIR wavelength is particularly hazardous, because the eye does not instinctively respond with its natural mechanisms (aversion, blinking and pupil contraction).

The pigmentation of the human iris consists mainly of two molecules: brown-black Eumelanin (over 90%) and yellow-reddish Pheomelanin [48]. Eumelanin has most of its radiative fluorescence under the visible wavelength (VW), which—if properly imaged—enables the capture of a much higher level of detail, but also of many more artefacts, including specular and diffuse reflections and shadows. Also, the spectral reflectance of the sclera is significantly higher in the VW than in the NIR and the spectral radiance of the iris in respect of the levels of its pigmentation varies much more significantly in the VW than in the NIR. These optical properties are the biological roots behind the higher heterogeneity of the VW iris images, when compared with the traditional NIR data. Also, the types and number of artefacts likely to appear in VW and NIR data are notoriously different, which justify the need for specialized recognition strategies.

Fig. 1 illustrates the variations in appearance of NIR and VW images, with respect to the levels of iris pigmentation. These images were acquired using a multispectral device, in a synchronous way. It is particularly interesting to observe the inverse relation between the levels of minutia captured in NIR and VW data, with respect to the levels of iris pigmentation: while for light pigmented irises, much more detail is perceived in VW than in NIR images, it occurs the opposite for heavily pigmented irises (leftmost image). Note that this is a particularly concerning problem, as the large majority of the world population has heavily pigmented irises.
Fig. 1 Comparison between the appearance of the iris texture acquired in a synchronous way, using multispectral sensors. The upper row gives the iris data in near-infrared (NIR) wavelengths, while the bottom row gives the corresponding data in visible wavelengths (VW). Note the inverse relationship in the NIR and VW data regarding the levels of iris pigmentation and the captured iris minutia.

2 VW Iris Recognition: Summary of Research Works

Tan et al. [80] performed biometric recognition according to both iris and periocular data. Global color-based features and local ordinal measures were used to extract discriminating data from the iris region, later fused to periocular data extracted from texton representations. Finally, fusion is performed by the sum rule using the normalized scores generated for the different types of features. Wang et al. [85] used an adaptive boosting algorithm to build a strong iris classifier learned from a set of bi-dimensional Gabor-based set of features, each corresponding to a specific orientation and scale and operating locally. Later, given the fact that the pupillary boundary is especially difficult to segment in VW data, the authors trained two distinct classifiers: one for irises deemed to be accurately segmented and another for cases in which the pupillary boundary was not accurately segmented. Santos and Hoyle [71] fused a set of recognition techniques that can be divided in two main categories: wavelet-based textural analysis methods applied to the iris region, complemented by distribution-based (histogram of oriented gradients and local binary patterns) and scale invariant feature transforms that analyze the periocular region, which was recently suggested as an important addition for handling degraded samples, essentially because it is less vulnerable to problems resulting from deficient illumination or low-resolution acquisition. Shin et al. [73] started by classifying the left and right eyes by their eyelash distributions, which they used to reduce the search space. Further, they coupled two encoding and matching strategies based in color and tex-
natural analysis to obtain multiple distance scores fused by means of a weighted sum rule, which is claimed to improve the separation between match and non-match distributions. Li et al. [33] used a novel weighted co-occurrence phase histogram to represent local textural features. This method is claimed to model the distribution of both the phase angle of the image gradient and the spatial layout, which overcomes the major weakness of the traditional histogram. A matching strategy based on the Bhattacharyya distance measures the goodness of match between irises. Finally, the authors concluded that the performance is improved when a simple image registration scheme accounts for the image deformation. Marsico et al. [35] proposed the use of implicit equations to approximate both the pupillary and limbic iris boundaries and perform image normalization. Next, they exploited local feature extraction techniques such as linear binary patterns and discriminable textons to extract information from vertical and horizontal bands of the normalized image. Li and Ma [32] introduced an image registration method based on the Lucas-Kanade algorithm to account for iris pattern deformation. Operating on the filtered iris images, this method divides the images into small sub-images and solves the registration problem for each small sub-image. Later, a sequential forward selection method searches for the most distinctive filters from a family of Gabor filters, concluding that a very small number of selected features are able to obtain satisfactory performance. Finally, Szewczyk et al. [77] presented a semi-empirical approach based on a reverse bi-orthogonal dyadic wavelet transform, empirically selecting a compactly supported bi-orthogonal spline wavelet for which symmetry is possible with FIR filters and three vanishing moments. The authors concluded that such a method produces a short biometric signature (324 bits) that can be successfully used for recognition under such challenging conditions, improving its reliability.

Du et al. [19] aimed at robustness and used the SIFT transform and Gabor wavelets to extract iris features, which were used for local feature point description. Then two feature region maps were designed to locally and globally register the feature points, building a set of deformable iris sub-regions that takes into account the pupil dilation/contraction and deformations due to off-angle data acquisition.

3 Data Acquisition: Frameworks and Major Problems

The term constraint refers to one of the factors that currently deployed systems impose, in order to perform recognition with enough confidence: subjects distance, motion and gaze direction and lighting conditions of the environment. These constraints motivate growing research efforts and became the focus of many recent proposals, among which the “Iris-on-the-move” project [46] should be highlighted: it is a major example of engineering an image acquisition system to make the recognition process less intrusive for subjects. The goal is to acquire NIR close-up iris images as a subject walks at normal speed through an access control point. Honeywell Technologies applied for a patent [28] on a very similar system, which was also able to recognize irises at a distance. Previously, Fancourt et al. [20] concluded that
it is possible to acquire sufficiently high-quality images at a distance of up to ten meters. Narayanswamy and Silveira [51] used a wavefront coded optic to deliberately blur images in such a way that they do not change over a large depth-of-field. Removing the blur with digital image processing techniques makes the trade-off between signal-to-noise ratio and depth of field linear. Also, using wavefront coding technology, Smith et al. [75] examined the iris information that could be captured in the NIR and VW spectra, addressing the possibility of using these multispectral data to improve recognition performance. Park and Kim [54] acquired in-focus iris images quickly at a distance, and Boddeti and Kumar [6] suggested extending the depth-of-field of iris imaging frameworks by using correlation filters. He et al. [23] analyzed the role of different NIR wavelengths in determining error rates. More recently, Yoon et al. [90] presented an imaging framework that can acquire NIR iris images at-a-distance of up to three meters, based on a face detection module and on a light-stripe laser device used to point the camera at the proper scene region. Boyce et al. [8] studied the image acquisition wavelength of revealed components of the iris, and identified the important role of iris pigmentation. Although concluding that illumination inside the 700-900 nm optimally reveals the richness of the iris structure, they observed that irises with moderate levels of pigmentation could be imaged in the visible light with good quality.

### 3.1 Proof-of-Concept

This section reports one possible solution for acquiring data of the ocular region from moving subjects in outdoor environments and large distances (between 10 and 40 meters), without requiring subjects’ willingness to be recognised. A prototype was developed, with two cameras mounted on the exterior wall of the SOCIA Lab.: Soft Computing and Image Analysis Lab.1, located in Covilhã, University of Beira Interior, Portugal. Cameras are at a first-floor level (approximately 5m above the ground), and pointing towards a parking lot. A master-slave configuration was adopted, i.e., a wide-view (static) camera (Canon VB-H710F in our prototype) covers the whole scene and provides data for human detection and tracking modules, which enables to point the PTZ camera (Hikvision DS-2DE5286-AEL) to subjects’ faces. Fig. 2 illustrates the environmental conditions in this prototype and the data acquired by both the wide-view and PTZ devices.

In order to automatically obtain information from the subjects faces / ocular regions, the whole processing chain is composed by five modules: 1) at first, the SOBS [45] is used to discriminate between the background / foreground objects in the scene. Next, 2) a human detection algorithm based in the widely known Haar-based Viola and Jones algorithm [83] enables to obtain a set of regions-of-interest (ROI), which feed a 3) object tracking module, based in the KLT algorithm [72] and in the omega-shape of the head and shoulders region as primary source of keypoints.

---

1 [http://socia-lab.di.ubi.pt](http://socia-lab.di.ubi.pt)
This tracker gives as output a set of point lists, each one describing the 2D position of one subject in the scene. Such positions are used by a 4) time-series predictor that estimates the subsequent positions of subjects in the scene, which is where the PTZ should be pointed to. A 5) camera calibration/synchronisation module is capable of accurately estimating the PTZ pan-tilt parameters without depending on additional constraints. Our approach exploits geometric cues to estimate subjects height and avoids depth ambiguity, obtaining the subject's 3D position in the scene. As main result of this processing chain, we get images similar to the ones illustrated in the bottom-left corner of Fig. 2.

In order to establish a baseline comparison between the performance attained by a face and a ocular recognizers, considering two classical algorithms: 1) the face recognition strategy proposed by Turk and Pentland [81], which introduced the concept of *eigenface* that became extremely popular in the computer vision literature; and 2) the periocular recognition algorithm due to Park et al. [55], defining a grid around the iris, from where histograms of oriented gradients, local binary patterns and scale-invariant features are extracted. Data from 25 persons were collected, with subjects from 15 to 30 meters away from the camera. In this case, only samples with relatively frontal pose (yaw $\pm \pi/10$) and *neutral* expression were kept, resulting in a total of 78,960 images pairwise comparisons. Fig. 3 compares the Receiver Operating Characteristic (ROC) curves obtained for the face and ocular recognition experts, being evident the better results of the ocular expert in the low false acceptances (FA) region, in contrast to the high FA region, where the facial recognition expert outperformed. In terms of the Area Under Curve (AUC) values, the ocular expert got 0.857, and the face expert 0.854, which are too close to provide statistically relevant conclusions about the best trait for this kind of environments. Note...
that in this experiment all subjects had neutral facial expression, which otherwise would decay more the recognition performance of face than of the ocular region. Anyway, the main purpose of this experiment was exclusively to obtain a baseline performance that could be substantially improved by using more sophisticated recognition algorithms.

Fig. 3 Comparison between the recognition performance observed for two classical face and ocular recognition algorithms, using data acquired in outdoor environments, under conditions that are currently associated to visual surveillance.

### 3.2 NIR vs. VW Data: Amount of Information Acquired

As illustrated in Fig. 4, four freely available data sets were selected for all experiments reported in this chapter, each one representing a different data acquisition setup / scenario:

- The University of Bath data set\(^2\) 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 for a recognition system. All irises are sharp, without relevant occlusions and in frontal view.
- The CASIA-Iris-Distance set\(^3\) was collected by the CASIA long-range device in a relatively unconstrained setup. Images feature blink, motion blur, off-axis gaze


\(^3\) [http://biometrics.idealtest.org/](http://biometrics.idealtest.org/)
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 [60] 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 heterogeneous 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 [56] 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 images from 868 classes were selected from this data set. All these images were reasonably segmented, according to visual inspection, and were converted to grayscale.

\[ h(I_{p \times p}) = - \sum_i P(I_{p \times p} = i) \log_2 (P(I_{p \times p} = i)), \]  

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

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
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.

resolution), the immediate conclusion is the higher homogeneity of values observed in NIR data than in the VW case. Note that the average values are 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 are due to frequent reflections not detected by the noise-free segmentation phase. Also, we noticed 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 cause shadows in these regions.
4 Iris Segmentation

4.1 Comparison of NIR vs. VW Issues

In order to acquire iris data from large distances and under unconstrained protocols, acceptable depth-of-field values demand high f-numbers for the optical system, corresponding directly (squared) with the amount of light required. Similarly, the motion factor demands very short exposure times, which again increases the amounts of light required. It is known that excessively strong illumination cause permanent eye damage and the NIR wavelength is particularly hazardous, because the eye does not instinctively respond with its natural mechanisms: aversion, blinking, and pupil contraction.

The above points were the major motivations for using visible-light to *in-the-wild* iris biometrics, even though such light spectrum increases the challenges in performing reliable recognition. As stated above, the pigmentation of the human iris enables to capture much higher level of detail in VW than in NIR, but also more noisy artefacts, including specular and diffuse reflections and shadows. In practice, this supports the uniqueness of the iris texture acquired in the visible-light spectrum (in a way similar to the empirically suggested for the near-infrared setup in previous studies [15]), but also stresses the difficulties in obtaining good quality data.

4.2 Why Is It So Difficult?

There are four families of factors that affect the quality of VW iris biometric data not acquired under the classical *stop-and-stare* protocol: A) blur; B) occlusions; C) perspective and D) lighting. By working in a broad range of distances and on moving targets, blurred (A.1) and low-resolution (A.2) images are highly probable. Also, portions of the iris texture are occluded by eyelids (B.1), eyelashes (B.2) and glossy reflections (B.3) from the surrounding environment. Camera-to-subject misalignments may occur, due to varying subjects gaze (C.1) and pose (C.2). Finally, variations in light intensity (D.1), type (D.2) and incident angles (D.3) reinforce the broadly varying features of this kind of data.

Considering that periocular biometrics uses data not only from the iris but also from the surroundings of the eye (e.g., eyelids, eyebrows, eyelashes and skin), particular attention should be paid to additional data degradation factors, such as (E.1) makeup, (E.2) piercings and (E.3) occlusions (e.g., due to glasses or hair).

Fig. 6 illustrates the four families of factors that primarily affect the quality of data that is not acquired under the classical *stop-and-stare* protocol. By working in a broad range of distances and on moving targets, blurred (I.a) and low-resolution images (I.b) are highly probable. Also, portions of the iris texture are occluded by eyelids, eyelashes (II.c) and by glossy reflections from the surrounding environment (II.d). Camera-to-subject misalignments might occur, due to varying subjects gaze...
(III.e) and pose (III.f). Finally, variations in light intensity, type and angle (IV. g and h) reinforce the broadly varying features of the resulting data.

### Fig. 6

Four major types of variability in ocular data acquired in non-constrained setups. The amount of information highly varies, due to optical defocus, motion blur and data resolution (group I). Portions of the iris texture are often occluded by eyelids, eyelashes and reflections (group II) and subjects are misaligned with respect to cameras (group III). Finally, light sources of different type, intensity and 3D angles may exist in the environment (group IV).

### 4.3 Iris Segmentation: Summary of Research Works

Segmentation is undoubtedly perhaps the most concerning phase of the processing chain, in terms of the ability of the whole system to deal with data that is degraded, due to the unconstrained acquisition setup. Also, as it is one of the earliest phases of the recognition process, it is the one that more directly has to deal with data variability and supports the whole process, with any error in segmentation (even small inaccuracies in one of the detected boundaries), easily propagating though the
processing chain and substantially increasing the recognition error rates [62]. Here we briefly summarise some of the most relevant research in the iris segmentation topic, not only covering methods for VW data, but also describing the approaches designed for NIR images, in order to stress the typical differences between both kinds of methods.

In Table 1 we give an overview of the main techniques behind several recently published iris segmentation methods. We compare the methods according to the data sets used in the experiments, categorized by the order in which they segment iris borders. The "Experiments" column contains the iris image databases used in the experiments. "Pre-processing" lists the image preprocessing techniques used before segmentation. Ord. Borders lists the order in which the iris borders are segmented, where $P$ denotes the pupillary borders and $S$ denotes the scleric iris borders ($x \rightarrow y$ denotes the segmentation of $y$ after $x$ and $x, y$ denotes independent segmentation, i.e., when no information from one parameterised border is used in the segmentation of the other). Pupillary Border and Scleric Border columns refer to the main methods used to segment that iris border.

Noting that the significant majority of the methods were designed to work with NIR images. These methods expect to find typically a high contrast between the pupil (almost black) and the iris, which justifies the order in which almost all of these NIR method segment both boundaries ($P \rightarrow S$). In contrast, methods that were particularly designed to handle VW data almost invariably segment the outer iris boundary first, and then use this information to constrain the region where the pupillary boundary is searched, as there is almost no contrast between the pupil and the iris, in case of heavily pigmented irises, imaged with reduced amounts of light. Among the relevant innovations in this topic, techniques such as the use of active contour models, either geodesic ([70]), based on Fourier series ([16]) or based on the snakes model ([3]) can be highlighted. Noting that these techniques require previous detection of the iris to properly initialize contours, they are associated with heavy computational requirements. Modifications to known form fitting methods have also been proposed, essentially to handle off-angle images (e.g., [95] and [82]) and to improve performance (e.g., [44] and [18]). Finally, the detection of non-iris data that occludes portions of the iris ring has motivated the use of parabolic, elliptical and circular models (e.g., [4], and [18]) and the modal analysis of histograms [16]. Even so, in unconstrained conditions, several authors have suggested that the success of their methods is limited to cases of image orthogonality, to the non-existence of significant iris occlusions, or to the appearance of corneal reflections in specific image regions.

5 Image Quality Assessment

The concept of good metric is not trivial to determine, although the best one should maximally correlate with recognition effectiveness. Previous studies reported significant decays in effectiveness when data is degraded by each of the factors listed
Table 1 Summary of the most relevant iris segmentation techniques.

<table>
<thead>
<tr>
<th>Method</th>
<th>Experiments</th>
<th>Preprocessing</th>
<th>Iris Border</th>
<th>Pupilary Border</th>
<th>Scleral Border</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu et al. [91]</td>
<td>CASA.1, ICE, WVU (NIR)</td>
<td>Spectral reflections detection (threshold), PDE and im-</td>
<td>P → S</td>
<td>Randomized Elliptical Hough Transform</td>
<td>Weighted integro-differential operator</td>
</tr>
<tr>
<td>Papadakis et al. [90]</td>
<td>UBIRIS (WV)</td>
<td>Image binarization (threshold), 2-D Fourier spectral loca-</td>
<td>P → S</td>
<td>Binarization (threshold), Circular Hough Transform</td>
<td>Geodesic Active Contours</td>
</tr>
<tr>
<td>Biva and Shah [30]</td>
<td>CASA.1, WVU (NIR), UBIRIS (WV)</td>
<td>Negative image, inpainting</td>
<td>P</td>
<td>Iterative expansion/dilation of the detected border based on morphological operators</td>
<td>-</td>
</tr>
<tr>
<td>Montillo et al. [99]</td>
<td>Non-specified (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Images difference</td>
<td>Images cascade at different scales, Scleral edge detection, elliptical form fitting</td>
</tr>
<tr>
<td>Car [57]</td>
<td>ICE (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Angular constrained Canny edge detection, Elliptical Hough-based transformation</td>
<td>Heights based on fitting, hypothesis and test process</td>
</tr>
<tr>
<td>Srinivas et al. [98]</td>
<td>BATH (NIR)</td>
<td>Integro-differential, image binarization (threshold), morphological operators, elliptical integro-differential operator</td>
<td>P → S</td>
<td>Height detection according to an elliptical model, followed by a modified Mumford-Shah functional</td>
<td>-</td>
</tr>
<tr>
<td>Pronoli and Almeida [58]</td>
<td>UBIRIS (WV)</td>
<td>Integro-differential operator</td>
<td>P → S</td>
<td>Elliptical integro-differential operator, integro-</td>
<td>-</td>
</tr>
<tr>
<td>Zain [91]</td>
<td>CASA.1 (NIR)</td>
<td>Morphological operators to eliminate cyclodesis</td>
<td>S → P</td>
<td>Split and merge process to localize regions of uniform intensity</td>
<td>-</td>
</tr>
<tr>
<td>Broussard et al. [93]</td>
<td>BATH (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Region extraction based on pixel neighborhood variance, Hough-based fitting</td>
<td>Previous method similar to the pupilary border</td>
</tr>
<tr>
<td>He and Shi [22]</td>
<td>ICE, CASIA.1 (NIR), WVU (NIR)</td>
<td>Image binarization, morphological operations</td>
<td>P → S</td>
<td>Region detection based on pixels + neighborhood operations</td>
<td>Previous method similar to the pupilary border</td>
</tr>
<tr>
<td>Bani and Javed [4]</td>
<td>BATH (NIR)</td>
<td>Image binarization, morphological operations</td>
<td>P → S</td>
<td>Iterative bijection-based method</td>
<td>Maximum of the difference of intensities of pupillary detection</td>
</tr>
<tr>
<td>Ovechkin and Yakubov [55]</td>
<td>CASA.1 (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Near circular active contour model (snake), interpretation process to improve performance</td>
<td>Integro-differential operator</td>
</tr>
<tr>
<td>Dongyan [10]</td>
<td>ICE (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Active contours based on Fourier series, modeled with 17 discrete Fourier coefficients</td>
<td>Active contours based on Fourier series, modeled with 4 discrete Fourier coefficients</td>
</tr>
<tr>
<td>Hui et al. [24]</td>
<td>CASA.1 (NIR)</td>
<td>Not described</td>
<td>P → S</td>
<td>Active contours based on Fourier series, modeled with 17 discrete Fourier coefficients</td>
<td>Active contours based on Fourier series, modeled with 4 discrete Fourier coefficients</td>
</tr>
<tr>
<td>Zhang et al. [33]</td>
<td>NTU (NIR)</td>
<td>Conversion into Hough space</td>
<td>P → S</td>
<td>Integral projection functions, median filtering, image segmentation</td>
<td>-</td>
</tr>
<tr>
<td>Lu and Xie [91]</td>
<td>CAS-PREAL (WV)</td>
<td>Not described</td>
<td>P → S</td>
<td>Integral projection functions, median filtering, image segmentation</td>
<td>-</td>
</tr>
<tr>
<td>Honey2Eye International [28]</td>
<td>CASA.1 (NIR)</td>
<td>Not described</td>
<td>P</td>
<td>Image clustering to perform segmentation (threshold), impainting</td>
<td>-</td>
</tr>
<tr>
<td>Meier et al. [15]</td>
<td>AIR, CVU (NIR)</td>
<td>Integro-differential, Canny edge detection</td>
<td>S</td>
<td>-</td>
<td>Canny edge detection, Angular constrained Hough transforms</td>
</tr>
<tr>
<td>Winkler et al. [74]</td>
<td>WVU (NIR)</td>
<td>Reflective specular reflections, elliptical integro-differential operator</td>
<td>P → S</td>
<td>Elliptical integro-differential operator, elliptical integro-differential operator</td>
<td>-</td>
</tr>
<tr>
<td>Tan et al. [73]</td>
<td>UBIRIS v1, UBIRIS v2</td>
<td>Image clustering to perform rough eye localization</td>
<td>P → S</td>
<td>Elliptical integro-differential operator</td>
<td>Elliptical integro-differential operator</td>
</tr>
<tr>
<td>Proenca [61]</td>
<td>UBIRIS v2, FIRE, FROG (WV), ICE / 2008 (NIR)</td>
<td>Scleral Detection</td>
<td>J → P</td>
<td>Local bin, blue luminance, red chroma, neural network classification, constrained polynomial fitting</td>
<td>Local bin, blue luminance, red chroma, neural network classification, constrained polynomial fitting</td>
</tr>
</tbody>
</table>

In Table 2, here we overview the main techniques used to assess iris image quality with respect to each factor and compare them according to the spectrum of light used, the type of analyzed data (raw image, segmented or normalized iris region) and their output (local or global), as they operate at the pixel or image level. We note that most of the methods operate on NIR images and assess quality in the segmented data (either in the cartesian or polar coordinate systems). Exceptions are usually related with focus measurement, obtained by one of two approaches: (1) measuring the high frequency power in the 2D Fourier spectrum through a high-pass convolution kernel or wavelet-based decomposition ([16], [31] and [11]); (2) analyzing...
the sharpness of the iris borders through the magnitude of the first and second order
derivatives ([1] and [92]). Another key characteristic is the level of analysis: some
methods operate globally (at the image level), usually to determine focus, gaze or
motion blur ([31], [38] and [84]). As image quality varies across the iris, others op-
erate at the pixel level to determine local obstructions ([1], [36] and [59]). Motion is
estimated by detecting interlaced raster shear that might be due to significant move-
ments during the acquisition of a frame ([17], [34], [86] and [97]). Other approaches
rely on the response of the convolution between the image and directional filters, be-
ing observed that linear motion blurred images have higher central peak responses
than sharp ones ([36] and [39]). Gaze is estimated by 3D projection techniques that
maximize the response of the Daugman’s integro-differential operator [36] and by
the length of the axes of a bounding ellipse [97]. Eyelids are detected by means of
line and parabolic Hough transforms [25], active contours [41] and machine learning
frameworks [59] [89]. The modal analysis of the intensities histogram enables the
detection of eyelashes [16] [25], as do spectral analysis [34] and edge-based meth-
ods [36]. As they usually are the brightest regions of images, specular reflections are
detected by thresholds [36], while diffuse reflections are exclusive of VW data and
more difficult to discriminate, being reported a method based in texture descriptors
and machine learning techniques [59]. Proença proposed a method [63] to assess the
quality of VW iris samples captured in unconstrained conditions, according to the
factors that are known to determine the quality of iris biometric data: focus, motion,
angle, occlusions, area, pupillary dilation and levels of iris pigmentation. The key
insight is to use the output of the segmentation phase in each assessment, which per-
mits to handle severely degraded samples that are likely to result of such imaging
setup.

6 Feature Encoding

Feature encoding is a particularly interesting sub-topic in the unconstrained recog-
nition domain, due to the reduced quality of the data that is expected to be acquired.
Here, a fundamental property of the iris texture should be considered, being one of
the major reasons that justify the interest on this trait for this kind of scenarios: most
of the discriminating information between the iris texture of different subjects lies
in the lowest and middle-low frequency components, which are (luckily) those that
are most easy to capture under outdoor environments and unconstrained acquisition
protocols.

A particularly interesting advance is the use of Multi-Lobe Differential Filters,
which are claimed to adapt better than the traditionally used Gabor filters to data
of reduced quality and can be used at reduced computational cost. On the other
way, they lie in a parameterisation space of much higher dimension than the one of
Gabor filters, making more difficult to obtain good parameterisations for a specific
recognition system / environment.
Table 2 Overview of the most relevant methods published to assess the quality of iris biometric data.

<table>
<thead>
<tr>
<th>Method</th>
<th>Experiments Data Sets</th>
<th>Images</th>
<th>Analysis</th>
<th>Quality Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahn et al. [24]</td>
<td>CASIA v3.0, WVU (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis), texture (wavelet-based descriptors)</td>
</tr>
<tr>
<td>Chen et al. [25]</td>
<td>CASIA v3.0, WVU (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis), texture (wavelet-based descriptors)</td>
</tr>
<tr>
<td>Huang et al. [26]</td>
<td>ICE-1 (NIR)</td>
<td>Raw data</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Huang et al. [27]</td>
<td>ICE-2 (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Shen et al. [28]</td>
<td>ICE-3 (NIR)</td>
<td>Segment and normalized</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>He et al. [29]</td>
<td>CASIA v2.0, WVU (NIR)</td>
<td>Segment</td>
<td>Local</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Hollingsworth et al. [30]</td>
<td>Univ. Notre Dame (NIR)</td>
<td>Segment</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Yang et al. [31]</td>
<td>WVU (NIR), UBIRIS v1 (VW) (NIR)</td>
<td>Raw data</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Kalka et al. [32]</td>
<td>CASIA v3, WVU, ICE (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Kang and Park [33]</td>
<td>CASIA v3, WVU (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Kang and Park [34]</td>
<td>CASIA v3 (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Krohn et al. [35]</td>
<td>WVU (NIR)</td>
<td>Segment and normalized</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Narukuma et al. [36]</td>
<td>UBIRIS v1 (NIR)</td>
<td>Segment and normalized</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Lee et al. [37]</td>
<td>CASIA v3 (NIR)</td>
<td>Segment and normalized</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Piramo and Alexandrov [38]</td>
<td>UBIRIS v1 (VW)</td>
<td>Segment and normalized</td>
<td>Local</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Price et al. [39]</td>
<td>UBIRIS v1 (VW)</td>
<td>Segment and normalized</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>War et al. [40]</td>
<td>StFV-10 (NIR)</td>
<td>Raw data</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Wei et al. [41]</td>
<td>CASIA v2 (NIR), WVU v1 (VW)</td>
<td>Raw data and segment</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Yi et al. [42]</td>
<td>CASIA, CASIA v2 (NIR)</td>
<td>Raw data</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Zhang and Segald-cell [43]</td>
<td>CASIA v2, WVU (NIR)</td>
<td>Segment</td>
<td>Global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
<tr>
<td>Zuo and Schmid [44]</td>
<td>ICE-1 (NIR)</td>
<td>Segment</td>
<td>Local, global</td>
<td>Feature extraction (frequency domain analysis)</td>
</tr>
</tbody>
</table>

6.1 Gabor vs. Multi-Lobe Differential Filters

The discriminating power provided by each region of VW and NIR iris images was assessed, with respect to two families of filters: 1) Gabor kernels, which faithfully model simple cells in the visual cortex of mammalian brains [13] 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 [76].
The impulse response of a Gabor kernel is defined by the multiplication of a harmonic and a Gaussian function:

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

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., [15]), allowing for different scales along the axes (\( \sigma_x \) and \( \sigma_y \)). In this chapter, to keep moderate the dimension of the parameterisation space, only filters with the same scale along the axes are considered.

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 are considered. Hence, the MLDF filters are expressed by:

\[
m[x_j, \mu_j, \sigma_j] = \sum_{j=1}^{t_l} \frac{(-1)^{j+1}}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(x_j - \mu_j)^2}{2\sigma_j}\right],
\]

where \( x_j = (x_j, y_j) \) is the centre of each of the \( t_l \) lobes. Next, \( k = \{m, g\} \) filters were convolved with each normalized iris image \( I \), providing a set of coefficients. The sign of the coefficients was obtained, i.e., \( C \) is the vector representation of \( \text{sgn}(I \ast k) \).

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 \( \sigma : [1, 2, 3, 4, 5, 6] \).

**Table 3** Types and range of the filters parameters varied in our experiments.

<table>
<thead>
<tr>
<th>Gabor Filters ( g )</th>
<th>MLDF Filters ( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (px.)</td>
<td>( \omega : [1 : 1 : 14] )</td>
</tr>
<tr>
<td>Orientation</td>
<td>( \varphi : [0, \pi/4, \pi/2, 3\pi/4] )</td>
</tr>
<tr>
<td>Gaussian Sigma</td>
<td>( \sigma : 0.65 \omega )</td>
</tr>
<tr>
<td>Num. Lobes</td>
<td>( n : [1/1, 2/2, 3/3, 4/4] )</td>
</tr>
<tr>
<td>Gaussian Sigma</td>
<td>( \sigma : [1, 2, 3, 4, 5, 6] )</td>
</tr>
</tbody>
</table>

Fig. 7 illustrates the filters used and Table 3 summarizes the range of parameters considered, with \( \{a : b : c\} \) denoting values in the \([a, c]\) interval, with steps of size \( b \).
Fig. 7 Illustration of the filters used in experiments (\{g, m\}) and of the filters that give the contribution of each position in the iris to the coefficient in the iris code \{|g|, |m|\}.

Fig. 8 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 artefacts. 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 are not so evident as in the case of Gabor kernels.

### 6.2 Analysis of Iris Codes: Comparison Between NIR and VW Data

The discriminability \(\tau\) of each bit extracted from NIR and VW images 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...
Fig. 8 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).

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:

\[
\Psi_{[x, y]} = \frac{\sum \left( |k_i[x - r_i, y - c_i]| \cdot \tau(i) \right)}{\sum |k_i[x - r_i, y - c_i]|},
\]

where \([r_i, c_i]\) is the central position of the \(i^{th}\) filter \(k_i\) and \(\tau(i)\) is the discriminability of the \(i^{th}\) bit, given by:

\[
\tau(i) = P(C_i^{(p)} \oplus C_i^{(q)} = 0 | H_a) - P(C_i^{(p)} \oplus C_i^{(q)} = 0 | H_0),
\]

with \(P(C_i^{(p)} \oplus C_i^{(q)} = 0 | H_a)\) expressing the probability that the the \(i^{th}\) bit of an iris code is equal in two inter-subject samples, and \(P(C_i^{(p)} \oplus C_i^{(q)} = 0 | H_0)\) expressing the same probability for intra-subject samples.

Fig. 9 gives the discriminability provided by each region of the iris in the Cartesian and polar coordinate systems, when using Gabor filters. Complementary, Fig. 10 expresses the similar statistics when using MLDF filters. The immediate
Fig. 9 Average bit discriminability $\Psi[x,y]$ across the iris, using Gabor filters as feature encoders. Values are given for the Cartesian and polar coordinate systems, for the four data sets considered: BATH and CASIA-Iris-Distance (NIR) and UBIRIS.v2 and FRGC (VW).

The conclusion is that the maximal values are observed for the NIR data sets, both for Gabor and MLDF filters. Interestingly, in all cases the lower parts of the iris are better than the upper parts, which are 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. 8, large filters tend to produce more discriminant bits, which accords with the results given in [27].

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 observed levels of linear correlation between variables $h[x,y]$ and $\Psi[x,y]$ are -0.12/-0.38 (Gabor/MLDF filters), and -0.40/-0.22 for the CASIA-Iris-Distance set. Regarding the VW data, values are 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.

In summary, MLDFs appear to 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.

Fig. 10 Average bit discriminability $\Psi[x,y]$ across the iris, using Multi-Lobe Differential Filters as feature encoders. Values are given for the Cartesian and polar coordinate systems, for the four data sets considered: BATH and CASIA-Iris-Distance (NIR) and UBIRIS.v2 and FRGC (VW).
6.3 Codes Quantization: Is Too Much Information Lost?

In the most acknowledged iris recognition algorithm, only phase information is used in recognition. Amplitude information is not considered reliable, as it depends on imaging contrast, illumination and camera gain. Accordingly, Hollingsworth et al. [27] 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.

\[
\begin{align*}
\text{A)} & \quad C^* = \text{sgn}(C) \\
\text{B)} & \quad C^* = \frac{1+\text{erf}(\alpha C)}{2} \\
\text{C)} & \quad C^* = \frac{1}{2} + \alpha C
\end{align*}
\]

Fig. 11 Three different strategies for code quantization: A) binary; B) sigmoid function; and C) linear mapping.

Even considering the above 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 are 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. 12 compare the recognition performance with respect to each quantization strategy and Table 4 summarizes the results, giving the Area Under Curve (AUC) and the decidability index \( d' \) that, as suggested by Daugman [14], measures how well separated the genuine / impostor distributions are:

\[
d' = \frac{|\mu_G - \mu_I|}{\sqrt{\frac{1}{2}(\sigma_I^2 + \sigma_G^2)}},
\] (6)
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 can be drawn: for Gabor filters, the best results are observed when using the traditional sign() quantization function. In this case, using scalars instead of sign bits even decreased the recognition performance. Oppositely, for MLDF filters, the best results are observed when using the proposed sigmoid function, i.e., when the coefficients of small magnitude are also considered for the matching process. This points toward 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 result from the reported AUC and $d'$ values, which in the large majority of the cases are in agreement. The exceptions occur mostly in cases where the shape of the genuine / impostor distributions are the farthest from Gaussian distributions. For these particular cases, we relied mostly on the AUC value, as it does not require a specific data distribution to report meaningful results.
As an attempt to increase the robustness of iris recognition in visible-light data, the concept of periorcular biometrics has emerged, which compensates for the degradation in iris data by considering the discriminating information in the surroundings of the eye (eyelids, eyelashes, eyebrows and skin texture). Currently, the most relevant algorithms work in a holistic way: they define a region-of-interest (ROI) around the eye and use a feature encoding / matching algorithm regardless of the biological component in each point of the ROI. However, this augments the probability of sensitivity to some data covariate and the correlation between the scores extracted from the different points in the ROI.

### 7.1 Weak / Strong Ocular Experts

Under an atomistic criterion, two experts that use disjoint data can be devised, with radically different recognition strategies and attaining very different effectiveness. Here, the term *weak* is employed to refer to a recognition system that yields a poor separable decision environment, i.e., where the distributions of the genuine / impostor pairwise scores largely overlap. The term *strong* refers to a system where the distributions of genuine and impostor scores almost don’t overlap, resulting in a clearly separable decision environment and low error rates.

In this dual ensemble, the strong expert analyses the multi-spectral information in the iris texture, according to an automatically optimised set of multi-lobe differential filters (MLDF). Complementary, the weak expert parameterises the boundary of the visible cornea and defines a dimensionless ROI that comprises the eyelids, eyelashes and the surrounding skin. This expert helps to discriminate between individuals and has three interesting properties: 1) it analyses data that has an appearance independent of the iris texture; 2) it shows reduced sensitivity to the most problematic iris image covariates; and 3) it exclusively analyses traits that cannot be easily forged by anyone not willing to be recognised, which is in contrast to the traits classically

### Table 4 Variations in recognition performance with respect to different strategies for code quantization.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Features</th>
<th>A) sign()</th>
<th>B) sigmoid()</th>
<th>C) linear (no quantization)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>d’</td>
<td>AUC</td>
<td>d’</td>
</tr>
<tr>
<td>BATH</td>
<td>Gabor</td>
<td>8.79 ± 0.01</td>
<td>0.994 ± 0.001</td>
<td>7.08 ± 0.01</td>
</tr>
<tr>
<td>BATH</td>
<td>MLDF</td>
<td>9.15 ± 0.01</td>
<td>0.994 ± 0.001</td>
<td>8.82 ± 0.01</td>
</tr>
<tr>
<td>CASIA-Iris-Distance</td>
<td>Gabor</td>
<td>3.20 ± 0.01</td>
<td>0.982 ± 0.001</td>
<td>3.16 ± 0.01</td>
</tr>
<tr>
<td>CASIA-Iris-Distance</td>
<td>MLDF</td>
<td>3.89 ± 0.01</td>
<td>0.990 ± 0.001</td>
<td>4.12 ± 0.01</td>
</tr>
<tr>
<td>UBIRIS.v2</td>
<td>Gabor</td>
<td>1.23 ± 0.01</td>
<td>0.813 ± 0.006</td>
<td>1.16 ± 0.02</td>
</tr>
<tr>
<td>UBIRIS.v2</td>
<td>MLDF</td>
<td>1.88 ± 0.01</td>
<td>0.904 ± 0.003</td>
<td>1.96 ± 0.01</td>
</tr>
<tr>
<td>FRGC</td>
<td>Gabor</td>
<td>1.12 ± 0.02</td>
<td>0.792 ± 0.006</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>FRGC</td>
<td>MLDF</td>
<td>1.74 ± 0.01</td>
<td>0.892 ± 0.006</td>
<td>1.88 ± 0.02</td>
</tr>
</tbody>
</table>
used in periocular recognition (e.g., the shape of eyebrows). We encode the shape of eyelids, the distribution and shape of the eyelashes and the morphology of the skin wrinkles / furrows in the eyelids, which are determined by the movements of the orbicularis oculi muscles family. Fig. 13 overviews the information sources used in such recognition ensemble.

![Discriminating Information](image)

Fig. 13 Overview of the components in the vicinity of the human eye that can be used to extract discriminating information, useful for biometric recognition purposes.

It is evident that using multiple sources for biometric recognition is not a new idea, and some controversy remains: is it actually an effective way to improve performance? It is argued that when a stronger and a weaker expert are combined, the resulting decision environment is averaged and the performance will be somewhere between that of the two experts considered individually [14]. Due to the way such a strong / weak ensemble was designed, our experiments support a radically different conclusion: even when the fused responses come from experts with very distant performance, the ensemble attains much better performance than the stronger expert (iris). This is due to the fact that both experts produce quasi-independent responses and are not particularly sensitive to the same image covariate, augmenting the robustness against degraded data.

### 7.2 Relevant Ocular Recognition Algorithms

Concluded in 2011, the *NICE: Noisy Iris Challenge Evaluation* [64] promoted the research about iris / ocular recognition in visible-light data. It received over one hundred participations and the best performing teams described their approaches
in two special issues of the *Image and Vision Computing*\(^4\) and *Pattern Recognition Letters*\(^5\) journals. This event has documented the state-of-the-art recognition performance, having the best algorithm achieved d-prime values above 2.57, area under curve around 0.95 and equal error rates of 0.12. This method (due to Tan et al. [80]) is actually a periocular recognition algorithm: texton histograms and semantic rules encode information from the surroundings of the eye, while ordinal measures and color histograms analyse the iris. The second best approach was due to Wang et al. [85] and is quite more classical: it employs an AdaBoost feature selection scheme from a large set of quantized Gabor-based features, matched by the Hamming distance.

The most relevant recognition algorithms for VW images can be divided with respect to their data source: 1) the iris; or 2) the periocular region. Regarding the first family, Raffei et al. [67] preprocessed the iris to remove reflections and represented the normalised data at multiple scales, according to the Radon transform. The score from each scale was matched by the Hamming distance and fused by weighted non-linear combination. Rahulkar and Holambe [68] derived a wavelet basis for compact representation of the iris texture (triplet half-band filters), with coefficients matched by the minimum Cambera distance. A post-classifier outputs a match when more than \(k\) regions give a positive response. Roy et al. [69] used a feature selection technique from game theory, based on coefficients from the Daubechies wavelet decomposition. The Hausdorff distance yields the matching score between two feature sets. Kumar and Chan [43] approached the problem from the data representation perspective, having used a quaternionic sparse coding scheme solved by convex optimisation. Quaternion image patches were extracted from the RGB channels and the basis pursuit algorithm used to find the quaternion coefficients. In another work [42], the same authors were based in the sparse representation for classification algorithm, using the output of a local Radon transform as feature space.

The second family of algorithms considers other data beside the iris (sclera, eyebrows and skin texture), and its popularity has been increasing since the work of Park et al. [55]. Bharadwaj et al. [5] fused a global descriptor (GIST) based on five perceptual dimensions (image naturalness, openness, roughness, expansion and ruggedness) to circular local binary patterns. The Chi-squared distance matched both types of features and a fusion scheme (score level) yielded the final matching value.

Crihalmeanu and Ross [12] used the sclera patterns as biometric trait. The sclera was segmented according to the pixel-wise proportion between the NIR and green channel values. After enhancing the blood vessels by a line filter, SURF, minutiae and correlation-based schemes produced the matching scores that were fused subsequently. Similarly, Zhou et al. [94] enhanced the blood vessels in the sclera by Gabor kernels and encoded features by line descriptors. The accumulated registration distance between pairs of line segments yielded the matching score. Also, Oh and Toh [52] encoded the information in the sclera by local binary patterns (LBP) in

\(^4\) http://www.sciencedirect.com/science/journal/02628856/28/2  
\(^5\) http://www.sciencedirect.com/science/journal/01678655/33/8
angular grids, concatenated in a single feature vector. Then, a normalised Hamming distance produced the matching score.

In terms of hybrid approaches, Oh et al. [53] combined the sclera to periocular features. Directional features from the former region were extracted by structured random projections, complemented by binary features from the sclera. Tan and Kumar [79] fused iris information (encoded by Log-Gabor filters) to an over-complete representation of the periocular region (LBP, GIST, Histogram of Oriented Gradients and Leung-Malik Filters). Both representations were matched independently and fused at the score level.

**Table 5** State-of-the-art algorithms for recognising degraded ocular data acquired in visible light environments.

<table>
<thead>
<tr>
<th>Method</th>
<th>Traits</th>
<th>Feat. Encoding</th>
<th>Feat. Matching</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bharadwaj et al. [5]</td>
<td>Periocular (Holistic)</td>
<td>GIST, CLBP</td>
<td>Chi-square distance</td>
<td>75% rank-1 (UBIRIS.v2)</td>
</tr>
<tr>
<td>Crihalmeanu and Ross [12]</td>
<td>Sclera</td>
<td>SURF, Minutiae (vessel bifurcations)</td>
<td>Euclidean distance, data correlation</td>
<td>EER &lt; 1.8% (Own dataset)</td>
</tr>
<tr>
<td>Kumar and Chan [43]</td>
<td>Iris</td>
<td>Quaternion Sparse Orientation Code</td>
<td>Shift Alignment</td>
<td>48% rank-1 (UBIRIS.v2)</td>
</tr>
<tr>
<td>Kumar et al. [42]</td>
<td>Iris</td>
<td>Radon local transform</td>
<td>Sparse Representation for Classification</td>
<td>40% rank-1 (UBIRIS.v2), 33% rank-1 (FRGC)</td>
</tr>
<tr>
<td>Oh and Toh [52]</td>
<td>Sclera</td>
<td>LBP</td>
<td>Hamming distance</td>
<td>EER 0.47% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Oh et al. [53]</td>
<td>Periocular (Holistic), Sclera</td>
<td>Multi resolution LBP (Sclera), Directional Projections (Periocular)</td>
<td>Hamming and Euclidean distance</td>
<td>EER 5% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Proença [65]</td>
<td>Periocular (Piecewise)</td>
<td>MLDF (iris), shape and texture descriptors (eyelashes, eyelids)</td>
<td>Modified Hamming distance, Histogram distance (eyelids, eyelashes)</td>
<td>EER 2.97% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Raffei et al. [67]</td>
<td>Iris</td>
<td>Multi-scale local Radon transform</td>
<td>Hamming distance, weighted non-linear score combination</td>
<td>AUC 88% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Rahulkar and Holmabe [68]</td>
<td>Iris</td>
<td>Triplet half-band filter bank</td>
<td>Canberra distance, k-out-of-n post classifier</td>
<td>Acc &gt; 99% (UBIRIS.v1)</td>
</tr>
<tr>
<td>Roy et al. [69]</td>
<td>Iris</td>
<td>Daubechies wavelet, Modified Contribution feature selection</td>
<td>Hausdorff distance</td>
<td>TPR 97.43% @ 0.001%FPR (UBIRIS.v1)</td>
</tr>
<tr>
<td>Tan and Kumar [79]</td>
<td>Iris, Periocular (Holistic)</td>
<td>Log-Gabor filters (Iris), SIFT, GIST, LBP, HOG and LMF (Periocular)</td>
<td>Chi-square and Euclidean distances</td>
<td>39.4% rank-1 (UBIRIS.v2)</td>
</tr>
<tr>
<td>Tan et al. [80]</td>
<td>Iris, Eye</td>
<td>Texton Histograms, Semantic information (Eye), Ordinal Filters, Color Histogram (Iris)</td>
<td>Chi-square, Euclidean, Diffusion and Hamming distances</td>
<td>AUC 95% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Wang et al. [85]</td>
<td>Iris</td>
<td>Gabor filters, AdaBoost feature selection</td>
<td>Hamming distance</td>
<td>AUC 88% (UBIRIS.v2)</td>
</tr>
<tr>
<td>Zhou et al. [94]</td>
<td>Sclera</td>
<td>Line (sclera vessels) description</td>
<td>Accumulated line registration cost</td>
<td>EER &gt; 3.85% (UBIRIS.v2)</td>
</tr>
</tbody>
</table>
Table 5 overviews the state-of-the-art algorithms in terms of biometric recognition from VW ocular data. It compares the analysed traits and summarises the techniques used in feature encoding and matching. The error rates reported by authors are also given (Performance column). However, note the above listed algorithms might have used different experimental protocols and data subsets, which turns the direct comparison of the error rates unfair.

Proença recently proposed a recognition ensemble [65] composed by two experts. The strong expert analyses the multi-spectral information in the iris texture, according to an automatically optimised set of multi-lobe differential filters (MLDF). Complementary, the weak expert parameterises the boundary of the visible cornea and defines a dimensionless ROI that comprises the eyelids, eyelashes and the surrounding skin. This expert helps to discriminate between individuals and has three interesting properties: 1) it analyses data that has an appearance independent of the iris texture; 2) it shows reduced sensitivity to the most problematic iris image covariates; and 3) it exclusively analyses traits that cannot be easily forged by anyone not willing to be recognised, which is in contrast to the traits classically used in periocular recognition (e.g., the shape of eyebrows).

The weak expert encodes the shape of eyelids, the distribution and shape of the eyelashes and the morphology of the skin wrinkles / furrows in the eyelids, which are determined by the movements of the orbicularis oculi muscles family. With respect to related works, the main advantage of this method is that the responses of the iris (strong) and ocular (weak) experts are practically independent, in result of the disjoint regions analyzed and in the fully disparate algorithms used in feature encoding / matching. This way, even by using relatively simple fusion techniques that work at the score level, it is possible to use the weak biometric expert as a valuable complement of the iris expert, particularly in cases where this expert produces matching scores that are near the borderline accept / reject regions. This kind of complementarity between experts is illustrated in Fig. 14, showing pairwise comparisons that are intra-subject (upper row) and inter-subject (bottom row), with $P_s$, $P_w$ denoting the posterior probability of acceptance (by the strong $s$ and weak $w$ experts) of the null hypothesis $H_0$ that both images are from the same subject.

![Fig. 14 Examples of image pairwise comparisons that fall in the uncertainty region of the strong (iris) biometric expert ($P_s(H_0|x) ≈ 0.5$). In most cases, the weak (periocular) biometric expert provides valuable information $P_w$ to distinguish between intra-subject (green frames) and inter-subject comparisons (red frames).](image-url)
8 Fusion of Multiple Recognition Systems

Considering that the type of biometric recognition systems discussed in this chapter should work covertly, meaning that no conscious human effort will be required of subjects during the recognition processes, there is a theoretically interesting possibility of using multiple recognition systems regularly spaced across an airport terminal hallway or a city street. This section reports the (optimistic) performance that such a recognition ensemble would attain, considering as baseline recognizers the current state-of-the-art solutions for non-cooperative ocular recognition.

It is known that not all subjects perform consistently in terms of false matches and non-matches of a biometric system. Based on their intrinsic features, some are difficult to match (goats), while others are particularly vulnerable to impersonation (lambs) and consistently increase the probability of false matches [88]. We oversimplify the problem and regard all subjects of a population $P = \{s_1, \ldots, s_n\}$ as sheep, i.e., subjects that tend to follow the system averages: they match relatively well against themselves and poorly against others. Let us consider $k$ ocular recognition systems with roughly similar performance, with a sensitivity of $\alpha$ at a false match rate of $\beta$. Here we introduce the concept of exogenous independence, hypothesizing that purposely changing the lighting conditions in the environment (by using different levels of light or types of illuminants) and the acquisition protocols (poses, distances) should potentiate the independence between the system outputs. Assuming that the independence of each system provides an upper bound on the performance that would be attainable by the fusion of multiple systems, the binomial distribution can be used to obtain the probability that a subject $s_i$ is screened by $k$ recognition systems and correctly recognized by $k'$ of these, $1 \leq k' \leq k$:

$$P(R_{k'}) = \frac{k!}{k'!(k-k')!} \alpha^{k'} (1-\alpha)^{k-k'},$$

(7)

For different values of $k'$, the probability that a reported match is false is given by $\beta^{k'}$, assuming that false matches in each of the $k$ recognition systems are independent events. Accordingly, a match will be reported iff a minimum of $k'$ recognition systems output a match:

$$P(R \geq k') = \sum_{j=k'}^{k} P(R_j)$$

$$= \sum_{j=k'}^{k} \frac{k!}{j!(k-j)!} \alpha^j (1-\alpha)^{k-j},$$

(8)

provided that all events are mutually exclusive. Considering the average performance for a baseline recognizer that fuses at the score level the responses given by four state-of-the-art algorithms ([80], [85], [71] and [73]), Fig. 15 gives the expected sensitivity of a multipoint biometric system, with respect to the number of baseline recognizers used, considering different false match rates. However, note...
that this analysis provides an upper bound estimate of the ensemble performance, as it assumes that the responses given by individual experts are independent. Even though, this optimistic assumption would enable to conclude that around five independent recognition systems would be enough to attain almost full sensitivity at a false acceptance rate $\beta$ of 0.01. This value substantially increases when a lower number of false alarms is convenient (large scale applications), requiring between thirteen and twenty three independent recognition systems to operate, respectively, at $\text{FAR} \approx 1e^{-4}$ and $1e^{-6}$.

![Graph showing expected sensitivity of an ensemble of ocular recognition systems working covertly and consecutively](image)

**Fig. 15** Expected sensitivity of an ensemble of ocular recognition systems working covertly and consecutively (e.g., in an airport terminal hallway or a city street), with different required values for the false acceptance rates. Note that this is an optimistic estimate of the ensemble performance, as it assumes that the responses given by baseline recognizers are statistically independent.

### 9 Conclusions and Current Challenges

There is no doubt that concerns about the security and safety of crowded urban areas have been increasing significantly, particularly due to terrorist attacks such as the 2001 New York 9/11, the 2004 Madrid train bombing and the 2013 Boston marathon attacks. These concerns raised the interests on biometrics and made it one the most popular topics in the Pattern Recognition / Computer Vision domains. However, there are still not biometric recognition systems that work effectively using data acquired in totally uncontrolled environments and without assuming subjects’ willingness to be recognized.
This chapter discussed such extremely ambitious kind of biometric recognition and advocated the use of the ocular region as basis trait, due to several reasons: being a naturally protected internal organ visible from the exterior, it has a near circular and planar shape that turns easier its segmentation and parameterization. Also, its texture has a predominantly phenotypic or chaotic appearance that is stable over lifetime, which - particularly important - discriminating information between subjects lies in the lowest and middle-low frequency components, i.e., those that are easier to capture in unconstrained data acquisition protocols. Finally, the ocular region is less sensitive to facial expressions (than the whole face), and has a relatively small probability of being occluded due to hair, facial hair and clothing.

We started by summarising the most relevant research works devoted to increasing the recognition robustness with respect to data of reduced quality, and hereinafter, focused particularly in the major issues behind each of the phases that compose the recognition chain: data acquisition, segmentation, feature encoding and matching. Also, we summarized some of the most relevant works in the periocular recognition domain. In this topic, we stressed two key properties of an ocular recognition ensemble: 1) the weak (periocular) recognizer should provide as much independent scores (responses) as possible with respect to the strong (iris) recognizer; and 2) experts should not share particular sensitivity to the same data covariates, in order to actually improve recognition robustness.

Finally, the obstacles remaining in every phase of a fully non-cooperative ocular recognition system were discussed, particularly the difficulty in real-time detecting and segmenting all the components in the ocular region, which is important not only for developing non-holistic feature encoding / matching strategies, but also to estimate pose and data quality.

Acknowledgements The financial support given by "IT: Instituto de Telecomunicações" in the scope of the UID/EEA/50008/2013 project is acknowledged.

References

47. Mansfield T., Kelly G., Chandler D., Kane J.: Biometric product testing final report. issue 1.0 (2001)