Periocular Biometrics: An Emerging Technology for Unconstrained Scenarios

Gil Santos and Hugo Proença
IT - Instituto de Telecomunicações
Universidade da Beira Interior
Covilhã, Portugal
Email: gmelfe@ubi.pt; hugomcp@di.ubi.pt

***Abstract*—The periocular region has recently emerged as a promising trait for unconstrained biometric recognition, specially on cases where neither the iris and a full facial picture can be obtained. Previous studies concluded that the regions in the vicinity of the human eye - the periocular region- have surprisingly high discriminating ability between individuals, are relatively permanent and easily acquired at large distances. Hence, growing attention has been paid to periocular recognition methods, on the performance levels they are able to achieve, and on the correlation of the responses given by other. This work overviews the most relevant research works in the scope of periocular recognition: summarizes the developed methods, and enumerates the current issues, providing a comparative overview. For contextualization, a brief overview of the biometric field is also given.**

I. Introduction

Due to increasing concerns on security and safety of modern societies, biometrics has emerged in the last decade as a major domain of knowledge and has been motivating significant research efforts. Considering the outstanding levels of performance that currently deployed biometric systems achieve, the interest now in putted in the development of systems able to work in uncontrolled acquisition environments, which significantly increases the challenges on reliable recognition. In this setup, alternatives are sought [1] by improving the existing algorithms, by using multi-modal systems or exploring new traits. Despite a broad variety of traits that has been researched, the classical traits to perform *at-a-distance* recognition are the face and the iris.

The face is the most widely used biometric trait. Everyday and even without noticing it, we all use facial information to recognize each other. Not only that, it become one of the most successful applications of image analysis and understanding. Being non-intrusive and allowing cover acquisition, it became preferable over very reliable traits like the iris or fingerprint when aiming at less constrained subject recognition. Several commercial face recognition systems are now available, and a lot of techniques were developed [2] for both still images and video. Face recognition approaches are either based on a global analysis of the whole region as a set of pixel intensities, or the relation between facial attributes, their location and shape.

The iris texture has a predominantly randotypic morphogenesis unique for each individual and allows very high recognition accuracy, which justifies the efforts being held on iris biometrics research [3] and its quick ascent as one of the most popular biometric traits. While most of the commercially deployed iris recognition systems work with constrained near-infrared (NIR) data that favors perception of its patterns whilst reducing the number of noise factors associated, literature on extending this biometrics usability to ”relaxed” visible wavelength (VW) setups has broaden [4]–[6]. However, iris performance as a biometric trait is severely impacted in non-ideal setups, and its relatively reduced size and moving profile make it difficult to image at-a-distance and without user cooperation.

The periocular region represents a trade-off between the whole face and the iris alone. Containing the eye and its immediate vicinity, it covers eyelids and eyelashes, nearby skin area and eyebrows. Its use as a biometric trait has emerged, constituting nowadays a strong alternative for less constrained environments, when image acquisition is not reliable, and to avoid spoofing of the iris patterns. It is easy to acquire without user cooperation and does not require a constrained close capturing, Also, this region is not so affected by the aging process as other facial regions are, as for instance the mouth and cheek whose skin become loosened over time.

Periocular biometrics can be used alone or complementary to iris recognition, considering that the use of multiple traits might be specially important to compensate for the adversity of the environments and uncooperative subjects. Mot times, the responses of periocular methods are fused at the score level to the corresponding iris scores, due to their spatial proximity and to the fact that a single camera might be able to acquire both traits. Being relatively stable and rarely occluded, it's particularly useful when the subject is wearing a mask or otherwise only exposing their eyes.

In terms of features of the periocular region, they can be divided into two levels, as suggested by [7]: the first level comprise the eyelids, eye folds, and eye corners; and the second level comprises the skin texture, wrinkles, color and pores. Analysis of those features can be carried on based on their geometry, texture or color.

As described by Park et al. [8], the problems that arise from periocular recognition can be summarized as follows:
Imaging: What would be the optimal spectrum band for periocular biometrics? Is VW, more advantageous on covert

biometrics, fit for this trait?
Region definition What are the actual "boundaries" of the periocular region? Should iris, sclera or the eyebrows be included or masked/cropped?
Encoding Which features would better describe and discriminate this region? How reliable would they be when relaxing imaging conditions?
Matching What's the best matching scheme for those features? Will coarse classification be of any use?
Fusion What would be the benefit on fusing periocular with other traits? Which ones, and how to fuse them?

The remainder of this paper is organized as follows: Sections II and III overview the recognition systems and existing datasets; Section IV comparatively details the relevant methods developed on periocular recognition; and finally Section V present some final considerations.

II. Biometric System

The importance of the biometric authentication system must not be disregarded, as it will be the responsible for carrying the whole process, from the data acquisition, to feature extraction, and matching against the database. Therefore, designing a system that adapts to its application scenario is most important. In a general way, a recognition system is composed of four modules [9]:

1) Sensor Module: A wide variety of sensors are available, depending on which biometric trait we are going to work with. Since most of the biometric traits consist on visual data, cameras will be used for acquisition. On real-time systems, the balance between the richness on detail of the acquired data and the acquisition rate is essential, and therefore choosing a proper camera also is. This module is strictly related with the first step of recognition systems (trait acquisition) and is where the trade-off between the quality of the gathered data and user cooperation is set.

2) Quality assessment and Feature extraction: Even with an optimal sensor setup, not always the acquired data is suited for feature extraction. Therefore, its quality is usually assessed, and the image discarded if no minimum requirements are met, thus saving time in additional processing. The trait needs to be properly located and segmented (specially useful to gather preferably "good" data), and then encoded as feature templates.

3) Matching and decision-making: In this module, features are matched against the templates in the database, thus deciding either to be in the presence of a genuine or impostor comparison.

4) System database: This module consists on the repository of user biometrics and other identification information, which is acquired during the enrollment stage, and used for later identification or verification of users' identity.

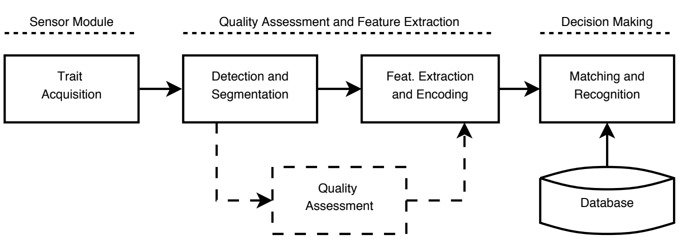

Fig. 1. General steps and elements of biometric recognition systems.

III. Datasets

Only a few public datasets were designed for the development of periocular recognition methods. Instead, face and iris databases are generally used for that purpose. The most commonly used databases for the evaluation periocular methods are now introduced[1], and their specifications summarized at Table I.

A. FERET

The Facial Recognition Technology (FERET) database [10] was designed as a standard for developing face recognition methods, and acquired at George Mason University over 11 sessions and a three years period (1993 to 1996). Initially released as low resolution (256×384 pixels) grayscale data, years later a high-resolution color version was also disclosed. A total of 14051 images were gathered from 1199 different subjects. Image acquisition protocol contemplates a semi-controlled environment, with strict expression, pose and illumination changes.

B. FRGC

Collected at the University of Notre Dame, the Face Recognition Grand Challenge (FRGC) database [11] consists of high resolution ($\approx 1200 \times 1400$ pixels) color still images, captured on both controlled and uncontrolled environments. The controlled subset was captured on a studio under uniform illumination, where subjects were required to stand still while looking straight at the camera and essay neutral and smiling expressions. As for the uncontrolled acquisition, images were shoot in different scenarios, disregarding both background and illumination. Data is split into a training partition of 12776 images from 275 subjects, and a testing partition of 24042 images from 466 subjects, 6 images per session for each subject in both partitions. Illumination is not regular, as the illumination bursts for a short period of time, and main noise factors are observable (eye blink, motion blur, occlusions, reflections). Acquired data is stored on 2048×2048, 15 frames per second (fps) AVI files, where iris spatial extension is about 120 pixels [12].

[1] Although not so common, the FC-NET database will be included by its relevant facial aging characteristics.

(a) FERET (b) FRGC (c) MBGC (d) UBIRIS2 (e) FG-NET

Fig. 2. Sample images from the commonly used datasets on evaluating periocular algorithms. Except from (d), data has been cropped for illustration purposes.

TABLE I. OVERVIEW OF DATABASE SPECIFICATIONS. VARYING ELEMENTS ARE DISTANCE (D), EXPRESSION (E), ILLUMINATION (I), OCCLUSION (O) AND POSE (P).

Name	Images	Subj.	Dimensions	Variations
FERET	14051	1199	512×768	E, I, P.
FRGC	36818	741	$\approx 1200 \times 1400$	E, I.
MBGC	149 AVI	114	2048×2048	D, E, I, O, P.
UBIRIS.v2	11102	261	800×600	D, O, I.
UBIPr	10950	261	Multiple	D, I, O, P.
FG-NET	1002	82	$\approx 400 \times 500$	D, E, I, P.

C. UBIRIS.v2

The UBIRIS.v2 is a unconstrained iris database [13], captured on the VW from moving subjects, at different distances and challenging illumination conditions, simulating realistic acquisition issues with the associated noise factors. Data for both eyes is separately available, as well as the surrounding periocular data, thus being prone to stress not only robust iris related methods for the visible spectrum, but periocular ones and their fusion as well. The 11102 acquired images represent a total of 261 subjects, from different ages and ethnicities.

D. UBIPr

This newly created UBI Periocular Recognition (UBIPr) database, by Padole and Proença [14], represent a renewed effort to advance periocular biometric research, providing new means of evaluating robust methods, at "higher levels of heterogeneity".

In opposition the most common datasets used for periocular method evaluation, noise factors were actually introduced through acquisition setup: varying acquisition distance, irregular illumination, pose and occlusion. In addition, database manual annotation include ROI and essential landmarks.

Dimensions vary, accordingly to the acquiring distance, between 501×401 (8m) and 1001×801 (4m).

E. FG-NET

FG-NET is a facial aging database with around one thousand images from 82 subjects, 0 to 69 years old. Captured with different acquisition setups and many years apart, subjects have clear changes in illumination, pose and expression. Images are 400×500 pixels in size, captured on VW, and for each one a 68 facial landmark points annotation is also provided.

IV. RELEVANT RESEARCH

In this section we will detail the relevant research on periocular biometric recognition, providing at Table III a summarized overview over the described methods and reported results.

A. Park et al. [8], [15]

Park et al. pioneering approach [15] explored the recognition capabilities of the periocular region. Feature extraction is divided in two approaches: local and global, as information concerns local regions, or is extracted from the whole image (or, in this case, several region of interest (ROI)).

For global feature extraction images are properly aligned using iris center and radius as reference. Although authors acknowledge eye corners to be more fit for such task [8], they claim that such points cannot be reliably determined. Then, two well-known distribution-based descriptors are employed, namely Histogram of Oriented Gradients (HOG) [16] and Local Binary Patterns (LBP) [17], [18]. Values are computed for a given ROI independently, and then quantized into 8-bin histograms. The ROI are contiguous squares, where the side equals in length the iris radius, forming a 7 by 5 grid centered on the iris. Those histograms, combining shape and texture information, are merged into a single-dimension array, easily matchable to an identical one (from another image) simply by computing the Euclidean distance.

As for the local features, Scale-Invariant Feature Transform (SIFT) [19] allowed the detection of a set of key-points, encoded with their surrounding pixels information, and compared against their counterparts from the testing image. SIFT offers invariance to translation, scaling and rotation.

Tests were conducted over a "small" (899 images, 30 subjects, 2 sessions) database of frontal periocular images, acquired in the VW. Although face matching achieving 100% rank-1 recognition accuracy, the reported recognition for periocular range from 62.5% when using HOG features, to 80.8% when fusing them with SIFT results. Curiously, combining the three descriptors didn't overcome those results, although joint performance was very close: 80%.

On their later work [8], authors went further on stressing periocular applicability for biometric recognition, analyzing the impact of diverse factors over performance: eyebrow inclusion or disguising, automatic segmentation, side information, iris and sclera masking and expression variation.

As expected, results highlighted eyebrow information importance, being more significant over SIFT where improvements reached almost 19%. Nonetheless, the eyebrow inclusion is more favorable over manual segmentation, as its performance degraded when using automatic segmentation through OpenCV, which was not observed on "eyebrow-less" data. Facial side information, on the other hand, can be considered almost irrelevant, since performance variation from both to same side matching didn't go behind 1% except for SIFT on 2 of the 48 test setups.

Changes in subjects' expression significantly lowered the performance of LBP and HOG, although on SIFT, more robust to distortions, a slightly increase was registered. Masking the iris and the entire eye also caused performance to decrease, this time being SIFT the more disfavored. Top accuracy for single classifiers was 79.49%, achieved through SIFT on unmasked periocular images, manually segmented with the eyebrow, when compared to an image captured from the same side and expression. As reported in their prior paper [15], score level fusion didn't represent a significant performance improvement.

The authors also simulated periocular recognition over non-ideal conditions, performing four simple tests: result comparison against recognition with partial (occluded) facial and periocular images; conducting cosmetic changes on the eyebrows; template aging; and perspective variations. For the first step, they used FaceVACS[2] face recognition system, whose 99.77% recognition accuracy on "clear" face images, dropped

[2] FaceVACS SDK available at: http://www.cognitec-systems.de

to 39.55% simply by occluding the lower region. Occluding the periocular region is also an element of concern, since relatively low occlusions lead to significant decay on performance. Without score fusing the feature encoding methods, 10%, 20% and 30% periocular occlusion led to accuracies no greater than 25.97%, 20.51% and 10.12% respectively (all with SIFT).

On eyebrow modifications, the TAAZ[3] tool was used to simulate eyebrow makeover, producing a decay from 7.5% on LBP to 10% on the other descriptors. The tests regarding pose effect were the ones with greater impact over periocular recognition accuracy, specially when using SIFT. Apart from frontal images, subjects shoot with $15°$ and $30°$ rotation of the head, produced a 35% and 45% decay on this method's accuracy, respectively. Finally, another concern the authors rise is the apparent tendency of the periocular region not to be stable over relatively small amounts of time. Images captured 3 months apart from each other have up to 15% less accuracy, and about 30% on only half an year.

As further work, multi-spectral analysis is suggested, along with improvements on the alignment and matching methods. Fusion with iris or face recognition is also not discarded.

B. Miller et al. [20], [21]

Miller et al. [20] analyze periocular skin texture using Uniform Local Binary Patterns (ULBP) alone, with some deeper insights on each region's impact on the recognition process. The ULBP, as it name states, is an LBP-based method, with "improved rotation invariance with uniform patterns and finer quantization of the angular space" [22].

At a first stage, the periocular region is cropped proportionally to the distance between the eyes, and scaled to 100×160 pixels. Then, a 7 by 4 grid of square ROIs is defined, centered on the eye, and iris and sclera texture effects are eliminated overlapping an elliptical neutral mask to the image. Each ROI's histogram is normalized, and ULBP calculated using an 8-pixel neighborhood. As such neighborhood produces 59 different possible results, 59-bin histograms are populated with the result count, and then merged to produce a single-dimension array as the final periocular signature. Manhattan distance is used for subject identification against the database.

Experiments were conducted on subsets of the FRGC and FERET databases, for the left and right eyes separately and both eyes together. Recognition rates were around 84% and 71% for each eye individually, and 90 and 74% for both eyes together, on FRGC and FERET respectively.

Further to this work, Miller et al. [21] conducted deeper analysis on image quality impact over periocular local texture based recognition, namely changing blur, resolution and illumination, while comparing the results with similar experiments conducted with the entire face.

As preprocessing, the periocular region was cropped from the FRGC database in proportion to the distance between the eyes, and then resized to a square region with 251 pixel long sides. Upon grayscale conversion, image histogram is equalized and the eye is masked. Texture is then encoded using LBP over a regular block division of the image, and values used to populate an histogram, similarly to other periocular approaches.

Image blurring was achieved through Gaussian filter convolution, and results showed that even though face being far less affected by small amounts of blur than periocular, this last trait is slightly better at high blur levels. As for resolution, images were down-sampled up to 40% its original size, and behavior was similar to the one of blurred images.

Illumination variation was not simulated, since the FRGC database already contains both controlled and uncontrolled acquired images. The low accuracy verified when matching pairs of images captured on uncontrolled setups suggest that local appearance approaches like LBP are not suited for irregular lighting conditions.

Finally, information differences from one color channel to the others were also analyzed. Conclusions show the green channel as the more discriminant, with accuracy levels $\approx 23\%$ higher than for the red channel (which is presented as the less discriminant). In fact, when fusing scores from all three channels, the red contribution only lowers the overall performance. Blue channel has similar texture information as the green one.

In a general way, periocular was proven to outperform face recognition in the stressed setups.

Further work includes conducting the same tests for different classification methods, possibly adapting Support Vector Machines (SVM) usage as suggested by Savvides et al. [23].

C. Adams et al. [24]

Adams et al. extended Miller's work [20], proposing the usage of a Genetic & Evolutionary Computing (GEC) method to optimize the original feature set.

The first stage of feature extraction was conducted as described by Miller et al. [20], and on the second stage the Genetic & Evolutionary Feature Extraction (GEFE) chosen was the Steady-State Genetic Algorithm (SSGA), as implemented by the NASA's eXploration Toolset for Optimization of Launch and Space Systems (X-TOOLSS)[4].

Reported results were about 86% accuracy for either eye on the FRGC database, and 80% on similar experiments for the FERET. Best results were obtained when using both eyes: 85% and 92% for those same datasets.

The usage of GEC represented an improvement of at least 10%, and only $49 \approx 52\%$ of the initial features were used. Nevertheless, the selected algorithm was not proven to be the optimal for that specific periocular features.

D. Xu et al. [25]

Inspired by the work of Park et al. [15], the authors decided to expand their experiments to less ideal imaging environments, evaluating the performance of different feature schemes over the FRGC database.

In addition to LBP and SIFT, both local and global feature extraction schemes were stressed: Walsh masks [26], Law's masks [27], DCT [28], DWT [29] Force Fields [30],

[3]Free virtual makeover took, available at http://www.taaz.com

[4]http://nxt.ncat.edu/

SURF [31], Gabor Filters [32] and Laplacian of Gaussian (LoG). The LBP itself was tested while applied over some of the other methods (Table II). For matching, different distance metrics were tested: Normalized Cosine, Euclidian and Manhattan.

TABLE II. RANK-1 ACCURACY FOR LBP FUSION WITH OTHER METHODS [25].

Fused methods	Accuracy
LBP + LBP	42.5%
Walsh Masks + LBP	52.9%
Laws' Masks + LBP	51.3%
Discrete Cosine Transform (DCT) + LBP	53.1%
Discrete Wavelet Transform (DWT) + LBP	53.2%
Force Field Transform + LBP	41.7%
Gabor Filters + LBP	12.8%
LoG Filters + LBP	30.9%

Experiments shown that best results were registered when using local descriptors, and the post-application of LBP was translated in a performance boost (Table II). Top accuracy of 53.2% was attained with DWT + LBP, followed closely when combining this last one with DCT (53.1%) and Walsh or Laws' Masks (52.9% and 51.3%).

Worst outcomes were registered for SIFT and Speed Up Robust Features (SURF), with a Verification Rate (VR) no greater than 1%, possibly due to the low resolution of the images.

E. Bharadwaj et al. [33]

Bharadwaj et al. propose a new global matcher (GIST), and its combination with ULBP for periocular recognition over VW uncooperative images from UBIRIS.v2 database.

The GIST algorithm consists on combining five perceptual dimensions, usually associated with scene description [34]: naturalness,openness, roughness, expansion and ruggedness.

When computing the global GIST descriptor, and to achieve local contrast normalization, the image is preprocessed with Fourier transform. Then, the spatial envelope is computed using a set of Gabor filters (4 scales $\times$ 8 orientations, producing a 1536 element GIST descriptor).

The ULBP is computed over the original image, sliced into 64 patches (producing a 64×256 descriptor).

For both descriptors, matching is computed using χ^2 distance, and min-max normalized results from both eyes are fused simply by using a weighted sum.

Results showed that GIST overperformed ULBP, with Rank-1 accuracy around 62% for the regions separately, and 70.82% for their fusion. The ULBP performance was around 53%, and 63.77% when fusing both region results. When combining both descriptors, accuracy was boosted to 73.65%.

F. Woodard et al. [7], [12]

In their work, Woodard et al. [7] aimed at evaluating periocular performance, thus determining its usability as a biometric trait over NIR and VW data. Their analysis is focused only on second level features (texture and color).

As pre-processing, periocular slice of images is cropped, and an elliptical mask overlapped to the iris and sclera region for "unbiased" periocular analysis. Cropped color images from the FRGC are scaled down to 100×160, while the periocular NIR frames from the Multi Biometric Grand Challenge (MBGC) are 601×601 pixel.

Texture features were encoded the same way for both databases, through LBP computation over a ROI grid, which was then quantized into histograms. As for the color information on FRGC images, it was encoded using color histograms for red and green channels. On this database, score level fusion was used to combine texture and color results. Matching was achieved using Manhattan distance for LBP and Bhattacharya distance for color histograms.

Results suggest texture information to be more discriminant that color, and score fusion only slightly improves overall performance. As a comparison term, reported texture based accuracy was around 90% and 88% on the VW, and 81% and 87% on NIR for the left and right periocular regions respectively.

On their later work, Woodard et al. [12] make use of the periocular region texture information to improve iris data reliability, aiming at overcoming the difficulties when dealing with non-ideal imaging.

Tests were conducted over MBGC that, although being a NIR database, is a challenging one for iris recognition due to at-a-distance in-motion subjects and illumination variations. Frames were treated as described above, with texture measured computing LBP the same way. Iris processing was as of Daugman's [35], except for the segmentation that was manually performed to avoid further errors. Both methods' results were then normalized using min-max scheme, and combined by a simple weighted sum.

Results demonstrate iris' poor accuracy ($10.1\% \approx 13.8\%$) to benefit from fusing with periocular results, raising rank-1 to 96.5%.

G. Padole and Proença [14]

Padole and Proença also stressed how noise factors deteriorate periocular recognition, using natural images where those factors were included by the acquisition framework instead of simulating them: pose variation, distance of the subject, pigmentation and occlusion.

Inspired by the work of Park et al. [15], they used the same feature extraction techniques, except that ROI center was computed with relation to eye-corners instead of iris center. This new alignment method led to most significant improvements, specially since in unconstrained biometrics gaze variations are more prone to happening.

On score level fusion, linear an non-linear methods were also tested: logistic regression [36] and Multi Layer Perceptron (MLP) respectively. Although the last one reported to lead to slightly better results, difference was not significant.

For the stressed covariates, interesting conclusions were reached. Results shown that closer acquired distances didn't led to better performance, and neither did very large ones. Worst results were obtained for images acquired at 4 m, and though highest stressed distance was 8 m, top performance was obtained at 7 m. Not surprisingly, pose variation impact on performance was in inverse proportion: higher tilting angle result in lower accuracy values. Same as for the occlusion.

Finally, iris pigmentation was reported to also impact periocular recognition performance, specially on heavily pigmented ones which lead to lower accuracy. Best results were obtained for medium pigmented irides.

Another interesting discovery was that subject gender affects recognition rates. More precisely, female subjects are easily identified using periocular biometrics than male ones.

H. Juefei-Xu et al. [37]

Juefei-Xu et al. address in their work the aging effect on periocular recognition, reported to be an issue by several authors (e.g. Park et al. [8]), even at relatively small time lapses (months). This important issue is not trivial, as modeling the aging process would require large datasets, and the decoding of its dependence on external factors, as ethnicity, gender, etc. The authors method was developed and validated on images from the FG-NET database, taken years apart at different acquisition setups, thus also dealing with illumination, pose and expression issues.

Their method starts by preprocessing the periocular region: pose is corrected through Active Appearance Models (AAM), illumination is dealt with anisotropic diffusion model, and region is normalized using the landmark points provided with the database. Next step is feature extraction using Walsh-Hadamard transform encoded LBP (WLBP), followed by unsupervised discriminant projection (UDP) [38] application that boosted results to very high performance levels.

Results show UDP to give better accuracy than Principal Component Analysis (PCA) and Locally Preserving Projections (LPP) by up to 40%. As for WLBP, results were 15% better than raw pixel intensity matching, and pose correction resulted in a 20% improvement. Finally, the proposed method for the tested images resulted in a complete 100% identification accuracy.

I. Hollingsworth et al. [39], [40]

The human ability to use contextual information and to ”disregard” most of noise factors adapting itself to surrounding conditions is outstanding, making it a harder task for machines to mimic. In fact, recognition algorithms should not try to just mimic the human perception system, but to understand its way of working, and then seek alternate strategies to tackle the same issues.

Hollingsworth et al. understood existing methods to have overlapped that step. Having that in mind, they [39] established parallelisms between human perception and automatic recognition systems, identifying which ocular elements humans find more useful for periocular recognition.

On their essay, 640×480 NIR images were acquired from 120 subjects using an iris camera (LG2200), and the iris was completely masked to avoid biased answers. Only periocular from eyes’ tight vicinity is visible, with some features used by other methods partially hidden (e.g. eyebrows). 80 pairs of images were presented to 25 human observers, who were asked to tell if they belong to the ”same person” or ”different people”, and how ”certain” they were. Further to that, the observer had to individually rate each one of the features’ helpfulness, in a three level scale. Results showed eyelashes to be the most helpful periocular feature, closely followed by the *medial canthus* and the eye shape. The observers based themselves on eyelash clusters, density, direction, length and intensity. To the human observers, skin was actually the less useful. Average human accuracy was 92%.

On their later work [40], similar tests with human observers were widen to the VW band, with a more extent study on new factors. The algorithms suggested by Park et al. [15] were also implemented for periocular performance comparison, and irides were evaluated using the IrisBEE biometric system from ICE [41].

Trial data was also widen to 210 subjects, imaged on the same controlled fashion with a setup as above, and on the VW using a Canon D80 camera. The amount of observers also increased to 56, to whom 140 pairs of images were presented for each one of the four sets of experiments built: NIR and VW, periocular and iris images. Test subjects could then rank their certainty of a positive match in a 5 level scale, and for the periocular images they had to specify how helpful individual features were (”eye shape”, ”tear duct”, ”outer corner”, ”eyelashes”, ”skin”, ”eyebrow”, ”eyelid”, ”color”, ”blood vessels” and ”other”).

Human NIR periocular recognition accuracy dropped to 78.8%, probably due to the different pairing system and limited observation time, and VW performance was set on 88.4%. Machine results were similar, within a 1% difference on overall accuracy. The features identified as fit for periocular NIR region were similar to the ones at [39], but for VW data changes occurred: blood vessels, skin and eye shape were reported to be more helpful than eyelashes.

When acquiring data on VW band, differences on acquired skin details are perceptible. Also with the LG2200 camera illumination, being designed for iris recognition, usually causes skin saturation. As so, VW band was found to be preferable for periocular recognition tasks.

Human perception of iris features is greater on NIR images, leading to 85.6% accuracy against 79.3% on VW. However, and unlike periocular, machines recognition was 13% better, on average, than human observers, with 100% and 90.7% accuracy for those same bands.

V. Conclusions

The interest on the periocular region as a biometric trait has justifiably increased over the last years, considering the pioneer approach of Park et al. [15] a starting point. Subsequently, even simple algorithms led to fair performance levels, and the surprisingly good response of LBP based methods (like ULBP and WLBP) is noteworthy.

The recently developed methods focus mainly on texture analysis and keypoint extraction. Periocular is currently regarded as specially suitable for unconstrained and uncooperative scenarios, where iris cannot be properly imaged and neither a full facial picture can be obtained. Also, results

TABLE III. Overview of the most relevant periocular recognition methods.

Approach	Features	Extract	Classifier	Dataset	Accuracy
Park et al. [15]	Shape, Texture, Key-Points	HOG, LBP, SIFT	Euclidean distance, SIFT matcher	899 VW images, 30 subjects, 2 sessions	HOG: 62.5%, LBP: 70.0%, SIFT: 74.2%, Best: 80.8%
Miller et al. [20]	Texture	ULBP	Manhattan distance	FRGC, FERET	FRGC: 89.8%, FERET: 74.1%.
Adams et al. [24]	Texture	LBP +GEFE	Manhattan distance	FRGC, FERET	FRGC: 92.2%, FERET: 85.1%.
Woodard et al. [7]	Color, Texture	RG color histogram, LBP	Bhattacharya, Manhattan distance	FRGC, MBGC	Left VW peri: 90% Right VW peri: 88% Left NIR peri: 81% Right NIR peri: 87%
Woodard et al. [12]	Texture	Daugman's irisCode, LBP	Hamming distance, Manhattan distance	MBGC	Left Iris: 13.8% Left Peri: 92.5% Fusion: 96.5% Right Iris: 10.1% Right Peri: 88.7% Fusion: 92.4%
Xu et al. [25]	Texture, Key-Points	Walsh Masks, Laws' Masks, DCT, DWT, Force Field Transform, Gabor Filters, LBP, SIFT, SURF,	Cosine distance, Euclidean distance, Manhattan distance	FRGC	DWT+LBP: 53.2% DCT+LBP: 53.1% Walsh+LBP: 52.9% Laws'+LBP: 51.3% ...
Bharadwaj et al. [33]	Naturalness, Openness, Roughness, Expansion, Ruggedness, Texture	GIST, ULBP	χ^2 distance	UBIRIS.v2	GIST: 70.82% ULBP: 63.77% Fusion: 73.65%
Juefei-Xu et al. [37]	Texture	WLBP+UDP	Cosine distance	FG-NET	100%
Hollingsworth et al. [39]	Human	Human	Human	NIR images, 120 subject	92%
Hollingsworth et al. [40]	Human	Human	Human	NIR and VW, 210 subjects	NIR Peri: 78.8% VW Peri: 88.4% NIR Iris: 85.6% VW Iris: 79.3%

favoring VW periocular over NIR also show its fitness for more relaxed setups and for its use based on conventional surveillance cameras.

However, some issues remain to be properly addressed, specially the about poses, occlusions and aging. Regarding the latter, extending Juefei-Xu et al [37] work to different scenarios should be considered.

The work of Hollingsworth et al. [39], [40] on human perception suggests that eye shape constitutes a powerful ally to the skin analysis methods on both spectral bands, thus making us rethink periocular recognition, possibly taking a leap away the overused texture methods. Eyelashes are also pointed as a good indicator, specially for NIR, keeping in mind that images differ from the "traditionally" used periocular images and the close capturing of the data could have biased the results. Those issues should be addressed in further work, as well as a more complete and uniform study of existent methods' performance over the UBIPr dataset.

Acknowledgment

The financial support given by "FCT - Fundação para a Ciência e Tecnologia"and "FEDER"in the scope of the PTDC/EIA/103945/2008 research project "NECOVID: Negative Covert Biometric Recognition"is acknowledge. Also, the support given by IT - Instituto de Telecomunicações in the scope of the "NOISYRIS"research project is acknowledge too.

References

[1] K. Ricanek, M. Savvides, D. Woodard, and G. Dozier, "Unconstrained biometric identification: Emerging technologies," *Computer*, vol. 43, no. 2, pp. 56 –62, February 2010.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, December 2003.

[3] K. Bowyer, K. Hollingsworth, and P. Flynn, "Image understanding for iris biometrics: A survey," *Comput. Vis. Image Underst.*, vol. 110, no. 2, pp. 281–307, May 2008.

[4] K. Y. Shin, G. P. Nam, D. S. Jeong, D. H. Cho, B. J. Kang, K. R. Park, and J. Kim, "New iris recognition method for noisy iris images," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 991 – 999, 2012.

[5] T. Tan, X. Zhang, Z. Sun, and H. Zhang, "Noisy iris image matching by using multiple cues," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 970 – 977, 2012, ¡ce:title¿Noisy Iris Challenge Evaluation II - Recognition of Visible Wavelength Iris Images Captured At-a-distance and On-the-move¡/ce:title¿.

[6] G. Santos and E. Hoyle, "A fusion approach to unconstrained iris recognition," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 984 – 990, 2012.

[7] D. Woodard, S. Pundlik, J. Lyle, and P. Miller, "Periocular region appearance cues for biometric identification," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, June 2010, pp. 162 –169.

[8] U. Park, R. Jillela, A. Ross, and A. Jain, "Periocular biometrics in the visible spectrum," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 1, pp. 96 –106, 2011.

[9] A. Jain, P. Flynn, and A. Ross, Eds., *Handbook of biometrics*. Springer, 2008.

[10] P. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The feret evaluation methodology for face-recognition algorithms," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 10, pp. 1090 – 1104, October 2000.

[11] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, June 2005, pp. 947 – 954 vol. 1.

[12] D. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, "On the fusion of periocular and iris biometrics in non-ideal imagery," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, August 2010, pp. 201 –204.

[13] H. Proença, S. Filipe, R. Santos, J. Oliveira, and L. Alexandre, "The ubiris.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 8, pp. 1529 –1535, August 2010.

[14] C. Padole and H. Proença, "Periocular recognition: Analysis of performance degradation factors," in *Biometrics (ICB), 2012 5th IAPR International Conference on*, April 2012, pp. 439 –445.

[15] U. Park, A. Ross, and A. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in *Biometrics: Theory, Applications, and Systems, 2009. BTAS '09. IEEE 3rd International Conference on*, September 2009, pp. 1 – 6.

[16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *In Proceedings of the International Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886–893.

[17] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision amp; Image Processing., Proceedings of the 12th IAPR International Conference on*, vol. 1, October 1994, pp. 582 –585 vol.1.

[18] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.

[19] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, November 2004.

[20] P. Miller, A. Rawls, S. Pundlik, and D. Woodard, "Personal identification using periocular skin texture," in *Proceedings of the 2010 ACM Symposium on Applied Computing*, ser. SAC '10. New York, NY, USA: ACM, 2010, pp. 1496–1500.

[21] P. Miller, J. Lyle, S. Pundlik, and D. Woodard, "Performance evaluation of local appearance based periocular recognition," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, September 2010, pp. 1 –6.

[22] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971 –987, July 2002.

[23] M. Savvides, R. Abiantun, J. Heo, S. Park, C. Xie, and B. Vijayakumar, "Partial holistic face recognition on frgc-ii data using support vector machine," in *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, June 2006, p. 48.

[24] J. Adams, D. Woodard, G. Dozier, P. Miller, K. Bryant, and G. Glenn, "Genetic-based type ii feature extraction for periocular biometric recognition: Less is more," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, August 2010, pp. 205 –208.

[25] X. J., M. Cha, J. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides, "Robust local binary pattern feature sets for periocular biometric identification," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, sept. 2010, pp. 1 –8.

[26] T. Beer, "Walsh transforms," *American Journal of Physics*, vol. 49, no. 5, pp. 466–472, 1981.

[27] K. Laws, "Rapid texture identification," in *Proc. SPIE Conf. Image Processing for Missile Guidance*, 1980, pp. 376–381.

[28] N. Ahmed, T. Natarajan, and K. Rao, "Discrete cosine transform," *Computers, IEEE Transactions on*, vol. C-23, no. 1, pp. 90 –93, January 1974.

[29] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674 –693, July 1989.

[30] D. Hurley, M. Nixon, and J. Carter, "A new force field transform for ear and face recognition," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 1, 2000, pp. 25 –28.

[31] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[32] D. Clausi and M. Jernigan, "Towards a novel approach for texture segmentation of sar sea ice imagery," in *26th International Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society*, Vancouver, BC, Canada, 1996, pp. 257–261.

[33] S. Bharadwaj, H. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, September 2010, pp. 1 –6.

[34] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, pp. 145–175, 2001.

[35] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 11, pp. 1148 –1161, November 1993.

[36] D. Hosmer and S. Lemeshow, *Applied logistic regression (Wiley Series in probability and statistics)*. Wiley-Interscience Publication, September 2000.

[37] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen, "Investigating age invariant face recognition based on periocular biometrics," in *Biometrics (IJCB), 2011 International Joint Conference on*, October 2011, pp. 1 –7.

[38] J. Yang, D. Zhang, J. yu Yang, and B. Niu, "Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 650 –664, April 2007.

[39] K. Hollingsworth, K. Bowyer, and P. Flynn, "Identifying useful features for recognition in near-infrared periocular images," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, September 2010, pp. 1 –8.

[40] K. Hollingsworth, S. Darnell, P. Miller, D. Woodard, K. Bowyer, and P. Flynn, "Human and machine performance on periocular biometrics under near-infrared light and visible light," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 2, pp. 588 –601, April 2012.

[41] P. Phillips, W. Scruggs, A. O'Toole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, "Frvt 2006 and ice 2006 large-scale experimental results," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 5, pp. 831 –846, May 2010.