Published in: Pattern Recognition Letters Volume 57, 1 May 2015, Pages 43-51

The Version of Record is available online at: https://doi.org/10.1016/j.patrec.2014.10.017

BIRD: Watershed Based IRis Detection for mobile devices ☆

Andrea F. Abate[a], Maria Frucci[b], Chiara Galdi[a], Daniel Riccio[c,*]

[a] Università degli Studi di Salerno, Fisciano, I-84084, Italy
[b] Istituto Calcolo e Reti ad Alte Prestazioni, CNR, Napoli, I-80131, Italy
[c] Università di Napoli Federico II, Napoli, I-80125, Italy

a b s t r a c t

Communications with a central iris database system using common wireless technologies, such as tablets and smartphones, and iris acquisition out of the field are important functionalities and capabilities of a mobile iris identification device. However, when images are acquired by means of mobile devices under uncontrolled acquisition conditions, noisy images are produced and the effectiveness of the iris recognition system is significantly conditioned. This paper proposes a technique based on watershed transform for iris detection in noisy images captured by mobile devices. The method exploits the information related to limbus to segment the periocular region and merges its score with the iris' one to achieve greater accuracy in the recognition phase.

1. Introduction

Iris detection is one of the most powerful biometrics for security applications, e.g., physical access control, as it is witnessed by the wide literature in the field [2–4]. In fact, iris acquisition devices, based on near infrared or visible light technology, are characterized by non-invasiveness. Moreover, the iris of any individual subject is characterized by uniqueness, external visibility and remains the same during the entire life of the subject. Consequently, iris recognition systems are generally preferable with respect to systems based on other biometrics.

It is possible to use simple image processing tools during acquisition if you are able to control the lighting conditions and the distance of the person whose iris must be recognized. Unfortunately, however, even in the case of a cooperative subject, the performance of an iris recognition system decays significantly, if the lighting conditions are irregular, the acquisition is accomplished on the move or mobile devices are used. Tan et al. [15] have shown that with regard to segmentation of iris and periocular area, the process of localization and transformation of objects of interest in images acquired with visible light in an uncontrolled environment presents many more pitfalls. This work, as well as [10], addresses the problem in the context of iris acquisition on the move, but differently from [10], this article is specifically oriented to the case of acquisition done by means of mobile devices.

Iris is delimited by sclera and pupil, which represent the lightest and the darkest parts of the eye, respectively. A number of tasks have to be accurately accomplished by an iris recognition system, such as acquisition of the eye image, iris segmentation, iris coding and recognition. In particular, accuracy is fundamental during iris segmentation, to generate an iris code able to improve the performance of the recognition system.

In this paper we introduce the watershed Based IRis Detection (BIRD) technique for smart mobile devices, which is the follow up of a technique we have recently suggested [7,8]. BIRD exploits the use of the watershed transform to identify more precisely the iris boundary and, hence, to obtain a more accurately computed code for iris recognition.

A positive feature of the watershed transform is that the contours delimiting the regions into which an image is divided are mostly placed where human observers perceive them. In fact, the watershed transformation is a growing process performed generally on the gradient image, where the edges are enhanced. This feature should allow to correctly detect the limbus boundary. In turn, a negative feature is over-segmentation, i.e., the image may be partitioned into a number of parts that is remarkably larger than expected. Over-segmentation is particularly evident when all the regional minima in the gradient image are considered as seeds for the growing process. A common strategy to overcome this drawback is to adopt region merging and/or seed selection to reduce the number of watershed regions. However, in the case of eye images, processes for over-segmentation reduction cannot be stressed. Otherwise, some weak boundaries between sclera and limbus (light eye case) or between eyelashes and limbus (dark eye case) might be no longer present in the segmented image.

BIRD performs a binarization of the watershed transform to obtain an image where large portions of the limbus boundary are better enhanced. In this way BIRD is able to exploit the positive features of the watershed transform independently of over-segmentation problem. The boundaries of the foreground region are then inputted to a circle detection process, which aims at finding the circle that best approximates the limbus boundary (limbus circle).

To refine further the limbus boundary, the regions of the watershed transform overlapping the limbus circle are analyzed. Circle fitting [16] is then applied within the iris boundary to identify the pupil boundary (pupil circle).

The information regarding position and size of the iris (its center and radius) constitute the starting point for the delimitation of the periocular region. Recent studies [11] showed how the latter could be considered itself a biometrics. BIRD checks out the periocular region and applies to it a transformation from Cartesian to polar coordinates. In this way, it is possible to apply to the periocular area a process of extracting and matching characteristics similar to that used for the iris. Iris and the periocular region are then fused at score level through a simple sum criterion in order to increase the accuracy of the recognition system.

BIRD has been tested on a database including 1500 eye images, taken from 75 individual subjects. Eye images are the results of outdoor and indoor acquisitions, accomplished by means of three mobile devices (tablet SamsungGalaxy, Apple iPhone 5, and SamsungGalaxy S4 smartphone). Eye images in the database are rather different from each other, due to the technical features characterizing each of the three cameras and due to the uncontrolled acquisition conditions.

The rest of the paper develops as follows. Section 2 describes a pre-processing phase that improves the quality of the input eye image and reduces the computational cost of the whole process. Section 3 is concerned with computation and binarization of the watershed transform. Section 4 describes the detection of the limbus circle and the refining process adopted to fit the real limbus boundary. Circle fitting is then accomplished also for pupil detection. The segmentation of the periocular region is then addressed in Section 5. Section 6 deals with iris recognition. Section 7 regards the experimental setup. The interoperability of BIRD is also investigated by performing cross-datasets experiments. Final remarks are given in Section 8.

2. Pre-processing

Uncontrolled iris acquisition may produce an image with local distortions due for example to shadows and different color temperature. A color/illumination correction is performed to reduce such local distortions, by processing separately the three RGB components of the eye image as gray level images. For each gray level image, a Gaussian filtered version is computed. A new image is built, where each pixel is set to the ratio between the value of the homologous pixels in the gray level image and in its filtered version. This ratio has the side effect to bring out the details in the image, so the kernel parameters of adopted Gaussian filter play a fundamental role. The parameters are the kernel size g_k, the average m_k and variance σ_k (they are mainly related to the resolution of the input image). In fact, a kernel too small excessively flattens the distribution of the colors within the image, while one that is too large will not produce any substantial correction on lighting and color distortions in it. In order to find a viable relationship between Gaussian kernel parameters to be adopted and the resolution of the image, was considered a set of pictures of irises at different resolutions $w_k \times h_k$ where $k = 1, 2, \ldots, n$, $w_{k+1} > w_k$ and $h_{k+1} > h_k$. The image resolution was represented by considering the value of the diagonal $d_k = \sqrt{w_k^2 + h_k^2}$. The optimal parameters for the Gaussian kernel were determined in terms of segmentation and recognition accuracy obtained on the set of images. It was observed that the relationship between g_k and d_k is quadratic, i.e. $g_k = \alpha_2 d_k^2 + \alpha_1 d_k + \alpha_0$, while $m_k = g_k$ and $\sigma_k = 0.1 g_k$. In this case, it was found that $\alpha_2 = -0.0001$, $\alpha_1 = 0.3064$ and $\alpha_0 = 11.1351$.

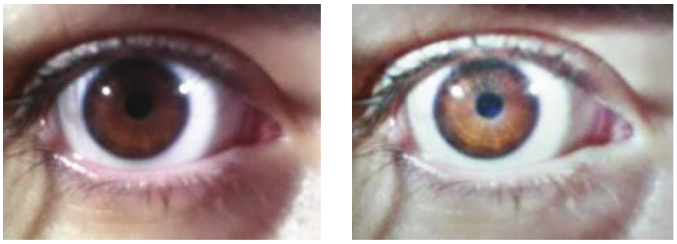

Fig. 1. An eye used as running example (left), and the result of color/illumination correction (right).

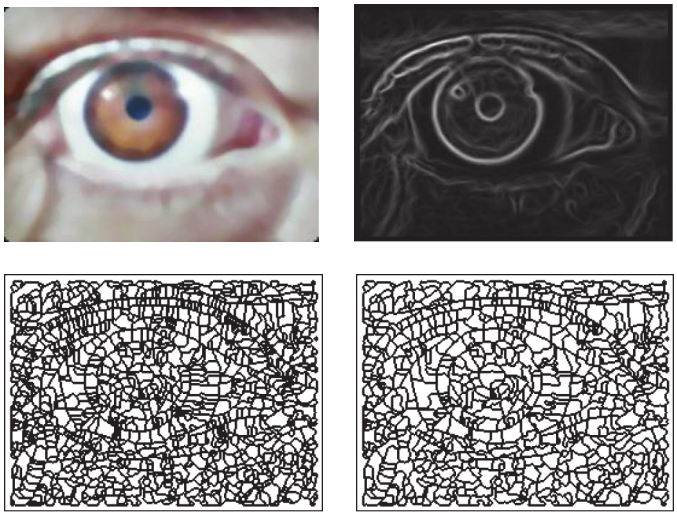

Fig. 2. From top left to bottom right: resized and smoothed image, gradient image, watershed transform, result of region merging.

A normalization process of pixel values is performed to map the values in the range [0, 255]. The combination of three obtained gray level images originates the color/illumination corrected image. The left image in Fig. 1 shows the original image, while the right one corresponds to the color/illumination corrected image. As BIRD is able to work even on low resolution images, it is possible to limit the computational cost of the method. The color/illumination corrected image is resized by using a linear interpolation method without changing the aspect ratio, in order to get an image of the eye in the foreground with a horizontal resolution of 200 pixels (vertical resolution depends on aspect ratio). As previously, the process of correcting lighting/color enhances the details in the image and these details are irrelevant for the segmentation. Thus a median filter is applied with a fixed-size window 7×7. The window size can be regarded as fixed, because the image is first brought to a standard resolution by step resizing. The resulting image I is shown in Fig. 2 top left.

3. Watershed transform and binarization

The effect of color/illumination correction is to generate an almost uniform image, independently of the acquisition conditions that are highly uncontrolled. To extract the region of interest, i.e., the iris, the watershed transformation is used to partition the image into regions, based on gradient information. As already pointed out, over-segmentation is likely to affect the obtained partition. Thus, a successive process is necessary to merge adjacent regions, characterized by a certain homogeneity. Though the watershed transformation is computationally heavy, its use significantly reduces the processing time of the remaining steps, which will involve operations to be applied

to a small number of regions rather than to their individual pixels. A color quantization process is also performed to associate a unique color to all pixels of the same region.

3.1. Watershed transform, region merging and color quantization

The watershed transformation performs a partition of an image into a number of disjoint regions; each of them is characterized by a certain degree of homogeneity according to a specific uniformity criterion. The partitioning process relies on a region growing strategy, where starting seeds are chosen as the regional local minima of the gradient image.

BIRD computes the gradient image from the color/illumination corrected image I. It first decomposes the color image I in its three RGB components and then applies the 3 × 3 Sobel edge filter to each of them, separately. The final gradient image is obtained by averaging gradients computed on the three channels (Fig. 2 top right). Then, the watershed transform W (Fig. 2 bottom left) is obtained by adopting the topographical distance approach [13].

Due to the very large number of sub-regions in W produced by the watershed transformation, a merging process is mandatory to reduce significantly their number. To this aim, a representative color is assigned to each watershed region and the merging criterion is based on the difference in representative color of adjacent regions.

Let R_i be a region of W and let C_i (r_i, g_i, b_i) be the representative color assigned to R_i, which is computed as the arithmetic mean of the colors of the pixels belonging to R_i. In the following, the representative color of R_i is denoted as C_i.

Two adjacent regions R_i and R_j are merged if it results:

$$d(C_i, C_j) < \delta,$$

where $d(C_i, C_j)$ represents the Euclidian distance between C_i and C_j. Thanks to the process of correcting the distortions of color and lighting applied during the preprocessing phase, the value for the threshold δ can be set permanently for all images and, in this application, it is experimentally set to 50. The result of merging is shown in Fig. 2 bottom right.

After merging, the representative colors of the so obtained watershed regions are updated originating a new version, Q, of the color-quantized image (Fig. 3 top left).

3.2. Binarization of watershed transform

It is worth to notice that, in general, the color of pupil and iris are darker than that of sclera and eyelids in an eye image. More precisely, the former two show colors closer to black, while the latter a color closer to white. BIRD strongly relies on this property to derive the binary image BW from the watershed transform. Indeed, all regions R_i whose representative color C_i is closer to black could be tentatively ascribed to the foreground, while regions with a representative color closer to white could be tentatively associated to the background. To this aim, information on the differences of the representative colors with respect to black and white should be properly taken into account to fix the threshold value able to cause the correct assignment of the regions to foreground or background.

In the RGB space, black is represented by (0, 0, 0) and white by (255,255,255), respectively. The Euclidean distances db_i and dw_i of all the representative colors C_i from black and white are then computed, as well as their arithmetic means, db and dw. Finally, the distance between black and white, dbw, is also computed.

In principle, if db_i db, R_i might be ascribed the foreground status. In turn, if dw_i dw, R_i might be associated to the background. However, a decision on the status of R_i cannot be taken in the following two cases:

(a) $db_i > db$ and $dw_i > dw$
(b) $db_i \leq db$ and $dw_i \leq dw$

The former case occurs if it results $db + dw < dbw$, while the latter case when it results $db + dw$ dbw. Thus, there are cases in which R_i would remain unassigned (case a), or R_i might be ascribed both the foreground and background status (case b).

To overcome the above problems, a binarization threshold T is used whose value is set to $db \times dbw / (db + dw)$, where the ratio $dbw/(db + dw)$ is a multiplicative weight for the arithmetic mean db. Thus, any region R_i of W such that db_i T $(db_i > T)$ is tentatively assigned to the foreground (background).

Let F and B be the sets of the regions tentatively ascribed the foreground status and the background status, respectively. Any region belonging to B is definitely considered as belonging to the background of BW, while regions belonging to F are further processed to take the final decision on their status. In particular, any region R_i belonging to F changes its status from foreground to background if it results:

$$d(C_i, c_B) \leq d(C_i, c_F)$$

where c_F and c_B, are the average foreground color and the average background color, respectively. The values c_F and c_B are computed as the arithmetic means of the colors associated to the regions of F and B, respectively. The final binary image BW is shown in Fig. 3 top right.

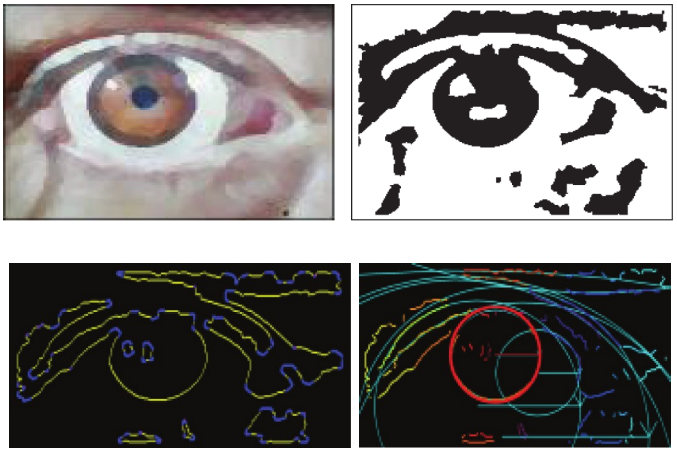

Fig. 3. From top left to bottom right: color quantized image, binarized image, foreground contours, best fitting circle (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Iris detection

Circle fitting is accomplished only on smooth contour components of the foreground F of BW. The obtained circles undergo a voting procedure, and the one getting the highest score is identified as the raw approximation of the limbus boundary. The boundary is refined by using information derived from the regions of W at least partially overlapping the circle. In order to locate the pupil region, circle fitting is applied again and circles found are scored according to two criteria, namely homogeneity and separability. Once again, the circle obtaining the best score is identified as the pupil boundary. Iris is then represented by the annular region enclosed by the limbus and the pupil circle.

4.1. Circle detection

The circle fitting procedure [10] searches for circumferences fitting the contour of the foreground region in BW. As it can be observed in Fig. 3 bottom left, the limbus represents only a small portion of the whole foreground. Thus, before performing the circle fitting procedure, a split operator is applied to the whole contour to partition it in smaller components. It is worth to notice that limbus boundary has

a very smooth circular shape, whose end points are characterized by a very high curvature. It comes out that curvature analysis may represents a suitable tool to detect end points of the limbus boundary, so as to isolate it from the whole contour just by breaking it at those points. In other words, the whole contour is divided into parts having a very smooth curvature (yellow parts in Fig. 3 bottom left), which are delimited by parts characterized by strong curvature changes (blue parts in Fig. 3 bottom left).

To estimate the curvature of a contour component CC, the sequence $S = p_1, p_2, \ldots, p_{|S|}$ of pixels of CC, detected by contour tracing, is considered. For each pixel p_i of S, a pixel p_{i+t}, with $t = 4 \times log_2(|S|)$ is selected. Let p_k be the contour point midway along the contour arc delimited by p_i and p_{i+t}, and let p_m be the mid-point of the straight line segment joining p_i and p_{i+t}. The curvature at p_i is estimated by dividing the distance $d(p_k, p_m)$ by the greatest of all distances computed for the pixels of CC. In this way the curvature values are normalized in the range [0,1], so that the points p_i, whose curvature value is greater than 0.5, are considered to be points of separation of two smooth contour components of CC.

For circle fitting, we follow the strategy described in [10] taking into account only contour components representing smooth curves and whose dimension is at least 5% of the whole contour. Many circles are generated, but only those included for at least 80% in the image are taken into account for the selection of the best fitting circle. A voting process that assigns a score to any generated circle identifies the best limbus approximation.

Let CL be a circle with radius r and let $CL1$ and $CL2$ be two additional circles having the same center as CL and radii $0.9r$ and $1.1r$, respectively. For each point in CL having polar coordinates (ρ,θ), two pixels, $p1$ and $p2$, located at the same angle θ on $CL1$ and $CL2$ are considered. The score for CL is computed as the sum of the differences between each pair of pixels corresponding to $p1$ and $p2$ in the binary image BW. The best fitting circle is represented by the circle with the maximal score, shown in red in Fig. 3 bottom right.

4.2. Limbus detection

The best fitting circle might not coincide completely with the limbus, because it has been generated starting from only a part of the limbus boundary. Moreover, the limbus is not always characterized by circular shape. Thus, the selected circle might include parts external with respect to the limbus (sclera, eyelashes and eyelids). To correctly identify the pixels actually belonging to iris and pupil, all the regions of W at least partially overlapping the circle (shown in purple in Fig. 4 left) are again taken into account to achieve a new segmentation BW' of W, where the foreground of BW' will include only iris and pupil. The remaining regions of W will be no more analyzed and assume the background status in BW'. The regions of W overlapping the limbus circle are divided in two different subsets RT and RP. RT is the set of regions of W totally overlapping the circle (red and blue regions in Fig. 4 middle), while RP is the set of regions of W only partially overlapping the circle (green regions in Fig. 4 middle). Note that not all the regions of RT and RP belong to the foreground of BW; for example the blue regions and some green regions in Fig. 4 middle belong to the background of BW.

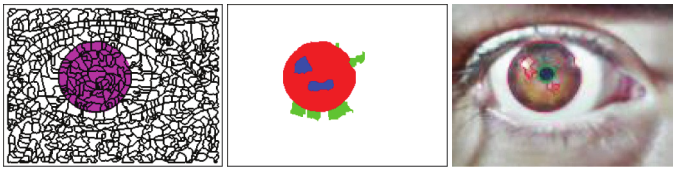

Fig. 4. From left to right: circle detected by circle fitting (purple) superimposed on the watershed transform, watershed regions totally overlapping the circle (red and blue) and partially overlapping the circle (green), red and green curves denote limbus and pupil boundaries. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To compute BW', the following features of the regions of RT and RP are taken into account: status in BW, representative color, and (total or partial) degree of overlapping with the circle.

Let H be the set of pixels within the circle and belonging to foreground of BW. Let K be the set of pixels that are outside the circle and have been assigned to the background in BW. Let c_{limbus} and c_{back} be the arithmetic means of the colors that the pixels of H and of K assume in the quantized image Q, respectively. The values c_{limbus} and c_{back} are taken as representing the average colors of limbus (iris and pupil) and of the background, respectively.

The first step of segmentation implies the analysis of the regions belonging to RT. In particular, any region R_i belonging to RT is ascribed to the foreground of BW' if at least one of the following conditions holds:

(1) $d(C_i, c_{back}) > d(C_i, c_{limbus})$.
(2) R_i belongs to the foreground in BW, and at least one of its adjacent regions belongs to RT, while no adjacent region belongs to RP.

Any region R_i belonging to RT and that is not yet been ascribed a status in BW', is analyzed again. In particular, R_i is ascribed to the foreground of BW' if the following condition is satisfied:

(3) It exists at least one adjacent region R_j assigned to the foreground of BW' such that $d(C_i, c_{back}) \quad d(C_i, C_j)$.

Otherwise, R_i assumes the status of background in BW'.

The second step of segmentation involves a change of W in correspondence of the regions belonging to RP. Any region R_i belonging to RP is divided into two sub-regions, respectively including pixels within the circle, and pixels outside the circle. The former sub-region replaces R_i in RP and the latter sub-region is ascribed to the background of BW'. Since W has been modified, the quantized image Q and BW are updated, in order to update also the values c_{limbus} and c_{back}. Any region R_i belonging to RP is assigned to the foreground if at least one of its adjacent regions has already been assigned to the foreground in BW', and the above condition (1) or condition (3) holds. Otherwise, the region R_i is assigned to the background in BW'. The boundary of the foreground in BW' is shown in Fig. 4 right as a red curve superimposed on I.

4.3. Pupil detection

Pupil detection is performed by applying a circle detection process similar to that adopted for limbus detection and described in Section 4.1. However, for pupil detection only the part of the eye image I delimited by the limbus boundary is taken into account. In fact, being the pupil always inside the iris, only the portion of I corresponding to foreground of BW' is considered as region of interest (ROI) for pupil detection.

The color ROI image is transformed into a gray level image and the Canny filter is applied to detect the edge pixels. Actually ten Canny filtered images are computed by adopting ten different thresholds $\{0.05, 0.10, 0.15, \ldots, 0.55\}$. In each Canny filtered image, circle fitting is accomplished for the connected components of extracted edge pixels including enough pixels (more than 5% of the pixels contained into the filtered image). The circle better fitting the pupil is found by taking into account only the circles entirely included in the ROI image. Two criteria are adopted to assign a score to a circle: homogeneity and separability. Homogeneity is evaluated in terms of the histogram of the region included by the circle. Separability is evaluated by adopting the same voting process described in Section 4.1. The score assigned to each circle is the sum of the scores on homogeneity and on separability. The best fitting circle is represented by the circle with the

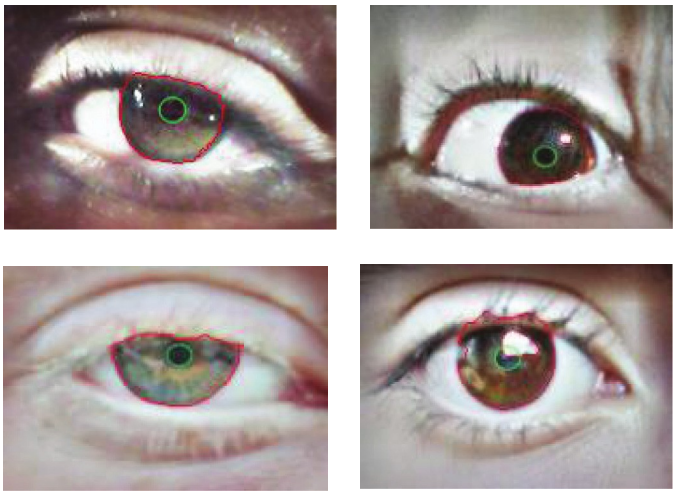

Fig. 5. Red and green curves denote limbus and pupil boundaries detected by BIRD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

maximal score, and it is shown, for the running example, as a green curve in Fig. 4 right.

In Fig. 5, some examples of the performance of BIRD are shown, where red and green curves delimit the limbus boundary and the pupil boundary, respectively.

5. Periocular region segmentation

When dealing with eye images captured at a distance, the quality of the iris texture is very poor due to noise, defocus and/or specular reflections. These image distortions represent a crucial problem as images acquired in visible spectrum are more affected by poor illumination conditions than those captured in infra-red. It has been demonstrated that features detected in the periocular region can be used as a soft-biometric to authenticate people [17] and significantly improve the recognition accuracy of iris recognition systems when dealing with iris images acquired in visible light [1,18]. By following this research line, BIRD also segments the periocular area and maps it to a rectangular region that undergoes a feature extraction procedure that is similar to the one adopted for the iris texture.

Starting from the approximating circle detected during the iris segmentation process, BIRD exploits the center coordinates $C_{Iris}(x_c, y_c)$ and radius r of the circle to construct two concentric ellipses E_1 and E_2 that enclose part of the area around the iris. Both E_1 and E_2 are centered in C_{Iris}, but they are characterized by different parameters (major and minor axes, a and b), which are respectively $(a, b) = (2 \times r, r)$ for E_1 and $(a, b) = (3 \times r, 2.5 \times r)$ for E_2 (see Fig. 6).

The area enclosed by E_1 and E_2 is mapped by BIRD to a rectangular region $I_{Periocular}$ by adopting a procedure that is similar to the Daugman' s rubber sheet model [5]. For each angle θ (ranging between 0 and 2π), two homologous points $P_1(\theta)$ and $P_2(\theta)$ are considered on E_1 and E_2 respectively, and pixels lying on the line ρ joining these points are mapped on the θth column of $I_{Periocular}$. The rectangular region resulting from this remapping procedure shows a resolution (256 × 32 pixel in this case), which depends on the granularity (discretization) chosen for θ and ρ (see Fig. 7).

6. Iris recognition

A further purpose of this paper is to investigate the interoperability of such an iris segmentation approach among different mobile devices.

To this aim, BIRD also implements an iris recognition technique that exploits both color and textural features to evaluate the similarity between two iris images. BIRD applies the rubber sheet model [5] to map the annular iris texture to a rectangular region. By passing from the Cartesian space to the polar space, the rubber sheet model is able to compensate for low off-axis effects and small displacements of the pupil circle. To remap the original annular iris region also offers the further advantage of simplifying the feature extraction process. BIRD extracts two kinds of features from the reshaped region, which characterize both the iris color and texture. First, it performs a color conversion from the RGB color space to HSV, with the aim of separating color information from that regarding luminance, which are then treated separately. Indeed, the cumulative sums (CSUM) [12] technique is applied to create an iris code for the structural pattern contained into the luminance channel, while a histogram based approach is used for color channels. As there are two color channels, BIRD combines them through a pixel by pixel product, so to obtain a single channel. A 64-bits histogram is computed on this channel and it is used as a color feature vector. The distance between two color vectors is computed by using the cosine dissimilarity, while textural feature vectors produced by the CSUM approach are compared by means of the Hamming distance.

The remapping procedure is applied to the periocular region to extract from $I_{Periocular}$ the same kind of color and textural features, which are compared according to the same distance measures. In order to assess the contribution of the periocular region to the recognition task, we considered two different scenarios. In the first one, only iris features are exploited by BIRD to match two irises, while in the second one it fuses both iris and periocular biometrics at a score level, by implementing a simple sum approach.

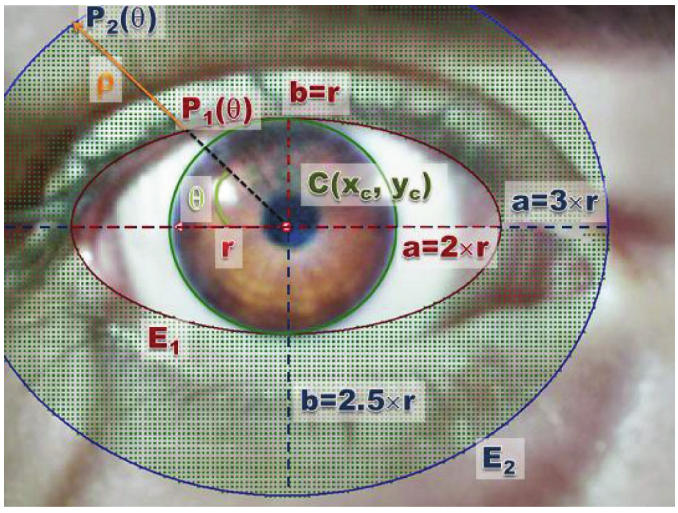

Fig. 6. The green grid denotes the periocular area delimited by ellipses E1 and E2 (red and blue curves) constructed around the circle detected during iris segmentation. The same color is used to show the ellipse E_1 (symmetrically E_2) and its corresponding parameters (semi-axis a and b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. The periocular region extracted by BIRD after applying the remapping procedure.

7. Experimental results

The experiments have been conducted on the MICHE dataset [19], which includes 3132 images of irises captured from 75 persons. The

Table 1
Segmentation accuracy measured in terms of percentage of error with respect to manual segmentation for the three tested approaches (ISIS, NICE-I, and BIRD) on probe/gallery images without color/illumination correction.

Device	Method		Iris *CX*	Iris *CY*	Iris rad.	Pupil *CX*	Pupil *CY*	Pupil rad.
IP5	ISIS out 33%	out	3 .16	5 .05	5 .09	3 .21	4 .47	2 .81
		nout	0 .62	1 .77	0 .79	0 .70	1 .24	0 .67
	NICE-I out 10%	out	1 .033	3 .11	1 .89	1 .26	2 .07	1 .03
		nout	0 .32	2 .77	0 .98	0 .55	1 .46	0.57
	BIRD out 12%	out	3 .30	4 .82	2 .76	3 .54	4 .66	1 .14
		nout	0 .61	1 .91	1 .21	1 .50	1 .56	0 .87
SG4	ISIS out 13%	out	4 .51	5 .33	6 .26	4 .61	5 .49	3 .57
		nout	0 .92	2 .06	1 .28	0 .75	1 .37	0 .75
	NICE-I out 33%	out	1 .27	3 .18	2 .03	1 .51	2 .26	1 .13
		nout	0 .33	2 .61	0 .93	0 .62	1 .53	0 .65
	BIRD out 36%	out	4 .09	4 .93	2 .79	4 .32	4 .87	1 .16
		nout	2 .52	2 .85	2 .14	1 .48	2 .62	0 .94
SGT	ISIS out 54%	out	5 .91	7 .57	10 .76	6 .02	7 .39	5 .94
		nout	3 .28	5 .69	5 .50	3 .46	5 .20	2 .47
	NICE-I out 21%	out	1 .60	3 .67	2 .71	1 .79	3 .47	1 .36
		nout	0 .36	2 .57	1 .12	0 .61	2 .00	0 .66
	BIRD out 34%	out	3 .54	4 .16	2 .55	4 .04	4 .36	1 .24
		nout	2 .86	2 .70	2 .22	2 .19	2 .60	0 .97

Table 2
Segmentation accuracy measured in terms of percentage of error with respect to manual segmentation for the three tested approaches (ISIS, NICE-I, and BIRD) on probe/gallery images with color/illumination correction.

Device	Method		Iris *CX*	Iris *CY*	Iris rad.	Pupil *CX*	Pupil *CY*	Pupil rad.
IP5	ISIS out 33%	out	2 .85	3 .98	4 .31	3 .03	3 .79	2 .04
		nout	0 .62	1 .75	0 .88	0 .71	1 .43	0 .70
	NICE-I out 46%	out	4 .84	6 .73	10 .93	4 .84	6 .99	5 .49
		nout	3 .27	4 .11	2 .51	3 .39	3 .34	1.07
	BIRD out 25%	out	3 .33	3 .69	1 .43	3 .51	3 .46	0 .82
		nout	0 .43	1 .58	0 .54	0 .58	1 .15	0 .59
SG4	ISIS out 32%	out	4 .12	5 .00	4 .39	4 .19	4 .77	1 .92
		nout	0 .62	1 .87	0 .83	0 .77	1 .35	0 .63
	NICE-I out 50%	out	5 .89	6 .69	12 .78	5 .76	6 .71	6 .46
		nout	3 .23	5 .22	2 .74	3 .15	4 .66	1 .27
	BIRD out 30%	out	3 .58	4 .07	1 .68	3 .70	4 .01	0 .81
		nout	0 .46	1 .69	0 .61	0 .68	1 .61	0 .58
SGT	ISIS out 54%	out	6 .03	7 .89	9 .46	6 .13	7 .92	5 .23
		nout	2 .67	4 .67	3 .69	2 .96	4 .59	1 .81
	NICE-I out 39%	out	7,45	6 .90	7 .86	7 .50	7 .62	3 .57
		nout	2 .90	4 .60	2 .40	3 .04	3 .97	0.93
	BIRD out 32%	out	3 .25	4 .18	2 .28	4 .40	6 .30	1 .09
		nout	1 .05	2 .06	0 .92	1 .30	2 .17	0 .83

images containing either the left or the right iris have been acquired by means of three mobile devices: (i) a Samsung Galaxy S4 (SG4), (ii) an iPhone 5 (IP5) and a Samsung Galaxy Tablet (SGT). Both the frontal and the rear camera were used with the two smartphones, whereas only the frontal camera was used with the tablet. At least 40 images per subject have been acquired during a unique session that corresponds to 4 snapshots per each of the 10 acquisition settings: (i) indoor/outdoor, (ii) frontal/rear, (iii) mobile device.

BIRD has been tested on a subset of MICHE (we refer to as MICH-Esub), which includes the first two images per subject in all 10 acquisition conditions, resulting in a total of 1500 iris snapshots. A manual segmentation of the whole MICHEsub dataset has been performed, in order to create a ground truth, which has been used for comparing BIRD performances with those of two state of the art methods, which are ISIS [6] and our re-implementation of that proposed by Tan et al. [15] that will be referred as NICE-I from now on. Experiments have been carried out to assess performances in terms of both segmentation precision and recognition accuracy.

To evaluate the segmentation precision, the segmented iris masks provided by the automatic methods ISIS, NICE-I and BIRD are compared with those provided by the manual segmentation into the ground truth, in terms of percentage of errors measured in pixels. In more details, circle fitting is applied to the binary masks in order to approximate the center and radius of both iris and pupil. The discrepancy in pixels is then computed between homologous parameters (centers and radii) extracted from each ground truth mask and the corresponding ones obtained by the testing approaches. A global estimation is obtained by averaging the discrepancy values over all images, after they have normalized with respect to the image resolution and multiplied by 100. There are cases in which the detection is completely wrong, so we can consider them as outliers when computing the global segmentation precision. Outliers can be detected as they generally provide a segmentation error higher than a threshold ε_{out}, which is computed as $\varepsilon_{out} = 6m_d$, where m_d is the median value computed over all the errors in the subset of tested images. Since the process of color/lighting correction plays an important role for the subsequent stages of segmentation, precision was evaluated in two cases, in which it is included or not in the process of preprocessing. Tables 1 and 2 show the results obtained in terms of accuracy of the three methods of segmentation (ISIS, NICE-I and BIRD) with three different devices (IP5, SG4 and SGT), excluding or including the color/lighting correction process from the pipeline. In both Tables 1 and 2, we provide results accounting for both cases, in which outliers have been considered or not in the evaluation. It comes out from these

Table 3
Recognition accuracy measured in terms of decidability, EER and AUC for the three tested approaches (ISIS, NICE-I and BIRD) on probe/gallery images acquired by the same mobile device.

Method		Device								
		IP5			SG4			SGT		
		Dec.	EER	AUC	Dec.	EER	AUC	Dec.	EER	AUC
ISIS	I	0 .562	0.393	0.651	0 .938	0.335	0 .742	0 .692	0 .388	0 .664
	P	0 .626	0.414	0.637	0 .894	0.349	0 .707	0 .681	0 .402	0 .671
	F	0 .634	0.446	0.632	1 .039	0.336	0 .729	0 .765	0 .378	0 .673
NICE-I	I	0 .787	0.386	0.6674	0 .869	0.409	0 .714	0 .672	0 .413	0 .681
	P	0 .366	0.470	0.581	1 .087	0.334	0 .753	0 .829	0 .334	0 .715
	F	0 .366	0.470	0.581	1 .086	0.334	0 .753	0 .829	0 .334	0 .715
BIRD	I	1.151	0.266	0.784	1 .083	0.314	0 .778	0 .880	0 .326	0 .707
	P	0 .949	0.295	0.736	1 .340	0.278	0 .798	1 .159	0 .342	0 .789
	F	1 .156	0.260	0.764	1 .290	0.322	0 .775	1 .103	0 .277	0 .765

results that color correction improves the segmentation accuracy of both ISIS and BIRD almost always. There are only few cases in which a very low decrease in precision is registered after color correction of the input image. In particular this mostly happens for ISIS on images acquired by SGT, and can be ascribed to the lower resolution of the camera mounted on this device. Indeed, on low resolution images after color correction ISIS undergoes more difficulties to correctly locate the pupil, so undermining all the subsequent steps. On the contrary, color correction seems to negatively affect the performance of NICE-I. This can be explained by considering that the first step in the NICE-I pipeline consists in a coarse clustering of pixels with low brightness to be considered as iris candidates. The color/illumination correction produces significant changes in the histogram of the iris image, so that clustering fails even jeopardizing subsequent steps in the remaining pipeline. Indeed, by inspecting images for which NICE-I fails, it comes out that the method do not output any circle neither for the iris nor for the pupil. This is also the motivation for which the number of outliers produced by NICE-I increase when color/illumination correction is included.

It is worth to notice that the amount of outliers produced by a testing approach is in itself a measure of its effectiveness. From results in Table 2, we observe that BIRD always outperforms ISIS in terms of both segmentation precision and number of outliers. From both tables, it is also noted that in most cases, when outliers are kept out of the computation, the increment in segmentation precision observed for BIRD is higher than that obtained by ISIS. However, BIRD produces more outliers than NICE-I. This happens because after the merging operation the iris boundary presents some cuts, which mislead the curvature analysis process and result in a mislocated iris/pupil circle. Furthermore, BIRD provides a higher accuracy in locating the iris than the pupil; this is because pupil is smaller than iris and then it is more affected by noise factors like bad illumination and specular reflection. This can also explain why BIRD outperforms ISIS. Indeed, BIRD first locates iris and then searches for pupil by restricting the region of interest (for the pupil) to the area included into the iris circle. On the contrary, ISIS performs in the real opposite way by first searching for pupil and then extending its searching window to neighboring regions. This jeopardizes the performances of ISIS, since for several images into the dataset the pupil is not completely visible, due to large occluded areas or severe out of focus conditions, which are caused by user holding the mobile device in his/her hands during the acquisition.

Results in Tables 1 and 2 show that NICE-I provide better performance when color/illumination correction is not included in the segmentation process. In this case, it also outperforms BIRD in terms of segmentation accuracy. Conversely, color/illumination significantly improves the performance of BIRD, so it must be considered as a fundamental part of the segmentation process implemented by BIRD. Indeed, by comparing results obtained by BIRD in Table 2 (its best performance) with those of NICE-I in Table 1 (its best performance) it is worth to notice that in most cases BIRD outperforms NICE-I.

Several experiments have also been conducted to assess recognition performances of testing approaches (ISIS, NICE-I and BIRD). In each test, two images per subject have been considered to build a gallery set (used to enroll subjects) and a probe set (used to test the recognition system). System performance was measured in terms of decidability, equal error rate (EER) and area under curve (AUC), the former being calculated according to the definition given by Proenca and Alexandre [10].

More specifically, given a set of user templates V, an "all-against-all" comparison provides a set of dissimilarity intra-class measures $D_I = \{D_{I1}, D_{I2}, \ldots, D_{Ik}\}$ and inter-class measures $D_E = \{D_{E1}, D_{E2}, \ldots, D_{Em}\}$, depending on whether templates belong to the same class or not. The decidability $d(D_I, D_E)$ of the set V is computed as:

$$d(D_I, D_E) = \frac{|avg(D_I) - avg(D_E)|}{\frac{1}{2}(\sigma(D_I)^2 + \sigma(D_E)^2)}$$

where the value of $d(D_I, D_E)$ ranges into [0,). Both EER and AUC are single descriptions related to the receiving operating characteristic curve (ROC), which is a graphical plot that characterizes the performance of a biometric verification system as its discrimination threshold is varied. The EER represents the value where false accept rate and false reject rate equal and can be considered such as a steady state point for the system. While ROC is a two-dimensional representation of a model's performance, the AUC distils this information into a single scalar and represents the probability that a classifier ranks a randomly chosen positive instance higher than a randomly chosen negative one.

In the first experiment, both gallery and probe images have been acquired with the same mobile device using the rear camera. The results were produced for iris recognition, for periocular region recognition and for merging of both. In Table 3, we report the numeric values for the various performance measures.

From the obtained results, it comes out that the segmentation approach implemented by BIRD induces a higher performance than ISIS in all cases. The reason underlying the larger increase in recognition accuracy produced by BIRD is that the precision it offers in segmenting the iris boundary (limbus) is generally higher than that provided by ISIS. Indeed, ISIS only approximates both iris and pupil boundaries by circles, while circle fitting represents only a preliminary step for the segmentation procedure implemented by BIRD. Moreover, BIRD further refines the representation of the limbus by considering subregions, which cross the iris circle and re-assigning them to the foreground/background depending on the classification of neighboring regions. Since results provided by BIRD outperform those obtained by ISIS, only the former has been considered in the following experiments, whose main goal is to assess also the robustness of BIRD with

Table 4
Recognition accuracy measured in terms of decidability, EER and AUC for the BIRD approach (without periocular information) on probe/gallery images acquired by different mobile devices.

Method		Gallery								
		IP5			SG4			SGT		
		Dec.	EER	AUC	Dec.	EER	AUC	Dec.	EER	AUC
Probe	IP5	1 .151	0 .266	0.784	0 .834	0 .333	0.745	0 .016	0 .485	0 .525
	SG4	0 .255	0 .414	0.601	1 .083	0 .314	0.778	0 .887	0 .729	0 .248
	SGT	0 .213	0 .534	0.451	0 .317	0 .430	0.593	0 .880	0 .326	0 .707

Table 5
Recognition accuracy measured in terms of decidability, EER and AUC for the BIRD approach (with periocular information) on probe/gallery images acquired by different mobile devices.

Method		Gallery								
		IP5			SG4			SGT		
		Dec.	EER	AUC	Dec.	EER	AUC	Dec.	EER	AUC
Probe	IP5	1 .156	0 .260	0 .764	0 .858	0 .309	0 .738	0 .374	0 .397	0 .614
	SG4	0 .375	0 .397	0 .630	1 .290	0 .322	0 .775	0 .296	0 .565	0 .406
	SGT	0 .124	0 .584	0 .466	0 .502	0 .367	0 .642	1 .103	0 .277	0 .765

Table 6
Recognition accuracy measured in terms of decidability, EER and AUC for the BIRD approach on probe/gallery images captured in different acquisition settings (outdoor/indoor).

Method		Device								
		IP5			SG4			SGT		
		Dec.	EER	AUC	Dec.	EER	AUC	Dec.	EER	AUC
BIRD	I	0.529	0 .401	0 .652	0.583	0 .386	0.665	0 .078	0 .472	0 .535
	P	0.246	0 .470	0 .581	0.633	0 .373	0.672	0 .596	0 .391	0 .658
	F	0.506	0 .399	0 .646	0.647	0 .367	0.689	0 .432	0 .430	0 .619

respect to adverse acquisition conditions as well as its interoperability among different mobile devices. It can be observed that, in many cases, the results provided by iris and periocular area are comparable, with regard to the other measures of performance. Moreover, it can be observed that while in some cases the fusion leads to an improvement in others it remains, albeit slightly, below the better of the two, when it is considered individually. This result can be explained by considering that a simple sum has been adopted for the score fusion step. However, resorting to a more refined approach [14] or postponing the fusion process to the decision level [9] could lead to larger improvements.

In the second experiment, the probe and gallery sets come from different devices, which show significant variation in image resolution, as they are equipped with very different sensors. Results when only the iris region is considered are reported in Table 4, while those obtained including also periocular information are reported in Table 5.

From results in Tables 4 and 5, it emerges that testing on cross-datasets is particularly difficult when images captured with an IP5 mobile device are used as probe. This can be explained by visually inspecting images acquired by different devices. Indeed, before the lighting/color correction, IP5 iris images show a higher sharpness and brightness than those acquired by both SG4 and SGT, independently of the used camera (either frontal or rear). Even so, BIRD achieves a better accuracy than ISIS further confirming that a higher precision in segmenting iris images greatly affects the recognition process.

The last experiment is devoted to assess performances of both testing approaches with respect to the capturing settings. In particular, images acquired indoor constitute the gallery, while those captured outdoor (probe images) are used to test the methods. This scenario is more plausible than the opposite one, since one can assume that the enrollment of a person is rarely performed (may be just once and in very controlled conditions), while testing occurs more often either in controlled or adverse conditions. Numerical results in terms of decidability, EER and AUC are reported in Table 6.

When only iris features are used to recognize people, there is no appreciable gap in performance of BIRD when using different device. However, when periocular features are also integrated, the accuracy obtained with IP5 remains almost the same, while that achieved with SG4 and SGT slightly improves.

Another important aspect in the evaluation of the performance of BIRD is the execution time given by the sum of the time needed to carry out the various stages of processing. In order to provide an estimate of the impact that each one has on the total time, the time necessary to accomplish each stage is reported as a percentage of the weight that it has on the entire running time:

(i) color/illumination correction (20%)
(ii) image resizing (4%)
(iii) median filtering (1%)
(iv) watershed (11%)
(v) merging (6%)
(vi) binarization (2%)
(vii) iris detection (40%)
(viii) pupil detection (16%)

The tests were carried out on a 64-bit system with Genuine Intel® processor U7300 1.3 GHz, with 4 GB of RAM and the total time was about 1.2 s, which is comparable with that required by ISIS while it is of an order of magnitude smaller than that required by NICE-I.

8. Conclusions and future work

This paper presents a new technique, named BIRD for segmenting iris images, which is based on the watershed transform. It is

principally devised for mobile devices, as it is designed to be robust with respect to many kinds of image distortions that typically occur when the user is holding a smartphone/tablet in his/her hands to perform acquisition (out of focus, occlusions, motion blur, non-uniform illumination, off-axis angles). BIRD exploits the watershed transform at two different levels. First it applies the watershed transform to binarize the image and then it further analyze the sub-regions produced during the first application of the watershed in order to refine their classification as either foreground or background. Starting from a circle fitting of the iris boundary, BIRD also segments the periocular region and polarizes it in a way that is similar to the rubber sheet model commonly adopted for the iris annular region, but specifically devised for this aim.

Both textural and color information are extracted from iris, as well as from the periocular region, in order to match biometric features, with the aim of quantitatively assessing the improvement that a better segmentation can induce in terms of decidability. BIRD has been tested on iris images contained in the MICHE dataset, which has been captured with three different mobile devices both in indoor and in outdoor settings. Results have been compared with those provided by two state of the art techniques, namely ISIS and NICE-I, showing that iris detection/recognition on mobile device is a challenging task, but also underlining the high potential of the proposed technique.

Given the high variability of the images captured by mobile devices, the segmentation process implemented by BIRD leaves room, however, for further improvement, especially for what concerns the periocular area. In particular, a future development of this method will aim to refine the segmentation technique of the periocular area, explicitly addressing the identification of the reflections and the region occupied by the eyelids. Another important point is the reduction of the number of outliers BIRD produces.

References

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

[2] K.W. Bowyer , K.P. Hollingsworth , P.J. Flynn , Image understanding for iris biometrics: a survey , Comp. Vis. Image Und. 110 (2) (2008) 281–307 .

[3] K.W. Bowyer , K.P. Hollingsworth , P.J. Flynn , A survey of iris biometrics research: 2008–2010 , in: M. Burge, K.W. Bowyer (Eds.), Handbook of Iris Recognition, Springer 2013, pp. 15–54 .

[4] M. Burge , K.W. Bowyer , Handbook of Iris Recognition , Springer, London, UK, 2013 ; K. Siddiqi, S.M. Pizer (Eds.), Medial Representations: Mathematics, Algorithms and Applications, Springer, 2008 .

[5] J.G. Daugman , High confidence visual recognition of persons by a test of statistical independence , IEEE Trans. Pattern Anal. Mach. Intell. 11 (15) (1993) 1148–1161 .

[6] M. De Marsico , M. Nappi , D. Riccio , IS_IS: Iris Segmentation for Identification Systems , in: Proceedings of the International Conference on Pattern Recognition (ICPR 2010) 2010, pp. 2857–2860 .

[7] M. Frucci , M. Nappi , D. Riccio , G. Sanniti di Baja , Using the Watershed Transform for Iris Detection , in: Proceedings of the International Conference on Image Analysis and Processing (ICIAP 2013) 2013, pp. 269–278 .

[8] M. Frucci , M. Nappi , D. Riccio , Watershed Based Iris SEgmentation , in: Proceedings of the 5th Mexican Conference on Pattern Recognition (MCPR 2013) 2013, pp. 204–212 .

[9] K. Kryszczuk , J. Richiardi , P. Prodanov , A. Drygajlo , Reliability-based decision fusion in multimodal biometric verification systems , EURASIP J. Adv. Sig. Process. (2007) 1–9 Article ID 86572 .

[10] H. Proenca , L.A. Alexandre , Towardcovert iris biometric recognition:experimental results from the NICE contests , IEEE Trans. Inform. Forensics Security 7 (2) (2012) 798–808 .

[11] H. Proença , G. Santos , J.C. Neves , Using Ocular Data for Unconstrained Biometric Recognition, Face Recognition in Adverse Conditions , in: M. De Marsico, M. Nappi, M. Tistarelli (Eds.), IGI Global 2013 .

[12] E. Raj , M. Chirchi , R.D. Kharadkar , Biometric iris recognition for person identification using cumulative sum algorithm , Int. J. Sci. Eng. Res. 3 (5) (2012) .

[13] J.B.T.M. Roerdink , A. Meijster , The watershed transform: definitions, algorithms and parallelization strategies , Fundam. Inform. 41 (1–2) (2001) 187–228 .

[14] N. Srinivas , K. Veeramachaneni , L.A. Osadciw , Fusing Correlated Data from Multiple Classifiers for Improved Biometric Verification , in: 12th International Conference on Information Fusion 2009, pp. 1504–1511 .

[15] T. Tan , Z. He , Z. Sun , Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition , Image Vis. Comput. 28 (2010) 223–230 .

[16] G. Taubin , Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations, with applications to edge and range image segmentation , IEEE Trans. Pattern Anal. Mach. Intell. 13 (13) (1991) 1115–1138 .

[17] D.L. Woodard , S.J. Pundlik , J.R. Lyle , P.E. Miller , Periocular Region Appearance Cues for Biometric Identification , in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2010, pp. 162–169 .

[18] D.L. Woodard , S.J. Pundlik , P. Miller , R. Jillela , A. Ross , On the Fusion of Periocular and Iris Biometrics in Non-ideal Imagery , in: International Conference on Pattern Recognition 2010, pp. 201–204 .

[19] http://biplab.unisa.it/MICHE/database/ .