

Iris Recognition: A Method To Segment Visible Wavelength Iris Images Acquired On-The-Move and At-A-Distance

Hugo Proença

Dep. of Computer Science, IT - Networks and Multimedia Group
Universidade da Beira Interior, Covilhã, Portugal
Email: hugomcp@di.ubi.pt

Abstract. The dramatic growth in practical applications for iris biometrics has been accompanied by many important developments in the underlying algorithms and techniques. Among others, one of the most active research areas concerns about the development of iris recognition systems less constrained to users, either increasing the image acquisition distances or the required lighting conditions. The main point of this paper is to give a process suitable for the automatic segmentation of iris images captured at the visible wavelength, on-the-move and within a large range of image acquisition distances (between 4 and 8 meters). Our experiments were performed on images of the UBIRIS.v2 database and show the robustness of the proposed method to handle the types of non-ideal images resultant of the aforementioned less constrained image acquisition conditions.

1 Introduction

Being an internal organ, naturally protected, visible from the exterior and supporting contactless data acquisition, the human iris has, together with the face, the potential of being covertly imaged. Several issues remain to achieve deployable covert iris recognition systems and, obviously, these type of systems will constitute a tradeoff between data acquisition constraints and recognition accuracy. This area motivates growing interests on the research community and constituted the scope of a large number of recent publications.

It is expectable that far less constrained image acquisition processes increase the heterogeneity of the captured images, according to varying lighting conditions, subjects' poses, perspectives and movements. In this context, the image segmentation stage plays a major role, as it is the one that more directly should handle this heterogeneity. Also, as it is one of the earliest stages of the complete recognition process, it acts as basis of any further stages and any failure will compromise the success of the whole process.

Figure 1 establishes a comparison between images that typically result of constrained image acquisition processes, on the near infra-red wavelength (figure 1a) and images acquired under less constrained imaging conditions, at-a-distance and on-the visible wavelength (figure 1b). Apart evident differences in the image homogeneity, several types of data obstructing portions of the iris texture in the visible wavelength



(a) Near infra-red image, acquired under high constrained conditions (ICE database [1]). (b) Visible wavelength image, acquired at-a-distance and on-the-move (UBIRIS.v2 database [2]).

Fig. 1. Illustration of the typical differences between close-up iris images acquired under high constrained conditions in the near infra-red wavelength (figure 1a) and images acquired in the visible wavelength, within less constrained conditions (figure 1b).

image can be observed, whose increase the challenges of performing accurate biometric recognition. Also, less constrained acquisition protocols lead to the appearance of non-frontal, defocused or motion blurred images.

This work focuses on the segmentation of visible wavelength close-up iris images, captured at-a-distance and on-the-move, under varying lighting conditions and with minimal image acquisition constraints. It can be divided into three parts: at first, we overview the most significant iris segmentation methods, specially those recently published, and establish some common and distinguishable characteristics between them. Later, we empirically describe some of the reasons that makes those methods less suitable for the type of images we aim to deal with. Finally, we give a new iris segmentation method based on the neural Pattern Recognition paradigm that, as our experiments indicate, is suitable to deal with the aforementioned type of images.

The remaining of this paper is organized as follows: Section 2 briefly summarizes the most cited iris segmentation methods, emphasizing those most recently published. A detailed description of the proposed method is given in Section 3. Section 4 reports our experiments and discusses the results and, finally, Section 5 gives the conclusions and points some directions for our further work.

2 Iris Segmentation

The analysis of the most relevant iris segmentation proposals allowed us to identify two major strategies to perform the segmentation of the iris: use a rigid or deformable iris template or use its boundary. In most cases, the boundary approach is very similar to the early proposal of Wildes [3]: it begins by the construction of an edge-map, followed by the application of some geometric form fitting algorithm. The template-based strategies are in general more specific, although having as common point the maximization of some energy model that localizes both iris borders, as originally proposed by Daugman [4]. These methods were though to operate in noise-free close-up iris images

captured in the near infra-red wavelength and, specifically the Daugman's integrodifferential operator, proved their effectiveness on multiple deployed systems that operate in constrained imaging conditions.

The increase of segmentation robustness to several types of non-ideal images motivated a large number of proposals in the last few years. However, the large majority of these methods operate in near infra-red images, whose typically have higher contrast between the pupil and the iris regions and induces the usual option of start by the segmentation of the pupillary border. Oppositely, visible wavelength images usually have less contrast between the pupil and the iris, which explains the inversion in the order of the borders' segmentation.

Regarding the basis methodologies, various innovations were proposed, as the use of active contour models, either geodesic (e.g., [5]), based on Fourier series (e.g., [6]) and on the snakes model (e.g., [7]). Here, the previous detection of the eye is a requirement to properly initialize contours and the heavy computational requirements can also be regarded as a weak point. Also, modifications to known form fitting processes were proposed, essentially to deal with off-angle images (e.g., [8] and [9]) and to improve performance (e.g., [10] and [11]).

The detection of non-iris data that obstructs the discriminating information motivated the use of parabolic, elliptical and circular models (e.g., [12], and [11]) and the modal analysis of histograms [6]. In this compass, several authors constraint the success of their methods to image orthogonality, to the non-existence of significant iris obstructions or to the appearance of corneal reflections in specific image regions.

3 Proposed Method

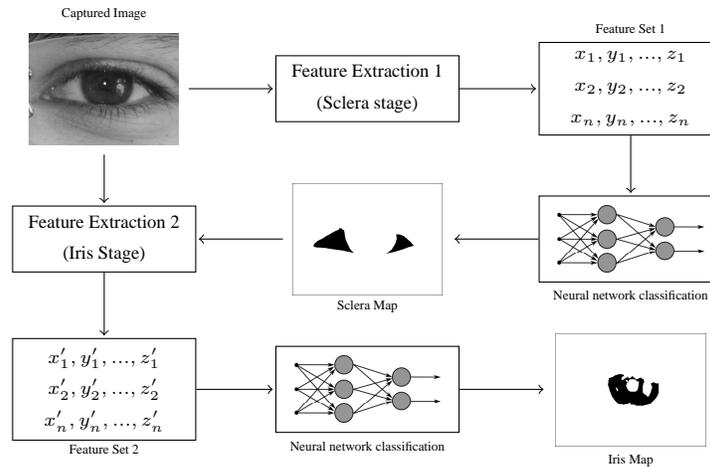


Fig. 2. Block diagram of the proposed iris segmentation method.

Figure 2 gives the block diagram of the proposed segmentation process, that can be divided into two major stages: sclera detection and iris segmentation. Having found that the sclera region usually remains as the most distinguishable under varying lighting conditions, we propose a feature extraction stage that will provide enough discriminant information to localize the sclera. Later, we take profit of the mandatory adjacency between the sclera and the iris and, together with a new feature set extracted from the original image, perform the detection of the regions that correspond to the noise-free iris pixels. It should be stressed that this process comprises three tasks that are typically separated in the specialized literature: iris detection, segmentation and detection of noisy regions. As it is shown in the experiments section, starting from a relatively small set of manually classified images that constitute the learning set, it is possible to use machine learning methods that will robust and quickly discriminate between the noise-free iris regions (used in the subsequent recognition stages) and the remaining data. In the following, we detail each feature extraction stage and the used classification models.

3.1 Feature Extraction

Regarding the feature extraction, we had a primary concern: to exclusively evaluate features that are possible to compute in a single image scan, which is crucial to enable the application of the method to real-time applications. Previously, Viola and Jones [13] proposed a set of simple features, reminiscent of Haar basis functions, and used an intermediate image representation to compute them in a single image scan. Based on their definition we propose the extraction of a set of central moments within small image regions, based on the pixels intensity in different color spaces.

For a given image I , Viola and Jones defined an *integral image* II

$$II(x, y) = \sum_1^x \sum_1^y I(x', y') \quad (1)$$

where x and y denote respectively the image column and row. Also, they proposed a pair of recurrences to compute the integral image in a single image scan

$$s(x, y) = s(x, y - 1) + I(x, y) \quad (2)$$

$$II(x, y) = II(x - 1, y) + s(x, y) \quad (3)$$

Based on the concept of integral image, the average intensity μ of the pixels within any rectangular region R_i , delimited by its upper-left (x_1, y_1) and bottom-right (x_2, y_2) corner coordinates, can be obtained accessing exclusively four array references. Let $T_i = (x_2 - x_1) \times (y_2 - y_1)$ be the total of pixels of R_i .

$$\begin{aligned}
\mu(R_i) &= \frac{1}{T_i} \sum_{x_1}^{x_2} \sum_{y_1}^{y_2} I(x, y) \\
&= \frac{1}{T_i} \left(\sum_1^{x_2} \sum_{y_1}^{y_2} I(x, y) - \sum_1^{x_1} \sum_{y_1}^{y_2} I(x, y) \right) \\
&= \frac{1}{T_i} \left(\sum_1^{x_2} \sum_1^{y_2} I(x, y) - \sum_1^{x_2} \sum_1^{y_1} I(x, y) \right. \\
&\quad \left. - \left(\sum_1^{x_1} \sum_1^{y_2} I(x, y) - \sum_1^{x_1} \sum_1^{y_1} I(x, y) \right) \right) \\
&= \frac{1}{T_i} \left(II(x_2, y_2) + II(x_1, y_1) - II(x_2, y_1) - II(x_1, y_2) \right) \quad (4)
\end{aligned}$$

Similarly, the standard deviation σ of the intensity of the pixels within R_i is given by

$$\begin{aligned}
\sigma(R_i) &= \sqrt{\frac{1}{T_i} \sum_{x_1}^{x_2} \sum_{y_1}^{y_2} (I(x, y) - \mu)^2} \\
&= \sqrt{\frac{1}{T_i} \sum_{x_1}^{x_2} \sum_{y_1}^{y_2} (I(x, y)^2 - 2 I(x, y) \mu + \mu^2)} \\
&= \sqrt{\frac{1}{T_i} \left(\sum_{x_1}^{x_2} \sum_{y_1}^{y_2} I(x, y)^2 - 2 \mu \sum_{x_1}^{x_2} \sum_{y_1}^{y_2} I(x, y) + T_i \mu^2 \right)} \\
&= \sqrt{\frac{1}{T_i} (II(x, y)^2 - 2 \mu II(x, y) + T_i \mu^2)} \quad (5)
\end{aligned}$$

where μ is given by (4), $II(x, y)$ is obtained from (3) and $II(x, y)^2$ is similarly obtained, starting from an image where the intensity values appear squared.

3.2 Sclera Stage

Based on the described average (μ) and standard deviation (σ) values within image regions, on the detection of the sclera we use a feature set with 11 components. For each image pixel we extract $\{x, y, h_{0,3,7}^{\mu,\sigma}(x, y), u_{0,3,7}^{\mu,\sigma}(x, y), cr_{0,3,7}^{\mu,\sigma}(x, y)\}$, where x and y denote the pixel's position, $h(\cdot, \cdot)$, $u(\cdot, \cdot)$, and $cr(\cdot, \cdot)$ denote regions of the hue, chrominance and red chroma image components and the subscripts denote the respective radius values of such regions, centered at the respective pixel. The used color components were empirically selected, based on observations of their discriminating capacity between the sclera and the remaining data, as illustrated in figure 3.

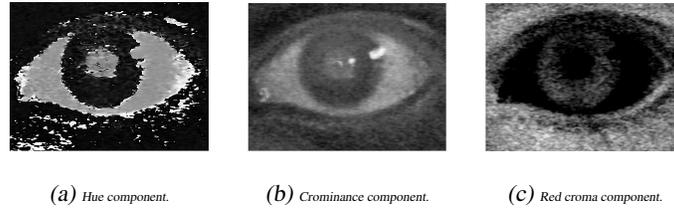


Fig. 3. Illustration of the discriminating capacity of the color components used in the detection of the sclera regions. It is evident that pixels belonging to the sclera have respectively higher (figures 3a and 3b) and lower (figure 3c) intensity values than the remaining pixels. Also, we observed that this separability tends to remain stable, even on high heterogeneous images as those that constitute the scope of this work.

3.3 Iris Stage

In the detection of the noise-free iris regions we also used the previously described average (μ) and standard deviation (σ) values, together with the information that came from the previous sclera detection stage, taking profit of the mandatory adjacency between the iris and the sclera. Here, our main concern was to use components of various color spaces that maximize the separability between the sclera and the iris. For each image pixel we extracted $\{x, y, s_{0,3,7}^{\mu,\sigma}(x, y), u_{0,3,7}^{\mu,\sigma}(x, y), sc_{\leftarrow, \rightarrow, \uparrow, \downarrow}(x, y)\}$, where $s(\cdot, \cdot)$, and $u(\cdot, \cdot)$ are regions of the saturation and color chrominance image components and the subscripts denote the respective radius values of such regions, centered at the respective pixel. Again, the used color components were empirically selected, based on observations of their respective discriminating capacity between the sclera and the iris, as illustrated in figure 4. $sc(\cdot, \cdot)$ denotes a feature map that measures the proportion of pixels belonging to the sclera in the left (\leftarrow), right (\rightarrow), upper (\uparrow) and bottom (\downarrow) directions, regarding the reference pixel (x, y) . This maps are specially relevant to provide information about the relative localization between the iris and the sclera, as it is illustrated in figure 4b.

$$sc_{\leftarrow}(x, y) = \mu\left(R_{sc}((1, y - 1), (x, y))\right) \quad (6)$$

$$sc_{\rightarrow}(x, y) = \mu\left(R_{sc}((x, y - 1), (W, y))\right) \quad (7)$$

$$sc_{\uparrow}(x, y) = \mu\left(R_{sc}((x - 1, 1), (x, y))\right) \quad (8)$$

$$sc_{\downarrow}(x, y) = \mu\left(R_{sc}((x - 1, y), (x, H))\right) \quad (9)$$

where $\mu(\cdot)$ is given by (4) and $R_{sc}((\cdot, \cdot), (\cdot, \cdot))$ denote regions of the sclera map (figure 4a) delimited respectively by its top-left and bottom-right corner coordinates. W and H are respectively the image width and height.

As we describe in the next section, for the purpose of classification we used feed-forward neural networks, whose are known to be extremely fast classification models.

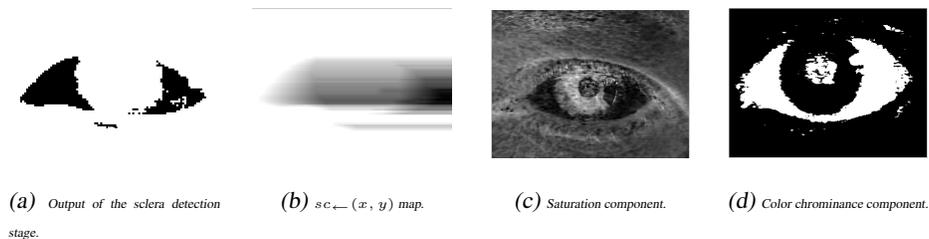


Fig. 4. Features used in the detection of the iris regions.

Thus, apart the accuracy and robustness factors, when comparing with the large majority of the iris segmentation methods described in section 2, the computational performance of the proposed method is regarded as a significant advantage.

3.4 Supervised Machine Learning Process

In both classification stages of the proposed segmentation method we followed the neural pattern recognition paradigm. This is justified by the networks ability to discriminate data in high complex hyperspaces, providing good generalization capacity and usually without requiring any user-parameterized thresholds.

We used multilayered perceptron feed-forward networks with one hidden layer, varying in our experiments the number of neurons of the hidden layer. We adopted the convention that the input nodes are not counted as a layer. All the used networks have as many neurons in the input layer as the dimension of the feature space and a single neuron in the output layer, due to the desired binary classification task.

Our choice is justified by three fundamental learning theory issues: model capacity, computational and sample complexity. At first, regarding model capacity, it is known that this type of networks can form arbitrary complex decision boundaries. Also, they are accepted as high performance classification models, which is not affected by the size of the training data. Finally, regarding the sample complexity, the use of the backpropagation learning algorithm propitiates good generalization capabilities using a relatively small set of images in the learning stages, as it is detailed in the experiments section.

4 Experiments

This section describes the experiments that were performed in the evaluation of the proposed classification method. We detail the used data sets and the process that enabled us to automatically obtain the error rates.

4.1 Data Sets

Our experiments were performed in images of the UBIRIS.v2 [14], a multi-session iris images database which - singularly - contains data captured in the visible wavelength, at-a-distance and on on-the-move subjects, being its images acquired within

non-constrained and varying lighting conditions. The significantly higher range of distances between the subjects and the imaging framework (from four to eight meters) is one of the major distinguishable points between the UBIRIS.v2 database and others with similar purposes. Through visual inspection, fourteen different factors that deteriorate the image quality were detected and classified into one of two categories: *local* or *global*, as they affect exclusively image regions or the complete image. The first category comprises iris obstructions, reflections and partial images, while the later comprises the poor focused, motion-blurred, rotated, off-angle, improper lighting and out-of-iris images. A comparison between a close-up iris image with good quality (upper-left image) and images that contain at least one of the aforementioned noise factors is given in figure 5.

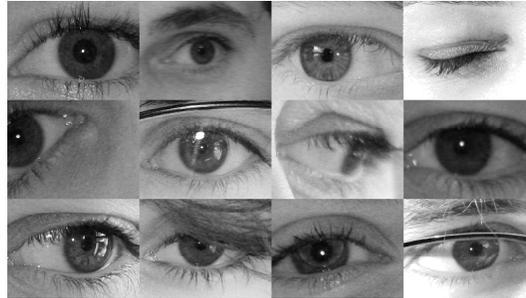


Fig. 5. Examples of close-up iris images acquired at varying distances (between four and eight meters), from on-the-move subjects, under high dynamic lighting conditions and without requiring subjects cooperation.

4.2 Ground Truth

In the evaluation of the proposed method we used the data sets delivered to participants of the NICE.I contest [2], whose are part of the complete UBIRIS.v2 iris image database. Both the learning and test data sets comprise 500 close-up iris images and 500 correspondent binary maps that were made by humans and distinguish between the noise-free regions of the iris and all the remaining types of data.

Images have fixed dimensions of 400×300 pixels, giving a total of 60 000 000 pixels per data set. Considering the process of noise-free iris segmentation as a binary classification task, this value allowed us to obtain 95% confidence intervals for the results given in this paper of approximately $\pm 1,29\% \times 10^{-2}$.

4.3 Learning Algorithms

Both learning stages of our classification models were based in the backpropagation algorithm. Originally, this learning strategy updates the network weights and biases in

the direction of the negative of the gradient, the direction in which the performance function E decreases most rapidly, being E a squared error cost function given by $\frac{1}{2} \sum_{i=1}^p \|y_i - d_i\|^2$, where y_i and d_i are respectively the network and the desired outputs and p the number of train patterns given to the network in the learning stage. There are many variations of the backpropagation algorithm, which fundamentally aim to increase the learning performance, resulting in the network convergence performance from ten to one hundred times faster.

Typical variants fall into two categories: the first one uses heuristic techniques, as the momentum technique or the varying learning rates. The second category uses standard numerical optimization techniques, as the search across the conjugate directions (with Fletcher-Reeves [15] or Powell-Beale [16] updates) or quasi-Newton algorithms (Broyden, Fletcher, Goldfarb, and Shanno [17] and one-secant [18] update rules) that, although based on the Hessian matrix to adjust weights and biases do not require the calculation of second derivatives, this matrix at each iteration of the algorithm.

In the following section we give results about the error rates obtained by each of the used backpropagation variants, both in the learning and classification stages, regarding the number of images used in the learning process and the networks' topology.

4.4 Results and Discussion

The method proposed in this paper has - at least - 3 parameters that have impact in its final accuracy: the used learning algorithm, the networks' topology and the amount of data (number of images) used to learn. As an exhaustive search for the optimal configuration leads to a 3D search and an impracticable number of possibilities to evaluate, we decided to start with the selection of the most suitable backpropagation learning algorithm for this type of problem. We built a set of neural networks with what we considered to be an a priori *reasonable topology* (3 layers with a number of neurons in the input and hidden layers equal to the dimension of the feature space) and used 30 images in the construction of the learning set, from which we randomly extracted 50 000 instances, half-divided between positives (iris) and negatives (non-iris) samples. Table 1 gives the obtained results. "Learning Error" columns give the averages errors obtained in the learning stages, "Time" the average elapsed time of the learning processes (in seconds), "Classification Error" the average error obtained in the test set images. "Sc" and "Ir" denote respectively the sclera and iris classification stages. All the values are expressed in confidence intervals of 95%.

The above described experiments led us to select the Fletcher-Reeves [15] learning method for the backpropagation algorithm and use it in subsequent experiments, namely in the search of the optimal networks' topology and of the minimum number of images required in the learning set.

Figure 6 gives 3D graphs that contain the obtained error rates in the test set, according to the number of images used in the training set (depth axis) and the number of neurons of the networks' hidden layer (horizontal axis). The error rates are averages from 20 neural networks and are expressed in percentage. Not surprisingly, we observed that the networks' accuracy have direct correspondence with the number of neurons in the hidden layer and with the number of images used in the learning process. However, we concluded that these error rates tend to stabilize when more than 40 images were

| Learning Algorithm | Time (Sc) | Learning Error (Sc) | Classification Error (Sc) | Time (Ir) | Learning Error (Ir) | Classification Error (Ir) |
|----------------------------|-----------------|-----------------------|---------------------------|------------------|-----------------------|---------------------------|
| Fletcher-Reeves [15] | 2808 ± 7.35 | $0.027 \pm 2.1E^{-4}$ | $0.029 \pm 2.7E^{-4}$ | 3320 ± 8.98 | $0.020 \pm 1.8E^{-4}$ | $0.021 \pm 1.8E^{-4}$ |
| Powell-Beale [16] | 2751 ± 8.20 | $0.026 \pm 2.3E^{-4}$ | $0.029 \pm 2.7E^{-4}$ | 3187 ± 9.30 | $0.020 \pm 2.0E^{-4}$ | $0.022 \pm 2.1E^{-4}$ |
| Broyden <i>et al.</i> [17] | 4807 ± 9.14 | $0.026 \pm 3.2E^{-4}$ | $0.031 \pm 3.5E^{-4}$ | 5801 ± 10.52 | $0.019 \pm 2.7E^{-4}$ | $0.023 \pm 2.9E^{-4}$ |
| One-secant [18] | 2993 ± 7.13 | $0.030 \pm 2.2E^{-4}$ | $0.034 \pm 2.4E^{-4}$ | 3491 ± 8.61 | $0.024 \pm 2.0E^{-4}$ | $0.031 \pm 2.1E^{-4}$ |

Table 1. Comparison between the average error rates observed in the learning and classification stages, regarding the variants of the backpropagation algorithm used in our experiments.

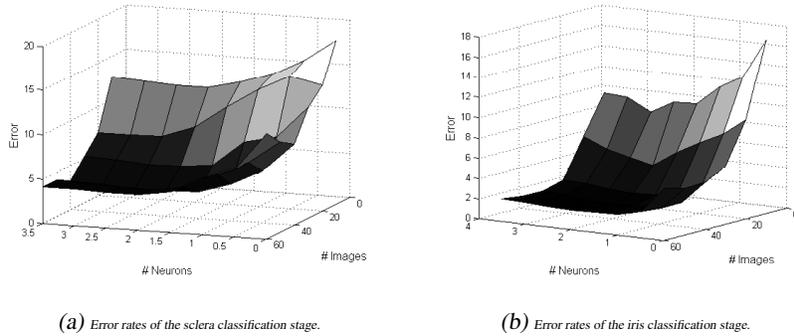


Fig. 6. Error rates obtained on the test data set, regarding the number of images used in the learning stage (depth axis) and the number of neurons in the network hidden layer (horizontal axis, expressed in proportion with the dimension of the feature space). The error rates are expressed in percentage and are averages of 20 neural networks with the respective configuration.

used in the training set, with a number of neurons in the hidden layer about 1.5 times the dimension of the feature space. Also, these observations were confirmed either in the sclera and in the iris classification models.

Interestingly, the lowest error rates were obtained in the iris classification stage, which we justified by the useful information that iris detection networks receive about the sclera localization and lessen the difficulty of their classification task. The lowest iris classification error obtained was about 1.87%, which we considered highly acceptable. This gives about 2244 misclassified pixels per image, whose are possible to reduce if basic image processing methods were applied to the network's output. For instance, morphologic operators (erosion) will eliminate small regions of black pixels separated from the main iris region whose often we observed to be cause of small errors, as it is illustrated in figure 7.

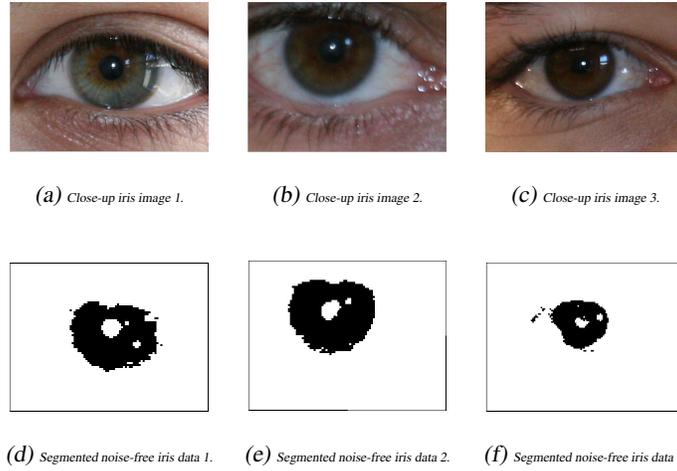


Fig. 7. Examples of the iris segmentation results obtained by the proposed method. The upper row contains the original images and the bottom row the correspondent segmentation of the noise-free iris regions.

5 Conclusions

Due to favorable comparisons with other biometric traits, the popularity of the iris has considerably grown over the last years and substantial attention was paid by both commercial and governmental organizations. Also, growing efforts are concentrated in finding the minimum level of image quality that enables recognition with enough confidence.

In this paper we used the neural pattern recognition variant to propose a method that performs the detection of the eye, the iris segmentation and the discrimination of the noise-free iris texture, analyzing images acquired on the visible wavelength under less constrained image acquisition processes. Our approach comprises two binary classification stages. At first, we used the HSV and YCbCr color spaces to provide us information about the sclera localization. This information is mixed with a new feature set to discriminate the noise-free iris regions of the images. We concluded that the proposed method accomplishes its major purposes and achieves very low error rates, even when starting from a relatively small set of images to learn appropriate classification models.

Acknowledgements

We acknowledge the financial support given by "FCT-Fundação para a Ciência e Tecnologia" and "FEDER" in the scope of the PTDC/EIA/69106/2006 research project "BIOREC: Non-Cooperative Biometric Recognition".

References

1. National Institute of Standards and Technology: Iris challenge evaluation (2006) <http://iris.nist.gov/ICE/>.
2. Proença, H., Alexandre, L.A.: The NICE.I: Noisy Iris Challenge Evaluation, Part I. In: Proceedings of the IEEE First International Conference on Biometrics: Theory, Applications and Systems (BTAS 2007), Washington (2007) 27–29
3. Wildes, R.P.: Iris recognition: an emerging biometric technology. In: Proceedings of the IEEE, vol. 85, no.9, U.S.A. (1997) 1348–1363
4. Daugman, J.G.: High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 11 (1993) 1148–1161
5. Ross, A., Shah, S.: Segmenting non-ideal irises using geodesic active contours. In: Proceedings of the IEEE 2006 Biometric Symposium, U.S.A (2006) 1–6
6. Daugman, J.G.: New methods in iris recognition. IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, vol. 37, no. 5 (2007) 1167–1175
7. Arvacheh, E., Tizhoosh, H.: A study on Segmentation and Normalization for Iris Recognition. Msc dissertation, University of Waterloo (2006)
8. Zuo, J., Kalka, N., Schmid, N.: A robust iris segmentation procedure for unconstrained subject presentation. In: Proceedings of the Biometric Consortium Conference. (2006) 1–6
9. Vatsa, M., Singh, R., Noore, A.: Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. IEEE Transactions on Systems, Man and Cybernetics - B, vol. 38, no. 3 (2008) ??–??
10. Liu, X., Bowyer, K.W., Flynn, P.J.: Experiments with an improved iris segmentation algorithm. In: Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies. (2005) 118–123
11. Dobes, M., Martineka, J., Dobes, D.S.Z., Pospisil, J.: Human eye localization using the modified hough transform. Optik, vol. 117 (2006) 468–473
12. Basit, A., Javed, M.Y.: Iris localization via intensity gradient and recognition through bit planes. In: Proceedings of the International Conference on Machine Vision (ICMV 2007). (2007) 23–28
13. Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision, vol. 57, no.2 (2002) 137–154
14. Proença, H., Filipe, S., Santos, R., Oliveira, J., Alexandre, L.A.: Toward covert iris recognition: A database of visible wavelength images captured on-the-move and at-a-distance. IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics (submitted to) (2008)
15. Fletcher, R., Reeves, C.: Function minimization by conjugate gradients. Computer Journal, Vol. 7 (1964) 149–154
16. Powell, M.: Restart procedures for the conjugate gradient method. Mathematical Programming, Vol. 12 (1977) 241–254
17. Dennis, J., Schnabel, R.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Englewood Cliffs, Prentice-Hall, E.U.A. (1983)
18. Battiti, R.: First and second order methods for learning: Between steepest descent and newton’s method. Neural Computation, vol. 4, no. 2 (1992) 141–166