

Semantic Segmentation of Color Eye Images for Improving Iris Segmentation

Dailé Osorio-Roig^(✉), Annette Morales-González^(iD), and Eduardo Garea-Llano

Advanced Technologies Application Center (CENATAV), 7a # 21406 e/ 214 y 216,
Rpto. Siboney, Playa, 12200 Havana, Cuba
{dosorio, amorales, egarea}@cenatav.co.cu

Abstract. Iris segmentation under visible spectrum (VIS) is a topic that has been gaining attention in many researches in the last years, due to an increasing interest in iris recognition at-a-distance and in non-cooperative environments such as: blur, off-axis, occlusions, specular reflections among others. In this paper, we propose a new approach to detect the iris region on eye images acquired under VIS. We introduce the semantic information of different classes of an eye image (such as sclera, pupil, iris, eyebrows among others) in order to segment the iris region. Experimental results on UBIRIS v2 database show that the semantic segmentation improves the iris segmentation by reducing the intra-class variability, especially for the non-iris classes.

1 Introduction

Iris segmentation refers to the task of automatically locating the annular region delimited between the pupillary and limbic boundaries of the iris in a given image [1]. Results in the NICE I dataset [2] show that iris segmentation under visible spectrum (VIS) is still a very challenging problem. According to [3], non-cooperative iris recognition refers to automatically recognize at-a-distance and dealing with several factors that deteriorate the quality of an iris image [4], such as occlusions, blur, off-axis, specular reflections, poor illumination among others.

Many algorithms have been proposed for separating the iris region from the non-iris regions on images affected by the aforementioned factors. One of the main approaches consists on boundary-based methods [5–7] which generally use techniques for detecting obstructions by specular reflections as first step, and then, they use methods based on the detection of two circular boundaries or non-circular trajectories with the aim of localizing both iris boundaries. After, techniques for detecting upper and lower eyelids and eyelashes are applied. The main drawback of these algorithms is that the detection of non-iris regions such as eyelids, eyelashes and specular reflections, does not preserve entirely the spatial relationship with respect to the iris region. Moreover, these approaches only take into account as non-iris regions, the eyelids, eyelashes and specular reflections disregarding other important non-iris regions such as eyebrows, pupil, skin, hair and other factors related to problems of non-cooperative environments such

as glasses and lenses. Other approaches are based on a combination of region and boundary-based methods, which initially find an approximate eye area that is discriminatory and adjacent to the iris region followed by a process of detection of circular boundaries with the aim of separating the iris region of the non-iris regions. Examples of these approaches are the methods based on the localization of the sclera region, followed by a fast circular hough transform [7], which locates the iris region by identifying the sclera as the most easily distinguishable part of the eye under varying illumination [8]. The main drawback of these methods is that they only take into account the discriminatory and adjacency notion of a single region (sclera) of the set of non-iris regions present in an eye image.

In general, the methods for iris segmentation mentioned above use only low level features for locating the iris region. They do not take into account the semantic information of all classes that conform an eye image. For this reason, we believe that, using more semantic classes (namely sclera, iris, pupil, eyelids, skin, hair, glasses, eyebrows and specular reflection) and not only iris and non-iris classes, we may improve the iris segmentation by reducing the intra-class variability and increasing the inter-class variability, especially for the non-iris class. On the other hand, by semantically segmenting color eye images (employing manual annotations in the learning process) it is possible to extend the applicability of iris segmentation algorithms to other soft-biometric features, such as the sclera and periocular regions, that can be useful to improve recognition performance on color eye images [1]. To the best of our knowledge, iris segmentation have not been addressed in terms of semantic segmentation in the literature. Summarizing, our contributions are: (1) creating a set of manual annotations for eye images containing 9 semantic classes, (2) exploring the performance of semantic segmentation for eye image classes and (3) extending the use of semantic segmentation for the specific case of iris segmentation.

The remainder of this paper is organized as follows. In Sect. 2, we present a description of the proposed method. The experimental evaluation, showing our results and a comparison with other state-of-the-art methods is presented in Sect. 3. Conclusions and future work are presented in Sect. 4.

2 Proposed Method

We propose to employ a semantic segmentation approach in order to perform iris segmentation. For this task we chose a semantic segmentation algorithm named HMRF-PyrSeg [9] that was originally designed for general images. To the best of our knowledge, there are no semantic segmentation approaches specifically designed for iris segmentation.

In Fig. 1, a description of this semantic segmentation process can be seen.

2.1 Semantic Segmentation Algorithm

The first step of the algorithm is to segment the image using the Felzenszwalb's segmentation algorithm [10] (Fig. 1b). This segmented image is represented as a

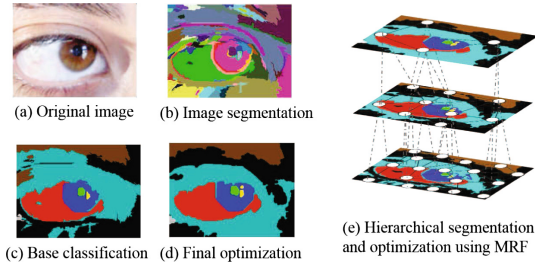


Fig. 1. General steps of the segmentation/classification process. (a) Original image, (b) initial segmentation (c) initial classification of the regions obtained by the base classifier, (d) final classification/optimization result and (e) overall example of hierarchical segmentation, classification and optimization by means of a hierarchical MRF approach (Best viewed in color).

Region Adjacency Graph (RAG), where each vertex represents a region and each edge represents the adjacency between regions. We obtain several segmentations by varying the algorithm parameters in order to avoid under-segmentation and over-segmentation problems. For all the regions in each segmentation, low-level features are extracted and a base classifier (it could be any classifier with probabilistic output) is trained taking into account the semantic classes presented in the dataset. All the image regions are classified using this base classifier (Fig. 1c). This will provide a base line classification/segmentation result for the entities present in the image.

We build a hierarchy on top of each initial segmentation, with the aim to improve it using the semantic information coming from the classification process. Given an initial segmentation, and an initial classification of its regions, we build a new segmentation level by contracting the edges of the graph, and joining the regions connected by these edges. In order to decide which edges should be contracted, we employ a criterion based on semantic information (i.e., the classes assigned to each vertex/region and their probability values), and low-level information given by the Canny edge detector to preserve important edges in the image. For more information regarding this criterion, please refer to [9]. After the new segmentation level is created, we perform the classification process on the level, and we can create a new level on top of this one using this information. This process can be repeated until no further contractions can be performed.

Once we have this segmentation hierarchy and a first classification of all its levels, a refinement is performed by means of a hierarchical Markov Random Field (MRF) approach (See Fig. 1e). Within the segmentation hierarchy, a MRF is built per level. Each MRF will receive information from the MRFs computed in adjacent levels in the hierarchy. In this way, it is possible to link spatial information within each level of segmentation and the hierarchical information between levels. Intuitively, one MRF per segmentation ensures a smoothing in terms of region labeling (classification), by taking into account the spatial relations among the regions. By using a hierarchical MRF, we are trying to add

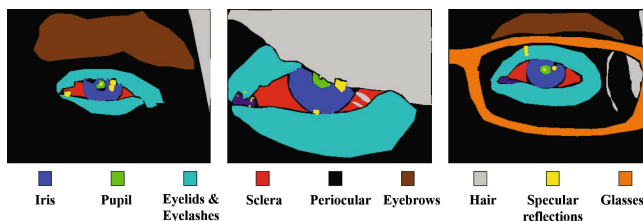


Fig. 2. Examples of manual annotation of eye images (Best viewed in color).

a hierarchical smoothing, which means, for instance, that two regions wrongly classified in one level can be corrected if their father (those two regions merged as one) in the upper level was correctly classified. In this way, we can combine local and global information at different levels of segmentation. At the end of this optimization process, the regions should be annotated with more coherent classes, according to their spatial relations and context (See Fig. 1d).

After a final classification is obtained, we add to the algorithm a post-processing step for the iris regions, which consists in a morphological close followed by a connected components analysis. The objective of this step is to remove the isolated pixels or regions that are not connected to larger objects.

2.2 Manual Annotation of Eye Images

In order to obtain a ground truth for training classifiers using the semantic classes presented in eye images, we manually annotated 53 images from the UBIRIS.v2 [11] training set. These annotations were done at pixel level, i.e. each pixel is annotated with its corresponding class id. Since this annotation process is a hard and time-consuming process, we conducted a data augmentation step, where we performed a horizontal flip on the 53 original images and their corresponding manual ground truths, thus increasing the training set to 106 eye images. Examples of these manual annotations can be seen in Fig. 2. The images were annotated with 9 semantic classes: sclera, iris, pupil, eyelids and eyelashes, skin, hair, glasses, eyebrows and specular reflections. We selected these labels with the aim of obtaining the biggest semantic information describing the different regions of an eye image.

3 Experimental Results

In this section we show the experimental results of our approach. The main objective is to show that the semantic information from eye images has advantages over the iris and non-iris classification, as well as the classification of other parts of an eye image, such as: sclera, skin, eyelids/eyelashes (E&E), pupil, eyebrows, specular reflections (SpRef), glasses and hair.

The proposed method was tested on the UBIRIS.v2 dataset employed in the NICE I competition [2]. It contains eye images acquired under unconstrained

conditions with subjects at-a-distance and on-the move. They present serious problems of occlusions, specular reflections, off-axis and blur.

We obtained, for each image, 10 initial segmentations (changing parameters from over-segmentation to under-segmentation). On top of each segmentation, we applied the hierarchical classification/optimization algorithm, obtaining 5 new segmentation levels in a hierarchy. We used Random Forest as base classifier and the low-level features employed to describe each segmented region are Geometric Context [12] and the texton-like features described in [13].

The first experiment was conducted to validate the semantic segmentation of all classes in terms of pixelwise annotation accuracy. For this experiment we employed 53 images (106 including the horizontal flip) with manual annotations (See Sect. 2.2). Those 53 images are a subset of the 500 UBIRIS.v2 images that were used for training in the NICE I [2] competition. We performed a hold out cross-validation, by splitting the set randomly in 70% for training and 30% for testing over 5 rounds. The results can be seen in Table 1. For the base classification, we show the results for level 3 and for optimization we show the results for level 1 in the hierarchy, since this configuration exhibited the best performance.

Table 1. Pixelwise annotation accuracy (%) for each semantic class

Algorithm	Pupil	Iris	Sclera	E&E	Eyebrows	SpRef	Skin	Glasses	Hair	Mean (Std)
Classification	56.80	82.68	73.92	80.09	72.28	14.55	79.38	22.92	22.19	75.00 (± 2.73)
Optimization	46.97	83.81	71.43	85.72	75.54	12.92	77.16	28.02	6.65	75.36 (± 2.97)
Post-processing	–	85.81	–	87.33	77.39	–	–	–	–	–

In Table 1 can be seen that, for some classes (iris, eyelashes & eyelids, eyebrows and glasses), the MRF optimization outperformed the base classifier, but for the other classes there was no improvement (rows 1 and 2). The reason for this might be that, for classes with small area (i.e. pupil, sclera, specular reflections, hair) the pairwise smoothing term of the MRF might induce them to change the initial label to that of the surrounding class region. Also, the specular reflections, glasses and hair classes obtained the lowest results because in the training set there are very few images annotated with these classes. Nevertheless, it is important to notice that for the best level of the base classifier the iris class obtained the best performance with respect to the rest of the classes. The third row shows the result of post-processing the iris, eyebrows and E&E classes since these classes are the more compact ones and with better defined shapes, and they usually appear as a single connected component.

The process of optimization was able to correct some errors of the initial semantic segmentation made by the base classifier. Some of these corrected errors are shown in the Fig. 3. Notice, for instance, iris (dark blue), eyelids/eyelashes (green), specular reflections (yellow) and eyebrows (brown) regions.

The second experiment was designed for evaluating the performance of iris and non-iris classification of the proposed iris semantic segmentation method.

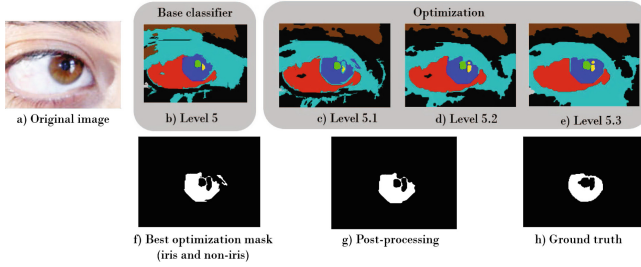


Fig. 3. Example of classification/optimization results on NICE I. (b) Level 5 of initial segmentation and base classifier result, (c–e) 3 sample levels of the optimization hierarchy built on top of Level 5, (f) mask obtained for iris/non-iris classification and (g) mask after post-processing (Best viewed in color).

We employed for training the 106 manually annotated images, and for testing we used the test set of 500 images from the NICE I competition (disjoint from the training set). We used the evaluation protocol of the NICE I competition and the average segmentation error (E_1) [2] as evaluation measure. In Table 2 we show these results for 10 initial segmentation levels with the base classifier (row 1) and their corresponding optimization (row 2).

Table 2. Comparison using E_1 (%) for different levels in classification and optimization on NICE I.

Algorithm	1	2	3	4	5	6	7	8	9	10
Base classifier	2.63	2.57	2.56	2.52	2.51	2.56	2.52	2.61	2.61	2.59
Optimization	2.39	2.39	2.39	2.39	2.38	2.49	2.52	2.55	2.68	2.69

Table 3 shows the comparison of our proposed approach with other state-of-the-art methods. It is important to notice that the number of training images used in our method is considerable smaller than other state-of-the-art methods. In Table 3 we can see that the E_1 of our method using only the base classifier was 2.51. When we used the hierarchical MRF optimization we decreased the error to 2.38, by choosing a fixed level of the optimized segmentation hierarchy. When we chose the best level of segmentation for each image, then the error was even smaller: 1.59, which outperforms most of the other methods. This last case implies selecting manually the best level of segmentation given the E_1 against the ground truth, which cannot be used in practice, but is a good indicative of the potential of our method. Automatically selecting these best levels is a good line for future improvements of our method.

Although our method was not able to outperform the one of [14], we improved some of the errors reported by that work. This can be seen in Fig. 4. In this figure we considered to show E_1 as well as the false positive and false negative

Table 3. Comparison of E_1 error in the NICE I dataset.

Method	NICE I error E^1 (%)
Our method (Base classifier)	2.51
Our method (Optimization)	2.38
Our method (Best selection)	1.59
Liu et al. [14]	0.90
Tan and Kumar [15]	1.72
Proenca [8]	1.87
Tan and Kumar [16]	1.90
Chai et al. [17]	2.80
Chen et al. [7]	2.97
Abdullah et al. [6]	2.95

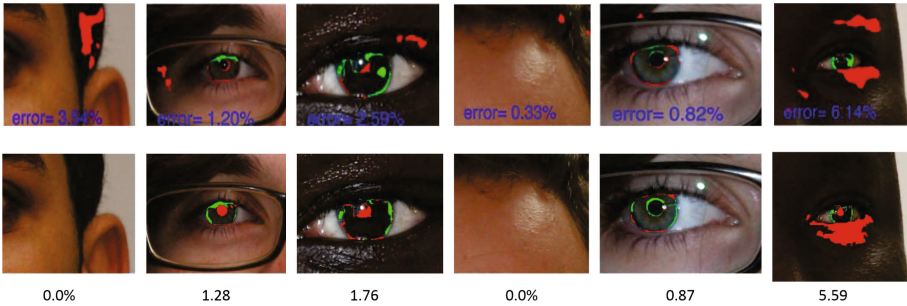


Fig. 4. Examples of classification errors. First row shows images obtained from [14] and second row shows our images. Red points denote false accept points (i.e. points labeled as non-iris by the ground truth but iris by our method). Green points denote false reject points (i.e. points labeled as iris by the ground truth but non-iris by our method) (Best viewed in color).

of those segmentations reported in [14]. We analyze these differences because this algorithm, based on multi-scale fully convolutional network (MFCNs), is the best result of iris segmentation under VIS on NICE I dataset. It can be noticed, particularly for columns 1 and 4, where there is no iris in the image, that our method was able to correctly detect that situation, while the other method detected some missclassified iris areas. Also for the case of dark skin, our method improved the recognition of these areas (see column 3).

4 Conclusions

In this paper we proposed a new iris segmentation algorithm for eye images acquired under VIS, which introduces the semantic information of the different

classes of an eye image. A set of manual annotations for eye images containing 9 semantic classes was created. Experimental results showed that using the semantic information of the different classes of an eye image is a promising path for iris segmentation. It is important to note that only with 53 annotated images the proposal achieved state-of-the-art iris segmentation accuracy. Also, by using our semantic segmentation proposal, it is possible to address other segmentation tasks like sclera and periocular segmentation, which are of interest of many researches. As future work we plan to automatically select the best segmentation level for each image, as well as to work on a fusion scheme that allows merging regions of different segmentation levels.

References

1. Jillela, R.R., Ross, A.: Segmenting iris images in the visible spectrum with applications in mobile biometrics. *Pattern Recogn. Lett.* **57**, 4–16 (2015)
2. Proença, H., Alexandre, L.A.: The NICE.I: noisy iris challenge evaluation-part I. In: *First IEEE International Conference on Biometrics: Theory, Applications, and Systems, BTAS 2007*. IEEE, pp. 1–4 (2007)
3. Proença, H., Alexandre, L.A.: Toward noncooperative iris recognition: a classification approach using multiple signatures. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4) (2007)
4. Rathgeb, C., Uhl, A., Wild, P.: *Iris Biometrics: From Segmentation to Template Security*, vol. 59. Springer Science & Business Media, New York (2012). <https://doi.org/10.1007/978-1-4614-5571-4>
5. Jeong, D.S., Hwang, J.W., Kang, B.J., Park, K.R., Won, C.S., Park, D.K., Kim, J.: A new iris segmentation method for non-ideal iris images. *Image Vis. Comput.* **28**(2), 254–260 (2010)
6. Abdullah, M.A., Dlay, S.S., Woo, W.L., Chambers, J.A.: Robust iris segmentation method based on a new active contour force with a noncircular normalization. *IEEE SMC* (2016)
7. Chen, Y., Adjouadi, M., Han, C., Wang, J., Barreto, A., Rishe, N., Andrian, J.: A highly accurate and computationally efficient approach for unconstrained iris segmentation. *Image Vis. Comput.* **28**(2), 261–269 (2010)
8. Proença, H.: Iris recognition: on the segmentation of degraded images acquired in the visible wavelength. *IEEE Trans. PAMI* **32**(8), 1502–1516 (2010)
9. Morales-González, A., García-Reyes, E., Sucar, L.E.: Improving image segmentation for boosting image annotation with irregular pyramids. In: Ruiz-Shulcloper, J., Sanniti di Baja, G. (eds.) *CIARP 2013*. LNCS, vol. 8258, pp. 399–406. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41822-8_50
10. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. *IJCV* **59**(2), 167–181 (2004)
11. Proença, H., Filipe, S., Santos, R., Oliveira, J., Alexandre, L.A.: The UBIRIS.v2: a database of visible wavelength iris images captured on-the-move and at-a-distance. *IEEE Trans. (PAMI)* **32**(8), 1529–1535 (2010)
12. Hoiem, D., Efros, A.A., Hebert, M.: Recovering surface layout from an image. *IJCV* **75**(1), 151–172 (2007)
13. Gould, S., Fulton, R., Koller, D.: Decomposing a scene into geometric and semantically consistent regions. In: *ICCV*, pp. 1–8. IEEE (2009)

14. Liu, N., Li, H., Zhang, M., Liu, J., Sun, Z., Tan, T.: Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In: ICB, pp. 1–8, June 2016
15. Tan, C.W., Kumar, A.: Towards online iris and periocular recognition under relaxed imaging constraints. *IEEE Trans. Image Process.* **22**(10), 3751–3765 (2013)
16. Tan, C.W., Kumar, A.: Unified framework for automated iris segmentation using distantly acquired face images. *IEEE Trans. Image Process.* **21**(9), 4068–4079 (2012)
17. Chai, T.Y., Goi, B.M., Tay, Y.H., Chin, W.K., Lai, Y.L.: Local Chan-Vese segmentation for non-ideal visible wavelength iris images. In: TAAI, pp. 506–511. IEEE (2015)