Pupil Localization Using Self-Organizing Migrating Algorithm

Radovan Fusek and Petr Dobeš

Technical University of Ostrava, FEECS, Department of Computer Science,
17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
radovan.fusek@vsb.cz, dobesp.nj@gmail.com

Abstract. In this paper, we propose a new method for pupil localization in images. The main contribution of the proposed method is twofold. Firstly, the method is based on the proposed eye model that takes into account physiological properties of eyes (i.e. reflects the properties of pupil, iris, and sclera). Secondly, the correct shape and the position of the model are determined using an evolutionary algorithm called Self-Organizing Migrating Algorithm (SOMA). Thanks to these ideas, the proposed method is faster than the state-of-the-art methods without reduction of accuracy. We evaluated the algorithms on two publicly available data sets in remote tracking scenarios (namely BioID [6] and GI4E [10]).

Keywords: SOMA, pupil detection, evolutionary algorithms, object detection, shape analysis

1 Introduction

In the recent years, two main sensor setups are widely used for image-based analysis of eyes (i.e. for iris, pupil, and eyelids localization or eye blink monitoring).

The first setup is represented by the head-mounted systems that are located very close to the human eyes. These systems are capable to capture the eye images in high resolutions. Thanks to this, the image-based approaches can precisely solved many important tasks like eye gaze estimation or eyelid localization.

The second setup is represented by the cameras that are located remotely from the subjects. In these systems, the eye covers only a small region of the image and the extracted images of eyes are usually of a lower quality than in the head-mounted systems in many cases. This fact makes the detection of eye parts, eye tracking and gaze estimation more difficult.

In addition to the detection accuracy, the performance in terms of computational time is equally important. For example, many vehicles are equipped with in-car cameras that monitor the driver fatigue. The fatigue of drivers represents a frequent cause of car accidents. To evaluate and prevent this situation, the approaches (e.g. pupil localization) must work very quickly.

These facts were a motivation for creating a new method for pupil localization that is faster and more precise than the existing state-of-the-art methods and works well in the remote camera systems with low resolution images. In the proposed method, the pupil center is localized by finding the global extrema of the proposed fitness function using an evolutionary algorithm called SOMA (Self-Organizing Migrating Algorithm) [12]. Thanks to SOMA and the proposed fitness function that is based on the eye model, the presented method achieved better computational time and recognition performance when compare to the existing methods.

The rest of the paper is organized as follows. The previously presented papers from this area are mentioned in Section 2. In Section 3, the main ideas of proposed method are described. In Section 4, the results of experiments are presented showing the properties of the new method.

2 Related Work

In the recent years, the detection and localization of pupil and iris became very important in many different areas (e.g. medicine, psychology, bio-metric, automotive). In the area of pupil localization, a popular and robust algorithm, named Starburst, was presented in [8]. The corneal reflection is located and removed from the image in the first step. Then the possible pupil edge points are marked using a set of rays (intensity changes along the rays are examined). In the last step, the marked points are used for ellipse fitting using RANSAC. Swirsky et al. [9] proposed the pupil tracking method that uses Haar-like features to estimate the pupil location combined with the k-means clustering to refine the pupil centre. In [2], Exclusive Curve Selector (ExCuSe) for pupil detection was proposed. This method is based on edge filtering and oriented histograms calculated via the angular integral projection function. Another pupil detection method know as SET was proposed in [5]. The SET method consists of image thresholding followed by segmentation to group related pixels. The segments are filtered and the border of each segment is used as an input for ellipse fitting. The next algorithm for pupil detection, named Ellipse Selector (ElSe), was proposed in [3]. The Canny edge filter and morphological operations are used to detect pupil related edges which is followed by ellipse fitting. In the case that no ellipse is found, an advanced blob detection is used to find the pupil center. The method for iris centre localization was proposed in [4]. In the first step, the coarse location of iris center is obtained using a fast convolution based approach. Afterward, the iris center location is refined using boundary tracing and ellipse fitting. In [11], the authors mentioned that the eyelids and eyelashes are noise factors for iris recognition. Therefore, in their paper, the eyelids and eyelashes detection algorithm based on the Canny edge detection technique with the Hough transform is used and the eyelids and eyelashes are removed from the iris images to increase the performance of the iris recognition system. In [7], a pupil localization method based on Hough regression forest was proposed. This method is based on the training process (supervised). A big evaluation of the mentioned pupil detection methods was presented in [1].

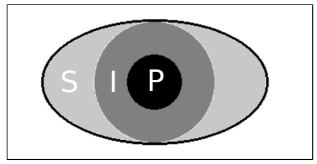

Fig. 1. The eye model for determining the proposed fitness function.

3 Proposed Method

In this section, we describe the main ideas and contributions of the proposed method for pupil localization.

The contribution of the proposed method is twofold. Firstly, the method is based on the proposed eye model (Fig. 1). Thanks to this model, the method gives a lower number of false positive detections compare to the state-of-the-art methods. However, determine the correct location and properties of this model (in real images) represents a challenging optimization problem. Therefore, the second contribution is represented by the use of an evolutionary algorithm SOMA (Self-Organizing Migrating Algorithm). Using SOMA, we are able to find the correct location of the pupil (model fitting) faster than the state-of-the-art methods without reduction of accuracy.

Since our method for pupil detection takes into account physiological properties of eyes, let us consider a following ideal model of eye in Fig. 1. It is important to note that in the flowing text, we suppose that the eye region is obtained beforehand (e.g. using facial landmarks or classical eye or face detectors). In the ideal model, we can assume that the pupil is represented by a dark area (area P in Fig. 1) surrounded by a slightly brighter area of iris (area I in Fig. 1) which is surrounded by a white sclera (area S in Fig. 1). In general, the goal of the pupil (iris) detection method is to find the appropriate location of these parts. As was note in the previous sections, the position of the pupil center is important information for gaze direction recognition, recognition of driver drowsiness, or pupil tracking systems.

As was mentioned, the parts of the eye (pupil, iris, and sclera) have different pixel intensity values in images. For example, we can suppose that the pupil area is the darkest area in the eye region (Fig. 1). However, the darkest area can be detected in the eyebrow. Which is a common problem of many pupil detection approaches. Therefore, in addition to the information of the darkest area of pupil, we use the pixel intensity values of iris (or sclera). It follows that the correct location of the pupil (iris) can be determined as the ratio of mean intensity values between the mentioned areas. Based on this fact, we propose the following function that describes the properties of the presented eye model, and that is designed to maximize the contrast between all three regions (Eq. 1).

$$f = \frac{mean_I}{mean_P} \cdot \frac{mean_S}{mean_P} \tag{1}$$

Let us consider that the eye model in Fig. 1 represents the image template, and we can perform a template matching procedure. It means that we have to compute the function (Eq. 1) of the template at each location of the input image to find the best match (the location with the maximum value of the proposed function Eq. 1). It is clear that in real images, the eye parts (pupil, iris, and sclera) have different sizes (parameters) and one fixed size template will not be enough for correct detection of the eye parts. It follows, that carry out the template matching procedure (model fitting) for each parameter (different template sizes) represents a time consuming operation.

Therefore, we propose a novelty way for model fitting (localization) of the proposed eye model. Our approach uses an optimization method based on an evolutionary algorithm called Self-Organizing Migrating Algorithm (SOMA [12]). The main idea of this algorithm can be described as individuals cooperating and wandering through the searched space. During the iteration process, the individuals affect each other so they can form the groups migrating together while searching for the global maximum.

Since we focus on the low resolution images (remote tracking scenario), and due to the fact that the computational complexity of pupil detection plays a crucial role (e.g. for the recognition of driver fatigue), the function in Eq. 1 is designed in such a way that it can also be computed very quickly with the use of simplified rectangle eye model. It follows that the proposed eye model in Fig. 1 can be transformed into a square shape model (Fig. 2). The experiments (in the following section) shown that the combination of this model and SOMA allows as to determine the correct location of pupil very quickly, especially in low resolution images, and the achieved results outperform the state-of-the-art methods. Therefore, in the following text, we use this square model.

For the convenience of readers, let us describe the whole SOMA process (with implementation details) in a following example (Fig. 3). In our case, the dimension of searched space is 5 because 5 parameters have to be obtained. The coordinates x and y of the pupil center (center of area P in Fig. 2), and lengths of the sides of pupil, iris, and sclera (lengths of areas P, I, and S in Fig. 2). Each

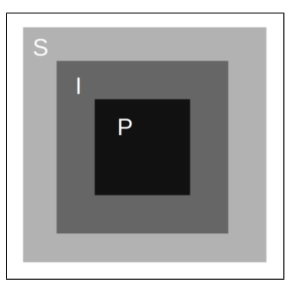

Fig. 2. The square eye model designed for fast computation of the proposed fitness function.

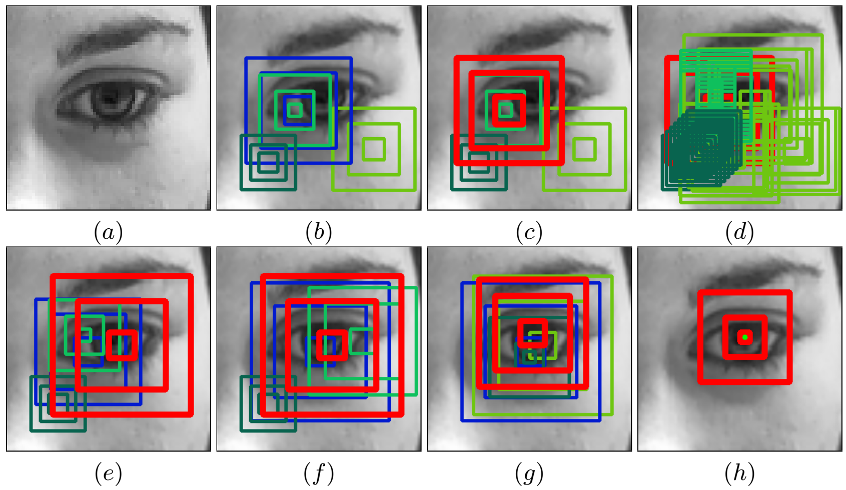

Fig. 3. The detection process of the proposed method. The input image (a). The initial population (the size of population is 4) of individuals (b). The first leader (depicted by the red color) based on the proposed fitness function (c). Examples of jumps of individuals (d). Examples of individuals with the new leaders after three iterations $(e-g)$. The final leader (after 12 iterations) that represents the correct area of pupil with the depicted pupil center (h).

individual is then defined as the vector of size equal to the dimension size of the searched space and represents the position in the searched space.

In the first step, an initial population (the size of population is 4) of individuals is generated (Fig. 3 (b)). At the beginning of every iteration, so-called *Leader* is found as the individual with the highest fitness function value (Eq. 1) in the population. In Fig. 3 (c), the first leader is depicted by the red color.

In the second step, the individuals begin to jump, towards the *Leader*. The jump means, that the individual is moved in every dimension. The jumps are defined by the fixed size step and by the number of jumps. Both parameters are defined in advance. An example of jumps is shown in Fig. 3 (d). In every new position (jump) of the individual, the fitness function is evaluated. After evaluating all jumps, the individual is moved to the position with the best fitness function value. After all individuals are processed this way, the algorithm continues with the next iteration by selecting the new *Leader*. Examples of individuals with the new leaders (after each iteration) are shown in Fig. 3 (e-g). The final position of the detected area of pupil (final leader after 12 iterations) with the pupil center (after described detection process) is shown in Fig. 3 (h).

It is important to note that the individuals begin to jump, towards the *Leader*, according to a perturbation vector. The perturbation vector is generated for every individual at the beginning of iteration. For every dimension of the searched space, a random number (in range $(0, 1)$) is generated and com-

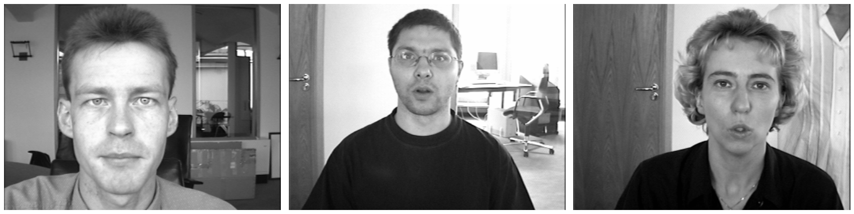

Fig. 4. Examples of BioID dataset images.

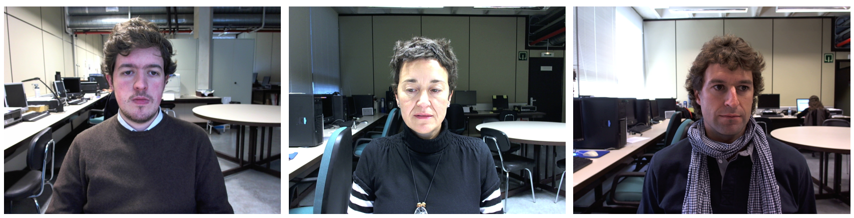

Fig. 5. Examples of GI4E dataset images.

pared to a control parameter PRT which is set in advance. If the number is greater than the constant PRT, the individual will move in this dimension. We also note that many versions of SOMA exist, in this work, we use a all-to-one version.

4 Experiments

Fig. 6. Examples of used eye regions. The first row: GI4E dataset, the second row: BioID dataset.

To evaluate the results of the proposed method, we used two public datasets; BioID [6] and GI4E [10]. The BioID dataset contains 1521 gray level images with the resolution of 384x286 pixels. Every image shows one of 23 persons with different illumination conditions in different indoor environments. The dataset was collected for the purpose of testing the face recognition algorithms. The

FGNet project, which focuses on face and gesture recognition provides manually marked up labels for the BioID dataset with several facial landmarks (including the pupil center positions). Fig. 4 shows example images from the BioID dataset.

The GI4E database provides 1339 images with the resolution of 800x600 pixels along with the manually labeled ground truth consisting of the iris center position and eye corners positions. In Fig. 5, example images from the GI4E dataset are shown. In the following experiments, the eye regions are selected according to the eye corner positions based on the provided ground truth information. Examples of eye regions are shown in Fig. 6.

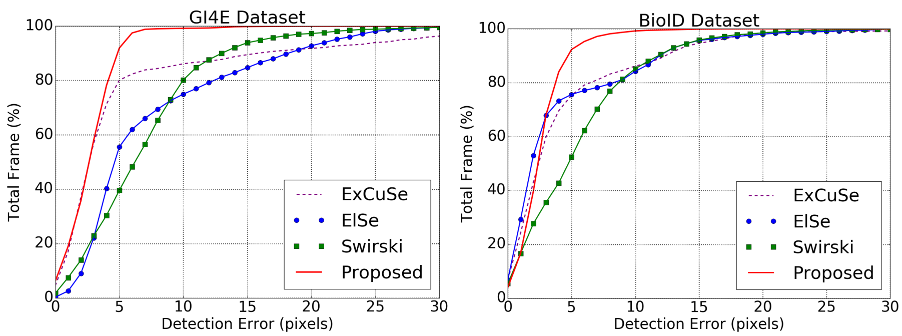

Fig. 7. The cumulative distribution of detection error. The error that is calculated as the Euclidean distance (in pixels) is in the x-axis. The y-axis shows percentage of frames with detection error smaller or equal to a specific error. The names of datasets are placed above the graphs.

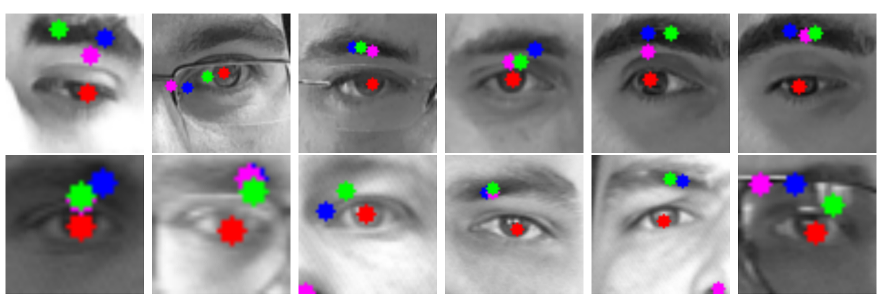

Fig. 8. Examples of images in which the proposed method performs better compared to other tested methods. The results of methods are distinguished by color: proposed method - red, ElSe - blue, ExCuSe - purple, Swirski - green. The first row: GI4E dataset, the second row: BioID dataset.

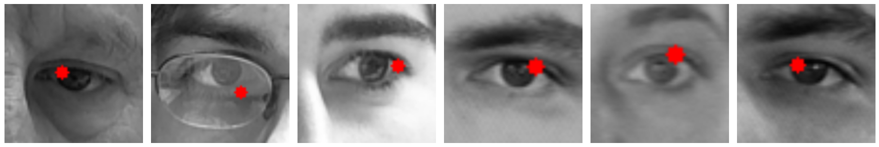

Fig. 9. Examples of images in which the proposed method fails.

Table 1. A comparison of time and errors.

	BioID Mean error (pixels)	GI4E Mean error (pixels)	Time per region (ms)
Proposed method	2.91	2.73	1.0
ElSe	4.17	7.62	2.3
ExCuSe	4.44	6.25	1.5
Swirski	5.58	6.97	7.6

To compare the proposed algorithm to state-of-the-art methods, we have chosen three methods, namely ElSe, ExCuSe and Swirski. Even though ExCuSe, ElSe and Swirski were primarily designed to work with high-resolution images acquired by head-mounted cameras, the experiments in [1] show that the methods can be used in the remote images as well. Since the methods that were compared require that the values of parameters are properly set, we paid attention to experimenting with their various values. For ElSe, we directly used the setting for remotely acquired images published by the authors of the algorithm.

The resulting plots of experiments are shown in Fig. 7. In the plots, we provide the cumulative distribution of detection errors for each method and dataset (percentage of frames with detection error smaller or equal to a specific error). The error is calculated as the Euclidean distance from the ground truth of the pupil center and the center provided by the particular detection method. In Table 1, we provide the average errors (in pixels) and average times needed for processing one eye region on an Intel core i3 processor (3.7 GHz).

Our results show that the proposed algorithm gives a stable and high detection rate among the evaluated approaches. Especially on the GI4E dataset, our method has a very small average error (2.73 pixels). Based on the results in Fig. 7, it can be observed that the proposed method is able detect approximately 90% of all frames with detection error smaller than 5 pixels. As is shown in Table 1, the proposed method is one of the fastest in the tests (only ExCuSe achieved similar time).

Fig. 8 shows several cases in which our method works better compared to other tested methods. Typical cases of errors are caused by the presence of black eyebrow or glasses. Based on the results, we can conclude that the proposed method is better in such cases. However, not all images are successfully detected by proposed method. Fig. 9 shows several examples of images in which the proposed method fails.

During the development of proposed method, several parameters were determined experimentally. The number of individuals and jumps was 10, the size step of jump was 2 pixels. The number of iterations of the SOMA algorithm was 12. The higher number of iterations does not lead to perceptible increase the performance. The PRT parameter was set to 0.5. The maximal sizes of the side lengths of pupil, iris, and sclera were determined based on the size of input image.

5 Conclusion

In this paper, we presented a new method for pupil detection in the remote images. The main idea is based on the fact that the pupil can be localized using the eye model that reflects the properties of eyes. Based on this model, we proposed the appropriate fitness function. To find the global extreme of the function, we used Self-Organizing Migrating Algorithm (SOMA). On the basis of the experiments, we can conclude that the newly proposed approach outperforms the state-of-the-art methods (namely ExCuSe, ElSe and Swirski).

In the current version of the presented approach, the rectangular eye model is used due to the fact that it can be computed very quickly. We leave the deeper experiments with another models (e.g. circular, elliptical) that can also be used for pupil detection (e.g. in head-mounted high resolution images) for future work.

References

1. Fuhl, W., Geisler, D., Santini, T., Rosenstiel, W., Kasneci, E.: Evaluation of state-of-the-art pupil detection algorithms on remote eye images. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. pp. 1716–1725. UbiComp '16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2968219.2968340
2. Fuhl, W., Kübler, T., Sippel, K., Rosenstiel, W., Kasneci, E.: Excuse: Robust pupil detection in real-world scenarios. In: Azzopardi, G., Petkov, N. (eds.) Computer Analysis of Images and Patterns. pp. 39–51. Springer International Publishing, Cham (2015)
3. Fuhl, W., Santini, T.C., Kübler, T.C., Kasneci, E.: Else: Ellipse selection for robust pupil detection in real-world environments. CoRR abs/1511.06575 (2015), http://arxiv.org/abs/1511.06575
4. George, A., Routray, A.: Fast and accurate algorithm for eye localization for gaze tracking in low resolution images. CoRR abs/1605.05272 (2016), http://arxiv.org/abs/1605.05272
5. Javadi, A.H., Hakimi, Z., Barati, M., Walsh, V., Tcheang, L.: Set: a pupil detection method using sinusoidal approximation. Frontiers in Neuroengineering 8, 4 (2015), https://www.frontiersin.org/article/10.3389/fneng.2015.00004
6. Jesorsky, O., Kirchberg, K.J., Frischholz, R.W.: Robust face detection using the hausdorff distance. In: Bigun, J., Smeraldi, F. (eds.) Audio- and Video-Based Biometric Person Authentication. pp. 90–95. Springer Berlin Heidelberg, Berlin, Heidelberg (2001)

7. Kacete, A., Royan, J., Seguier, R., Collobert, M., Soladie, C.: Real-time eye pupil localization using hough regression forest. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1–8 (March 2016)
8. Li, D., Winfield, D., Parkhurst, D.J.: Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops. pp. 79–79 (June 2005)
9. Świrski, L., Bulling, A., Dodgson, N.: Robust real-time pupil tracking in highly off-axis images. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 173–176. ETRA '12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2168556.2168585
10. Villanueva, A., Ponz, V., Sesma-Sanchez, L., Ariz, M., Porta, S., Cabeza, R.: Hybrid method based on topography for robust detection of iris center and eye corners. ACM Trans. Multimedia Comput. Commun. Appl. 9(4), 25:1–25:20 (Aug 2013), http://doi.acm.org/10.1145/2501643.2501647
11. Wagh, A.M., Todmal, S.R.: Article: Eyelids, eyelashes detection algorithm and hough transform method for noise removal in iris recognition. International Journal of Computer Applications 112(3), 28–31 (February 2015)
12. Zelinka, I.: SOMA — Self-Organizing Migrating Algorithm, pp. 167–217. Springer Berlin Heidelberg, Berlin, Heidelberg (2004)