# Accurate and Robust Pupil Positioning Algorithm Using Adaboost Cascade Detector and Improved Starburst Model

Bing Liu<sup>1, a</sup>, Zhiyi Qu<sup>1, b</sup>, Yueqing Ren<sup>1, c</sup> and Ruicheng Wang<sup>1, d</sup> <sup>1</sup>School of Information Science & Engineering, Lanzhou University, Lanzhou 730000, China.

<sup>a</sup>liub13@lzu.edu.cn, <sup>b</sup>quzy@lzu.edu.cn, <sup>c</sup>renyq13@lzu.edu.cn, <sup>d</sup>wrch1218@126.com

Keywords: Pupil positioning, eye tracking, Adaboost, starburst model.

**Abstract.** With the rapid development of human-computer interaction systems, eye tracking algorithm has become a very popular research field. As the most important part in eye tracking, pupil positioning algorithm is still facing many challenges, especially in processing non-ideal eye images. In this paper we propose an accurate and robust pupil positioning algorithm using Adaboost cascade detector and improved starburst model. Firstly, we use an average linear interpolating method to interpolate the specular reflections. Secondly, we combine HAAR features and Adaboost cascade detector to extract the region of interest (ROI) from the eye image (after specular reflections removal). Lastly, the improved starburst model is adopted to locate the pupil contour and the pupil center. In order to evaluate the performance of proposed method, we collect several eye moving videos using the specified camera device and the experimental results prove that the accuracy and robustness of proposed algorithm are better.

## Introduction

The video-based pupil center positioning technology, whose goal is to detect the eye moving trajectory accurately, has been used in many areas in recent decade, such as the human-computer interaction, the lie detection and the auxiliary diagnosis of vertigo disease, etc. The main work of pupil positioning algorithm is to approximate the pupillary boundary with an appropriate curve and based on which to calculate the position of pupil. Usually, however, the eye images of the video are not very satisfactory and they may have more or less noises from eyelids, eyelashes and shadows, so there are still many challenges for the researchers to find an accurate, robust and fast method to locate the position of pupil.

In recent years, lots of researchers have put forward many viable ideas in this field. In 2005, Li et al. [1] first proposed the classical starburst model to detect the pupillary boundary. Based on Li's work Yan et al. [2] made an improvement through using the gray projection and edge enhance strategies. But their works does not mention the processing for the non-ideal eye images (eyes with long eyelashes or the pupil were sheltered by eyelids, etc.). Soltany et al. [3] used the gray projection method to locate the center of the pupil coarsely and then the circular Hough transform was adopted for pupillary boundary fitting. It is obvious that it will generates serious errors when the pupil is not a standard circle. Yadav et al. [4] adopted a fusion strategy, which combined the circular Hough transform, gray projection algorithm and coarse positioning method, to improve the robustness and accuracy of the locating results. The huge calculations may reduce the efficiency of the algorithm and for non-ideal eye images, the result may not be satisfying.

In this paper, we propose an accurate and robust pupil positioning algorithm using Adaboost cascade detector and improved starburst model. The rest of this paper is structured as follows: Section 2 provides the processing for specular reflections removal of the original eye image. In section 3, Adaboost cascade detector is adopted for extracting the region of interest (ROI). Section 4 represents the details of pupillary boundary locating and pupil positioning using improved starburst model. The experimental results is given in section 5. Section 6 is the conclusions of this paper.

#### **Specular reflections removal**

As most of eye images are collected under the illuminating of infrared light source [5], it is hardly inevitable to generate some specular reflections on the cornea and in some cases, the structure of pupils are damaged by the reflections, which may causes a series of influences for the subsequence processing. So we need to eliminate the specular reflections at first.

Here we adopt an average linear interpolating method to interpolate the reflections. Above all, convert the original image (the left part of Fig.1a) into binary pattern, as shown in Fig.1b. It is obvious that the highlight regions correspond to the specular reflections in the original image. Then, substitute the grayscale value of the pixel point in the reflections with the information of average grayscale values from up, down, left and right neighborhood regions.

We take Fig.1a as an example to illustrate the details of interpolating process. Assume that the distance of P to the upper boundary of the reflection is  $d_{up}$  and to the lower boundary is  $d_{down}$ , the average grayscale value of upper neighborhood region (the region adjacent to the reflection and contains 5 to 7 pixels along the direction of P to the upper boundary point ) of original image is  $Avg_{up}$  and the average grayscale value of lower neighborhood region (similar with the definition of the upper neighborhood region but at the lower part) of original image is  $Avg_{down}$ . Then we can get the value of vertical direction as follows:

$$r = \frac{d_{up}}{d_{up} + d_{down}} \tag{1}$$

 $P_{vertical} = (1 - r) \times Avg_{up} + r \times Avg_{down}$ <sup>(2)</sup>

Further, we can get the value of horizontal direction  $P_{horizontal}$  in a similar way as that of vertical and the final interpolating pixel value is calculated as follow:

$$P_{pixel\_value} = \frac{P_{vertical} + P_{horizontal}}{2}$$
(3)

Then after traversing each pixel in the reflection regions with this interpolating method, the reflections are removed. The advantage of this way to interpolate the reflections is to make the side which the interpolating point more close to it will provide more information for interpolating while the further provide less. This is a more reasonable method compared with other simple inpainting methods.





Fig.1 Sample to illustrate the specular reflections removal. (a) The diagram of interpolating. (b) The binarized eye image. (c) The eye image after specular reflections removal.

#### Extraction to the region of interest

In order to reduce the burden of compute and make the subsequence processing more focused, here we need to eliminate the irrelevant regions from eye image as far as possible and segment out the region of interest. In recent years, the Adaboost algorithm [6] has been proven to be a good strategy in processing two-class classification problems and has an excellent performance in detecting well-structured objects such as human face, hand gesture etc. The core idea of Adaboost algorithm is to combine a series of weak classifiers into a strong classifier.

Before training the Adaboost cascade detector, we need to extract the features from each of image sample, the most popular available features contain: HAAR feature, LBP (Local Binary Pattern) feature and HOG (Histogram of Oriented Gradient) feature. The HAAR feature and LBP feature are mostly used in human face detection and the HOG feature is mostly used in pedestrian detection. However, for pupil detecting there is no existing proven to be excellent feature to use, we need to select a proper one from the three features mentioned above through experimental analyzing.

In experiments, we collect 4720 samples and separate these samples into two sets: the training set (1920 samples) and the testing set (2800 samples). The training samples include 780 positive samples, as shown in Fig.2a, and 1140 negative samples, as shown in Fig.2b. In addition, all the positive samples are normalized into a size of 32\*32 while the negative samples without any further processing.



Fig.2 Some of the training samples. (a) Part of the positive samples. (b) Part of the negative samples. In order to get the best-fit detector, we extract all three features from the training samples and use them to train the Adaboost cascade detectors respectively. Through analyzing the experimental results we find that the detector trained by HAAR features have a higher performance in detecting the pupil structure compared to the other two. Therefore, we decide to use the detector trained by HAAR features as the best-fit detector to detect the pupil region in experiments. Some samples after ROI extracting using the Adaboost cascade detector are shown in Fig.3.



Fig.3 Some samples after Adaboost cascade detector detecting. The white boxes in (a) - (c) denote the ROI of the eye images and (d) is the ROI extracted from the eye image (c).

Through this step, the algorithm can skip the images without eyes (e.g. blink, or the eye covered by the eyelids seriously), so it can reduce the time consuming in processing. Besides, the ROI makes the subsequence processing more focused.

### Pupil positioning using improved starburst model

In 2005, Li et al. [1] developed an open-source eye-tracking system, in the system they first proposed the classical starburst algorithm for pupil locating and tracing and got a good experimental result. They regard the center of the first frame eye image as the starting point to detect the candidate edge feature points, iteratively, until the starting point converges and then adopt the Random Sample Consensus (RANSAC) algorithm to find a best-fit ellipse model for pupil contour fitting.

In order to get a better performance, here we make an improvement based on the classical algorithm. Before using the starburst model to detect the candidate edge points, we firstly get a coarse position of the pupil center as follows:

1. *Binarization*. To binarize the ROI image (the pupil image extracted in the above section) using an adaptive threshold. The threshold can be get by using the discrete gray gradient differential thresholding method [7]. As shown in Fig.4. Here we limit the adaptive threshold in the range of

$$[0, \frac{2}{3} Average]$$
, where Average is the average value of all the pixel values of the ROI image.

Then the median filter is used for eliminating the noises of the binarized pupil image.

2. *Finding Maximum Contour*. To find out the contour of the maximum connected domain and record the coordinate values of the contour points in a set as follow:

$$P = \{ (x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n) \}$$
(4)

Where  $(x_i, y_i)(i = 1, 2, 3, ..., n)$  denotes the position of the point i on the contour (See Fig.5).

3. *Coarse Positioning*. To calculate the average values of the horizontal and the vertical coordinates values of the maximum contour respectively:



Fig.4 Pupil image binarization using the discrete gray gradient differential thresholding method.

$$Avg_{horizontal} = \frac{\sum_{i=1}^{n} X_{i}}{n}$$

$$Avg_{vertical} = \frac{\sum_{i=1}^{n} Y_{i}}{n}$$
(6)

Where n is the number of the points on the contour.



Fig.5 The coarse positioning result. The white points denote the maximum connected domain contour in the binarized pupil image and the position of the crossing denote the coarse center of pupil.

Then we can coarsely get the center point of the pupil as  $(Avg_{horizontal}, Avg_{vertical})$ , which can be seen in Fig.5 (marked with a white crossing).

Next, we regard the coarse position of the pupil center as the starting point and follow the rays extending from the starting point to calculate the candidate feature points following the steps described in [1] iteratively (See Fig.6a), until the starting point converges (See Fig.6b). Then using the RANSAC algorithm to search for the best-fit elliptical model as the contour of the pupil. So the elliptical center can be regarded as the accurate center of the pupil, as shown in Fig.6c.



Fig.6 Sample to illustrate the positioning process using starburst model. (a) The first iteration using the coarse pupil center as the start point to search the feature points along the direction of rays. (b) The detected feature points when the start point converges. (c) To fit the pupil contour using an elliptical model and regard the center of the ellipse as the center of the pupil.

Here we using the coarse center of the pupil as the start point to detect the feature points on the edge of pupil contour can greatly reduce the iteration times and improve the positioning accuracy.

### **Experimental results**

In order to test the efficiency of the proposed algorithm, we perform the algorithm on a PC with Intel i3 CPU at 3.30GHz speed and with 4.0G RAM.

The test videos are collected by the camera device provided by China Medical Group, as shown in Fig.7. Each image of the video with a high resolution of 1080p and in experimental we contract it into a size of 384\*216. Some eye image samples collected by the camera device are shown in Fig.8.



Fig.7 The camera device.



Fig.8 The eye image samples collected by the specified camera device.

In order to select the most proper feature from HAAR, LBP and HOG, we extracted all three features from the training samples and used them to train the Adaboost cascade detectors, respectively. Then perform the detectors on the test set and the comparison is listed in Table 1.

Table 1 The performance results comparison of the detectors trained by different features.

Features	Right	Wrong	No result	Accuracy(%)
HAAR	2732	23	45	97.6%
LBP	2641	47	112	94.3%
HOG	2709	43	48	96.8%

In Table 1 we can obviously see that the classifier trained by HAAR feature has better performance compared with the other two.

In experiments we selected 5 representative eye videos to evaluate the performance of proposed algorithm and in order to make a comparison we re-implemented the algorithm described in [1].

The frames count of all 5 videos is 5589. The statistic comparison is listed in Table 2.

Table 2 The performance comparison of algorithms				
Methods	Mean Frames Per Second	Accuracy (%)		
Li et al.[1]	13	89.0%		
Proposed Method	11	96.3%		

In table 2, the mean frames per second denotes that the proposed method has a similar performance in time consuming compared with [1], however, the accuracy of our method is much higher than that

of [1]. So, on the whole, the performance of our method is better. Some positioning result samples are shown in Fig.9.



Fig.9 Samples of positioning result using the algorithm proposed in this paper.

## Conclusions

Eye tracking technology has been used in many human-computer interaction areas and continue to improve the processing speed and the accuracy of positioning, especially for the non-ideal eye images, is the pursuing of every researcher in this field. In this paper we proposed an accurate and robust pupil positioning algorithm using Adaboost cascade detector and improved starburst model. Firstly, we use an average linear interpolating method to interpolate the specular reflections. Secondly, through comparing we select the best-fit HAAR feature to train an Adaboost cascade detector to extract the ROI from the original eye image (after reflections removal). This step makes the subsequence processing more focused. Then the improved starburst model is used for pupil contour fitting and pupil center positioning. Lastly, the experimental results show that our algorithm proposed in this paper is accurate and robust.

In the future work we have to do our best to improve the performance of the algorithm in processing real-time video images and further enhance the robustness for extremely non-ideal eye images, such as, the eye sheltered by long eyelashes and eyelids etc.

## References

[1] Li, D., Winfield, D., and Parkhurst, D.J.: 'Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches'. Proc. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops - Volume 032005 pp. Pages

[2] Bei, Y., Xigong, Z., and Lifeng, G.: 'Improvement on Pupil Positioning Algorithm in Eye Tracking Technique', in Editor (Ed.)<sup>(Eds.)</sup>: 'Book Improvement on Pupil Positioning Algorithm in Eye Tracking Technique' (2009, edn.), pp. 1-4

[3] Soltany, M., Zadeh, S.T., and Pourreza, H.-R.: 'Fast and accurate pupil positioning algorithm using circular Hough transform and gray projection', in Editor (Ed.)^(Eds.): 'Book Fast and accurate pupil positioning algorithm using circular Hough transform and gray projection' (2011, edn.), pp. 556-561

[4] Yadav, M.R., and Shivdas, S.S.: 'Novel Method to Localize the Pupil in Eye Gaze Tracking Systems ', International Association of Scientific Innovation and Research (IASIR) (An Association Unifying the Sciences, Engineering, and Applied Research) 2014, 8(1), pp. 52-57

[5] Chennamma, H., and Yuan, X.: 'A survey on eye-gaze tracking techniques', arXiv preprint arXiv:1312.6410, 2013

[6] Freund, Y., and Schapire, R.E.: 'A decision-theoretic generalization of on-line learning and an application to boosting', Journal of computer and system sciences, 1997, 55, (1), pp. 119-139

[7] Lin, Y., Qu, Z., Zhang, Y., and Han, H.: 'A Fast and Accurate Pupil Localization Method Using Gray Gradient Differential and Curve Fitting', in Wong, W.E. (Ed.): 'Proceedings of the 4th International Conference on Computer Engineering and Networks' (Springer International Publishing, 2015), pp. 495-503