Proceedings of the 37th Chinese Control Conference
July 25-27, 2018, Wuhan, China

Research on Pupil Center Localization in Eye Gaze Tracking System

Xing Jin, Zhicheng Li, Jingjing Zhang*, Xiang Yang

1. School of Automation, China University of Geosciences, Wuhan 430074, China

2. Hubei key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China
E-mail: work.zhang@cug.edu.cn

Abstract: The human-computer interaction technology which based on the gaze tracking system is convenient and fast. It can achieve the purpose of sight and computer interaction. Based on the relatively static head tracking system, we make improvements for OTSU algorithm to separate the binarized pupil image and the background image completely. I wrote a function to delete a small area that gets a complete and clear binarized pupil image for the existence of small area noise clump. Finally, using the contour extraction method and the three - point circle method to complete the task of Pupil location.

Key Words: Gaze tracking system, Otsu algorithm, Pupil location

1 Introduction

With the rapid development of computer technology, human-computer interaction technology has evolved from the initial single-channel method of mouse and keyboard to multi-channel interaction mode which combining voice, gesture, posture and even biological myoelectric signals. Human-computer interaction technology based on the gaze tracking system has been applied in various fields in recent years due to its convenience and rapidity[1].Pupil location as an important research topic in line-of-sight tracking, its accuracy and effectiveness directly affect the pros and cons of the entire gaze tracking system[2].

In general, camera pixels that capture human eye images are low, while displays for interactive functions tend to have higher resolution. It may cause the calculated gaze point position deviation of dozens of pixels if the results of pupil positioning have a pixel deviation. Therefore, the accuracy of pupil location is crucial to the overall performance of the system. Most methods are based on the detection of a circle which mainly considering the geometric characteristics of the pupil. Commonly used methods include Hough transform[3] method and ellipse fitting method[4][5].Hough transform method is to determine a circle by parameter search. Ellipse fitting mainly uses the least square method or other methods to perform ellipse fitting. At first Hough transform and ellipse fitting method need to binarize the eye grayscale image and then extract the edge information of the binary image to locate its edge information. However, it is difficult to select the threshold value when obtaining the binary image, especially the selection of the threshold under the infrared light is not easy. Wildes[3] and Tisse[6] through the Hough transform voting mechanism to find the optimal value in the parameter space. Takeshi Takegami[7] proposed Hough transform to fit ellipses to locate ellipses. Shuqiang Zhang, Yueshen Lai[8] and others using Sobel operator combined with Hough circle detection method to locate the pupil. Changyuan WANG, Xueying Shi[9] proposed a method of locating the pupil using Hough Transform to detect a circle based on strong and weak edges. Liuliu Zhu, Guangming Lu[10] proposed a method which based on the detection of the longest chord boundary in the iris localization. The method uses Otsu algorithms to realize the threshold extraction in the image binarization. The Otsu algorithms improves the adaptiveness of image threshold selection. However, it is easy to cause more black areas, and the operation time of the whole method is longer in low light conditions. In general, its effect isn’t very well.

Various disturbing factors(such as IR spotlight on the cornea, eyelashes, eyelids, uneven lighting, hair and glasses, and pupil ellipse in many cases) will affect the pupil edge extraction in the process of positioning the pupil. This paper explores the removal of external disturbing factors and the extraction of complete and clear pupil images contrary to the influence of various disturbing factors on the conventional pupil locating algorithm. At first, Extract binarized pupil image by using improved Otsu algorithms. Then, a function to delete a small area and fill the hole is written to eliminate the appearing small area spots in the extracted image. Finally using the method of contour extraction and three-point fixed circle to realize the precise pupil location.

2 Principle of Experiment

The appearance of the human eye is constituted by the sclera, iris and pupil. The iris and pupil form the eyeball. The sclera is the white part of the eye's periphery. Eye center is the part of pupil. The iris is located between the sclera and the pupil. The gray value of the pupil is smallest. The gray value of the sclera is largest. The gray value of iris Between the two. The gray image information can be used to analyze the eye image and completing pupil location because of the large difference between the gray levels of different parts of the eye. So the content of my research is mainly divided into two aspects: the human eye and pupil center positioning.

(1) The main goal of eye positioning is to detect the position of the eye from the input image or video stream. Many researchers have used various methods to complete

*This work is supported by the National Key Research and Development Program of China (2016YFF010020002).

gaze tracking system. For example, detect the face from the image, and then detect the human eye from the human face. Although this approach is very accurate, and the accuracy is relatively high, but also increase the amount of computing system, affecting the system real-time. At the same time, people's head and camera are fixed relative motionless. It is completely unnecessary to make complex human eye positioning operation. So this article uses the region of interest extraction method to locate the human eye.

(2) Pupil location is based on the eye positioning including the image binarization, threshold selection, image morphology filtering operation through the gray image of the eye image to complete the precise positioning of the pupil. The use of contour extraction method and the three-point fixed circle method to determine the pupil coordinates.

3 Eye Location

This system fixed the camera and head relative position during the period of design the hardware of the system. So the relative position of the eyes and the camera is unchanged even if the head has a slight disturbance. That is to say, we only need to circle a part of the eye area accurately which we want from the one image frame. Area of interest (ROI) is an image area that selected from the image. This area is the focus of image analysis. Using ROI delineation can reduce image processing time and increase accuracy.

There are two ways to define ROI regions: One is to use Rect for representing a rectangular area. It defines a rectangle by specifying the rectangle's upper left corner coordinates and the length and width of the rectangle. Another way is to specify the range of rows or columns of interest to define ROI. This article uses the first method. The image size captured by the camera is 640*480 by taking a photo test. The relevant parameters of the ROI are obtained after comparing the position of the human face and the human eye image. Finally the eye region map is obtained as shown in the following figure 1.

Fig. 1 Eye area map

4 Pupil Location

4.1 Threshold Selection

It is necessary to obtain the best binary image that can distinguish the pupil portion from other portions in order to split out the pupil portion. Therefore, the selection of the threshold is the key to the whole process. The Otsu algorithms is a commonly used automatic pupil threshold segmentation algorithm, The following is the basic introduction of this algorithm.

4.2 Otsu Algorithm

The basic idea of Otsu algorithms is: The final binarized segmentation threshold is the pixel value with the largest difference between the two categoriescan which be determined by counting the pixel values of each point in the whole image and dividing it into two categories.

T is the segmentation threshold of the foreground and background of the picture. The number of foreground points in the whole image are w_0. Its gray value of the image is u_0. And the ratio of the background points of the image are w_1, and the average gray value of the image is u_1 .Then we can calculate the average gray of total image is u . T is determined by traversing from the minimum gray value to the maximum gray value. T is the optimal threshold for segmentation when the variance between classes is maximized. The image size is $M \times N$ pixels. L is the image gray level. The number of pixels whose gray value is less than T in the image is denoted as N_0 , and the number of pixels whose gray level is greater than T is denoted as N_1 . Then we can be calculated:

$$w_0 = \frac{N_0}{M \times N} \tag{1}$$

$$N_0 N_1 = M \times N \tag{2}$$

$$w_0 + w_1 = 1 \tag{3}$$

$$u = w_0 u_0 + w_1 u_1 \tag{4}$$

$$g = w_0 (u_0 - u)^2 + w_1 (u_1 - u)^2 \tag{5}$$

Final decision: $g = w_0 w_1 (u_0 - u_1)^2$ (6)

When T changes from 0 to 255, the T value is required to obtain the initial threshold for pupil segmentation which can make the overall variance reach the maximum. The threshold segmentation results shown in Figure 2 below.

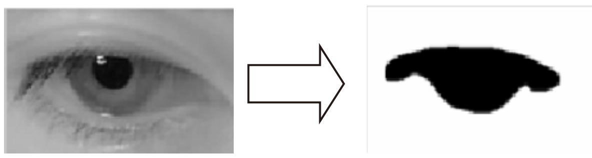


Fig. 2 Pupil binarization

From the figure above, we can see that simply using the maximum inter-class variance method to calculate the pupil segmentation threshold bigger than the true threshold in this case. It only remove most of the facial non-human eye area and can not accurately separated the pupil and the iris. There are also the effects of noise shadows on the eyelids and eyelashes. Therefore, we make improvements to this method to achieve complete segmentation of the pupil image and the background image.

4.3 Improved Otsu Algorithm

1. According to the histogram of the grayscale distribution of the image, the gray value of the non-occurring pixel is removed and the calculation range is redrawn.

During the experiment, It will increase the amount of unnecessary computation if the variance is also calculated for a pixel with a gray value of zero. Therefore, these zero pixel gray values may not be included in the calculation of the inter-class variance. Only the inter-cluster variance of the non-zero number of pixels in the image may be calculated. According to this idea, we can greatly reduce the number of variance calculations and improve the running speed.

2. Using the traditional OTSU algorithm to calculate the initial threshold t_1. Taking the first peak as the center and the initial segmentation threshold to obtain the final pupil segmentation threshold according to the histogram of the gray distribution. the specific steps are as follows.

(1) Beginning with the newly demarcated histogram grayscale minimum and traversing the histogram from the left half to the right half the size of the initial threshold to find the threshold peak center in the region.

(2) Taking this peak as the search center of the pupil region. The same length is extended to the right from the starting point to the search center to determine the pupil region segmentation compensation threshold t_2.

(3) Calculate the average of t1 and t2, using this value as the segmentation threshold t .

Pupil threshold segmentation results shown in Figure 3. The figure shows that the use of the proposed automatic pupil threshold segmentation algorithm can be precisely pupil segmentation. Effectively avoiding the Purkinchkin and upper and lower eyelid occlusion pupil area for the pupil region segmentation. However, a small area of noise clusters still exist in the binarized image of the pupil.

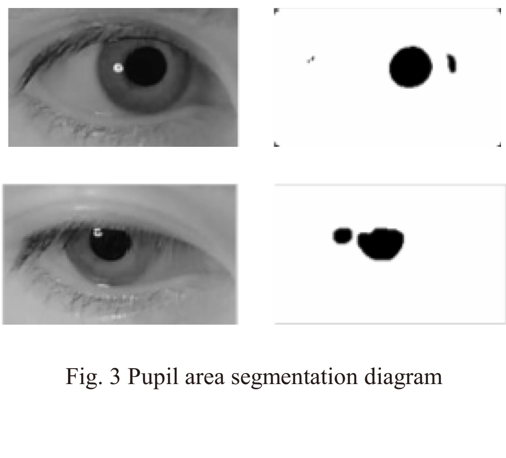


Fig. 3 Pupil area segmentation diagram

4.4 Elimination of Small Area and Hole Filling

Binarizing the eye image will appear eye corner and eyelash interference if we use the above method because the gray value of the eye approximate iris gray value, eyelash gray value approximate pupil gray value approximation. I wrote a self-excluding a small area and fill the hole function in order to remove the corner of the eyelashes and other parts and other interference. For binarized images, the method used to remove the hole is actually the same as removing the small area. So it is entirely possible to do this with the same function. The effect of the treatment as shown in Figure 4 below. These two functions can be achieved by using the regional growth method. The basic steps of the regional growth method are:

1. Select seed point p_0. Use the stack to represent the seed area and push the seed point to the seed stack.
2. Points out the first seed in the seed stack of the stack, and traverses the center 8 neighborhood pixels centered on this point.
3. Judging whether the traversed pixel (x, y) is already in the seed area. If not, determining whether the traversed pixel satisfies the similarity of adjacent seed dots. Pushing to the stack if the pixel satisfies the similarity.
4. Repeat steps 2-3 until the seed stack is empty.

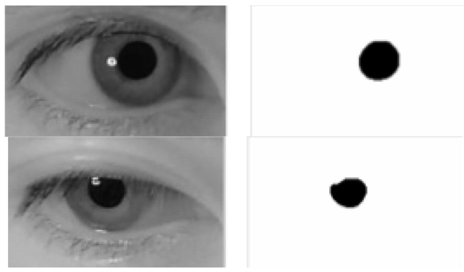


Fig. 4 Binarized image and Processed binarized image

5 Determine the Pupil Center

the accurate positioning of the pupil is achieved by the contour extraction method[11] and the three-point fixed circle on the basis of removing the small area of the image and the hole and the separation pupil and the puerchin spot. We can see that the background pixel gray value is 255 and the pupil pixel gray value is 0 based on the threshold extraction. The pupil contour extraction steps are as follows:

(1) The center pixel value is always reserved no matter what the value of the remaining eight pixels if the central pixel is 255.

(2) The central pixel value is changed to 255 if the central pixel is 0 and the remaining eight adjacent pixel values are 0.

(3) All the center pixel value is changed to 0 in the remaining cases.

We can obtain the pupil center by using the three-point ellipse method based on obtaining the pupil contour. Three points are randomly drawn from the outline of the contour. The center of the circle is obtained by using the three-point fixed circle formula. And the center of the obtained fitted circle is saved. Under the condition of guaranteeing the participation of new points, extract three points from the edge point again, confirm and save the center of circle again. The above operation is repeated N times. Let N be equal to 1.5 times the edge point. The resulting N sets of center-point data are averaged as the final result.

Grant the three coordinate points taken from the edge are respectively: $x_1(x_1, y_1)$、$x_2(x_2, y_2)$、$x_3(x_3, y_3)$, The center coordinates $x_0(x_0, y_0)$ can be obtained by the following formula.

$$x_0 = ((B-C)\times y_1 + (C-A)\times y_2 + (A-B)\times y_3) \ / \ (2\times G) \tag{7}$$

$$y_0 = ((C-B)\times x_1 + (A-C)\times x_2 + (B-A)\times x_3) \ / \ (2\times G) \tag{8}$$

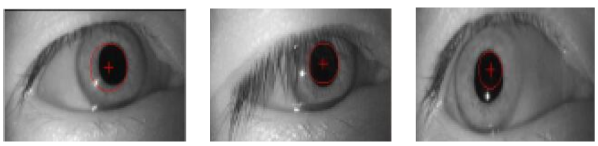
$$A = x_1^2 + y_1^2 \quad B = x_2^2 + y_2^2 \quad C = x_3^2 + y_3^2 \tag{9}$$

$$G = (y_3 - y_2)\times x_1 + (y_1 - y_3)\times x_2 + (y_2 - y_1)\times x_3 \tag{10}$$

You should maintain a certain distance to ensure that the three edge points drawn from the edge can not be too close in practice. Then the third one should be taken from the right edge if two points are taken from the left edge. The data in the center of the circle should be excluded which is farther from the medianand that is not involved in the calculation of the final result when calculating the average of the center of the circle. The specific effect as shown below:

6 Experimental Results

This article contrast test the method in the literature[8], the method in the literature[10] and the accuracy of this method to locate the pupil in order to investigate the effectiveness of the algorithm. From the literature[8] method of positioning results can be seen that the center of the pupil deviation largerly from the actual pupil center value. The results of the method[10] are more accurate. But there is still a small amount of pupil edge loss due to the image pre-processing. However, the method of this paper obtains a better binary image by removing external disturbances and Puerchin spot and realizes the complete extraction of pupil edge information which makes the pupil localization result more accurate. The specific effect as shown below:



the method in the literature[8]

the method in the literature[10]

This article method

Fig. 5 Contrast effect of pupil center localization

6.1 Method Timeliness

Table 1 has shown for us that this paper takes less time and more timeliness after comparison with the first method, second method and this article method in the method of timeliness .

Table 1: Method Timeliness

Method(Unit (ms))	Pic. 1	Pic. 2	Pic. 3
the method in the literature[8]	350	360	345
the method in the literature[10]	244	239	231
This article method	129	121	115

7 Conclusion

In this paper, We improve the traditional Otsu algorithm and separate the pupil image with the background image. A method of removing small areas and filling holes has been written to eliminate the effect of small area on binarized pupil images. Finally, the center coordinates of the pupil are calculated quickly and accurately by the contour extraction method and the three-point fixed circle method, and the precise positioning of the pupil is achieved. Practice has proved that this method is an effective method of pupil location which is a good extraction of the pupil binary image and pupil center location .

References

[1] Jung J W, Sato T, Bien Z. Dynamic footprint-based person recognition method using a hidden Markov model and a neural network[J]. International Journal Intelligent System, 2004, 19(11):1127-1141.

[2] Wildes R P. Iris recognition： an emerging biometric technology[J]. Proceedings of IEEE, 1997, 85(9): 1348-1363.

[3] Liu L Z, Guang M L. An Improved Rapid Iris Location Algorithm[J]. Computer Application,2008,28(3):674-676.

[4] Tisse C, Martin L, Torres L. Person identification technique using human iris recognition[C]. Proc Vision Interface.[S.l.] IEEE Press,2002:294-299.

[5] Carlos H Morimoto, Marcio R M Mimica. Eye gaze tracking techniques for interactive applications[J].Computer Vision and Image Understanding, 2005; 98: 4~24.

[6] Otsu N.A threshold selection method from gray-level histogram[J].IEEE Trans on Actions on SMC，1979，9：652-655.

[7] Feng H Wu. A Study of Image Contour Extraction Based on Computer Vision Measurement Technology [J] .Journal of Measurement and Measurement, 2007,28 (1): 18-22.

[8] Lin Y T, Lin R Y, Lin Y C, etal. Real-time eye-gaze estimation using a low-resolution webcam[J]. Multimed Tools Appl,2013,65:543-568.

[9] Xiaohui Yang, Jiande Sun, Ju Liu, Jinyu Chu, Wei Liu, and Yuling Gao, “ A gaze tracking scheme for eye-based

intelligent contorl”, Proceedings of the 8th World Congress on Intelligent Control and Automation, Jinan, China.
[10] J. Sigut, and S. Sidha, “Iris center corneal reflection method for gaze tracking using visible light”, IEEE Trans. Biomedical Eng., vol. 58, no. 2, Feb. 2011, pp. 411-419.
[11] E.D. Guestrin, M. Eizenman. General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections[J]. IEEE Trans. Biomedical Eng., vol. 53, 2006, no. 6: 1124-1133.