A Method for Assessing the Pupil Center Coordinates in Eyetracking with a Free Head Position

Gennadiy I. Gromilin* and Nikolay S. Yakovenko

Institute of Automation and Electrometry SB RAS, 1 Academician Koptyug Ave., Novosibirsk 630090, Russia

* e-mail: gromilin@iae.nsk.su

Abstract. IR-illuminated Eyetracking systems include cornea reflection and pupil center coordinates detection to calculate the operator's gaze fixation point. When you turn a view of a large angle, some of the frames are blurred, and the coordinates are unreliable. The article describes a method for determining the center of the pupil in the gaze fixation system for operation at an increased camera frame rate. Comparison with known algorithms is given. The algorithm execution average time is about 1.2 ms on a typical office computer by processing images in fragments of the order of 340x240 pixels. © 2021 Journal of Biomedical Photonics & Engineering.

1 Introduction

Person's gaze determining systems are beginning to be applied in various areas of activity. In technology – the devices control, first, remotely unmanned devices [1, 2]. In medicine, it is the detection of the human condition in psychology [3], the study of the central nervous system [4], the assessment of visual acuity in ophthalmology [5]. In marketing, Eyetracking is used to study how to inspect a scene when promoting products [6]. The considered system is designed to device control.

The most widespread approaches are systems with IR illumination of the eyes, which provide the greatest contrast between the pupil and the iris and the least influence of extraneous illumination. For each video stream frame in such systems, the main algorithms are executed:

- finding eye cornea reflections from illuminators and determining their coordinates;
- finding the pupil and determining its center coordinates;
- calculation gaze points from the cornea reflection and the pupil center coordinates.

With free head position, the size of the eye area is significantly smaller than the size of the frame. To reduce the processing time, the eye area is localized first. Region of interest (ROI) selection is used when processing various types of images, for example, medical [7, 8]. In this paper, the task is simplified by presence of the cornea reflections from IR illuminators, which is brighter than the rest of the image. Cornea reflections are easy to detect, but you need to sort them into true and parasitic ones.

During saccades movements eye movement speed reaches 450–700 deg/sec. In Eyetracking system with a frame rate of 30–50 Hz the eye image is blurred, which prevents the necessary accuracy of the gaze direction registration.

The proposed algorithm uses threshold selection of the pupil boundary and adaptive threshold estimation, which require significantly less computational resources.

2 Materials and Methods

When the camera is not mounted on a person's head, it is necessary that the camera image capture an area approximately twice the distance between the eyes in order to track the operator head's movement. In this case, it is necessary to determine the eyes position in the image before the pupil searching.

The system uses IR LED illumination, which increases the contrast between the pupil and the iris. In addition, bright small size corneal reflections are formed on the cornea of the eye. Therefore, it is possible to allocate the fragments including the eye image, and image processing should be carried out in the fragments. To reduce the computational resources needs, the cutoff threshold carries out the pupil boundaries selection.
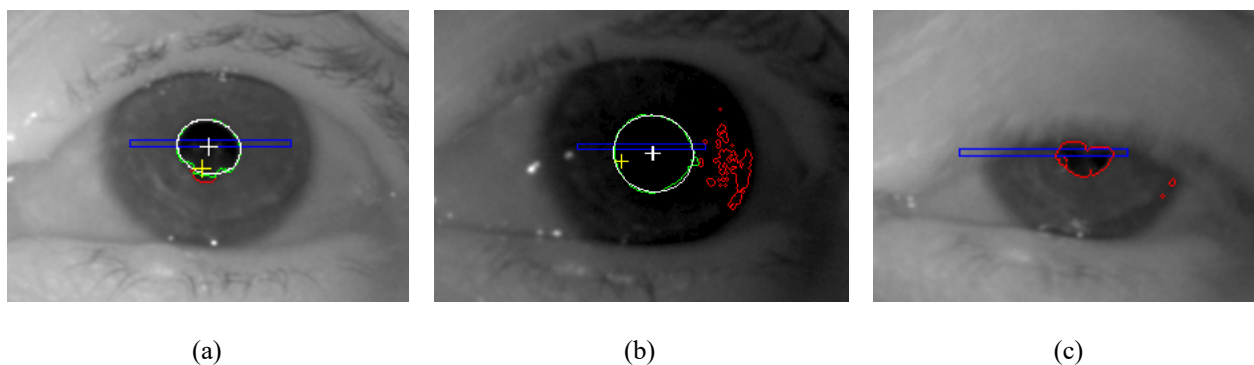
(a)   (b)   (c)

Fig. 1 Selection of the pupil boundaries: (a) a fragment with good selection, (b) elimination of unnecessary contours, (c) reduction of the area when winking. Fragment sizes are about 240×180 pixels.

While the difference in the pupil and the iris brightness is small, and the lighting conditions may be changed, an adaptive threshold is applied, which is adjusted at each frame in which the pupil is detected. Fig. 1a shows the stages of eye image processing.

Start point for pupil search is marked with a blue cross. The threshold level is estimated inside the blue rectangle for red outlines. RANSAC (Random Sample Consensus) algorithm [9] has left a green contour, along which the white ellipse parameters were estimated.

After the detection and selection of corneal reflections (Fig. 2), the fragment with the eye image is allocated and the procedure for isolating the pupil boundaries and the ellipse parameters is performed (Fig. 1a).

To obtain the coordinates of the pupil center, the following steps are performed:

- thresholding the image;
- searching areas boundaries and selection a specified range length contour (Fig. 1b);
- selecting boundary points inliers by the RANSAC [9];
- refinement the ellipse parameters by the least squares method [10];
- estimation current frame threshold value and recursive filtering by time (frame number).

At the first pupil search, the threshold is estimating before threshold procedure. The assessment is provided on the area outlined by a blue rectangle. The area is getting out above patches of corneal reflections as IR LEDs are below the monitor. After filtration for noise reduction, the threshold level is getting out in the middle between a pupil and an iris brightness. This approach works both at dark and at light pupils (Fig. 2).
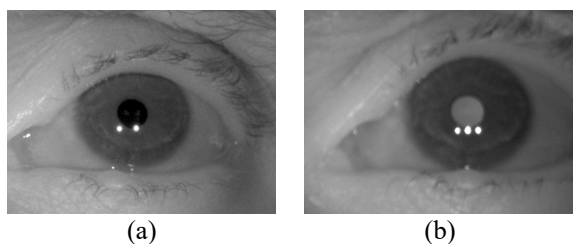
The binary image is obtained by using the threshold. The boundary contours of the regions are found and highlighted by OpenCV findContours() method, described in the Ref. [11], and discarded the contours with boundary length outside the specified range.

The chosen contour boundary points are checked by RANSAC for belongs to the ellipse model. The parameters of the ellipse are clarified by the least square method from OpenCV fitEllipseDirect() function.

A change of the pupil area is also checked, and if the change is greater than the specified one, the frame is discarded. Fig. 1c shows the change in pupil size when winking.

After obtaining the corneal reflection and the pupil center coordinates, the three-dimensional cornea center coordinates and the gaze direction are determined. The intersection of the gaze vector with the monitor plane defines the gaze fixation point.

3 Results

Fig. 3 shows the layout diagram of the breadboard device. The camera and IR LEDs are located under the industrial monitor. Left and right side LEDs creates cornea reflections to determine the three-dimensional eyes coordinates. Center LEDs around the lens are used to work with light pupil. The lighting current levels are adjustable.

Fig. 3 Device layout diagram.

(a)   (b)

Fig. 2 Fragments about 280×220 pixels of an eye images with a dark (a) and light (b) pupil.
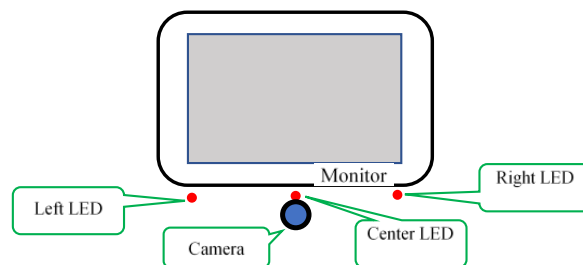
The gaze detection system used a fully programmable camera with USB 3.0 output and frame size up to 1440×1080 pixels. Most of the frame time on the

computer is used to take the image from the camera. There is no threading parallelization.

The software modules are written in C++, image processing was performed using the OpenCV library [12].

The determination of the pupil center was carried out in fragments of about 350×240 pixels, including the eye image (right or left by choice).

Since described image frame processing algorithms contain iterative procedures, the processing time depends on the scene in the frame. During the experiments, the procedure of calibration, calibration test and two seconds fix point hold was carried out. The sequence contains a sufficient variety of scenes. The real time frame processing is carried out to obtain the coordinates of the gaze fixation point and the frames recording at camera frame rate. Then, the recorded to a hard disk sequence may be processed from the disk at the fastest speed for a computer. By disabling the given algorithm, it is possible to estimate the average running time of the algorithm by the difference in processing time.

Frame file reading and entire image processing lasts about 8 ms on a 3.6 GHz processor, and the pupil center coordinates are obtained in an average of 1.2 ms. Depending on the image quality, the edge points outliers cutoff time by the RANSAC algorithm may be significantly increased in some frames, but the image arrays buffering evens out such deviations. Frame loss was also monitored. The system was tested at frequencies of 50 Hz and 100 Hz. At a frequency of 200 Hz, the entire frame time is occupied by the exchange with the camera and processing in real time is impossible. Executing the program in multiple threads allows you to bypass these limitations.

With a camera operating frequency 50 Hz, the transfer of the gaze fixation point from one monitor corner to another takes 2–3 frames. The eye is blurred and the frames have to be discarded. At a frequency 100 Hz, one image is blurred, and the pupil center position is determined on it, albeit with an error.

During the processing a real sequence, it is impossible to separate the algorithm operation influence error and fast eye movements. The system noise effect on the accuracy of the ellipse center coordinates determination was evaluated on a sequence formed from one real frame with different realizations of noise. The noise level, estimated from the difference between adjacent real frames, is 1.37 image sampling level. The ellipse center coordinates standard deviation Xe and Ye for 100 sequence frames at different noise levels is given in the Table 1.

Table 1 shows that stability is maintained up to high noise levels. When operating at a frequency 100 Hz, the noise level is slightly higher, because when the exposure time is decreased, the camera gain must be increased.

The system operation was tested on operators of different ages, different irises colors, and different pupil sizes both in laboratory and in the field.

Table 1 The noise value effect on the accuracy on the pupil center coordinates estimating.

Noise STD	Xe STD, pix.	Ye STD, pix.
1.0	0.06	0.06
2.0	0.09	0.07
3.0	0.14	0.13
4.0	0.37	0.24
8.0	0.34	0.29

4 Discussion

The well-known algorithm for pupil finding and determining its center coordinates Starburst [13] was developed for a head-mounted device. The MatLab version works well but slowly. There is approximately one second per frame of the camera image. A simplified version in C++ manages to process at 50 fps frame rate. Nevertheless, the pupil's border is very noisy because the radius derivative is evaluated in 4–6 pixels increments to highlight the level increment at low contrast.

Attempts to speed up the detection for the pupil center have been made several times. For example, in Ref. [14], the pupil contour is approximated by a sinusoid (SET method). The comparison of the accuracy of determining the pupil center by the Starburst and SET methods is carried out. However, approximation methods generally use iterative approximation and are slow. The algorithm was tested on small images containing the eye area that is typical for head-mounted devices.

The proposed algorithm selects a smoother border that is close to the real one. However, it uses the starting point after corneal reflection highlighting. StarBurst can find the border without specifying the starting point but only in images slightly larger than the eye size. You still need to find the eye area and select a fragment.

The proposed algorithm works on the both light and dark pupil. Many algorithms work with light pupil. We focused on dark pupil option. The hardware implementation is simpler, does not require a small diameter lens with central LEDs around it. Only two side LEDs are required to calculate the 3D eye position. There is also no need to synchronize the illuminators switching with camera frame changes.

5 Conclusion

The proposed algorithm for the determination of the pupil center coordinates provides stable results with a small scatter. The system was tested at frame rates of 50 and 100 Hz. The algorithm can be used at higher camera frequencies when using a camera with a higher channel bandwidth. Frames image processing in parallel with obtaining the next frame will also allow the pupil center coordinates assessing at higher frequencies.

It is also necessary to reduce computational resources requirement of the cornea reflection extraction and selection algorithm, which takes up most of the processing time.

As already mentioned, the main requirement is the IR illumination presence, which increases the contrast between the pupil and the iris. Hence the main limitation: the direct sunlight exposure lack. The use of the system is limited to rooms and moving objects with closed volumes.

A free head position reduces operator fatigue. In medical applications, it allows the study of patient behavior *in vivo*.

Disclosures

All authors declare that there is no conflict of interests in this paper.

Acknowledgements

References

1. P. Biswas, J. DV, "Eye Gaze Controlled MFD for Military Aviation," In 23rd International Conference on Intelligent User Interfaces, 7–11 March 2018, Tokyo, Japan, 79–89 (2018).
2. J. P. Hansen, A. Alapetite, I. S. MacKenzie, and E. Møllenbach, "The use of gaze to control drones," Proceedings of the ACM Symposium on Eye Tracking Research and Applications, 27–34 (2014).
3. T. E. Petrova, E. I. Riekhakaynen, and V. S. Bratash, "An Eye-Tracking Study of Sketch Processing: Evidence From Russian," Frontiers in Psychology 11, 297 (2020).
4. E. A. Novikov, I. A. Vakoliuk, R. D. Akhapkin, I. A. Varchak, I. G. Shalanginova, D. A. Shvaiko, and E. A. Budenkova, "Automation method of computer oculoghaphy for research of the central nervous system based on passive video analysis," Machine Learning and Data Analysis 1(12), 1–12 (2015).
5. L. Cercenelli, E. Marcelli, "Eye Tracking in Ophthalmology: A Glimpse Towards Clinical Practice," EC Ophthalmology ECO 01, 16–18 (2017).
6. P. Chandon, J. W. Hutchinson, E. T. Bradlow, and S. H. Young, "Measuring the Value of Point-of-Purchase Marketing with Commercial Eye-Tracking Data," SSRN Electronic Journal (2007).
7. R. S. Hessels, C. Kemner, C. van den Boomen, and I. T. C. Hooge, "The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli," Behavior Research Methods 48, 1694–1712 (2016).
8. D. S. Raupov, O. O. Myakinin, I. A. Bratchenko, V. P. Zakharov, and A. G. Khramov, "Multimodal texture analysis of OCT images as a diagnostic application for skin tumors," Journal of Biomedical Photonics & Engineering 3(1), 010307 (2017).
9. M. A. Fischler, R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM 24(6), 381–395 (1981).
10. A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," IEEE Transactions on Pattern Analysis and Machine Intelligence 21(5), 476–480 (1999).
11. Satoshi Suzuki, K. Abe, "Topological structural analysis of digitized binary images by border following," Computer Vision, Graphics, and Image Processing 30(1), 32–46 (1985).
12. OpenCV (accessed 24 May 2021) [https://opencv.org/].
13. D. Li, D. J. Parkhurst, "Starburst: A robust algorithm for video-based," Elselvier Science (2005).
14. A.-H. Javadi, Z. Hakimi, M. Barati, V. Walsh, and L. Tcheang, "SET: a pupil detection method using sinusoidal approximation," Frontiers in Neuroengineering 8, 4 (2015).