

# Head-mounted eye tracker based on android smartphone

*Alexander V. Fisunov<sup>1,\*</sup>, Maxim S. Beloyvanov<sup>1</sup>, and Iakov S. Korovin<sup>2</sup>*

<sup>1</sup>Don State Technical University, 346500 Shakhty, Russia

<sup>2</sup>Southern Federal University, 344006 Taganrog, Russia

**Abstract.** This paper presents a wearable eye tracker that tracks points of interests of user at videostream showed at smartphone screen. The system consists head-mounted case for smartphone, point of interest detection algorithm, the software developed for this purposes, and Android smartphone used to show videostream, estimate point of interests at video, and log estimated data into device internal memory.

## 1 Introduction

Eye trackers have been developed for a long time, and their area of application has a wide spectrum: from the study of human behavior and thought processes to the developing human-computer interfaces for contactless control of multimedia systems [1, 2, 3, 4, 5]. One of the applications of such systems is gathering information about where a person is looking, and what objects in a video or image are interesting to him. Currently, technologies that allow to assess the direction of a person's gaze are being actively developed. Many different devices appear, as a rule, these devices (glasses) are located close to the human eye, require calibration for their work, and additional separate device for performing estimation (personal computer or laptop), in some cases additional sensors or lights are added. With this approach, to solve the problem of tracking points of interest, you need to collect specialized equipment, and purchase software that can process the data received. On the other hand, as a video capture device, screen for show video, data processing, backlight, the most common smartphone on the Android platform can be used, and only need the case to hold the smartphone near the eyes.

## 2 The proposed system

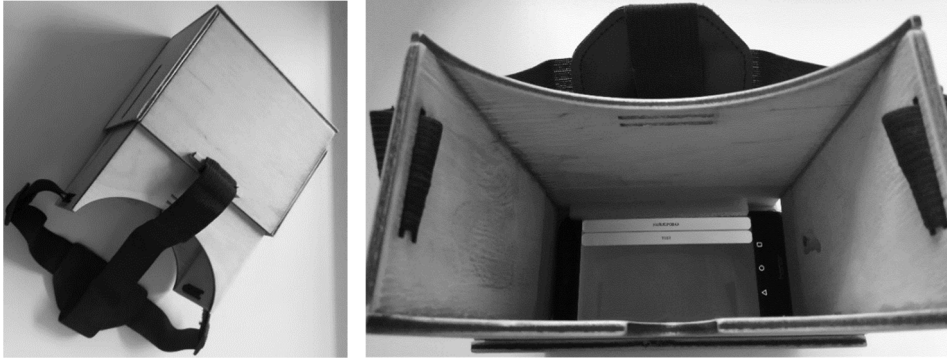
The paper proposed a device for tracking and logging points of interest in the video displayed on the smartphone screen. The device consists of a case for fixing the smartphone in the eye area that can be easily printed on a 3D printer, a bluetooth joystick for launching the application and initial calibration, and the smartphone itself. Smartphone is used to store and display video (from device gallery), capture video stream from the frontal camera and to track

---

\* Corresponding author: [fisunov.a.v@gmail.com](mailto:fisunov.a.v@gmail.com)

the center of the pupil, and estimate the point of interest on the displayed video. Currently, even the inexpensive smartphone on the Android platform allows to solve these tasks.

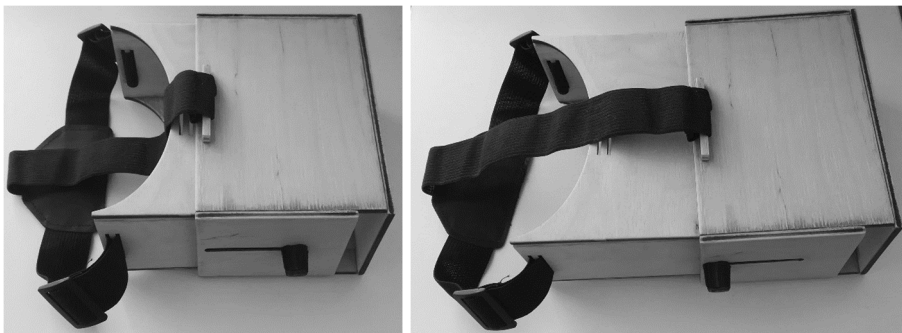
To fix the smartphone in the eye area, it is proposed to use a special case into which the smartphone is inserted, and which can be fixed on the head with the belts. The case design, with the smartphone installed, is shown in figure 1.



**Fig. 1** Case with a smartphone installed.

The bluetooth joystick is used to launch the application and calibrate the system. As can be seen from figure 1, the case does not have slots for the admission of natural light, and it does not provide additional lighting. Additional lighting is not required because the application itself adjusts the brightness level when calibrating and displaying video. When a video is shown, a special frame is displayed around the edges of the image, the brightness of which is determined by the average brightness of the scene being shown.

Depending on the person's vision, it may be necessary for the phone to be located closer or farther from the eyes, for this purpose a sliding case design is provided, as can be seen at figure 2.



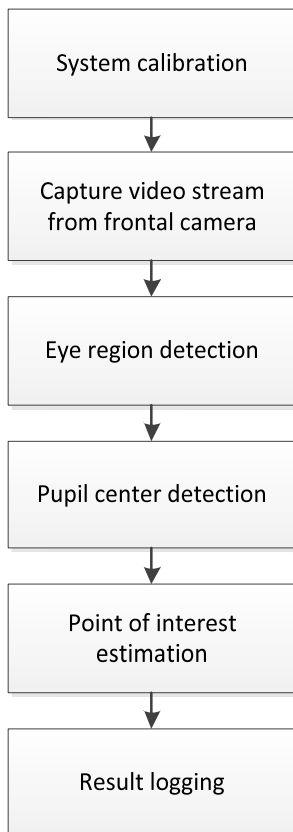
**Fig. 2** Adjusting the distance between the eyes and the smartphone.

The principle of work of the device is as follows, the user inserts a smartphone into the case and fixes the case on the head. Then starts the application with the bluetooth joystick, and perform calibration. The calibration process consists of a consistent look at a set of points.

After completing the calibration, it is possible to select and show a video from the smartphone gallery, tracking the points of interest of the user, and logging this information into internal memory of the device. Consider now the algorithm for the tracking points of interest.

### 3 Points of interest detection algorithm

The estimation of points of interest in the video is performed in several stages. The algorithm of the POI tracker is presented in the figure 3.



**Fig. 3** Proposed algorithm for point of interest detection

At the first stage, the video stream is captured from the frontal camera of the smartphone. The camera can have several standard locations (left, center, right), so the program offers a choice current camera location.

In addition to displaying the video received from the frontal camera, as mentioned above, a special luminous frame is added to adjust the level of illumination (this is necessary when watching videos with dark scenes). Given the location of the camera, the coordinates of the frame area in which the image of the eye is located are calculated. Eye region is copied into a buffer for the further processing. The center of the pupil estimation is performed only within the buffer, and not the entire received frame, which is done in order to increase the algorithm performance.

To detect the center of the pupil, the Fabian Timm and Erhardt Barth method [6] is used, which allows real-time monitoring of the centers of the pupils on a video stream.

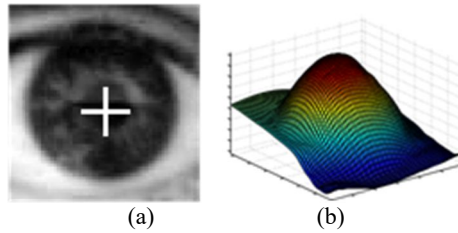
The coordinates of the pupil center as follows from article [6] are the intersection of the image gradients, that is, there is necessary to analyze the vectors field of the image whose maximum is the eye center orienting.

$$c^* = \operatorname{argmax}_c \left\{ \frac{1}{N} \sum_{i=1}^N (d_i^T g_i)^2 \right\} \tag{1}$$

$$d_i = \frac{x_i - c}{\|x_i - c\|_2}, \quad \forall i: \|g_i\|_2 = 1 \tag{2}$$

where  $c$  represents possible center of pupil and  $g_i$  the gradient vector at position  $x_i$ .

The displacement vectors  $d_i$  are normalized to obtain an equal weight for all positions of the pixels. The resistance increase to linear changes in lighting and contrast is accomplished by scaling the gradient vectors up to the block length. An approximate estimate sum of point products for different centers is shown in Figure 4, where the objective function determines the maximum in the center of the pupil.



**Fig. 4** Pupil center detection using evaluation (1). (a) detection results (b) 3D plot with the objective function achieves a maximum at the center of the pupil [6]

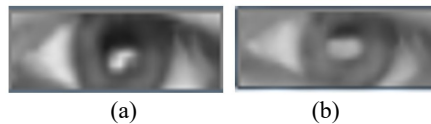
The assessment of the center of the pupil can be estimated with an error caused by the dominance of the eyelids, eyelashes, wrinkles, in combination with the low contrast between the iris and the sclera. To solve this problem, the author [6] uses the weight function  $w_c$  for each possible center  $c$ , since it is assumed that the probability of a dark center of the pupil is higher. Then the weight function is written as [6]:

$$\operatorname{argmax}_c \frac{1}{N} \sum_{i=1}^N w_c (d_i^T g_i)^2 \tag{3}$$

$$w_c = I^*(cx, cy) \tag{4}$$

where  $I^*$  represents smoothed by Gaussian filter and inverted input image, and  $w_c$  represents value of  $I^*$  at point  $cx, cy$ .

Due to the small distance between the smartphone screen and the eyes, situations are possible when the flare occupies a significant part of the pupil, which adversely affects the detection of its center. Examples of images of the eye taken from the front camera of the smartphone are shown in the figure 5.



**Fig. 5** Significant flare area on the pupil.

This problem is partially solved by correcting the brightness of the displayed video. Also brightness of the rectangular frame used as the backlight is changed, which was discussed above.

The result of the above algorithm will be the coordinates of the center of the pupil. The next step is to transform these coordinates into the coordinates of the point of interest in the video. A separate application module receives data from the frontal camera and uses the algorithm described above to determine the center of the pupil. At this time, the user is

consistently shown a set of points in different parts of the screen. When looking at a point, the user presses the joystick button and the module remembers the coordinates of the pupil when looking at the specified point and show the next point. After performing this operation for all points, the calibration data is memorized and used to track the point of interest.

After calibration has been completed, the third module of the system allows the user to select a video from the device's gallery, start showing this video, tracking points of interest and logging their coordinates into the device's internal memory.

## 4 Conclusion

The paper proposed a system for assessing points of interest on video using a smartphone on the Android platform. The proposed system allows the user to watch the video from device gallery, track the center of the pupil using frontal smartphone camera, estimate the point where the user is looking at each moment of time and log received data.

The paper is published due financial support of RFBR (Russian Foundation for Basic Research), project 18-29-03225 mk.

## References

1. K. Holmquist, *Eye Tracking: A comprehensive guide to methods and measures* (OUP Oxford, 2011)
2. A.T. Duchowski, *Behavior Research Methods, Instruments, and Computers*, **34**, 4, 455 (2002)
3. R.J.K. Jacob, K.S. Karn, *Minds Eye*, 2 (3), 573 (2003)
4. D. Hansen, A. Pece, *Computer Vision and Image Understanding*, **98**, 1, 155 (2005)
5. H. S. Al-Khalifa, P.G. Remya, *eLearn Magazine*, **2010**, 6, 8 (2010)
6. F. Timm, E. Barth, *VISAPP*, **11**, 125 (2011)