International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 15 (2017) pp. 4985-4988


# Remote Eye Tracking Method for UHD Monitor Based on High-resolution Infrared Imaging

**Eui Chul Lee**[1]

*Intelligent Engineering Informatics for Human, Sangmyung University, Seoul, Korea.*

[1]*Orcid: 0000-0001-6504-3333*

**Abstract**

This paper proposes an eye tracking method that uses a remote camera and an infrared light illuminator. As the remote camera can capture five-megapixel images of eight-bit gray at 15 frames per second through a USB 3.0 interface, it guarantees a sufficiently high spatial resolution of the region occupied by the eye to continuously track the gaze. Based on a mapping matrix of the pupil-reflection vectors and the four corners of the monitor acquired from a four-point initial user-dependent calibration, the positions of the gaze are successively calculated by analyzing the magnitude and direction of the pupil-reflection vector, where the reflection is due to the infrared light illuminator. The user does not need to wear a device on the head in the proposed method, and no calibration is needed to obtain the relations among the positions of the camera, the illuminator, and the monitor. Since the camera has a high spatial resolution and a wide field of view, and as the proposed method tracks only one eye, some facial movement is allowed. The average error incurred by the method for gaze point estimation was approximately 0.9 degrees at a Z-distance of 100 cm on a 42-in LCD monitor with UHD resolution.

**Keyword:** Eye tracking; USB 3.0 camera; Infrared light imaging

## INTRODUCTION

Eye tracking is used to acquire the position of a user's gaze as this may reflect the object of his or her interest at any given time. A considerable amount of research has been devoted to developing convenient and accurate eye tracking systems that have been applied to human-computer interface, Virtual Reality, immersive gaming, video conferencing, computer interfaces for the disabled, and a variety of other areas [1-7].

Prevalent camera vision-based eye tracking methods can be categorized into two groups, remote camera-based and wearable camera-based methods.

In remote camera-based methods, a remote camera with a zoom lens and a narrow-view camera track and capture the user's eye region based on a panning and tilting mechanism that uses information about the position of the eye acquired from the wide-view camera [7-9]. The methods in this category are convenient because they do not require that the user wear a device. However, the cost, complexity, and motor operational latency of the pan-and-tilt camera are disadvantages. A 3D analysis-based eye tracking method has been proposed [8]. Although it can generate an accurate gaze vector based on a 3D model of the eye, it is expensive and requires an elaborate stereo camera system that involves a complex calibration procedure.

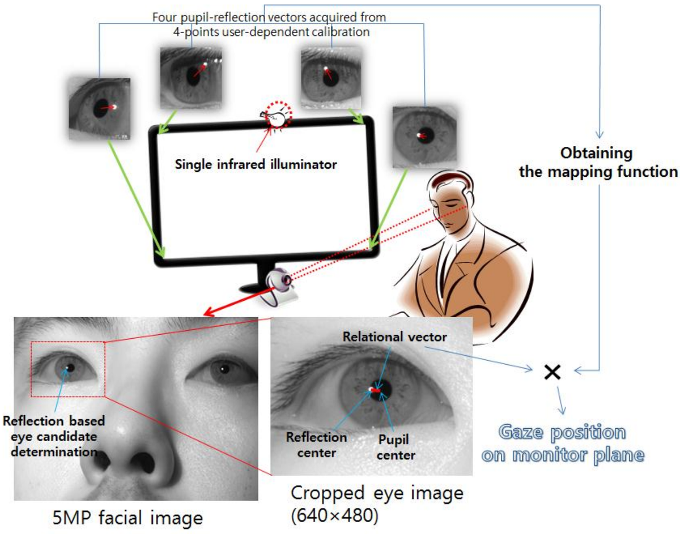
Wearable camera-based methods are independent of the user's facial movements because the wearable camera can be moved accordingly. In this class of methods, images of the eye are captured up close [10][11]. However, once the wearable device is slipped down after initial user-dependent calibration, the accuracy of the gaze position calculated using the initial calibration-based mapping function can degrade. The wearable device is also inconvenient for the user.

To solve the above problems, this paper proposes a non-wearable eye tracking method. To acquire high-resolution images of the eye at a distance, a five-mega pixel (MP) camera is used in a USB 3.0 interface. The relevant regions of the face and the eye are sequentially extracted using the Adaboost method, and the pupil center - reflection center vector is obtained from this information. Based on an initial four-point user-dependent calibration and the corresponding mapping function, the gaze is calculated on a monitor.

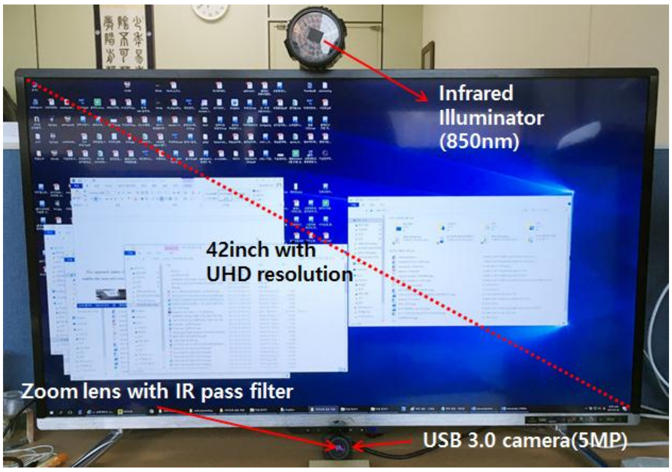
The remainder of this paper is structured as follows: The proposed method and the devices used are described in Section 2, and the experimental results and the conclusions of this study are provided in Sections 3 and 4.

## PROPOSED SYSTEM

The proposed eye tracking system contains an infrared light illuminator and an infrared camera, as shown in Fig. 1 (a) and (b). The infrared illuminator consists of 72 IR-LEDs (infrared-light-emitting diodes) with peak wavelength and illumination angles of 850 nm and 10 degrees, respectively. To capture images of the dark pupil and the bright iris irrespective of the amount of melanin in the latter, the infrared camera is made by combining the USB-3.0 interface camera with a zoom lens and a band-pass filter with a wavelength of 850 nm.




(a)




(b)

**Figure 1**. Proposed remote eye tracking system. (a) Overview of the proposed method. (b) Actual configuration of proposed eye tracking system

The camera could capture images at a resolution of 5 MPs (2448 (H) × 2048 (V)) of eight-bit gray at 15 frames per second. To transfer the large amounts of image data to the computer, the USB 3.0 interface is used because of its bandwidth of 625 MB/s. The zoom of the controllable lens was assembled in the camera. To allow only infrared light to pass, the band-pass filter was attached to the front of the zoom lens.

In the proposed method, a four-point initial user-dependent calibration was performed, and no other calibration was needed to configure the geometric relations among the camera, the monitor, and the illuminators. However, to guarantee optimal and stable specular reflection (SR) with no distorted eye images, the infrared illuminator and the infrared camera were located at the top and the bottom, respectively, relative to the horizontal center of the monitor, as shown in Fig. 1.

**PROCEDURE**

The proposed system proceeds according to the stages shown in Fig. 2.

A sub-sampled image (306 (H) × 256 (V) pixels) captured by the 5-MP camera is first generated to minimize the computational complexity of Adaboost-based face and eye detection. This is shown in step (a) in the flowchart in Fig. 2 (a) and Fig. 3 (a).

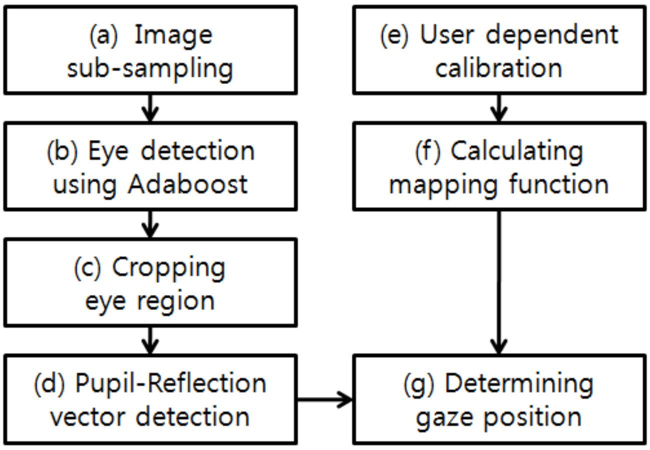



**Figure 2**. Processing stages of the proposed method

The region occupied by the eye in the sub-sampled image is detected by using the Adaboost method [12], which is commonly used to detect regions of interest (ROI). This is shown in step (b) in Fig. 2 (a) and in Fig. 3 (a), where the training information of the eye region is obtained from 500 infrared light images of the eye.

Based on the center of the eye box in the sub-sampled image, a high-resolution image is cropped from the original 5-MP image, as shown in step (c) of Fig. 2 (b) and in Fig. 3 (b). The cropping size was empirically determined to be 640 (H) × 480 (V) pixels.

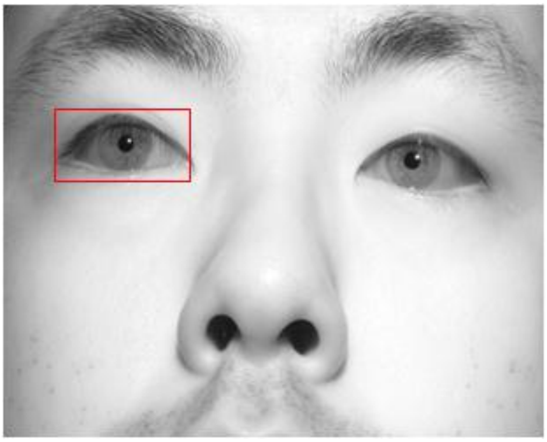
For the cropped high-resolution eye image, the circular edge detection method is used to detect the center of the pupil, and is expressed by the following equation [13]:

$$argmax_{(r,x_0,y_0)}\left|\frac{\delta}{\delta r}\oint_{r,x_0,y_0}\frac{I(x,y)}{2\pi r}ds\right| \qquad (1)$$

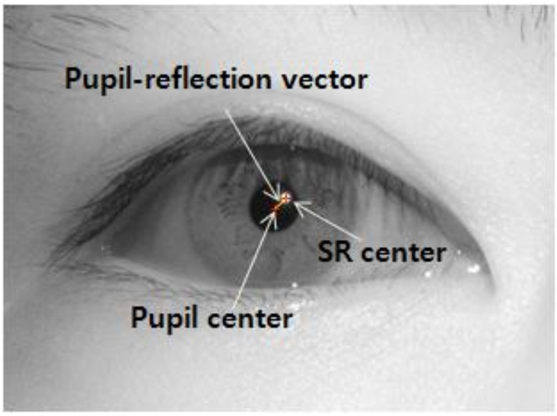
In Eq. (1), $I(x, y)$ represents intensity at $(x, y)$. $(x_0, y_0)$ and $r$ represent the center and the radius of the circular template, respectively. That is, the maximum difference between the summations of the intensities of the two circles is determined to be the center of the pupil $((x_p, y_p))$ as shown in step (d) in Fig. 2 and in Fig. 3 (b).

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 15 (2017) pp. 4985-4988


Following this, in the 150 × 150-pixel local region based on the center of the pupil, the center of specular reflection (($x_r$, $y_r$)) was detected as in step (d) of Fig. 2 and in Fig. 3 (b). To this end, binarization (threshold: 253), component labeling, and size filtering are sequentially performed because specular reflection is significantly brighter than the rest of the eye region in the local square.



(a)




(b)

**Figure 3**. Example of processing results: (a) A sub-sampled image, and the results of eye detection. (b) Cropped eye image and the detected pupil-reflection vector.

Two coordinates ($x_p$, $y_p$) and ($x_r$, $y_r$) are measured relative to the origin of the initial full image because the position of the cropped eye image continually changes. From the two geometric centers ($x_p$, $y_p$) and ($x_r$, $y_r$), the pupil-reflection vector (($x_{pr}$, $y_{pr}$)) is obtained.

To calculate the position of the gaze, user-dependent calibration is performed at the initial stage as shown in step (e) in Fig. 2. In our method, four-point calibration was used. Assuming that the four pupil-reflection vectors acquired when gazing at four points (($m_{x1}$, $m_{y1}$), ($m_{x2}$, $m_{y2}$), ($m_{x3}$, $m_{y3}$), and ($m_{x4}$, $m_{y4}$)) are ($x_{pr1}$,$y_{pr1}$), ($x_{pr2}$,$y_{pr2}$), ($x_{pr3}$,$y_{pr3}$), and ($x_{pr4}$,$y_{pr4}$), an eight-DOF (degree-of-freedom) geometric transform matrix was calculated as follows [14]:

$$M = T \cdot V$$

$$\begin{bmatrix} m_{x1} & m_{x2} & m_{x3} & m_{x4} \\ m_{y1} & m_{y2} & m_{y3} & m_{y4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{pr1} & x_{pr2} & x_{pr3} & x_{pr4} \\ y_{pr1} & y_{pr2} & y_{pr3} & y_{pr4} \\ x_{pr1}y_{pr1} & x_{pr2}y_{pr2} & x_{pr3}y_{pr3} & x_{pr4}y_{pr4} \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{2}$$

To calculate the eight parameters of the transform matrix ($T$), the inverse matrix of $V$ was multiplied to both sides.

Based on the acquired transform matrix ($T$), the position of the gaze on the monitor plane (($G_x$, $G_y$)) is calculated as follows:

$$\begin{bmatrix} G_x \\ G_y \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{pr} \\ y_{pr} \\ x_{pr}y_{pr} \\ 1 \end{bmatrix} \tag{3}$$

## EXPERIMENTAL RESULTS

Our eye tracking system was tested on an CPU with Intel Core i7 (3.6 GHz) with 16 GB of RAM. An eye tracking program was developed on Visual C++ using the Intel OpenCV library. The size of the monitor was (diagonally) 42 in (16: 9), and it had a UHD (Ultra High Definition) resolution of 3840 × 2160.

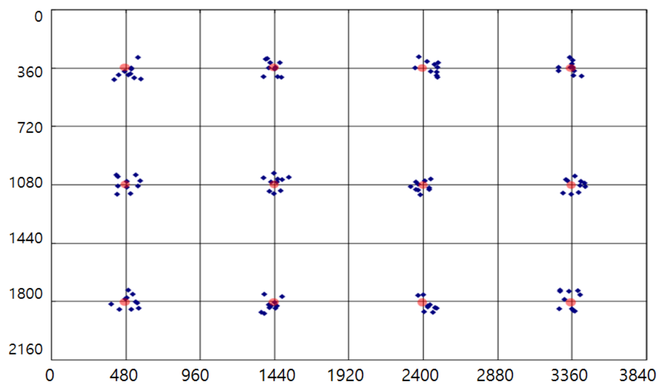



**Figure 4**. Ten examples of the test for measuring gaze estimation accuracy (red and blue dots represents the reference points and the calculated gaze positions, respectively)

To test the accuracy of the proposed eye tracking method, we performed a test involving gazing at 12 reference points. Ten subjects were recruited, and each performed the test five times.

The results showed that the root mean square (RMS) error between the reference points and the gaze points was approximately 94 pixels, which corresponded to the angular error of 0.9 degrees.

## CONCLUSION

In this paper, we proposed an eye tracking method that uses a camera and an infrared light illuminator. The method does not require that the user wear any device on his/her head or face. As the USB-3.0 interface-based 5-MP camera can acquire high-resolution eye images at a distance, gaze was estimated comparatively accurate on a UHD resolution display.

In future work, we intend to compensate for error caused by facial movements of users by tracking other feature points, such as the nose, the mouth, or the corners of the eye. More experiments involving a comparison with previous eye tracking methods will also be performed under the same environmental and systematic conditions to further test the proposed method.

## REFERENCES

[1] Yamato, M., Monden, A., Matsumoto, K., Inoue, K., and Torii, K., 2000, “Quick Button Selection with Eye Gazing for General GUI Environments,” in Proc. of International Conference on Software: Theory and Practice. pp. 712-719.

[2] Lee, E. C., and Park, K. R., 2009, “A Robust Eye Gaze Tracking Method based on Virtual Eyeball Model,” Machine Vision and Applications, 20(5), pp. 319-337.

[3] Azarbayejani, A.,Starner , T., Horowitz, B., and Pentland, A., 1991, “Visually Controlled Graphics,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(6), pp. 602-605.

[4] Gemmell, J., Toyama, K., Zitnick, C., Kang, T., and Seitz, S., 2000, “Gaze Awareness for Video Conferencing: A Software Approach,” IEEE Transactions on Multimedia, 7(4), pp. 26-35.

[5] Hutchinson, T. E., White, K. P., Martin, W. N., Reichert, K. C., and Frey, L. A., 1989, “Human-Computer Interaction Using Eye Gaze Input,” IEEE Transactions on Systems, Man, and Cybernetics, 19(6), pp. 1527-1534.

[6] Matsumoto, Y., Ino, T., and Ogasawara, T., 2001, “Development of Intelligent Wheelchair System with Face and Gaze-based Interface,” In Proc. of IEEE 10th International Workshop on Robot and Human Interactive Communication, pp. 262-267.

[7] Wang, J. G., and Sung, E., 2002, “Study on Eye Gaze Estimation,” IEEE Transactions on Systems, Man and Cybernetics, Part B, 32(3), pp. 332-350.

[8] Shih, S. W., and Liu, J., 2004, “A Novel Approach to 3D Gaze Tracking Using Stereo Cameras,” IEEE Transactions on Systems, Man and Cybernetics, Part B, 34(1), pp. 234-245.

[9] Yoo, D. H., and Chung, M. J., 2004, “Non-intrusive Eye Gaze Estimation without Knowledge of Eye Pose,” In Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition, p. 785.

[10] Lee, E. C., and Park, K. R., “A Study on Eye Gaze Estimation Method based on Cornea Model of Human Eye,” 2007, Lecture Notes in Computer Sciences, 4418, pp. 307-317.

[11] Ko, Y. J., Lee, E. C., and Park, K. R., 2008, “A Robust Gaze Detection Method by Compensating for Facial Movements Based on Corneal Specularities,” Pattern Recognition Letters, 29(10), pp. 1474-1485.

[12] Viola, P., and Jones, M., 2004, “Robust Real-time Face Detection,” International Journal of Computer Vision, 57, pp. 137-154.

[13] Cho, D. H., Park, K. R., and Rhee, D. W., 2005, “A Study on Robust Iris Detection for User Authentication in Mobile Phone,” In Proc. of 4th International Conference on Asian Language Processing and Information Technology, pp. 61-66.

[14] Cho, C. W., Lee, J. W., Lee, E. C., and Park, K. R., 2009, “A Robust Gaze Tracking Method by Using Frontal Viewing and Eye Tracking Cameras,” Optical Engineering, 48(12), p. 127202.