Research on the technologies for Human-computer interaction Based on the Eye-tracking

Jingjing Zhou *
College of Computer Science and Technology, Hubei Business College, Wuhan, 430070, China
281128544@qq.com

Yi Zhu
College of Computer Science and Technology, Hubei Business College, Wuhan, 430070, China
39440102@qq.com

ABSTRACT

Among the new human-computer interaction technologies such as eye tracking, gesture recognition, and EEG tracking, the human-computer interaction technology based on eye movement improves the intelligence, naturalness and efficiency of human-computer interaction. In this paper, a geometric model of a single-camera eye-tracking system under natural light source is established, and the reasonable placement area of the camera when the user's head is stationary and the maximum rotation range allowed when the user's head is free to rotate are discussed. The function of eye-moving mouse is realized.

CCS CONCEPTS

• **Human-centered computing** → Interaction design; Systems and tools for interaction design.

KEYWORDS

Eye tracking, human-computer interaction, Eye tracking interaction, gaze tracking, eye movement recognition

ACM Reference Format:

Jingjing Zhou * and Yi Zhu. 2024. Research on the technologies for Human-computer interaction Based on the Eye-tracking. In *5th International Conference on Computer Information and Big Data Applications (CIBDA 2024), April 26–28, 2024, Wuhan, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3671151.3671164

1 EYE MOVEMENT INTERACTION

At present, the research on eye tracking technology is relatively common, and the research on this technology mainly focuses on exploring the method of eye movement collection and how to analyze the collected movement information. The use of eye sight information as an interactive input to control computer programs has slowly emerged in recent years, and the prospects are very broad. Since the eyes are the main means for people to obtain information, about 80%-90% of the information can be obtained through the visual channel, and the visual channel has characteristics that other channels cannot have: directness, naturalness, and bidirectionality etc [1]. People can freely control the eye movement to look at different objects, so as to obtain external information. This makes the visual channel-based eye-tracking input have the incomparable advantages of other emerging inputs: it is simple, fast and easy to be grasped by the user, and it does not require the movement of human limbs to reduce the physical load of the user. Therefore, eye movement interaction has received more and more attention in the field of human-computer interaction[2].

Common eye movement interactions include the following.

(1) Dwell time trigger

In fixation behavior, the duration of the fixation point is called dwell time. Dwell time triggering means that when the dwell time of the gaze point reaches a certain level, the line of sight can be used instead of traditional input devices such as mouse clicks or keyboard buttons to trigger corresponding execution operations. Dwell time trigger is mostly used to control the graphical interface or locate the mouse, cursor, etc[3]. It is a relatively popular way of eye movement interaction. It can also reflect the user's conscious control intention to better complete the interaction.

(2) Smooth follow movement

Smooth-following motion mostly occurs when there are slow-moving objects or targets in the observation scene, and the line of sight will produce a smooth-following motion state. Smooth following motion is a state of continuous feedback [4]. The eyes capture the signal of the moving target, feed back information such as the speed, direction, and angle of the target movement to the brain, and then control the eyeball to follow the target object for relative movement.

(3) Eye blinking

Eye blinking is a physiological behavior of human beings in a natural state, which is unconscious. Under normal circumstances, the duration of each blink of an ordinary person is about 200-400ms, with an average of 10-20 blinks per minute. When interacting with blinking behavior, it is necessary to recognize conscious blinks, such as blinking frequency exceeding a certain level, or eye closure time exceeding a certain threshold during a blink. Blink triggering is relatively simple, but when the human eye is closed for a long time, since the eye tracker cannot capture the pupil, the gaze point may be lost, which will affect the accuracy of the eye control system to a certain extent.

(4) Gaze gesture

Gaze Gesture is proposed on the basis of saccade, but it is different from saccade in that saccade is often an unconscious shift of sight that occurs when people observe a scene or object, and the start and end of saccade are Unknown, relies on human visual attention.

Gaze gesture is defined as a series of ordered gaze trips, each trip being an intentional gaze movement of two fixed fixation points or

CIBDA 2024, April 26–28, 2024, Wuhan, China

ACM ISBN 979-8-4007-1810-6/24/04
https://doi.org/10.1145/3671151.3671164

regions of fixation. Therefore, eye gestures, as a new way of eye movement interaction, can reflect people's conscious triggering intentions. The itineraries of different paths can define different eye gestures, and different eye gestures can be mapped to different interactive instructions.

2 EYE-MOVEMENT MODEL

With the development of image processing technology and the improvement of the performance of camera equipment and the reduction of cost, most eye tracking systems today use cameras to capture images, and use image processing technology and machine vision software and hardware to achieve. No direct contact with the user is required, allowing the user to be monitored in a completely "natural" state. The current eye tracking system model requires the participation of infrared LED light sources to facilitate detection and feature extraction, but it is required to be performed indoors, and the outdoor environment is unstable to LED light sources. In addition, the initialization process of a multi-camera system is cumbersome, requiring calibration and calibration of all cameras. These have greatly hindered the development of eye-tracking systems towards practicality.

The single-camera system can reduce the system cost and simplify the calibration process of the system. The natural light source does not limit the use environment of the system, and the system can be used indoors and outdoors. Moreover, the defects of relatively low contrast of images collected under natural light sources can be compensated by certain image processing algorithms. By establishing a geometric model of a single-camera eye-tracking system under natural light sources, this paper discusses the reasonable placement area of the camera when the user's head is stationary, and the maximum rotation range allowed when the user's head is free to rotate.

2.1 When user's head is still

When the user's head is still, the camera is located directly in front of the human eye as shown in the figure 1, where point C represents the position of the camera, θ is the horizontal angle of view of the camera. The camera angle of view and the lateral canthal angle of the user's left and right eyes intersect at point L and R respectively, r is the vertical distance between the camera and the center of the user's head, w is the length between the lateral canthal angles of both eyes. The shortest distance of r is r_{min}.

$$r_{min} = \frac{w}{2\tan\left(\frac{\theta}{2}\right)} \tag{1}$$

The average length of an adult's single eye is 30mm, and the interpupillary distance between the two eyes of an adult female is 64mm. From the figure, it can be deduced that the length between the lateral canthal angles of both eyes w = 30mm + 64mm=94mm. The horizontal field of view of the camera selected in this article is $\theta = 20^0$. Substitute the data into Equation 1 to get $r_{min} \approx 267$mm.

In practical applications, when the camera is placed directly in front of the eyes, it will block the line of sight, which will affect the reliability and accuracy of the eye tracking system application. The camera should be placed on the side of the person, and the camera should be moved from point C in front of the human eye to the left or to the right, but it cannot move infinitely. When the deflection angle of the camera is greater than the threshold, the camera cannot capture complete and clear binocular image.

Since the left and right horizontal movements of the camera are symmetrical, take the camera moving to the right as an example to discuss the maximum lateral movement angle of the camera and its shortest distance at each lateral position. In the geometric model shown in the figure 2, point C is the camera, θ is the horizontal angle of view of the camera, C' is the new position of the camera after moving to the right, β is the horizontal deflection angle of the camera, the angle γ and the angle δ are formed between the two edge lines of the horizontal viewing angle of the camera and the lateral canthal angles of the left and right eyes.

Deduced from the geometric model, when the camera moves from the front of the eyes to the right, the minimum distance between it and the midpoint O of the line connecting the eyes is r, camera horizontal deflection angle is β.

$$r = \sqrt{\frac{W^2}{4}\left[\left(4\cot^2\theta + 1\right)\cos^2\varphi \pm 4\cot\theta\sin\varphi\cos\varphi + \sin^2\varphi\right]} \tag{2}$$

$$\beta = \sin^{-1}\left[\frac{w}{2}\frac{\cos\varphi}{r}\right] \mp \varphi \tag{3}$$

At r, the camera's position coordinate is $(r\sin\beta, r\cos\beta)$. According to the value of r, β, the motion trajectory of the camera can be determined.

2.2 When the user's head turns freely

In practical applications, the user's head should be allowed to rotate freely within a certain range to improve the user's comfort experience and allow them to obtain a wider field of view. However, if the user's head is turned too far, the camera may not be able to capture a complete image of both eyes, so that the eye movement gesture cannot be recognized. So, we need to discuss the maximum rotation range of the user's head. The rotation diagram is shown in Figure 3.

O is the midpoint of the line connecting the eyes, and point C is the camera position, the camera angle of view and the lateral canthal angle of the user's left and right eyes intersect at point L and R respectively, $L'R'$ is the position after the eyes are rotated horizontally to the right angle α with the head synchronously, $M'N'$ is the projection of $R'N$ in the horizontal direction. d is the length of the human eye. After the head is turned, in order for the camera to still capture a clear eye image, the following conditions should be met:

$$d\cos\alpha \geq \frac{d}{2} \tag{4}$$

The value range of the horizontal rotation angle of the head to one side is $[0^0, 60^0]$.

3 ALGORITHM OF EYE MOVEMENT RECOGNITION

At present, the commonly used eye movement recognition algorithms mainly include: Velocity Threshold Recognition, Hidden Markov Model Recognition [5]. Dispersion Threshold Recognition and Minimum Spanning Tree Recognition, etc, these algorithms can more accurately identify the gaze behavior and saccade behavior of people in the natural state.

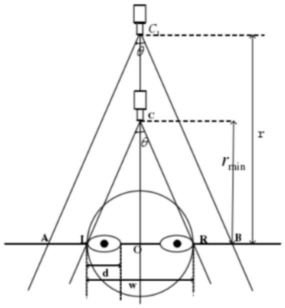

Figure 1: The camera is located directly in front of the human eye.

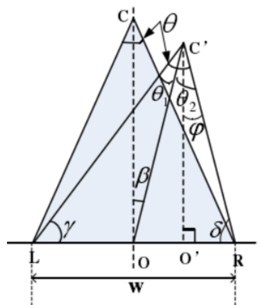

Figure 2: The geometric model

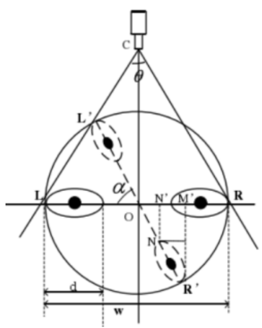

Figure 3: The rotation diagram.

Velocity threshold identification is a relatively easy to understand and implement identification method in identification algorithms [7]. I-VT is a velocity-based method that separates fixation and saccade points based on point-to-point velocity. There are basically two velocity distributions of eye movements: low-velocity gaze movements and high-velocity saccade movements.

Hidden Markov Model recognition uses probabilistic analysis to determine the most likely eye movement behavior. The I-HMM uses a two-state HMM, where the states represent the velocity distributions of saccade and fixation, respectively. The I-HMM method is more reliable than the fixed threshold method of I-VT.

In contrast to the velocity-based recognition of I-VT and IHMM, the principle of discrete threshold recognition is to exploit the fact that fixation points tend to be clustered closely together due to

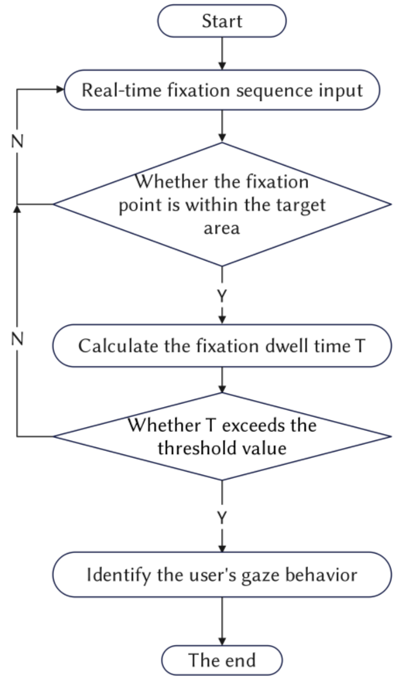

Figure 4: The flow chart of the real-time recognition algorithm of gaze behavior.

their low velocities. Since fixation behaviors typically have dwell times of at least 100ms, discrete threshold identification techniques typically include a minimum dwell time threshold of 100-200ms.

The I-DT algorithm uses a moving window spanning consecutive data points to examine underlying gaze behavior. The moving window initially spans a minimum number of points, determined by a given dwell time threshold and sampling frequency. I-DT then checks the dispersion of the points in the middle of the window by summing the difference between the maximum and minimum x and y values of the points, if the dispersion is higher than the dispersion threshold, the window does not indicate fixation behavior and the window is to the right Move a point. If the dispersion is below the dispersion threshold, the window represents fixation behavior.

Accurately identifying the user's eye movement behavior is the most important link in the eye movement human-computer interaction process, and it is also the premise of correctly completing the interactive instructions.

The flow chart of the real-time recognition algorithm of gaze behavior is shown in Figure 4.

The flow chart of the real-time recognition algorithm of eye gesture behavior is shown in Figure 5.

The flow chart of the real-time recognition algorithm of conscious blinking behavior is shown in Figure 6.

4 IMPLEMENTATION OF THE FUNCTION FOR EYE MOVEMENT MOUSE

Eye movement simulation of mouse and keyboard are two basic studies of single eye movement interaction [8]. When the mouse and keyboard cannot be used in some special scenarios, for example,

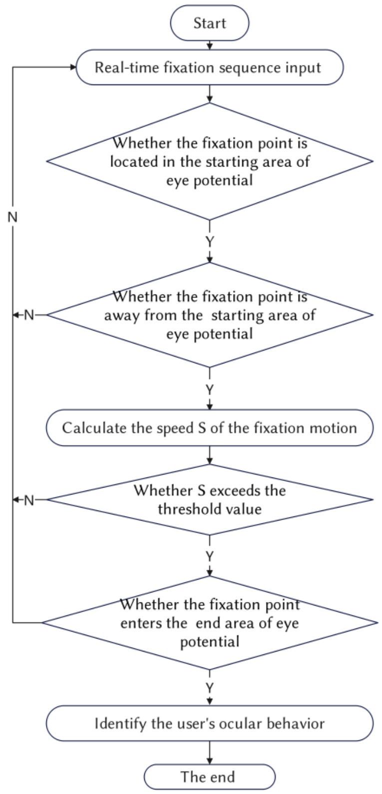

Figure 5: The flow chart of the real-time recognition algorithm of eye gesture behavior.

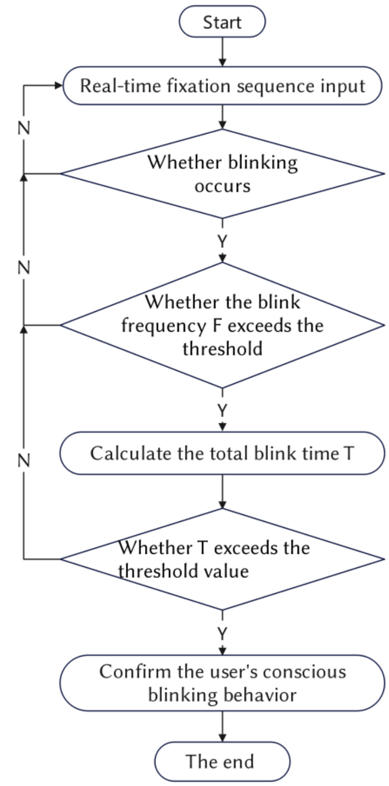

Figure 6: The flow chart of the real-time recognition algorithm of conscious blinking behavior.

when the user's hands are occupied, or the user has a physical disability, the mouse cannot be used, keyboard and other devices for input, the eye control interface is a new interactive choice, which can increase the naturalness and intelligence of human-computer interaction [6].

Using eye posture and gaze tracking, it can replace the cursor movement, scrolling, clicking, double-clicking, dragging and other functions of the ordinary mouse [9]. For users who can blink, the equivalent of a traditional mouse can be accomplished, including moving the pointer, left and right clicks, double clicks, and dragging. For users who cannot blink but can control blinking, move the pointer, left-click. The function description of the eye-tracking mouse is shown in Table 1.

Since the user will also produce certain eye gesture behaviors in the process of observing the mouse, in order to distinguish the user's eye movement manipulation from the unconscious observation, that is, to avoid the "Midas touch" problem, this paper first sets the conscious blink to control the eye. control operation on and off.

The frequency of involuntary blinks is about 10-20 times per minute, and the closing time of each blink is about 200-400ms. Therefore, conscious blinking can be identified by setting the blink frequency threshold F_d and the blink duration threshold T_d, When the blink frequency $F>F_d$ and the total duration of continuous blinking $T>T_d$, it is judged as a conscious blinking behavior, and the eye control operation is turned on. The blink frequency threshold set in this paper is F_d=30 times/min, Blink duration threshold T_d=800ms.

Select 10 people to participate in the test and complete the control of five basic commands of left-click, right-click, left-click, drag, and scroll, and each command is completed 30 times. The test data is shown in the Table 2.

Table 1: Features of Eye Tracking Mouse.

Function	Behavior of the eye movement
left clicking	Single blink of left eye, closing time for more than 0.5s
right clicking	Single blink of the right eye, closing the eye for more than 0.5s
Double left clicking	Single blink of left eye, closing time for more than 1s
dragging	Left blink for more than 1.5s, the mouse is pressed, and left blink again for more than 1.5s, the mouse is released
scrolling	Up eye movement triggers scroll up, blink down eye movement triggers scroll down

Table 2: Eye tracking mouse test results.

user ID	accuracy of Left-clicking (%)	accuracy of Right clicking (%)	accuracy of double Left clicking (%)	accuracy of Dragging (%)	accuracy of scrolling(%)
1	86.7	90.0	80.0	83.3	73.3
2	83.3	80.0	73.3	93.3	86.7
3	80.0	93.3	86.7	83.3	80.0
4	86.7	83.3	83.3	80.0	86.7
5	93.3	90.0	86.7	83.3	73.3
6	90.0	86.7	83.3	93.3	80.0
7	80.0	90.0	93.3	73.3	90.0
8	83.3	73.3	86.7	80.0	86.7
9	93.3	83.3	73.3	86.7	93.3
10	73.3	93.3	83.3	80.0	83.3

5 CONCLUSION

This paper expounds the technical principles of eye control systems, and analyzes the implementation methods of eye movement related technologies in detail. This paper analyzes and discusses the problems existing in gaze interaction based on gaze input and eye gesture interaction based on saccade input. In the research on the input characteristics of eye gestures, firstly, the basic eye gesture interaction time threshold is studied, and secondly, the saccade speed is proposed as an index to divide the priority of different monocular gestures, and the priority order between monocular gestures is obtained as follows: The multi-eye gesture action design provides an objective theoretical basis. Five specific functions: left-click, right-click, left-double-click, drag, and scroll are verified by experiments, and the correct rate of eye movement test is obtained.

REFERENCES

[1] Pengbo Su; Kaifeng Liu. Effects of anthropomorphic design on comprehension of self-monitoring test results: Integrating evidence of eye-tracking and event-related potential. Journal [J] Displays. Volume 82, Issue. 2024.

[2] Sandra Caloca Amber; Elba Mauriz; Ana M. Vázquez Casares. Exploring eye-tracking data as an indicator of situational awareness in nursing students during a cardiorespiratory arrest simulation.Nurse Education in Practice, 2024.

[3] Klaib A F, Alsrehin N O, Melhem W Y, *et al.* Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies. Expert Systems with Applications, 2021, 166:114037.

[4] Liu Hui-Jia, Chi Jian-Nan and Yin Yi-Xin. A Review of Feature-based Gaze Tracking Methods. Acta Automatica Sinica, 2021, 47(2): 252-277.

[5] Zhang R, Walshe C, Liu Z, *et al.* Atari-head: Atari human eye-tracking and demonstration dataset. In: Proceedings of the AAAI conference on artificial intelligence. New York, USA: AAAI, 2020. 34(4): 6811 -6820.

[6] Brousseau B, Rose J, Eizenman M. Hybrid eye-tracking on a smartphone with cnn feature extraction and an infrared 3d model. Sensors,2020, 20(2): 543.

[7] Hsu W Y, Chung C J. A novel eye center localization method for head poses with large rotations. IEEE Transactions on Image Processing, 2020, 30: 1369-1381.

[8] Yiu Y H, Aboulatta M, Raiser T, *et al.* DeepVOG: Opensource pupil segmentation and gaze estimation in neuroscience using deep learning. Journal of neuroscience methods, 2019, 324: 108307.

[9] Martinikorena I, Cabeza R, Villanueva A, Urtasun I, Larumbe. Fast and robust ellipse detection algorithm for head-mounted eye tracking systems. Machine Vision and Applications, 2018, 29(5): 845-860.