Research and Application of Eye Movement Interaction based on Eye Movement Recognition

Shu Wang^{1,a}, Qing Wang¹ and Hong Chen¹

¹China Agricultural University, 100083 No. 17 Qinghua East Road, China

Abstract. Generally, human-computer interaction is an interaction and operation between users and machine hardware. The user submits instructions to the machine, and the machine outputs the processed data and results to the user after receiving the instructions from the user. Mouse, keyboard, etc. are common input channels. With the maturity of eye tracking technology and the development of equipment miniaturization, turning eye movements into human-computer interaction input channels has become a hot spot in the field of human-computer interaction. Therefore, this paper analysed the physiological characteristics of eye movement, proposed the design principles and framework of eye movement interaction, and designed three kinds of eye movement recognition algorithms of fixation, saccade and blink. On this basis, using Unity 3D cross-platform development engine as a development tool, a children's attention training game application based on eye movement interaction is designed. The game is designed to combine eye movement interaction technology with attention training mode, simplify the control mode of the game, get attention feedback at the first time, achieve better training effect and improve the efficiency of human-computer interaction.

1 Introduction

The keyboard and mouse have always been traditional tools for manipulating computers, but with the development of computer technology, users need to communicate with new methods to improve work efficiency. For example, touch screen technology makes electronic products more friendly. The appearance of somatosensory technology makes it possible for limbs to participate in interaction. The addition of speech recognition technology makes interaction a step further in convenience. It can be seen from this that human-computer interaction technology is moving towards the goal of human oriented, natural and efficient development. Among all kinds of human-computer interaction, eye-tracking technology is a more natural new type of human-computer interaction using eye movement. Eyes are an important organ for human beings to obtain information, and about 80 % of external information is obtained through eyes. In the cognitive process, vision and hearing are the starting points for understanding the world. Human beings acquire external information through their ears and eyes, and form motion instructions through cognitive processors, thus guiding other forms of interaction, such as clicking the mouse and touching the screen. At the same time, the fine muscles of eye movement have better immune characteristics to fatigue, and can still ensure accuracy and operability in long-term interactive experience. In addition, eye movement interaction is based on non-interference control and will not have any negative impact on the people around you^[1]. Based on the above factors, more and more researchers

use eye movement as an input channel for human-computer interaction, which can complete the operation of the computer through eye movement and better assist other forms of input, ensuring the fluency and sustainability of input without increasing the user's operation burden.

Eye movement interaction depends on the eye-tracking technology. Early eye-tracking methods mainly include observation and mechanical recording. In the past 100 years, there have been technical methods to objectively observe and record eye movement with hardware devices^[2]. Among all kinds of methods, the optical recording method has the longest research history and the most research achievement is Feng Fu, which is the most important application type in modern eye tracker^[3]. In the optical recording method, the most popular method is to calculate the fixation vector^[4] by Ohno's method, that is, using the principle that the reflectivity of invisible infrared rays after passing through the cornea and pupil is different, it is detected by an infrared detector placed at the front of the eye at a certain distance, and the pupil coordinates and the centre position of corneal curvature are located by image processing technology to determine the position vector connecting the two points, thereby calculating the vector coordinates of the gaze point.

Domestic and foreign scholars' interest in eye movement interaction research mainly focuses on the optimization of Midas contact problem, the expansion of scene application, and the design and implementation of eye movement interaction system. However, the design of eye movement interaction and the definition and

^a Corresponding author: 460594932@qq.com

recognition of eye movement have not yet formed a unified standard. Therefore, this paper analysed the physiological characteristics of eye movement, proposed the design principles and framework of eye movement interaction, and designed three kinds of eye movement recognition algorithms: fixation, saccade and blink. Using Unity 3D cross-platform development engine as a development tool, a children's attention training game application based on eye movement interaction is designed and implemented. Simplify the game control mode and strengthen the user's attention training. The test results showed that this method has good attention training effect and human-computer interaction efficiency.

2 Physiological analysis of eye movement

As one of the most important sensory organs of human body, eyes are very important for human to obtain information. According to whether the two eyes rotate in the same direction, the eye movement can be generally divided into two types: conjugate eye movement and unconjugate eye movement [5]. The classification of eye movements is shown in Figure 1.

Figure 1. Eye movement classification diagram

Considering the characteristics of human-computer interaction, we often use fixation and saccade in conjugate eye movements and blink eye movement to identify and interact.

Fixation: The human retina has one of the most sensitive special areas, the macular area, with a diameter of about 3 - 5 mm and a small depression, or fovea, in the centre of the macular area. The process of aligning the fovea of the eye with the target stimulus for more than 100ms can be called fixation. However, in the human-computer interaction, in order to distinguish the conscious and unconscious fixation process of the user, i.e. Midas contact problem, the action with fixation time longer than 300ms is often defined as fixation[6].

Saccade: Saccade refers to the rapid jump between gaze points. The eyes do not always stay at one point of interest, but jump between several regions of interest in a scene, and then these discontinuous points of interest recombine in the brain to form a picture. Some scholars have verified through experiments that the maximum speed of saccade can reach 1000°per second. The starting time of a general saccade is at least 200ms, and the duration is determined by the intensity of the task to be completed, generally 30 to 120 mm or more [7].

Blink: Blink is a group of activities that the upper and lower eyelids close and open quickly, and it is a normal physiological activity of human beings. The common blinking consists of two ways, one is called protective blinking. The other is the usual blink of an eye, which aims to distribute tears evenly over the cornea and conjunctiva to ensure that they do not dry and to give temporary rest to the retina and eye muscles. According to relevant data, normal people blink more than 10 times per minute, each blink takes 0.3-0.4 seconds, and every second is separated by about 2.8-4.0 seconds, showing individual differences.

3 Design Principles of Eye Movement Interaction

Based on the hardware of eye movement equipment and adhering to the concept of user-oriented design, this paper puts forward the principles to be followed in eye movement interaction design, including the following four aspects:

Goal-driven design: With the development of eye-movement interaction technology, its interaction object is no longer only for the disabled. The change of demand and the maturity of technology lead to eye movement interaction as a daily interaction channel to serve normal people. Therefore, before designing the eye movement interaction system, people with different cognitive abilities should be divided to realize the goal-driven design.

Multi-channel constraint: Midas's contact problem is unavoidable in the eye movement interaction process, and the computer cannot accurately distinguish "intentional" eye movement from unintentional eye movement. In addition to using the traditional methods of time delay, spatial proximity and widget - based interaction context, reducing Midas effect through multi-channel combination of keyboard, voice and body feeling has greatly improved the interaction efficiency.

Define eye movement interaction set: The eye movements defined for interaction should have a higher degree of discrimination, reduce ambiguity and conform to the user's usage habits, thus reducing cognitive burden and learning time.

Visual feedback: Visual feedback can judge whether the system recognizes the user's current eye movements in real time, effectively improve the accuracy of interaction and help relieve visual fatigue. For example, when watching a button, the current watching action is recognized through the state change of the button.

4 Eye movement interaction framework

According to the above interaction principles, an application framework based on eye movement interaction (FEBI) is proposed, as shown in figure 2.

Figure 2. Eye movement interaction framework

This application framework is from the bottom to the up, which includes the system layer, the middle layer and the application layer. At the bottom, the system layer mainly provides the camera and infrared devices by hardware and the corresponding parameters are set. The middle layer is mainly composed of calibration module of eye location module, eye action recognition module and other interactive event encapsulation modules. Firstly, the user's eye space coordinate system is established in the line of sight positioning module, and the user's gazing direction $p(x, y, z)$ in the space coordinate system will be obtained in real time. Secondly, in the calibration module, it will be converted to the $g(x, y, t_i)$ in the screen scene coordinate system according to the mapping relationship, in which X and y are the user's fixation points on the screen scene. t_i represents the time point of the system. Thirdly, in the eye movement recognition module, the user eye interactive action set is defined by combining with the corresponding eye movement input mechanism, so as to determine whether the current existence and what type of eye movements exist according to a certain context. Finally, in the interactive event encapsulation module, eye movement assistance event mapping is generated and encapsulated by combining other channel input commands such as keyboard and gesture. In the application layer, the real-time eye movement coordinate and the event message can be called is encapsulated, so as to replace the mouse movement and activate the corresponding interactive function. Therefore, the interactive operation and control of the application can be realized.

5 Eye movement recognition

According to the types of eye movements mentioned in this article and the design principles of eye movement interaction, three types of eye movements will be defined, including fixation, saccade and blink.

Table 1. Types of eye movement interaction.

Movement	Describe
Blink	Blink deliberately
Fixation	Gaze the target area for a while
Saccade	Jump within the target areas

5.1 Blink recognition

According to statistics, normal people blink more than 10 times per minute, taking 0.3-0.4 seconds for each blink, with 2.8 - 4 seconds between each blink. Therefore, the choice can be realized by a binocular blink with deliberate long pauses (as opposed to a normal blink).

The legal values of both eyes can be obtained by the eye tracker, thereby constructing the legal sequence of time t :

$$\langle l_1, l_2, t \rangle \quad (1)$$

l_1 represents the left eye, l_2 represents the right eye, the value is true when eyes are open, and the value is false when eyes are close.

The conditional formula (2) is used to determine whether there is a legal sequence transformation (3).

$$\forall t_j - t_i \in \varepsilon_T \quad (2)$$

Where ε_T is a custom time domain (which should be greater than the normal blink time to distinguish the normal blink time).

$$\langle 1, 1, t_i \rangle - \langle 0, 0, t_k \rangle - \langle 1, 1, t_j \rangle \quad (3)$$

If it exists, it means that a binocular blink occurred in the current time and then triggered the corresponding instruction.

5.2 Fixation recognition

The speed threshold recognition method (I-VT)^[8] is used to obtain the basic attributes of eye movement data and calculate the average speed of the current gaze point. Comparing and judging the calculated speed with the set threshold, if it is lower than the threshold, it will be classified as fixation, if it is higher than the threshold, it will be classified as saccade. When the gaze coordinate has been in a certain area for a certain period of time, it is judged as gaze. The specific method is as follows:

$\{S_n\}$ is the n target interest areas that predefined by the system. If the condition (4) and (5) is satisfied in $\langle t_i, t_j \rangle$, there is a gaze of ε_T in the region of interest S_m .

$$\exists t_j - t_i = \varepsilon_T \quad (4)$$

$$gazePoint_{t_k} \in S_m, k \in [i, j], m \leq n \quad (5)$$

Where $\langle t_i, t_j \rangle$ represents the time series from t_i to t_j . $gazePoint_{t_k}$ indicates the gaze point at the time of t_k . ε_T is the threshold value of the custom gaze time.

5.3 Saccade recognition

In order to have good human-computer interface interaction, we have defined five target areas, which are located at the four vertices and the center of the screen. As shown in figure. 3.

Figure 3. 5-Target areas

Saccade is a series of gaze judgments on the target area. Taking a two-point saccade as an example, the recognition process starts when the user watches any target area for more than a defined threshold (400ms). During $<t_i, t_j>$ period, t_i time gaze point coordinate is $P_i(x_i, y_i, t_i)$ and t_j time viewpoint coordinate is $P_j(x_j, y_j, t_j)$. If S_1 and S_2 are two target areas predefined by the system, S_1 and S_2 are expressed as:

$$S_1 = \begin{cases} x_1 \in [x_{\min 1}, x_{\max 1}] \\ y_1 \in [y_{\min 1}, y_{\max 1}] \end{cases} \quad (6)$$

$$S_2 = \begin{cases} x_2 \in [x_{\min 2}, x_{\max 2}] \\ y_2 \in [y_{\min 2}, y_{\max 2}] \end{cases} \quad (7)$$

If the conditions (8) and (9) are met, it means that the gaze point has jumped from the target area S_1 to the target area S_2 within the predefined time threshold, and the system successfully recognized the saccade action and triggered the corresponding instruction.

$$\exists t_j - t_i \leq \mathcal{E}_T \quad (8)$$

$$\begin{cases} x_i \in [x_{\min 1}, x_{\max 1}] \\ y_i \in [y_{\min 1}, y_{\max 1}] \end{cases} \wedge \begin{cases} x_j \in [x_{\min 2}, x_{\max 2}] \\ y_j \in [y_{\min 2}, y_{\max 2}] \end{cases} \quad (9)$$

where $\mathcal{E}_T$ is a predefined saccade time threshold.

6 Application and Verification

6.1 Application of children's attention training game based on eye movement interaction

Based on FBI framework and Unity 3D cross-platform development engine as development tool, this paper designed and implemented a children's attention training game application based on eye movement interaction. This application trains children's attention through the Shulte square training method. The game chose a design that combines eye movement interaction technology with attention training game, simplified the control mode of the game, obtained attention feedback at the first time, and can achieve better training effect and operability.

After starting the game, the user will prompt through the interface, move the line of sight and complete the corresponding operation. The game interface is shown in Figure 4-5. The pink squares located in the four corners of the screen are the starting target area of the saccade action command.

Figure 4. User-interface

When the line of sight stays at the number 9, a visual feedback effect appears. The square where the number 9 is located slowly becomes larger, and the color changes from white to light red to show that the current eye movement is recognized, and subsequent operations can be carried out.

Figure 5. Visual feedback

The eye movement interaction instructions involved in this system are shown in Table 2. The user can lock a certain option through the gaze action and confirm the locking option with two consecutive blinks. You can pause, exit or restart the game at any time during the game.

Table 2. Eye movement instruction list

Type	Eye interaction movement	Instruction
Fixation	Fixation	Lock or release
Blink	Blink deliberately	Confirm
Saccade	Gaze from the top of right corner to the centre area	Exit
Saccade	Gaze from the top of left corner to the centre area	Return to the main menu
Saccade	Gaze from the bottom of left corner to the centre area	View menu
Saccade	Gaze from the bottom of right corner to the centre area	Set options

Eye movements based on saccade are illustrated in figure. 6. Under Windows, the exit option is in the upper right corner of the interface, the start option is in the lower left corner, and the setting option is in the lower right corner. Therefore, the eye movement interaction designed by this application is more in line with people's habits and can reduce learning time and cognitive burden.

Figure 6. (a) Exit-eye movement; (b) Return to the main menu-eye movement; (c) View menu-eye movement; (d) Set options-eye movement.

6.2 Experimental Verification

We tested the application from two aspects. First, evaluate the interactive usability of this game application. The second aspect is to evaluate the efficiency of attention. We selected 12 preschool children aged 4 to 6 as subjects. In the interactive usability evaluation experiment, 6 children were randomly selected from 12 children to play independently without the guidance of relevant personnel. The results showed that only one child sought the help of relevant personnel while the rest of the children could play the game independently until the end. At the end of the experiment, all the children were asked how they felt about this interaction pattern. No one felt dizzy. The experimental results show that the interactive mode has good interactive usability. In terms of attention enhancement, 12 children were randomly divided into two groups for training. The first group used the way of eye movement interactive operation, the second group used traditional interactive method as the control group, in which the training period was 7 days and the training frequency was three times a day. The completion time T (s) and T'(s) of the Schulte form before and after the training period were recorded, and the results were tested by T-test. At the same time, the results were shown in the table.

Table 3. Test record(s)

	C-G			E-G		
	T	T'	ΔT	T	T'	ΔT
1	42	38	4	38	29	9
2	40	36	4	37	28	9
3	60	55	5	35	26	9
4	36	30	6	30	24	6
5	46	40	6	38	25	13
6	53	48	5	44	27	17

Table 4. P value of T test

	T	ΔT
P	0.048714357	0.007018846

In the table, the P value in T-test for the completion time of the Schulte form in the two groups before training was 0.048714357, which was less than 0.05. This means that there is a significant difference in the use of different interaction methods for training under the condition that there is no training, and the completion time of eye movement interaction is generally shorter than that of the traditional control group. The P value in T-test of the difference between the completion time of Schulte table before and after training in the two groups was 0.007018846, which was less than 0.05, which showed that there was significant difference between the two groups after training, and the experimental group with eye movement interaction had better effect on attention enhancement.

7 Conclusions and discussion

Based on the analysis of eye movement characteristics, this paper put forward the design principles and interaction framework of eye movement interaction application, and used Unity 3D cross-platform development engine as development tool to design and

implement a child attention training game application based on eye movement interaction. Experiments showed that the attention training game application using eye movement interaction has good interactive usability. The author believes that future eye movement interaction will be studied in depth in the following aspects:

1. With the popularization of the Internet of Things, the number of electronic devices we need to interact with will also increase day by day. As a new man-machine interface technology, eye movement interaction in the future will be applied to various electronic products to realize intelligent life.

2. In order to reduce Midas' contact effect, eye movement interaction norms should be defined, not only to adapt to people's inherent interaction habits, but also to have a unified operation paradigm in all kinds of platforms and applications.

3. The popularization of interaction will inevitably lead to confusion in development. A unified eye movement standard library should be defined and middleware should be developed.

References

1. Shu Wang, Qing Wang, Hong Chen. Design and verification of children-oriented educational application based on eye movement interactive technology. *Computer and Modernization*, 62-67(2018)
2. Juan Xu. Development and performance comparison about eye trackers. *China Modern Educational Equipment*, 16-18(2012).
3. Yuchang Han. The Developmental History of the Eye Movement Apparatus and the Experimental Method of Eye Movement. *Journal of Psychological Science*, 454-457(2000).
4. Ohno T, Mukawa N, Yoshikawa A. FreeGaze: a gaze tracking system for everyday gaze interaction. *Eye Tracking Research & Application Symposium*, 125-132(2002).
5. Carpenter R H S. Movements of the Eyes. *Quarterly Review of Biology*, 62(1988).
6. Wei Liu, Xiugan Yuan. Human vision - eye movement system research. *Chinese Journal of Ergonomics*, 41-44(2000).
7. Purves D, Augustine G J, Fitzpatrick D. *Neuroscience*. 2nd edition. National Center for Biotechnology Information's Bookshelf, 2001.
8. Salvucci D D. Identifying fixations and saccades in eye-tracking protocols. *Eye Tracking Research and Applications Symposium*. 71-78 (2000).