ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0013

Research on Virtual Simulation based on Eye Movement Tracking

Shaopeng Li[1,2,*], Yuming Qi[1,2], Qiangqiang Li[1,2], Ming Zhang[1,2], Pengbo Feng[1,2]

[1] Institute of Robotics and Intelligent Equipment, Tianjin Vocational and Technical Normal University, Tianjin, 300222, China

[2] Tianjin Key Laboratory of Intelligent Robot Technology and Application Enterprise, Tianjin, 300350, China

*wy875987670@163.com

Abstract

In view of the current eye tracking technology has a direct, can replace traditional human-computer interaction operation sensitivity and natural features, this article to control science, cognitive science, computer science, behavioral science theory as the instruction, constructs the eye movement data model and eye movement behavior model, designed the eye movement interaction mechanism and real-time recognition algorithm of eye movement behavior, Virtual keyboard eye control system and unmanned platform eye control system are realized. Firstly, in order to reflect the temporal and spatial characteristics and variation trend of eye movement information, a spatio-temporal cube model based on eye movement data was constructed. Then, region-based eye-movement interaction mechanism was designed, and three eye-movement behavior models of fixation, eye potential and conscious blinking were constructed based on this, and real-time recognition algorithms were designed respectively for these three kinds of eye-movement behavior. Then, according to the principle of eye movement interaction, a human-computer interaction system design framework based on eye movement is proposed, and the interaction instructions are formally represented.

Keywords

Eye Tracking; Human-computer Interaction; The Eye Control Systerm; Interactive Mechanism.

1. Introduction

With the rapid development of computer, Internet, artificial intelligence, information technology and the rapid popularization of some intelligent devices, human-computer interaction technology is widely used in military, industry, education, medical treatment, home and other application fields[1]. Now, the human-computer interaction mode is no longer the traditional keyboard and mouse interactionn[2] , but has been further developed into a multi-channel interaction mode integrating new interaction modes such as eye movement, gesture, speech and EEG, thus improving the intelligence and efficiency of human-computer interaction[3]. In recent years, with the emergence of gesture recognition equipment eye tracker, EEG tracking device[4], VR headset and other intelligent interactive products, interaction methods have been greatly improved and enhanced[5].

People can obtain external information through many ways, among which vision is one of the most efficient ways for people to obtain external information and perceive external environment. About 80% of such external information is captured by eyes [6]. In addition, due to the more efficient and agile interaction characteristics of eye-captured information[7], relevant researchers have strong

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0013

research interest in it and hope to develop eye-movement interaction into a new interaction channel. Therefore, eye tracking plays an irreplaceable role in the field of interaction[8].

In the current human-computer interaction system on the market, visual interaction is often used to receive the information output of the computer[9]. However, there are often important emotional states of users and hidden values of interaction hidden in people's eyes[10]. Therefore, it is of great scientific research value to focus on the in-depth study of eye movement interaction mechanism. The earliest eye tracking technology focused on eye movement behavior during reading, visual search and other aspects[11]. It is devoted to the research of collecting user eye movement data and analysis, so as to explore the connection between human eye movement behavior and interest preference[12]. In recent years, more and more scholars have taken eye movements as a new interaction mode to study how to make interactive objects deeply understand human interaction behaviors and then control computers through eye movements[13].

2. Identification of Eye Movement Data

2.1 Identification of Dispersion Threshold (I-DT)

Different from traditional data analysis, eye movement interaction process involves time dimension. Eye movement viewpoint represents (xi,yi,zi). In this process, ti represents the sampling time of the ith viewpoint, and (xi,yi,) is the coordinate under the ith viewpoint. Figure 1.2 shows the eye movement trajectory in a specific sampling time, which is a series of ordered motion data sets.

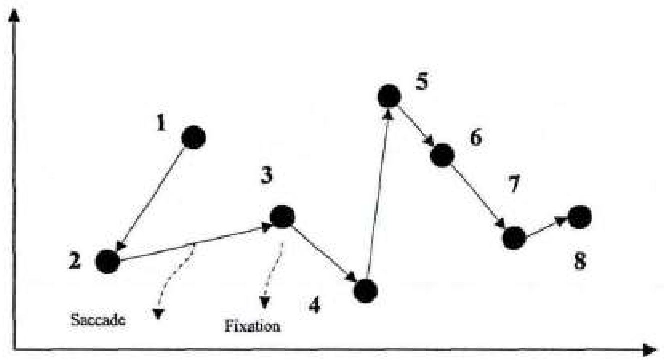

Fig. 1 Time space trajectory diagram.

For discretization, the common algorithms are discretization threshold (I-DT) and minimum spanning tree identification (I-MST). For velocity identification, the common algorithms are velocity threshold identification (I-VT), hidden Markov model identification (I-HMM), Kalman filter identification (I-KF).

In this algorithm, it is necessary to first set the minimum sustained fixation time tmin here (usually 100-300ms), and then calculate the distance D of different viewpoints in the data window from the discretization threshold, that is, the discretization degree. If D is less than Dt, then the identified viewpoint is the fixation point. Otherwise, it is the jumping point, and then input the viewpoint for judgment. The calculation of dispersion can be summarized as follows:

① Distance threshold (DT): calculate the distance between adjacent viewpoints, and compare the set threshold to judge fixation.

② Distance dispersion measure (DD): calculate the distance between the two viewpoints, and compare the gaze with the threshold set beforehand.

③ Fixation radius (Rd): the first step to calculate the center of the window line of sight: then calculate the distance between the view point and the center point, compared with the radius threshold, the fixation point and the beating point are judged by the calculation results.

④ Salvucci dispersion (Slv): By calculating the maximum vertical distance and maximum horizontal distance D between viewpoints, compared with the previous radius threshold, the fixation point and jumping point are judged by the final value.

$$D=[(MaxX-MinX)+(MaxX-MinY)]/2 \quad (1)$$

⑤ Standard deviation (SD): calculate the center of the viewpoint, calculate the standard deviation of the distance between the viewpoint and the center of the viewpoint, and compare with the threshold to determine the fixation point and the jumping point.

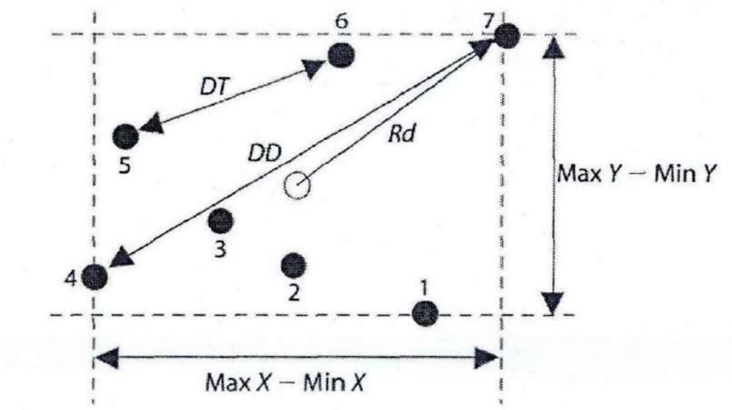

Fig. 2 Calculation method of dispersion

2.2 Speed Threshold Identification Method (I-VT)

Speed threshold recognition (I-VT) is a gaze classification algorithm for speed, which is relatively simple. The usual idea of I-VT is to classify the signal speed, and distinguish the fixation and beating in the data by the speed of human eye movement. The simplest way to calculate speed is to calculate the distance between two viewpoints and divide it by time:

$$V_{t1t2} = \frac{S_{t1}-S_{t2}}{t_1-t_2} \quad (2)$$

The algorithm is the following process:

Step 1: The first step is to calculate the viewpoint speed.

Step 2: Secondly, compared with the speed threshold set by the system. If the value is higher than the speed threshold, it means the speed is fast and is classified as jumping. If the value is lower than the speed threshold, it is regarded as gaze behavior.

Assume that point coordinates of a computer for si position (xi, yi), then the point behind the eye position is si + 1 (, yi xi + 1 + 1), by using Euclidean distance to represent the two fixation point spacing, namely Di = | | (xi, yi), (, yi xi + 1 + 1) | |, d refers to the computer screen and the subjects' spacing, Calculate the Angle of sight change between two points:

$$\theta_i = 2\arctan(\frac{D_i}{2d}) \quad (3)$$

Angular velocity of view:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0013

$$\theta = \theta_i / \Delta t \tag{4}$$

Compared with the angular velocity threshold, if the Angle of sight change is smaller than the angular velocity threshold, the fixation point can be identified. Otherwise it's a beat.

2.3 Hidden Markov Model Identification (I-HMM)

The hidden Markov model identification algorithm refers to the identification algorithm of velocity, which is more complicated than I-VT identification algorithm. In the third stage of I-HMM, Baum-Welch is used to reevaluate the algorithm. This algorithm reevaluates the initial parameters and minimizes the errors as much as possible. The parameters are largely executed by Baum-Welch. Compared with the I-VT identification, the identification result of the I-HMM identification algorithm with concept parameters is relatively better, but the process is more complicated.

2.4 Kalman Filter Identification (I-KF)

The Kalman filter refers to a recursive estimator that computes a future estimate of a dynamic system from incomplete and noisy measurements. The Kalman filter can minimize the mean of the squared estimation error and then predict the state of the system. The state estimation is calculated using the previous time step and the most recent measurement to reestimate the state of the dynamic system.

Speed and position are two important parameters in the I-KF identification algorithm. By analyzing the viewpoint coordinate system, I-KF can predict the eye movement velocity. The chi-square test can be used to calculate the predicted eye movement velocity to obtain the classification states such as fixation and beating. For the Kalman filter, the observation variables given by the gaze tracking system are gaze coordinates and time.

$$x^2 = \sum_{i=1}^{p} \frac{(\theta^{-}{}_{i} - \theta_i)^2}{\delta^2} \tag{5}$$

If it is smaller than the threshold value, the viewpoint is the fixation point, and vice versa.

3. Real-time Eye Movement Data Processing

3.1 Real-time Identification Algorithm based on Eye Movement Interaction

For the related data of smooth viewpoint, it is necessary to determine that the latest viewpoint belongs to the jump-start point or the continuous fixation point or the current viewpoint outlier to some extent, and it is necessary to determine the user's eye movement.

Three parameters, Eps(distance), MinTime(minimum fixation time) and θr(speed threshold), are prespecified. Eps refers to distance related parameters; MinTime refers to the time-dependent parameter, and the size of the data value is determined according to the minimum fixation time.

The algorithm process mainly includes two aspects: eye-saccade behavior identification and eye-saccade behavior identification. The specific steps of the algorithm are as follows:

Step 1: The initialization process of the system. The current state is recorded as the fixation state, and the measured first view P is taken as the note;

The initial point of view window, represented by C, input sequence in sequence;

Step 2: Calculate the current cluster center P. The cluster center is calculated by the mean value of a series of points in the window, and the weight is assigned to the positions of viewpoints in the window:

$$P_c = \left(\frac{\sum_{i=c}^{s} p_i}{S-C+1} = \frac{\sum_{i=c}^{s} x_{pi}}{S-C+1}, \frac{\sum_{i=c}^{s} y_{pi}}{S-C+1}, \frac{\sum_{i=c}^{s} t_{pi}}{S-C+1} \right) \tag{6}$$

Calculate the distance between the viewpoint and the middle point of the window. If the distance is satisfied, then the viewpoint is still fixation.

Step 3: If, then the line of sight track will change. On the one hand, the change of line of sight is caused by noise caused by line of sight oscillation, on the other hand, it is caused by intentional beating. It is classified as a potential window and returns to the clustering center at the same time.

Step 4: Judge the change of viewpoint. If the point speed of sight is smaller than the threshold value, then it is still a fixation behavior. Add the viewpoint to the current fixation window. If the point speed of the line of sight is larger than the threshold value, then when the point of view is added to the potential fixation window, the potential fixation window will become the current fixation window, and the system will switch to the jumping mode, which will become the starting point of the current fixation window.

Step 5: Calculate the cluster center point that appears in the cluster. If, and, then the gaze will appear, otherwise, it will still show the beating state.

Step 6: Repeat the above steps to continuously judge the state of eye movement until the viewpoint is entered and all processing is completed.

3.2 Viewpoint Filtering Algorithm

In order to eliminate noise in eye tracking system, many researchers use average filtering for real-time detection. In the latest round of detection process, the first data in the window is removed, and the remaining N-1 data are moved in order. At the same time, new sampled data is added, and filtered data is obtained after calculation. After all data collection is completed, the algorithm is completed.

$$y[n] = \frac{x[n] + x[n-1] + \ldots x[n-N+1]}{N} = \frac{1}{N}\sum_{k=0}^{N-1} x[n-k] \tag{7}$$

k refers to the data of the KTH eye movement viewpoint, x[k] refers to the information of the KTH viewpoint, and y[n] refers to the filtered output after the NTH sampling.

According to the characteristics of eye movement behavior, an adaptive eye smoothing filtering algorithm is proposed to reduce noise and eye fluctuation caused by tremor. The algorithm includes two kinds of positions: one is the original viewpoint position, the other is the filtered viewpoint position.

Current clustering centers:

$$P_m = \frac{1}{\sum_{K=1}^{N} E_{i-k}} \sum_{j=1}^{N} p_{i-j} * E_{i-j} \tag{8}$$

The ith viewpoint position:

$$\overline{P_i} = \frac{1}{\sum_{K=1}^{N} E_{i-k+1}} \{\sum_{j=1}^{N-1} p_{i-j} * E_{i-j} + P_i * E_i\} \quad (9)$$

Represents the ith viewpoint, refers to the input value; Represents the output and refers to the ith filtered viewpoint position. Is an impact factor that assigns an impact factor of 0 or 1 to each new viewpoint based on its distance from the cluster center. If, then the new viewpoint is the normal viewpoint; If, then it indicates that the new viewpoint has an effect on fixation, and the purpose is to get rid of the outliers caused by or inadvertently caused by eye movement, and threshold is its judgment threshold.

The relevant algorithm process of adaptive line-of-sight smoothing filtering is as follows:

Step 1: Set the collected data as the initial viewpoint and calculate the midpoint of the cluster.

Step 2: Determine the influence factor according to the distance between the viewpoint and the midpoint of the cluster, and perform the filtering processing of the latest N viewpoint data.

Step 3: Judge the viewpoint. If, then do the first step and the second step, if, then no change, if the system needs to restart the gaze, restart the new cluster, repeat the first step and the second step;

Step 4: Output smoothing viewpoint.

4. Experiment and Analysis

In the process of the movement of the car, to reside in the vision to control the movement of the camera, can achieve the rotate around, for real-time monitoring of the surrounding environment, if want to get more important environmental information, can again through the eyes of the more complicated the potential schedule done filming for the materials collection, which verify the feasibility of the experiment.

In order to evaluate the performance of the system, subjects were given an obstacle avoidance detour task experiment. The start and end points were specified, and the subjects had to walk through the eye-controlled car until the end point. The study included 10 participants, all of whom were participating in the experiment for the first time. Three strategies (keyboard, visual dwell time and eye potential) were selected for research and analysis, and the selection strategies of keyboard and eye-controlled interactive system were compared. In the experiment, the objective test (task selection completion time and average error) was carried out at the beginning, and one-way ANOVA was used to verify the effectiveness of the system. Finally, the subjective test is conducted through the usability evaluation survey of the system.

① Select the task completion time

Keyboard and eye motion: during the control of car operation, the average task completion time of keyboard selection strategy was (M=50.25s), and the average task completion time of eye motion selection strategy was (M=68.03S).

Keyboard and visual dwell: Cradle control, dwell time threshold can be set by its own conditions, usually set to 400ms.

② Average error rate

The average number of errors in a task is the number of errors in choosing a strategy for each task during the completion of several task strategies. There are the following situations: the number of command failures, the number of times that obstacles are not avoided and the number of times that they are triggered by mistake. The following statistics were made on the experimental situation of the subjects.

The keyboard had the lowest average number of errors when compared with the eye and visual dwell mechanisms for both the running tasks and the control tasks of the gimbal. Compared with the number

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0013

of keyboard errors, the error rate of visual dwell is a little higher, which is reflected in the continuous visual dwell operation, which brings uncomfortable experience to the subjects, which will lead to the error of sending commands or the sight shaking unable to trigger the operation.

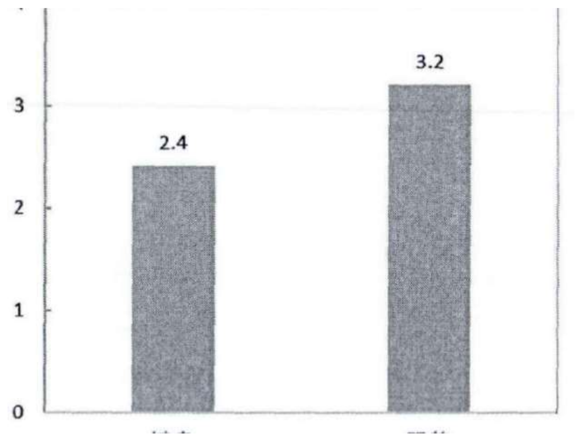

Fig. 3 Average number of task errors

In order to evaluate the usability of the system, participants were selected to conduct a questionnaire survey. The objective usability and subjective satisfaction of the system were measured by the survey. The conclusion is as follows:

● quickly complete the calibration and eye tracking, the sensitivity of the system is relatively high;

● The interface is more friendly, the visual resident selection and eye potential interaction mode of learning time is relatively short;

● Point of view into the completion area, this area highlighted, the user feedback this execution eye potential has ended, the car to enter;

The feedback greatly improves the interaction with the user and improves his accuracy rate.

● When the recognition rate of interaction becomes higher, eye potential can reduce fatigue, respond faster, and be more sensitive than visual dwell selection.

Through the above analysis, the following conclusion can be drawn: in the performance evaluation of the system, although the effect of the two strategies is not as good as that of the keyboard in terms of task completion and error rate, the obstacle avoidance operation can be completed efficiently with the cooperation of the two strategies. The usability evaluation of the system reflects the interactive system, timely feedback, and can reduce the mental pressure of the subjects, who can pay attention to their intention position randomly, so the selection of eye potential has a wide application prospect in remote interaction.

Acknowledgments

Item no.21KPXMRC00040.

References

[1] Zhao Ying, Liang Jinpeng, Yan Liqi, et al. Research on the Influence of Sign Setting on Tourists' Wayfinding Behavior: An experimental analysis based on eye tracking [J]. Tourism Tribune, 2020, 35(9): 63-73.

[2] Xue Yaofeng, Zeng Zhitong. Study on eye movement model of learners with different cognitive styles for adaptive learning [J]. Modern Educational Techniques, 2020, (8) : 91-97.

[3] Nunneley C E, Fishman M, Sundheim K M. Leading synchronous virtual teaching sessions[J]. The Clinical Teacher, 2020, (in Chinese) (3) : 231-235.

[4] Xue Yaofeng, Li Zhuowei. Research on quantitative model of cognitive load in online learning based on eye-tracking technology [J]. Modern Educational Technology, 2019, (7) : 59-65.

[5] Hu Weifeng, Zhang Kun, Zhang Zhongya. Design Evaluation of Fracturing vehicle based on eye tracking technology [J]. Journal of Machine Design, 2015, (06) : 109-112.

[6] YU Mingxuan. Research and Implementation of Key Technologies of Eye Tracking System [D]. Xidian University, 2013.

[7] Chen Chen. Introduction of Eye movement Technology and its application in Advertising Psychology [J]. Graduate Journal of Sun Yatsen University (Science Subject). Med Ed.), 2014, (04) :1-5.

[8] Lin Min. Research on Key Technologies of Human-computer Interaction Based on Eye Movement Information [D]. Shanghai University, 2014.

[9] Chen Jiaxu, Zhang Shaowei, Zhou Ying, Yang Changwei, Meng Tian, Chen Qiuting, Xu Xianchun, Hu Wenhan. Usability evaluation of medical device instruction manual based on eye movement technology [J]. China Health Industry, 2016, (13) : 35-41.

[10] Lin Shutan, Yin Changqing. Digital Gesture Recognition Based on LeapMotion [J]. Computer Knowledge and Technology, 2015, (35) :108-109.

[11] Sun Wencai, Yang Zhifa, Li Shiwu, Xu Yi, Guo Mengzhu, Wei Xuexin. DBSCAN-MMC method for Driver Gaze region Division [J]. Journal of Zhejiang University (Engineering Science), 2015, (08) : 1455-1461.

[12] Kangas J,Rantala J,Akkil D,et al.Vibrotactile Stimulation of the Head Enables Faster Gaze Gestures[J]. International Journal of Human-Computer Studies, 2016, 98: 62-71.

[13] Yang X, Li D, Liu X.Learner behaviors in synchronous virtual learning sessions[J].The Clinical Teacher, 2020.