Application of Eye Tracking in Intelligent User Interface

Yongqiang Liang^a^, Wei Wang^b^, Jue Qu^c^ and Jie Yang^d^  
Air Force Engineering University Xi’an, Shaanxi, 710051, China  
E-mail: ^a^1102268250@qq.com, ^b^85689437@qq.com, ^c^qujue402@sina.com, ^d^1102268250@qq.com

Abstract. Aiming at the requirement of intelligent user interface state awareness and intention prediction, a method of interactive state classification and intention prediction based on eye tracking is proposed. 5 states of interaction are defined, which are monitoring state, tracking state, decision state, burst state and off-loop state, design induced experiments were conducted to collect eye movement data in each state, a single factor analysis of variance shows that the 7 eye movements in the 5 interactive states has a significant difference. The prediction accuracy of the SVM classification model under category 5 conditions is 77.2%, while the accuracy rate of the classification under category 4 conditions is 85.9%, individual differences also have important effects on prediction accuracy, the accuracy rate of single subjects under category 5 conditions are more than 84%, while under category 4 conditions up to 90%. The research results are of reference value to the design and application of intelligent user interface.

1. Introduction
Intelligent user interface (IUI) is a new type of human-computer interaction interface which aims at high efficiency, effectiveness and naturalness, using artificial intelligence technology to construct the dynamic change interactive model which adapts to the demand of users. Through perception, reasoning and expression, according to users’ characteristics, background, task state, dialogue model and media form to achieve human-computer intelligence interaction [1].

A key aspect of intelligent user interface is “perception of user’s current state and prediction of the next operation”, state understanding and intent identification are the bases of intelligent user interface construction. Recently, an important research is conducted to identify or understand user intentions by analyzing the user’s behavior or state changes. For example, Pynadath et al. [2] used contextual relationships to solve the problem of planning identification in traffic monitoring through a generic Bayesian network. Bui et al. [3] proposed an on-line probability recognition strategy based on the abstract hidden Markov model (AHMM), and achieved the scalability of the strategy identification in the model by using the approximate reasoning scheme.

On the other hand, in cognitive psychology and cognitive neuroscience, human motivation can be analyzed and identified by various physiological signals such as EEG, eye movement and EMG. With the rapid development of eye tracking technology, the acquisition of eye movement signals is getting lower, and the accuracy and reliability of eye tracking are getting higher. Compared with EEG signals, the acquisition of eye movement signals has advantages of non-invasive, non-contact and simple operation, and based on the fact that eye movement can largely reflect human motives, it is an ideal method to use eye movement signals to identify human operating states and intentions. Recently, eye tracking technology research is widely used in psychology, medical, aerospace, traffic, military and other areas involving human-computer interaction.
In the human-computer interaction, explicit intentions such as open, close, delete are easily understood and executed by the computer, but implicit intentions usually involve people’s inner activities, therefore, they are difficult to understand and recognize directly by the computer [4]. However, the implicit intentions of humans can often be reasoned by facial expressions, physical movements, and eye features, specially, eye movement is an indispensable activity in the cognitive process of humans [5].

Eye movements are closely related to people’s cognitive processes, therefore, people’s visual features include rich and complex information about their interests, preferences and intentions. There are some uncertainties in the eye movement data. Wang et al. [6] proposed a method of intention identification based on eye tracking and fuzzy reasoning, but due to human factors leading to low degree of system intelligence, accuracy and real-time is not high. Lee et al. [7] compared the three different types of intentions of “navigation intention”, “low specific intention” and “high specific intention”. He found that there were no significant differences in the eye movement patterns of the three intentions before the visual scene changes, but after the visual scene changes there are significant differences between the three, indicating that the type of intention will affect the eye movement mode in the change detection task. Park et al. [8] subdivides the specific intentions of humans into cognitive intentions and affective intentions, and it is found that eye movements corresponding to affective intention have more fixations than cognitive intention, and the average pupil diameter larger. Jang et al. [9] classifies the implicit intentions of visual stimuli into information intention and navigation intention. The two were classified by the number of fixation points, the amplitude of the eye and the size of the pupil.

2. Experimental design

2.1. Experimental design background
Human-computer interaction processes such as aerospace, medical imaging, nuclear power operations, and military allegations are often complicated, time-critical. The fault-tolerant rates are low, and operators often face higher workloads. The intelligent user interface is designed to adapt the user’s current state and predict its operational intent, and adaptively adjusts its interface presentation to provide the user with the most appropriate control interface at the right time.

Analysis and summary of the human-computer interaction process, the operation of the user’s possible state and intent are defined as five categories. Monitoring state can be seen as some of the scholars referred to the low specific intent, the operator in this state has no specific target intention, only a rough access to the interface of visual information, and can be free to focus on any content of their own interest, such as a smooth flight of the aircraft or a radar monitor when there is no enemy. Tracking state is a high specific intention, in this state, the operator often needs to focus on some objects according to the task, and try not to be interrupted. For example, the aircraft in the takeoff and landing process need to focus on observation speed, height, pressure and other indicators of the instrument. Medical personnel in the rescue process need to continue to pay attention to the patient vital signs parameters on the measurement instrument interface. Decision state can be interpreted as a decentralized intent, according to the theory of information processing, decision-making is the mapping relation of information to multiple pairs of reactions, to produce a choice often need to obtain a lot of information and make their analysis and evaluation, like air defense operations, the commander needs to quickly make a threat to determine the situation based on the enemy’s target characteristics and equipment and timely issued a fire command in a short time. Sudden state is a problem that any human-computer interaction process may face, such as aircraft flight process by the meteorological factors abnormalities, sudden failure of nuclear power equipment and radar monitoring in the release of interference, which may all cause unexpected events so that operators face a greater mental load and overwhelmed. The last off-loop state means that the operator is out of the task loop, it can be understood as absent-minded or distracted, for example, an excessive degree of automation or a
long boring job may leave the operator out of the task loop, under this state the operator’s mental model is disrupted, it may cause serious security risks.

2.2. Subjects
Students at the Air Force Engineering University in China (N = 18) participated in this experiment, including 10 males and 8 females, aged 18 to 22 years old (SD = 3.08). All subjects were in good health, and had normal or corrected-to-normal vision, no color blindness, color weakness or other characteristics. All subjects had more than 5 years of computer experience, and mastered basic knowledge of military operations.

2.3. Materials
The main experimental stimuli was shown in Figure 1, which stimulate the radar surveillance and command decision process in air defense operations, the entire interface can be divided into five areas, the black circular area above the left is an analog radar surveillance area, the different colors and symbols in the area represent the targets of both sides, and the specific meanings are explained in the lower left of the experiment interface, this region can be a rough expression of the target height, orientation and quantity and distribution. The upper right of the interface is the specific value of the orientation, height, distance, and speed of all the targets on the left side, the real-time status of the target can be accurately obtained through the area table. The bottom right of the interface is composed of the information prompt area and the operation button area. The prompt content of the information prompt area is the operation instruction that needs to be executed by subjects in the experiment, the area where the operation is performed is the operation button area at the bottom right of the interface, all 8 buttons can be clicked by the mouse. All of the elements of the stimuli (graphic symbols, tables, buttons) are not intended for any existing display interface, all are experimental needs simulation production.

Figure 1. Experimental interface stimuli
2.4. Experimental tasks and processes
In order to collect the eye movement data of the subjects in five kinds of interactive states, the experiment requires each subject to complete the experimental task under five state preset scenarios, the experimental task was designed to induce the subjects to enter 5 preset interactive states, for example, by asking the subjects to “review the meals in the past two days” during the experiment, the subjects were separated from the task loop and entered the off-loop state, the specific task instructions are shown in Table 1.

The experiment program is made by E-prime software, and the specific process is to complete the operation requirements of the 5 simulation interactive state tasks successively, under each state of the task first read the instructions to the subjects and emphasized the experimental notes, and then calibrate the eye movement point of view. After calibration, press the space bar to enter the point of view aggregation interface, continuous 2000ms automatically into the experimental stimulation interface. Then the subjects completed the experimental task prompted on information prompt area, and press the space bar to enter the end of the interface, so that an experimental task is completed, lasts 2000ms automatically into the next experimental task until all the experimental tasks are completed.

Table 1. Experimental tasks introduction.

| Interactive state | Experimental tasks |
|-------------------|--------------------|
| monitoring state  | Please observe the contents of the interface at will, without any action. |
| tracking state    | 1. Please report the enemy bomber’s position, height, distance and speed.  
|                   | 2. Please report the number of enemy air raid missiles in each area and find the lowest missile target. |
| decision state    | Assumption: threat degree = (number of enemy cruise missiles + number of air raid missiles + 2 × bombers) - (our air defense missile number + 2 × fighter number)  
|                   | 1. Please determine the highest threat area, and click on the blow area.  
|                   | 2. There are 20 fighter reinforcements, please arrangements for the number of reinforcements in the region reasonable. |
| burst state       | The task is the same as the "decision state", but the interface will simulate the abnormal interference, and some areas of the interface become visual blind. |
| off-loop state    | Please review your meal in the past two days. |

2.5. Apparatus
Experimental apparatus, including a Lenovo PC, RED5 eye tracker, the highest sampling frequency is 500Hz. The host display resolution is set to 1280×1024 pixels, the screen brightness is adjusted to 300cd/cm2, the distance between the subject and the screen is about 60cm, The center of the screen is substantially flush with the eyes of the subjects.

3. Experimental results

3.1. Difference analysis of eye movement characteristics in 5 interactive states

The relevant eye movements data in 5 interactive states are shown in Table 2. In order to facilitate the expression, eye movements indexes are expressed in short form, the specific rules are as follows: AFD=Average Fixation Duration, APX=Average Position X, APY=Average Position Y, APS=Average Pupil Size, APD=Average Pupil Diameter, ASD=Average Saccade Duration, Amp=Average Amplitude, ASA=Average Saccade Acceleration, ASV=Average Saccade Velocity. Similarly, State 1 to 5 represent monitoring state, tracking state, decision state, burst state and off-loop state.
Table 2. Experimental tasks introduction.

|       | State1 | State2 | State3 | State4 | State5 |
|-------|--------|--------|--------|--------|--------|
| AFD[ms] | 372.6  | 393.6  | 385.0  | 325.4  | 329.8  |
| APX    | 734.5  | 830.2  | 607.4  | 798.8  |        |
| APY    | 423.2  | 445.5  | 406.7  | 399.0  | 195.4  |
| APS[mm] | 11.92  | 12.65  | 13.84  | 14.01  | 14.24  |
| APD[mm] | 3.07   | 3.41   | 3.58   | 3.61   | 3.73   |

In order to compare the difference of eye movement characteristic data in five kinds of interactive states, the corresponding bar graphs are shown in Figure 2, and the abscissa 1 to 5 in the figure represent the five interactive states. It can be seen from Figure 2, APS, Amp, ASA, and ASV have a great difference in five kinds of interactive states, other eye movement characteristics difference is not obvious. The results of ANOVAs showed that the average fixation duration \(F(4,85)=3.74, p<0.05\), the average pupil size \(F(4,85)=28.48, p<0.001\), the average pupil diameter \(F(4,85)=6.52, p<0.05\), the average saccade duration \(F(4,85)=3.27, p<0.05\), the average amplitude \(F(4,85)=12.25, p<0.001\), the average saccade acceleration \(F(4,85)=6.39, p<0.001\), the average saccade velocity \(F(4,85)=12.57, p<0.001\) have a significant differences in 5 kinds of interactive states. However, the average position X \(F(4,85)=2.61, p=0.27\) and the average position Y \(F(4,85)=1.69, p=0.09\) of the fixation points don’t have a significant difference in the five interaction states.

Figure 2. Eye movement characteristic data statistics

According to the above analysis results, the seven eye movement characteristics of the average fixation duration, average pupil size, average pupil diameter, average saccade duration, average amplitude, average saccade acceleration and average saccade velocity can be classified as five kinds of interactive states feature components. It will be validated by using SVM later.
3.2. SVM classification solution

![Figure 3](image-url)

Figure 3. Classification prediction results of test set

The average fixation duration, average pupil size, average pupil diameter, average saccade duration, average amplitude, average saccade acceleration and average saccade velocity, these seven eye movement indexes were chosen as characteristic components to classify the five interactive states. Since the AOI data does not have real-time measurements, it is not used as a characteristic component. MATLAB solution is Accuracy = 77.2% (193/250) (classification).

As the eye movement feature fractal visualization shows that the individual eye movement characteristic component is not sufficient to classify the sample, while the accuracy rate of classification is 77.2% for the combination of seven eye movements characteristic. Obviously this figure is not very high. For further analysis of the classification of test samples shown in Figure 3, it can be seen that in the decision state (category 2) a considerable part of the test data is wrongly classified as off-loop state (category 4), the other categories of classification’s confusion is not obvious, within the acceptable range.

4. Discussion and analysis

4.1. The influence of class number on the accuracy of classification forecast

![Figure 4](image-url)

Figure 4. Classification prediction results of test set

The interaction state of 5 categories for classification prediction only to obtain the correct rate of 77.2%, after analyzing the test set data, it is found that the main reason for the classification error is that the SVM classifier incorrectly classified the decision state as off-loop state. The reason may be
that the 2 states require more intense brain cognitive activity, and some eye movements may be similar. Considering that the off-loop state is not the primary state in the interaction, the discussion and analysis section removes this state, and then proceeds to discuss the interactive state of the 4 classification problem. The part of the off-loop state in the experimental data is all removed, similarly, the algorithm and program part can be modified to get the solution model of the 4 classification problem. The test set classification result of the 4 categories is shown in Figure 4, and the prediction accuracy of classification is 85.9%. This result is clearly much higher than the 5 classification.

4.2. The influence of individual differences on classification accuracy

Researchers (Park et al., 2016) have pointed out that differences in the eye movement patterns of the subjects individual will result in a reduction in the accuracy of the classification prediction. The conclusion is similarly validated for the experimental task of this paper. The experimental data of 5 subjects were randomly selected from 18 subjects, each subject’s experimental data were separately verified by SVM, the results showed that the accuracy of prediction was more than 84% in the case of interactive state 5 classification, and 90% in the case of interaction state 4 classification, specific results are shown in Table 3.

| Table 3. Experimental tasks introduction. |
|-------------------------------------------|
| S1     | S2     | S3     | S4     | S5     |
|--------|--------|--------|--------|--------|
| 5 classification | 0.862   | 0.845   | 0.852   | 0.897   | 0.853   |
| 4 classification | 0.903   | 0.915   | 0.917   | 0.908   | 0.922   |

5. Conclusions

In this paper, we took intelligent user interface as the research background, aiming at the demand of intelligent user interface for operator’s state awareness and intention prediction, and propose a method of interactive states classification based on eye movement characteristics. First, the interactive state (intent) was summarized as monitoring state, tracking state, decision state, burst state and off-loop state. By designing experimental scenarios, the subjects were induced to enter the corresponding interactive state and collected real-time eye movement data. And then use the ANOVAs to determine the seven eye movement characteristics which have a significant difference in five interactive states. Finally, SVM classification prediction model was established, and discussed the prediction accuracy of 5 classification and 4 classification, it is found that the accuracy of 4 categories was significantly higher than the 5 classification, and proved that individual cognitive differences have important influence on prediction accuracy. The results of the study have some reference value for intelligent user interface state perception and intention prediction module design.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (51675530).

References

[1] Maybury, M. (1999). Intelligent User Interfaces: An Introduction. Proceedings of the 4th international conference on Intelligent user interfaces, California, USA.

[2] Pynadath, D. V., & Wellman, M. P. (1995), Accounting for Context in Plan Recognition, with Application to Traffic Monitoring. Proceedings of the Eleventh International Conference on Uncertainty in Artificial Intelligence, Montréal, Canada.

[3] Bui, H. H. (2003). A General Model for Online Probabilistic Plan Recognition. Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico.

[4] Ma Lisha, Lv Jian, Pan Weijie, Sfan Junjun, & Ping Zhengqiang. (2017). Research on Implicit Intention Recognition and Classification Based on Eye Movement Pattern. JOURNAL OF GRAPHICS. 38(3), 332-340.
[5] Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye Movements and Perception: A Selective Review. Journal of vision, 11(5), 89-91.

[6] Wang, M., Maeda, Y., & Takahash, Y. (2012). Human intention recognition via eye tracking based on fuzzy inference. Joint International Conference on Soft Computing & Intelligent Systems, Kobe, Japan.

[7] Lee, S., Park, H., Lee, M., & Kwak, H. W. (2015). Effects of search intent on eye-movement patterns in a change detection task. Journal of Eye Movement Research, 8(2), 1-10.

[8] Park, H., Lee, S., Lee, M., Chang, M. S., & Kwak, H. W. (2016). Using eye movement data to infer human behavioral intentions. Computers in Human Behavior, 63, 796-804.

[9] Jang, Y. M., Mallipeddi, R., Lee, S., Kwak, H. W., & Lee, M. (2014). Human intention recognition based on eyeball movement pattern and pupil size variation. Neurocomputing, 128, 421-432.