Classification of Cognitive Load Based on the Pilot' Visual Scan Patterns

Jie Yao[1], Yanyu Lu[2,*], Zhenjie Zhong[1], Shan Fu[2]

1. China Southern Airline

2. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, P.P. China

* Corresponding author: Yanyu Lu, luyanyu@sjtu.edu.cn

abstract

In aviation accidents, human factors account for the majority causes instead of the technical factors like the system failure and especially. One of the aviation human factors researches is the analysis of pilot tasks and their effect on the cognitive load of the pilots. Since the visual channel is an important channel for the pilot to obtain information and feedback, this study focuses on the differences between the pilots' visual scan patterns in normal tasks and emergency tasks in order to reveal how pilots' basic visual state changes during flight tasks. Ten airline pilots were invited to participate the experiment to complete flying normal and abnormal tasks in a simulated cockpit. The study examines the differences of scan patterns during different flight tasks by eye movement features-average fixation duration, saccade frequency and entropy. The results from the statistical analysis showed that these eye movement features were significantly different among tasks of different cognitive load levels. Besides, this paper uses a variety of classification algorithms to test the classification ability of different eye movement features in different tasks, including the algorithms of K-means, SVM (support vector machines) and LR (logistic regression). The results show that the SVM and LR with supervised learning have higher classification accuracies, respectively) compared with the unsupervised K-means when combining with three eye movement metrics.

Keywords: pilots' cognitive load, visual scan pattern, human factors.

Introduction

In aviation accidents, human factors account for the majority causes instead of the technical factors like the system failure[1]. Increasingly complex human-computer interfaces (HCIs) and tasks are creating the challenges to the cognitive load of the pilots that could affect their decision and behaviors during tasks. Therefore, an important research direction of aviation human factors is the analysis of pilot tasks and their effect on the cognitive load of the pilots.

The task process is closely related to the information obtained by the pilot, in which the visual channel is an important channel for the pilot to obtain information and feedback. Pilots’ visual scan pattern can reflect pilots’ attention allocation, cognitive states and their tasks being performed. Besides, pilots’ visual scan pattern can be captured continuously and measured objectively through tracking eye movements without interrupting the pilots’ activities[2]. Therefore, pilots’ visual scan pattern based on the eye movements is an important content of cockpit human factors research.

Researchers focused on the pilots’ visual scanning behaviors during a flight task to study the pilots’ mental and physical workload, situation awareness, cognitive states and task performance, and then to evaluate the design of the HCIs. Sirevaag et al [3] studied the indices of oculomotor activity of helicopter pilots during simulated low-level flight task, and revealed that task time resulted in longer blink duration and fewer and later reactive saccades. Merwe et al [4] reported that eye movement could be an indicator of situation awareness during a simulated flight task, where the fixation rate and dwell time reflected the information acquisition, and entropy reflected the new information acquisition activities. Yu et al [5] measured the effect of visual scanning patterns on situation awareness through the use of eye-tracking tools in a flight simulator and found that pilots with better situation awareness had lower perceived workload. Lu et al [6] found that the pilots had different visual scanning mode according to the flight mode and their experience during simulated flight task. Schriver et al [7] investigated expertise differences in pilot decision making by measuring a hypothesized attention-action link. Jin et al [8] studied how expert and novice pilots can distribute their visual attention to improve flight performance.

Pilots’ visual scan pattern is influenced by the task procedures and task demands, in additional to an indicator of the pilot cognition state and performance. The pilots scan different displays and obtain the relevant information in terms of the different tasks. Their visual scan path was dependent on the task procedure. Pilots' gaze sequence and EEG were used to identify different flight tasks in Yang et al study[9]. Doane and Sohn [10]found that fixation time was much more in the uncertain and high-risk task than that in the certain task. Li et al [11] examined the pilots' attention distributions between chasing a moving target and a stationary target, and found pilots' visual behavior is significantly associated with task characteristics.

Pilots' visual scan patterns are composed of fixations, saccade and other basic visual information, which reflect pilots' basic states and visual requirements from the task. We hypothesize that the combination of multiple eye movement metrics can provide a reliable means to identify complex tasks. This study focuses on the differences between the pilots' visual scan patterns in normal tasks and emergency tasks in order to reveal how pilots' basic visual state changes during flight tasks.

Experiment

Subjects

Nineteen male airline pilots with 30-50 years were invited to participate the experiment. They were paired into 10 flight crews in which one pilot served as the Pilot Flying (PF) and the other pilot as the Pilot Monitoring (PM). They were informed of the purpose and procedures of the experiments and signed an informed consent form prior to participation.

Apparatus

The experimental apparatus included a A320 full-flight simulator and an eye tracker. The simulator was a qualified simulator (level D) conforming to the guidelines presented in the Federal Aviation Administration Advisory Circular (AC120-40B) – Airplane Simulator Qualification, which has also been used for pilot training for commercial airlines. The eye tracker, Tobii glasses 3 (Tobii Technology, Stockholm, Sweden) at a sampling rate of 100Hz, is used to track the pilot's eye behavior and record eye movement data.

Task procedure

The pilots are asked to complete flying different tasks in a simulated cockpit. These tasks include normal and abnormal tasks, which could impose different cognitive load on the pilots.

During the normal task, the pilots should perform an altitude-keeping task, while during the abnormal task, a system failure with engine on fire, occurred during the cruise phase, and the pilots should deal with the failure.

Before the experiment, all pilots were trained with a flight profile to help themselves familiar with the experiment procedure and the eye tracker.

Data process

The metrics of eye movements

The study examines the differences of scan patterns during different flight tasks by eye movement metrics-average fixation duration, saccade frequency and entropy. Eight areas of interest (AOIs) are developed according to the functions and information displayed on the flight deck: primary flight display (PFD), multi control panel (MCP), electronic centralized aircraft monitoring (ECAM), flight control panel (FCP), control display unit (CDU), standby (STANDBY), out of the windows (OTW) and others.

Figure 1. An example of fixation position and fixation transition sequence in one minute

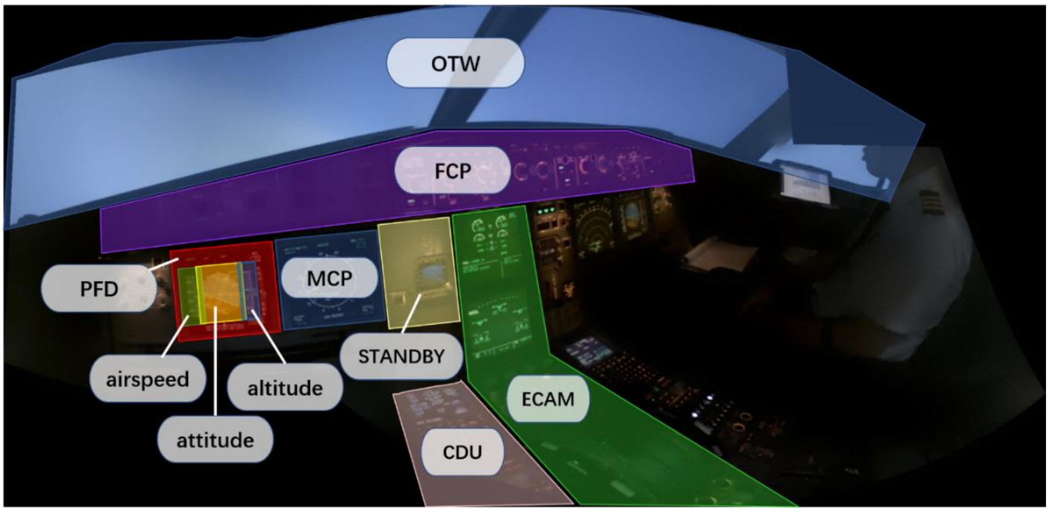

Figure 2. Areas of interests of the flight deck

The fixations and saccades are identified according to the Velocity-Threshold Identification algorithm in the eye tracker software. The average fixation duration per area of interest can show the pilot's scan pattern and gaze allocation, which may have great differences among individuals with the change of cognitive load. The average fixation duration in an AOI was calculated as

$$F_{dwell} = \frac{\sum_{i}^{N} end\_t_i - start\_t_i}{N} \quad (1)$$

where $start\_t_i$ was the start time of the i^{st} fixation and end_t_i was the end time. N was the fixation number in the AOI.

The saccade frequency can measure the ability and requirement of information searching, which was calculated as the saccade number each minute.

Entropy, a concept from information theory, represents the randomness of the scan pattern of the pilot. Entropy is computed by the transition matrix as follows:

$$H(x) = -\sum_{i=1}^{n} p_i \sum_{j=1}^{n} p(i,j) \log_2 p(i,j) \quad (2)$$

where p_i is the proportion of sequence x in state i and $p(i,j)$ represents the probability of the transition from state i to state j, which means the probability of transition from one AOI to another in this study. The larger the Entropy value, the more complex the scan pattern and the harder it is to predict.

The metrics of the eye movements from different tasks are analyzed with paired-T test to examine the statistically difference.

The classification models

After the statistical analysis, three algorithms f K-means clustering, SVM (support vector machines) and LR (logistic regression) are applied to establish a classification model of the cognitive load:

1) K-means clustering

K-Means Clustering is an Unsupervised Machine Learning algorithm, which groups N data points into K clusters depending on their distance from the center of the clusters. The algorithm works as follows:

[1]. k points is randomly selected as initial cluster centroids $C = \{c_1, c_2, \dots, c_k\}$.

[2]. each item in the dataset is categorized to its closest mean;

[3]. the mean's coordinates is updated, which are the averages of the items categorized

in that cluster $c_i = \frac{1}{|c_i|}\sum_{x \in c_i} x$.

[4]. we repeat the process for a given number of iterations and at the end.

2) SVM

SVM is a supervised machine learning algorithm for classification and regression. It classifies data by constructing an optimal hyperplane that maximizes the distance between each class. The expression of the hyperplane is

$$\boldsymbol{\omega}^T \boldsymbol{x} + b = 0 \tag{3}$$

where $\boldsymbol{\omega}$ is the normal vector, $\boldsymbol{x}$ is the training sample, and b is the displacement term. Since SVM solves the convex optimization problem, the objective function is

$$\begin{aligned} &\max_{\omega,b} \frac{1}{\|\omega\|^2} \\ &\text{s.t.} \quad y_i\left(\omega^T x_i + b\right) \geq 1 \end{aligned} \tag{4}$$

Through the Karush-Kuhn-Tucker condition and Lagrange multiplier method, its dual problem is

$$\begin{aligned} &\min_{\alpha} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n}\alpha_i \\ &\text{s.t.} \quad \sum_{i=1}^{n}\alpha_i y_i = 0 \\ &\qquad \alpha_i \geq 0, \quad i = 1,2,\ldots,n \end{aligned} \tag{5}$$

The selection of the sample will affect the classification accuracy in the supervised machine learning algorithm. the optimal visual scan patterns from the experienced pilots are selected in the presented study to train the model.

3) LR

LR uses a logistic function to model a binary values (0 or 1) classification. The logistic function is

$$g(z) = \frac{1}{1+e^{-z}} \tag{6}$$

The hypothesis function of LR is：

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}} \tag{7}$$

where $\boldsymbol{x}$ is training data and $\boldsymbol{\theta}$ is a coefficient that must be learned from the training data. This is done using maximum-likelihood estimation, as follows:

$$L(\theta)=\prod_{i=1}^{k} h(x_i)\prod_{i=k}^{n}(1-h(x_i)) \tag{8}$$

The gradient descent algorithm is used to reach the optimal parameter $\boldsymbol{\theta}$。

Results

The results of test

The results from the statistical analysis showed that these eye movement features were significantly different between tasks of different cognitive load levels. In the normal tasks where the cognitive load of the pilot is higher, the average fixation duration is significantly longer (P = 0.049, Figure 3), and the saccade frequency is significantly lower (P = 0.041 , Figure 4), which suggest that the pilots pay more attention to obtain the information and deal with the failure. The entropy is significantly less (P = 0.015, Figure 5) during the abnormal task. Obviously, when dealing the failure, the scan patterns of the pilots are more regular since they pay more attention to relevant displays such as ECAM, while in the normal task, they should scan different displays to monitor the condition of systems. Therefore, the scanning patterns with these features could have the potential to classify the different cognitive load.

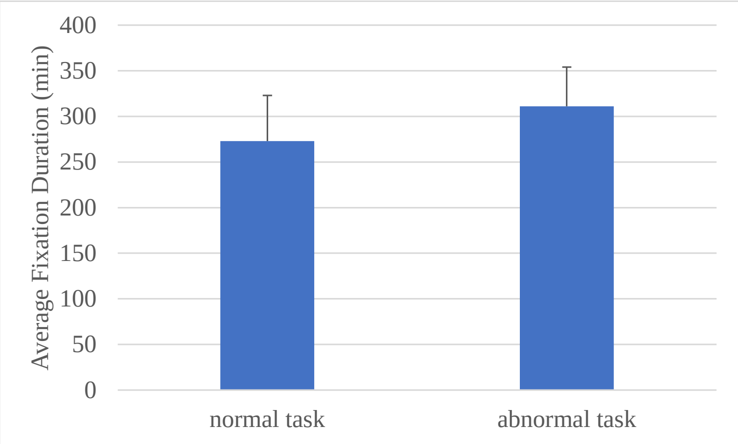

Figure 3. the average fixation duration of the pilots during the normal and abnormal tasks

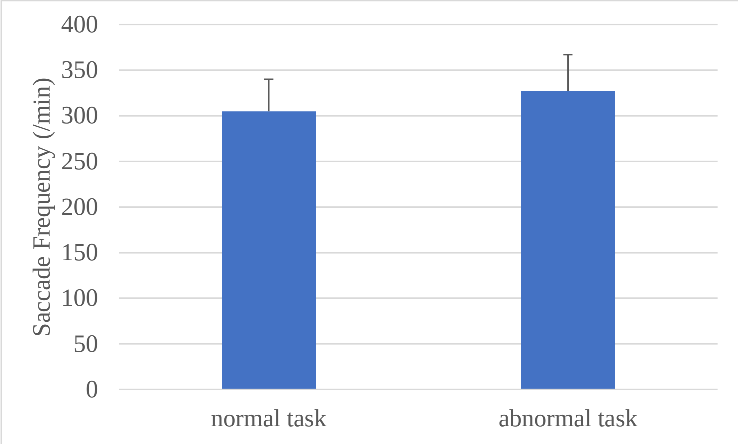

Figure 4. the saccade frequency of the pilots during the normal and abnormal tasks

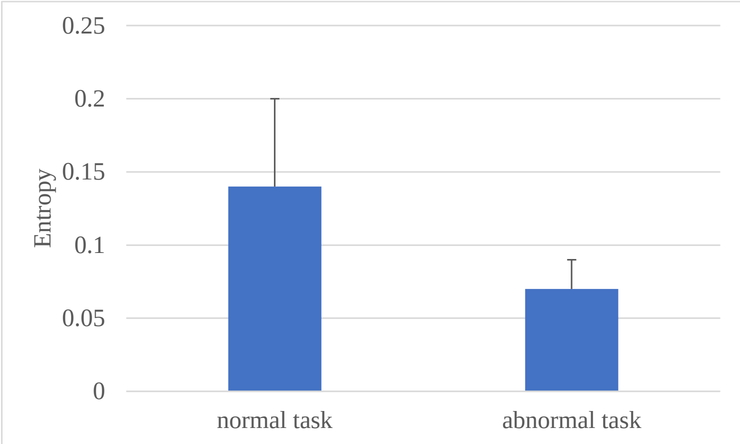

Figure 5. the entropy of the eye movement of the pilots during the normal and abnormal tasks

Classification model

When using the single feature of eye movement for the cognitive load classification, the results show that whichever algorithm is used, classification accuracy with the entropy is better than those with fixation duration and saccade (Figure 6-8), which means that the entropy of the eye movement could describe the scanning pattern better than the fixation duration and saccade.

Furthermore, we conduct the classification models based on the combination of multiple eye movement metrics (average fixation duration, saccade frequency and entropy). The classification results reveal that the combination of multiple eye movement metrics can do better than a single eye movement metric in the SVM and LR models. Comparing three classification algorithms, the SVM and LR with supervised learning have higher classification accuracies (83.07% and 83.85%, respectively) compared with the unsupervised K-means (79.72%).

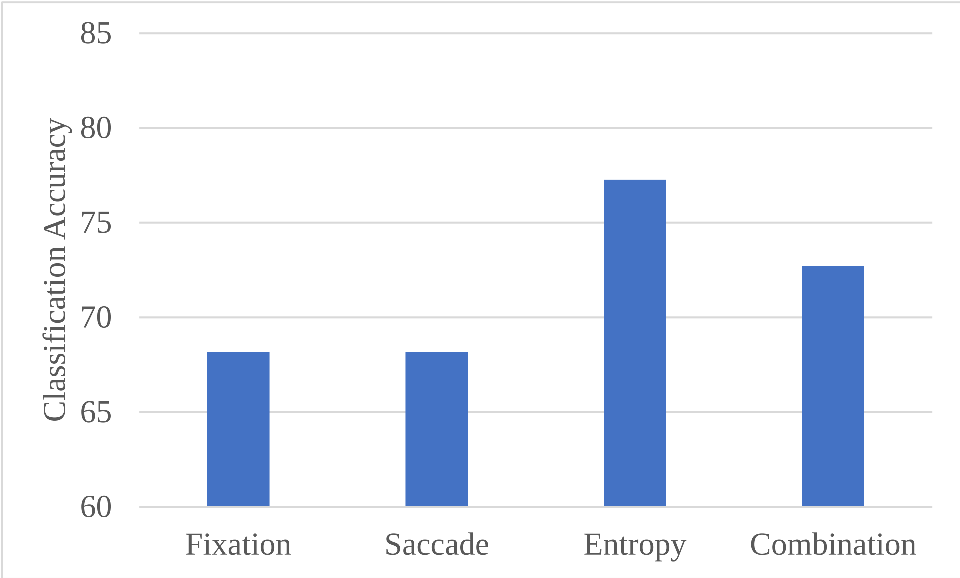

Figure 6. the classification accuracy of visual scanning patterns using the K-Means algorithms

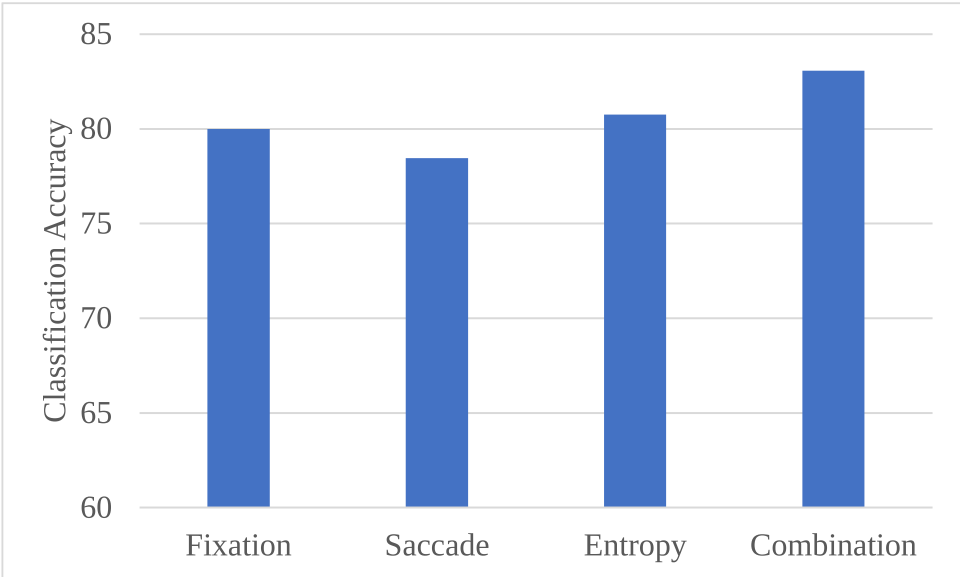

Figure 7. the classification accuracy of visual scanning patterns using the SVM

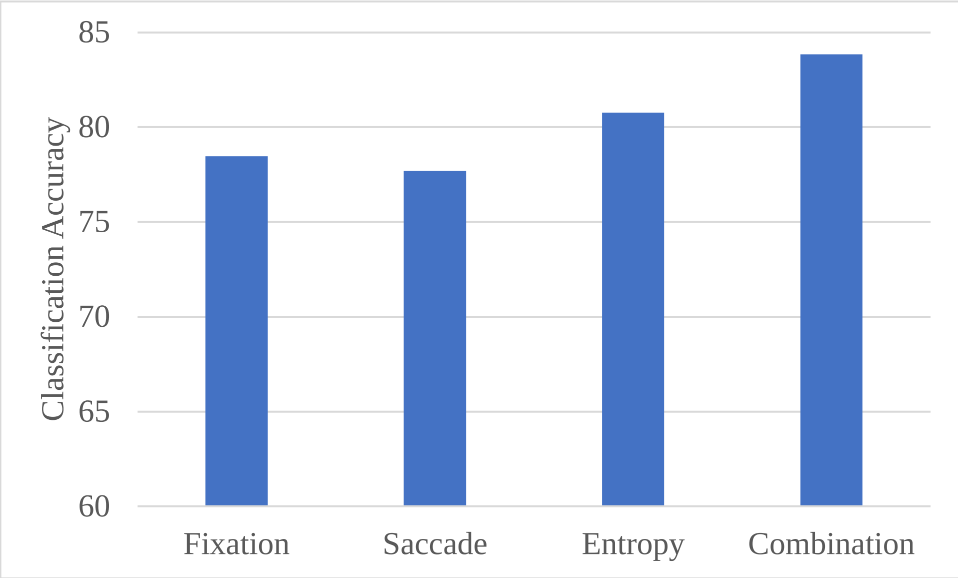

Figure 8. the classification accuracy of visual scanning patterns using the LR

Conclusions

In conclusion, the present study explores the visual scanning pattern during the normal and abnormal tasks in the simulator. The eye movement metrics comprise the fixation duration, saccade frequency and entropy, which can represent the visual workload and the regularity of attention distribution. The results reveals that the pilots had different scanning patterns according to the flight tasks, which could indicate their different cognitive load. We expect that pilots' visual scanning behaviors during tasks will help the training and the design of the human-machine interaction. Considering the existed research, the future work is expected to take a further insight into the relationship of eye movement data and brain activities. The eye tracker is a more acceptable tool with less interference compared with an EEG cap, although the brain activities are the direct indicators of the cognitive states. Multiple physiological signals should be considered to establish better classification models with high accuracy and better reliability.

References

[1] Sant'Anna, D. A. L. M. d.,Hilal, A. V. G. d. The impact of human factors on pilots' safety behavior in offshore aviation companies: A brazilian case[J]. Safety Science, 2021, 140 (4): 105272.

[2] Wickens, C. D.,McCarley, J. S. Applied attention theory[M]. CRC press, 2007.

[3] Sirevaag, E. J.,Rohrbaugh, J. W.,Stern, J. A., et al. MULTI-DIMENSIONAL CHARACTERIZATIONS OF OPERATOR STATE: A VALIDATION OF OCULOMOTOR METRICS[J]. Cardiovascular System, 1999.

[4] Merwe, K. V. D.,Dijk, H. V.,Zon, R. Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment[J]. International Journal of Aviation Psychology, 2012, 22 (1): 78-95.

[5] Chung-San,Yu,Eric, et al. Pilots' visual scan patterns and situation awareness in flight operations[J]. Aviation Space & Environmental Medicine, 2014.

[6] Lu, Y.,Zheng, Y.,Wang, Z., et al. Pilots' Visual Scanning Behaviors During an Instrument Landing System Approach[J]. Aerospace Medicine and Human Performance, 2020, 91 (6): 511-517.

[7] Schriver,A.,T., et al. Expertise Differences in Attentional Strategies Related to Pilot Decision Making[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2008, 50 (6): 864-878.

[8] Jin, H.,Hu, Z.,Li, K., et al. Study on How Expert and Novice Pilots Can Distribute Their Visual Attention to Improve Flight Performance[J]. IEEE Access, 2021, PP (99): 1-1.

[9] YANG, K. W., Hao-ran. Behavior Pattern Recognition for Pilots Using Head up Display[J]. Science Technology and Engineering, 2018, 18 (29): 6.

[10] Doane, S. M.,Sohn, Y. W. ADAPT: A Predictive Cognitive Model of User Visual Attention and Action Planning[J]. User Modeling and User-Adapted Interaction, 2000, 10 (1): 1-45.

[11] Li, R.,Cui, J.,Gao, R., et al. Situation Awareness Recognition Using EEG and Eye-Tracking data: a pilot study[M]. 2022; 209-212.

Copyright Statement