Experiment on Cognitive Load of Gesture Interaction Based on Semiotics

Wei Sun¹, Yudi Jin¹, *, JinLong Wang¹ and Hanbing Yang¹

¹Beijing University of Posts and Telecommunications, China

*Corresponding author e-mail: 2013212217@bupt.edu.cn

Abstract. Objective: To explore the amount of information contained in gesture symbols and its relationship with cognitive load, and to provide theoretical guidance for gesture interaction design based on semiotics. Methods: Firstly, a user-defined method was combined with oral analysis to collect gestures and establish a corresponding semantic knowledge base. Based on the knowledge base, the icon algebra theory is used to redesign the gesture symbols. Then through experiments combined with the NASA-TLX scale to measure the user's cognitive load, and through data analysis to explore the relationship between the amount of information contained in the gesture symbol and it. Conclusion: The amount of information contained in gesture symbols has a significant correlation with the user's cognitive load and response time. As the amount of gesture symbol information increases, the user's cognitive load and response time also increase accordingly. Based on this, a regression equation of user cognitive load and information amount of gesture symbols is established, which provides theoretical support for the design of gesture symbols based on semiotics.

Keywords: Gesture Interaction, Cognitive Load, Information Theory

1. Introduction

Gesture interaction is an interaction method that relies on hand movements or shapes to convey information [1], and has natural and intuitive characteristics [2]. With the rapid development of computer technology, traditional mouse-key input and screen-based information input methods are gradually unable to adapt to modern human-computer interaction systems [3], and the human-computer interaction methods will increasingly include Natural interactions including gestures [4-5].

As a theory of research significance [6], the introduction of semiotic theory into the design of gesture interaction can enhance the rationale of gesture design, facilitate understanding and recognition, and reduce the user's memory burden [7]. Previous studies of gesture interaction based on semiotics mainly focused on how to use semiotics' extension meaning and connotative meaning for symbolization [8-9], and did not quantitatively describe the cognitive load of gesture symbols.

As a functional symbol, gesture symbols should be as simple and easy to remember as possible, and should not consume too much physical and psychological resources of users. In this paper, information theory-related theories are used to quantify the amount of information contained in the
gesture symbols, and the participants' cognitive load is measured through experiments and the NASA-TLX scale to further study the relationship between the two.

2. Experimental Study on the Relationship between Information Volume and Cognitive Load

2.1. Gesture Symbol Information Quantization

The experiment first invited 6 participants (4 males and 2 females) and asked them to make as many different gestures and video as possible in 10 minutes without any help. By analyzing the collected gestures, experiments have found that within the error range, a person can reach approximately 18 spatial locations with one hand under a straight arm. Among them, there are 7 coronal planes (excluding the origin), 5 sagittal planes, 3 45° oblique cut planes in front left and 3 45° oblique cut planes in front right. Since the human hand will naturally face forward when reaching the highest point, the highest points of the faces other than the coronal plane are not counted. On this basis, combined with 4 types of arm bends (including clamp arms and 3 types of 90° elbow arms of normal, eversion and varus), a total of approximately 90 spatial positions can be reached with one hand. After reaching the space position, a person's shoulders, elbows, wrists, and fingers can perform a total of 28 limb movements as shown in Table 2, and any two of these movements will not be performed simultaneously in a natural state. In addition, human palms also have three hand shapes: natural, curved and fist. Combining the above-mentioned spatial locations, hand shapes, and limb movements, almost all one-hand gestures of a person can be made.

Table 1. The space locus.

| The space locus | 45°, 90°, 135°, 180° |
|-----------------|---------------------|
| Left coronal plane | 45°, 90°, 135° |
| Right coronal plane | 45°, 90°, 135° |
| Anterior sagittal plane | 45°, 90°, 135° |
| Posterior sagittal plane | 45°, 90° |
| 45° front left oblique section | 45°, 90°, 135° |
| 45° front right oblique section | 45°, 90°, 135° |

Table 2. Body movements.

| Body movements | Left, right, up, down, left-right, up-down, clockwise, counterclockwise |
|----------------|------------------------------------------------------------------|
| Full arm       | Left, right, up, down, left-right, up-down, clockwise, counterclockwise |
| Forearm        | Left, right, up, down, left-right, up-down, clockwise, counterclockwise |
| Palm           | Left, right, up, down, left-right, up-down, clockwise, counterclockwise, 90° |
| Finger         | Grip (natural hand shape), inward (natural hand shape), release (fist shaped hand shape), outward (curved hand shape) |

According to information theory, the calculation formula for gesture information is

\[ H(X) = \sum_{i=1}^{m} P_i(X) \log_2 P_i(X) \]  \hspace{1cm} (1)

Where \( H(X) \) is the amount of information contained in gesture \( X \); \( m \) is the total number of states that can occur; \( P_i(X) \) is the probability of occurrence of the \( i \)-th state. Assuming that each state has the same probability, that is, \( P_i(X) = 1/m \), the above formula can be simplified to

\[ H_i(X) = \log_2 m \]  \hspace{1cm} (2)

For example, a common waving goodbye gesture contains \(-\log_2 (1/90 \ast 1/3 \ast 1/26) \) (approximately 12.78bit) information. When both hands participate in gesture movement, the amount of information
contained in both hands needs to be calculated separately and then summed. In addition, it is necessary to consider the mutual constraints between limb shape and movement. For example, when the straight arm is in the forearm, it is almost impossible to perform limb movements such as up and down; when the palm is in a natural state, it is impossible to release and outward limb movements.

2.2. Gesture Design based on Semiotics
The experimental task scene was set up as an operation video, and six experimental gestures were determined: start, pause, fast forward, rewind, return to the previous level, and return to the main page. In the experiment, six participants were invited to design gestures in a user-defined way [10-13]. At the same time, participants were required to report and record their own thinking processes during the design process. After the experiment, the semantic analysis of the user-defined gesture symbols based on the records was performed to establish a knowledge base, and the icon algebra theory [14, 15] was used to redesign the gesture symbols based on the knowledge base. As shown in the figure below, the semantics a of X and the semantics b of Y are combined into new semantics through operators. COM stands for combine, CON stands for context, ENH stands for enhance, MAR stands for MARK, INV means invert, IDX means index.

![Icon algebra](image)

**Figure 1** Icon algebra.

By redesigning, the experiments show six gestures as shown in Figure 2. The meaning of the "pause" gesture symbol in real life is to remind the following car or the person to pay attention, and it is extended to "pause" in the task scene; the meaning of the "start" gesture symbol in real life is to touch or press a button. The semantic meaning in the task scene is "start"; "fast forward" and "back". The gesture meaning in real life is left and right, which is extended to look at the picture frame in the task scene; "return to the previous level" The meaning of the gesture symbol in real life is erasing, which is extended in the task scene to drag the screen to the upper right corner to hide; the gesture symbol of "Back to the main page" uses the method of figurative body [8], use "fork" Means to exit completely and return to the main page.

![Gesture symbol](image)

**Figure 2** Gesture symbol.
2.3. Experimental Subjects
The experiment invited 10 subjects, including 4 men and 6 women, with an average age of 22.7 years. All subjects were right-handed and in good physical and mental condition.

2.4. Experimental Device
The experiment uses an HP Pavilion laptop with a screen resolution of 1366 * 768. The experiment was performed using E-Prime2.0 software, and the computer's own keyboard was used for the reaction operation. The experimental material is a set of Chinese gesture pictures containing all 11 gesture symbols. The picture is black on a white background and the font size is 32 points. It was generated using Photoshop.

2.5. Experiment Process
Before the experiment, the participants need to fill out a questionnaire containing information such as gender, age, right-handedness, drinking, and illness. All unqualified subjects will be excluded from this experiment.

After completing the questionnaire, participants will be given simple pre-experiment training, including familiarization with the experimental process and the gestures and semantics involved in the experiment. After training, rest for 20 minutes.

Before the training module begins, participants are required to turn off all communication equipment and keep the eyes 50cm away from the center of the screen. At the same time, inform the participants to relax as much as possible and choose a comfortable posture. If the participants are all ready, press the "Q" key to enter the training module. During the training process, the staff will provide all necessary assistance to the subjects, and the staff will not interfere with the subjects after the formal experiment begins. After training, if the participants are familiar with the experimental process, press the "P" key to start the formal experiment, otherwise they can press the "Q" key to train again.

After the start of the formal experiment, a 32-point red "+" symbol will be displayed in the center of the screen to remind the participants to pay attention. The display time is 800ms. Then the clue of the gesture symbol will be displayed in the center of the screen. If the subject recalls the corresponding gesture, they need to make the gesture and press the "J" key as soon as possible, otherwise the clue will automatically disappear after 8 seconds. After pressing the "J" key (or disappearing automatically), the participants need to complete the corresponding NASA-TLX scale, and then press the "F" key to enter the next experiment.
Figure.3 Experimental stimulus flow

3. Experimental Study on the Relationship between Information Volume and Cognitive Load
First, two experts were invited to score the factors affecting cognitive load, and then the analytic hierarchy process was used to obtain the weight coefficients shown in Table 3. The consistency ratio RI of the judgment matrix is 0.074, which is less than 0.10, so the judgment matrix is considered to have satisfactory consistency. Then use this coefficient to calculate the score of the NASA-TLX scale, and get the overall cognitive load.

| Mental demand | Physical demand | Time requirement | Effort level | Frustration |
|---------------|----------------|-----------------|--------------|-------------|
| 0.22314       | 0.148798       | 0.085001        | 0.058405     | 0.484657    |

As shown in Table 4, the Pearson correlation coefficient between the amount of gesture symbol information and cognitive load is 0.506, and the P value is less than the significance level of 0.05, which is significant. In addition, as shown in Table 5, there is also a correlation between the amount of gesture symbol information and the user's reaction time.

| Amount of information | Cognitive load |
|-----------------------|---------------|
| Pearson correlation   | 1             | .506          |
| Sig.                  |               | .000          |
| Number of cases       | 60            | 60            |
| Pearson correlation   | .506**        | 1             |
| Sig.                  | .000          |
| Number of cases       | 60            | 60            |

| Amount of information | Response time |
|-----------------------|---------------|
| Pearson correlation   | 1             | .540**        |
| Sig.                  |               | .000          |
| Number of cases       | 60            | 60            |
| Pearson correlation   | .540**        | 1             |
| Sig.                  | .000          |
| Number of cases       | 60            | 60            |

According to the relevant parameters shown in Table 6, we can establish the following regression equation: Cognitive load of user's gesture symbol = 2.409 + 0.173 * information amount of gesture symbol.

| Model | Unstandardized coefficient | Standardized coefficient | Collinear statistics |
|-------|-----------------------------|--------------------------|----------------------|
|       | B   | Standard error | Beta | t | Saliency | Tolerance | VIF |
| 1     | (Constant) | 2409 | .488 | 4.934 | .000 | 1.000 | 1.000 |
| 2     | Amount of information | 173 | .039 | .508 | 4.471 | .000 | 1.000 | 1.000 |

4. Conclusion
This experiment explores the relationship between the amount of gesture symbol information and the user's cognitive load. The experimental results show that there is a significant correlation between the amount of gesture symbol information and the user's cognitive load and response time. Know the
regression equation between loads. However, because the experimental sample size is too small, the equation is not very predictive, and further experimental data needs to be accumulated in the future. In addition, the experiment only explored the correlation between the amount of gesture symbol information and the user's cognitive load, and did not further investigate causality.

References

[1] Sun Xiaohua, Zhou Bo, Li Tong. Design elements and principles of air gesture interaction [J]. Packaging Engineering, 2015, 36 (08): 10-13.

[2] Jahani H, Kavakli M. Exploring a user-defined gesture vocabulary for descriptive mid-air interactions[J]. Cognition Technology and Work, 2017, 20(8):1-12.

[3] Song Yifan, Zhang Peng, Liu Libo. Human-computer interaction system based on visual gesture recognition [J]. Computer Science, 2019,46 (S2): 570-574.

[4] Tan Zhengyu, Yang Wenling. Study on the design of sound synaesthesia for natural interaction [J]. Packaging Engineering, 2018,39 (08): 68-73.

[5] Yang Xian, Zhang Shuo, He Hanwu, Wu Yueming. Natural Gesture Interaction in a Virtual Experimental Environment [J]. Lab Research and Exploration, 2019,38 (03): 90-94.

[6] Zhao Yiheng. Redefining Symbols and Semiology [J]. International Journalism, 2013,35 (06): 6-14.

[7] Zhang Yi, Qin Jingyan, Li Wei. Research on user behaviour of multi-touch interactive system based on sign language semantics [J]. Packaging Engineering, 2008 (05): 144-145 + 156.

[8] Yin Chao, He Renke. Application of Semiotics in VR Gesture Interaction Design [J]. Packaging Engineering, 2013,34 (22): 13-15.

[9] Zhang Jiaqi, Wen Hao, Zhou Kun, Li Caijie. Research on Gesture Trajectory of Vehicle Gesture Interaction Based on Semiotics [J]. Packaging Engineering, 2019,40 (16): 187-193.

[10] Zhao Xingzhi. The development trend of important new schools of semiotics since the 21st century [J]. Journal of Southwest University for Nationalities (Humanities and Social Sciences), 2019,40 (08): 184-191.

[11] Xiao Huali, Lin Li, Yan Longhua, Chen Zhiang. A tourism souvenir design based on the analysis and evolution of Miao silver ornaments based on design semiotics [J]. Packaging Engineering, 2018,39 (14): 221-226.

[12] Zhu Shangshang, Luo Shijian. Reconstruction of Cultural Elements Based on Design Semiology in Product Design [J]. Journal of Zhejiang University (Engineering Science), 2013,47 (11): 2065-2072.

[13] Zhou Yuxiao, Luo Shijian, Chen Gencai. Icon Design Based on Design Semiology [J]. Journal of Computer-Aided Design & Computer Graphics, 2012,24 (10): 1319-1328.

[14] Lu Nengchao, Cao Yue, Qin Ling, Wu Chaozhong. Study on the Effectiveness of Driving Load Based on Traffic Sign Information [J]. China Journal of Highway and Transport, 2018,31 (08): 165-172.

[15] Hou Wenjun, Wu Chunjing. Research on Intuitive Interaction of Gesture of Smart Watch Based on Data Analysis [J]. Packaging Engineering, 2015,36 (22): 13-16 + 21.