



Study on Spatiotemporal Characteristics of Gaze Gesture Input

Wen-jun Hou^{1,2}, Si-qi Wu^{1,2}(✉), Xiao-lin Chen^{1,2},
and Kai-xiang Chen^{1,2}

¹ School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing 100876, China

1519390005@qq.com

² Beijing Key Laboratory of Network Systems and Network Culture, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. The gaze gesture input has the advantages of high bandwidth, high efficiency, no misoperation and high customization, but it also has the disadvantages of high cognitive load and high fatigue degree. At present, few researches have thoroughly explored the regularity of gaze gesture input, which cause designers lack of design guidance, and it is difficult to reasonably apply gaze gesture input in interaction. This paper focuses on the study of the spatiotemporal features of the gaze gesture input. An experiment was designed so that the input region size, the feedforward type and the input gesture shape were within-subject experimental variables. By analyzing 1200 trajectory data of gaze gestures to exploring the effects of these three variables on the performance of and the subjective satisfaction of the gaze gesture input, the details of the three major characteristics of gaze gesture-based interaction were summarized, which are the input continuity, the implicit interaction and the real-time feedback. These may provide design guidance for gaze gesture interaction designers.

Keywords: Eye movement-based interaction · Gaze gesture input · Spatiotemporal feature · Feedforward

1 Introduction

The explosive growth of total information makes human-computer interaction more frequent and complex. Eye movement-based interaction has attracted more and more attention due to its high bandwidth, continuous input and natural interaction. As early as the early 1990s, the development of eye position input system based on eye tracking system has attracted much attention [1]. The improvement of accuracy and resolution of eye tracking technology makes eye movement-based input possible. Especially with the development of real-time eye tracking measurement technology, eye movement-based interaction has become a useful human-computer interaction mode [2, 3]. Many studies have confirmed that eye movement-based interaction has a faster targeting speed [2–4] than traditional interaction methods (such as mouse). Moreover, because it can only use eye movement to interact with machines, even people with limb defects

can easily use it, which also makes the research of eye movement-based interaction have high social value.

The essence of eye movement-based interaction is to record and recognize the movement mode of human eyes through devices, and take specific movement mode as input signal to control specific tasks. In human-computer interaction, blinking, gazing and saccade scanning are usually used as input signals.

At present, the application and research of eye movement-based interaction mainly focus on blink input [5, 6] and gaze input [7–9]. These two input modes require high spatiotemporal characteristics of user interface, and often accompanied by low efficiency, narrow bandwidth, easy misoperation and other usability problems. This also makes the related research of eye movement interaction stagnate, and fails to get real application and promotion from the lab to the market. The definition of gaze gesture comes from the direction or amplitude of the saccade, so the requirement for the spatiotemporal characteristics of the user interface is very low, which is not easy to cause Midas contact problem, and has the advantages of high bandwidth, high efficiency and high customization. However, gaze gesture input also has the remarkable characteristics of high cognitive load. If we can overcome these shortcomings, give full play to the advantages of gaze gesture input, and apply it reasonably to design, it will bring high application value and release the potential of eye movement-based interaction. Table 1 shows the difference between blink input, gaze input and gaze gesture input.

Table 1. Comparison of three input mechanisms for eye movement-based interaction

| | Blink input | Gaze input | Gaze gesture input |
|---|---|-----------------------------------|--|
| Parameter | Blink duration/Blink frequency | Fixation duration/Fixation field | Saccade length/Saccade duration/Saccade velocity |
| Bandwidth | Low | High | Lowest |
| Efficiency | Fastest | Slow | Fast |
| Requirements for spatiotemporal characteristics of interfaces | High | Very high | Low |
| Naturalness | Quite natural | Natural | Not very natural |
| “Midas contact” problem | Appears often | Appears very often | Appears rarely |
| Application status | A long history, obvious shortcomings and few applications | A long history, many applications | A short history, few applications |

However, the research on gaze gesture input is relatively less, and mainly focuses on the study of simple gaze gesture input symbols. The performance of experimental

task completion is almost the only research indicator. No one has ever studied the appropriate input area size for gaze gesture, and no one has ever set the experimental conditions for gaze gesture input in the eye-friendly zone (ignoring the ergonomics to develop input performance). Moreover, few people have studied the effect of feed-forward types on gaze gesture input. To sum up, there are many factors that affect the performance of gaze gesture input, but many previous studies are too narrow to have broad practical significance.

In this paper, the spatiotemporal characteristics of gaze gesture input were focused. The effects of input area size, feedforward type and input gesture shape on the performance of gaze gesture input and the subjective satisfaction of users were synthetically explored. Then, the performance of gaze gesture input in the input continuity, the implicit interaction and the real-time feedback were summarized.

2 Related Work

Eye movement-based interaction takes vision as input channel and has the potential to replace traditional mouse-like pointing devices (such as mouse, stylus, finger touch screen). The difference in application between eye movement and mouse-like pointing devices makes the spatiotemporal characteristics of eye movement-based interaction significantly different from traditional ones. The time characteristic in eye movement-based interaction refers to the time threshold (residence time, blink time, etc.) required to trigger an interface instruction using eye movement data (including eye gaze, blink, eye gesture, etc.). The spatial characteristics in eye movement-based interaction refer to the spatial accuracy needed to select the interactive objects on the interface based on eye-movement interaction (including the size of the interactive objects, the distance between objects, the arrangement of objects, etc.).

In the study of spatiotemporal characteristics of eye movement-based interaction, Feng (2006) et al. found that for human-computer interface based on eye-movement gaze input technology, the horizontal arrangement of objects was significantly better than the vertical arrangement [10]. Zhu (2014) et al. found that in the touch screen interaction system, by observing the eye movement data, the user's visual focus in the operation process is mainly concentrated on the position of the object and the target, while only a small amount of attention is paid to the process, so there is enough data and time for eye movement behavior to be an input mechanism [11].

In addition to spatiotemporal characteristics, the study of feedback in eye movement-based interaction can also help to improve the availability of eye movement-based interaction. Feng (2004) et al. found that introducing visual feedback in eye movement-based interaction can improve the efficiency of searching, locating and activating the target objects on the interface, while introducing visual display of the border of the interactive objects in eye movement-based interaction interface has no significant impact on user's performance [12]. Zhu (2014) found that the availability of the touch screen interactive system has been significantly improved when the eye movement-based assisted operation was introduced. Moreover, if appropriate interference indication is introduced into the system, it will help users to establish target objects faster [11].

The above studies mainly focus on gaze gesture input. Gaze gesture input, also known as saccade input, is of great research value because of its advantages of fast speed, low requirement for spatiotemporal characteristics of the interface and difficulty in misoperation. Firstly, the fastest speed of saccade up to 400° – 600° per second which means that gaze gesture input can reach 1° to 40° viewing angle within 30–120 ms which is much faster than a standard gaze input unit time 300–500 ms. Secondly, as gaze gesture input does not require specific interactive controls and elements on the interface, the interactive time is also relatively high robustness and it does not necessarily require an accurate response time, interface design will be easier and faster because of the low requirements of the spatiotemporal characteristics of the interface. Thirdly, since gaze gesture input is sequence-based and does not require a precise starting point and ending point, it is naturally more advantageous for Midas contact than blinking and gazing [12].

There are relatively few studies on spatiotemporal characteristics of gaze gesture input, and there is no decisive conclusion. Xin (2015) et al. studied the efficiency of gaze gesture and only analyzed the input performance of subjective defined short-range and long-range gaze gesture [13]. Møllenbach et al. found that short-range gaze gesture trajectories triggered faster than long-range gaze gesture trajectories, and the speed of horizontal trajectories performed better than vertical trajectories' [14]. However, there are some inaccuracies in this experiment. The experiment was carried out on a rectangular screen. In order to control the distance of gaze gesture in horizontal and vertical directions equally, the size of the trigger area of gaze gesture in these two directions is different, which may affect the validity of the experimental results. The flaws in this experimental design can be supported to a certain extent in the study by Heikkilä (2012) et al. In their study, similar experimental content has reached the opposite conclusion that gaze gesture moves faster in the vertical direction than in the horizontal direction [15]. Heikkilä et al. also found that there was no significant difference in time between short-range and long-range gaze gesture when eye movements were performed using closed eyes to end eye movement-based control. They think it is possible that in their experimental design, users only need to move their eyes in the right direction without precise eye movement control, which makes the conclusion different from that of Møllenbach [16].

In fact, the gaze gesture input mechanism relies on saccade, and the speed of saccade is extremely fast, the fastest speed can reach $400^{\circ}/s$ – $600^{\circ}/s$. Therefore, if there is no significant difference in length between the same type of gaze gesture, performance should not vary too much. In ergonomics, there is a comfort zone when the human eye rotates. When rotating in the comfort zone, the muscle burden is small, it is not easy to fatigue, and the speed of eye movement is fast, and vice versa. However, previous researchers often neglected this point when they studied the performance of gaze gesture. We can assume that the reason for the significant difference in the performance of long-range and short-range gaze gesture in some experiments may not be simply because of the difference in the length of gaze gesture, but because long-range gaze gesture may have exceeded the comfort zone of gaze while short-range gaze gesture not. As shown in Fig. 1, there is a rotational comfort zone in eye rotation. The optimum upper and lower regions: upper 25° + lower 30° = 55° . The best left and right regions : left 15° + right 15° = 30° .

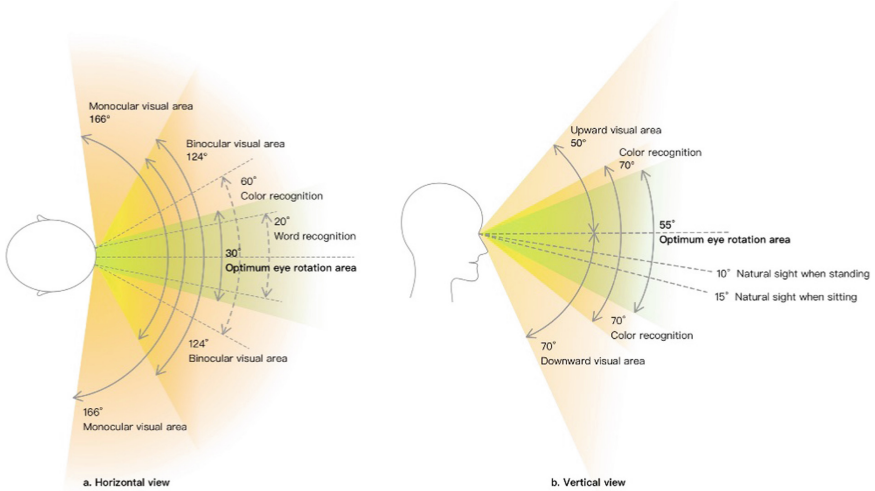


Fig. 1. Rotational comfort zones in eye rotation. a. horizontal view b. vertical view

Gaze gesture input has not been fully applied yet because of its obvious disadvantages. How to design gaze gestures as input symbols is a very difficult research topic. If the gaze gesture is too simple, it is easy to overlap with the unconscious eye movement, leading to misoperation; but if it is too complex, it will increase the user's learning cost, memory burden and cognitive load, which is contrary to the original intention of natural interaction. Previous studies have mostly focused on simple gaze gestures (single-step long gaze gestures), which are suitable for completing simple human-computer interaction tasks. Although multi-step gaze gesture (multi-step or radian gaze gesture) is difficult for users to learn and has a high degree of input fatigue, it is more suitable for completing complex human-computer interaction tasks because of its higher bandwidth. Istance et al. designed a series of composite gaze gestures for World of Warcraft games. Experiments have shown that the use of composite gaze gestures can interact more accurately than a single gaze gesture, but it will occupy a large number of cognitive channels [17]. Feedforward (giving guidance before the user performs the operation) as a special form of feedback, can reduce the user's learning burden and operating burden, so it is worth considering.

In summary, this article conducted research on the following four aspects:

- (1) Gaze gesture input area size was taken as the control variable to study its influence on input performance.
- (2) Input gesture type was taken as the control variable to study its influence on input performance.
- (3) Feedforward type was taken as the control variable to study its influence on input performance.
- (4) A multi-factor analysis of the above three factors and input performance was conducted to investigate whether there is a cross-effect effect.

3 Experiment

3.1 Participants

20 participants (10 males and 10 females) were voluntarily recruited, ranging in age from 20 to 25 years old (mean = 22.2 years old and SD = 1.67 years). All of the participants had visual acuity or corrected visual acuity of 5.0 or above and the corrected visual acuity of the participants was all within 200° , which ensured that the lenses they wore would not be too thick to affect the detection of eye tracker. All participants successfully passed the eye tracker calibration test, and the calibration accuracy in the X and Y directions was less than 0.5° .

3.2 Device

The eye tracker used in the experiment was the German SMI iView RED non-contact eye tracker with a sampling frequency of up to 500 Hz, a tracking resolution of 0.1° , and a gaze positioning accuracy of 0.5° – 1° . The dedicated display for the eye tracker was a DELL 22-inch display. The physical size of the electronic screen was 475 mm long, 298 mm wide, and the resolution was $1680\text{pi} * 1050\text{pi}$, with a ratio of 16:10. In terms of software, the software used to control the eye tracker device was iView X version and the analysis software was BeGaze. Experimental staff for the experimental included an operator, a recorded and a host.

3.3 Selection of Tasks

Firstly, the experimental level of input region size variables was determined. The relationship of human visual angle range is shown in Fig. 2 where α is the horizontal view size, β is the vertical view size, L is the horizontal view range length, W is the vertical view range length, and o is the center point of the screen. When the human eye's line of sight is perpendicular to o point in the center of the screen, H is called the line of sight distance from the human eye to the screen.

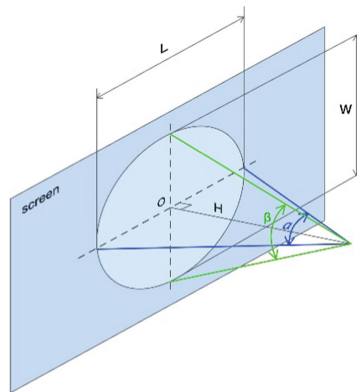


Fig. 2. View angle size and target size diagram

The relationship between the size of the perspective, the range of the perspective, and the line-of-sight distance is shown in formula 1 and formula 2.

$$\alpha = 2\arctg \frac{L}{2H} \tag{1}$$

$$L = 2H\arctg \frac{\alpha}{2} \tag{2}$$

According to the physical size and pixel size of the screen used for the observed materials in the experimental equipment, 10 levels of experiments as shown in Table 2 have been defined (the data have been rounded). The determination of the maximum viewing angle depends on the maximum height of the experimental screen, that is 298 mm. On the basis of 25°, the level of other variables is determined by decreasing the angle of view by 5°. When the angle is less than 5°, the decrease is changed to 1°. All horizontal viewing angles, both left and right rotation and up and down rotation, are in the comfortable region of rotation, as detailed in Table 2.

Table 2. Ten levels of the input area

| Pixel (px) | View angel (°) | Physical dimension (mm) | Other instructions |
|--------------|----------------|-------------------------|---|
| 1050 * 1050 | 27.89 | 298 | The best eye movement area (within 30°) |
| 937 * 937.37 | 25 | 266.03 | The best eye movement area (within 30°) |
| 745 * 745.54 | 20 | 211.59 | The best eye movement area (within 30°) |
| 556 * 556.65 | 15 | 157.98 | The best eye movement area (within 30°) |
| 369 * 369.92 | 10 | 104.99 | The best eye movement area (within 30°) |
| 184 * 184.61 | 5 | 52.39 | Parafoveal (2°–5°) |
| 147 * 147.65 | 4 | 41.90 | Parafoveal (2°–5°) |
| 110 * 110.72 | 3 | 31.42 | Parafoveal (2°–5°) |
| 73 * 73.80 | 2 | 20.95 | Foreal area (1°–2°) |
| 36 * 36.90 | 1 | 10.47 | Foreal area (1°–2°) |

Secondly, the experimental level of input gesture type variables was determined. Previous studies have been divorced from reality, and most of them were researches of single-step gaze gesture, and did not depend on specific application scenarios. The experimental level of the input symbol type variable in this experiment was derived from the action in the consensus set obtained in the previous experiment, from which the two representative inputs of square and circle were selected.

Finally, the experimental level of input feedforward type variables was determined. Considering the attention allocation mechanism of human eye movement, the experimental level was designed as follows: line-like feedforward, no feedforward and point-like feedforward.

In summary, there were three control variables in the experiment, including 10 levels, 2 levels and 3 levels, totaling 60 input tasks. In order to facilitate the experimental operation, during the actual experiment, the tasks were divided into 6 groups, as shown in Table 3.

Table 3. Details of 60 tasks in 6 groups

| Group | Input gesture type | Feedforward type | Gaze gesture input area size |
|--------|--------------------|------------------------|---|
| Group1 | Square | Line-like feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |
| Group2 | Square | No feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |
| Group3 | Square | Point-like feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |
| Group4 | Circle | Line-like feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |
| Group5 | Circle | No feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |
| Group6 | Circle | Point-like feedforward | 28°–1° (All ten levels:28°, 25°, 20°, 15°, 10°, 5°, 4°, 3°, 2°, 1°) |

According to the experiment task, 60 stimulating materials were developed. The stimulating materials were on the grey background, the input area was white background, the feed-forward reminder was located in the white background, and the inner margin was controlled between 0.5°–1°.

3.4 Procedure

Firstly, the participants adjusted their sitting posture to maintain a relatively comfortable sitting posture, with their eyes just perpendicular to the center of the display screen, and their heads aligned and fixed. The distance between their eyes and the screen was measured and determined to be about 600 mm. Specific experimental scenario is shown in Fig. 3.

The experiment was introduced to the participants before the experiment was officially started. A training opportunity was provided before each group of experiments begins. Between groups of experiments, three minutes' rest time was provided, and within each group, 20 s' rest time was provided.

After all the experiments were completed, participants were asked to fill in questionnaires to evaluate the overall satisfaction of each task, that was, "I feel satisfied with the overall performance of this gaze gesture input". Satisfaction refers to whether users feel it is easy, accurate and fast to complete the task with gaze gesture input. The design



Fig. 3. Experimental scene

of the questionnaire was based on the Likert 10-point scale, with 10 representing “very satisfied” and 1 representing “very dissatisfied”. The Likert 10-point scale can increase the rating, and the participants are more likely to give a score and improve the discrimination of the results [18], which is suitable for experiments with up to 60 samples in this evaluation. The whole experiment lasted about 60 min.

4 Results

The results of the experiment totaled 1200 input trajectories, of which 3 participants made obvious errors in the trajectory results of their individual tasks. There were two types of errors: one was systematic errors, that was, the device did not accurately record the input trajectory; the other was that users were distracted for a long time, and the characteristics of the input trajectory were obviously deviated from the whole. The number of valid input trajectories collected was 1020.

4.1 Qualitative Analysis of Input Trajectory

Rule1: It is most accurate to draw the square with gaze gesture input, when the feedforward form is point-like feedforward.

As shown in Fig. 4, the different feedforward types had very significant differences for the square trajectories ultimately drawn by the participants. The comprehensive performance of point-like feedforward was the best, while that no feedforward was the worst, mainly reflected in two points: first, the positioning of the four key points was not accurate; second, during the scanning process, there were a lot of additional fixation behaviors.

Rule2: The fourth step of drawing a square shows an offset.

When drawing a square, the fourth step of the trajectory generally showed an inclination of about 8° instead of vertical. When drawing a square, the trajectory of the fourth step was inclined about 8° , not vertical. The trajectory of the second step also showed a certain slope, but it was not obvious as the fourth step. The same situation did not occur in the first and third steps.

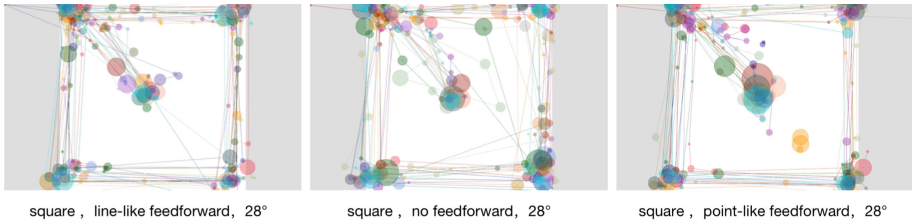


Fig. 4. All trajectories of the square

By analyzing the set of trajectory diagrams, we can find that there were two cases of b and c in Fig. 5b showed that when the third step was drawn, the position of the scanning end point exceeded the target position, and the participant added a fixation point to modify the figure in the fourth step. In the case of c, when the third step was drawn, the scanning end point positioning failed to reach the target position, and the participant added a fixation point for correction, but often the corrected point exceeded the target position. As a result, most of the graphs obtained in the final drawing present the situation of the fourth inclined step.

In conclusion, it can be speculated that the positioning accuracy of horizontal saccade is less than that of vertical saccade, while the positioning accuracy of left-to-right saccade is better than that of right-to-left saccade.

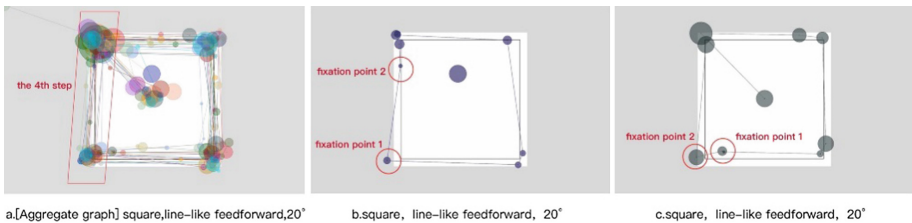


Fig. 5. An illustration of the fourth step of drawing a square to show the offset

Rule3: Point-like feedforward has a negative effect on the accuracy of drawing a circle.

Different feedforward types had a significant impact on the final circular trajectory. The performance of line-like feedforward was the best, while that of no feedforward was the worst. The fixation points of the circular trajectory drawn with line-like feedforward were evenly distributed on the circular border, compared with the figure drawn without feedforward, which deviates a lot from the circle. The result of point-like feedforward was like a diamond. After the experiment was completed, many of the participants bluntly stated that they needed to allocate extra attention to prevent them from simply connecting the points with straight lines, as shown in Fig. 6.

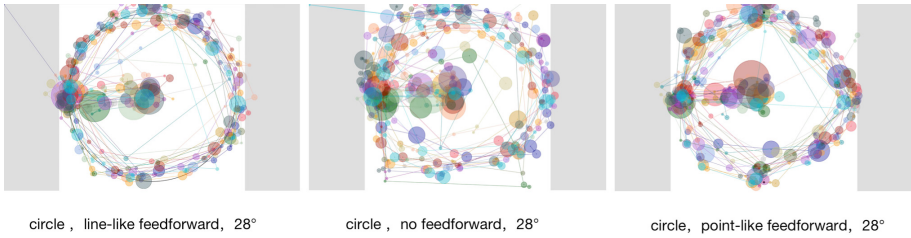


Fig. 6. All trajectories of the circle

Rule4: The participants automatically add key points to correct the drawing of the graph.

In Rule 3, point-like feedforward was counterproductive for guiding the user to draw an accurate circle. If we look at each participant's drawing process, as shown in Fig. 7, we can find that most of the participants unconsciously completed the presentation with one saccade, but also subconsciously added a fixation point between the two key points to correct their drawing.

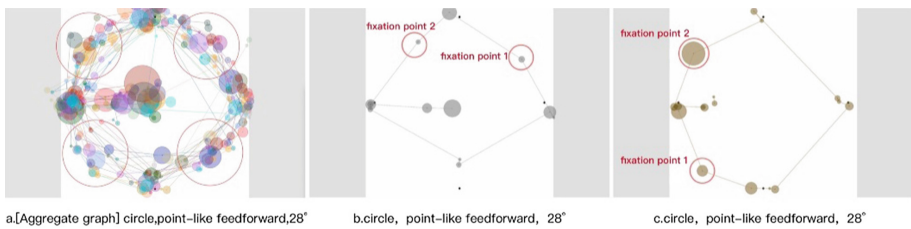


Fig. 7. An illustration that participants automatically add fixation points to correct their drawing

Rule5: The accuracy of symbol rendering decreases significantly when the input area is less than 5° .

The experimental results showed that when the input area decreased to less than 5° , the accuracy of drawing symbols decreased significantly, and the accuracy of drawing square was better than that of drawing circle. Usually when the input area dropped to 3° , it was very difficult to distinguish the approximate figure from the trajectory (as shown in Fig. 8c and d). Of course, a small number of excellent participants also showed good accuracy when using gaze gestures to draw in small areas (as shown in Fig. 8c and d). Combined with the results of the loud thinking of the participants, when the input area was less than 3° , the participants felt that the input experience was extremely poor. First, it was necessary to concentrate a large amount of attention to control the unconscious shaking of the eyeball. Second, without feedback, participants felt frustrated if the drawing is inaccurate.

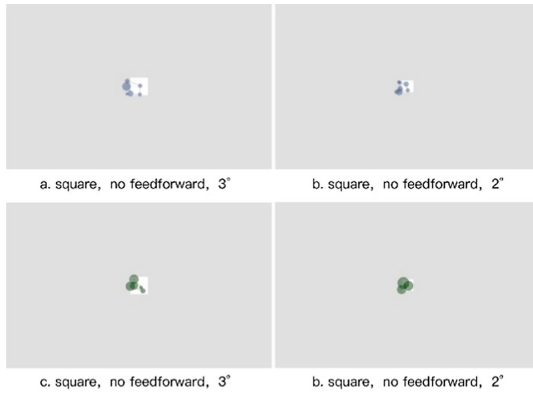


Fig. 8. Trajectory with too small input area

4.2 Quantitative Analysis of Input Trajectory

1020 valid sample data collected by eye movement equipment were quantitatively analyzed for input performance. Table 4 shows the average time-consuming statistics of 60 input tasks (accurate to milliseconds).

Table 4. Average time-consuming of 60 gaze gesture input tasks

| (ms) | Square + line-like feedforward | Square + no feedforward | Square + point-like feedforward | Circle + line-like feedforward | Circle + no feedforward | Circle + point-like feedforward |
|------|--------------------------------|-------------------------|---------------------------------|--------------------------------|-------------------------|---------------------------------|
| 28° | 3716 | 3905 | 4109 | 4849 | 5590 | 5343 |
| 25° | 3656 | 3958 | 3895 | 4860 | 5294 | 5395 |
| 20° | 3577 | 3912 | 3895 | 4654 | 5324 | 5594 |
| 15° | 3487 | 4188 | 3586 | 4022 | 4884 | 5236 |
| 10° | 3105 | 3950 | 3642 | 3604 | 4662 | 4649 |
| 5° | 3085 | 3808 | 3335 | 3461 | 4114 | 4270 |
| 4° | 3066 | 3771 | 3272 | 3396 | 3923 | 4042 |
| 3° | 3029 | 3614 | 3206 | 3216 | 3956 | 3820 |
| 2° | 2870 | 3524 | 3016 | 3082 | 3759 | 3322 |
| 1° | 2838 | 3410 | 2831 | 2785 | 3443 | 3070 |

Figure 9 was obtained by visualizing the above table data. From the results, as the input area became smaller, the overall input performance increases; the overall performance of the square input was better than the circular input; the performance of the line-like feedforward was the best. Simply analyzing the impact of each factor, we found that the three factors had significant impact on the input time (the P values of the three factors are less than 0.001).

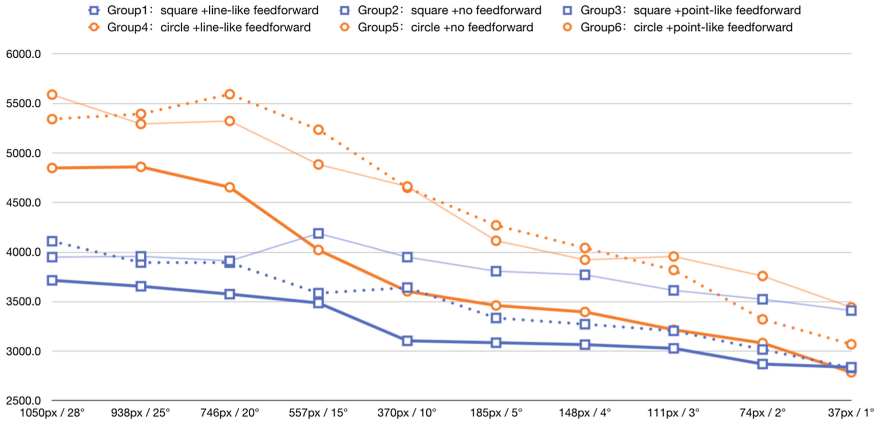


Fig. 9. Mean input time of six groups of tasks with different input area sizes

However, the eye movement-based speed (average eye movement angle per second) continued to decline as shown in Fig. 10. Eye movement-based speed reflects the degree of fatigue to a certain extent. The faster the eye movement-based speed is, the less effort the user pays to concentrate on the task, so user’s fatigue is lower. When input area size was larger than 15°, the eye movement speed of the user was faster and would not cause obvious fatigue.

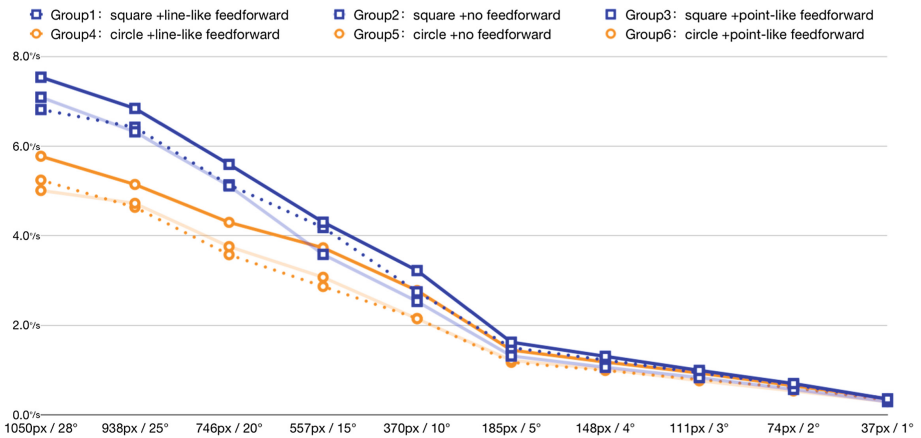


Fig. 10. Eye movement-based speed

In addition, it can be seen from the figure that each group of data had obvious variation rules and cross effects. In order to verify the above judgment, the data results were analyzed by univariate multivariate analysis of variance. The results are shown in Fig. 11.

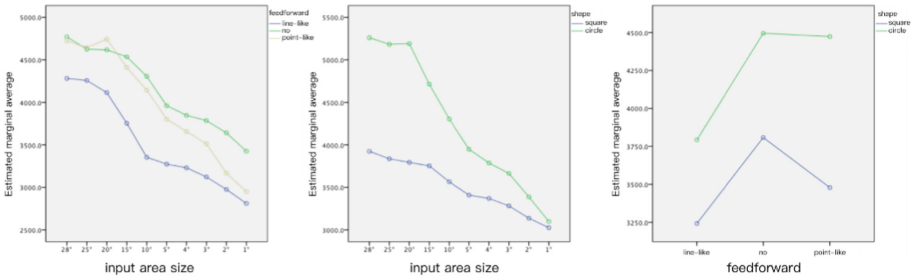


Fig. 11. Multivariate cross-effects of input time (a.size & feedforward b.size & shape c. feedforward & shape)

4.3 Change Law of the Number of Fixation Points

The variation of the number of fixation points was studied. Table 5 shows the statistics of the average number of fixation points for 60 input tasks.

Table 5. The average number of fixation points for 60 input tasks

| (number) | Square + line-like feedforward | Square + no feedforward | Square + point-like feedforward | Circle + line-like feedforward | Circle + no feedforward | Circle + point-like feedforward |
|----------|--------------------------------|-------------------------|---------------------------------|--------------------------------|-------------------------|---------------------------------|
| 28° | 13.1 | 11.9 | 13.4 | 16.8 | 18.9 | 18.4 |
| 25° | 11.2 | 12.4 | 11.6 | 16.9 | 17.6 | 18.2 |
| 20° | 10.5 | 12.5 | 11.3 | 16.1 | 16.8 | 18.1 |
| 15° | 10.6 | 11.8 | 10.1 | 13.6 | 15.3 | 16.2 |
| 10° | 9 | 14.1 | 9.4 | 11.2 | 12.5 | 12.4 |
| 5° | 7.3 | 8.6 | 8.1 | 8.9 | 9.4 | 9.2 |
| 4° | 7.1 | 8.1 | 7.5 | 7.2 | 8.2 | 7.8 |
| 3° | 6.1 | 7 | 6.6 | 6.8 | 7.2 | 6.5 |
| 2° | 3.1 | 5.3 | 3.8 | 4.8 | 6.2 | 4.4 |
| 1° | 2.8 | 4.1 | 2.9 | 4.8 | 4.9 | 3.8 |

The above table data was visualized and shown in Fig. 12. From the results, as the input area became smaller, the total number of fixation points decreased; the number of fixation points used for square input was generally less than the circular input, but this situation became inconspicuous when the input area began to be less than 5°. Simply analyzing the impact of each factor, we found that all three factors had a significant impact on the number of input fixation points (the P values of all three are less than 0.05).

In addition, it can be found from the figure that each group of data had obvious variation rules and cross effects. In order to verify the above judgments, the data results were analyzed by univariate multivariate analysis of variance. The results are shown in Fig. 13.

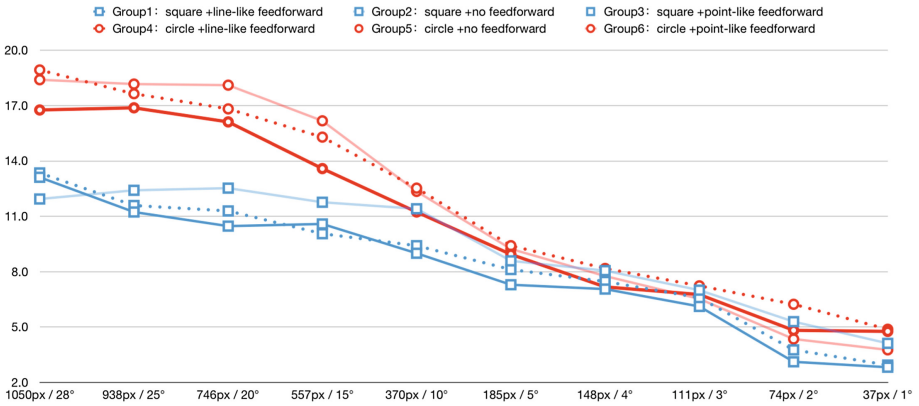


Fig. 12. The mean number of fixation points for six tasks with different input sizes

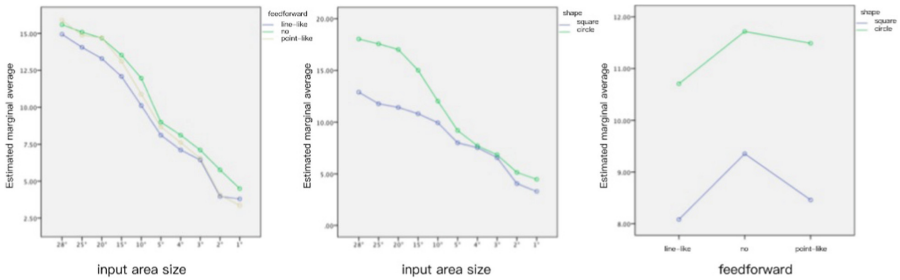


Fig. 13. Multivariate cross-effects of fixation points (a.size &feedforward b.size &shape c. feedforward &shape)

The results of multivariate cross-analysis are as follows:

Firstly, there was no significant cross-effect between feedforward and input area size ($p = 0.974 > 0.05$). In general, no matter what kind of feedforward, their trend of the number of fixation points changed consistently, which decreased with the decrease of the size of the input area.

Secondly, there was a significant cross-effect between shape and size of input area ($p = 0.000 < 0.001$). When the input area size was greater than 5° , the number of square input fixation points was less than that of the circular input; when the input area started to be less than 5° , the difference of the input symbols had little effect on the number of fixation points.

Thirdly, there was no significant cross-effect between feedforward and shape ($p = 0.405 > 0.05$). Generally speaking, both line-like feedforward and point-like feedforward could reduce the number of fixation points used in input, which was not affected by the type of input shape.

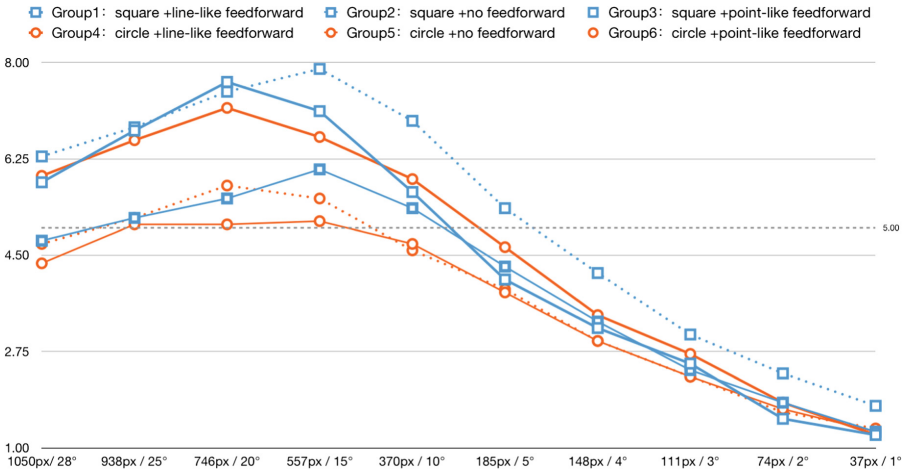


Fig. 14. Mean satisfaction of six tasks with different input sizes

Finally, from the results of descriptive statistics and cross-analysis, the effects of the three factors on the input duration and the number of input fixation points were very close, so Pearson correlation analysis of the two factors showed that they were positively correlated, and the correlation coefficient was 0.0651 ($p < 0.001$), that was, input performance had a strong positive correlation with the number of fixation points, and the more input fixation points were, the longer it took.

4.4 Subjective Satisfaction Evaluation

The data of 17 users collected were summarized and counted to obtain the comparison chart of the average satisfaction in Fig. 14. It was found that the overall satisfaction was higher when the input area was between 10°–20°. When the input area began to be less than 5°, the satisfaction decreased rapidly. The overall satisfaction of square input was better than that of circular input, but this phenomenon was not obvious when the input area was less than 5°. Point-like feedforward improved the satisfaction of square input but reduced the satisfaction of circular input. Simply analyzing the impact of each factor, we found that all three factors had a significant impact on satisfaction (the P values of all three are less than 0.001).

In addition, it can be found from the figure that each group of data had obvious variation rules and cross effects. In order to verify the above judgments, the data results were analyzed by univariate multivariate analysis of variance. The results are shown in Fig. 15.

Firstly, feedforward and input area size had significant cross-effects ($p = 0.009 < 0.001$). When the size of the input area was larger than 10°, the satisfaction when there was feedforward was much higher than that without feedforward, and the satisfaction of point-like feedforward was higher than that of line-like feedforward. When the input area began to be less than 10°, the existence of feedforward had little effect on the input satisfaction.

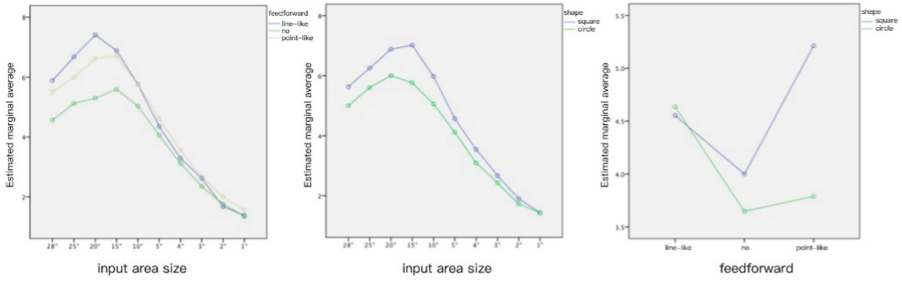


Fig. 15. Multivariate cross-effects of satisfaction (a.size & feedforward b.size & shape c. feedforward & shape)

Secondly, there was no significant cross-impact effect between shape and input area size ($p = 0.15 > 0.05$). In general, whether the input is square or circular, their satisfaction trends are consistent, and they all change with the size of the input area.

Thirdly, the feedforward and shape factors had significant cross-effects ($P = 0.000 < 0.001$). When there was line-like feedforward, the satisfaction of circle and square had little difference; when there was no feedforward, the satisfaction of square was obviously higher than that of circle; when there was point-like feedforward, the satisfaction of square was much higher than that of circle. In general, feedforward can improve the user's satisfaction at gaze gesture input, but different feedforward types have different help for different graphics. For example, point-like feedforward plays a great role in promoting input square, but it has very limited help for input circle.

5 Conclusion

In this paper, the spatiotemporal characteristics of gaze gesture input were focused. The conclusions of this study include the following aspects:

In terms of real-time feedback:

- (1) Feedforward is very important, which can reduce the number of fixation points and improve performance.
- (2) Drawing straight lines with point-like feedforward is more accurate, while drawing curves with line-like feedback has better accuracy.
- (3) When the user's eye movements have sight deviates or the number of points in the point-like feedforward is insufficient to describe feature of the gaze gesture, the user will automatically correct the data by adding fixation points.

In terms of implicit interaction:

- (1) When user performs gaze gesture, it is easier to lead to visual fatigue which tend to cause fuzzy input which has low accuracy.
- (2) When the input area is smaller than 5° , the drawing accuracy of gaze gesture is degraded, and the satisfaction is very low.

- (3) When the feedforward information is insufficient, the user's input will be ambiguous.

In terms of input continuity:

- (1) In the gaze gesture interaction, the positioning accuracy of horizontal saccade is less than vertical saccade, and the positioning accuracy of left-to-right saccade is better than right-to-left saccade.
- (2) When the input area is larger than 5° , the performance and satisfaction of square are generally better than circle.
- (3) With the decrease of input area size, the number of fixation points decreased significantly, and the performance of gaze gesture interaction also increased significantly. However, the eye movement-based speed (average eye movement angle per second) continued to decline. Eye movement-based speed reflects the degree of fatigue to a certain extent. The faster the eye movement-based speed is, the less effort the user pays to concentrate on the task, so user's fatigue is lower. When input area size is larger than 15° , the eye movement speed of the user is faster and will not cause obvious fatigue. Satisfaction is higher when the gaze gesture interaction area is 10° – 28° .

According to the research on the spatiotemporal characteristics of gaze gesture interaction, the following reference points in the gaze gesture interaction design are obtained:

- (1) Considering the accuracy of gaze gesture, the vertical saccade accuracy is better than the horizontal saccade, and the positioning accuracy of left-to-right saccade is better than right-to-left saccade.
- (2) Gestures consisting of only straight lines should be adopted as far as possible in the design of gaze gesture, and feedforward with sufficient information should be used to guide users.
- (3) When the gestures consisting of curves have to be drawn, line-like feedforward or point-like feedforward with sufficient information should be used to reduce the cognitive burden of user.
- (4) It is not advisable to set the eye gesture interaction area smaller than 5° , and the gaze gesture interaction task should be designed as far as possible between 10° and 28° .

Through interviews with the participants, one of the reasons for the influence of the gaze gesture input performance was that once the user is psychologically aware that he was using the gaze gesture to input, the spirit would be tightened. In order to pursue the accuracy of the input, the user's eye fatigue would increase dramatically, and the psychological pressure would double, which in turn would affect the performance of continuous gaze gesture input.

According to the experimental results, the relationship between the degree of eye load and the degree of attention distribution is roughly described as shown in Fig. 16. The Y-axis represents load, including physiological load and psychological load. When the user is in unconscious eye movement, the load is very low, and is not affected by the complexity of eye movement. When the user is in the subconscious eye movement

area, the load is affected by the complexity of the eye movement. The X-axis represents attention resources. When users realize that they are using gaze gesture to input, they begin to allocate attention resources, and the allocation is affected by the complexity of the eye movement. The unconscious eye movement area, which is the user's natural eye movement area, is usually used as the data source for the implicit input of eye movement in human-computer interaction. Subconscious eye movement area is usually used as a data source for explicit input in human-computer interaction.

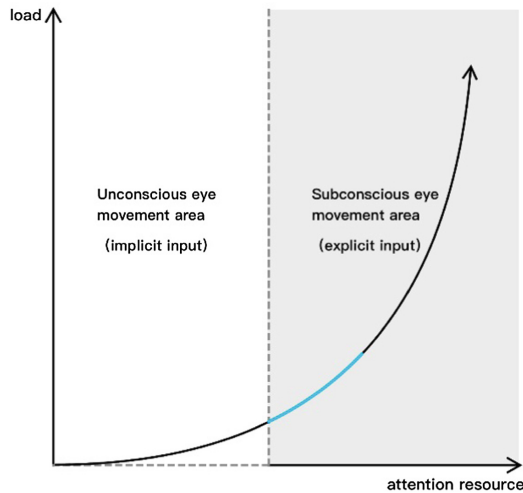


Fig. 16. Relationship between attention distribution of gaze gesture input and load

Not all eye movements are suitable as input instructions. The blue line in Fig. 16 is used to represent the appropriate input state. In the specific application process, we should follow the principle of attention distribution and reasonably consider the complexity of input symbols.

There is a long way to go in gaze gesture input and eye movement interaction interface research. The work done in this paper is only a little bit. Specifically, inspired by the research results of this paper, there are already several clear directions, which are worthy of further research:

- (1) The input performance can be studied when the input area is larger than 30° (outside the comfort zone of the eyeball rotation).
- (2) It is possible to explore whether there is a better performance feedforward type for input symbols such as circles.
- (3) It is possible to define the “complexity” metric for eye movements and the range of complexity metrics that are appropriate for eye movement explicit input.

References

1. Farwell, L.A., Donchin, E.: Taking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.* **70**(6), 510–523 (1988)
2. Sibert, L.E., Jacob, R.J.K.: Evaluation of eye gaze interaction. In: *Proceedings of CHI2000*, The Hague, The Netherlands, 281–288 (2000)
3. Murata, A.: Eye-gaze input versus mouse: cursor control as a function of age. *Int. J. Human-Computer Interact.* **21**(1), 1–14 (2006)
4. Zhai, S., Morimoto, C., Ihde, S.: Manual and gaze input cascaded (MAGIC) pointing. In: *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, pp. 246–253, New York, USA (1999)
5. Grauman, K., Betke, M., Gips, J., et al.: Communication via eye blinks - detection and duration analysis in real time. In: *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, Kauai, HI, USA (2001)
6. Kaufman, A.E., Bandopadhyay, A., Shaviv, B.D.: An eye tracking computer user interface. In: *Proceedings of Research Frontiers in Virtual Reality Workshop*, pp. 120–121. IEEE Computer Society Press (1993)
7. Jacob, R.J.K.: What you look at is what you get: eye movement- based interaction technique. In: *Proceedings of Conference on Human Factors in Computing Systems*, pp. 11–18 (1990)
8. Hyrskykari, A.: Gaze control as an input device. In: *Proceeding of ACHCI 1997*, University of Tampere (1997)
9. Corno, F., Farinetti, L., Signorile, I.: An eye-gaze input device for people with severe motor disabilities (2002)
10. Cheng-Zhi, F., Mo-Wei, S.: Task efficiency of different arrangements of objects in an eye-movement based user interface. *Acta Psychol. Sin.* **04**, 515–522 (2006). (in Chinese)
11. Zhu, H.-J.: Research on eye movement control of mobile terminal system and prototype design of interaction framework. Master thesis, Shanghai Jiao Tong University (2014) (in Chinese)
12. Feng, C.-Z., Shen, M.-W., Chen, S., Su, H.: Influence of border and eye cursor on task performance in eye movement interaction, *Chinese J. Appl. Psychol.* **01**, 9–12+17 (2004) (in Chinese)
13. Xin, Y.-X.: Design and implementation of text input system based on eye movement. Master thesis, Nanjing University (2015) (in Chinese)
14. Møllénbach, E., Lillholm, M., Gail, A., et al.: Single gaze gestures. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 177–180 (2010)
15. Heikkilä, H.: Simple gaze gestures and the closure of the eyes as an interaction technique. In: *Symposium on Eye Tracking Research and Applications*, vol. 85, pp. 147–154 (2012)
16. Mollenbach, E., Hansen, J.P., Lillholm, M., et al.: Single stroke gaze gestures. *CHI 2009 Extended Abstracts on Human Factors in Computing Systems*. ACM, pp. 4555–4560 (2009)
17. Istance, H., Hyrskykari, A., Immonen, L., et al.: Designing gaze gestures for gaming: an investigation of performance. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 323–330 (2010)
18. Yu-Hui, L., Bei-Ping, T., Yun, W., Liang, S.: Comparative study on variant degree likert scale on satisfaction research. *J. Data Anal.* **1**(2), 159–173 (2006). (in Chinese)