Investigating the Effect of Targets' Spatial Distribution on the Performance of Gesture Interaction in Virtual Reality Environment

Kai Chen, Xiaozhou Zhou, and Chengqi Xue

School of Mechanical Engineering, Southeast University
211189 Nanjing, China

ABSTRACT

In virtual reality(VR) environment, how to accurately and effectively select target objects is an important part of interactive tasks such as virtual simulation and virtual assembly. Gesture is one of the most crucial interaction technologies in human-computer interaction. This study conducted a multi-factor experiment in VR environment to explore the effects on the performance of pointing task under different depths and different perspectives. Participants need to point to the target objects in different viewing angle areas at different depth levels with HTC-Vive headset and Noitom HI5 gloves. The results show that the higher the depth level and the closer the target is to the visual center, the higher the pointing accuracy, and the pointing deviation is smaller on the dominant hand side than the non-dominant hand side. The research results have reference value for the spatial distribution of target objects for

gesture interactive tasks in VR environment.

Keywords: Human-computer interaction, Gesture interaction, Spatial distribution, Virtual reality

INTRODUCTION

Human-computer interaction technology refers to the technology that enables people to communicate effectively with machines through devices that can be used for interaction (Hoc, 2000). Gesture interaction is one of the most widely used human-computer interaction methods in the three-dimensional environment, especially the virtual reality(VR) environment. According to whether the device is in contact with the body, it can be divided into contact-based gesture interaction and vision-based gesture interaction. Researchers had developed a set of gesture recognition technology based on data gloves (Grimes, 1993). The continuous grasping sequence could be recognized by hand shape and contact point information obtained by data glove and tactile sensor, and a multi-user grasping action system was designed by using hidden Markov model (Bernardin et al. 2005). The recognition rate of the grasping action reached 95% accuracy. Generally, Gesture recognition based on data gloves can more intuitively obtain the information of the user's hand posture changes, and has a higher accuracy rate and faster recognition speed.

Pointing task is one of the most basic interactive tasks in the VR environment, which had been studied by many scholars on the factors that affect the task performance. It is found that light and cursor control could greatly improve the precision of short-range target selection by comparing light projection and light and cursor control (Baloup et al. 2019). Researchers investigated the effects of spatial reference system, motion path curvature, and target position on target selection performance under immersion environment. The results show that the performance of task with body as reference frame is better than that with head as reference frame (Zhou et al. 2020). The accuracy of the pointing task was the worst in the depth direction, but it was much better in the horizontal and vertical directions (Liang et al. 2017).

The main factors of the pointing performance considered in this research were the size and position of the target object, the target distance and reach area, input and output devices, the degree of freedom of the input device, and the user's preference (Argelaguet and Andujar, 2013). According to Fitts' Law, the size and spatial position of the target object directly affect the pointing performance, so this study mainly investigated the influence of the target object's size and the spatial position on the pointing performance in the VR environment (Leitkam et al. 2017).

METHODOLOGY

This study set up a multi-factor experiment to explore the impact of target objects in different depths and perspectives on task performance through ray selection tasks in VR environment. The following section introduces the details of the experiment.

Experiment Environment and Apparatus

In this experiment, the gesture recognition was based on data gloves, so we used the Noitom HI5 gloves to capture, collect and record the gesture data. The experiment equipment included a pair of Noitom HI5 gloves, a USB2.0 transceiver, two Vive trackers and an HTC-Vive headset. Each glove is equipped with seven 9-axis high-performance inertial sensors, which can accurately capture the entire hand movement and the hand joints to recognize gestures (see Figure 1).

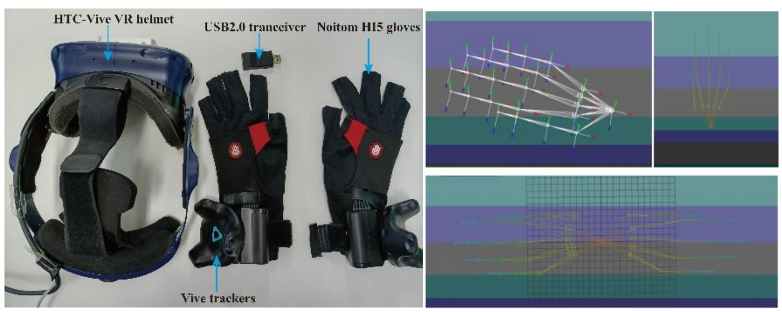

Figure 1. The experimental apparatus. (left) HTC-Vive VR headset and Noitom HI5 gloves. (right) The joint local axis of Noitom HI5 gloves.

Experiment Task

The main task of the experiment was the pointing task, which means that the subject used the ray-casting to point to the target object. The target object was a circular plane of the same size, and the depth and the viewing angle determined its position in the virtual scene. The experiment set three depth levels including 1m, 2.5m and 6m, which was the distance between the target object and the participant's eyes. At the same depth level, we set a total of 9 circular target objects, which were in vertical and horizontal viewing angles. The viewing angles were arranged evenly at 20° intervals, and the size of the target objects at different depth levels was also kept the same (see Figure 2). During the experiment, target objects in different depth levels would appear one after another. Participants needed to use a virtual ray to point to the center of the target object as soon as possible after the target object appeared and pressed the keyboard space bar to confirm their operation.

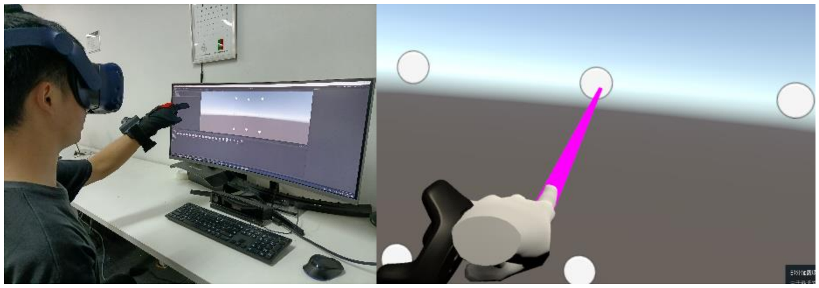

Figure 2. The pointing task based on HTC-Vive VR headset and Noitom HI5 gloves.

Participant

This experiment recruited 14 participants, including 7 males and 7 females. They were all between the ages of 23-26, and each participant had normal or corrected normal vision. None of the participants had any experience in using VR equipment.

Experiment Procedure

Before the practice experiment started, the participants put on Noitom HI5 gloves and HTC-Vive VR headset with the experimenter's help, adjusted the headset's field of view in the virtual scene, and then performed the Noitom HI5 gloves according to the standard calibration procedure. After that, the participants conducted a practice experiment. They would complete the same number of trials as the formal experiment. The participants needed to confirm that there was no obvious deviation in the pointing direction of the index finger and then completed a round of practice experiment with the guidance of the experimenter. After the participants were familiar with the equipment operation and experimental tasks, they could start the formal experiment. In the formal experiment, target objects of different depths would appear in sequence in the virtual scene. Participants needed to use gestures to point to the center of the target object as soon as possible after it appeared and pressed the space key to confirm their operation. After each round of trials was completed, they had a five-minute break. The entire experimental process lasted about half an hour.

RESULTS

The definition of pointing accuracy is the average value of the total distance between the coordinates pointed by the participant and the target object’s center location at each position. The smaller the value is, the higher the accuracy is. Blue points on the figures were the point position of all participants, and the points which had too much deviation from the center position of the target object were considered as the operation error and removed. The pointing accuracy was related to the size of the target object in the human field of vision, and the radius of the target object in the three levels of depth is equal (see Figure 3), the pointing accuracy was proportional to the distance from the participant's viewpoint to the target object, and the closer the target objects were to the central axis, the higher the accuracy would be. In addition, when the angle between the pointing point and Z-axis was 20° in the horizontal direction, the pointing accuracy on the right side of the central axis was less than that on the left side.

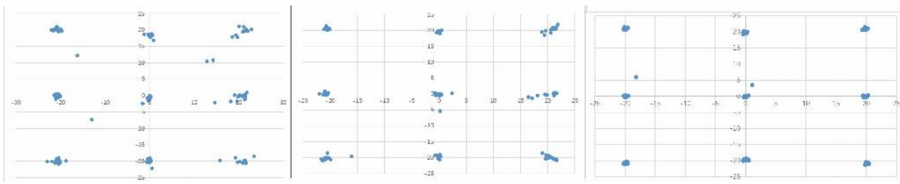

Figure 3. Distribution of actual pointing position at different depth levels. (left) depth=1m. (middle) depth=2.5m. (right) depth=6m.

Compared the coordinates of the intersecting point and the target coordinates of the ray interaction to explore the change of pointing accuracy under different positions. The coordinates were expressed in a three-dimensional polar coordinate system with the eye center as the spherical center, the deviation angle was in degrees, and the depth was in meters. Firstly, the data were taken out to compare the accuracy differences at different depths. It showed that there was a significant difference in pointing accuracy between 1m, 2.5m, and 6m, $F(2,376) = 18.137$, sig=0.000<0.01; The angle error of horizontal direction and vertical direction had significant difference. In the horizontal direction, $F(2,376) = 12.134$, sig=0.000<0.01; In the vertical direction $F(2,376) = 16.796$, sig=0.000<0.01. There were error distributions of ray interaction at different depths in horizontal and vertical directions (see Table 1). It could be seen that with the increase of depth, error on the deviation of viewing angle tended to decrease. The variation between 1 m and 2.5 m was very significant ($t(250)=4.000$,sig=0.0010<0.01), the change between 2.5 m and 6 m was also statistically significant ($t(253)=3.353$,sig=0.001<0.01), but the trend was not as large as the former.

Table 1: Horizontal and vertical error distributions of gesture interaction with ray feedback at different depths

Depth (m)	**Error mean of the vertical plane(degree)**	**Error standar-d deviatio-n of the vertical plane(degree)**	**Error mean of the horizontal direction(degree)**	**Error standard deviation of the horizontal direction(degree)**	**Error mean of the vertical direction(degree)**	**Error standard deviation of the vertical direction(degree)**
1	3.2189	6.38219	1.8524	4.25059	2.2267	4.96498
2.5	0.9459	0.76137	0.4750	0.55991	0.6997	0.66751

6	0.6427	0.67951	0.5542	0.67987	0.2118	0.23095

In the horizontal direction, there was no significant difference in pointing accuracy between the left and middle sides and right side, F (2,376) =1.374, sig=0.254>0.05, and there was no significant difference in the vertical direction, F (2,376) =1.230, sig=0.294 >0.05. However, there was a significant difference in the horizontal direction of 0.05 level, F (2,376) =3.647, sig=0.027 >0.05. The deviation for target widgets on the right side is significantly higher than for the left and middle sides (see Table 2).

Table 2: Horizontal and vertical error distributions of gesture interaction with ray feedback with different horizontal deviation

Horizont-al deviation (de-gree)	**Error mean of the vertical plane(deg)**	**Error standar-d deviatio-n of the vertical plane(deg)**	**Error mean of the horizontal direction(deg)**	**Error standard deviation of the horizontal direction(deg)**	**Error mean of the vertical direction(deg)**	**Error standard deviation of the vertical direction(deg)**
-20	1.3844	3.13089	0.5164	0.37856	0.7113	1.89494
0	1.3272	3.11606	0.9479	2.99644	1.2922	3.57723
20	2.0496	5.00614	1.3820	3.16878	1.0942	3.19123

In the vertical direction, there was no significant difference in pointing accuracy between the up row and middle row target widgets and the down target widget, F (2,376) =0.296, sig=0.744 >0.05. The horizontal direction (F (2,376) =0.831, sig=0.436 >0.05) and vertical direction(F (2,376) =0.562, sig=0.571 >0.05) and there was no significant difference (see Table 3).

Table 3: Horizontal and vertical error distributions of gesture interaction with ray feedback with different horizontal deviation

Vertical deviation(de-gree)	**Error mean of the vertic-al plane(d-eg)**	**Error standar-d deviatio-n of the vertical plane(deg)**	**Error mean of the horizontal direction(deg)**	**Error standard deviation of the horizontal direction(deg)**	**Error mean of the vertical direction(deg)**	**Error standard deviation of the vertical direction(deg)**
-20	1.7672	5.08715	0.7681	1.89491	0.8475	1.93609
0	1.5968	3.33621	1.1709	3.32977	1.0219	3.89984

20	1.3886	2.76568	0.9064	2.15838	1.2503	2.74584

DISCUSSION

During the experiment, participants needed to point to the center point of the target object appearing in different positions. And the actual pointing position would deviate from the target center point. The pointing point that deviates too much from the target center position is an operating error and not included in the data. In terms of different depth levels, the farther the depth distance is, the smaller the deviation between the actual pointing point and the target center position. In other words, the farther the target object is from the participant, the higher the pointing accuracy. At the same depth level, the closer the target object is to the participant's visual center, the higher the pointing accuracy. Moreover, the pointing accuracy in the horizontal direction is higher than that in the vertical direction. This is due to the participants' reading habits. People's usual reading speed is higher in the horizontal direction than in the vertical direction.

In addition, all participants used their dominant hand in this experiment, so the deviation between the actual pointing point on the dominant hand side and the target center position is smaller than the deviation on the non-dominant hand side. If the participant is right-handed, the deviation of the data on the right is smaller than that of the left; if the participant is left-handed, the deviation of the data on the left is smaller than that of the right. The overall analysis of all pointing point data reveals that the actual pointing point of the subject is lower than the target object position.

There were many limitations in this experiment. Participants did not have appropriate feedback when pointing to the center of the target object. Most participants proposed that the feedback mechanism with ray reduced the comfort and subjective willingness of interactive operations. Subsequent research would consider replacing ray feedback with a cursor or color change, or certain vibration feedback is generated when the pointing point falls into the appropriate area of the target object.

CONCLUSION

This study verifies that the pointing task performance is related to the target object's spatial position and depth level through gesture interaction experiments. The two variables that determine the pointing accuracy in the experiment are depth and the spatial position of the target object. The pointing accuracy is inversely proportional to the angle between the target object in the horizontal direction and the Z-axis and the angle between the vertical direction and the Y-axis. The closer the target control is to the center of the field of view, the higher the pointing accuracy. And compared to the horizontal direction, The pointing accuracy is more affected in the vertical direction. The pointing accuracy is proportional to the depth. The farther the target object is from the participant, the higher the pointing accuracy. At the same time, the actual pointing point of the participants is lower than the target center position, and the pointing deviation is smaller on the dominant hand side than the non-dominant hand side. The research results provide a reference for the target spatial arrangement of the pointing task in the VR environment. In the future, we would also further optimize and set more feedback methods.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (No. 71901061).

REFERENCES

Hoc, J.M. (2000). From human–machine interaction to human–machine cooperation. Ergonomics, vol. 43, pp. 833--843

Grimes, G. J. (1993). Digital data entry glove interface device. Bell Telephone Laboratories, Murray Hill, New Jersey. US Patent(4,414,537)

Bernardin, K., Ogawara, K., Ikeuchi, K., Dillmann, R. (2005). A sensor fusion approach for recognizing continuous human grasping sequences using hidden markov models. IEEE Transactions on Robotics, vol. 21, pp. 47-57

Baloup, M., Pietrzak, T., Casiez, G. (2019). Raycursor: A 3d pointing facilitation technique based on raycasting. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-12

Zhou, Q., Yu, D., Reinoso, M. N., Newn, J., Goncalves, J., Velloso, E. (2020). Eyes-free Target Acquisition During Walking in Immersive Mixed Reality. IEEE Transactions on Visualization and Computer Graphics, vol. 26(12), pp. 3423-3433

Liang, Z., Xu, X., Zhou, S. (2017). The smallest target size for a comfortable pointing in freehand space: Human pointing precision of freehand interaction. Universal Access in the Information Society, vol. 16, pp. 381-393

Argelaguet, F., Andujar, C. (2013). A survey of 3D object selection techniques for virtual environments. Computers and Graphics, vol. 37, pp. 121-136

Leitkam, S.T., Applegate, M., Thomas, J.S. (2017). Full-body reaching assessment of Fitts' Law in a virtual environment. International Conference on Virtual Rehabilitation, 1-6