Evaluation of pointing techniques for ray casting selection in virtual environments

SangYoon Lee, Jinseok Seo, Gerard Kim, Chan-Mo Park

SangYoon Lee, Jinseok Seo, Gerard Jounghyun Kim, Chan-Mo Park, "Evaluation of pointing techniques for ray casting selection in virtual environments," Proc. SPIE 4756, Third International Conference on Virtual Reality and Its Application in Industry, (1 April 2003); doi: 10.1117/12.497665

Event: Third International Conference on Virtual Reality and Its Application in Industry, 2002, Hangzhou, China

Evaluation of Pointing Techniques for Ray Casting Selection in Virtual Environments

Sangyoon Lee, Jinseok Seo, Gerard Jounghyun Kim, Chan-Mo Park*
Virtual Reality Laboratory
Department of Computer Science and Engineering
Pohang University of Science and Technology (POSTECH)

ABSTRACT

Various techniques for object selection in virtual environments have been proposed over the years. Among them, the virtual pointer or ray-casting is one of the most popular method for object selection because it is easy and intuitive to use and allows the user to select objects that are far away. Variants of the virtual pointer metaphor include the Aperture [5], Flashlight [3], and Image plane method [1] as categorized as such by [10]. In a monoscopic environment, these methods are essentially 2D interaction techniques, as the selection is made effectively on the image plane. Such a 2D based selection (or more generally, interaction) method has an added advantage in that it can find many good uses in 3D environments ranging from a simple 2D oriented subtask (object selection on a constrained surface, menu selection) to a siutation where a whole 2D application (e.g. sketching tool, desktop manager) is embedded in the 3D environment. In this paper, we experimentally compare the performance of four different virtual pointer implementations, namely, the direct image plane selection, head-directed pointer, hand directed pointer and head-hand directed pointer. The experimental results revealed that the direct image plane selection produced the best performance among the four in terms of both task completion time and the pixel-level pointing error.

Keywords: Virtual pointer, 2D based interaction, pointing, ray casting, task performance, object selection

1. INTRODUCTION

Pointing is often necessary to select objects that are not within the work volume of the user in virtual environments. Various techniques for object pointing and selection in virtual environments have been proposed over the years. Perhaps the most commonly used method is the virtual ray casting (or flashlight) [2][3] where the first object intersected with the virtual ray (whose direction and origin are determined by a 6D tracking device) gets selected. This method is quite intuitive in the sense that it is a natural abstraction of the real world finger pointing. Pierce [1] has proposed a set of image plane interaction techniques in which the user interacts with the 2D projections that 3D objects in the scene make on one's image plane. Thus, in a monoscopic environment, all these methods are essentially 2D interaction techniques, as the selection is made on the image plane, as also categorized as such by [10]. For instance, the virtual ray casting is a technique that uses the concept of projection along the direction of the ray. The cursor position is obtained by a function that maps a 3D position in the real world (i.e. from the tracking device) to a point on the image plane. When using a virtual ray, the "cursor" position is the point on the image plane collided with the ray. As illustrated in Figure 1, unless the object (or point) to be selected lies on the image plane, pointing toward the projected point is not equivalent to the 3D pointing. However, we can easily find the corresponding 3D point (or object) using the viewpoint. Rendering of the ray itself presents no problem as the ray to the projected point and the 3D virtual ray are projectively equivalent, thus the user should detect no visual difference (we do not consider the effect of depth perception for the moment).

There are variety of ray casting methods depending on how the origin and direction of the ray is determined, using the hand position/orientation (hand directed), hand orientation and head position (head-hand directed), head

* Futher author information
Sangyoon Lee: E-Mail: *sylee@postech.ac.kr*, Address: *San 31, Hyoja-dong, Pohang, Kyoungbuk 790-784 South Korea*
Jinseok Seo: E-Mail: *jsseo@postech.ac.kr*, Address: *San 31, Hyoja-dong, Pohang, Kyoungbuk 790-784 South Korea*
Gerard Jounghyun Kim: E-Mail: *gkim@postech.ac.kr*, Address: *San 31, Hyoja-dong, Pohang, Kyoungbuk 790-784 South Korea*
Chan-Mo Park: E-Mail: *parkcm@postech.ac.kr*, Address: *San 31, Hyoja-dong, Pohang, Kyoungbuk 790-784 South Korea*

Third International Conference on Virtual Reality and Its Application in Industry,
Zhigeng Pan, Jiaoying Shi, Editors, Proceedings of SPIE Vol. 4756 (2003)

position/orientation (head directed) (See Figure 2 and 4), and thus the "cursor" position [11]. For direct image plane pointing, the cursor is drawn by projecting the user's fingertip position orthographically on to the image plane (See Figure 4).

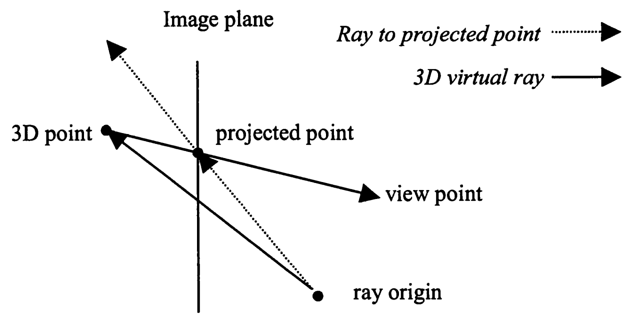

Figure 1. 3D virtual ray casting and direct image plane selection.

Such 2D based selection (or more generally, interaction) method has an added advantage in that it can find many good uses in 3D environments ranging from a simple 2D oriented subtask (object selection on a constrained surface, menu selection) to a situation where a whole 2D application (e.g. sketching tool, desktop manager) is embedded in the 3D environment. As fast and accurate pointing is the first step toward effective object selection in virtual environments, we, in this paper, experimentally compare the object selection performance of four different virtual pointer implementations. The four methods are the direct image plane selection, head-directed pointer, hand directed pointer and head-hand directed pointer, all different in the way of generating the cursor position. A large projection display was used for the experiment to neutralize the problems with weight of HMD's, and thus, the full head tracking was not used (assumed fixed head position). The respective performance in pointing was measured in terms of the task completion time and the pixel-level pointing error.

This paper is organized as follows. First, the next section briefly reviews some related work. Section 3 explains the details of each of the four cursor-based pointing method, and Section 4 describes the usability experiment. Section 5 presents the results and discussion, and Section 6 concludes the paper with a short summary of the work presented.

2. RELATED WORK

While it is well known that the 2D mouse works the best in 2D desktop environments, no definite results exist for achieving 2D pointing tasks in the 3D environments [12]. Bowman et al [11] gives a good review of various selection methods developed and tested for virtual environments. The idea of the ray casting method dates back to the early 80's [2]. Liang and Forsberg proposed the spot light approach to object selection [3][5]. Bowman and Poupyrev made several usability studies for object selection in the virtual environment, and also developed the GO-GO method in which the virtual hand extended in a non-linear fashion to reach far away objects [4][6]. Fukumoto et al proposed a pointing method in which the ray origin (called the virtual projection origin), different from the viewpoint, was derived [8]. Pierce et al [1] proposed a set of image plane based interaction techniques, including the "Head Crusher", "Sticky Finger", "Lifting Palm", and the "Framing Hands". However, no usability studies to compare its relative performances were made to this date despite potential utility in the VE.

3. THE FOUR POINTING TECHNIQUES

3.1. Head-Hand directed

This technique uses a ray whose origin starts at the head (or viewing) position, and whose direction is determined by the vector formed by the hand and head positions. The cursor position is where this ray collides with the cursor plane (See Figure 2). The cursor plane is perpendicular to the viewing direction, and its distance to the viewing position is dependent on the viewing angle and the image space size (See Figure 5). Note that one unit of cursor movement is projected to one pixel on the image plane. This rule is similar to the "Sticky Finger" of Pierce et al [1] to reflect direct manipulation of the cursor. This method is very intuitive and the aiming metaphor works well with the human's hand-eye coordination.

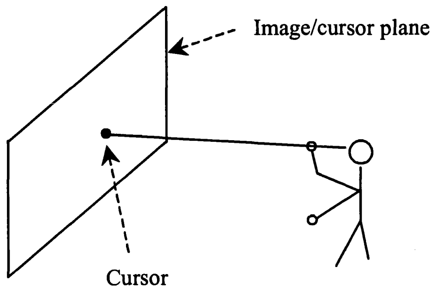

Figure 2. Head-hand directed technique.

3.2. Image Plane Selection (2D Virtual Mouse)

The 2D virtual mouse uses a 3D-to-2D orthographic mapping from a hand position to a point on the cursor plane. Therefore, it uses just two degrees of freedom (See Figure 3), therefore, is fundamentally different from the ray approach where higher degrees of freedom are needed. This method is easy to understand and use, due to the familiarity with and user's prior exposure to the popular desktop/mouse environment.

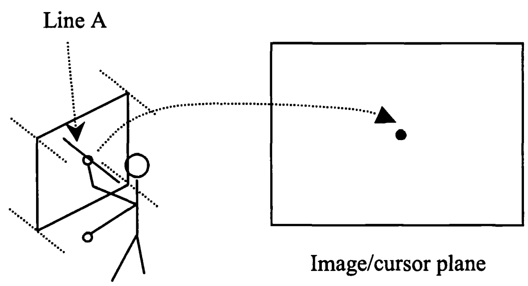

Figure 3. Direct image plane selection (2D virtual mouse). All of the hand positions along line A map to one point on the cursor plane.

3.3. Hand-directed

This technique uses both the position and orientation of one's hand as the ray's origin and direction respectively (See Figure 4). Users are able to point by changing the hand direction and the hand position. This method is known to be difficult to use for novice users, and may be tiring to use as the head/neck is not used.

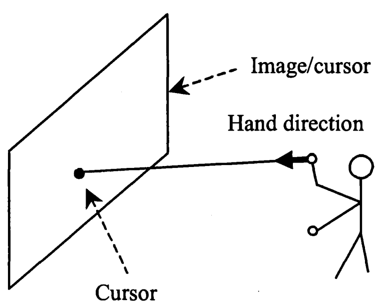

Figure 4. Hand-directed technique.

3.4. Head-directed

The head-directed technique is the same as the hand-directed one except that it uses the head position and orientation. Users are able to point by changing the direction and position of the head. This is similar to the gaze directed method, if we can assume the head direction is where the gaze is pointed at. However, the coupling of looking and selection (also for head-hand directed case) is known to cause usability problems. Precise pointing is also known to be difficult as it requires subtle movement using the neck.

4. THE EXPERIMENT

4.1. The Task

The task used for the usability experiment is illustrated in Figure 5. At first, the user was presented with one colored sphere (the 3D spheres appear as circles in different sizes according to their depths). The user was to select the same color from the menu. If the correct color was selected, then, a second white sphere appeared to which the user had to draw a line. The line was drawn by selecting the first sphere and dragging it to the second white sphere. If this task was successfully accomplished the white sphere became filled with a color and the overall task repeated, until the user completed a predefined number of line segments, which was 10 in this experiment. The user was asked to complete task as fast and accurate as possible, and the task completion time and pixel level error (from the center of the projected sphere) were measured. Cross marks were used on circles (or projected spheres) to indicate the center of the sphere (as a target for the pixel error subtask). The task for the experiment is mostly similar to the one carried out for the comparison of two handed interaction techniques [9].

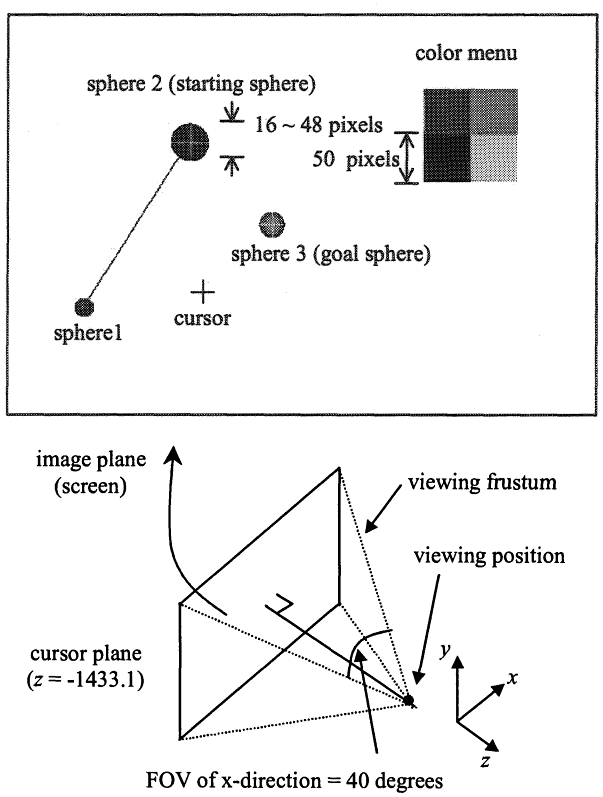

Figure 5. The task used in experiment and setting of cursor plane. At first the user was presented with one colored sphere. The user was to select the same color from the menu. Then, a second white sphere appeared to which the user had to draw a line. Once the line was drawn, the white sphere became filled with a color and the overall task repeated few times.

4.2. The Experiment Environment

In the experiment, we used a 120 inches projection display, the Logitech 3D Mouse for tracking positions/orientation of the hand, and the Polhemus FASTRAK for tracking orientations of the head. The resolution of the image plane was set at 1024 x 768.

4.3. The Experiment Procedure

A subject performed 1 actual trial for each pointing technique, a total of 4 trials in a random order, after 1 training trial for each technique. As for the data set used, the positions of the sphere were chosen randomly, however, the same data set was used for all four pointing methods. The number of the subjects was 20 (their age was between 24 and 28 years). All the subjects had used the desktop computer systems daily with 2D mice and half of the subjects had some or extensive experience in using 3D interfaces for computer graphics/games. We asked the subjects to stand while performing the task without any supporting structure. The subjects pointed using trackers attached to one's dominating hand/head and clicked on a button using one's non-dominant hand. In the experiment, all the subjects used their right hands as the dominant hand.

We assumed that the head position was fixed. We asked the subject not to change the position of their heads. Nevertheless, when using the head-hand/head directed technique, there was no subject who complained about the tracker position not matching with the cursor position. Figure 6 is a photograph that was taken when a subject performed a trial using hand-directed technique. Finally, users were asked to complete a subjective questionnaire asking preference, difficulty and the level of fatigue.

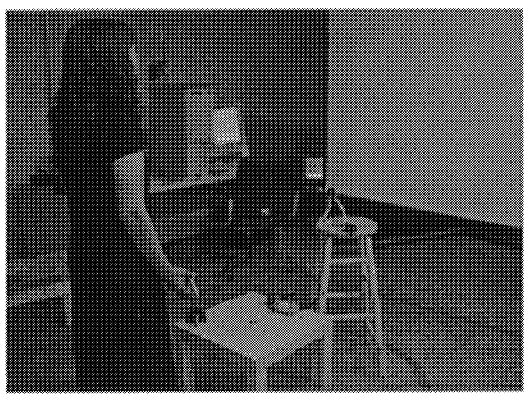

Figure 6. A subject trying out the hand-directed method.

5. RESULTS AND DISCUSSION

Table 1 shows the task completion time for the experiment and its within-subject ANOVA result. The Tukey HSD post-hoc analysis has revealed significant statistical differences among the techniques except for the case between 2D virtual mouse and hand-directed ($p < 0.01$). The 2D virtual mouse turned out to be generally the fastest method.

The pixel error measurements and the within-subject ANOVA result are shown in Table 2, and similar results were obtained. We observed that the aiming action of the head-hand directed method required the users to maintain their hands almost at the height of their eyes, and the subjects shook their hands more than when using their hands (at a lower position) for the hand-directed technique. The shaking prevented users from precise pointing and its effect was more apparent on the head-hand-directed technique, and the performance results seem to reflect this phenomenon.

On the other hand, we did not observe any shaking of the hands when using the 2D virtual mouse. It was because the subjects kept their hands pulled to their bodies to make their hands stable and to reduce arm-fatigue. In fact, the subjects tried to pull their hands to their bodies when using the head-hand-directed technique as well, however, but they soon realized it was difficult to use the head-hand-directed technique with such a pose and resumed to the normal pose. The relatively lower performance trend of the hand directed method compared to that of the 2D virtual mouse seems be due to the difficulty in resolving the natural hand-eye coordination cue in controlling "end-of-the-ray", plus subjects usually felt easier to "move", rather than orient, to a position on the screen.

The head-directed technique, as expected, came out the worst both task completion time and accuracy, most likely due to the strain and difficulty in controlling the head precisely. It is our general observation and conclusion that in our experiment, the 2D virtual mouse performed the best task completion time wise and accuracy wise. However, it was found that the method may perform differently for different image/cursor plane distance. For a relatively close image plane, the hand position has more influence to the cursor position, and conversely, for far image planes, the hand orientation exerts more influence. This issue needs further investigation as a future work.

Figure 7 shows the preference of the subjects. The subjects who preferred the 2D virtual mouse reported that the technique was easy to understand and control. On the other hand, subjects who chose the hand-directed technique as their best reported that they did not feel any fatigue in it hand position could be kept stable. The subjects who did not like the head-hand-directed nor the head-directed technique acknowledged and confirmed that the former gave a lot of arm-fatigue, and that the latter was too difficult to control precisely.

Table 1. The task completion time ($F_{3,236} = 7.39, p < 0.0005$)

Techniques	The average of task completion time (sec)	The standard deviation (sec)
M1: Hand-head-directed	98.0	23.8
M2: 2D virtual mouse	83.3 *	13.4
M3: Hand-directed	92.2	17.2
M4: Head-directed	106.4	19.4

Table 2. The pixel-level error ($F_{3,236} = 7.12, p < 0.0005$)

Techniques	The average of pixel-level error	The standard deviation
M1: Hand-head-directed	5.0	0.98
M2: 2D virtual mouse	4.4 *	0.98
M3: Hand-directed	4.8	0.96
M4: Head-directed	5.2	1.06

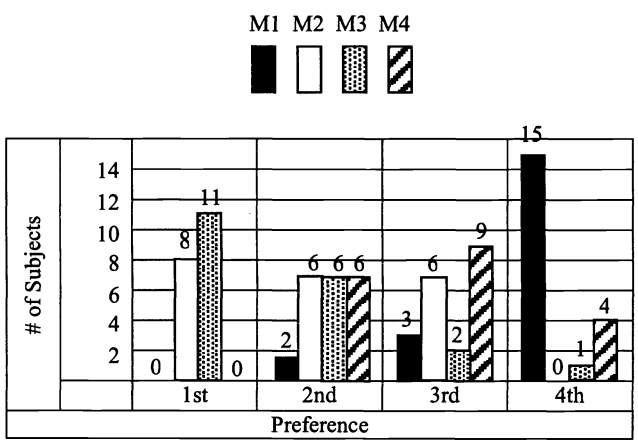

Figure 7. The preference of the subjects (the number of subjects rating each method to be their first, second, third and fourth choice). (M1: Head-hand, M2: 2D Virtual Mouse, M3: Hand, M4: Head)

6. CONCLUSION AND FUTURE WORK

Fast and accurate pointing is important for effective object selection in virtual environments, and thus, we compared four representative virtual pointer implementations based on how the cursor (end of the ray) is formed on the image/cursor plane. We found a strong trend that the 2D virtual mouse performed the fastest and most accurate. Such an interface design guideline should prove valuable for 2D subtasks or 2D applications embedded in VE's (we make a note that the cursor plane does not have to coincide with the image plane).

We plan to continue the usability study under other conditions, for instance, with head tracked HMD's, or for stereoscopic display. We are also interested in applying the interaction method for image-based or hybrid (image + 3D geometry) virtual environments. Finally, the image plane interaction can be enacted easily by the foot since the foot naturally has a 2D constrained surface. It would be interesting to test if the foot-based 2D interaction would be helpful as an auxiliary interaction channel, when the dominating hand is occupied for different tasks.

ACKNOWLEDGEMENT

The work presented in this paper has been supported in part by the Korea Ministry of Education's BK21 Project and Korea Science and Engineering Foundation (KOSEF) supported Virtual Reality Research Center.

REFERENCES

1. J. S. Pierce et al., "Image Plane Interaction Techniques in 3D Immersive Environments", Proc. of the 1997 Symposium on Interactive 3D Graphics, 1997.
2. R. Bolt, "Put-That-There: Voice and Gesture at the Graphics Interface", Proc. of the Siggraph '80, 1980, pp. 262-270.
3. J. Liang and M. Green, "JDCAD: A Highly Interactive 3D Modeling System", Computer and Graphics, Vol. 18, No. 4, 1994, pp. 499-506.
4. D. Bowman et al., "The Art and Science of 3D Interaction", IEEE VR 99 Tutorial Notes, 1999.
5. A. Forsberg, K. Herndon and R. Zeleznik, "Aperture based Selection for Immersive Virtual Environments", Proc. of the UIST '96, 1996, pp. 95-96.
6. I. Poupyrev et al., "Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR", Proc. of the UIST '96, 1996, pp. 79-80.
7. D. Bowman and L. Hodges, "An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments", Proc. of the 1997 Symposium on Interactive 3D Graphics, 1997.
8. M. Fukumoto et al., "Finger-Pointer: Pointing Interface by Image Processing", Computer and Graphics, Vol. 18, No. 5, 1994, pp. 633-642.
9. P. Kabbash, W. Buxton and Abigail Sellen, "Two-Handed Input in a Compound Task", Proc. of CHI: Human Factors in Computing Systems '94, 1994, pp. 417-423.
10. I. Poupyrev, et al, "Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques", Eurographics '98, Vol 17, No. 3, 1998
11. D. Bowman and L. Hodges, "Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments:, Journal of Visual Languages and Computing, Vol 10, pp. 37-53, 1998
12. S. Card, T. Moran, and A. Newell, "The Psychology of Human-Computer Interaction", Hillsdale, Lawrence Erlbaum, 1983