Gaze- vs. Hand-Based Pointing in Virtual Environments

Nathan Cournia, John D. Smith, Andrew T. Duchowski
Department of Computer Science, Clemson University
{ *acnatha* | *jsmith6* | *andrewd* }*@vr.clemson.edu*

ABSTRACT

This paper contributes to the nascent body of literature on pointing performance in Virtual Environments (VEs), comparing gaze- and hand-based pointing. Contrary to previous findings, preliminary results indicate that gaze-based pointing is slower than hand-based pointing for distant objects.

Introduction

In their study of interactive object selection in Virtual Environments (VEs), Tanriverdi and Jacob tested gaze-based pointing against more conventional hand-based pointing using a magnetic tracker worn on a ring around the participant's finger [4]. Experimental procedures included selection of "close" and "distant" objects in VR. Results indicated that performance with gaze-based pointing was significantly faster in the distant VE (no performance difference was found in the close VE).

While Tanriverdi and Jacob's experiment is one of the first to objectively examine gaze-based interaction in VR, hand-based interaction appears to have been based on the arm-extension paradigm. Comparison of pointing based on ray-casting (gaze-based) with arm-extension (hand-based) may not necessarily be fair (e.g., Bowman and Hodges classify and test interactive pointing techniques in terms of arm-extension and ray-casting but do not report any speed advantage between the two methods [1]).

We extend the work of Tanriverdi and Jacob and compare gaze-based and hand-based pointing replicating their experimental paradigm except for two important distinctions. First, we employ a ray-casting approach for hand-based selection. Ray-casting techniques make use of a virtual light ray to select an object, resulting in the user's wielding of a "virtual wand" with which s/he can select objects at near or far distances (see Figure 1). Second, instead of a planar target arrangement, as what appears to have been modeled in Tanriverdi and Jacob's VE, we distribute our virtual objects about a 360° truncated sphere surrounding the participant.[1]

[1] A planar arrangement of objects may not afford full head or body rotation, at least in the distant environment, in our opinion negating the immersive benefits of VR.

Figure 1: User wearing HMD wielding "virtual wand".

Methodology

Apparatus: Our VE is driven by a 1.5 GHz dual-CPU Linux PC with 1 G RAM and an NVidia GeForce4 Ti 4600 graphics card. Multi-modal components include a binocular eye tracker mounted within a Virtual Research V8 Head Mounted Display (HMD). HMD position and orientation is provided by an Ascension 6 Degree-Of-Freedom (6DOF) Flock Of Birds (FOB). A 6DOF tracked, hand-held mouse provides a means to represent a virtual tool for the user in VR (see Duchowski et al. [2] for details).

Interaction Techniques: We use `OpenGL`'s selection buffer to implement gaze-based pointing from eye tracked Point Of Regard (POR) coordinates. Visual feedback is given in the form of a small dot (similar to a laser pointer) at the user's POR. Figure 2 (left) shows an instance of a selected object with the gaze dot just visible on the middle part of the letter target. Dwell time is used to indicate selection, avoiding the well-known Midas Touch problem [3].

Hand-based pointing is implemented following the ray-casting approach. Users hold a 3D-tracked mouse and select objects in VR by pointing the virtual wand (see Figure 1). The tip of the wand changes size in relation to its distance from the user and changes color when the wand's intersection with an object is detected. To select a currently intersected object, the user presses the right button on the 3D-tracked mouse. The wand is represented by a blue line and can be seen in Figure 2 (middle).[2]

[2] In preliminary trials we observed that a simple dot (as used for gaze-based feedback) gave users trouble controlling the depth of the wand

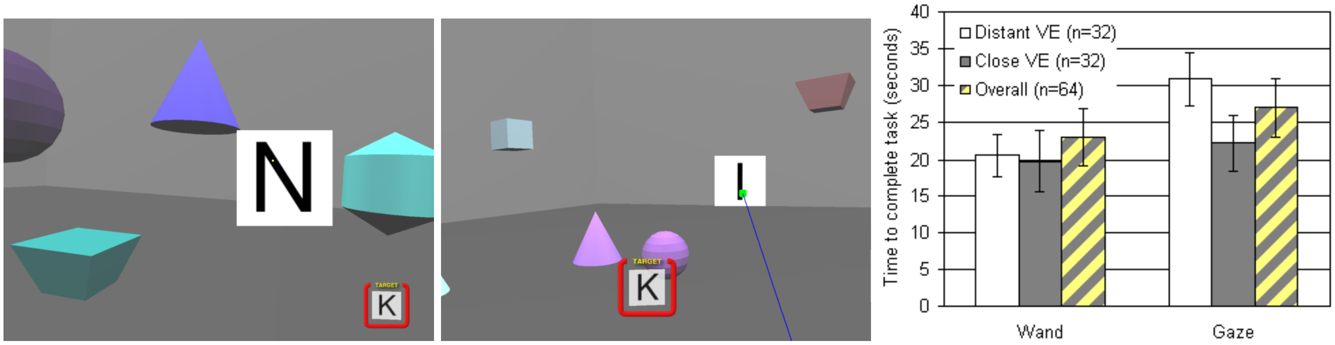

Figure 2: Gaze- (left) and hand-based selection with "virtual wand" (middle); search time performance results (right).

Stimuli: Our VE consists of 20 objects rendered at randomly predetermined locations equidistant from the user. Close objects are distributed with constant radius roughly corresponding to arm's length; distant objects are distributed with constant radius beyond arm's reach. Objects randomly take on color and one of 5 possible shapes: frustum, sphere, cube, cone, or an elongated rhomboid (resembling a top). When selected, an object's shape fades to reveal a flat square mapped with a randomly assigned sans-serif black-on-white letter (ranging from A-N), as shown in Figure 2.

Participants: Twenty-four participants (17 m/7 f) volunteered to take part in the experiment, all undergraduate or graduate students, with ages ranging from 19 to 34 (average age 23). All successfully completed the experiment (none were eliminated due to unsuccessful eye tracker calibration).

Experimental Design: Independent variables controlled in the experiment were selection modality (gaze, wand) and object distance (close, distant). A within-subjects design was used for interaction techniques and a between-subjects design was chosen for object distance, i.e., half the participants (randomly assigned) interacted with close objects and half with distant objects. A Latin square was used to counteract interaction modality ordering effects. Each participant completed 4 trials per interaction condition resulting in 8 trials per person.

Procedure: Unlimited training time was provided prior to commencement of each trial. Instructions were given for the participant to attempt to locate the given target letter as quickly as possible. For gaze-based pointing, a short 5-point eye tracker calibration sequence was performed before display of the VE. At each task commencement, the participant was asked to assume an initial position and orientation by aligning a blue ball target in the center of view.

pointer—this is probably due to the wand's introduction of additional degrees of freedom that the user must control. Moreover, unlike with gaze, the user must compute additional geometry transformation in their minds to point the wand pointer at the desired target object. For this reason, we provided additional feedback by drawing the wand vector in an attempt to equate the two ray-casting pointing modalities.

Results

Figure 2 (right) gives descriptive statistics from measurements of the dependent variable (time to complete task). Comparing performance scores using pair-wise ANOVA, we noticed a significant difference in performance in the distant environment between gaze- and hand-based pointing ($F[1,190] = 5.15$, $p < 0.05$), suggesting slower gaze-based performance.[3]

Discussion and Conclusions

We did not find a clear performance advantage of gaze-based pointing over hand-based pointing; instead a significant performance decrease was observed in the distant object condition. We believe our results contradict previous research primarily because of our use of ray-casting for both gaze- and hand-based pointing, i.e., our results espouse the similarity between the two pointing styles when both use ray-casting.

Based on preliminary analysis, our recommendation for pointing in VR is as follows: for performance comparable to gaze-based pointing with the hand (e.g., when an eye tracker is unavailable), use ray-casting but not arm-extension.

REFERENCES

1. BOWMAN, D. A., AND HODGES, L. F. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Symposium on Interactive 3D Graphics* (1997), IEEE, pp. 35–38.

2. DUCHOWSKI, A., MEDLIN, E., GRAMOPADHYE, A., MELLOY, B., AND NAIR, S. Binocular Eye Tracking in VR for Visual Inspection Training. In *VRST '01 Proceedings* (2001), ACM, pp. 1–8.

3. JACOB, R. J. What You Look at is What You Get: Eye Movement-Based Interaction Techniques. In *CHI '90 Proceedings* (1990), ACM, pp. 11–18.

4. TANRIVERDI, V., AND JACOB, R. J. K. Interacting with Eye Movements in Virtual Environments. In *CHI '00 Proceedings* (2000), ACM, pp. 265–272.

[3] Although a similar trend is evident in the close VE, no signifi cant difference in performance was observed.