GENERALIZEABILITY OF LATENCY DETECTION IN A VARIETY OF VIRTUAL ENVIRONMENTS

Stephen R. Ellis [1], Katerina Mania [2], Bernard D. Adelstein [1], Michael I. Hill [1]
[1] NASA Ames Research Center, USA
Human Factors Research and Technology Division

[2] University of Sussex, UK
Department of Informatics

User perceptual sensitivity to changes of system latency was tested in three simple virtual environments: one with only a foreground object, a second with only a background object, and a third that combined both of these elements. Prior psychophysical measurements of sensitivity, Just Noticeable Difference; and bias, Points of Subjective Equality, from our laboratory are confirmed with measurements in 13 subjects. Our measurements indicate that perceptual stability across a variety of virtual environments will require latencies less than 16 ms. We discount a possible explanation that the differences between our results and those from a study by Allison *et al*. could be related to a visual capture effect initially reported by L. Matin. Instead, the differences may be due to the type of psychophysical judgment rendered by the subjects and the degree to which subjects were instructed and practiced.

INTRODUCTION

Excessive system latency is a well-known defect of current virtual environment (VE) and teleoperation systems. It is particularly troublesome for head-tracked systems since delays in head orientation measurement give rise to errors in presented visual direction. But it is doubly confusing for hand-held virtual objects because the reversed error signs for hand and head tracked elements result in spurious, complex, apparent movement to which users have difficulty adapting (Smith, McCrary & Smith, 1962; Cunningham et al. 2001). Depending upon the amount of error and the user's specific behavior, these dynamic registration errors can lead to visual instability and can degrade performance and simulation fidelity (Ellis, Bréant, Menges, Jacoby & Adelstein, 1997). In severe cases, they can lead to disorientation and/or "simulator sickness," making the use of a VE problematic (Frank, Casali, & Wierwilli, 1988). Visual motor performance is also well known to be degraded by transport delays (Ferrell & Sheridan, 1963), the type of latency that arises due to computation and rendering lags in virtual environment systems.

Latency in VEs has consequently been the object of measurement, and software and hardware optimization (Jacoby, Adelstein, & Ellis, 1996; Hill, Adelstein, & Ellis, 2004; Regan, Miller, Rubin & Kogelnik, 1999). Because latency is unavoidable in some systems such as long-distance space teleoperation, software or hardware techniques cannot eliminate it totally from real-time operation. Techniques for managing it, thus, necessarily must be informed by user studies of its performance and perceptual impact.

Earlier studies have examined overall user performance (e.g., Ellis, *et al*., 1997; Watson, Walker, Woytiuk & Ribarsky, 2003). Some have investigated users' ability to detect latency during hand motion of virtual objects (Ellis et al., 1999a). Others have investigated user sensitivities to latency changes during head rotations (Ellis, *et al*., 1999b; Regan et al., 1999). Many of these studies suggest that representative user sensitivity to latency is quite acute with latency difference becoming noted when on the order of 10-20 ms. These studies show also that the latency discrimination process does not follow Weber's Law, that it cannot be based on a purely temporal judgement, and that it depends to at least some degree upon image slip during movement (Adelstein, Lee & Ellis, 2003).

A recent study by Allison *et al* (2001), however, suggests that visual instability due to system latency during periodic head movements is not apparent at a 50% detection rate until latencies are on the order of 200 ms. This value, which is the Point of Subjective Equality (PSE), conflicts with measurements made in our laboratory . Our measurements indicate user sensitivity to latency to be much more acute during similar head rotations. Although there are technical differences in apparatus and psychophysical method between the two laboratories' studies, the virtual environments should have had at least similar visual quality since both used the same Virtual Research V8 head mounted display (HMD).

One major difference, however, is that in prior studies (Ellis *et al*., 1999ab; Adelstein et al., 2003) we have asked our subjects to perform a two-interval judgment, comparing the apparent stability of two instances of system latency while viewing a foreground virtual object presented against a black background. In contrast using a single interval judgment, Allison et al.'s subjects judged the visual stability of a patterned background surface alone against their internal sense of what constituted stability Allison *et al*. only measured their observer's bias (PSE) [1], while Adelstein et al. reported both JND[2] and PSE.

The contrast in viewing conditions between Allison's and our prior studies could be determinative since there are well known differences between the apparent stability of the visual direction of small foreground objects versus that of backgrounds. For example, Matin *et al*. (1982) (see also Stark & Bridgeman, 1983) have shown that while the efference associated with a person's attempt to move a paralyzed eye can lead to apparent movement and displacement of a fixated small isolated, object, such illusory movement is not noted if the subjects instead observe a fully enveloping environment. In a sense, the full environment "captures" the users' visual reference frame and thereby obscures apparent displacement otherwise associated with the efference. The illusory motion is caused in Matin's case by a mismatch between the subject's efferent signal to move their eyes and the failure of the eyes to move because of local paralysis.

A similar contrast between the illusory perception of object or environment movement could underlie the explanation of the differences between our and Allison et al.'s measurements of latency detectability. We have generally examined the stability of small isolated objects, while Allison et al. looked at environment stability.

Accordingly, we designed the following experiment to investigate the generality of our previous measures of user ability to discriminate latency changes. One condition replicates the environment used by Allison et al. (2001). A second condition replicates the one used in our previous studies. The third combines features of the two and allows examination of whether previous determinations of latency generalize to a more complex environment exhibiting relative motion between objects at different depths during observer movement, a cue that could increase user sensitivity to latency changes. In all cases, we use a two down-one up adaptive staircase procedure (Levitt, 1970) to determine subjects' psychometric functions, a psychophysical method similar but more efficient (Watson & Fitzhugh, 1990) than that in our previous studies.

METHODS

Subjects

Twelve subjects (9 male, 4 female, with normal or corrected normal vision; ages 21-44 years) naïve with respect to the experimental hypothesis participated in a repeated measures Latin square experiment. In addition, one of the authors (KM) provided an additional replication of one of the six cells in the Latin square design.

Apparatus and software

The experiment's virtual environment (VE) simulation was run under Windows 2000 on a with Dell Precision 530 workstations (Dual 2.4 GHz Xeon processors, NVidia GeForce4 MX-440 graphics card). The VE system employed in this work included a single receiver Polhemus Fastrak running at 120 Hz for motion sensing and a Virtual Research V8 HMD. The experiment software is built as two separate Visual C++ applications. One is a Windows service (AuAST, custom written by AuSIM Inc.), through which the Fastrak and continuously refreshes head position and orientation data that are stored into shared memory locations. The second application, developed by co-author M. Hill, models and renders the visual content of the experimental VE based on the most recent Fastrak data retrieved from the AuAST shared memory locations.

Figure 1. Virtual environment display equipment.

Participants viewed one of three VEs. The first was an empty, pink environment, containing only a red octahedral frame, built from back-to-back right pyramids joined at their 15 cm square base, with a total height of 15 cm. The second was a large faceted, spherical surface (100 cm radius), which was viewed from inside and texture-mapped with a red and white checkerboard-like pattern, defined by twelve regular longitudinal and five latitudinal divisions. This environment matched the one used by Allison *et al*. (2001). The third environment combined the octahedron and the faceted sphere of the first two conditions. All three conditions were adjusted for approximately constant, mid-photopic space averaged luminance.

The position of the virtual octahedron was fixed in world coordinates at eye-height ~43 cm in front of the seated viewer's eye-point. This position placed it 57 cm in front of the background checkerboard when it was present. At this distance, the octahedron occupied a horizontal visual angle of 20°. The custom tracker driver and a multi-processing, shared-memory architecture described above ensured less than a 9 ms base

[1] The latency difference which is detected 50% of the time when actually present in a two alternative forced choice situation is the point of subjective equality (PSE).

[2] The change in latency required to increase or decrease detection 25% from the PSE is the just noticeable difference (JND).

latency[3] and a constant 60 Hz update rate, providing the very high dynamic performance that made the following experiment possible.

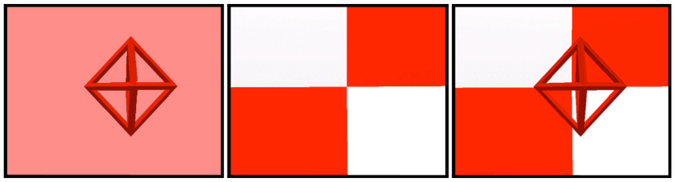

Figure 2. Screen images capturing part of the three alternative virtual environments as seen by the subject through the HMD).

Procedure and experimental design

The subjects were instructed to yaw their head smoothly and sinusoidally from side to side (30° end-to-end). They were paced by computer-generated beeps every 1 s, marking four intervals for two full back and forth motions. If they turned too far, the scene darkened by 58% to signal excessive rotation. Each visual stimulus remained fully visible while spanning the HMD's 48° horizontal FOV.

The subject's judgments, signaled by button pushes on a hand-held input device, advanced the stimulus levels according to an adaptive staircase algorithm, incorporating descending or ascending 8.3 ms latency steps. The experiment comprised a single scripted set of 18 staircases, comprising 6 staircases per visual condition (3 starting at the max level of and initially descending 3 starting at zero and initially ascending). A set of two staircases, one initially ascending and the other descending, both with the same visual stimulus condition were interleaved to prevent subjects from tracking and predicting the sequence of comparisons. Thus there were three sets of interleaved staircases per visual condition. The reference latency condition (R) was fixed at 8.5 ms. and compared with a probe (P) condition that was experimentally varied according to the staircase algorithm. The experiment's visual conditions were presented to all subjects in counterbalanced order using a Latin square design, with a pair of subjects being assigned to each of the six possible sequences of the three viewing environments.

RESULTS

Psychometric functions were determined for each viewing condition for each subject by a least-squares fitting procedure that estimated the point of subjective equality (PSE), and the just noticeable difference (JND). The former reflects individual subjects' biases in determining the perceptual equivalence of differing physical latencies; the latter reflects subjects' sensitivity to latency differences.

As can be seen in Figure 3 and Table 1, the fits of the psychometric functions modeled as cumulative gaussians were generally quite good. The overwhelming majority of fits for most subjects were statistically highly with the poorer fits being distributed roughly evenly across the experimental conditions.

When subjected to a two-way parametric ANOVA of sequence and viewing condition, the two estimated parameters showed no statistically significant main effects or interaction (JND: Viewing condition: $F_{2,14} = 0.36$, ns; Sequence: $F_{5,7} = 0.69$, ns.; Condition X Sequence: $F_{10,14} = 0.61$, ns. PSE: Viewing condition: $F_{2,14} = 0.47$, ns; Sequence: $F_{5,7}=1.668$, ns.; Condition X Sequence: $F_{10,14} = 0.97$,ns.). The lack of a main effect of condition was also confirmed by a Friedman ANOVA (JND: $\chi^2_2 = 4.2$, ns; PSE: $\chi^2_2 = 2.0$, ns.).

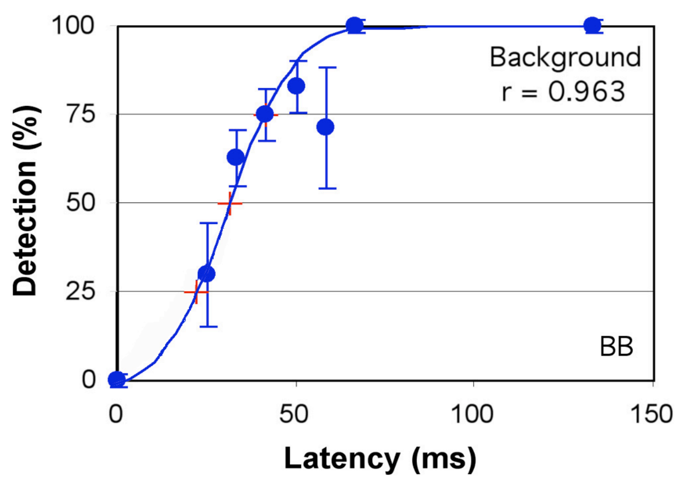

Figure 3. Psychometric function for the detection of latency differences for one subject under the *Background* condition.

Subject	Correlation with best fitting Ogive		
	Objec	*Background*	*Both*
AG	**0.993**	**0.955**	**0.933**
BB	**0.992**	**0.964**	0.670
BM	**0.878**	**0.933**	**0.947**
CM	**0.981**	**0.992**	**0.998**
EF	**0.965**	**0.984**	**0.843**
JA	0.698	*0.742*	**0.935**
KM	**0.988**	**0.990**	*0.714*
LL	**0.871**	**0.932**	**0.980**
MS	**0.969**	*0.788*	**0.961**
ME	0.529	**0.937**	0.280
PH	**0.992**	**0.919**	**0.987**
SF	**0.997**	**0.990**	**0.994**
TL	**0.987**	**0.985**	**0.991**

Table 1. Correlations between the standard cumulative gaussian model for a psychometric function and subject's data for each condition. (**p < 0.01**, $p < 0.05$, ns.)

Additionally, there was no effect of order as determined by an independent one-way ANOVA (JND: $F_{2,24} = 1.757$, ns.; PSE: $F_{2,24} = 0.05$, ns.). All subjects' individual parameter estimates and group medians and inter-quartile ranges (IQR) corre-

[3] The end-to-end VE system latency was verified for all environments using the method described by Hill, Adelstein, and Ellis, (2004)).

sponding to these analyses' viewing conditions are presented in Figures 4 and 5. Note that were the precision of the estimates (corresponding standard errors of means) for each condition to be plotted, they would in general be hard to see on a scale large enough to present all of the individual subject data.

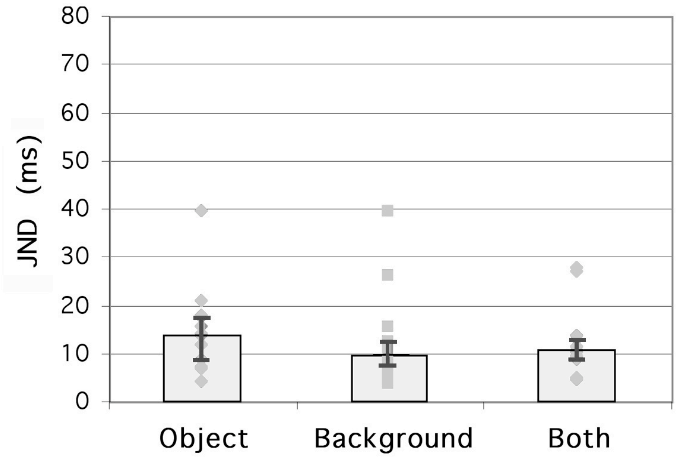

Figure 4 Group mean just noticeable differences (bars), interquartile ranges (error bars) and individual data points.

DISCUSSION

Our results present a precise and striking contrast with those of Allison *et al*. (2001), especially for the *Background* condition that is the replication of their experiment. Our much more detailed results show that subject sensitivity and bias to latency change to be quite acute. In particular, our measured bias reflected in the PSE is much lower than Allison et al.'s reports of 200 ms and higher. Furthermore, our results are consistent with our previous studies and those of Regan et al. (1999), whose experiments were conducted with a very different visual display format that presented space-stabilized virtual objects via desktop CRT.

We believe that there are several possible explanations for our subjects lower bias (PSE). One is that we made sure through training that they understood what latency "looked like" in the display before we began the experiment. Additionally, our subjects had prolonged exposure spread across several sessions to the experiment conditions. In Allison et al.'s study, subject exposure was limited to only two blocks on two consecutive days (Allison et al, personal communication). Finally, our experiment employed two-interval psychophysical judgments in which subjects always compared a probe stimulus against a constant, low latency reference. Allison et al. used a single interval task to elicit an evaluation of whether the virtual scene appeared stable with respect to whatever preconceived notion of stability was held by subjects without prior HMD-based VE experience (Allison, personal communication)

We found that unpracticed subjects initially required careful description to understand exactly what latency meant visually, but that once they understood it, they were sensitive observers. The reliability of our subjects is confirmed by the precision of our population parameter estimates and by the absence of sequence or order effects in our data. We also can see that increasing the complexity of the environment with the combination of a foreground with a background did not increase the subjects' sensitivity to detecting a latency change. This equality suggests that the shear evident in the *Both* condition does not necessarily aid latency discrimination.

Some insight into why the added complexity in the image combining a foreground and background may be found in comments solicited from subjects regarding how they made their discriminations. Several subjects remarked in a subsequent study (Mania et al., 2004) that they focused on one feature in the display such as a particular corner or edge, and then judged its stability during their head movements. Clearly, such a strategy could bypass information about appropriate movement timing by the presentation of shear.

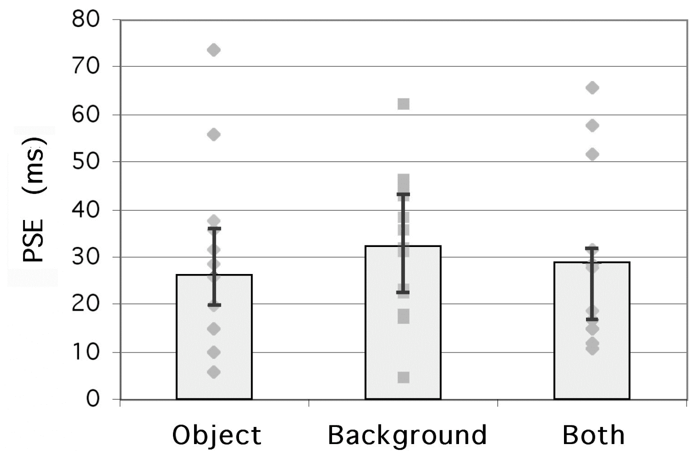

Figure 5. Group mean points of subjective equality (bars), interquartile ranges (error bars) and individual data points.

It is also possible that the head rocking used in the experiment did not produce sufficient lateral translation of the eye's nodal points to create enough additional shear movement. In planning the experiment, we verified that the placement of the test object 57 cm in front of the background would produce easily visible relative movement during head rotation. Since subjects learned to avoid the maximum rotation cue by rotating less than the maximum 30° allowed, our estimates were that a typical head movement during the experiment would have produced a relative rotation of ~20° .At the radius of the background surface, this rotation would produce apparent relative motion equal to the width of a rectangle at the equator of the sphere's checkerboard pattern (Figure 6.) While a background with finer details might have made the relative motion more salient, we are certain for the conditions used that the motion was well above threshold visibility and would have increased the amount of apparent movement associated with system latency.

It should be noted additionally that the head movements studied in this experiment may not have been optimized for latency detection. During initial data collection, one subject not following instructions, had to be reminded not to use a deliberately jerky head movement. Analysis of these data, which were later replaced by rerunning the conditions, showed that when he used these jerky movements, he had a dramatically better JND than the otherwise. Consequently, we expect that subjects using such head movements might be able to discriminate smaller latency differences that those we report. While complete psychometric functions were not determined in their study, several such very short JND's below 10 ms have already been reported by Regan *et al*. (1999).

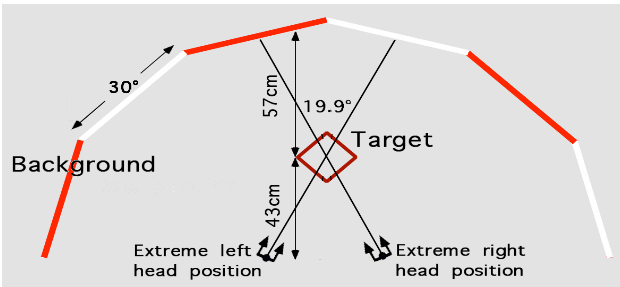

Figure 6. Illustration of maximum relative motion of target due to estimated average actual head rocking during the experiment.

Another incidental observation leads us to an additional caveat concerning our results. Our observations were made with the subject's full focal attention directed towards judging the visual stability of specific display elements. On one occasion during pilot testing, one of the authors (KM) was acting as a subject but also discussing the experiment and other unrelated matters during testing. The experiment monitor noticed that when talking she quite often did not report latency differences that she otherwise rarely missed. Accordingly, we believe latency changes may be far harder to detect when visual attention is distracted.

Finally, it should be noted that our study of the role of increased spatial complexity could be pursued with a more complicated, photorealistic environment with many objects and depth planes. Such as study has been completed and will be reported elsewhere (Mania, Adelstein, Ellis, & Hill, 2004). It confirms that increased scene complexity does not affect latency sensitivity for the kind of focused attention discrimination that we have been testing.

ACKNOWLEDGEMENTS

This work was supported by a grant from NASA Headquarters Code UB, the Advanced Human Support Technology program.

REFERENCES

Adelstein, B.D., Lee, T.G., Ellis, S R.(2003) Head tracking latency in virtual environments: psychophysics and a model. *Proceedings, 47th Annual Meeting HFES*, pp. 2083-2987.

Adelstein, B.D., Johnston, E.R., Ellis, S. R.(1996) Dynamic response of electromagnetic spatial displacement trackers. *Presence: Telepresence and Virtual Environments*, *5*(3), 302-318.

Allison, R.S., Harris, L.R., Jenkin, M., Jasiobedzka, U., & Zacher, J.E. (2001) Tolerance of temporal delay in virtual environments, *Proceedings, IEEE Virtual Reality 2001*, pp. 247-253.

Azuma, A., & Bishop, G. (1994) Improving static and dynamic registration in an optical see-through display. *Proceedings, SIGGRAPH*, pp. 197-204.

Ellis, S.R., Bréant, F., Menges, B.M., Jacoby, R.H. & Adelstein, B.D. (1997) Operator interaction with virtual objects: effects of system latency. *Proceedings, HCI'97 International*. San Francisco, CA., pp. 973-976.

Ellis, S.R., Young, M.J., Adelstein, B.D., & Ehrlich, S.M. (1999b) Discrimination of changes in latency during head movement. *Proceedings, Computer Human Interfaces,* pp. 1129-1133.

Ellis, S.R., Young, M.J., Ehrlich, S.M., & Adelstein, B.D. (1999a) Discrimination of changes of rendering latency during voluntary hand movement. *Proceedings, 43rd Annual Meeting HFES*, pp. 1182-1186.

Frank, L.H., Casali, J.G., & Wierwill, W. (1988). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. *Human Factors*, *30*, 201-217.

Held, R., Efsathiou, A., & Greene, M. (1966) Adaptation to displaced and delayed visual feedback from the hand. *J. Exp. Psychology*, *72*(6), 887-891.

Hill, M.I., Adelstein, B.D., & Ellis, S.R. (2004) Achieving minimum latency in virtual environment applications. *Proceedings, IMAGE Society Annual Meeting*, Scottsdale AZ.

Jacoby, R.H., Adelstein, B.D., & Ellis, S.R. (1996) Improved temporal response in virtual environments through system hardware and software reorganization. *Proceedings, SPIE Conference on Stereoscopic Displays and Applications VI*, *Vol. 2653,* pp. 271-284.

Jung, J.Y., Adelstein, B.D., & Ellis, S.R. (2000) Discriminability of prediction artifacts in a time delayed virtual environment. *Proceedings, 44th Annual Meeting HFES*, pp. 499-502.

Levitt, H. (1970) Transformed up-down methods in psychoacoustics. *J. Acoustical Soc. of America, 49*(2), 467-477.

Mania, K., Adelstein, B.D., Ellis, S.R., & Hill, M.I. (2004) Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. *Proceedings, First ACM Symposium on Applied Perception in Graphics and Visualization*, Los Angeles, August 7-8, 2004.

Matin, L., Picoult, E., Stevens J. K., Edwards, M.W. Jr., Young, D., & MacArthur, R (1982) Oculoparalytic illusion: visual field dependent spatial mislocationzation by humans partially paralyzed with curare. *Science, 216,* 198-201.

McCandless, J.W., Adelstein, B.D., Ellis, S.R. (2000) Localization of a time-delayed monocular virtual object superimposed on a real environment. *Presence: Teleoperators and Virtual Environments*, *9*(1), 15-24.

Meehan, M., Razzaque, S., Whitton, M., Brooks, F. (2003). Effect of latency on presence in stressful virtual environments. In *Proceedings, IEEE Virtual Reality 2003*, pp. 141-148.

Regan, M., Miller, G., Rubin, S., & Kogelnik, C. (1999) A real-time low-latency hardware light-field renderer. *Proceedings, SIGGRAPH*, Los Angeles, CA, pp. 287-290.

Sheridan, T.B., & Ferrell, W.R. (1963) Remote manipulative control with transmission delay. *IEEE Transactions Human Factors in Electronics, HFE-4(*l), 25-29.

Smith, W.M., McCrary, J.R. & Smith, K.U. (1962) Delayed visual feedback and behavior. *Science*, *132*, 1013-1014

Stark L. & Bridgeman, B. (1983) Role of corollary discharge in space constancy, *Perception and Psychophysics, 43*, 371-380.

Watson, A. B., & Fitzhugh, A. (1990). The method of constant stimuli is inefficient. *Perception & Psychophysics, 47*(1), 87-91.

Watson, B., Walker, N., Woytiuk, P., Ribarsky, W. (2003). Maintaining usability during 3D placement despite delay. *Proceedings, IEEE Virtual Reality 2003*, pp.133-140.