An Evaluation of Input Devices for 3-D Computer Display Workstations

Robert J. Beaton, Richard J. DeHoff, Novia Weiman, and Peter W. Hildebrandt

Imaging Research Laboratory, Tektronix Laboratories, P.O. Box 500, Mail Stop 50-320
Tektronix, Inc., Beaverton, Oregon 97077

Abstract

This paper reports several results from an on-going research program designed to examine the utility of alternate input device technologies for 3-dimensional (3-D) computer display workstations. In this paper, operator performance levels on a 3-D cursor-positioning task were compared using three input devices: (1) a trackball that allowed unrestricted (i.e., free-space) movements within the display space, (2) a mouse that provided selectable two-axis (i.e., plane) movements, and (3) a set of thumbwheels that provided separate controls for orthogonal single-axis (i.e., vector) movements. In addition, the input device evaluation was conducted for two operationally distinct 3-D display techniques: (1) a linear perspective encoding of image depth information and (2) a field-sequential stereoscopic encoding of depth information. Results are discussed in terms of input device selection and general design considerations for the user interface to 3-D computer workstations.

Introduction

There is growing interest among vendors and users of high-performance computer workstations in the use of 3-D information display systems for several widely disparate application areas. These range from engineering and scientific data analysis systems through business and graphical arts systems to computer-assisted manufacturing and inspection systems. Much of this interest stems from the increased demand by users of these extremely powerful, yet relatively inexpensive, workstations for enhanced fidelity in graphical renditions of complex data structures. Because of the vast "local" processing capabilities afforded by these workstations, operators often are required to use and visually interpret data structures too complex for meaningful presentation on conventional 2-dimensional display systems. Examples of a few application areas requiring efficient display of complicated data structures include: (1) thermal and stress analysis of mechanical components, (2) multidimensional statistical and signal-processing applications, and (3) computer modeling of real-world objects. It is a widely held position that 3-D display techniques, employing various combinations of linear perspective, intensity shading, or true volumetric display formats, may provide the additional "information" dimensions to operators for improved productivity.

The successful integration of 3-D workstations into existing as well as many new applications areas will depend upon several additional facets of the entire workstation system. It has been shown repeatedly in the marketplace for traditional computer-based workstations that design and implementation of the user interface directly influences wide-spread acceptance of the product concept. Although human factors experts have identified numerous components of the user interface, a most critical component involves the "ease-of-use" or efficiency of communication between the operator and the controlling hardware, firmware, or software logic of the workstation. This statement implies that a complete, well-designed workstation must incorporate additional tools for the user that simplify user-machine communication protocols. Hence, a large and varied gamut of input devices has become available for conventional information display workstations.

Due to the potential of input devices to add significant value to computer-based workstations, considerable efforts have been invested by human factors researchers to optimize input device designs. However, much of the previous research has been confined to user interactions with conventional 2-dimensional display systems, and very few published reports pertaining specifically to 3-D display systems are available. Therefore, the present research attempts to fill a noticeable void in the human factors literature by contributing a systematic evaluation of three currently-popular input device technologies adapted for use in 3-D computer display workstations. In addition, it is recognized that several 3-D display techniques exist for presenting image depth information. Therefore, to provide widely applicable results, input device comparisons were performed for two prominent techniques: a linear perspective 3-D display and a field-sequential stereoscopic 3-D display.

Methods

Equipment

Input devices. Three input device technologies were investigated in this evaluation: (1) a trackball (Measurement Systems, Model 636-G1535), (2) a mouse (Mouse Systems, Model M3), and (3) a custom-designed set of thumbwheels (Tektronix, prototype).

The trackball device was mounted in a small tabletop cabinet. The forward-backward and sideways motions of the trackball were mapped into the usual 2-dimensional display movements of up-down (Y-axis display-space coordinate) and left-right (X-axis display-space coordinate) on the display screen, respectively. Display screen movements in depth (Z-axis display-space coordinates) were accomplished by rotary motions of the trackball about its vertical axis. In this manner, the trackball device provided simultaneous three-axis or "free-space" movements on the display screen.

The mouse device also was designed for tabletop operation and was configured functionally to provide simultaneous movements in two of the three display-space coordinate axes; that is, by pressing one of three push-button switches on the device housing, subjects activated cursor movements on either the X-Y, X-Z, or Y-Z display planes. In the more familiar X-Y mode, left-right and up-down screen movements were accomplished by sideways and forward-backward motions of the mouse, respectively. However, in the X-Z mode, the forward-backward motions of the mouse corresponded to movements along the displayed depth dimension and, in the Y-Z mode, left-right sideways motions of the mouse corresponded to screen movements in depth. Given these functional characteristics, the operational mode of the mouse device may be described as providing users the capability for "plane" movements on the display screen.

The final input device examined consisted of a set of three thumbwheels, mounted in a small tabletop cabinet and arranged physically to correspond to the orientation of the display-space coordinate system. That is, the rotational dimension of one thumbwheel was aligned to the left-right or X-axis of the display space, the rotational dimension of a second thumbwheel was aligned to the up-down or Y-axis of the display space, and the third thumbwheel was aligned to the in-out or Z-axis of the display space. The entire set of thumbwheels was sized appropriately for single-handed operation, usually by the user's thumb, forefinger and little finger. Since three separate thumbwheels were used for 3-D cursor positioning, the operational mode of this device may be described as providing "vector" movement capabilities.

In accord with arguments made previously in the human factors literature on input device evaluation, [1,3,4,5,6] an attempt was made to optimize the control-to-display ratio of each input device prior to conducting the experiment. In this manner, optimal configurations of each input device technology were compared. Since the experimental task described below primarily involved fine adjustments of cursor position, it was sufficient to employ a single-valued control-to-display ratio determined separately for each device.

3-D display system. The computer workstation used was based on a field-sequential stereoscopic display technology that is extensively described elsewhere.[7] The display system consisted of a high-resolution color CRT (Mitsubishi, Model C-9918N, 0.31 mm mask pitch, P-22 phosphor set), an active liquid crystal polarizing filter (Tektronix) with accompanying passive polarizing eye spectacles, and a microcomputer equipped with a dual-image frame buffer (Tektronix, Model 3-D/AT). The frame buffer consisted of two separate image fields, each containing 512 by 512 by 4-bit pixels, and it operated the CRT display at a non-interlaced refresh rate of 120 Hz per image field. Although only one image field was displayed at a time (i.e., during the vertical raster period), each image field was displayed during alternate vertical raster periods. In other words the two fields were multiplexed successively in time. The frame buffer also controlled the active liquid crystal filter, which was mounted over the CRT display faceplate. With this display system configuration it was possible to select the polarization angle of light emitted from each image field. Thus, when the display screen was viewed through the spectacles, which incorporated asymmetrical polarizers for each eye, the left eye of the subject received one image field and the right eye received the other image field. The refresh rate of the display was sufficient to avoid temporal flicker.[2]

Stimulus conditions. On each trial in the experiment, the subject viewed a 3-D workspace that resembled a large volumetric cube inscribed within the boundaries of the CRT display screen. From the subjects viewpoint, the front and rear sides of the workspace were transparent, whereas the left-right and top-bottom sides of the workspace were outlined by a wire-frame of narrow lines. The lines forming the wire-frame were intensity-encoded to demarcate three distinct depth planes. The wire-frame lines closest to the observer were brighter than those lines at the central depth plane of the display space, which, in turn, were brighter than those wire-frame lines furthest away.

The display screen also contained a diagonally-oriented "crosshair" cursor and a wire-frame rectangular box centered around a single pixel target. On each trial, the cursor initially appeared at the center of the 3-D workspace; however, the location of the target box was determined randomly as described below. To help the subject locate the position of the single pixel target in depth, "shadow lines" from the rectangular box were projected onto the left-right and top-bottom sides of the workspace.

Twenty-six target locations were used in each session. Target locations were determined by combining three planar positions, one from each orthogonal plane (i.e., X-Y, X-Z, and Y-Z planes) in the 3-D workspace. From each plane, eight positions at a constant distance around the midpoint were sampled. Combining these planar positions yielded an assortment of target locations at a constant 3-D distance from the initial position of the cursor, but that required either a direct single-axis (i.e., vector) cursor movement, a two-axis (i.e., plane) cursor movement, or a three-axis (i.e., free-space) cursor movement. This construction of target locations was intended to facilitate evaluation of the various operational modes inherent to the input devices examined.

For the perspective display conditions, the depth dimension of the workspace was encoded by a linear transformation that assumed a fixed observer viewing distance of 122 cm. All stimulus displays under this condition were composed of two identical image fields in the dual-image frame buffer.

For the stereoscopic display conditions, the perspective encoding of depth within each image field was augmented by a binocular disparity transformation. This transformation used an average intraocular separation of 6.35 cm and was computed separately for each image field in real-time by the frame buffer.

Procedure. A 3-D cursor-positioning task was devised for the experiment. Each experimental trial began with the subject viewing the crosshair cursor through the front, transparent side of the workspace. Once the subject pressed a "Start-Stop" key, the rectangular box surrounding the single pixel target appeared on the display screen. Then, using the input device, the subject manipulated the cursor through the 3-D workspace until it overlaid the single pixel target. After positioning the cursor, the subject pressed the "Start-Stop" key again to indicate the end of the trial. Following each trial the cursor positioning time and positioning error was computed and recorded. Positioning time (in milliseconds) was defined as the interval between the onset of the target box and the termininating key press. Positioning error (in graphical coordinate system units) was defined as the 3-D distance between the target location and the intersecting lines of the crosshair cursor. No feedback regarding the subjects positioning time or positioning error was given at any time during the trials.

Before beginning the experimental trials for each 3-D display condition, the subject received a series of practice trials in which the stimulus conditions closely resembled those employed for the actual trials. The purpose of practice was three-fold: (1) to familiarize subjects with the use of a particular input device under each display condition, (2) to ensure high levels of task proficiency during the experimental trials, and (3) to minimize the "carry-over" effects from participation in prior sessions of the experiment. The practice was conducted in blocks of 10 trials, and the subject was required to meet preselected performance criteria for positioning time and positioning error. If necessary, the practice trial blocks were repeated up to three times. Subjects were dismissed from the entire evaluation if they did not meet all performance criteria.

For each subject the entire experiment was conducted in three separate sessions. All subjects used the trackball device in their first session, the mouse device in their second session, and the thumbwheels in their third session. Within each session a subject completed the cursor-positioning task for one display condition before starting the same task for the other display condition. Presentation order of the two display conditions was counterbalanced across subjects. Under both display conditions the cursor-positioning task was replicated twice and the presentation order of target locations was randomized uniquely for each subject (except for the constraint that all 26 target locations were presented before any repeated observations).

Subjects. Sixteen individuals (2 females) employed by Tektronix, Inc. volunteered to participate in the evaluation. Although each individual worked daily with computer-based workstations, no one was experienced with the use of 3-D display systems. Over the course of the experiment it was necessary to dismiss two subjects; one individual was unable to complete all experimental sessions due to conflicting personal priorities, the other experienced difficulty achieving prerequisite performance criteria. For the remaining 14 participants, the experimental sessions were completed over a 3 month period. Each session, including practice, experimental trials, and rest breaks required 1–2 hours.

Results and Discussion

Three separate statistical analyses were performed on the data obtained in the evaluation. In the first two, positioning time and positioning error responses were subjected to separate Analysis of Variance (ANOVA) procedures. In the third analysis, to assess the effects of the experimental conditions upon a composite measure of operator performance, a simple linear combination of positioning time and error was formulated and analyzed by an ANOVA procedure. Results from these statistical analyses are described briefly below.

Positioning error

Mean cursor positioning errors, averaged over repeated observations for each observer, were computed for 30 experimental conditions. These consisted of the factorial combination of three levels of input device (i.e., trackball, mouse, and thumbwheels), two levels of display mode (i.e., perspective vs. stereoscopic), and five levels of movement (i.e., vector, X-Y plane, X-Z plane, Y-Z plane, and free-space). Levels of the movement factor were derived from the 26 target locations used in the experimental trials. Since the movement factor resulted in an unbalanced design (vector, N=6; X-Y plane, N=4; X-Z plane, N=4; Y-Z plane, N=4; free-space, N=8), a general linear models solution to the three-factor within-subjects ANOVA procedure was used.[8] The main effects of input device { $F(2,26)=7.59, p=0.0025$ } and display mode { $F(1,13)=41.75, p=0.0001$ } were significant, as was the two-factor interaction between input device and display mode { $F(2,26)=3.86, p=0.0339$ }. None of the remaining ANOVA effects were significant.

Figure 1 illustrates the main effect of input device. A post-hoc Newman-Keuls test indicated that positioning error associated with the thumbwheel device was lower than that associated with trackball and mouse ($\alpha=0.05$). It is clear from the figure that the thumbwheels led to more than a two-fold increase in positioning accuracy compared with either the trackball or the mouse.

Figure 2 shows the main effect of display mode, in which the positioning error associated with the field-sequential stereoscopic display is about 60% less than of the positioning error associated with the perspective display.

The presence of a two-factor interaction between input device and display mode indicates that the relative differences among the three input devices varied differentially across the two display conditions. The mean positioning errors observed in stereoscopic and perspective display conditions, respectively, were 4.0 and 8.4 for the thumbwheels, 7.5 and 19.0 for the trackball, and 8.2 and 19.1 for the mouse. In general, therefore, positioning accuracy for the three input devices increased about 51–60% with the stereoscopic display compared to the perspective display. The highest positioning accuracy was obtained with the thumbwheels.

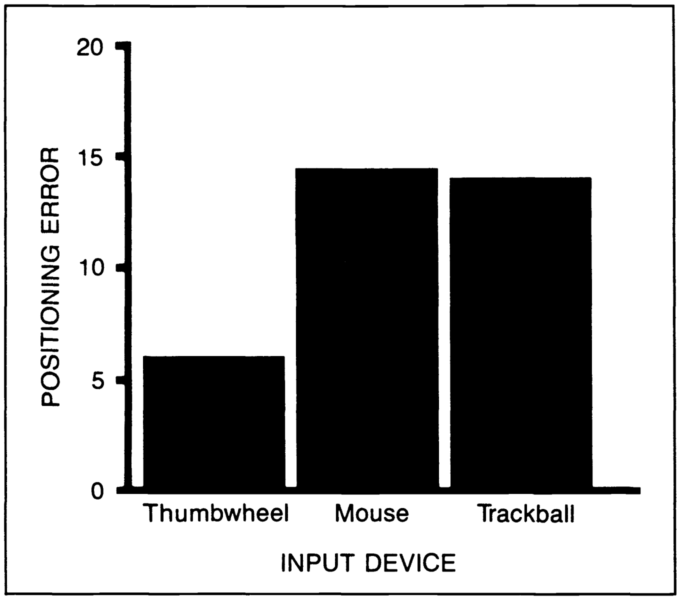

Figure 1. Main effect of input device, averaged over 14 subjects, 2 display modes, and 5 movements, on cursor positioning error. Error values are expressed as 3-D distances in graphical coordinate system units.

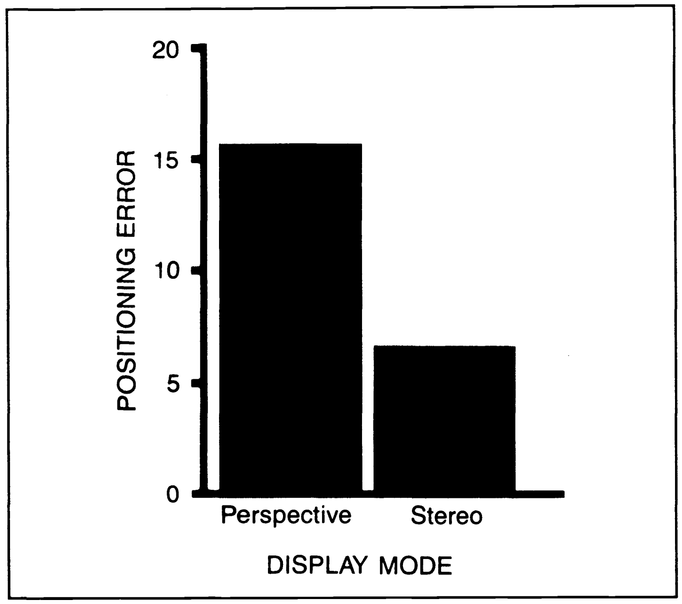

Figure 2. Main effect of 3-D display mode, averaged over 14 subjects, 3 input devices, and 5 movements, on cursor positioning error. Error values are expressed as 3-D distances in graphical coordinate system units.

Positioning time

Mean cursor positioning times, averaged over repeated observations for each subject, were computed for the same 30 experimental conditions and were subjected to the same general linear models ANOVA procedure as used in the preceding positioning error analysis. The main effects of input device { $F(2,26)=9.74, p=0.0007$} and movement { $F(4,52)=20.95, p<0.0001$} were significant. Additionally, several two-factor interaction effects were significant: the input-device-by-display-mode interaction { $F(2,26)=2.93, p=0.0710$}, the input-device-by-movement interaction { $F(8,104)=2.33, p=0.0242$}, and the display-mode-by-movement interaction { $F(4,52)=7.23, p<0.0001$}. No other ANOVA effect was significant.

Figure 3 shows the main effect of input device. A Newman-Keuls test indicated that positioning time associated with the mouse was longer than positioning times associated with the trackball and thumbwheels ($\alpha=0.05$). As shown in the figure, cursor positioning with either the trackball or thumbwheels was about 23% faster than positioning with the mouse.

Figure 4 illustrates the main effect of movement. Results of a Newman-Keuls test showed that positioning times for vector and X-Y plane movements were faster than those for X-Z plane movements which, in turn, were faster than those for either the Y-Z plane or free-space movements ($\alpha=0.05$). Although similar positioning times for vector and X-Y plane movements were expected (since these movements occur on conventional 2-dimensional displays), the long positioning times associated with Y-Z plane movements was not anticipated. Some insight into the nature of this data trend was revealed by the significant two-factor interactions.

In general, the input-device-by-movement interaction and the display-mode-by-movement interaction do not alter the overall performance trends observed. However, the input-device-by-movement interaction indicated that positioning time with the mouse on the Y-Z plane required considerably longer than did any other combination of input device and movement (i.e., 32 vs. 14–23 seconds, respectively). Also, the display-mode-by-movement interaction indicated that Y-Z plane positioning under the field-sequential stereoscopic display condition was slower than the other combinations of display mode and movement (i.e., 29 vs. 18–25 seconds, respectively). Both of these findings point out a general difficulty associated with rapid cursor movements on the Y-Z display plane. It is possible that operational characteristics of the 3-D mouse account for these findings since depth movements on the Y-Z plane were accomplished by incongruent left-right motions. Moreover, the influence of this "cognitive" inconsistency upon subject's performance may have increased under the field-sequential stereoscopic condition due to enhanced fidelity in the portrayal of the 3-D workspace. This contention is supported by the input-device-by-display-mode interaction, in which positioning time for the mouse under the field-sequential stereoscopic condition was longer than any other device display combination (i.e., 28 vs. 19–24 seconds, respectively).

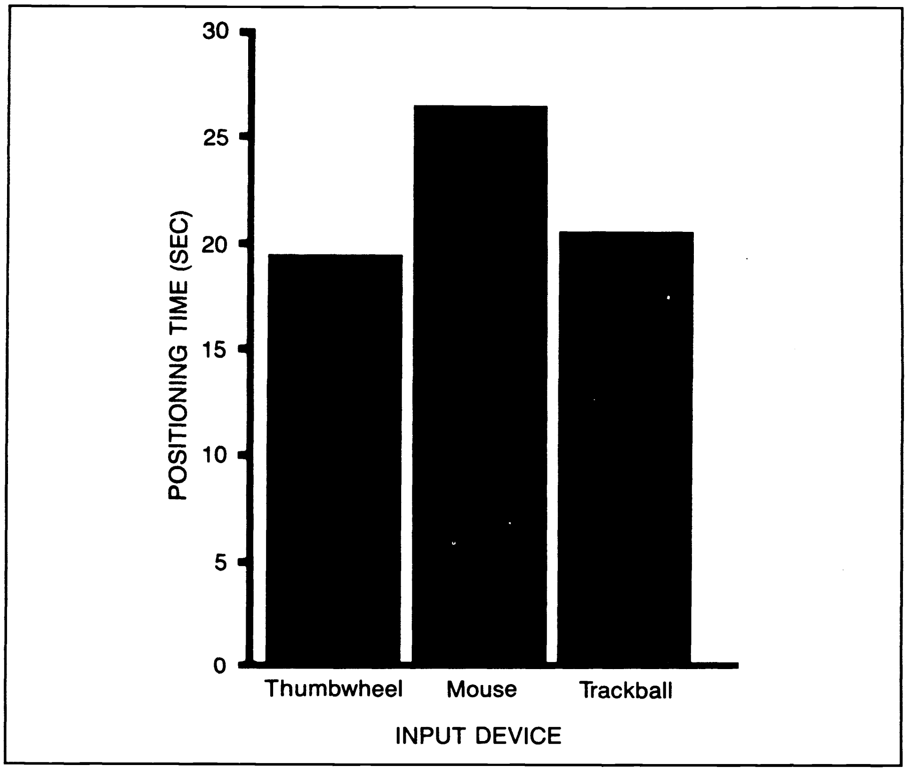

Figure 3. Main effect of input device, averaged over 14 subjects, 2 display modes, and 5 movements, on cursor positioning time.

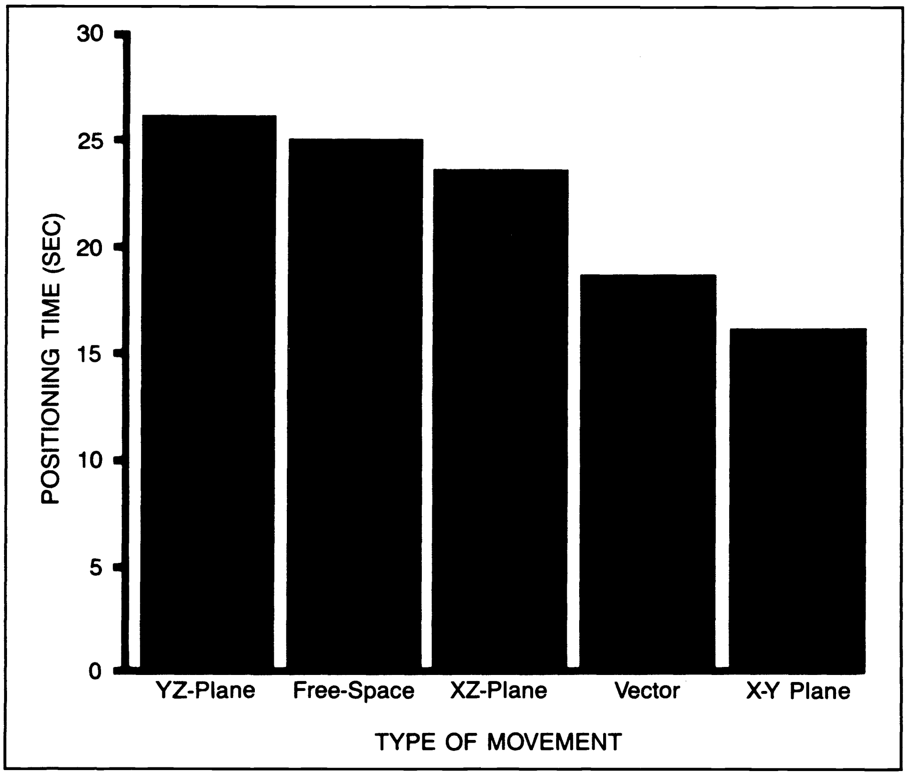

Figure 4. Main effect of movement, averaged over 14 subjects, 2 display modes, and 3 input devices, on cursor positioning time.

Performance

In a real-world task environment, the concept of operator performance is characterized by numerous parameters. Although only two parameters (positioning time and error) were examined in this evaluation, an assessment of performance, "operationalized" by a composite index, was desired. Therefore, based on the viewpoint that operator performance is negatively correlated with both positioning error and positioning time, the following index was formulated:

$$Performance = \frac{1}{\beta_0 Error + \beta_1 Time}, \quad (1)$$

in which the constraint

$$\sum_{i=0}^{i=N} \beta_i = 1 \quad (2)$$

allows proportional weighting of time and error, as well as additional factors if included for other purposes. In the present analysis, $\beta_0 = \beta_1 = 0.5$ since positioning time and positioning error were believed to be equally important.

Mean performance scores, averaged over repeated observations for each subject, were computed for the 30 experimental conditions described earlier in the positioning time and positioning error analyses, and the mean performance scores were analyzed with the same ANOVA model. The main effects of input device { $F(2,26)=4.85, p=0.0162$}, display mode { $F(1,13)=7.26$, $p=0.0184$}, and movement { $F(4,52)=31.41, p<0.0001$} were significant, as were the two-factor interactions between input device and display mode { $F(2,26)=2.89, p=0.0735$} and between input device and movement { $F(8,104)=8.23, p<0.0001$}. The remaining ANOVA effects were not statistically significant.

Figure 5 shows the performance scores for the main effect of input device. A Newman-Keuls test indicated that performance was about 20% higher with the thumbwheels than with either the trackball or mouse ($\alpha=0.05$). Although this effect was uncovered in the positioning time and positioning error analyses, the present performance analysis provides a more conservative estimate of its relative magnitude.

The main effect of display mode is illustrated in Figure 6. This figure clearly indicates that overall performance was approximately 25% higher in the field-sequential stereoscopic display condition than in the perspective display condition. Note that the magnitude of the performance analysis result is more conservative than the corresponding finding from the error analysis (see Figure 2). As a technical aside, conservative estimates of human performance provided by the proposed performance index can be desirable for certain cost-benefit analyses since decisions based on the magnitude of an effect are less likely to reflect measurement error.

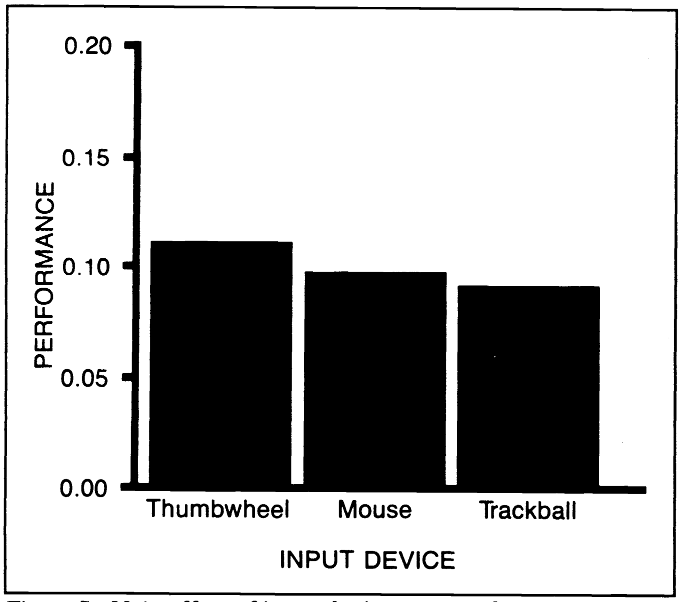

Figure 5. Main effect of input device, averaged over 14 subjects, 2 display modes, and 5 movements, on composite performance.

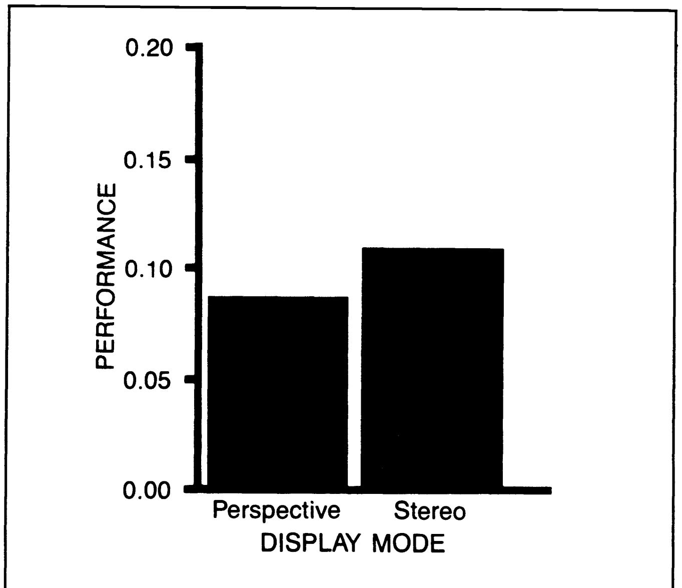

Figure 6. Main effect of 3-D display mode, averaged over 14 subjects, 3 input devices, and 5 movements, on composite performance.

Figure 7 shows the main effect of movement on cursor positioning performance. A Newman-Keuls test found that performance during vector and X-Y plane movements exceeded performance during either X-Z plane, Y-Z plane, and free-space movements ($\alpha=0.05$). Interpretation of this finding is straightforward compared to the corresponding result from the positioning time analysis. As shown in the figure, those movement conditions requiring cursor motions in depth were associated with relatively low performance levels. Due to the indicated magnitude of this effect (i.e., $\approx 50\%$), careful consideration by designers of user interfaces to 3-D display systems is warranted to avoid operator task dependencies involving accurate cursor positioning in depth.

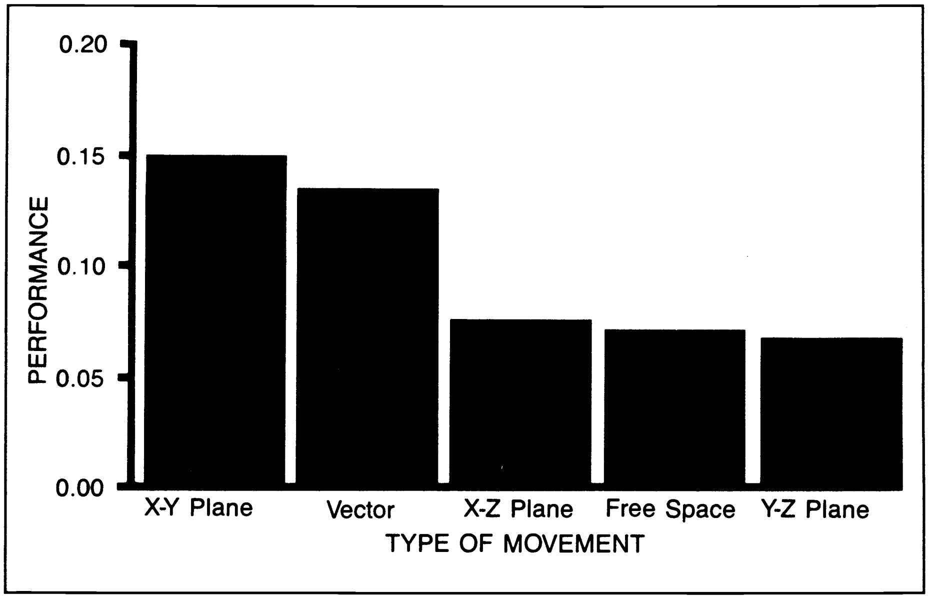

Figure 7. Main effect of movement, averaged over 14 subjects, 2 display modes, and 3 input devices, on composite performance.

The two-factor interaction effect between input device and display mode is illustrated in Figure 8. A Newman-Keuls test indicated that this interaction effect was attributable to the relatively high performance levels obtained with the trackball device under the stereoscopic display condition ($\alpha=0.05$). However, overall performance trends (observed in the main effects of input device and display mode) remain unaltered.

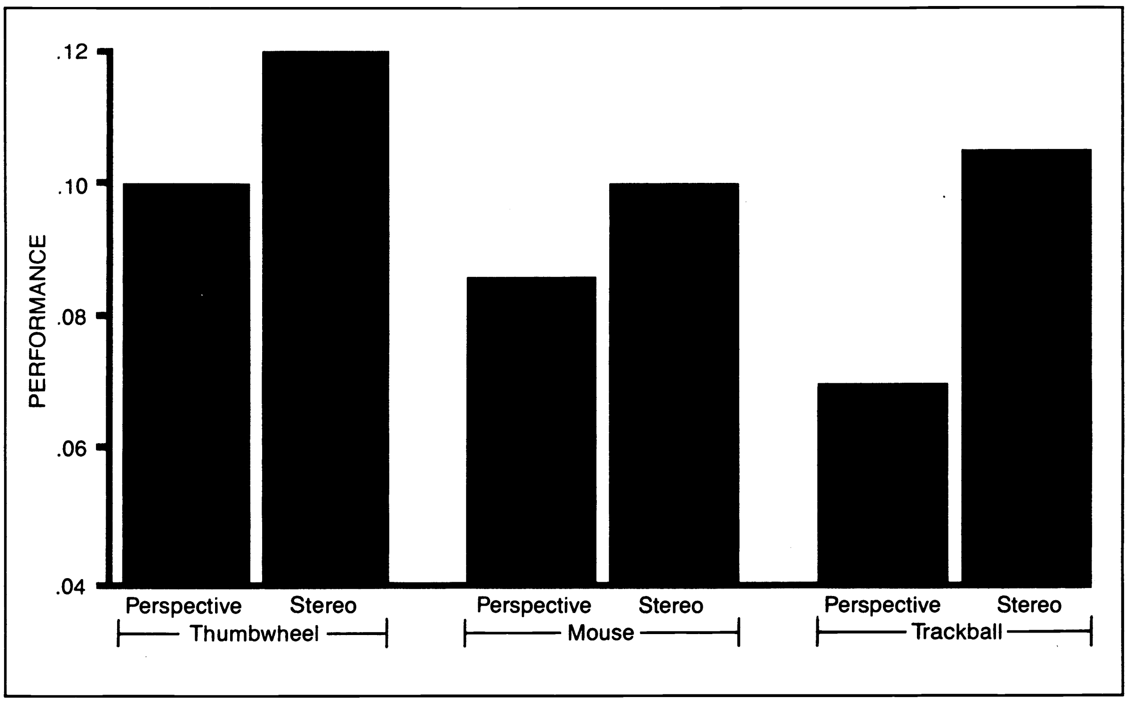

Figure 8. Two-factor interaction effect of input device and 3-D display mode, averaged over 14 subjects and 5 movements, on composite performance.

Finally, Figure 9 shows the two-factor interaction between input device and movement. Results of a Newman-Keuls test indicated that the overall effect of movement resembled the trend shown in Figure 7, but that the ordering of input devices (in terms of performance) differed across the movement levels. For vector movements, the thumbwheels produced about 54% higher performance than either the mouse or trackball. For X-Y plane movements, performance was about 45% higher with the thumbwheels and mouse than with the trackball. For X-Z and Y-Z plane movements (as well as for free-space movements), performance was about 25% higher with the thumbwheels and trackball than with the mouse. These trends indicate that input devices based on vector and plane operational modes produce higher performance for single- and two-axis cursor positioning, respectively. However, these trends also indicate that the free-space mode of operation does not provide a practical advantage over the vector mode of operation for three-axis cursor positioning.

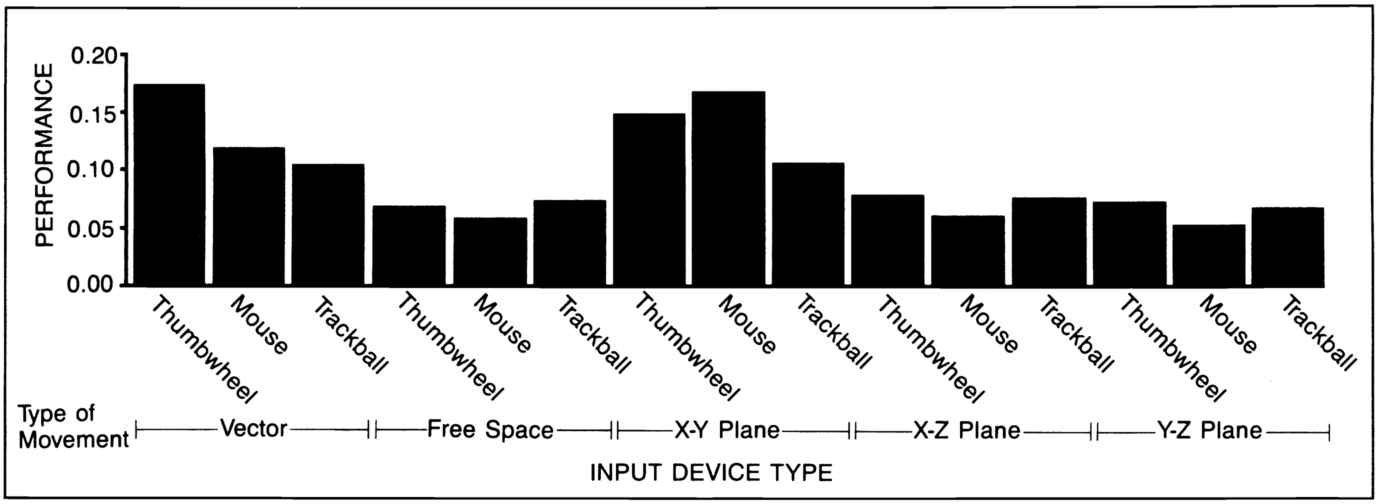

Figure 9. Two-factor interaction effect of input device and movement, averaged over 14 subjects and 2 display modes, on composite performance.

Conclusions

Results of this evaluation support several conclusions regarding the selection of input devices for 3-D computer workstations and, in general, concern design principles for the construction of user interfaces to these workstations.

First, it is clear that the thumbwheels were the "best" device overall for cursor positioning across the stereoscopic and perspective 3-D display conditions, at least in terms of positioning accuracy, positioning time, and the composite performance index. Performance comparable to the thumbwheels was obtained with the mouse only during X-Y plane movements and with the trackball only during free-space movements.

Second, regardless of the input device or 3-D display technique used, cursor positioning in depth was more difficult than conventional X-Y plane cursor positioning. This statement applies to simple vector movements in depth, such as moving directly toward or away from the observer, as well as to two- and three-axis movements. From a practical applications point of view, this suggests that compensatory mechanisms need to be developed and incorporated into the design of 3-D display systems. Perhaps changes in target size and use of extensive visual feedback may prove to be necessary features of 3-D task environments in which operators are expected to perform cursor positioning equally well across all display dimensions.

Third, since cursor positioning performance consistently was higher under the field-sequential stereoscopic display conditions compared to the perspective display conditions, it is reasonable to conclude that the former display provided a more efficient or "easier-to-use" interface to the workstation system. The enhancement of operator performance with the stereoscopic display most likely is due to the high-fidelity rendition of the 3-D workspace. Therefore, as a general guideline, field-sequential stereoscopic displays are preferable to perspective displays for presentations of complex 3-D data structures.

References

1. Arnaut, L.Y. and Greenstein, J.S., Two experiments investigating the effects of control-display gain and method of cursor control on touch tablet performance. *Proceedings of the Human Factors Society—29th Annual Meeting*, pp. 122-126, 1985.

2. Beaton, R.J., DeHoff, R.J., and Knox, S.T., Revisiting the display flicker problem: Refresh rate requirements for field-sequential stereoscopic display systems. *Proceedings of the Society for Information Display*, Vol. XVII, pp. 150-152, 1986.

3. Beaton, R.J. and Weiman, N., Effects of touch key size and separation on menu-selection accuracy. *Proceedings of the Society of Photographic Scientists and Engineers*, Advances in Display Technology V, Vol. 526, pp. 8-12, 1985.

4. Becker, J.A. and Greenstein, J.S., A lead-lag compensation approach to display/control gain for touch tablets. *Proceedings of the Human Factors Society—30th Annual Meeting*, pp. 332-336, 1986.

5. Epps, B.W., Comparison of six control devices based on Fitts' law models. *Proceedings of the Human Factors Society—30th Annual Meeeting*, pp. 327-329, 1986.

6. Epps, B.W., Snyder, H.L., and Muto, W.H., Comparison of six cursor devices on a target acquisition task. *Proceeding of the Society for Information Display*, Vol. XVII, pp. 302-305, 1986.

7. Haven, T.J., A liquid crystal video stereoscope with high extinction ratios, a 28% transmission state, and 100 μs switching. *Proceeding of the Society of Photographic Instrumentation Engineers, O-E/LASE'87*, 1987.

8. Winer, B.J., *Statistical principles in experimental design*. (Second Edition), McGraw-Hill: New York, pp. 402-430, 1971.