Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces

Ravin Balakrishnan[1,2] and Gordon Kurtenbach[2]

[1]Dept. of Computer Science
University of Toronto
Toronto, Ontario
Canada M5S 3G4
ravin@dgp.toronto.edu

[2]Alias|wavefront
210 King Street East
Toronto, Ontario
Canada M5A 1J7
{ravin | gordo}@aw.sgi.com

ABSTRACT

We explore the use of the non-dominant hand to control a virtual camera while the dominant hand performs other tasks in a virtual 3D scene. Two experiments and an informal study are presented which evaluate this interaction style by comparing it to the status-quo unimanual interaction. In the first experiment, we find that for a target selection task, performance using the bimanual technique was 20% faster. Experiment 2 compared performance in a more complicated object docking task. Performance advantages are shown, however, only after practice. Free-form 3D painting was explored in the user study. In both experiments and in the user study participants strongly preferred the bimanual technique. The results also indicate that user preferences concerning bimanual interaction may be driven by factors other than simple time-motion performance advantages.

Keywords

Bimanual input, 3D interfaces, camera control, interaction techniques, empirical evaluation

INTRODUCTION

Several user interface researchers over the past decade, having recognized that in the physical world people often use both hands to cooperatively perform many tasks, have explored the possibility of using both hands simultaneously in the computer interface. In an early study, Buxton and Myers [4] showed that in a compound task, a one-handed interface (i.e. the status-quo) was inferior to a two-handed interface which split the compound task into two subtasks that could be performed in parallel by both hands. Kabbash, Buxton, and Sellen [13] came to a similar conclusion, however, they also showed that two hands could be worse than one if an inappropriate interaction technique is employed, particularly when cognitive load is increased.

Building partly on this empirical work, several researchers have demonstrated systems with compelling bimanual interfaces for both 2D [2, 15] and 3D [5, 9, 17, 19, 23] applications. However, apart from some fundamental work by Hinckley [9, 10, 11], little formal evaluation of bimanual 3D interfaces has been carried out.

Published in Proceedings of 1999 ACM Conference on Human Factors in Computing Systems (CHI'99), pp 56-63.

In this paper, we describe and evaluate a bimanual interaction technique for desktop 3D graphics applications which not only increases the input control bandwidth but also enhances user perception of the virtual 3D scene. Essentially, we propose using the non-dominant hand to operate the virtual camera controls typically found in 3D graphics applications, thus freeing the dominant hand to perform other manipulative tasks in the 3D scene. Other researchers [5, 9, 19, 23] have demonstrated camera operations using the non-dominant hand but have either done so in concert with the dominant hand (i.e., both hands are used to specify camera parameters) [5, 23] or attempted to directly mimic the real world [9, 17], using higher (>2) degree-of-freedom input devices more suited to virtual reality applications. We focus our attention on mouse and keyboard based *desktop* 3D environments which form the basis of current commercial 3D graphics applications for modeling, design, and animation.

In order to motivate our work, we first briefly review the various depth cues used in 3D displays, followed by a discussion of a current theoretical model of bimanual interaction. We then discuss how one of the most powerful 3D depth cues can be enhanced by following the principles of this bimanual interaction model. Our proposed bimanual interaction technique is then evaluated for a range of typical 3D tasks.

BACKGROUND

Depth Cues in Virtual 3D Scenes

3D graphics applications typically utilize a variety of depth cues to enhance user's perception of the virtual 3D scene. These cues, whose origins can be traced to the human visual perception literature, include *perspective, occlusion* or *interposition, light* and *shadows, relative size, textual gradient, proximity-luminance covariance, relative motion gradient, retinal binocular disparity,* and *motion parallax* (see [8, 22] for a review).

In 3D graphics, *perspective* is one of the most commonly employed cues, as evident in the ubiquitous wireframe "groundplane" present in most 3D applications. Also important are *occlusion* cues which are implemented via hidden line and surface removal techniques. *Lights* and *shadows* are less frequently used in interactive 3D graphics because of the high computational cost involved, although this is changing with ever faster graphics engines. *Stereopsis*, which results from retinal binocular disparity, is a strong depth cue and has been investigated extensively in the virtual reality domain, but is not commonly used in desktop 3D graphics because of the need for expensive and cumbersome viewing apparatus. Also, based on an extensive review of the role of the various depth cues in 3D perception and 3D display design, Wickens et. al. [21, 22] concluded that while stereopsis, motion, and occlusion are all salient cues, *motion* (e.g., the kinetic depth effect [3, 6] generated when the user's view of the scene is continuously varied by manipulating the virtual camera) is particularly important in creating a sense of three-dimensionality since stereopsis may provide no benefit when motion cues are present. Other evidence [3, 6] also demonstrate and emphasize the importance of motion cues.

A Model of Bimanual Interaction

Much recent work in bimanual user interfaces [2, 5, 10, 11, 13, 15, 16] has been guided by the theoretical work of Guiard [7]. In his Kinematic Chain (KC) model of skilled bimanual action, the two hands are thought to be two abstract motors assembled in a serial linkage, thus forming a cooperative kinematic chain. Three general principles emerge from this model:

1. *Dominant-to-Non-Dominant Spatial Reference:* The non-dominant hand sets the frame of reference relative to which the dominant hand performs its motions.

2. *Asymmetric Scales of Motion:* The two hands operate in asymmetric spatial-temporal scales of motion. For instance, when writing on a piece of paper, the motion of the non-dominant hand controlling the position of the paper is of lower temporal and spatial frequency than the writing movements of the dominant hand which nonetheless depends on the non-dominant hand's movement for spatial reference.

3. *Precedence of the Non-Dominant Hand:* Contribution of the non-dominant hand to a cooperative bimanual task starts earlier than the dominant hand. In the handwriting example, the dominant hand starts writing *after* the paper has been oriented and positioned by the non-dominant hand.

This model has been explored and largely validated in the virtual manipulation arena by Hinckley [10, 11]. Leganchuk, Zhai, and Buxton [16] also used this model to help reason about the manual and cognitive benefits they found in an experimental study on bimanual input.

ENHANCING DEPTH PERCEPTION VIA BIMANUAL INTERACTION

In desktop 3D graphics applications, moving the virtual camera enables the user to view different parts of the 3D scene. In addition to the obvious purpose of bringing once occluded objects into the forefront, camera manipulation also serves a less obvious but very important purpose: enhanced depth perception through motion. As discussed earlier, this motion depth cue, called the kinetic depth effect [3, 6], is critical in enabling the user to accurately perceive the virtual 3D scene. As Kirsh and Maglio have described [14], humans perform actions not only to bring them closer to the physical goals of a task (*pragmatic action*), but also to facilitate perception and cognition (*epistemic action*). Thus, one finds users of unimanual interfaces to 3D graphics applications constantly switching between the epistemic action of camera manipulation for depth perception and pragmatic actions to perform manipulative tasks on objects in the scene. Based on these theories and observations, it is likely that allowing users to perform the pragmatic actions via one input stream (i.e., the mouse in the dominant hand as in the status quo) while the often epistemic actions of camera control are performed via a second input stream (i.e., an input device in the non-dominant hand) will result in both improved time-motion task performance and an enhanced sense of perception (or sense of engagement) of the 3D scene. This style of interaction also squares nicely with Guiard's KC model.

In order to explore the benefits of using the non-dominant hand to operate camera controls in typical 3D tasks, we conducted two formal experiments and one informal user study. In addition to the primary goal of quantitatively and qualitatively evaluating this style of interaction, we also wanted to explore how performance and user preference changed as the complexity of the task increased.

This is the first of a series of planned experiments in this area. At this early stage, we are mainly concerned with how users perform when the operation of camera controls are moved from the dominant hand to the non-dominant hand. While there are several camera control metaphors commonly used in 3D graphics applications, we chose to do all our experiments using one typical metaphor. The issue of which camera control metaphors are better suited to the non-dominant hand, or if several control techniques can be interchangeably used, is left for later investigation. Similarly, numerous different input devices could conceivably be used in either hand. We chose to use a standard two degree-of-freedom mouse in each hand for several reasons. First, the mouse is the status-quo input device for the dominant hand in desktop 3D graphics applications (see [1] for a discussion of why the mouse dominates, despite the availability of higher degree-of-freedom input devices). Second, this is a reasonable configuration for a practical, low cost bimanual interface. Third, using a mouse in both hands means that our experiments measure only the effects of moving camera controls to the non-dominant hand and are not confounded by participants having to learn to use an unfamiliar input device.

EXPERIMENT 1: SELECTION

To begin our evaluation of non-dominant hand camera control, we felt it would be best to start with a simple canonical task, and if the results were promising, we could then move on to more complex tasks. Accordingly, we chose 3D target selection as our first experimental task. Target selection is one of the simplest tasks typically used in studying human performance in computer input control. Other typical tasks like object docking, path following, and pursuit tracking, are considerably more difficult.

Method

Task and Stimuli

Participants were asked to select targets which appeared on the surface of a large cubic object in the 3D scene. As illustrated in Figure 1 (colours in the figure have been changed to accommodate greyscale printing), the scene consisted of the cubic object in the centre of the display and a light grey wireframe grid at the bottom of the display. The purpose of this grid (often called the "groundplane" in 3D graphics parlance) was to provide an additional perspective depth and occlusion cue. The cubic object was an opaque, pink coloured Gouraud shaded cube whose faces were divided into nine equal sized square sections. The target to be selected was a flat, yellow coloured disk which appeared on one of the nine sections of five faces of the cubic object (4 side and 1 top face; the bottom face of the cubic object was not used since one would have to look through the groundplane to view that face). Thus, there are 9x5=45 different locations where the target could appear. Since the cubic object was opaque, not all of its faces are visible in a given view. In order to see the other faces in search of the target, the view of the scene had to be changed by manipulating the virtual camera. To further encourage camera manipulation, "raised walls" were placed on the boundaries around the nine sections of each face of the cubic object. These "raised walls" obscured the sections such that one had to view a section almost "head on" to see if a target was on it, thus necessitating frequent camera movement.

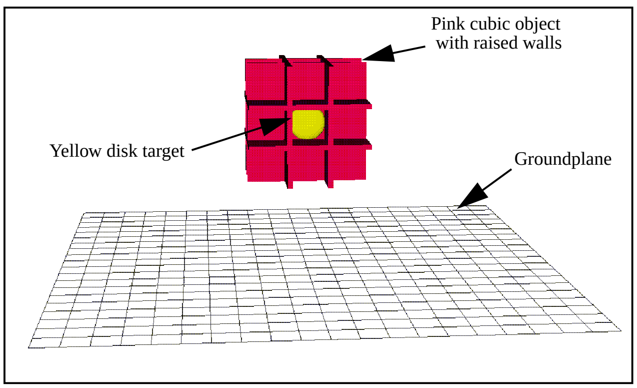

Figure 1. Stimuli for Experiment 1.

The camera control metaphor used is often referred to as "tumbling" the camera, and is analogous to holding and manipulating a turntable (represented by the groundplane in the graphics scene) in one's hand. The turntable can be rotated about its normal axis as well as the horizontal screen axis. Technically, this requires revolving the camera about the centre of the scene by varying the azimuth and elevation angles in the perspective view. This allows objects in the middle of the 3D scene to be viewed from any direction. The viewing distance from the object, as well as the view angle (or focal length of the camera) is kept constant. This camera control metaphor is ideal when the object(s) of interest are located, as in this experiment, in the centre of the 3D scene. It is one of the most frequently used camera controls in mainstream 3D applications, others such as panning (moving the centre of interest), zooming/dolly (moving closer or further away from the centre of interest) are important but less frequently used when working on a single object in the scene.

Selection of the target was done by using a mouse to move a 2D selection cursor in the plane of the screen such that the cursor was over the target (in line of sight) and clicking the left mouse button. This "ray casting" method of selecting 3D targets using a 2D cursor is widely employed in 3D graphics applications and has been shown to be superior to selection using 3D cursors [12, 20]. If the target was successfully selected, it disappeared and a new target appeared 500ms later at another location. Errors could not occur since the next target would not appear until the current one had been selected. The participant thus had to manipulate the camera to locate the target, and then select the target using the selection cursor.

The experiment compared task performance using a one-handed (1H) vs. a two-handed (2H) technique. In the 1H technique, participants used their dominant hand to operate a mouse which controlled both the selection cursor and the camera. Clicking on the target selected the target, clicking and dragging anywhere else in the scene moved the camera in the appropriate direction. Thus, participants had to constantly switch between camera control and selection in order to perform the task.

In the 2H technique, participants used their dominant hand to operate a mouse which controlled the selection cursor (as in the 1H technique), while their non-dominant hand operated a second mouse which controlled the camera. In this case, both the camera and the selection cursor could be operated simultaneously. There was no cursor attached to the non-dominant hand mouse. Also, no button presses were required since it was permanently attached to controlling the camera.

Experimental Hypotheses

Our hypotheses were developed from our informal early prototype use of the non-dominant hand for camera control and the formal framework provided by Guiard's KC model. The experimental task using the 2H technique nicely adheres to all three principles of the KC model: 1) moving the camera sets the frame of reference for the selection cursor to select the target; 2) camera control is a coarse grain task, whereas selection is a fine grain task; 3) the camera movement must precede selection. With the 1H technique however, the dominant hand has to perform both camera control and selection - constantly switching between them. Accordingly, we hypothesize that:

H1: The 2H technique will be faster than the 1H technique, primarily because the mode switching time present in the 1H technique is eliminated in the 2H technique. While it is true that in the 2H technique the participant has to switch between using the dominant hand and the non-dominant hand, this switching time should be negligible compared to that of the 1H technique because the non-dominant hand is "ready to go" the moment the dominant hand has completed its task and vice-versa.

H2: Participants will subjectively prefer the 2H technique since it more closely follows their natural real world expectations of holding an object in one hand and manipulating it with the other hand.

Apparatus

The experiment was conducted on a Silicon Graphics Indigo2 Extreme workstation with a 19 inch colour display. Two standard serial/PS2 mice set to the same gain were used as input devices. The workstation ran in single-user mode, disconnected from all network traffic.

Participants

10 right-handed volunteers participated in the experiment.

Design

A within subjects repeated measures design was used. All participants performed the experiment using both techniques (1H and 2H). The presentation order of the two techniques was counterbalanced across the participants. For each technique, participants performed 3 blocks of trials. Each block consisted of 1 trial for each of the 45 possible positions that a target could appear on the cubic object, presented in a constrained pseudorandom order within the block. The constraint imposed was that the target always appeared on a different face of the cubic object from the previous target. This ensured that participants had to manipulate the camera in order to select each target. In target selection experiments, the size of the target is typically manipulated as an experimental factor. However, in pilot testing of our experiment, we found that target size had no effect on the relative performance between the 1H and 2H techniques (i.e., there was no Target Size x Technique interaction). Therefore, we used a single target size in this experiment.

Participants were given eight practise trials to familiarize themselves with the task. They were allowed breaks after each block of 45 trials. The experiment consisted of 2700 total trials, as follows:

10 participants x

2 techniques (1H and 2H) x

3 blocks of trials for each technique x

45 trials per block

= 2700 total trials.

For each subject, the experiment was conducted in one sitting and lasted under half an hour. Subjects were alternately assigned to one of two experimental orders: 1H technique followed by 2H (1H/2H) or 2H first (2H/1H).

A short questionnaire designed to elicit participants' subjective preferences for the two techniques was completed by participants at the end of the experiment.

Results and Discussion

Trial Completion Time

Figure 2 compares participants' mean trial completion time for both techniques over the three blocks of trials. Trial completion time was measured beginning when the target first appeared on the cubic object and ending when the target was selected. Repeated measures analysis of variance with trial completion time as the dependent variable was conducted on the data. As hypothesized (H1), a significant

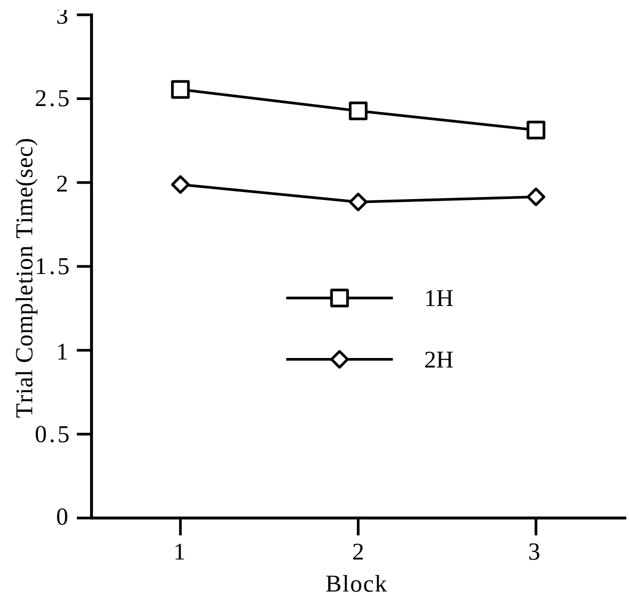

Figure 2. Experiment 1: Mean trial completion time for both techniques over the course of three experimental

main effect was found for the technique used (1H or 2H) ($F_{1,8}$ = 17.62, $p < .01$). Overall, the 2H technique was 20% faster than the 1H technique.

The order of presentation (1H/2H or 2H/1H) had no significant effect ($F_{1,8}$ = 0.12, $p > .5$). This, coupled with the absence of any Technique x Order interaction ($F_{1,8}$ = 2.78, $p > .1$), effectively rules out the possibility of asymmetrical skill transfer – an often overlooked artifact of within-subjects designs [18]. Learning across the three blocks of trials was not significant ($F_{2,16}$ = 3.50, $p > .05$). This supports our observations during the experiment that the task was elemental enough that participants had little difficulty performing the task quickly right from the beginning. No other significant interactions were observed.

Subjective Evaluation

At the end of the experiment, participants were asked to rate their preference for each technique on a scale of -2 (very low) to 2 (very high). The results, summarized in Table 1, validate our second hypothesis (H2) and is consistent with the quantitative trial completion time data.

Technique \ Rating	-2 very low	-1 low	0 ok	1 high	2 very high
1H technique (mean score: -0.4)		6	3		1
2H technique (mean score: 1.6)			1	2	7

Table 1. Subjective preferences in Experiment 1. Each cell contains the number of subjects with that rating.

EXPERIMENT 2: DOCKING

Experiment 1 showed that operating camera controls in the non-dominant hand is beneficial in a 3D selection task. However, the selection task was relatively lightweight in terms of both motor and cognitive effort required of the participant. Few epistemic actions were required to get an understanding of the 3D scene. An obvious question, therefore, is whether similar benefits can be realized in a more demanding task. To answer that question, we ran a second experiment using 3D object docking as the experimental task.

Method

Task and Stimuli

The 3D object docking task required participants to select an object in one corner of the virtual 3D scene and place it inside a target object located at the diagonally opposite corner.

As shown in Figure 3 (colours have been changed to accommodate greyscale printing), the scene consisted of two objects and a groundplane (identical to the one used in experiment 1) in the middle of the virtual scene. The object to be manipulated was a blue coloured sphere. The target was a purple cube with translucent faces. Colours and transparency effects were chosen to ensure that participants were not hindered in their task by insufficient visual cues. The manipulated object was two thirds the size of the target object.

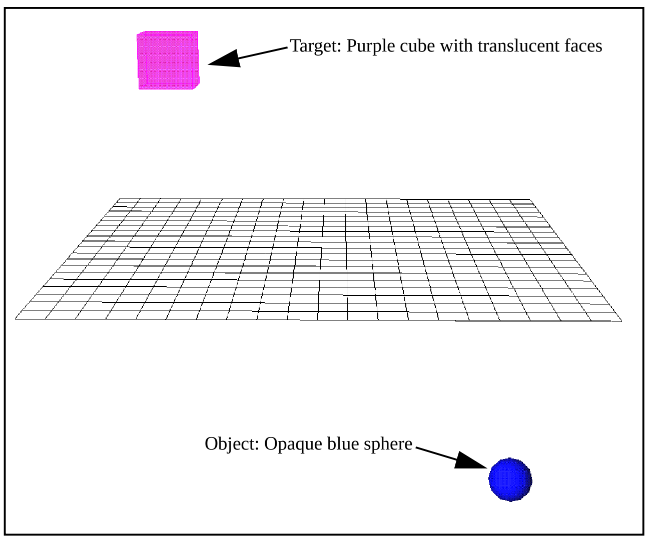

Figure 3. Stimuli for Experiment 2.

As in Experiment 1, we compared task performance using a one-handed (1H) vs. a two-handed (2H) technique. In the 1H technique, participants used their dominant hand to operate a mouse which controlled both the selection cursor and the camera. Clicking and dragging on the object selected and moved the object; clicking and dragging anywhere else in the scene moved the camera in the appropriate direction. When selected, the object could be moved in two dimensions at a time, always parallel to the plane of the screen (i.e., in the screen's x-y plane). In order to move the object along the z-axis in the virtual scene, the camera ideally has to move 90 degrees such that the virtual scene's z-axis became parallel to the screen's x or y axis. This "screen space" or "image plane" style of object movement is commonly employed in 3D graphics applications which use the 2 degree-of-freedom mouse as the primary input device. It works reasonable well, but as discussed in the introduction, requires constant switching between camera control and object manipulation in order to move an object in 3D space.

In the 2H technique, participants selected and manipulated the object with the dominant hand mouse, while the non-dominant hand operated a second mouse which controlled the camera. In this case, both the camera and the object could be manipulated simultaneously. As a result, it becomes possible to move the object into the target in a single movement if the non-dominant hand controlling the camera can coordinate its movements with the dominant hand controlling the object (one way of visualizing this movement is to think of the camera being moved such that the target is being brought closer to the viewer, while the object is also being moved such that it is also being brought closer to the viewer. At some point in the middle, the object and target will meet). Of course, an alternate strategy is to simply move the camera first, followed by the object, and keep alternating between the two until the task is completed. This is similar to the strategy that has to be used in the 1H technique, except that no explicit switching of modes from camera control to object manipulation is required in the 2H technique since each task is assigned to a different hand.

The camera control metaphor was identical to that used in Experiment 1.

When the object was within the target's boundaries, the target turned bright green. Participants released the dominant hand left mouse button while the object was within the target to indicate completion of a trial.

Experimental Hypotheses

Our hypotheses were developed from the results of Experiment 1, and once again the formal framework of Guiard's KC model. If the experimental task using the 2H technique is performed one hand at a time (asymmetric interaction), it adheres to all three principles of the KC model. The results of Experiment 1 indicates that this will outperform the 1H technique. However, if the task is performed by moving both hands simultaneously (symmetric interaction), it may no longer be conceptually perceived as "move camera, then move object"; rather it becomes move camera (or effectively, move the target) and object simultaneously. Although Guiard's KC model does not address the issue of symmetric interaction, we nonetheless expect to see some performance improvement over the 1H technique if this strategy is employed.

Formally, we hypothesize that:

H1: Regardless of the manipulation strategy used, the 2H technique will be faster than the 1H technique, primarily because the mode switching time present in the 1H technique is eliminated in the 2H technique.

H2: Participants will subjectively prefer the 2H technique since it (a) more closely follows their natural expectations for performing these types of tasks in the real world, and (b) lowers the cost of performing epistemic actions, thus providing a greater sense of "engagement" with the virtual world.

Apparatus

The apparatus was identical to that used in Experiment 1.

Participants

10 right-handed volunteers participated in the experiment. Prior to participating in this experiment, they all participated in Experiment 1. Any skill transfer from Experiment 1 to Experiment 2 should therefore be symmetrical for all subjects and not adversely affect the validity of Experiment 2.

Design

A within subjects repeated measures design was used. All participants performed the experiment using both techniques (1H and 2H). The presentation order of the two techniques was counterbalanced across the participants. For each technique, participants performed 5 blocks of trials. Each block consisted of eight conditions presented at random: we tested participants' ability to move an object from each of the eight corners of the virtual scene's viewing volume to a target located at the diagonally opposite corner. Subjects performed four trials for each of the eight conditions.

Prior to performing the experiment with each technique, participants were shown how to do the task using that technique. For the 2H technique, they were shown how to do the task by simultaneously moving both hands, and also by moving one hand at a time. Participants were given two practice trials for each condition to familiarize themselves with the task. They were allowed breaks after each set of four trials per condition. After completion of a trial, there was a 500ms pause before the next trial began.

The experiment consisted of 3200 total trials, as follows:

10 participants x

2 techniques (1H and 2H) x

5 blocks of trials for each technique x

8 conditions per block x

4 trials per condition

= 3200 total trials.

The experiment was conducted in one sitting and lasted under an hour per subject. Subjects were alternately assigned to one of two experimental orders: 1H technique followed by 2H (1H/2H) or 2H first (2H/1H).

A short questionnaire designed to elicit participants' subjective preferences for the two technique was completed by participants at the end of the experiment.

Results and Discussion

Trial Completion Time

Figure 4 compares participants' mean trial completion time for both techniques over the five blocks of trials. Trial completion time was measured beginning when the object and

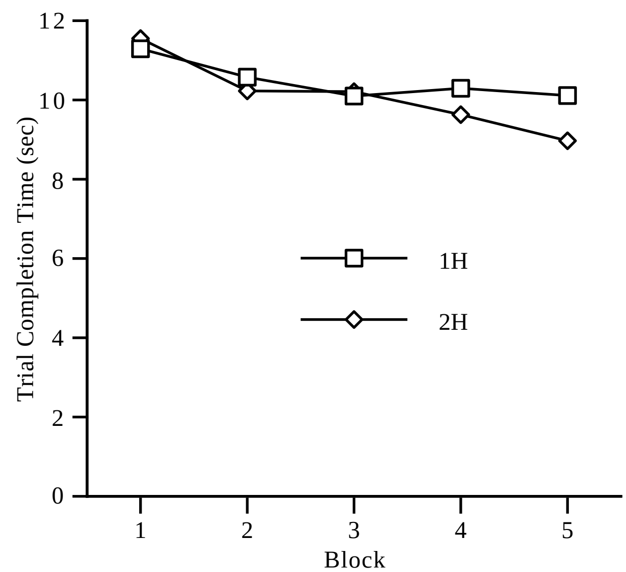

Figure 4. Experiment 2: Mean trial completion time for both techniques over the course of five experimental blocks. Data from all 10 participants.

target first appeared in the scene and ending when the object was successfully placed in the target. Repeated measures analysis of variance with trial completion time as the dependent variable was conducted on the data. Overall, there was no significant difference between the two techniques (1H or 2H) ($F_{1,8} = 0.70$, $p > .1$). This is a somewhat surprising result, especially given the significant performance gains observed in Experiment 1 for the 2H technique. Possible explanations for this result can be found in two observations we made while participants were performing the experiment as well as our own experience with the task.

First, we observed that participants were largely trying to use both hands simultaneously in the 2H technique. When the task is performed in this symmetric manner, it appears to become more difficult than the 1H technique. There are three likely reasons for this: 1) both the target and object have to be monitored continuously, dividing attention and increasing the cognitive load on the participant; 2) four degrees-of-freedom – two controlling the object, two controlling the camera – have to be simultaneously controlled, increasing the load on the participant's motor system; and 3) the geometric transformation that has to be mentally computed in order to bring object and target together is non-trivial, especially for the novice user. The 1H technique, in contrast, time-multiplexes between controlling the camera and controlling the object. This imposes a lighter cognitive and motor load at any one time. From our results, it is clear that the sum of the two subtasks (symmetric strategy in 2H technique) has a greater cost than its parts (1H technique). As noted in the introduction, Kabbash, Buxton, and Sellen [13] also found that increased cognitive load resulted in reduced performance time in some bimanual tasks.

A second observation was that because in the 2H technique there was no explicit switching cost involved in manipulating the camera, participants tended to perform more epistemic actions than in the 1H technique. While this results in participants getting a better perception of the 3D scene, the time cost incurred adds to the overall time taken to perform the pragmatic task of placing the object in the target. In a sense, while the design of the 2H technique was motivated by the desire to facilitate epistemic actions, it appears that in certain situations too much of a good thing can be bad!

Given these observations, the temporal performance result is not surprising. If participants had performed the task asymmetrically (i.e., move camera, then move object) and/or with fewer epistemic actions, we might have seen a performance gain similar to that obtained in Experiment 1. Moving to a more parallel, symmetric style of interaction, as well as performing more epistemic actions, clearly results in a performance cost in the pragmatic task. However, as expert users of our experimental system, we found that we could perform the task using the 2H technique in a symmetric manner much faster than using the 1H technique.

Now, the question is whether the experimental data supports our personal experience that symmetric 2H performance improves with practice. Further data analysis showed a significant learning effect across the five blocks of trials ($F_{4,32} = 17.52$, $p < .001$). By the time participants reached the last block of trials (block 5), the difference between the two techniques became statistically significant ($F_{1,8} = 5.72$, $p < .05$), thus indicating that as participants get more expert at the task, the cognitive and motor loads discussed earlier are reduced. In terms of magnitude of difference, in block 1 the 2H technique was marginally (2%) slower than the 1H technique, while in block 5 the 2H technique was 11% faster than the 1H technique. No other significant interactions were observed in the data analysis.

Subjective Evaluation

As in Experiment 1, at the end of Experiment 2 participants were asked to rate their preference for each technique on a scale of -2 (very low) to 2 (very high). The results, summarized in Table 2, shows that despite their relatively poor initial temporal performance with the 2H technique, participants strongly preferred it over the 1H technique. This validates our second hypothesis (H2).

Technique \ Rating	-2 very low	-1 low	0 ok	1 high	2 very high
1H technique (mean score: -0.5)	1	5	3		1
2H technique (mean score: 1.2)		1		5	4

Table 2. Subjective preferences in Experiment 2. Each cell contains the number of subjects with that rating.

INFORMAL STUDY: PAINTING

In Experiments 1 and 2 we formally studied users' performance using 1H and 2H techniques for 3D selection and docking tasks. Another task that could benefit from non-dominant hand camera manipulation is 3D painting (projective paint or paint on surface) or sculpting. Several commercially available packages (e.g., Amazon's 3Dpaint, Alias|wavefront's Maya) provide 3D painting/sculpting functionality, but generally use the dominant hand for both camera control and painting. We feel that moving the camera controls to the non-dominant hand would provide a greater sense of directness to the task, and also facilitate epistemic actions that enable better visualization of the painting/sculpture being created. Unfortunately, painting or sculpting are tasks where obtaining quantitative performance metrics is difficult. Thus, we informally asked five volunteers who had experience with 3D paint packages to try out a simple 3D painting system we developed. They were asked to paint a "cartoonized" head onto a plain 3D sphere, and to do it with 1H and 2H techniques in turn. In the 1H technique, the dominant hand used a pen on a digitizing tablet to paint on the sphere as well as to control the camera (the "ALT" key on the keyboard was held down to switch into camera control mode – this is the status-quo technique used in commercial packages). In the 2H technique, a mouse in the non-dominant hand controlled the camera while the dominant hand painted using the digitizer pen. The camera control metaphor was identical to that used in Experiments 1 and 2.

The participants were asked to rate their preference for the two techniques. They overwhelmingly preferred the 2H technique (Table 3), despite the fact that they all had prior experience with the 1H status-quo technique. Comments included "I feel like I'm really painting on the sphere", and "wish I had this in Maya".

Technique \ Rating	-2 very low	-1 low	0 ok	1 high	2 very high
1H technique (mean score: -1.8)	4	1			
2H technique (mean score: 2)					5

Table 3. Subjective preferences in painting study. Each cell contains the number of subjects with that rating.

CONCLUSIONS

Our experiments and informal study have shown that having the non-dominant hand operate a subset of possible camera controls in 3D graphics interfaces can be beneficial over a range of tasks. The results of Experiment 2, however, caution that when the interaction style deviates from Guiard's KC model and both hands begin to operate in a symmetric manner, temporal benefits may not be immediately apparent. Of particular interest is the strong preference shown by participants for the two-handed technique regardless of their temporal performance in the task. Because subjective preferences cannot be quantified as reliably as, say, time-motion performance, less weight tends to be placed on such data. While there is a possibility that some of this subjective data suffers from the "good participant" effect (where participants will rate highly experimental conditions which they perceive are favoured by the experimenter, even if the favoured conditions are not explicitly revealed to the participants), we believe, however, that the subjective preference data is in some ways more valuable than quantitative data.

The creative people (artists, modelers, animators, designers) who use 3D graphics applications want interfaces that "feel right", and don't necessarily place much importance on speed advantages. If speed is everything, then one could argue that command line interfaces which experts can often operate much faster than GUIs would still dominate the industry. Clearly, GUIs predominate for reasons other than speed efficiency. As discussed in the introduction, there is a large perceptual component to many 3D graphics tasks and frequent epistemic actions are required to gain a good perceptual understanding of the scene. We believe that this translates into the user getting a better or faster understanding or evaluation of the results of their pragmatic actions. Non-dominant hand operation of camera controls, in addition to speed advantages in some tasks, reduces the cost of epistemic actions and provides the user with a greater sense of engagement with the 3D scene – a step in making 3D graphics interfaces "feel right".

ACKNOWLEDGMENTS

We thank George Fitzmaurice, Bill Buxton, and Russell Owen for valuable discussions and assistance during the course of this work. We also thank all the volunteers who participated in our experiments. The support of Alias|wavefront is gratefully acknowledged.

REFERENCES

[1] Balakrishnan, R., Baudel, T., Kurtenbach, G., & Fitzmaurice, G. (1997). The Rockin'Mouse: Integral 3D manipulation on a plane. *Proceedings of the CHI'97 Conference,* 311-318, ACM

[2] Bier, E. A., Stone, M. C., Pier, K., Buxton, W., & DeRose, T.D. (1993). Toolglass and magic lenses: The see-through interface. *Proceedings of the ACM Siggraph Conference,* 73-80, ACM.

[3] Braunstein, M.L. (1976). *Depth perception through motion.* Academic Press.

[4] Buxton, W., & Myers, B. A. (1986). A study in two-handed input. *Proceedings of the CHI'86 Conference,* 321-326, ACM.

[5] Cutler, L.D., Frohlich, B., & Hanrahan, P. (1997). Two-handed direct manipulation on the responsive workbench. *Proceedings of 1997 Symposium on Interactive 3D Graphics,* 107-114, ACM.

[6] Cutting, J.E. (1986). *Perception with an eye for motion.* MIT Press.

[7] Guiard, Y. (1987). Asymmetric division of labour in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behaviour, 19,* 486-517.

[8] Haber, R.N., & Hershenson, M. (1973). *The psychology of visual perception.* Holt, Rinehart, and Winston.

[9] Hinckley, K., Pasuch, R., Goble, J.C., & Kassell, N.F. (1994). Passive real-world interface props for neurosurgical visualization. *Proceedings of the CHI'94 Conference,* 452-458, ACM.

[10] Hinckley, K., Pausch, R., Proffitt, D., Patten, J., & Kassell, N. (1997). Cooperative bimanual action. *Proceedings of the CHI'97 Conference,* 27-34, ACM.

[11] Hinckley, K., Pausch, R., & Proffitt, D. (1997). Attention and visual feedback: The bimanual frame of reference. *Proceedings of the 1997 Symposium on Interactive 3D Graphics,* 121-126, ACM.

[12] Johnsgard, T. (1994). Fitts' Law with a virtual reality glove and a mouse: Effects of gain. *Proceedings of Graphics Interface 1994,* 8-15, Canadian Information Processing Society.

[13] Kabbash, P., Buxton, W., & Sellen, A. (1994). Two-handed input in a compound task. *Proceedings of the CHI'94 Conference,* 417-423, ACM.

[14] Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18(4),* 513-549.

[15] Kurtenbach, G., Fitzmaurice, G., Baudel, T., & Buxton, W. (1997). The design of a GUI paradigm based on tablets, two-hands, and transparency. *Proceedings of the CHI'97,* 35-42, ACM.

[16] Leganchuk, A., Zhai, S., & Buxton, B. (in press). Manual and cognitive benefits of two-handed input: An experimental study. To appear in *ACM Transactions on Computer Human Interaction*, ACM.

[17] Multigen Inc., *SmartScene*. http://www.multigen.com/

[18] Poulton, E.C. (1989). *Bias in quantifying judgements.* Lawrence Erlbaum Associates.

[19] Sachs, E., Roberts, A., & Stoops, D. (1991). 3-draw: A tool for designing 3D shapes. *IEEE Computer Graphics and Applications, 11(6),* 18-26.

[20] Ware, C., & Lowther, K. (1997). Selection using a one-eyed cursor in a fish tank VR environment. *ACM Transactions on Computer Human Interaction,* 309-322, ACM.

[21] Wickens, C.D., Todd, S., & Seidler, K. (1989). Three dimensional displays: Perception, implementation, and applications. *CSERIAC Rep: CSERIAC-SOAR-89-001.* Wright-Patterson Air Force Base, Ohio.

[22] Wickens, C.D. (1992). *Engineering psychology and human performance.* Harper Collins.

[23] Zeleznik, R. C., Forsberg, A. S., & Strauss, P. S. (1997). Two pointer input for 3D interaction. *Proceedings of the 1997 Symposium on Interactive 3D Graphics,* 115-120, ACM.