Rotating Virtual Objects with Real Handles

COLIN WARE and JEFF ROSE
University of New Brunswick

Faculty of Computer Science,
University of New Brunswick,
P.O Box 4400, Fredericton, NB.
CANADA E3B 5A3
Phone: (506) 458-7283
cware@unb.ca

Colin Ware is a Professor in the Computer Science at the University of New Hampshire, but this work was carried out when he was at the University of New Brunswick. He holds a PhD in Experimental Psychology from the University of Toronto and an MMath in Computer Science from the University of Waterloo. His research interests include 3D interaction techniques and the effective visualization of information. He is particularly interested in developing new techniques based on theories of perception.

Jeff Rose currently works with SkyRidge Systems Inc. This research was carried out while was a co-op research Assistant in the Faculty of Computer Science at the University of New Brunswick. His professional area of interest is systems analysis and design

Abstract

Times for virtual object rotations reported in the literature are of the order of ten seconds or more and this is far longer than it takes to manually orient a “real” object, such as a cup. This is a report of a series of experiments designed to investigate the reasons for this difference and to help design interfaces for object manipulation. The results suggest that two major factors are important. Having the hand physically in the same location as the virtual object being manipulated is one. The other is based on whether the object is being rotated to a new, randomly determined orientation, or is always rotated to the same position. Making the object held in the hand have the same physical shape as the object being visually manipulated was not found to be a significant factor. The results are discussed in the context of interactive virtual environments.

Categories and Subject Descriptors: H.1.2 [**Models and Principles**]: User/Machine Systems – human factors; I.3.6 [**Computer Graphics**]: Methodology and Techniques – interaction techniques.

General Terms: Human Factors, Experimentation.

Additional Key Words: Input devices, Virtual reality, Two handed input, Direct manipulation, 3D object manipulation, 3D rotation.

1. INTRODUCTION

Many studies have established that it typically takes about a second to move a cursor to an object on a computer screen using a device such as a computer mouse [Card et al., 1983; MacKenzie, 1992]. This is despite the fact that the device held in the hand is nothing like the cursor on the screen, and perhaps more importantly, the hand is usually held to the side of the computer and not on the screen of the computer where the cursor appears. Ware and Balakrishnan [1994] found that 3D object positioning could be accomplished in about two or three seconds so long as there was little lag between the object held in the hand and the 3D cursor feedback. In this case the hand holding the input device was also held by the user's side and not in the same space with the 3D cursor. Graham and MacKenzie [1996] suggest this displacement may be a factor in increasing selection times, but the evidence shows that it is not a large effect. Thus, we can conclude that translating virtual objects is essentially a solved human factors problem. Given a suitable device, the task can be done rapidly and efficiently.

But what about rotating virtual objects? Picking up a small object, such as a cup, and reorienting it in some way should convince the reader that this task can also be performed in a second or two in the real world. But the studies that have been done of virtual object rotations have reported much longer times. Ware [1990] found that 3D placement (both position and orientation) took as much as 55 seconds when subjects were instructed to place an object "as accurately as possible". In a condition where speed was emphasized orientation matches were reduced to 13.4 seconds on average.

Zhai et al. [1996] studied an orienting task with two devices, a glove and a finger ball. In early trials they measured times of 18 and 14 seconds respectively but these improved with practice to 11 and 5 seconds by the end of 5 blocks of trials.

Hinckley et al. [1997] also examined the effect of input device shape on object rotation. Two input devices were used; one was a two-inch diameter ball while the other was the basic unadorned sensor sold by Polhemus. They reported rotation times of 20.7 seconds for females, 14.9 seconds for males with the ball device, and 21.5 seconds for females, 15.9 seconds for males with the tracker device. These times are all extraordinarily long for a task that seems simple. However, under at least some circumstances, object rotations can be fast. Wang et al.[1998] measured the time taken to align a small wooden cube with a computer graphics generated cube presented in the same location. This was done using a half-silvered mirror so that the user could see the computer graphics generated imagery of the virtual cube (reflected from the mirror) superimposed on the subject's hand holding an actual wooden cube. The task was to rotate the real cube to match the position and orientation of the virtual cube. Wang et al. [1998] measured rotation and translation times in the range of half a second to one full second, depending on the distance traveled. Moreover, the time taken for the rotations was actually less than the time taken for translation. One difference between Wang et al's study and previous ones was that the rotations required were relatively small, 22.5 and 45 degrees. Nevertheless, this result shows that fast rotations can be achieved under the right circumstances.

Although one must be cautious about cross experiment comparisons there would appear to be a large discrepancy between studies that show real object rotation to be fast and virtual object rotation to be slow. This discrepancy is not as large for translations. The following is a list of factors that may cause the discrepancy.

1) Haptic and visual object shape mismatch. Typically in VR environments the user holds some manipulator object in their hand and this becomes attached or "bound" to some virtual object in the 3D environment. There is usually no physical shape correspondence between the object that is grasped by the hand and the computer graphics object that is manipulated. A major exception to this was a study by Hinckley et al. [1994] in which they used a doll's head, held in the hand, to control a virtual head representing an MRI scan of a human head. In this case the object in the hand bore a close resemblance in shape to the object being manipulated and they reported that this provided a natural, easy to use interface, although they did not provide performance figures.

2) Lateral displacement. With real object manipulation, there is perfect correspondence between the object haptically manipulated and the object visually perceived. In many computer systems the hand is displaced from the virtual object being manipulated. Although, as discussed above, displacing the hand does not appear to have a large effect on object translations, it may be significant when rotations are performed.

3) Degree of practice. Using a computer mouse is a heavily overlearned skill for many people. In contrast, people rotating objects in virtual environments are often performing the task for the first time. The power law of practice holds that performance improves as a logarithmic function of the number of times a task is performed. However, Card et al. [1983] found cursor positioning using a mouse could be done in about 2.5 seconds on the very first block of 20 trials. This is still much faster than the object rotation times that have been reported. For example, Zhai et al. [1996] reported mean times of approximately 15 to 18 seconds on the first block of trials depending on the input device used, and mean times of approximately 7 to 10 seconds after 5 blocks of 5 trials.

4) Lag. There is no lag in the visual feedback from real object manipulations, but typically between 50 and 200 milliseconds lag exists in virtual object manipulation systems. This has been shown to be an important factor in the rate at which 2D and 3D translations can be carried out [Mackenzie and Ware, 1993; Ware and Balakrishnan, 1994].

5) One handed versus two handed input. A number of studies have suggested that two handed input can be more natural in object orientation. A key theoretical insight is that of Guiard [1987] who proposed differences in the roles of the left and right hands. The left hand generally holds the workpiece while the right hand manipulates tools within the frame of reference provided by the left hand.

6) Virtual image quality. The real world contains spatial cues not found in VR. Sometimes wire frame objects are manipulated (for example, Zhai et al. [1996]). In most cases cast shadows are not computed despite the fact that they are important depth cues [Wanger et al., 1992]. Stereoscopic viewing has been found to be advantageous in 3D object positioning tasks [Ware, 1990; Wanger et al., 1992].

7) Task degrees of freedom. One way of analyzing interaction is though the number of degrees of freedom that constrain the task. Many of the object positioning tasks that have been studied have involved only object translation in the plane, which involves only two degrees of freedom. Orienting and positioning an object in space involve six degrees of freedom, three for translation and three for rotation. Experimental results have shown that rotations and translations can be carried out in parallel [Ware, 1990; Wang et al., 1998] and under some circumstances increasing the number of degrees of freedom may not incur any performance penalty. Nevertheless the task degrees of freedom must always be borne in mind as a factor in task performance.

8) Ratcheting. A large rotation cannot always be accomplished in a simple rotation of the hand due to joint constraints. Some interfaces allow the subject to make ratcheting movements, using a button as a clutch and incrementally rotating the virtual object to its target destination [Ware, 1990]. Zhai et al. [1996] suggested that a small ball input device may be optimal since it can be rolled between the fingers eliminating the need for ratcheting.

9) Axis of rotation. Parsons [1995] showed that subjects are unable to judge the proper angle required to rotate an object to another arbitrary orientation. They can only make this judgement when the rotation is about the axis of symmetry of the object. Zhai and Milgram [1998] showed that virtual object rotations were faster about axes of symmetry compared to arbitrary axes.

10) Miscellaneous task requirements. There are many task demands that may influence positioning time. For example whether movement is repetitive or whether a different movement is required on each trial is likely to be a major factor, as is the precision of the movement that is required. Some experiments have required precise tolerances on the orientation whereas others have only asked for "quick and dirty" positioning. Zhai et al. [1996] required subjects to match a tetrahedron object so that all four corners matched within a specified tolerance. Another factor is that having a movement from a random start position to a fixed target position is faster than having a random new location on each trial given a fixed start position over a series of trials (Stelmach, 1978).

In this paper we present four experiments designed to investigate the factors that contribute to the differences between virtual and real object rotations. The experimental task in all cases was to make one 3D object parallel to another 3D object by rotating from one random orientation to another. The first experiment is designed to establish baseline times for the task by having subjects make a wooden "handle" object held in the right hand parallel to an identical object held in the left hand. The second experiment investigates the effects of the displacement of the hand from the virtual object being manipulated. The third experiment investigates three additional factors: a comparison between real and virtual visual feedback, a comparison between two handed and one handed manipulation, and a comparison between repeatedly moving something to the same target orientation (from a random start orientation) and moving it to a new random orientation. The ultimate goal of these experiments is to help design better object positioning interfaces and the relevance to this is discussed in the conclusion as are more theoretical issues. In all cases rotation was about arbitrary axes. The fourth experiment investigates the importance of the shape of the object held in the subject's hand.

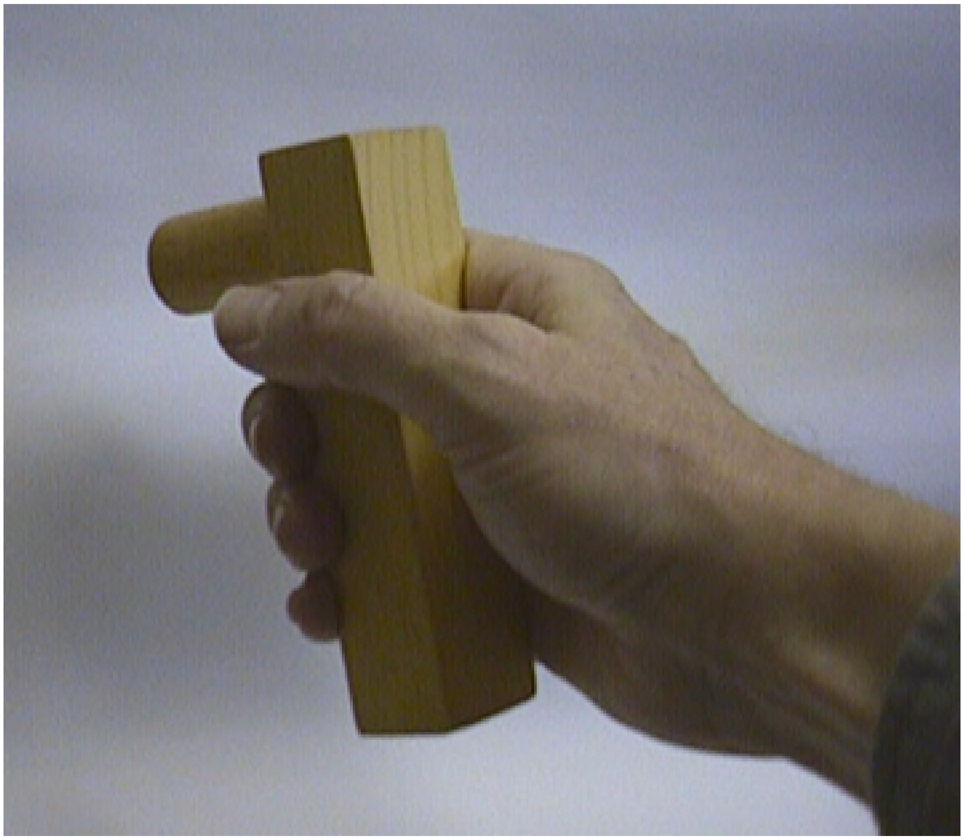

Figure 1. One of the handles used in the studies. Computer graphics images matched in size, shape and texture.

2. EXPERIMENT 1: REAL OBJECT ROTATION

The task chosen for the series of experiments reported here was that of making one 3D object have the same orientation as another. To this end, a 3D object was designed so that its orientation would be easy to perceive unambiguously, by its shape alone. By this criterion, a cube or tetrahedron object would be a bad choice because they can be rotated to create something visually identical. The objects manipulated were as shown in Figure 1. Each object had a box shaped part with a cylinder protruding from one end as shown. For the remainder of this paper these objects are referred to as “handles”.

The first experiment had a primary goal of establishing baseline performance in the task of bringing two real objects to a matching orientation with normal visual feedback. A secondary goal was to evaluate accuracy and speed of manually rotating an object with only haptic, and not visual feedback. It is known that visual feedback increases the accuracy of translational movement [Prablanc et al, 1979]; we were interested in the extent to which it affects accuracy and speed in rotation movements. Our hypothesis was that having visual feedback would reduce rotation time and increase accuracy.

Condition (i) Visual feedback. Subjects were required to hold a handle in the left hand, pick up a second handle in the right hand and match the orientations. Subjects could look at their hands and the wooden handles while performing the task. The start orientations were randomly determined.
Condition (ii) Blindfolded. Same as first condition only wearing a blindfold.

2.1 Method for Experiment 1

Apparatus

The physical objects that were manipulated were made of a light colored varnished wood. The box shaped part of the object was 2.8 x 3.4 x 13.2 centimeters. Protruding from one end was a 2.5 centimeter diameter dowel with a length of 4.7 centimeters. The purpose of the dowel was to give the object a completely unambiguous orientation. Each wooden handle had a Polhemus Fastrak tracking device embedded in its center. We estimated that under the conditions of the experiment we were able to measure absolute positions for the sensors to within 5 millimeters and differences between the orientation of two sensors within 1.0 degree. The foot switch used to terminate trials was constructed by modifying a standard computer mouse.

Procedure

Throughout the course of the experiment subjects held one of the handle objects in the left hand. There were 16 different starting orientations for the right hand object generated as shown in Figure 2. At the start of each trial the experimenter placed the object on a specially made block in one of the 16 orientations. The block was 10 centimeters square in plan view and 5.5 centimeters high. It was placed on the monitor table to the right and in front of the monitor. Once the object was positioned the subject loosely grasped it with his or her right hand without changing its position. The subject was instructed to manually find the position of the dowel so that they would know the orientation of the object at the start of each trial. On a tone signal the subject fully grasped the right hand object and rotated it to make it parallel with the left hand object. The grip they adopted in taking hold of the object was a kind of pistol grip as shown in Figure 1 with the rectangular portion representing the handle and the cylindrical portion representing the barrel. As part of the operation, some small rotation of the left hand object typically occurred but most of the rotation was done with the right hand. Once the subject was satisfied that the objects were parallel, he or she pressed the foot switch to terminate the trial. Subjects were instructed to perform the operation as fast and accurately as possible.

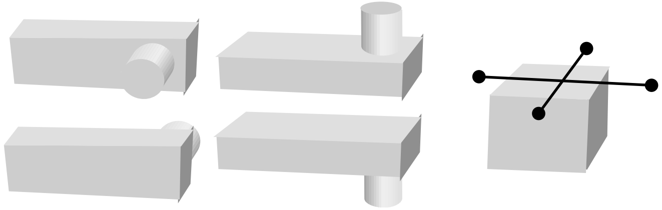

Figure 2. Four of the 16 starting orientations for the handle. The other orientations where generated by symmetry about the two axes indicated on the right. At the start of each trial the handle was placed on the block in one of the 16 orientations .

There were five practice trials before the experiment started and two practice trials at the start of each trial block. There were 16 experimental trials in each trial block, consisting of the 16 start orientations given in a random order. Four blocks of trials were conducted for each subject, either in the order Visual, Blind, Blind, Visual, or the order Blind, Visual, Visual, Blind.

The 13 subjects were either undergraduate or graduate students paid for participation. All were self proclaimed right handers and able to see stereoscopic depth in Julesz stereograms.

2.2 Results of Experiment 1

Because the data were highly skewed all analysis was carried out on log transformed raw data. The means that are reported are based on the antilog of averages obtained from the transformed data.

The time to perform the task in the visual feedback condition was 1.64 seconds on average. This data establishes a baseline for performing a rotation task under “real world” conditions. The time taken to perform the task in the blindfolded condition was about 17% longer, 1.92 seconds on average. An analysis of variance showed this difference to be significant ($p < 0.01$).

Quaternions were used to measure angular changes. A quaternion can be used to define the rotation from an arbitrary orientation of a 3D object into any other arbitrary orientation [Pletinckx, 1989]. A quaternion defining such a rotation can be decomposed into an axis of rotation and an angle of rotation about that axis. This angle component is always the smallest angle that is required to rotate the object to the other orientation. The orientation error was calculated using the normalized angle component of the quaternion defining the difference in orientation between the two handles.

For both conditions the range of initial rotation mismatches between the two handles measured during the course of the experiment was from 75 degrees to 180 degrees. The average angle of rotation from the start position to the end position was 128 degrees. Subjects were almost twice as accurate with visual feedback as without visual feedback; the mean angular error was 4.64 degrees in the visual feedback condition and 9.21 degrees without visual feedback. This difference was also highly significant ($p < 0.01$). These results confirm the hypothesis that both time and accuracy are improved with visual feedback.

There were significant individual differences in both time and accuracy data. For example, the fastest subject took 0.93 seconds on average in the visual condition while the slowest subject took 3.3 seconds in the visual condition. In terms of accuracy the most accurate subject achieved a mean error of 2.4 degrees while the least accurate averaged over 10 degrees. When mean time is plotted against error, a clear speed-accuracy tradeoff is revealed as shown in Figure 3. The fastest subjects were the least accurate and vice versa.

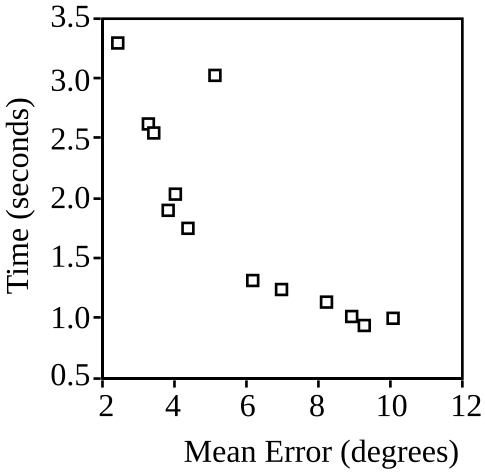

Figure 3. Each point represents average data from a single subject in the visual feedback condition of Experiment 1. There is a clear speed accuracy tradeoff. Subjects who performed the task faster also did it less accurately.

3. EXPERIMENT 2: HAND DISPLACEMENT

The second experiment is a comparison between rotations performed with the hand in the same space as the virtual object being manipulated versus the hand held at some distance to the object being oriented. In many systems, subjects typically hold a controller by the side of the body while positioning an object on screen. As a result there is a translational mismatch between the object that is manipulated and the object that is held. In this experiment we investigated the effect of eliminating the mismatch. To simplify the task all interaction was one handed, not two handed as in Experiment 1.

The mirror setup illustrated in Figure 4 was used to put an image of a virtual handle in a place accessible by the subject’s right hand. The subject saw two identical virtual handles of the same size and shape as the wooden handles they held. The virtual handles were texture mapped with a similar wood grain as illustrated in Figure 5. The left of these was set to a randomly determined target orientation, the right was controlled by the subject through one of the handles containing a Polhemus sensor. The subjects’ task was to use the wooden handle to make one virtual handle match the other target handle in orientation. Our hypothesis was that subjects would perform both faster and more accurately when the visual and haptic objects were spatially superimposed. There were two conditions.

Condition (i) Position match. The system was calibrated so that the virtual image of one of the handles closely matched the actual position of the wooden controller handle. The subject reached around behind the mirror with the handle in his or her right hand to perform the orientation matching task.

Condition (ii) Position mismatch. In this case the subject held the input handle in a comfortable position by the right side of his of her body in order to place the virtual handle in front of them. There was approximately a 60 centimeter displacement in the calibration relating the controller handle position to the virtual handle position.

3.1 Method for Experiment 2

Apparatus

The wooden handles were the same as for Experiment 1. Computer graphics were done using a Silicon Graphics Indigo2 Maximum Impact. All viewing was in frame sequential stereo at a 120 Herz update rate (60 Herz to each eye). CrystalEyes liquid crystal shutter glasses were used to separate the images to the two eyes. We estimated the mean system lag (between the time a movement of the Polhemus tracker was reflected in the computer graphics imagery) to be approximately 75 milliseconds. See Ware and Balakrishnan [1994] for calibration.

Procedure

In most respects the procedure was the same as that for Experiment 1 except that the number of trials in a trial block was increased to 24. There were 24 target orientations like those illustrated in Figure 2 but with the addition of 8 orientations generated using the third axis symmetries. The virtual objects were rotated so that all target orientations were at 45 degrees to the x,y and z axes defined in screen coordinate space. The 24 target orientations thus generated were given in a different random sequence for each subject. The 13 subjects were the same as those used in Experiment 1. Each was paid to participate.

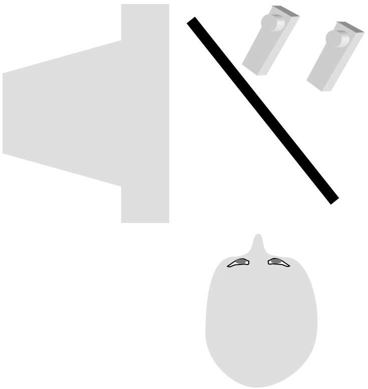

Figure 4. View from above of setup used in Experiment 2. A mirror 45 degrees to the screen was used to place a stereoscopic virtual handle in the space as shown. In one condition a virtual handle was calibrated to appear in the same position and orientation as a real wooden handle.

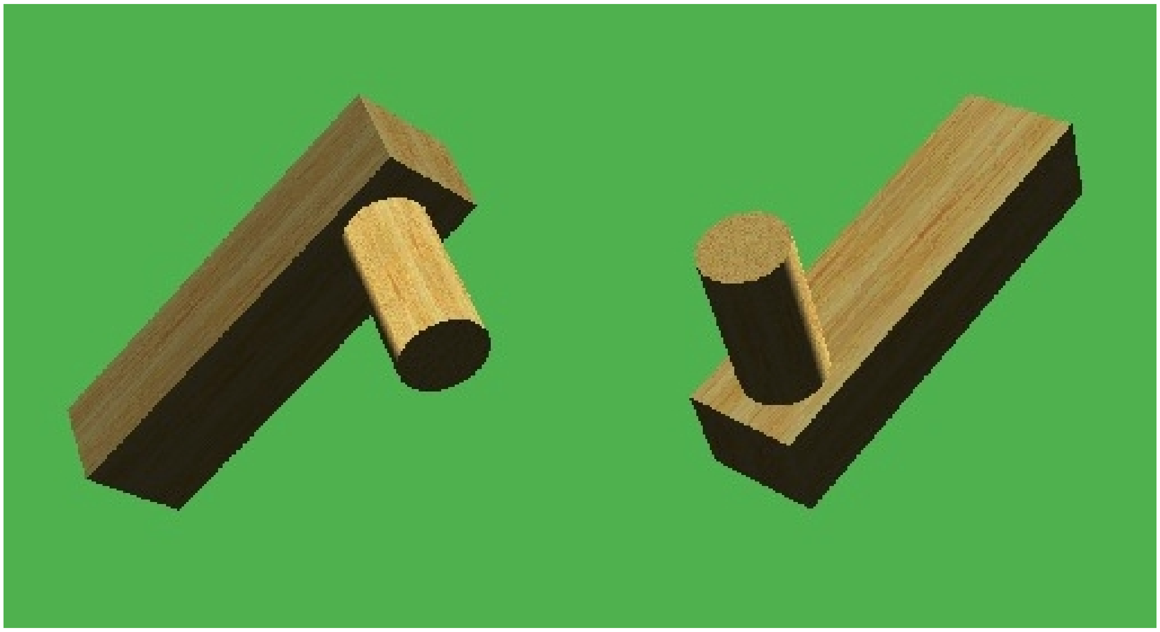

Figure 5. The virtual handles had the same size, shape and texture as the wooden handles held by the subjects.

3.2 Results of Experiment 2

On average subjects took significantly less time to orient the object with the controller object in the same place as the virtual object (3.70 seconds compared to 4.96 seconds, $p < 0.01$). This 35% difference confirms the hypothesis that it is easier to rotate an object if the hand is in the same spatial location as the object being rotated. There was no significant difference in mean accuracy (5.8 degrees versus 6.0 degrees).

Overall the times were longer and the errors larger than with Experiment 1. There was no evidence for a speed-accuracy tradeoff like that found in Experiment 1. The average rotation angle for the task was 125 degrees, slightly smaller than Experiment 1 (128 degrees). The fact that performance was considerably slower than that measured in Experiment 1 might be due to any of the following differences between Experiments 1 and 2. Input was one handed versus two handed and virtual instead of real objects were used. In addition, the task for this experiment involved the subject rotating an object to a new, randomly different orientation on each trial. This was different from Experiment 1 where the subjects started each trial with a random orientation on each trial but always ended the trial with the object oriented the same way. These differences are each addressed in Experiment 3.

5. EXPERIMENT 3: REAL VERSUS VIRTUAL, TWO HANDED VERSUS ONE HANDED

The third experiment in our series had three objectives. The first was to make a direct comparison between real and virtual object manipulation. Graham and MacKenzie [1996] compared moving a virtual pointer with pointing directly with the finger. The conditions differed mainly in the kind of visual feedback that was provided. In one case feedback was from the direct observation of the subject's own hand. In the other case, feedback was given using computer graphics in the same spatial location as the subject's hand. The superimposition was done using a mirror. They found a small time difference between the two conditions that increased as the task difficulty increased. We believe that this difference can be explained by the lag that is inevitable from providing computer graphics feedback [MacKenzie and Ware, 1993]. The purpose of the present experiment was to similarly compare performance when subjects were allowed feedback from viewing their own hand compared to feedback provided using computer graphics imagery for feedback. In a sense this is a kind a control condition; if there are adequate visual spatial cues then there is no obvious reason why the computer graphics feedback should not be as effective as direct visual feedback, except for the small amount of lag that is introduced by the use of computer graphics and the hand tracking apparatus. The first of the two conditions (i) was identical to the first condition of Experiment 1. In this condition, the subject brought two wooden handles into an identical orientation, under natural visual feedback. In the second condition (ii) two wooden handles were held in the same way but visual feedback was provided by two virtual handles in the same position and orientation. In other words, computer graphics imagery replaced the real-world imagery.

The second objective was to compare two handed and one handed input. To accomplish this, a third condition was added, similar to condition (ii). The third condition was the same as the second condition except that the left virtual handle was maintained in a fixed position instead of being controlled by the left hand. Since Experiment 1 showed that most of the rotation was done with the right hand, in reality this was a relatively small change. Thus condition (iii) was like condition (ii) except that only one of the two handles was held. Our hypothesis was that two handed input would be faster (condition ii would be faster than condition iii).

The third objective was to examine the effect of rotating to a particular fixed target orientation, versus rotating to a random target orientation. In Experiment 1, in bringing the two handles into alignment, the subjects always ended up with the handles in the same orientation. However, in Experiments 2 and 3 the subject always rotated one of the objects to a new and different target orientation. It seemed likely that rotating something many times to the same orientation from different starting orientations (as in Experiment 1) would be easier than rotating an object to a randomly different orientation on each trial. Studies of linear and lever positioning tasks have shown that a random start is generally faster than a random end condition (Stelmach, 1978). Conditions (i), (ii), and (iii) all involved rotating to a fixed orientation (like Experiment 1). To investigate the effects of random start versus random end condition (iv) was added. This condition was identical to condition (iii) except that it involved matching random target orientations generated as described in Experiment 2. Our hypothesis was that the condition (iv) task would take longer than the condition (iii) task.

Condition (i) Two handed real. Subjects were required to hold a handle in the left hand, pick up a second handle in the right hand and match the orientations. Subjects could look at their hands and the wooden handles while performing the task. This was the same as condition (i) in Experiment 1. The start orientations were randomly determined.

Condition (ii) Two handed virtual. Subjects looked at two virtual handles calibrated to be in the same place as the real handles that they held in each hand. The task was the same as for condition (i). Visual feedback was provided by virtual images of the two handles. The start orientations were randomly determined.

Condition (iii) One handed virtual. The task was the same as for condition (ii) except that the matching was one handed. The left hand virtual handle was in a fixed position throughout. The subject rotated the right hand handle from the start position to match the fixed handle in orientation. The start orientations were randomly determined.

Condition (iv) One Handed virtual, random end. The task was like condition (iii) except that the target virtual object was given a orientation on each trial. The starting hand positions were fixed. The end orientations were randomly determined.

5.1 Method for Experiment 3

Apparatus

The computer graphics was generated in the same way as described for Experiment 2. A mirror arrangement was constructed to place the wooden handles in the virtual workspace. This was a variation on that created for Experiment 2 (which only allowed one hand to be placed in the virtual workspace). In the new arrangement the subject looked down as illustrated in Figure 6. This allowed the user to place both hands (holding wooden handles) in the virtual workspace.

Procedure

For conditions (i), (ii) and (iii) the starting positions for the handle held in the subjects' right hand were the same as those used for Experiment 1. Subjects started each trial with a hand placed on the handle in the same ways as for Experiment 2. For condition (iv) the end positions were the same as those used for Experiment 2. Each condition was presented in a single trial block, and trial blocks were given in a different random order for each subject. Only 12 of the original 13 subjects participated. Each was paid to participate.

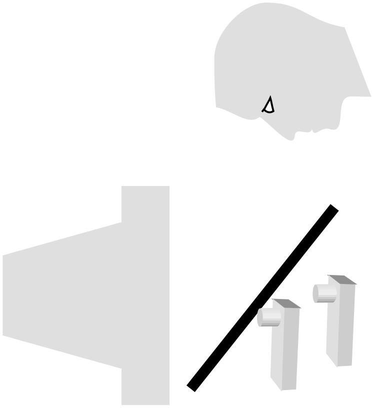

Figure 6. Apparatus setup for experiment 3. A mirror rotated 45 degrees from the vertical allowed subjects to look down on a virtual workspace and place their hands in the same place as the computer graphics imagery.

5.2 Results of Experiment 3

The two handed real object rotations with visual feedback took 1.80 seconds on average. This time is similar to the 1.63 seconds obtained in Experiment 1. The average angle of rotation on each trial was 125 degrees. The two handed rotations with virtual visual imagery was done in 2.25 seconds, about 450 milliseconds longer. This difference was significant ($p < 0.01$). Ware and Balakrishnan [1994] developed a model for the effects of lag in a 3D placement task.

$\text{MeanTime} = 0.739 + 1.59(0.266+\text{lag})\text{ID}$

Where $\text{ID} = (\log_2(\text{distance}/\text{width} + 1.0)$

Despite the fact that this is a rotation task and not a translation task it is interesting to apply this analysis using the mean rotation of 125 degrees divided by the mean error (approximately 4 degrees) to give an index of difficulty (approximately 5 bits). According to this rough analysis a 75 millisecond lag should add about 600 millisecond to this task. This is actually more than the 450 milliseconds difference that we found. Although we acknowledge that applying Fitts' law to a rotation task in this way is problematic, this analysis does add plausibility to the idea that the main factor in the difference between real object visual feedback and virtual object visual feedback was lag.

Although two-handed input was slightly faster than one handed input (2.25 versus 2.43 seconds) this 8% difference between conditions (ii) and (iii) failed to reach significance. It cannot be concluded from this that there are no benefits from two handed inputs, since in this task the left hand actually did very little work. However, the results do show that at least for some tasks there is no compelling advantage to using two hands to position objects.

The hypothesis that having a different random target orientation on each trial would result in slower performance was confirmed by the difference between conditions (iii) and (iv). This 37% highly significant ($p < 0.01$) difference was the largest obtained in this experiment. This difference surprised us since we had added this condition more in the interests of completeness than in the expectation that this would be an important factor. This result shows that task demands are higher when moving to a new orientation on each trial. Since this is probably a more realistic model than moving repeatedly to the same orientation (or position) it may be more useful for obtaining estimates of task difficulty.

4. EXPERIMENT 4: OBJECT HELD UNLIKE OBJECT SEEN

The idea of “props” in graphical user interfaces [Hinckley et al. 1994] is to create a hybrid environment in which real three dimensional objects manipulated by the user are used to control virtual objects in a 3D space created using computer graphics. Props are usually similar in shape to the virtual object being manipulated. Hinckley et al. used a doll’s head to manipulate MRI scan images of the human head. Most computer interfaces to CAD packages or drawing systems are not like this, the object held in the user’s hand is different from the object manipulated on the screen. The object held in the hand is usually a mouse, or the stylus of a digitizing tablet, while the virtual object manipulated on the screen is sometimes a large thing like an entire window that is being moved, but most often is a small cursor. This is unlike the situation in the real world where the object that is viewed has the same dimensions, shape and position as the object that is manipulated with the hand. In this experiment we compared performance on the task of matching orientations when the haptic shape did match the visual shape with performance when the haptic shape did not match the visual shape.

In the shape mismatch condition we chose to use a small ball as the control device, mostly because Zhai et al. [1996] had showed this to be an effective input device for a rotation task. Our hypothesis was that faster performance would result when the haptic and visual objects matched compared to when they did not match.

There were two experimental conditions.

Condition (i) Shape match. Subjects used the wooden handle to manipulate a virtual handle shaped object.
Condition (ii) Shape mismatch. Subjects used a rubber ball to manipulate a virtual handle shaped object.

4.1 Method for Experiment 4

Apparatus

The computer graphics was generated in the same way as described for Experiment 2. In addition to the apparatus already described, the Polhemus sensor was embedded in a 5 centimeter diameter rubber ball for condition (ii).

Procedure

A subject used his or her right hand holding a ball or handle object in front of them so that they could see both the virtual object being manipulated and the object they held. However, no mirror was used and the physical and virtual objects were not spatially superimposed. The number of trials was reduced to 8 per block. The target orientations on each trial were randomly selected from the 24 described for Experiment 2. Only 12 of the original 13 subjects were used. Each was paid to participate.

4.2 Results of Experiment 4

Mean rotation times were 3.83 seconds in the shape match condition and 4.09 seconds in the shape mismatch condition. This 7% difference just failed to reach significance. There was no significant difference in accuracy.

Zhai et al. [1996] argued that a small ball is an effective input device because it can be rotated using fine finger movements. In the case of the wooden handle, it was also observed that the subjects started the rotation of the object using finger manipulation before adopting the grip illustrated in Figure 1. Our hypothesis was based on the supposition that a mismatch between shape of the object haptically manipulated and the one visually perceived would cause cognitive problems that would degrade performance. We found no evidence to support this.

6. DISCUSSION AND CONCLUSION

For ease of comparison, the results from all four experiments are summarized in Table 1 and in Figure 7.

	Conditions	Time (seconds)	Mean Error (degrees)
Experiment 1	i) Visual feedback RS	1.64	4.64
	ii) Blindfold RS	1.92	9.21
Experiment 2	i) Position match RE	3.70	5.77
	ii) Position mismatch RE	4.96	5.98
Experiment 3	iii) Real two handed RS	1.80	3.71
	iv) Virtual two handed RS	2.25	3.73
	iii) Virtual one handed RS	2.43	3.65
	iv) Virtual one handed RE	3.32	5.64
Experiment 4	i) Shape match RE	3.83	5.38
	ii) Shape mismatch RE	4.09	5.65

Table 1. Summary of results. RS denotes a random start orientation. RE denotes a random end orientation.

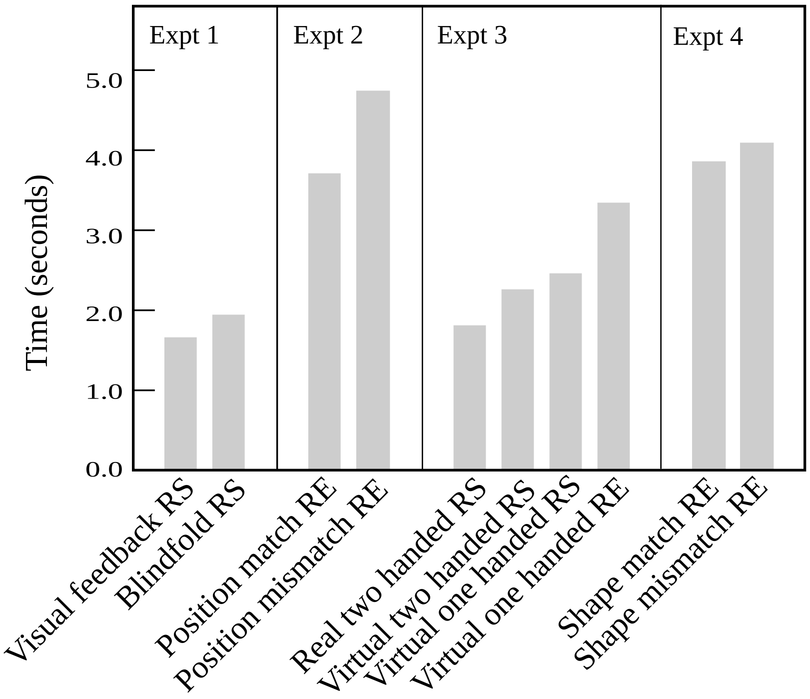

Figure 7. Summary of timing results for all four experiments.

The broad justification for virtual reality is that it should be easier to perform tasks if they can be carried out in the same way as in the “real” world. However, there are few experimental results showing the advantages of virtual reality systems, with a few exceptions such as the recent study of target searching by Pausch et al. [1997]. Experiment 2 in the present study provides additional evidence that, at least for the rotation task, creating conditions for natural eye-hand coordination can aid performance.

Overall, the results support the idea that placing the hand in the working space will improve interaction. In Experiment 2 subjects performed 35% faster despite the fact that the hand position was quite awkward in the position match condition. Schmandt [1983] was the first to investigate the technique as an aid to interacting with 3D virtual objects. He used a half-silvered mirror to place the user hand in the same virtual space as computer graphics imagery. However, because the mirror was only half-silvered the result was a blending of the two types of imagery without preservation of the all-important occlusion depth cue. Thus, more recent incarnations of this type of setup, for example Serra et al. [1997], used a fully silvered mirror so that only computer graphics imagery is seen. In general, the use of a stereoscopic display offers the capability of providing a relatively low cost virtual reality display with greater quality than helmet mounted displays [Deering, 1992; Ware et al., 1993]. The results of Experiment 2 suggests that interaction in such an environment, particularly object rotations, can be enhanced by placing the hands in the same physical location as the virtual objects being manipulated.

We have no adequate theoretical explanation for why rotations are affected by a translational displacement of the hand, whereas translations, apparently, are not. A possible explanation might be constructed based on the frame of reference of the hand. Since the resting position of the hand along with the forearm is rotated when the hand is moved from a position in front to the body to a position by side of the body and this may affect the frame of reference for rotations more than translations. We are grateful to an anonymous reviewer for this speculation.

The results failed to support the idea that it is important for the shape of an object held in the hand to be the same as the shape of the object seen. However, we cannot conclude from this result that any arbitrarily shaped object is a good manipulator. A sphere is a rather special shape and thus it would be unwise to generalize. Indeed, one interpretation of our results is that a sphere provides a very good input device, almost as good as holding the actual object manipulated. Supporting this interpretation, Zhai et al. [1996] found a glove based device to be worse than a sphere for a rotation task. However, as they noted, the sphere also has a disadvantage in that it lacks an obvious position for button placement. Indeed the fact that one or more buttons are required for most manipulation interfaces make the sphere design less attractive than it would otherwise seem.

Our results from Experiment 3 also failed to show an advantage for two handed input over one handed input. Without exception our subjects carried out the majority of the rotation with the right hand, even though they could have chosen to do part of the rotation with the left hand. This is consistent with the theory of Guiard [1987] that the left hand provides an anchor and reference for right hand orienting and positioning movements.

One of the larger effects that we had not anticipated at the start of the series of experiments was the effect of random start versus random end positions. To place this in context, the original Fitts' Law positioning experiments [Fitts, 1954] involved highly repetitive tasks. Subjects were required to tap from one vertical bar to another many times in rapid succession. Thus what was being measured was not a single execution of a task, but a single part of a composite task. It is hardly surprising that task performance will be longer in experiments where there is a different new position and/or orientation on each trial. The subject does not only have to make the correct hand movement but has also to execute a movement that is different from the preceding one. In a similar way, always completing a movement with the hand in the same position should be easier than always completing a movement with the hand in a new position. This aspect of positioning performance has received little attention but our results show it to be an important factor.

The results from Experiment 3 show that virtual object orientation can, under the right circumstances, approach real object rotation in speed. The difference between the real and virtual object manipulation was about 0.45 seconds and we attribute this to the 75 milliseconds estimated system lag multiplied by a number of iterations through an eye-hand coordination feedback loop. As Ware and Balakrishnan [1994] showed, the effects of lag may be multiplied several times in a task that involves closed loop control with visual feedback.

However, we have not accounted for the fact that even our longest times were shorter than the best virtual object rotation times reported by others. The probable reason for this lies in the task demands. Most previous studies have required that specific accuracy targets be met. For example Zhai et al. [1996] required that all four corners of a tetrahedron object being manipulated fell within a specified range of a target tetrahedron, although it is not clear what angular constraints were imposed by this.

To restate our main finding, our results provide empirical support for at least a limited form of virtual reality interface. One of the essential affordances of virtual reality systems is to provide natural eye-hand coordination for object manipulation and our work shows that this does indeed improve performance when the task is object rotation.

ACKNOWLEDGEMENTS

The authors wish to thank Christine MacKenzie and two anonymous reviewers for their details and insightful comments. This work was funded through an NSERC Strategic Grant to the first author.

REFERENCES

CARD, S.K., MORAN, T.P., AND NEWELL, A. 1983. The Psychology of Human-Computer Interaction. Lawrence Erlbaum.

DEERING, M. 1992. High resolution virtual reality. Proceedings of ACM SIGGRAPH'92. Computer Graphics 26(2) 195-202.

FITTS, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology: Human Perception and Performance, 7, 1019-1030.

GRAHAM, E.D. AND MACKENZIE, C.L. (1996) Physical versus virtual pointing. Proceedings of ACM CHI96, 292-299.

GUIARD, Y. 1987. Asymmetric division of labor in human skilled bimanual action. Journal of Motor Behaviour, 19(4) 487-517.

HINCKLEY, K., PAUSCH, R., TULIO, J., PROFFITT, D., KASSELL, N. 1997. Usability analysis of 3D rotation techniques. Proceedings of ACM UIST' 97, 1-10.

HINCKLEY, K., PAUSCH, R., GOBLE, J., AND KASSELL, N. 1994. Passive Real-World Interface props for neurosurgical visualization. Proceedings of ACM CHI'94, 452-458.

MACKENZIE, I.S. 1992. Fitts' Law as a research and design tool in human-computer interaction. Human-Computer Interaction, 7, 91-139.

MACKENZIE, I.S. AND WARE, C. 1993. Lag as a Determinant of Human Performance in Interactive Systems, Proceedings of ACM INTERCHI'93 Technical Paper. 488-493.

PARSONS, L. 1995. Inability to reason about an object's orientation using an axis and angle of rotation. Journal of Experimental Psychology: Human Perception and Performance, 21(6), 1259-1277.

PAUSCH, R. PROFFITT, D. AND WILLIAMS, G. 1997. Quantifying immersion in virtual reality, Proceedings of ACM SIGGRAPH'97, 13-18.

PLETINCKX, D. (1989) Quaternion calculus as a basic tool in computer graphics. The Visual Computer. 5, 2-13.

PRABLANC, C. ECHALLIER, J.E. JEANNEROD, M AND KOMILIS, E. 1979 Optimal response of eye and hand motor systems in pointing at a visual target II: Static and dynamic visual cues in the control of hand movement. Biological Cybernetics, 35, 183-187.

SCHMANDT, C. 1983. Spatial input-display correspondence in a stereoscopic computer graphic workstation. Proceedings of ACM SIGGRAPH'83, Computer Graphics 17,3, 253-259.

SERRA, L., NG H., CHOON, C.B. AND POSTON, T. 1997. Interactive vessel tracing in volume data. Proceedings of ACM SIGGRAPH Symposium on Interactive 3D graphics. 131-137.

STELMACH, G.E. 1978. Information Processing in Motor Control and Learning. Acacemic Press: New York.

WANG, Y., MACKENZIE, C., SUMMERS, V.A., AND BOOTH, K.S. 1998. The structure of object Transportation and Orientation in Human-Computer Interaction. Proceedings of ACM CHI'98 , 312-319.

WANGER, L., FERWANDA, J., AND GREENBERG, D. 1992. Perceiving spatial relationships in computer generated images. IEEE Computer Graphics and Applications (May) 44-58.

WARE, C. ARTHUR, K. AND BOOTH, K.S. 1993. Fishtank Virtual Reality. Proceedings of ACM INTERCHI'93 Technical Paper. Proceedings 37-42.

WARE, C. 1990. Using hand position for virtual object placement. The Visual Computer, 6, 245-253

WARE, C. AND BALAKRISHNAN, R. 1994. Object acquisition in VR displays: Lag and frame rate. ACM Transactions on Computer Human Interaction. 1(4), 331-357.

ZHAI, S., AND MILGRAM, P. 1998, Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. Proceedings of ACM CHI'98 Proceedings, 320-327.

ZHAI, S., MILGRAM, P., AND BUXTON, W. 1996. The effects of using fine muscle groups in multiple degree-of-freedom input. Proceedings of ACM CHI'96 308-315.