# Frames of Reference in Virtual Object Rotation

Colin Ware and Roland Arsenault
Data Visualization Research Lab, Center for Coastal and Ocean Mapping
University of New Hampshire

## Abstract

It is difficult with most current computer interfaces to rotate a virtual object so that it matches the orientation of another virtual object. Times to perform this simple task can exceed 20 seconds whereas the same kind of rotation can be accomplished with real objects and with some VR interfaces in less than two seconds. In many advanced 3D user interfaces, the hand manipulating a virtual object is not in the same place as the object being manipulated. The available evidence suggests that this is not usually a significant problem for manipulations requiring *translations* of virtual objects, but it is when *rotations* are required. We hypothesize that the problems may be caused by frame of reference effects – mismatches between the visual frame of reference and the haptic frame of reference. Here we report two experiments designed to study interactions between visual and haptic reference frames space.

In our first study we investigated the effect of rotating the frame of the controller with respect to the frame of the object being rotated. We measured a broad U–shaped relationship. Subjects could tolerate quite large mismatches, but when the orientation mismatch approached 90 degrees performance deteriorated rapidly by up to a factor of 5. In our second experiment we manipulated both rotational and translational correspondence between visual and haptic frames of reference. We predicted that the haptic reference frame might rotate in egocentric coordinates when the input device was in a different location than the virtual object. The experimental results showed a change in the direction predicted; they are consistent with a rotation of the haptic frame of reference, although only by about half the magnitude predicted. Implications for the design of control devices are discussed.

**CR Categories:** H.5.2 [Information Interfaces and Presentation]: User Interfaces–Theory and Method; I.3.6 [Computer Graphics] Methods and Techniques–Interactive Techniques.

**Keywords:** 3D Interaction, Input devices, Direct manipulation, virtual reality.

## 1 Introduction

Object rotations are an important challenge for 3D user interface designers because, using most current desktop interfaces, rotating a virtual object can take between ten and thirty seconds [Chen et. al. 1988, Hinkley et al. 1997]. This is much slower than real object rotations that take between one and two seconds [Wang et al. 1998, Ware and Rose, 1999]. Thus, there is room for an order of magnitude improvement. It is also in marked contrast to the situation with object translations, where times of less than two seconds are the norm, even with simple mouse-based interfaces.

In previous work, we and others have found a number of factors to be important in determining the speed and accuracy of virtual object rotations, including the type of input device, the mapping of hand motion to object motion and system lag. One factor that has been implicated, but not thoroughly investigated, is the interplay between the various frames of reference that come into play in object manipulation. There are three that seem especially significant: the egocentric head/torso centered frame of reference that defines visual space; the haptic/hand arm centered frame of reference defined by the input device; and the object centered frame of reference of the object being manipulated. In this paper, we discuss frame of reference issues and we report experimental results showing that orientation mismatches between haptic and visual frames of reference can have large effects on object rotation times. But first, we review some of the prior work related to the problem of virtual object rotations.

### 1.1 Object Rotation Interfaces

Two of the best mouse-based interfaces for object rotations are the arcball [Shoemake, 1994] and the virtual trackball [Chen et al. 1988]. In both, the mouse movement is mapped to the rotation of a virtual sphere causing a coupled virtual object to be reoriented. However, rotation times are long. Chen measured times of 17.5 seconds to make complex rotations with the virtual trackball, while Hinckley et al. [1997] measured times of around 28 seconds for both the arcball and the virtual trackball. These results are puzzling because people can rotate real objects quite rapidly. Wang et al. [1998] measured times of less than a second for simple rotations of a real cube through about 45 degrees about a vertical axis. Ware and Rose [1999] measured rotation times of 1.64 seconds when subjects rotated real objects by up to 180 degrees.

An obvious problem with mouse interfaces is that a computer mouse does not measure its physical orientation. Some indirect mapping must be made between the mouse translation and object rotation. To improve the situation a number of researchers have worked with hand-held input devices that sense rotations as well as translations. These are usually based on Polhemus™ or Flock of Birds™ systems. However, even when the rotation of an object is mapped with a one-to-one correspondence to the rotation of a 6 degree-of-freedom device held in the user's hand, the time to complete an object rotation may be substantial. For example, Ware [1990] reported 13.5 seconds when speed was emphasized and 28 seconds when accuracy was emphasized in an object rotation task. Zhai and Milgram, [1998] measured 18 seconds with a glove device and 13 seconds with a small ball device, declining with practice to 11 and 5 seconds respectively. Hinckley et al. [1997] reported 10.7 and

14.9 seconds using a ball shaped device depending on whether the subjects were male or female or male. In all of these cases, the virtual object being manipulated rendered at a different location to the hand-held input device, but device rotations were mapped directly to object rotations. In most cases, the virtual object was seen on the monitor, in front of the user, whereas the input device was held somewhere in front and to the side of the monitor. Thus, there was a translational mismatch between the two frames of reference and the rotational frame of reference was simply translated to the new position. We will return to this point.

By using virtual reality techniques, it is possible to co-locate the object being manipulated with the user's hand, and fast rotations result. Ware and Rose [1999] carried out an experiment in which the subject manipulated a real object but received visual feedback from a virtual object that was co-located and had the same physical shape. Thus subjects saw a virtual object but felt a physical object having the same shape in the same place. The rotation times were only slightly slower than those obtained with real objects and the difference could be explained by system lag. In another experiment, the authors compared having the control device held at the user's side with having it co-located with the object being manipulated. They found that co-locating the input device with the virtual object significantly improved performance. This result was one of the main motivating factors in the present research. We wished to discover more about what performance problems might arise when various types of frame of reference mismatch occurred.

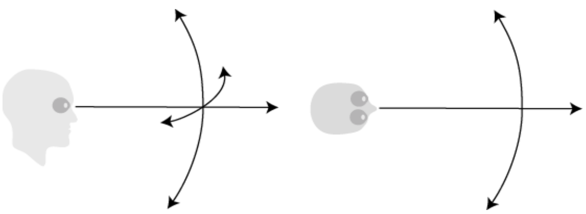

Figure 1. The egocentric perceptual reference frame provides a radial coordinate space centered at the user's head.

## 1.2 Frames of Reference

Computer interfaces have incorporated a large range of different mappings between hand motion and virtual object positioning. In this section, we introduce some of the important frames of reference.

*1) Egocentric reference frame.* This roughly corresponds to our conscious perception of the space around us; the location and orientation of objects are judged with respect to the observer's head. Because we scan the world by means of pan and tilt movements of the head, the egocentric reference frame can be regarded as providing a kind of radial coordinate space as Figure 1 illustrates. There is also evidence that we remember spatial layouts in an egocentric frame [Shelton and McNamara, 2001]. Generally, both psychophysical and neurophysiological studies suggest that it is the most important perceptual frame of reference, although many others play a role in cognition, such as eye-centered, body-centered, and arm-centered [Colby, 1998]. In the present work we are particularly interested in how visual egocentric space interacts with the haptic space of object manipulation.

*2) Room-centered reference frame.* Many modern environments have a strongly rectilinear layout, for example, the walls of a room [Pani and Dupree, 1994]. It is easier to remember the layout of spatial features if they are aligned with this frame. The gravitational upright is an especially important axis for imagining rotation in this reference frame.

*3) Object reference frame.* If an object has a clearly defined axis, we can more easily imagine rotations about that axis [Parsons, 1995]. A number of studies have shown that it can be very difficult for people to judge the axis required to rotate an object into another arbitrary orientation. In fact, near- chance performance occurs, except when an object's major axis coincides with the axis of rotation.

*4) Haptic reference frame.* Considerable attention has been given to the way the non-dominant hand (usually left), holding an object, can provide a frame of reference for the dominant hand (usually right) [Guiard, 1987]. A second issue is that, for biomechanical reasons, certain movements are more easily controlled than others. For example, certain wrist rotations are easier than others [Ware and Baxter, 1989]. Also, the shape of the input device can have an effect. Zhai and Milgram. [1998] suggested that a small ball shaped input device was to be preferred because it can be rotated with the finger tips, whereas a sensor attached to the back of the hand requires ratcheting movements for large rotations.

The four reference frames listed above are by no means all that exist. For example, some people are capable of mentally imagining 3D structures and manipulating these structures in external (allocentric) coordinates. However, people generally find this difficult, and the ability varies widely from one individual to another.

In most computer systems, the input device is separated spatially from the visually presented virtual objects that are being manipulated. This is a translation mismatch between the haptic frame of reference and the visual frame of reference. There can also be a rotational mismatch between the haptic frame of reference and the visual frame of reference. For example, rotating a control knob about a vertical axis might cause a virtual object to rotate about a horizontal axis. Our particular interest in the present paper is the consequences of rotational mismatches in reference frames and interactions between rotational and translational mismatches.

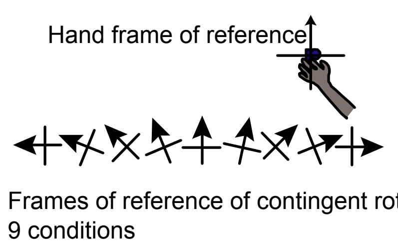


Figure 2. In the first experiment the frame of reference for motion of an object was distorted by rotating its reference frame about a vertical axis. The reference frames for the nine experimental conditions are shown in top-down view.

## 1.3 Study Goals

The goal of the two studies reported here was to explore the consequences of altering the haptic frame of reference with respect to the visual frame of reference. Studies of object translation positioning have shown little ill effect due to having the input device in a different place to the object being manipulated. However it seems likely that creating a rotational mismatch is more likely to cause problems. This can occur, for example, when an input device is misaligned with respect to a monitor.

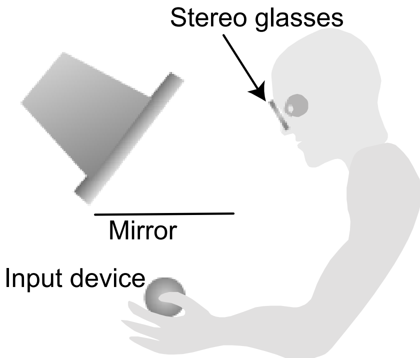


Figure 3. The user's hand can be placed in the same location as the object manipulated using a mirror placed as shown. The subject's actual hand is hidden but he or she sees a virtual image of the object being held in the appropriate place.

## 2 Experiment 1

The goal of the first experiment was to explore the effect of a rotational mismatch between the visual and haptic frame of reference. To study this the haptic frame of reference for object rotation was rotated about a vertical (Y) axis as shown in Figure 2. In the most extreme conditions rotating the input device about the X axis resulted in the virtual object being rotated about the Z axis. In each case the input device was located in the same place as the object being manipulated using the mirror setup illustrated in Figure 3.

### 2.1 The Task

In both of the experiments the task was always to make two virtual objects parallel. One of the objects (the target) was always in a fixed orientation (This is illustrated in Figure 4). The other object (the test) was rotated about a *randomly specified axis* by 90 degrees relative to the target object. The subject then used a ball-shaped input device held in the hand to rotate the test object until it matched the target in orientation. The position and orientation of this ball was continuously measured using a Polhemus FastTrack sensor embedded in it. At the start of each trial, the orientation of the ball became locked though software to the position and orientation of the virtual object. Rotating the ball rotated the virtual object by the same amount, subject to the rotational frame of reference manipulations that changed the mapping from device to virtual object. The subject's task was to rotate the ball so as to make the two virtual objects become parallel. When the object came to within 5 deg. of parallel, the trial was automatically terminated and the results were recorded.

### 2.2 Apparatus

The apparatus is illustrated in Figure 3. The monitor and the mirror were mounted on a custom made stand that reflected the computer graphics imagery. NuVision stereo glasses were used for frame-sequential stereo. A Polhemus FastTrack sensor was placed inside a 5.5 cm diameter ball to act as an input device. Computer graphics was generated using a Wildcat III. and maintained at a 60 Hz update rate.

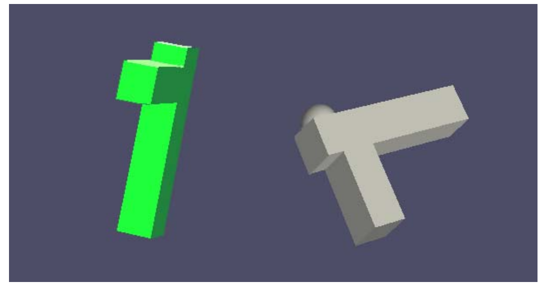

Figure 4. The task was for the subject to make the test object on the right have the same orientation as the target object on the left.

### 2. 3. Conditions

There were 9 different conditions corresponding to rotations of the haptic frame of reference about a vertical axis between –90 and + 90 degrees in 22.5 deg increments.

[-90.0, -67.5, -45.0 -22.5, 0.0, 22.5, 45.0 67.5, 90.0]. Subjects were initially given a training set in which they were given 3 trials in each of the experimental conditions in a random order.

### 2.4 Trials

The experimental trials were given in blocks of five trials in each of the 9 conditions, with conditions randomly ordered. This set of conditions was repeated a total of 4 times to make up an experimental session (resulting in 20 trials per condition). The entire sequence was then replicated in a different session on a different day to yield 40 trials per condition per subject.

### 2.5 Subjects

The 14 subjects were all right handed and had normal stereoscopic depth perception. They were paid for participating.

### 2.6 Experiment 1: Results

Figure 5 summarizes the results. This shows the mean time to make a match for the first and second sessions respectively. As can be seen the data show a broad U-shaped curve, with relatively little ill effect from the rotational mismatch until the mismatch exceeded 50 degree. As the mismatch approached 90 degrees the performance deficit became very large. Thus in the first session, the average time with no orientation mismatch was 4.5 sec but with 90 deg mismatch it was 24.6 and 19.5 seconds depending on whether the mismatch was clockwise or counter clockwise. An analysis of variance revealed highly significant effects for orientation ($F(1,13)=56.0$; $p < 0.0001$), session ($F(1,13) = 26.4$; p

< 0.0001) as well as a session/angle interaction ($F(8,104) = 13.5$; $p < 0.0001$).

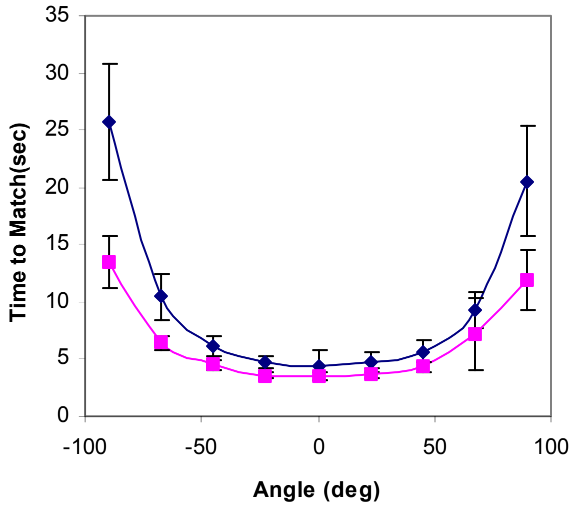


Figure 5. The time to make and orientation match is plotted against the angular mismatch between visual and haptic frames. The upper curve represents data from the first session. The lower from the second session. Vertical bars show 95% confidence intervals (based on subject means).

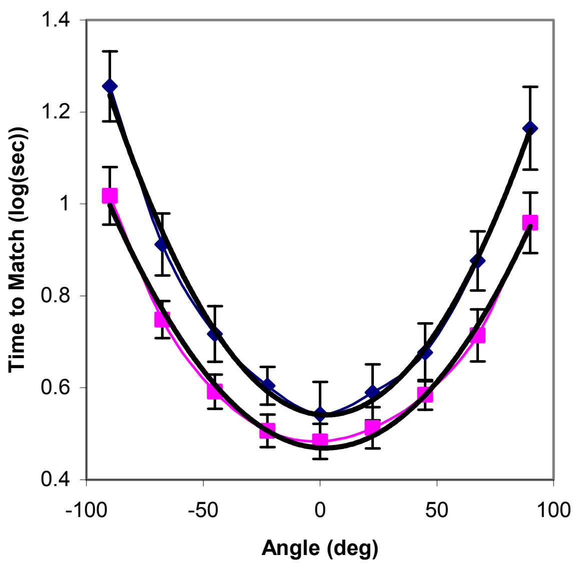


Figure 6. The mean log time to make a match is plotted against the frame of reference mismatch. Smooth curves show a quadratic fit to the data. Bars 95% confidence intervals (based on subject means).

In attempting to find a simple function describing the data we discovered that if the time data were log transformed the results can be closely approximated by quadratic polynomials. The curves fitting the log transform data are shown in Figure 6 and they account for about 99% of the variance due to orientation.

There is considerable improvement between the first and second sessions, especially when there was an extreme mismatch between the visual and haptic displays (-90 deg., +90 deg). Since each of the two session consisted of 4 replications we can generate a time series showing how performance improved with practice. This is illustrated in Figure 7 for each of the different orientation mismatches. As can be seen, performance was continuously improving through the experiment. The best performance was a mean time of 2.4 sec on the last trial set when the visual and haptic reference frames matched.

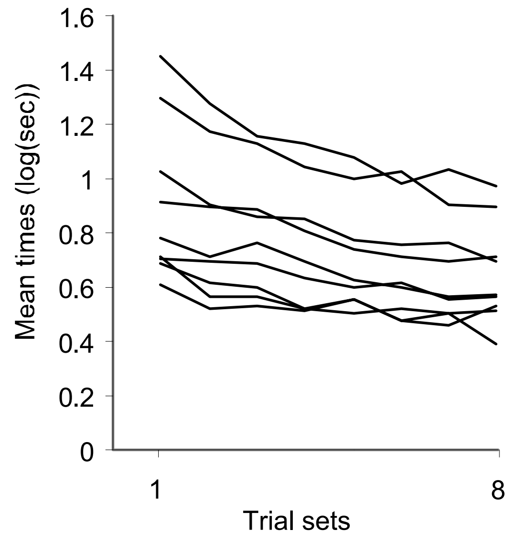


Figure 7. Performance improved over the course of the experiment. The upper curves represent large orientation mismatches while the lowest curve represents small orientation mismatches.

## 3. Experiment 2

Our second experiment had the goal of discovering a possible interaction we hypothesized might occur with both rotational and translational mismatches between reference frames. In many interfaces, the hand holding the input device is held to the side and in front of the monitor where the virtual object is visually displayed (a translational mismatch). It is always assumed, to our knowledge, that the rotational frame of reference should be simply transferred to the new location but there are other possibilities. It might be the case that the way the brain copes with this mismatch is not as a simple Cartesian translation of coordinate systems, but instead involves a rotation of mapping between the haptic and visual coordinate systems in egocentric space. This idea is illustrated in Figure 8.

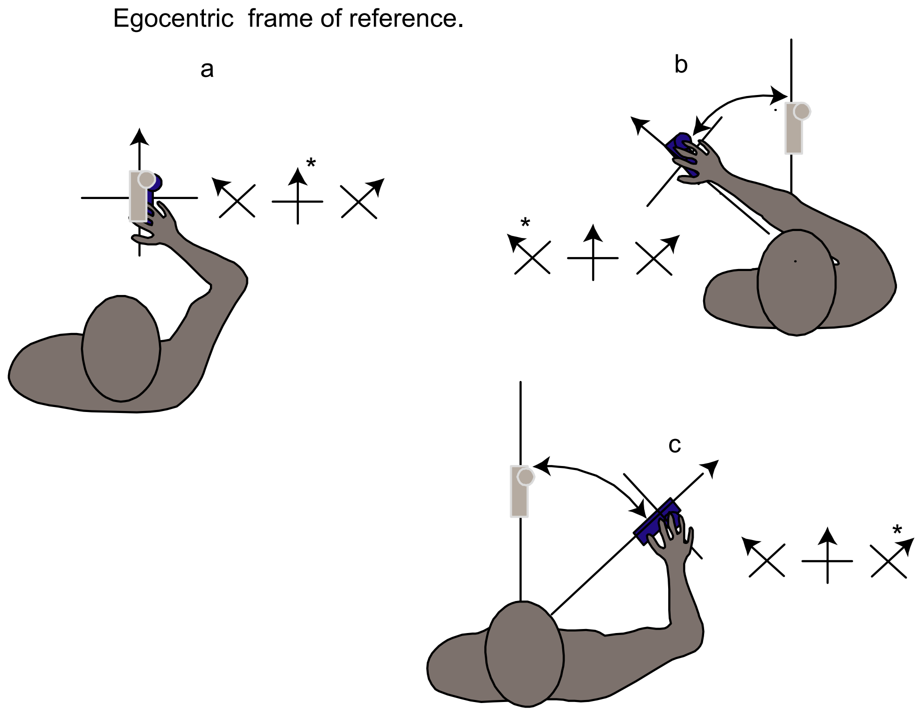


Figure 8. (a) the object manipulated is at the same place as the object seen. (b) the object manipulated is to the left of the object seen. (c) The object manipulated is to the right of the object seen. If there is an egocentric rotation of the haptic coordinate space, relative to the visual coordinate space, the starred conditions may be the easiest.

Our hypothesis was that the rotational frame of reference might become rotated in the egocentric coordinate space as though the subject had turned his or here head to impose the visual image of the object being manipulated on the input device (the starred options in Figure 8). If this were the case then the optimal interface should be one in which the rotational frame of reference of the input device is itself rotated. Our second experiment was designed to evaluate the possibility against a simple translation of frames of reference.

In the second experiment we varied both the position and the orientation of the input device relative to the visual position of the object being manipulated as illustrated in Figure 8. The idea, as discussed, was to look for evidence for a rotation in the mapping between the haptic and the visual coordinated system when the input device and the virtual object were in different locations.

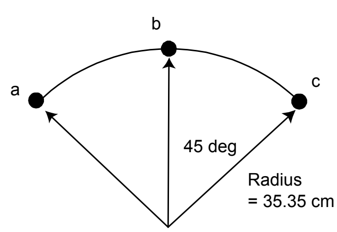


Figure 9. The geometry of the three hand positions used for Experiment 2.

### 3.1 Conditions and Trials

The second experiment had 9 conditions as illustrated in Figure 8. and listed in Table 1. In this experiment, the virtual object to be manipulate was always placed in the same position (the same as that used in E1). However, the subject was required to place their hand in three different locations to make the matching object appear next to the target object. These positions were arranged on a circle having a radius of 35.35 cm. centered approximately below the subject's eye position as illustrated in Figure 9.

Subjects were initially given a training set in which they were given 3 trials in each of the experimental conditions in a random order. As with experiment 1, the experimental trials were given in blocks of five trials in each of the 9 conditions, with conditions randomly ordered. This set of conditions was repeated a total of 4 times to make up and experimental session (resulting in 20 trials per condition). The entire sequence was replicated in a different session on a different day to yield 40 trials per condition per subject.

| Position | Orientation |
|---|---|
| Left (a) | -45 |
| | 0.0 |
| | 45.0 |
| Center (b) | -45 |
| | 0.0 |
| | 45.0 |
| Right (c) | -45.0 |
| | 0.0 |
| | 45.0 |

Table 1. The nine conditions for Experiment 2.

### 3.2 Experiment 2: Results

The results from the second experiment are illustrated in Figure 10. These show log transformed time-to-match data fit with quadratic curves. The way the curves shift based on position clearly suggest that there is a rotation in the mapping between haptic space and visual space in the direction hypothesized. However, the shift is not by a full 45 degrees, but appears to be more like 15-20 degrees, although with so few data points it is not possible to make a more precise determination.

We ran an analysis of variance on the data and found highly significant effect for Position ($F(2,26) = 40.2$; $p< 0.0001$), Session ($F(1,13) = 26.8$; $p< 0.0001$) and Angle ($F(2,26) = 89.6$; $P < 0.001$). The evidence for rotation of the coordinate system is found statistically in an interaction between position and angle ($F(4,52) = 100.5$) $P < 0.0001$). There were no other significant interactions.

### 4. Conclusion

When we initially conceived of these experiments we had expected that users would be quite sensitive to orientation mismatches between the visual and haptic frames of reference and supporting this we found that large angles of mismatch approaching 90 degrees are extremely disruptive resulting in times that are 4 or five-times as long. However, smaller angles of mismatch have a relatively modest impact.

One of the surprising results was the exactness of the quadratic fit to the log time data from Experiment 1. Unfortunately we do not have a theoretical reason for the shape of this curve. Nevertheless, the fact that such a simple function approximates the data may be useful as a practical modeling tool for input device designs.

The most interesting result we obtained was the evidence for rotation of the haptic reference frame relative to the egocentric reference frame. To our knowledge this is a new finding. We found less than half the amount of rotation than might have been expected based on the geometry of our experiment, but always the direction predicted by a rotation in egocentric coordinates. However, the geometry of egocentric space is not well defined, therefore to expect a specific angle of rotation about the head position would be unreasonable. It is even possible to think of our result as a contribution to the definition of egocentric coordinate space.

The result has implications for the way input devices should be physically designed and positioned with respect to the visual workspace. When input devices are designed for rotations of virtual objects, careful consideration should be given to where the device is placed with respect to the display screen. One way of avoiding alignment problems is to use a mirror arrangement to place the hand in the workspace. If this cannot be done then the device should be placed as close as possible to the center of the screen. If the device is to be displaced by a large amount from the virtual object then our results suggest that a rotation, either of the device itself, or of the frame of reference in which it operates, may make it easier to use.

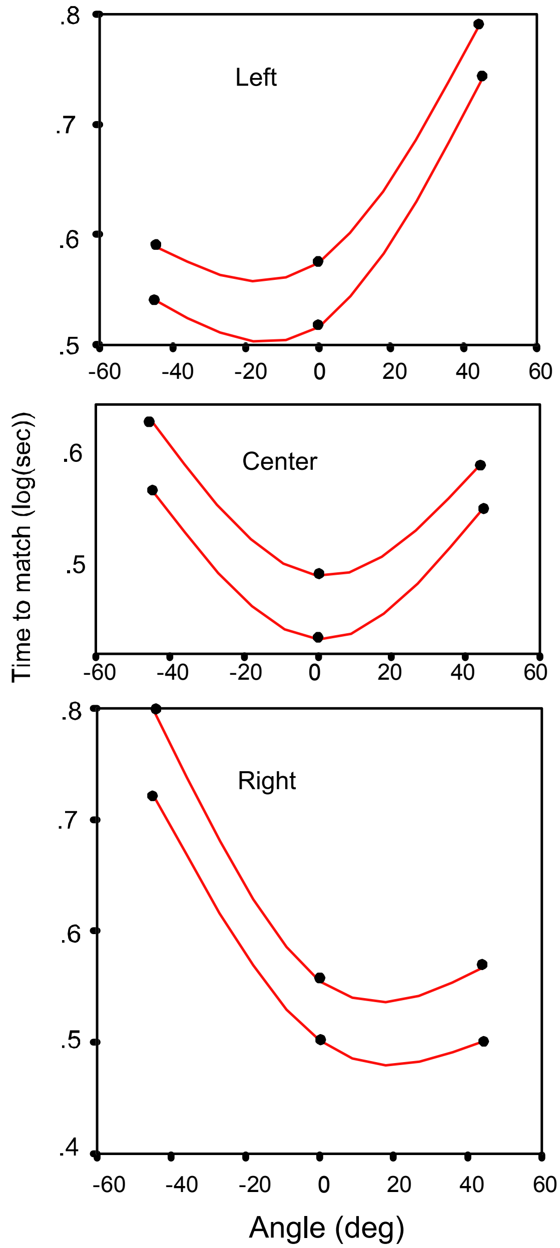


Figure 10. A summary of the results for Experiment 2. There is an apparent shift in the reference frame for rotation as a function of hand position.

Our study raises many questions about the nature of visual-haptic mappings that may occur when the input device is not in the same. We only investigated rotations, however similar effects may occur for translations. We have very sparsely sampled the set of possible manipulations. For example, we only looked at the effects of a mismatch around a vertical axis. The frame of reference may also be rotated about other axes. Only more empirical work will tell.


### Acknowledgements

This work was funded by the National Science Foundation Grant NSF 0081292 and by support from NOAA. We are grateful to Hannah Sussman for help with the experiments.

## References


BALAKRISHNAN, R., AND HINKLEY, K. 1999. The role of kinesthetic refernce frames in two-handed input performance. Proceedings ACM UIST'99 171-178.

CHEN, M., S. J. MOUNTFORD, AND SELLEN, A. 1988. A Study in Interactive 3-D Rotation Using 2-D Control Devices. Computer Graphics 22(4): 121-129.

COLBY, C.L. 1998. Action-Oriented Spatial Reference Frames in Cortex. Neuron, 20 15-24.

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., AND KASSELL, N. 1997. Usability analysis of 3D rotation techniques. Proceedings of ACM UIST'97, 1-10.

PANI, J.R. AND DUPREE, D. 1994. Spatial reference systems in the comprehension of rotational motion. Perception, 23, 929-946.

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

SHELTON, A.L., AND MCNAMARA, T.P. 2001. Systems of Spatial Reference in Human Memory. Cognitive Psychology, 43, 274-310.

SHEPARD, R.N. AND METZLER, J. 1971 Mental rotation of three dimensional objects. Science, 191, 952-954.

SHOEMAKE, K. 1992. ARCBALL: A User Interface for Specifying Three-Dimensional Orientation, Graphics Interface '92.

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.

WARE, C. AND BAXTER, C. 1989. Bat Brushes: On the uses of six position and orientation parameters in a paint program. SIGCHI. 189 Austin, Texas. Proceedings and Special issue of SIGCHI Bulletin. 155-160.

WARE, C. AND ROSE, J. 1999. Rotating Virtual Objects with Real Handles, ACM Transactions on CHI. 6(2) 162-180.

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

ZHAI, C. 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.