Eye-Hand Co-ordination with Force Feedback

Roland Arsenault and Colin Ware

Faculty of Computer Science
University of New Brunswick
Fredericton, New Brunswick
Canada E3B 5A3

Abstract

The term Eye-hand co-ordination refers to hand movements controlled with visual feedback and reinforced by hand contact with objects. A correct perspective view of a virtual environment enables normal eye-hand co-ordination skills to be applied. But is it necessary for rapid interaction with 3D objects? A study of rapid hand movements is reported using an apparatus designed so that the user can touch a virtual object in the same place where he or she sees it. A Fitts tapping task is used to assess the effect of both contact with virtual objects and real-time update of the center of perspective based on the user's actual eye position. A Polhemus tracker is used to measure the user's head position and from this estimate their eye position. In half of the conditions, head tracked perspective is employed so that visual feedback is accurate while in the other half a fixed eye-position is assumed. A Phantom force feedback device is used to make it possible to touch the targets in selected conditions. Subjects were required to change their viewing position periodically to assess the importance correct perspective and of touching the targets in maintaining eye-hand co-ordination, The results show that accurate perspective improves performance by an average of 9% and contact improves it a further 12%. A more detailed analysis shows the advantages of head tracking to be greater for whole arm movements in comparison with movements from the elbow.

Keywords
3d interfaces, haptics, interaction techniques, force feedback, virtual reality.

INTRODUCTION

One of the key arguments for virtual reality systems is that if artificial environments can be constructed that are like the real physical world, then we will be able to apply our everyday life skills in manipulating objects. Thus we will be able to learn to use computer software more rapidly and effectively Applications that could benefit include including 3D CAD, animated figure design for the entertainment industry, and interactive visualization of 3D data spaces. In the present paper we report a study that invstigates the value of eye hand coordination and simulated object contact in a limited, but high fidelelity virtual workspaces.
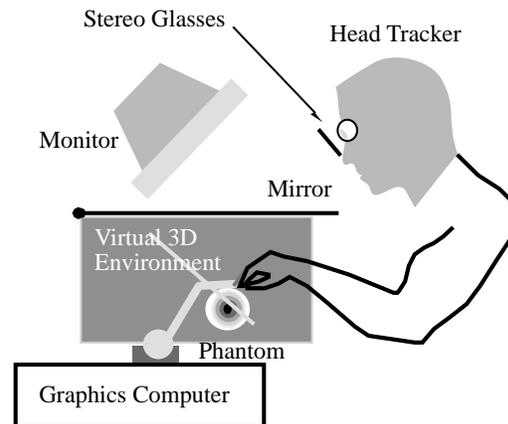

Figure 1. The apparatus used to create a small, high quality virtual environment that can be touched as well as seen.

Fish tank VR is a non-immersive type of virtual reality where a 3D virtual environment (VE) is created using a monitor display [3,17]. In order to create a correct stereoscopic view of a small virtual environment, the user's head position is tracked, from this their eye positions are calculated, and using this information a correct stereoscopic image can be displayed and continuously updated. In essence this involves making the center-of-perspective for the computer graphics coincide with the actual viewpoint for each eye. Using this technique it is possible to create a small, high quality VR environment located just behind and just in front of the monitor screen. With the addition of mirror to reflect the monitor, as shown in Figure 1, the user's hand can be placed in the same location as objects in the VE.

One of the thorniest problems in VR is the fact that although visual information and sound information can be simulated with reasonable fidelity, providing good touch information remains a problem. Recently, force feedback devices have become available that can provide a limited, but reasonable precise sense of touch, but only within a small working volume. The PHANToM, by Sensable Technologies, mechanically measures the

position of a finger tip in 3D space and also applies a force vector to the finger tip [15]. This allows the haptic simulation of solid objects and various force-related effects, such as springs and inertia.

Because of the similarity in the working volume, fish tank VR and Phantom force feedback would seem to be complementary technologies making it possible to combine visual and haptic images. Thus we place a Phantom Force feedback device as shown in Figure 1 to create a local high fidelity VE that can be both seen and touched.

Our goal in the research presented here has been to determine the value of providing real-time head-coupled perspective and of simulated object contact for a simple task. We first review some of the perceptual issues and results from the human factors literature that are relevant to this task

Adaptation

In perception research a number of studies have investigated how eye-hand coordination changes when there is a mismatch between feedback from the visual sense and the proprioceptive sense of body position. A typical experiment involves subjects pointing at targets while wearing prisms that displace the visual image relative to the proprioceptive information from their muscles and joints [7,8,13]. Subjects adapt quite rapidly to the prism displacement and point accurately. Also, after they remove the prisms (having worn them for an extended period) subjects make large errors pointing at targets before recoving. The usual explanation for this is that the mapping between eye and hand has become recalibrated in the brain (although there is much debate as to the exactly where and how this takes place). Recent work by Rosetti et al. [13] suggests that there may be two mechanisms at work in prism adaptation, a long-term slow acting mechanism that is capable of spatially remapping mis-aligned systems, and a short-term mechanism that is designed to quickly optimize accuracy in situations involving temporary misalignmnts.

There is also evidence that certain misalignments are readily compensated for, whereas other are not. Subjects seem to rapidly adapt to small lateral displacements of the visual field, but other distortions, such as inversion of the visual field can take months to adapt to, and adaptation may never be complete [7].

Adaptation experiments, such as those described above are relevant to the present study because we are interested in how useful virtual reality techniques are in making it easier for people to perform certain tasks. If it it possible to adapt quickly and completely to mismatches between hand postion and visual

information, then the case for VR seems much weaker. For example, if objects in small monitor-based virtual environments can be adequately manipulated using the hand placed off to the side, and viewed from a point that is not the center of perspective, then the required equipment will be cheaper and easier to configure. On the other hand, if placing the hand in the same location as a virtual object improves performance then a stronger case can be made that 3D design systems should use VR technologies.

Perspective Distortions

For every perspective picture there is a point, called the *center of perspective* viewed from which, the picture mimics the pattern of light from a scene. When an image is viewed from a point that is different from the correct centre of perspective, the laws of geometry suggest that distortions should occur as shown in Figure 2. However, although people report seeing some distortions when looking a moving pictures from the wrong point they rapidly become unaware of these distortions. Kubovy (1986) called this the *robustness of linear perspective*. One of the mechanisms that can account for this lack of perceived distortion may be based on a built-in perceptual assumption that objects in the world are rigid. If the object shown in Figure 2 were to appear to change shape when the viewpoint was changed, then it would be perceived as elastic and non-rigid. A perceptual rigidity assumption may account for the fact that we perceive stable rigid 3D virtual environments under a wide range of incorrect viewpoints.

Nevertheless, even though the brain appears to compensate for an incorrect viewpoint, there will still be a discrepancy between the visual image and the haptic image if an apparatus such as that shown in Figure 1 is used. As shown in Figure 2, if the displayed object is behind the virtual picture plane, the hand must reach to a different position to be coincident with a virtual object when the viewpoint is not correct. However, a 3D cursor used to make the selection will also be distorted in the same way and this may reduce the ill effects because the *relative* position between the cursor and the object will only be distorted by a small amount. But the extent to which off-axis stereo viewing of a 3D target disrupts target selection has not, prior to the present study, been experimentally investigated.
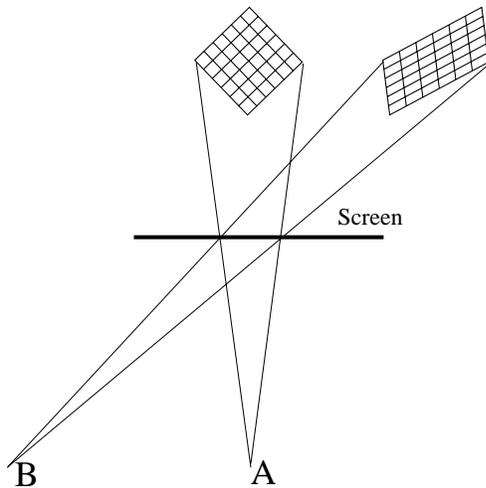
Figure 2. If an image computed to be viewed from position A is actually viewed from position B distortions occur as shown.

Previous results from VR research
There have been a number of studies reported in the human factors and virtual reality literature that bear on the importance of correct viewpoint and haptic feedback.

Ware and Franck showed that accurate perspective based on head position tracking assisted in the task of tracing paths in complex 3D networks [18]. However, they also showed that this is more likely to be a product of motion parallax information than correct perspective; hand linked motion of the virtual scene improved performance as much as providing head- coupled perspective. Recently, Pausch et al showed that using natural head movement to perform a visual search of an immersive environment can result in more rapid searches [12] under certain conditions. Also, head coupled perspective gives a strong sense of the three dimensionality of the virtual space [1].

We may be quite insensitive to translation mismatches between visual and proprioceptive information. In fact the normal practice of placing the mouse at the side of the computer is evidence for this. But there may be a significant advantage to placing the hand in the virtual workspace for object rotations. Ware and Rose [16] found that placing the subject's hand in a virtual workspace improved performance for object rotation, compared to having the subject's hand held to the side of the body.

System lag is likely to be a critical variable in how quickly people adapt to situations in which there is a mismatch between visual and haptic imagery. Held [8] found that the ability to adapt declined rapidly as lag increased beyond about 100 msec.

Simulated touch in object manipulation tasks can improves performance on a number of tasks [14]. Hannaford et al. [6] showed that force feedback reduced errors substantially in the task of placing pegs in holes, and Meek et al [10] showed that the ability to grasp and lift breakable objects was markedly improved with force feedback.

The prior work that comes closes to our present study is an experiment by Boritz and Booth [2] who evaluated a reaching task for targets with and without stereo viewing and with and without head tracked perspective. They found that stereoscopic viewing did improve performance but found no effect for head tracking. However, in this experiment the default head position of the subjects appears to have been close to the correct centre of perspective, thus there may have been little difference between the head-tracked condition and the non head-tracked condition. In addition, the fact that their subjects took several seconds to carry out a simple positioning task suggests that fluid interaction was not possible in their system, perhaps due to system lag.

Although fish tank VR, as described, can provide an accurate correct perspective view calculated from the user's actual viewpoint, this is not always possible or desirable. Head tracking is expensive and requires extra apparatus. Users are generally much more accepting of interfaces where they are unencumbered. On the other hand, when an artist is working on a sculpture or a mechanic is working on an engine they may often change head position to get a better view of what they are working on. Enabling this kind of viewpoint control may be useful and an added benefit to any improvement in eye-hand co-ordination.

In addition, there is the interesting question of whether, simulated contact with virtual objects may make the ability to adapt to an incorrect viewpoint more rapid or complete.

EXPERIMENT

In order to investigate the effects of accurately estimated eye position, and simulated contact we chose a task that could be performed rapidly. In this way we hoped to understand more about skilled fluid performance. The task chosen was the classic Fitts [4] tapping task whereby subjects tap back and forth between two targets. Fitts found that each reciprocal movement, from one target to the next, could be accomplished in less than half a second. However, we did not vary target width and target separation, as in a typical Fitts' law experiment since we were more interested in varying

other task parameters. Although this task is highly artificial, it requires a skill that might be used to rapidly press buttons in a 3D environment. This may become common if VR systems evolve like desktop systems.

The experiment described here had two primary objectives. The first was is determine if head tracking is advantageous when performing rapid, visually guided hand movements. More precisely, does the distortion caused by off-axis viewing of a projected image degrade eye-hand co-ordination? A second issue is whether feedback from physical contact with a target improves performance on the same task.

Method

Apparatus. A virtual environment with a coinciding haptic and visual display was constructed for this experiment. A Phantom 1.0 from Sensable Technologies was used to provide a haptic workspace of 5" x 7" x 10" (12.7 cm x 17.8 cm x 25.4 cm). The Phantom consists of a mechanical arm which tracks the fingertip's position and applies a force vector to the fingertip of the tip of a hand-held stylus [15]. A frame was built above the Phantom to support an upside-down video monitor tilted 45° towards the user to provide an image which was reflected on a mirror placed horizontally between the virtual workspace and the video monitor. The result, when viewed through the mirror is a video display tilted 45° away from the user. This virtual image coincides with the PHANTOM's workspace as shown in Figure 1.

Stereoscopic vision, using LCD shutter glasses, is used throughout. Head tracking is achieved by attaching a sensor from a Polhemous 3Space Isotrack to the stereo shutter glasses. By tracking the position and orientation of the shutter glasses, the position of each eye is calculated and used to provide a correct perspective image to each eye.

The coordinate system used to place objects originates from the center of the workspace with the X axis increasing towards the right, the Y axis increasing in the up direction and the Z axis increasing towards the user. The units of measure used are centimeters. The screen of the visual display can be seen as a plane centered at (0,0,0) with a normal vector perpendicular to the X axis and 45° to the Y and Z axis. A simple 45° rotation in software around the X axis aligns the visual and haptic workspaces.

Calibration of the virtual workspace is verified by replacing the mirror with a pane of glass. It is now possible to place a physical object it the workspace and have a virtual object of similar dimensions superimposed on the physical object. When properly calibrated, the virtual and physical objects remain it the same position when the head is moved.

Task. Subjects are asked to alternately tap the tops of two cylindrical targets. The targets are cylinders oriented such as the flat faces are parallel to a checkerboard ground plane as illustrated in Figure 3b. The cylinders can be seen visually and felt with haptic feedback. The cylinders have a radius of 1 cm each and are separated horizontally by 6.75 cm. Two sets of positions for the cylinders are used for this experiment as illustrated in Figure 3a. For right handed subjects, with position 1, tapping can be accomplished mainly by arm rotations of the forearm about the elbow. Whereas in position 2 subjects move their entire arm from the shoulder in order to tap back and forth. The overall location of both targets is randomly changed on successive trials by up to 1.0 cm on each axis, but the relative position of the two cylinders to each other is unchanged.

In all conditions subjects held the Phantom stylus in their right hand and used it to tap back and forth touching to tops of the targets in succession.

Subjects are required to change their viewpoint between trials. A 10 cm wide obstacle placed on the mirror which prevents the subject from viewing the targets from a central position. In order to view the target objects the subject moves his or her head approximately 18 cm left or right of the center. Since the subjects eye point is typically about 55 cm from the target area this results in a line-of-sight about 18 degrees off-axis. A signal in the form of sphere on the upper left or upper right portion of the display appears to indicate from which side of the obstacle the subject should look at the targets. The side is changed after every trial of 12 taps and three trials per side are run for every condition.
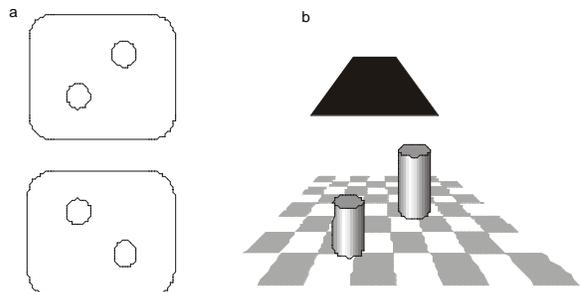

Figure 3, (a) Two sets of target positions are used in a right oblique and left oblique configuration. (b) a physical barrier, placed horizontallon on the mirror above the virtual targets forces the subject to look from one side or the other.

Conditions

There are three major independent variables.

Head-tracked vs non head-tracked. In the head tracked condition, the center of perspective is based on the users eye position (computed based on their measured head position). In the non head tracked condition a default center of perspective is used for each eye. This is at the mid point of the normal range of head movement.

Touch vs no touch.
In the force condition, force feedback is provided by the Phantom to provide a sense of contact with a hard surface. In the no-force conditions visual feedback for contact with the target is provided by making the target flash to a higher color intensity for a single frame of animation at the moment of contact.

Target position. The two sets of positions for the target cylinders are as illustrated in Figure 3a and described above.

All combinations of the three independent variables are tested giving a total of 8 conditions.

A trial consists of twelve successive taps back and forth between the tops of the two targets, six taps on each. On alternate trials subject change their head position alternately looking at the target from the right or left of the barrier. A trial block consists of six successive trials, that are the same with respect to head tracking (or not) and virtual contact (or not). A run consists of all possible trial blocks occurring in random order. The experiment consists of two runs per subject in one sitting for a total of 96 trials.

The subjects are allowed to try the task before measurements were made to familiarize themselves with the virtual environment. Once ready, the subjects are instructed to tap the targets, always starting with the green one (target 0). They are instructed to tap as fast as possible back and forth until a beep is heard. At that point, the subjects are asked to move their head position to view from the other side of the obstacle, as indicated by a red sphere appearing in the top of the workspace. At this point, before the targets where touched again, the user can take a small break to rest if desired.

Subjects. 13 subjects were chosen from within and outside the university population. 2 subjects had previous experience with the virtual environment. All subjects were right handed.

	No Force Feedback	Force Feedback	Average
No Headtracking	612	537 (-12%)	574
Headtracking	557(-9%)	491 (-20%)	524 (-9%)
Average	584	514 (-12%)	549

Table 1 - Average time (ms) for various conditions.

RESULTS

Table 1 shows the mean inter-tap interval averaged across all subject and all trials for the two main conditions. The overall mean interval was 549 ms. Using head tracking to compute the correct veiwpoint resulted in a reduction of 9% in the mean inter-tap interval. Using force feedback resulted in a reduction of 12% in the inter-tap interval. Both of these differences are highly significant ($p<0.01$). There was no significant interaction between them.

Each trial actually consisted of 12 taps giving 11 inter-tap intervals. Figure 4 shows a time series of inter-taps intervals averaged across all subjects and other conditions for head tracked and non head tracked conditions. Figure 5 shows the same series comparing performance both with and without force feedback. As can be seen over the course of each series the inter-tap interval decreased over the first four taps and then levelled off, but against our expectations there is no closing of the gap that might be expected from a rapidly acting eye-hand re-calibration.
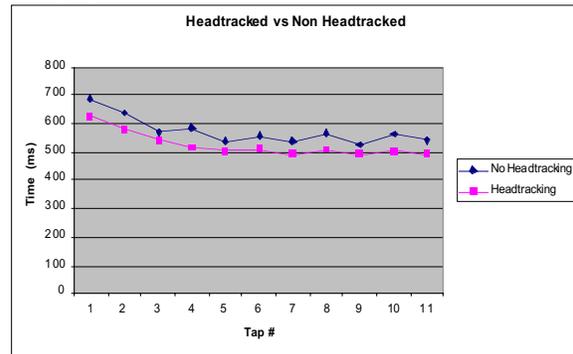

Figure 4. The average time series of inter-tap intervals is given both with and without head-tracked perspective.
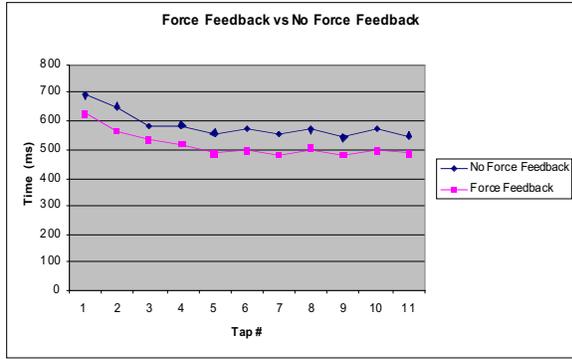
Force Feedback vs No Force Feedback

Figure 5. The average time series of inter-tap intervals is given both with and without force feedback.

For each condition there were six trial blocks divided into two runs and over the course of the experiment subjects speeded up from a mean inter-tap interval of about 605 ms. to about 535 ms.

Head tracking improved performance more for target positions 2 than for target positions 1. Figures 6 and 7 show the results both with and without head tracking with targets in positions 1 and 2 respectively. As can be seen there was approximately a 25 ms benefit for head tracking with the targets in position 1 and an 80 ms benefit in position 2. All of the subjects were right handed and position 2 required whole arm movements from the shoulder, whereas position 1 only required movements of the forearm.
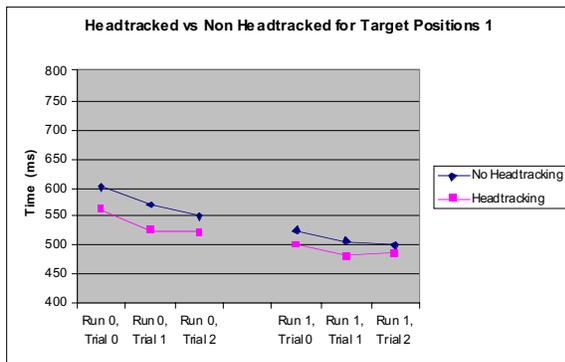
Headtracked vs Non Headtracked for Target Positions 1

Figure 5. The results are plotted over the time course of the experiment for targets in position 1.

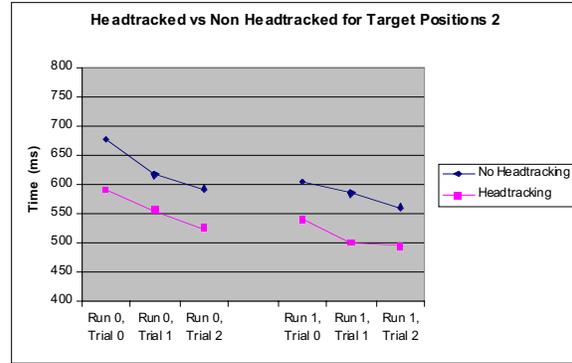**Headtracked vs Non Headtracked for Target Positions 2**

Figure 6. The results are plotted over the time course of the experiment for targets in position 2.

Errors occurred when a subject failed to make contact with a target yet kept on tapping. This resulted in inter-tap intervals approximately 3 times as long as the norm as two extra movement were required before the "next" target was registered (i.e. the one that had been missed). We devised the following post processing strategy to deal with these occurrences. If the individual time was greater than 2.25 times the average, this time was treated as an error and corrected by dividing it by 3.

Table 2 shows the errors broken down by the major conditions. The largest effect was that there were fewer errors with force feedback than without force feedback. This difference was significant ($p < 0.05$).

There is a method, originally developed by Welford, whereby error rates can be combined with tapping times to create a single unified metric of performance [11]. When this method is applied to the force data it shows an additional 2.8% advantage to using force feedback. Thus we get an almost 15% overall benefit.

	No Force Feedback	Force Feedback	Total
No Headtracking	3.29%	2.74%	3.02%
Headtracking	3.50%	2.80%	3.15%
Total	3.39%	2.77%	3.08%

Table 2 - Errors detected

CONCLUSIONS

The ultimate goal of virtual reality systems is to allow people to work naturally and efficiently at a variety of tasks. The contribution of this paper has been to show that for a rapid tapping task, having a perspective view computed for the observers actual eye position can

speed-up performance, although by a relatively small amount. In addition, making simulated contact with the targets also improves performance.

Our results differ from those of Boritz and Booth [2] in that we found an effect of head-tracker perspective, whereas they did not. One likely reason is mentioned in our introduction; they did not require their subjects to make head movements. Since the viewpoint in their non-headtracked condition was presumably quite close to the correct centre of perspective, there may have been very little difference between what the subjects saw in their headtracked and non-headtracked conditions. Thus their result cannot be taken as evidence that viewing a perspective image from an incorrect viewpoint has no ill effects. We forced head movements in the task that we devised and found a clear effect.

We measured an advantage for simulating contact using the Phantom. We are grateful to Christine MacKenzie (personal communication) for pointing out to us that the tapping task with force feedback engaged is actually a rather different task to the task without force feedback. In the no-feedback mode, subjects actually made the cursor move *through* the disc shaped target region and back in order to register a target hit; this required less effort than moving and bringing the cursor to a halt in the target centre. Conversely, in the force enabled condition. The cursor could be bounced off a target actually speeding its progress back to the other target. However, this should not be regarded as necessarily a flaw in the design. The constraints provided by the physical environment alter the characteristics of many real-world tasks, often making them easier. Exploiting such synergies may be the most compelling reason for introducing force feedback into virtual environments.

Our results only show quite small benefits to providing a correct perspective view and force feedback, and thus might not seem to warrant the considerable technology involved. However, skilled designers can take advantage of excellent tools. Taken together, including both head tracking and force feedback, improved tapping performance by 20% and, in addition, reduced errors. If the goal is to achieve fluid and highly responsive environment, this advantage may be accrued in every interaction; such small gains can easily make the difference between an environment that is a pleasure to use and one that is barely acceptable. In our experience the combinations of these technologies provides a compelling localised virtual reality experience. We are confident as costs drop and the systems improve, this kind of apparatus may provide effective support for designing virtual sculpting systems such and 3D CAD operations.

REFERENCES

1. Arthur, K. W., Booth, K. S., and Ware, C. (1993). Evaluating 3D Task Performance for Fish Tank virtual Worlds. *ACM Transactions on Information Systems,* 11(3) 239-265.
2. Boritz, James and Booth, Kellogg S., A Study of Interactive 3D Point Location in a Computer Simulated Virtual Environment, In *Proceedings of ACM Symposium on Virtual Reality Software and Technology '97*, Lausanne, Switzerland, Sept., pp. 181-187.
3. Deering, M. (1992). High Resolution Virtual Reality. *Proceedings of ACM SIGGRAPH '92, Computer Graphics*, **26:2**, 195-202.
4. Fitts, P.M. (1954) The information capacity of the human motor system in controlling the amplitude of movement. Journal of Expeimental Psychology, 47(6) 381-391.
5. Graham, E.D. and MacKenzie, C.L. (1996) Physical versus virtual pointing. Proceedings of ACM CHI'96. 292-299.
6. Hannaford, B. Wood, L. Guggisberg, b., McAffee, D. and Zack, H. (1989) Performance evaluation of a six-axis universal force-reflecting hand controller. In Proceedings of the 19th IEEE Conference on Decision and Control. Albuquerque, NM, Dec, 1197-1205.
7. Harris, C. S. (1965). Perceptual adaptation to inverted , reversed, and displaced vision. *Psychological Review,* **72**, 419-444.
8. Held, R., Estanthiou, A. and Green, M. (1966) Adaptation to displaced and delayed visual feedback from the hand. Journal of Experimental Psychology, 72, 887-891.
9. Kubovy, M. (1986) The psychology of perspective and renaissance art. Cambridge University Press.
10. Meek, S.G., Jacobson, S.C., and Goulding, P.P. (1989) Extended physiologic taction, design and evaluation of a proportional force feedback system. Journal of Rehabilitation Research and Development, 2693) 53-62.
11. MacKenzie, I.S. (1992) Fitts' Law as a Research and Design tool in human-computer interaction. Human-Computer Interaction, 7(1) 91-139.
12. Pausch, R., Proffitt, D., & Williams, G (1997) Quantifying immersion in virtual reality, ACM SIGGRAPH`97 Conference Proceedings, Computer Graphics.

13. Rossetti, Y., Koga, K., and Mano, T., (1993) Perception and Psychophysics, 54(3) 355-364.
14. Sheridan, T.B. (1992) Telerobotics, Automation and Human Supervisory Control, MIT Press: Cambridge, Mass.
15. Tan, H. Z., Srinivasan, M. A., Eberman, B., Cheng, B. (1994). Human Factors for the Design of Force-Reflecting Haptic Interfaces. *Dynamic Systems and Control,* 55:1
16. Ware, C. and Rose, J. (in press) Rotating Virtual Objects with Real Handles, ACM Transactions on Comptuter-Human Interaction
17. Ware, C. Arthur, K., and Booth, K.S. (1993) Fish tank virtual reality. In Proceedings of ACM INTERCHI'93, 37-42.
18. Ware, C., and Franck, G., Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three-Dimension. ACM Transactions on Graphics, 15(2) 121-140.