Paul Richard
Georges Birebent
Philippe Coiffet
Laboratoire de Robotique de Paris
UPMC-UVSQ-CNRS
10-12 Avenue de l'Europe
78140 Velizy, France

Grigore Burdea
Daniel Gomez
Department of Electrical and Computer Engineering
Rutgers—The State University of New Jersey
Piscataway, NJ 08855-0909

Noshir Langrana
Department of Mechanical and Aerospace Engineering
Rutgers—The State University of New Jersey
Piscataway, NJ 08855-0909

Presence, Vol. 5, No. 1, Winter 1996, 95–108

Effect of Frame Rate and Force Feedback on Virtual Object Manipulation

Abstract

Research on virtual environments (VE) produced significant advances in computer hardware (graphics boards and i/o tools) and software (real-time distributed simulations). However, fundamental questions remain about how user performance is affected by such factors as graphics refresh rate, resolution, control latencies, and multimodal feedback. This article reports on two experiments performed to examine dextrous manipulation of virtual objects. The first experiment studies the effect of graphics frame rate and viewing mode (monoscopic vs. stereoscopic) on the time required to grasp a moving target. The second experiment studies the effect of direct force feedback, pseudoforce feedback, and redundant force feedback on grasping force regulation. The trials were performed using a partially-immersive environment (graphics workstation and LCD glasses), a DataGlove, and the Rutgers Master with force feedback. Results of the first experiment indicate that stereoscopic viewing is beneficial for low refresh rates (it reduced task completion time by about 50% vs. monoscopic graphics). Results of the second experiment indicate that haptic feedback increases performance and reduces error rates, as compared to the open loop case (with no force feedback). The best performance was obtained when both direct haptic and redundant auditory feedback were provided to the user. The large number of subjects participating in these experiments (over 160 male and female) indicates good statistical significance for the above results.

1 Introduction

Virtual environments (VE) enhance user–computer interaction by providing a real-time, multisensorial interface to the simulation. It is important to understand how people perceive and respond to events in synthetic environments in order to develop better VE technology.

Over the past 5 years virtual environment research has resulted in significant advances in computer hardware [three-dimensional (3D) graphics boards and i/o tools] and software (real-time, multiuser distributed simulations). Little research has been conducted to date concerning human performance issues in the context of this new technology. Thus fundamental questions remain about how people interaction with VE is influenced by factors such as graphics refresh rates, mono/stereo viewing, control latencies, and multimodal feedback.

Human interaction with the environment is primarily visuomotor, and vision is the dominant source of feedback concerning the effects of actions. For instance, over the course of a normal day, a person makes hundreds of limb movements to manipulate objects in space. To produce such actions one must use visual information about the position, velocity, and acceleration of the object of interest (Georgopoulos, Kalaska, & Massey, 1981; Atkeson & Hollerbach, 1985).

In virtual environments the perceived object position and velocity with respect to the user are subject to perceptual distorsions. One of the most important parameters that creates such perceptual distorsions is the graphics frame rate—the maximum rate at which new virtual scenes are presented to the user. Low frame rates make objects appear to move in saccades (discrete spatial jumps in 3D space). The visual system thus has to bridge the gaps between perceived positions by using spatiotemporal interpolation.[1]

The effect of video frame rate on human performance has previously been investigated by Ranadive (1979) and Massimino & Sheridan (1989). However, these studies involved manipulation of static objects using servo manipulators and video displays. The frame rate had no effect on visual perception of static objects in the environment, since only motion perception was affected by frame rate. Airey, Rohlf, & Frederick (1990) later showed that a frame rate of 6 frames per second (fps) is the minimum rate for an interactive VE system. However, this study involved self-motion perception in a static virtual environment.

The belief of the VE research community is that the higher the frame rate, the better the visual feedback (typically 30 fps is desired). These high frame rates coupled with today's slow graphics rendering do, however, limit the complexity of the virtual world. The more realism required by the VE application, the more complex the world model is, and the smaller the number of frames per second displayed by the computer. It is, therefore, of interest to study user performance at slow frame rates. A related issue is whether at low frame rates stereo visual feedback is better than monoscopic viewing. The first study presented in this article considers the influence of frame rate and viewing mode on a tracking and grasping of a moving object.

One other key problem when interacting with virtual objects is the influence of haptic feedback. Being able to touch, feel, and manipulate objects in the environment, in addition to seeing (and/or hearing) them, gives a sense of compelling immersion that is otherwise not possible. Haptic feedback is not necessary in all applications, but it is mandatory for dark or visually occluded scenes.

Earlier research on integrating force feedback to VE simulations was done by the robotics community, which adapted existing teleoperation servo-arms (Brooks, Ouh-Young, Batter, & Jerome, 1990; Jacobsen, Inverson, Davis, Potter, & McLain, 1989; Jau, 1993). Steps toward less complex systems were taken by integrating force feedback to joysticks (Minsky, Ouh-Young, Steele, Brooks, & Behensky, 1990; Schmult, 1990; Iwata, 1993). Unfortunately, force feedback joysticks limit the hand motion to a small volume close to the desk, thus reducing the user's freedom of motion. Additionally, joysticks cannot provide force feedback to independent fingers, thus are not well suited for highly dextrous tasks, such as precision grasping of virtual objects (Cutkosky & Howe, 1990). Most of the studies reported to-date involve wrist force feedback (Millman, Stanley, & Colgate, 1993; Ellis, Ismaeil, & Lipsett, 1993). Much less work has been done for dextrous VE with nonportable haptic feedback, where interaction is at the finger tips (Massie & Salisbury, 1994; Hashimoto, Boss, Kuni, & Harashima, 1994).

To provide finger force/touch feedback to the user without sacrificing freedom of motion, force/tactile feedback interfaces have to be portable. Such portable prototypes have been developed in recent years (Patrick, 1990; Stone, 1992; Cutt, 1993; Bouzit, Coiffet, & Burdea, 1993). Portable masters need to be light, to prevent user fatigue during prolonged simulations (Burdea & Langrana, 1993). This is a design challenge due to the poor power/weight ratio of most of today's actuators. The Rutgers Master (Burdea & Zhuang, 1992) and Rutgers Master II (Gomez, Burdea, & Langrana, 1994) are force feedback devices that use light, silent, and safe pneumatic actuators. Previous tests with the Rutgers Master showed that haptic feedback enhanced manipulating task performance when interacting with virtual objects by about 50%, and reduced the learning time by 50% (Richard, Burdea, & Coiffet, 1993). These early

1. Spatiotemporal interpolation creates the impression of motion from a sequence of stationary images by reconstructing the motion path in between the stations actually presented.

tests were done with a monoscopic display and low graphics frame rates (6–7 fps). In the second part of this article we repeat earlier experiments, but using a higher frame rate (27 fps), a stereo display, and feedback to four instead of three fingers.

Providing force feedback to the operator's hand can have its disadvantages. For example, significant static friction in the haptic master may not allow the user to feel feedback forces during manipulation of very soft virtual objects. Moreover, presenting force feedback in the presence of even small delays has been shown to create operator-induced instabilities (Ferrel, 1966). In such cases, sensory substitution or information redundancy could be used to open the power loop or to help the operator apply small stable forces on the manipulated object. Therefore the present study includes the influence of redundant force feedback (visual or auditory cues) on subject's performance.

2 Temporal Resolution of VE Visual Displays

The temporal resolution of VE visual displays (number of updated images presented per unit time) is affected by a number of factors, such as hardware refresh rate, graphics frame rate, and delays. The hardware refresh rate is the number of images seen by each eye every second, and varies between 60 and 120 Hz. It differs from the graphics frame rate, which is the maximum rate at which new virtual scenes are presented to the user. For instance, if a display system has a refresh rate of 60 Hz, but a frame rate of only 4 fps, then the system presents 15 consecutive identical images before any changes are made in the scene.

Low frame rates produce annoying illusory motion artifacts, which are compounded by the system delays (from sensing, communication, graphics, and dynamics computations). The computational delay is a major factor in determining the frame rate of a display system. Sensorial delays represent the time between user input (position or force) and the corresponding sensor output. Sensorial, communication, and computational delays are additive. For instance, a system that has a 12 fps, 10 msec communication delay and 150 msec sensor delay, would produce a maximum delay of 260 msec whenever the viewer has moved enough to require the synthesis of a new image of the virtual world. If the delays or simulation "lag" are larger than 300 msec, then the user loses the immersive feeling in the simulation (Wloka, 1995). The more dynamic the virtual environment, the more important the simulation delays become. A case in point (and one that is the subject of the present study) is tracking and capturing a moving object while subject to low graphics refresh rates.

There are two types of visual displays used in VE simulations: monoscopic and stereoscopic. In a monoscopic display, a single 2D image is presented to both eyes. The human brain interprets the 2D monoscopic image as a 3D image by utilizing depth cues (perspective, occlusions, motion parallax, and so on). In stereoscopic vision the brain fuses two slightly different views of the world (seen by the two eyes) into one coherent 3D image. Stereoscopic displays provide distinct depth cues when the scene of interest is close to the user (Kim et al., 1985, 1987). It is of interest to find out if viewing a dynamic scene on a stereoscopic (as opposed to monoscopic) display can help the user regain some of the immersion feeling in the presence of low refresh rates. To do that the user's performance needs to be measured using objective criteria (such as error rates, task completion time, and task learning time).

3 Force Feedback and Sensorial Substitution

The visual and auditory sensorial channels have one-way, information-only flows. This involves collection and analysis of photons (for vision) and sound waves (for hearing) coming from the environment. Although head movements play a key role in 3D visual and aural perception, they do not involve energetic interactions with the physical world. By contrast, haptic sensations fundamentally involve a bidirectional flow of energy between the human and the environment (National Research Council, 1995).

The visual and auditory senses are of interest for haptic feedback research since they provide nonreactive representations of force feedback.[2] Thus, unlike force feedback, they do not cause instabilities in the presence of time delay. In addition, vision and hearing are allocated relatively large areas in the sensory cortex (Shreeve, 1993), suggesting that visual and auditory displays could be successful in presenting force feedback information.

Bach-y-Rita, Webster, Tompkins, and Crabb (1987) define sensory substitution as the provision to the brain of information that is usually in one sensory domain by means of the receptor, pathways, and brain areas of another sensory system. Sensory substitution has been successfully used for many years in helping people who are fully or partially deficient in one or more of their sensory systems (Mann, 1974; Bach-y-Rita et al., 1987). Sensory substitution has also been used in the field of teleoperation (Bejczy & Paine, 1982; Corker, Bejczy, & Rappaport, 1985; Reger, 1987; Ouh-Young, Pique, Hughes, Srinivasan, & Brooks, 1989; Massimino & Sheridan, 1994) to provide pseudohaptic feedback through auditory or visual cues. This was necessary due to the lack of adequate haptic feedback hardware. Sensory substitution may also serve as a redundant cue in the simulation. Redundant force feedback (by adding an auditory cue) may be used to accentuate certain simulation events, thus improving user's efficiency. Too much redundancy is detrimental, as it may lead to sensorial overload and user disorientation (Burdea & Coiffet, 1994). It is thus necessary to study sensorial substitution in general, and force feedback substitution and redundancy in particular, to determine what is beneficial and what is detrimental.

4 Experimental Study

The above mentioned problems related to simulation delays and haptic feedback substitution led us to perform a study of dynamic VE interaction considering the problems of 1) low refresh rates and 2) force feedback sensorial substitution and redundancy. We present first the experimental system, followed by a description of the experimental protocols and test results.

4.1 Experimental System

The Rutgers VE system has a loosely coupled, client-server architecture, as illustrated in Figure 1. This approach allows for the distribution of computation for the simulation on two workstations (Burdea, Roskos, Gomez, & Langrana, 1993). The first is a Sun 4 dedicated to reading, calibrating glove data, and maintaining state information on all virtual objects. An HP 755-48Z workstation is dedicated to graphics rendering and display of the virtual world at rates up to 28 fps. Our experiments used a single screen with time-multiplexed left and right views, and LCD shutter glasses. Alternating left and right images are synchronized with the shutter glasses through an infrared (IR) controller (Lipton, 1991). The HP755 monoscopic display (1280 × 1024 pixel resolution) was replaced by a special stereo monitor with 120 Hz refresh rate required for the use of LCD glasses.

A DataGlove™ (VPL, 1987) measures hand gestures and wrist 3D position using a Polhemus™ tracker (Polhemus Navigation, 1987). The tracker accuracy (Isotrack™ model) is about 2.5 mm translation, and 1° rotation. The Rutgers Master (Burdea, Langrana, Roskos, Silver, & Zhuang, 1992) is used for force feedback to the user's fingers. Its compact structure fits in the palm of the DataGlove, as shown in Figure 2. The feedback actuators consist of four pneumatic microcylinders that press against the finger tips. The lightness of the feedback structure (about 100 g) is important to reduce operator fatigue during the simulation.

The feedback pressure is controlled by analog proportional regulators (PPR) housed in an interface box. The box displays virtual force feedback: (1) haptically (through the feedback actuators), (2) visually through four sets of 20 LEDs located on the front panel (one set per finger), and (3) auditory through headphones. The number of LEDs "on" is proportional to the level of the virtual force feedback computed by the simulation. We decided to display auditory force feedback by controlling

2. Nonreactive sensory modalities do not induce operator finger movements. Such movements may be undesirable in certain situations.

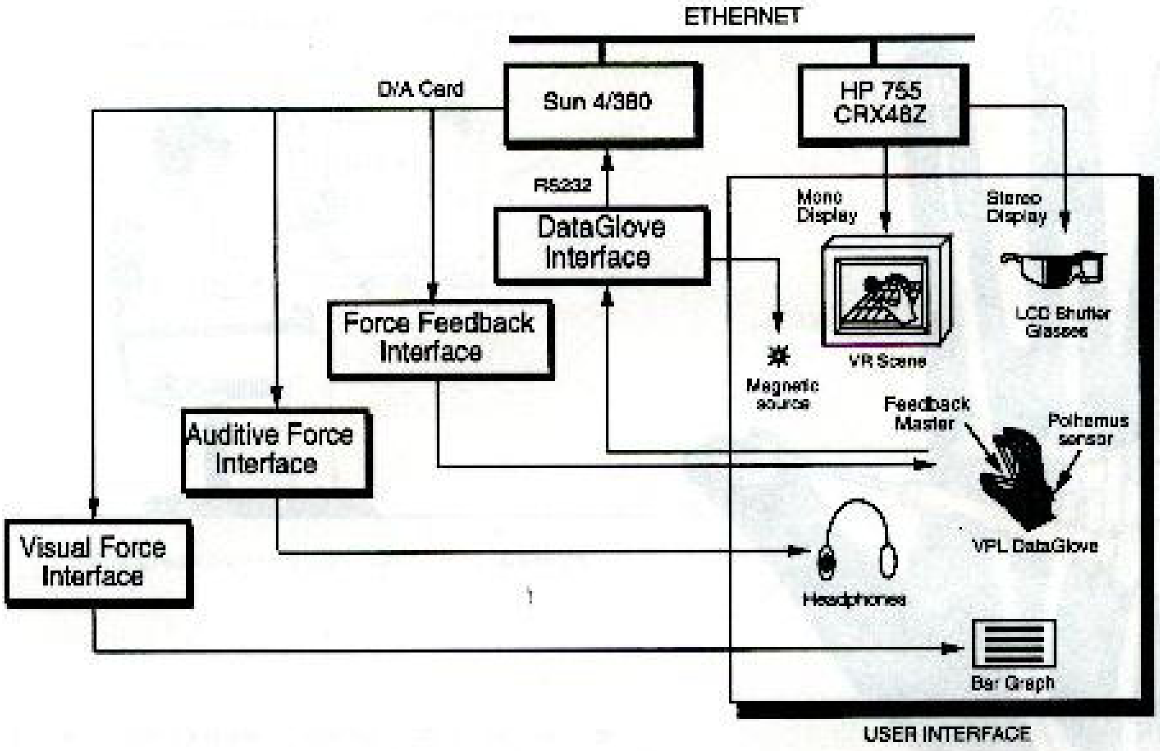

Figure 1. The experimental system architecture

the frequency (between 200 Hz and 1 kHz) of the audio signal since the minimum detectable change in this case is 2 Hz. The frequency of the audio signal is controlled by the average of the four virtual haptic feedback signals. A force feedback of 0.2 N corresponds to one LED "on" when the force feedback is displayed visually, or to a frequency of 50 Hz for auditory feedback.

The experimental virtual environment (illustrated in Fig. 3) is composed of black walls with perspective grids, a virtual hand, and a deformable red ball (target). Visual depth cues such as projection of object "shadows" can be introduced in the environment to improve depth perception. The volume of the virtual room is about 1 m^3. The virtual hand is made of 129 polygons and has similar kinematics to the human hand (less finger adduction/abduction). The ball is made of 72 polygons. The simulation uses the Starbase Graphics Library (Hewlett Packard, 1991), double buffering, and Gouraud shading with one light source. The ball, depending upon its compliance, deformed approximately as it would in the physical world.

Feedback forces (F_i) were calculated individually, according to the simple Hooke's law ($F_i = k\Delta x_i$). Here Δx_i represents the deformation at the contact point between finger i and the ball. In this way the equation is kept simple for rapid computation while still retaining the ability to model objects of varying stiffnesses k. Orientation of the contact forces was kept normal at the four grasping points on the ball. The maximum deformation of the ball corresponds to a force of approximately 4 N/finger.

5 Experiment 1 (Tracking and Grasping Task)

This experiment concerns the effect of frame rate on human performance in a tracking/grasping task of a 3D moving target, for both monoscopic and stereoscopic displays. Two groups of 42 subjects each (42 males and 42 females) were used in the experiment. The first group performed the tracking task using the mono-

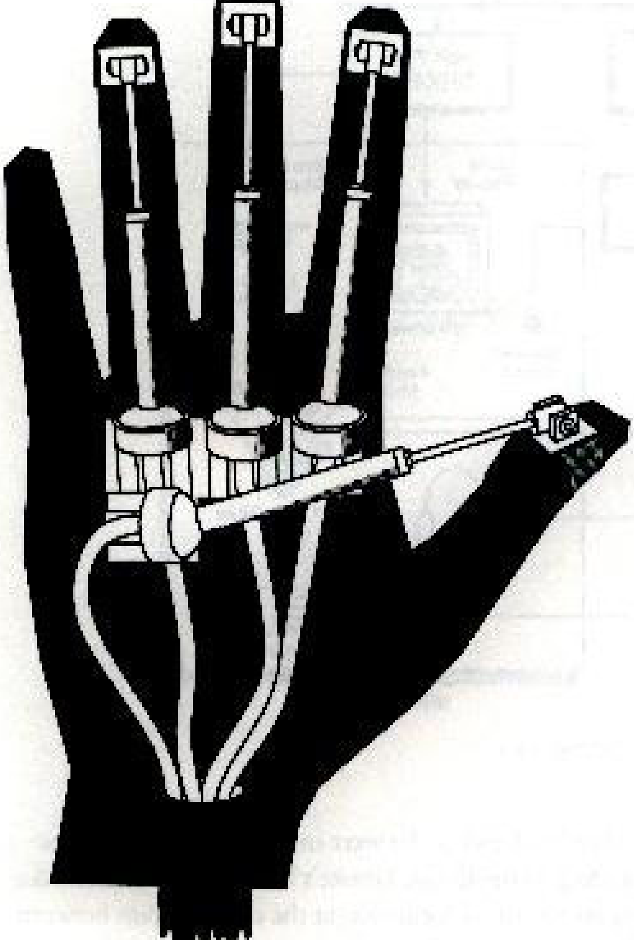

Figure 2. The Rutgers Master (Burdea et al., 1992)

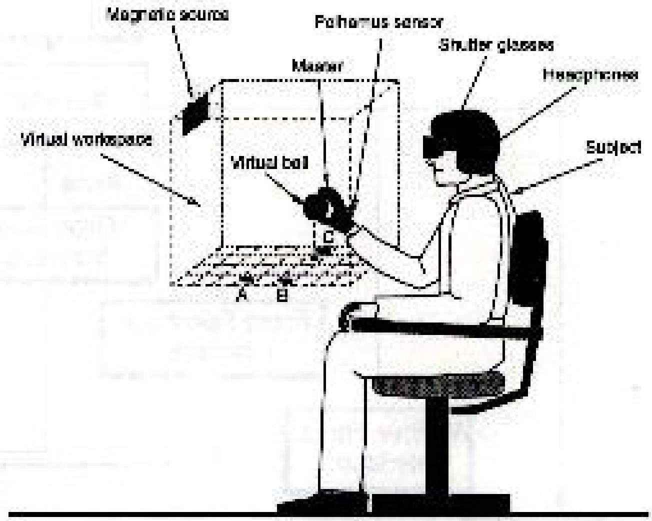

Figure 3. Experimental virtual environment configuration.

scopic display, while the second group performed the same task using the stereoscopic display. Each group was divided into 6 subgroups (G1, to G6) of 7 subjects each. G1 practiced capturing the target at a frame rate r1 = 28 fps, G2 at r2 = 14 fps, G3 at r3 = 7 fps, G4 at r4 = 3 fps, G5 at r5 = 2 fps, and G6 at r6 = 1 fps.

5.1 Experimental Protocol

Each trial was run as follows: A red target ball (of 7 cm diameter) appeared in the virtual room (of 1 m^3) at the same location (marked "C" in Fig. 3) and with the same initial velocity (25 cm/sec). The initial ball direction was changed randomly between trials, within a 45° cone. The base of the cone was facing the bottom wall of the virtual room. So, the target path varied between trials, but its velocity was always the same (25 cm/sec). During collisions with the virtual walls, the ball bounced without loss of energy, so that task difficulty did not drop with time.

Subjects were seated so that the eye-to-monocular display distance was approximately 60 cm. In the stereoscopic viewing mode, stereo glasses were worn. The DataGlove® was fitted to their right hand, then a quick calibration was done. Subjects performed one trial session to familiarize themselves with the limits of the virtual room and the grasping (gesture recognition) technique. Grid lined perspective walls were superimposed on the virtual walls. Laboratory room illumination was kept low to increase the contrast between the display and its immediate surroundings.

Subjects were then instructed to track the moving target and grab it as quickly as possible. For a given experimental session, subjects performed with a 15-sec rest period between trials. Each session consisted of 10 trials. Task completion time (time to grasp the target) was recorded after each trial. The experimental measurements are shown in Tables 1 and 2.

Table 1. *Mean Task Completion Time and STD When Using Monoscopic and Stereoscopic Visual Display, for Different Graphics Update Rates*

	28 fps	14 fps	7 fps	3 fps	2 fps	1 fps
Mono						
Mean	1.09	1.45	3.18	5.52	9.87	14.57
SD	0.34	0.42	1.48	2.61	4.35	5.97
Stereo						
Mean	0.99	1.11	1.58	2.26	4.36	6.96
SD	0.22	0.25	0.59	1.13	1.94	2.35

Table 2. *Mean Task Completion Time and SD When Using Monoscopic and Stereoscopic Visual Display*[a]

	28 fps	14 fps	7 fps	3 fps	2 fps	1 fps
Mono						
BT	1.26	1.80	4.72	8.49	14.77	21.70
ET	0.89	1.24	2.06	3.42	6.11	9.33
Stereo						
BT	1.12	1.32	2.18	3.48	6.49	9.38
ET	0.88	0.92	1.10	1.36	2.46	4.94

[a]BT, at the beginning of the training period; ET, at the end of the training period, for different graphics update rates.

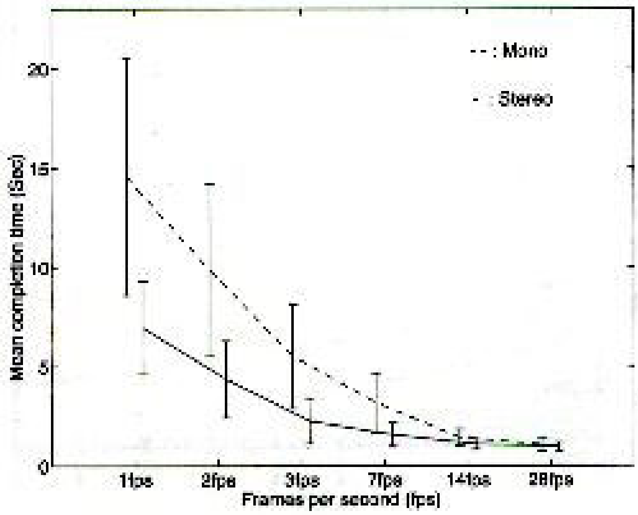

Figure 4. *Tracking task mean completion time for different graphics frame rates (monoscopic and stereoscopic displays).*

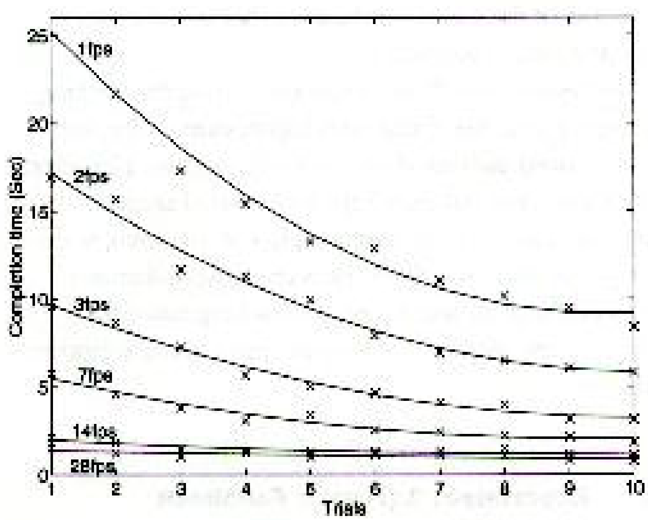

Figure 5. *Tracking task learning process for different graphics frame rates, (monoscopic display).*

5.2 Results

Results show that performance was stable from 28 down to 14 fps when using the monoscopic display, and from 28 down to 7 fps when using the stereoscopic display (Fig. 4). Performance is not statistically different between 28 and 14 fps for both viewing conditions. However, the stereo mode increased performance by 50% for update rates lower than 7 fps. It appears that stereo viewing provides more reliable results since the standard deviation (SD) was 50% smaller for update rates of 1–7 fps, and 30% smaller for frame rates of 7–28 fps.

Figures 5 and 6 plot the capture time for repeated trials. The user's learning process explains the decrease in the total capture time with task repetition. For monoscopic viewing (Fig. 5) and for low frame rates (1–7 fps) learning seems to stop after 10 trials (flat curve). There is less learning for low frame rates in stereoscopic viewing (Fig. 6). For higher frame rates the performance did not increase much for either monoscopic or stereoscopic viewing (curves are almost flat from the beginning of the trial sequence). These results confirm that high frame rates and stereoscopic viewing require less adaptation on

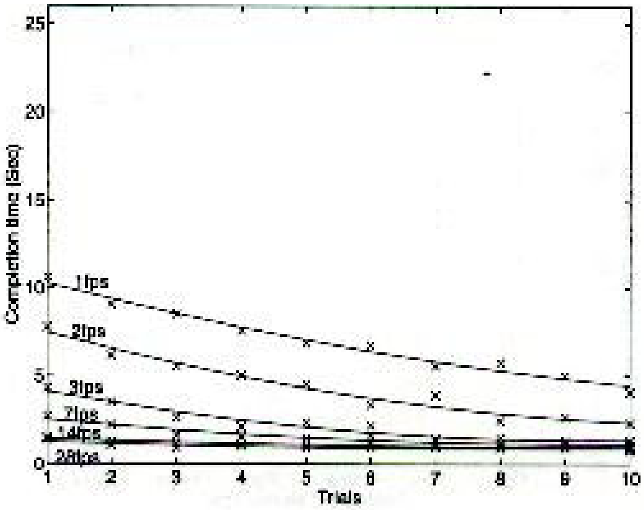

Figure 6. *Tracking task learning process for different graphics frame rates (stereoscopic display).*

the part of the user since these conditions are closer to natural human interaction.

Subjects used different strategies during the tracking/grasping task. For frame rates higher than 14 fps, subjects moved their hand continuously in space. However, for frame rates less than 7 fps they exerted saccadic movements (for both stereoscopic and monoscopic displays). At these low frame rates the subjects found it more difficult to accomplish the tracking task with a monoscopic display, resulting in longer average capture times.

6 Experiment 2 (Haptic Feedback Modalities)

This experiment investigates the effect of sensory substitution for grasping force regulation for a dextrous manipulation task. Six groups of 14 subjects each (42 males and 42 females) were used for the experiment. The first group (N) performed the task using only graphics feedback, while the second group (V) had graphics feedback of the virtual world and visual feedback for grasping forces (through the interface LED bar graphs). The third group (A) had graphics and auditory force feedback through headphones, while the fourth group (H) had graphics and haptic force feedback with the Rutgers Master. The fifth group (H-V) had graphics, haptic, and redundant visual force feedback, while the sixth group (H-A) had graphics, haptic, and redundant auditory force feedback. The frame rate (28 fps) and virtual world complexity (hand, sphere, and walls) were kept constant for all groups. The first 7 subjects of each group performed the task using a monoscopic graphics display, while the last 7 subjects performed the task using the stereoscopic display.

The force-feedback structure was integrated with the DataGlove[28] throughout the experiment. It was energized only for groups (H), (H-V) and (H-A). This allowed a true comparison between visual, auditory, and haptic force feedback, since friction and the master's weight were always present.

6.1 Experimental Protocol

Subjects were instructed to reach and pick up a red virtual ball (of 7 cm diameter) at location "A," then move it to "C" going through location "B," while keeping the vertical projection of the ball between the two lines. During the arm movement subjects had to apply a low, stable amount of deformation (10% of the sphere radius). The task had to be completed in less than 15 sec, however, subjects were told that the most important thing was to control the amount of deformation, not the completion time. Task completion time (time to move the ball from location A to location C) and mean deformation of the ball (mean of average amount of deformation at each contact location with the ball) were recorded after each trial. Subjects had no explicit information about the relationship between the number of LEDs "on" and percentage of deformation required (10%).

Subjects were seated so that the eye-to-monocular display distance was approximately 60 cm. Stereo glasses were worn by subjects throughout the experiment. The Rutgers Master integrated with a DataGlove[29] was fitted to their hand and calibrated. Subjects had one trial session to familiarize themselves with the limits of the virtual room and the sensory force feedback from the Rutgers Master (corresponding to 10% object deformation). Then, subjects performed the task in 10 trials with 15-sec rest periods in between. Perspective grid lines were superimposed on the virtual walls, and a

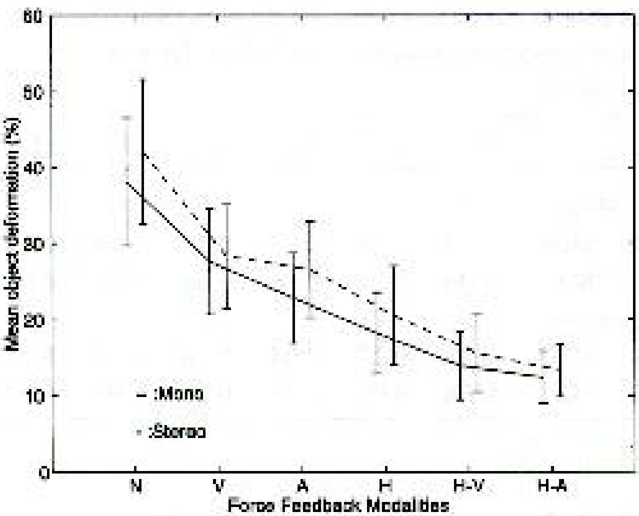

Figure 7. *Manipulation task mean error for different force feedback modalities: N, none; V, visual; A, auditory; H, haptic.*

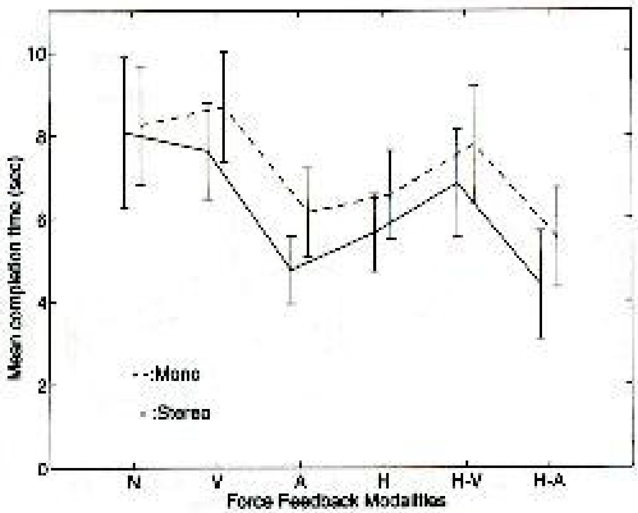

Figure 8. *Manipulation task mean completion time for different force feedback modalities (monoscopic and stereoscopic displays).*

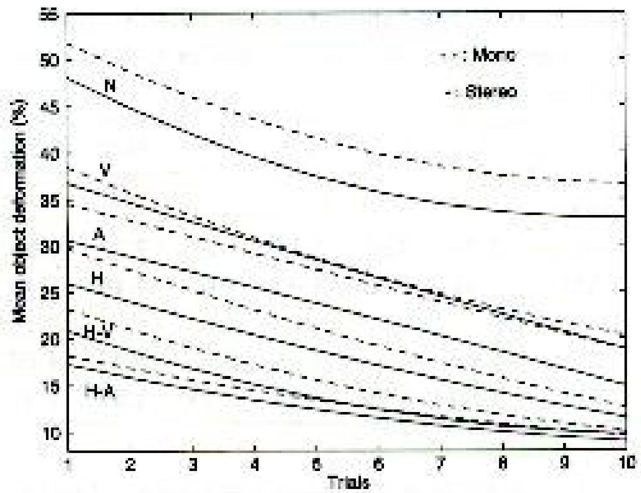

Figure 9. *Manipulation task learning process for different force feedback modalities (monoscopic and stereoscopic display).*

vertical projection of the hand was displayed throughout the experiment.

6.2 Results

Results show that in all cases, stereoscopic viewing did not improve performance significantly. This could be expected given the presence of strong monoscopic depth cues. When no force feedback cues were present subjects deformed the ball by about 42% while the required amount was 10%. This highlights the need for force feedback when interacting with virtual objects. We observed that when visual (V), auditory (A), and haptic (H) force feedback were present, performance (regulation of grasping forces) was increased by 33.3% (V), 35.7% (A), and 52.3% (H), respectively (Fig. 7). Redundant force feedback information (visual and auditory feedback) reduced ball deformation by 61.9% (H-V) and 69.0% (H-A), respectively, as compared to the open loop case (group N).

Auditory and haptic representation of feedback forces seems to decrease operator mental workload, as compared to the open-loop case, since task completion times (time to reach location "C") were reduced by about 30% (Fig. 8). Redundant visual presentation of forces increased the mean completion time again (group H-V). The (H-A) group had the shortest average completion time.

Figures 9 and 10 show the subject's learning process measured by the decrease in object deformation over repeated trials. These graphs show that the viewing condition (mono or stereo) was important only for group (N) and less important for visual, auditory, and haptic modalities. Redundant visual and auditory presentations led to the required deformation (10% of the radius) af-

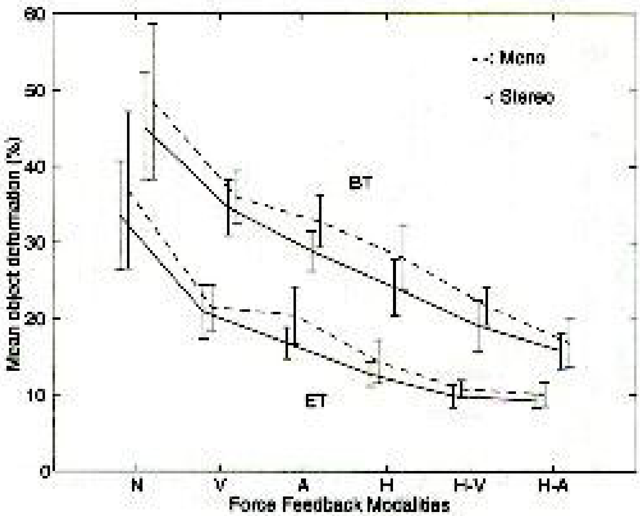

Figure 10. *Manipulation task performance at the beginning (BT) and at the end (ET) of the training session.*

Table 3. *Mean Task Completion Time and SD When Using Monoscopic and Stereoscopic Visual Display, for Different Force Feedback Modalities*

	N	V	A	H	H-V	H-A
Mono						
Mean	8.26	8.69	6.14	6.57	7.77	5.53
SD	1.42	1.33	1.06	1.07	1.40	1.20
Stereo						
Mean	8.08	7.63	4.75	5.65	6.84	4.39
SD	1.80	1.16	0.82	0.93	1.30	1.31

ter 10 trials. None of the subjects needed more than 15 sec to complete the task. Statistically, the standard deviation was smallest at the end of the trial sequence and for the groups (H-V) and (H-A). The largest variability (SD) was shown by subjects in group (N), with little benefit from task learning. Results are summarized in Tables 3, 4, and 5.

7 Discussion

The experiments presented above were aimed at understanding the influence of both visual and force feedback modalities in the performance of two typical tasks. The first was tracking and capturing of a moving virtual object (sphere), while the second was manipulation with minimal deformation of the same object. While these clearly represent only a small part of the manipulative tasks possible in virtual environments, tracking and object deformation are clearly high-frequency events that warrant careful study.

Table 4. *Mean Deformation and SD When Using Monoscopic and Stereoscopic Visual Display, for Different Force Feedback Modalities*

	N	V	A	H	H-V	H-A
Mono						
Mean	0.42	0.28	0.27	0.20	0.16	0.13
SD	0.09	0.07	0.06	0.06	0.05	0.03
Stereo						
Mean	0.38	0.27	0.23	0.18	0.14	0.12
SD	0.08	0.07	0.06	0.05	0.04	0.03

Table 5. *Mean Deformation and SD When Using Monoscopic and Stereoscopic Visual Display, for Different Force Feedback Modalities*[a]

	N	V	A	H	H-V	H-A
Mono						
BT						
Mean	0.48	0.36	0.32	0.28	0.22	0.17
SD	0.10	0.03	0.03	0.04	0.02	0.03
ET						
Mean	0.37	0.21	0.20	0.14	0.11	0.10
SD	0.10	0.03	0.04	0.03	0.01	0.01
Stereo						
BT						
Mean	0.45	0.35	0.29	0.24	0.19	0.16
SD	0.07	0.03	0.02	0.03	0.03	0.02
ET						
Mean	0.33	0.21	0.17	0.13	0.10	0.09
SD	0.07	0.03	0.02	0.02	0.02	0.01

[a]BT, at the beginning of the training period; ET, at the end of the training period.

The present study is important to the research community in at least two respects. First, unlike previous studies that used a small number of subjects, the tests reported here used over 160 subjects. Thus, while the results do not claim to be overly general, they do, however, have good statistical significance. The present study is also important since it is among the few to focus on dextrous manipulation with a portable force feedback master.

The first experiment showed that tracking performance (measured as the time-to-capture) degrades for low graphics refresh rates (fps). Stereo proved to be beneficial for low frame rates, where it reduced capture time by 50% compared to the monoscopic display. Our results are similar to those obtained by Ranadive (1979), who used a video display and an Argonne E-2 seven-degree-of-freedom servo manipulator. Massimino and Sheridan (1989) later compared manipulation capability for direct vision vs. video feedback in simple block-insertion tasks for different graphics frame rate. They observed that mean completion times dropped dramatically as the frame rate fell below 14 fps, similar to the results of our study in a partially immersive virtual environment.

The second experiment was aimed at dextrous manipulation of a deformable sphere with pseudo-force feedback, actual force feedback, and redundant force feedback. The manipulated object fitted in the palm and had no weight, thus the present results should not be extended to objects with weight and large volume. Data obtained with a Rutgers Master showed that haptic force feedback led to better performance in terms of object deformation [52.3% (H) versus 33.3% (V), and 35.7% (A)]. However, results concerning the auditory force feedback greatly depend on the task. The sphere was symmetrical, so a single audio channel sufficed, but this cannot be generalized to other tasks that require simultaneous control of different forces.

In the absence of haptic feedback, visual representation of feedback forces proved beneficial compared with graphics feedback alone. Similar results were obtained by Reger (1987), who studied the effect of visually displayed force feedback in delayed and nondelayed bilateral teleoperation. Ouh-Young et al. (1989) studied operator performance in a 6D docking task using haptic and visual force feedback. They observed that even though force feedback was more effective, the task could be reliably done with the visual force display alone.

The smallest error rates in the present study were obtained when subjects had haptic and redundant audio feedback. This is due to the nature of the task, which required careful deformation of the virtual object. The task was difficult, since the small desired object deformation Δx produced small feedback forces (due to Hooke's law). The static friction present in the Rutgers Master actuators could mask such small forces. Under these conditions sound proved to be useful in providing a cue for the start of the object deformation phase. Thus the group that had both haptic and redundant auditory feedback performed best. The role of sound feedback could have been played by tactile feedback actuators such as voice coils placed at the fingertip. Tactile feedback was, however, not available in the present study. The above results are applicable to other dextrous and nondextrous force feedback masters that have typical static friction on the order of 5% of their output force. All these devices will benefit from the integration of tactile feedback with force feedback during the simulation.

While redundant force feedback proved useful for group (H-A) in terms of object deformation and task completion time, it was not so for group (H-V). Here, poorly designed redundant information hampered rather than helped the task. The constant shifting of attention from the screen to the LED display and back resulted in increased task completion time (Fig. 8). This can be remedied by displaying the visual force cues (such as vectors) on screen, colocated with the virtual object being manipulated. Another approach is to color the finger phalanges and the palm of the virtual hand according to the magnitude of contact forces (Bergamasco, 1992).

Learning through task repetition clearly improved performance for both experiments. While stereo viewing had a big impact on learning in the first experiment, it had less impact for the second experiment. This is explicable because the second experiment relied primarily on haptic rather than graphics feedback. While the learning process ended after 10 trials for groups (H-A) and (H-V), it was continuing for groups (V), (A), and (H) involved in the second experiment. This means that the absence of redundant force feedback information lengthenced the learning process. Intersubject variability was largest for group (N), which had no force feedback cues, and smallest for groups (H-V) and (H-A). This variability was smallest at the end of the trial sequence, indicating subject adaptability through learning.

8 Conclusion

The present study investigated human performance for tracking/grasping and dextrous manipulation of a virtual object. The study was aimed at determining the influence of the visual and haptic feedback channels on task performance (measured by completion time and percent object deformation). Tests reported here used over 160 subjects (both male and female). Thus, while the results do not claim to be overly general, they do, however, have good statistical significance. The present study is also important since it is among the few to focus on dextrous manipulation with a portable force feedback master.

In the first experiment both the viewing conditions (stereo and mono display) and graphics refresh rates were varied. Object capture time was found not to be statistically different between 28 and 14 fps (for both viewing conditions). However, stereo mode increased performance by 50% for update rates lower than 7 fps.

Results of the second experiment showed that direct haptic feedback was superior to pseudo-feedback through visual or auditory cues. However, the best results (lowest object deformation and task completion time) were obtained when both haptic and redundant auditory feedback were present. Here auditory feedback serves as a substitute for tactile feedback in marking the moment of initial contact between the virtual hand and ball. These results are extendible to other force feedback interfaces where static friction may mask small feedback forces during initial object deformation. These interfaces would benefit from integrating tactile and force feedback modalities. It is also necessary to use low-friction, low inertia actuators whenever possible. This was the approach taken in the design of the Rutgers Master II, which uses noncontact sensing and custom force feedback actuators integrated in one structure. Future tests will use this more advanced device to repeat the experiments described here.

Acknowledgments

Work reported here was supported by grants from the Centre de Robotique Intégrée d'Ile de France (CRIIF), the Center for Computer Aids for Industrial Productivity (CAIP) at Rutgers—The State University of New Jersey—and the New Jersey Commission on Science and Technology. We would like to thank all the volunteers who participated in these experiments.

References

Airey, J., Rohlf, J., & Frederick, P. (1990). Towards image realism with interactive update rates in complex virtual building environments. *Symposium on Interactive 3D Graphics*, 41–50.

Atkeson, C. G., & Hollerbach, J. (1985). Kinematic features of unrestrained vertical arm movement. *Journal of Neuroscience, 5*, 2318–2230.

Bach-y-Rita, P., Webster, J. G., Tompkins, W. J., & Crabb, T. (1987). Sensory substitution for space gloves and for space robots. *Proceedings of Workshop on Space Telerobotics, 2*, 51–57.

Bejczy, A., & Paine, O. (1982). Manual control of manipulator force and torques using graphics display. *Proceedings of IEEE Conference on Systems, Man, and Cybernetics*, October.

Bergamasco, M. (1992). The GLAD-IN-ART project. *Proceedings of Imagina '92 Conference*, France.

Bouzit, M., Coiffet, P., & Burdea, G. (1993). The LRP dextrous hand master. *VR Systems Fall '93*, New York City, October.

Brooks, F., Ouh-Young, M., Batter, J., & Jerome, P. (1990). Project GROPE—haptic displays for scientific visualization. *Computer Graphics, 24*(4), 177–185.

Burdea, G., & Zhuang, J. (1992). Actuator system for providing force feedback to a dextrous Master Glove. U.S. Patent 5,143,505, September 1, 1992.

Burdea, G., Langrana, N., Roskos, E., Silver, D., & Zhuang, J. (1992). A portable dextrous master with force feedback. *Presence: Teleoperators and Virtual Environments, 1*, 18–28.

Burdea, G., & Langrana, N. (1993). Virtual force feedback—

lessons, challenges and future applications. *Journal of Robotics and Mechatronics, Japan, 5*(2), 178–182.

Burdea, G., Roskos, E., Gomez, D., & Langrana, N. (1993). Distributed virtual force feedback. *IEEE Workshop on "Force Display in Virtual Environments and Its Application to Robotic Teleoperation,* May 2, Atlanta, GA.

Burdea, G., & Coiffet, P. (1994). *Virtual reality technology*. New York: John Wiley.

Corker, K. M., Bejczy, A., & Rappaport, B. (1985). Force/torque display for space teleoperation control experiments and evaluation. *Proceedings of the 21st Annual Conference on Manual Control,* Los Angeles, CA, October 15–17, 465–477.

Cutkosky, M., & Howe, R. (1990). Human grasp choice and robotic grasp analysis. In S. Venkataraman and T. Iberall (Eds.), *Dextrous robot hands* (pp. 5–31). Berlin: Springer-Verlag.

Cutt, P. (1993). Tactile displays: Adding the sense of touch to virtual environments. *Proceedings of Virtual Reality Systems '93 Conference,* New York City, 120–122.

Ellis, R., Ismaeil, O., & Lipsett, M. (1993). Design and evaluation of a high-performance prototype planar haptic interface. *Proceedings of ASME WAM, DSC, 49,* 55–64.

Ferrel, W. R. (1966). Delayed force feedback. *Human Factors, October,* 449–455.

Georgopoulos, A. P., Kalaska, J., & Massey, J. (1981). Spatial trajectories and reaction times of aimed movement: Effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology, 46,* 725–743.

Gomez, D., Burdea, G., & Langrana, N. (1994). The second-generation Rutgers Master (RM-II). *The Third International Conference on Automation Technology,* Taipei, Taiwan, July 7–10.

Hashimoto, H., Boss, M., Kuni, Y., & Harashima, F. (1994). Intelligent cooperative manipulation system using dynamic force simulator. *Proceedings of IEEE International Conference on Robotics and Automation,* May, 2598–2603.

Hewlett-Packard (1991). *Starbase display list programmers manual,* 1st Edition.

Iwata, H. (1993). Pen-based haptic virtual environment. *Proceedings of IEEE Virtual Reality Annual International Symposium (VRAIS '93),* Seattle, WA, 287–292.

Jacobsen, S., Inverson, E., Davis, C., Potter, D., & McLain, T. (1989). Design of a multiple degree of freedom force reflective hand master slave with a high mobility wrist. *Third topical meeting on robotics and remote systems,* Charleston, SC, March.

Jau, B. M. (1993). Man-equivalent telepresence through four finger human-like hand system. *Proceedings IEEE International Conference on Robotics and Automation,* Nice, France, 834–848.

Kim, W. S., Ellis, S., Tyler, M., & Stark, L. (1985). Visual enhancement for telerobotics. *Proceeding of 1985 International Conference on Systems, Man and Cybernetics,* Tucson, AZ.

Kim, W. S., Ellis, S., Hannaford, B., Tyler, M., & Stark, L. (1987). A quantitative evaluation of perspective and stereoscopic display in three-axis manual tracking tasks. *IEEE Transportation Systems, Man and Cybernetics, SMC-17*(1), 16–72.

Lipton, L. (1991). *The CrystalEyes Handbook.* San Raphael, CA: StereoGraphics Corporation.

Massie, T., & Salisbury, K. (1994). The PHANToM haptic interface: A device for probing virtual objects. *ASME Winter Annual Meeting, DSC, 55*(1), 295–300.

Massimino, M., & Sheridan, T. B. (1989). Variable force and visual feedback effects on teleoperator man/machine performance. *Proceedings of NASA Conference on Space Telerobotics,* Pasadena, CA (1), 89–98.

Massimino, M., & Sheridan, T. B. (1994). Sensory substitution for force feedback in Teleoperation. *Presence: Teleoperators and Virtual Environments, 2*(4), 344–352.

Millman, P., Stanley, M., & Colgate, E. (1993). Design of a high performance haptic interface to virtual environments. *IEEE Virtual Reality Annual International Symposium (VRAIS),* Seattle, WA, 216–222.

Minsky, M., Ouh-Young, M., Steele, O., Brooks, F. Jr., & Behensky, M. (1990). Feeling and seeing: Issues in force display. *Computer Graphics, 24*(2), 235–243.

National Research Council. (1995). Haptic Interfaces. In N. Durlach and A. Mavor (Eds.), *Virtual reality. Scientific and technological challenges* (pp. 161–187). Washington, DC: National Academy Press.

Ouh-Young, M., Pique, M., Hughes, J., Srinivasan, N., & Brooks, F. Jr., (1989). Using a manipulator for force display in molecular docking. *Proceedings of the IEEE Robotics and Automation Conference,* Philadelphia, PA, April, 1824–1829.

Patrick, N. (1990). *Design, construction, and testing of a fingertip tactile display for interaction with virtual and remote environments.* Masters Thesis, Department of Mechanical Engineering, MIT.

Polhemus Navigation Science Division. (1987). *Space Isotrack user's manual.* Colchester, VT.

Ranadive, V. (1979). *Video resolution, frame-rate, and grey scale* *tradeoffs under limited bandwidth for undersea teleoperation,* SM Thesis, MIT.

Reger, J. (1987). Visually displayed force feedback in delayed and non-delayed bilateral teleoperation. *SID International Symposium—Digest of Technical Papers,* New Orleans, LA.

Richard, P., Burdea, G., & Coiffet, P. (1993). Human performance in tasks involving virtual objects with force feedback. *Proceedings of Interface to Real and Virtual Worlds Conference,* Montpellier, France, 229–238.

Schmult, B. (1990). Data analysis with a force-reflecting joystick. *Internal report.* Holmdel, NJ: AT&T Bell Laboratories.

Shreeve, J. (1993). Touching the phantom. *Discover,* June, 35–54.

Stone, J. R. (1992). Virtual reality: A tutorial. *Proceedings of MICAD 1992,* Paris, France, 2–16.

VPL Research Inc. (1987). *DataGlove model 2 operating manual,* Redwood City, CA.

Wloka, M. (1995). Lag in multiprocessor virtual reality. *Presence: Teleoperators and Virtual Environments, 4*(1), 50–63.