Some of What One Needs to Know About Using Head-Mounted Displays to Improve Teleoperator Performance

Andrew Liu, Gregory Tharp, Lloyd French, Stephen Lai, and Lawrence Stark, *Fellow, IEEE*

Abstract— **The head-mounted display (HMD) system developed in our lab would be a useful teleoperator system display if it increased operator performance of the desired task. It could, however, degrade performance because of display update rate limitations and communication delays. Display update rates are slowed by communication bandwidth and/or computational power limitations. Communication delays occur because command and feedback signals must travel large distances. 3-D tracking and pick-and-place tasks were simulated to characterize performance levels for a range of update rates and delays. How performance changes as the display is altered to maintain the update rate is also examined. The results of our update experiments indicated that performance levels degrade only slightly until the update rate is below 1 or 2 Hz. The attempts to maintain update rate by eliminating display features such as stereo and occlusion cues showed that the presence of either cue alone was sufficient to maintain performance levels in our simple, short-term tasks. Delays also degraded performance in both manipulation tasks. These HMD performance problems were not lessened by the presentation of a predictive feedback display. Instead, the delay in the head movement tracking loop must be eliminated.**

I. Introduction

HEAD-MOUNTED Display (HMD) systems have become increasingly popular as an interface for many applications including simulation, molecular modelling, entertainment, and telerobotics [1]–[4]. Stereoscopic display, head movement tracking, and powerful graphics computers have made the HMD a very realistic and intuitive man–machine interface. The correspondence to how humans naturally acquire visual information makes HMDs easy to learn how to use and adds to the realism of the display. The realism may even be sufficiently compelling to elicit " telepresence," which is the feeling that the wearer is physically present in the remote or virtual environment. For applications such as space telerobotics, where the remote environment can change rapidly and unpredictably, a good interface allows operators to adapt to the new conditions and maintain their performance of a given task.

Very little research has objectively assessed the effectiveness of HMD systems for space teleoperation. Most studies have examined the effects of design parameters such as interocular distance or field of view [5], [6]. Using HMDs in telerobotics applications presents other constraints that could adversely affect task performance. The flow of information during telerobotic manipulation can be visualized with a block diagram (see Fig. 1). At the local site, the human operator gets information from the display and controls the head angle and joysticks. The head angle and joystick signals which control camera angles and manipulator position, respectively, are transmitted to the remote site. The joystick signal can also be fed back through a local model and implemented on the display as a predictor of manipulator position. Visual information about the target is acquired by the cameras and shown on the display.

A. Update Rate

The update rate represents the frequency at which the image of the remote site is captured and subsequently displayed to the operator on the HMD. It can be affected by a number of factors, such as the bandwidth of the communication channels. As the bandwidth of the video signal is increased, the update rate for a fixed bandwidth communication channel will decrease. Ranadive and Sheridan [7] studied the tradeoff between video frame rate, resolution and gray scale and found that subjects could perform reasonably well with low frame-rates. Cole and Kishimoto [8] examined the effect of bandwidth-limited TV displays and other display variables on operator performance for underwater applications. They studied methods for optimizing the viewing system to minimize performance degradation. Both studies used fixed monoscopic displays.

A telerobotic system based on computer generated virtual environments is an alternative to video-based visual displays [9]. In this case, a smaller bandwidth signal consisting of information such as robot position or joint angles is sent to the local site from which a virtual representation of the remote site could be constructed. The update rate of this system represents the frequency at which modelled and measured parameters are passed through the rendering process and subsequently displayed. The computational and rendering speeds of the

---

Manuscript received February 3, 1992; revised April 23, 1993.

A. Liu was with the Bioengineering Graduate Group & Telerobotics Unit, 483 Minor Hall, University of California, Berkeley, CA, 94720 and is now with the Nissan Research & Development, Inc., Cambridge Basic Research, Four Cambridge Center, Cambridge, MA 02142.

G. Tharp is with the Robotic Systems and Advanced Computer Technology Section, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109–8099.

L. French and S. Lai are with the Telerobotics Unit, 483 Minor Hall, University of California, Berkeley, CA, 94720.

L. Stark is with the College of Engineering and the School of Optometry, 481 Minor Hall, University of California, Berkeley, CA 94720.

IEEE Log Number 9212014.

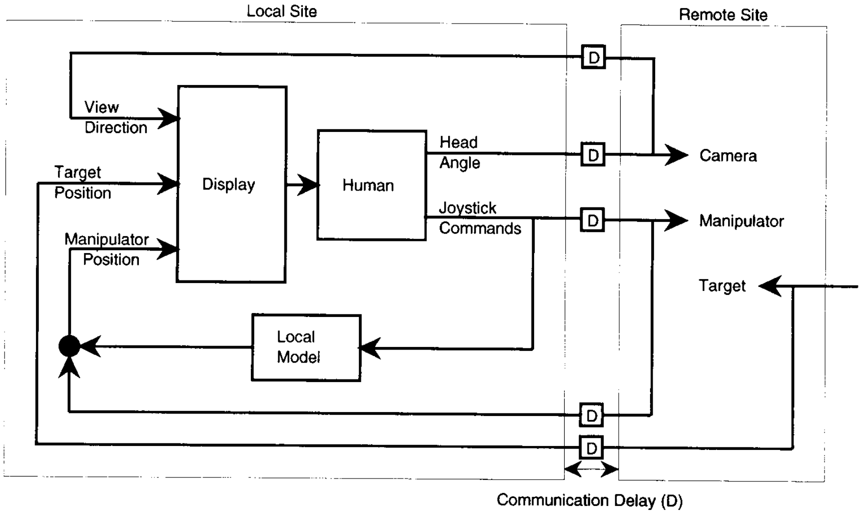


Fig. 1. Block diagram of simulated telemanipulation. The communication delay (D) represents the time to transmit the command and feedback signals between the remote and local sites.

graphics systems are the major determinants of the update rate. To reduce the computational load, scenes might be rendered with lines rather than polygons or bioptic images might be used instead of stereo image pairs. Many studies have shown that stereo is a valuable depth cue for manipulation [10]–[12]. However, other cues such as occlusion or motion parallax may still provide enough depth information to complete the task. Previous studies in our laboratory have indicated that performance may not be seriously affected for update rates above 1 or 2 Hz [13]. Therefore, it might be possible to maintain task performance by limiting the number of depth cues provided in the display that would maintain a suitable update rate.

The update rate has been demonstrated as an important parameter in the definition of lag times for virtual environment systems [14]. These system lags affect the apparent behavior of input devices such as position and orientation trackers. Slow update rates and long lag times uncouple the natural correspondence between head movements and image motions and can degrade task performance.

B. Delays

While computational operations introduce delays into the display of visual information, the source of most delay is nearly always communication [15], [16]. Communication delays occur both in sending signals to the remote site as well as receiving signals from afar. As a consequence, the operators see the position of the target after a one way communication delay (D), but must wait for a round trip delay (2D) before seeing the results of their manipulator and head angle commands (see Fig. 1).

The effect of delay on telemanipulation has been widely studied [15], [17]–[19]. The general conclusion is that increasing time delays has a detrimental effect on task performance. As in the case of slow update rates, the delay also affects the correspondence between head motions and the displayed scene in the HMD. The resulting disorientation may compound the problems already shown in the previous studies. If operators using a HMD can adapt only to a limited amount of delay, it might constrain the effective distance over which a HMD would be useful for telemanipulation. If, however, the HMD facilitates an operator's adaptation to delays, that would be a strong argument for its use.

Time delays also have a strong effect on tracking targets with head movements [20]. For low values of time delay, head tracking performance is fairly good; however, for delays greater than 1 s, performance degrades extremely quickly. The performance is also strongly dependent upon the frequency which the subject must move their head. As the frequency increases, the tracking performance gets worse. In these tracking experiments, the targets moved in a sinusoidal path which was predictable. Tracking performance for unpredictable pseudorandom motions should be worse for a given value of delay. The effects of time delays on more complicated manipulation tasks should occur at shorter delays than seen in these experiments.

Preview or predictor displays have been tested as a possible means for overcoming delay. Experiments with these types of displays have shown that a predictive display is helpful for manipulation tasks [21], [22] but only marginally for tracking tasks [16]. Predictive displays might be less helpful when a HMD is used under time-delayed conditions, because the head motions are also delayed. As a result, the operator might not be able to track the targets sufficiently to use the information provided by the predictor.

C. Research Objectives

We are interested in the utility of HMDs for telemanipulation, and our experiments were designed to measure operator performance under a range of display timing con-

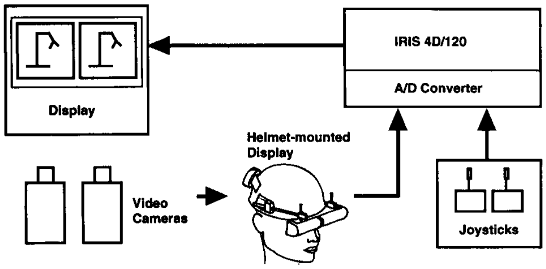


Fig. 2. Schematic diagram of the experimental apparatus. The image pair is captured by the video cameras and sent to the HMD. The computer collects data from the joysticks and head tracking coil in the helmet.

ditions. Specifically, we examined how display update rate and delay affected the performance of tracking and pick-and-place tasks while using a HMD. We also studied how different representations of a virtual environment scene would affect telemanipulation performance. We implemented a version of the predictive display previously described and tested whether it aided task performance with the HMD.

II. METHODS

A. Equipment

The HMD and head tracking system used in these experiments were designed and built in our laboratory [5], [23]. The simulation system consisted of the HMD, a high-speed graphics computer (Silicon Graphics 4D/120), and analog input devices (see Fig. 2). Two 1-in CRT viewfinders (Sony VF-208) provided a monochrome, 22–degree field of view to each eye with sufficient resolution and contrast for stereopsis. The viewfinders, which were mounted on a helmet worn by the subject, were driven by two video cameras (Sanyo Vision8) positioned in front of the computer monitor. The computer rendered a stereo image pair of the modelled world in two sections of its screen, one for each camera/eye pair. A 12-bit A/D board (Metrabyte VMECAI-16) mounted in the computer's VME bus read signals from a pair of displacement joysticks, a foot-activated switch, and the head tracker that represented the operator's control of the simulated manipulator and the operator head orientation, respectively. The head tracker continuously measured the pan and tilt of the helmet as fast as 100 Hz. However, the synchronization of the drawing routine with the data acquisition limited the data update to the video rate of 30 Hz. This still provided a signal for controlling the display direction with negligible delay.

B. Tasks

We used two tasks that were relevant for telemanipulation: 3-D tracking and pick-and-place. Each task required different control techniques and had different performance measures. The basic difference lay in the fact that the dynamic response of the subject was forced at different levels during the two tasks [15]. Previous studies have shown that similar qualitative effects on performance occur under various display conditions [11], [12]. This suggests that results from these tasks may be generalizable. Both tasks occurred in a virtual remote space simulated by the computer. The dimensions of the virtual space for the tracking task are shown in Fig. 3. The virtual space for the pick-and-place task was identical except that the checkerboard wall was not present. The position limits represented the volume within which the operator could move the manipulator.

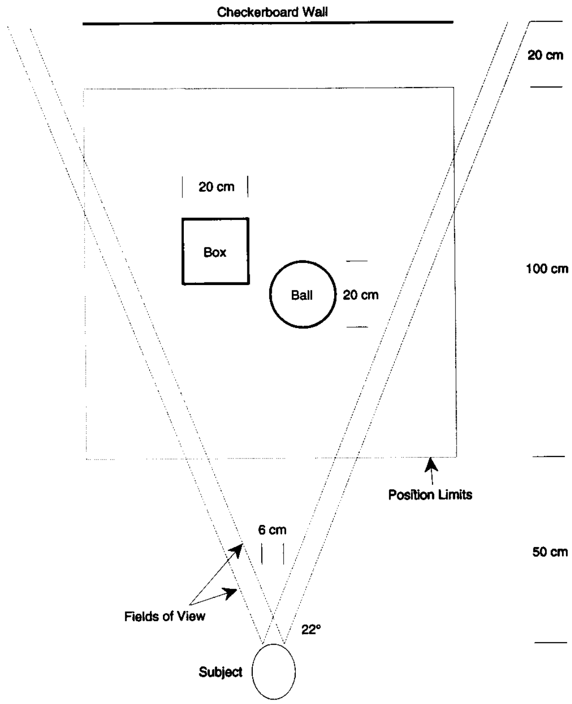
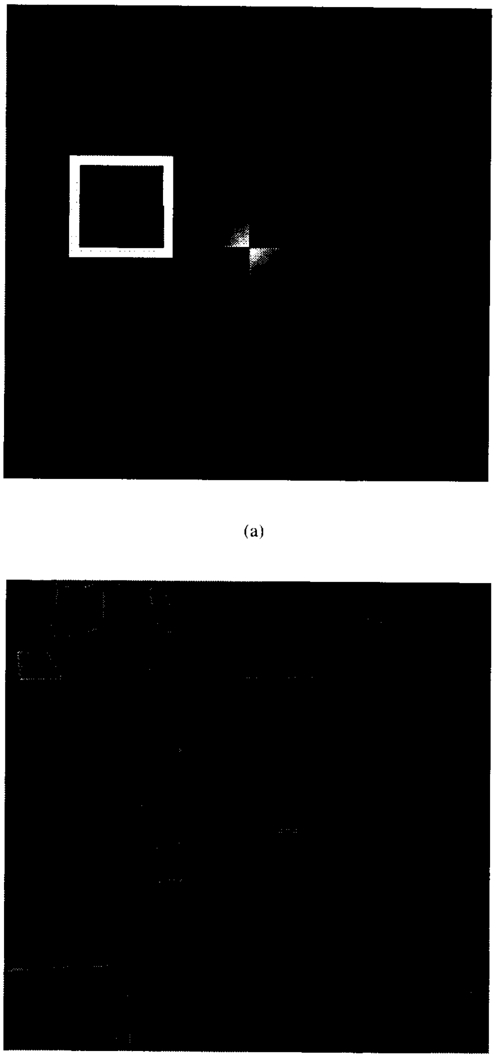


Fig. 3. Layout of the virtual workspace for 3-D tracking. The same position limits are used for the pick-and-place task.

3-D Tracking: In the 3-D tracking task (see Fig. 4(a)), the goal of the task was to keep a ball inside of a box. The box moved in an unpredictable 3-D trajectory within the virtual workspace for approximately 64 s. Each target trajectory component was a sum of three nonharmonically related sinusoids with frequencies ranging from 0.11 Hz to 0.36 Hz. The magnitudes were chosen so that the ratio from highest to lowest frequency components was 10:7:4. The subjects controlled the position of the ball in three dimensions with a pair of joysticks. Because of the limited field of view and the extent of the box trajectory, subjects must turn their heads to maintain the box in view on the display.

Task performance was measured by the normalized RMS error between a one-dimensional trajectory and the response trajectory along the corresponding axis. This provided three measures that allowed us to differentiate between a subject's performance in the display plane (x- and y-axes) and in depth (z-axis) [24]. For example, the normalized RMS error between the x-axis target trajectory $p_x(i)$ and the response trajectory


(a)

(b)

Fig. 4. Monoscopic views of the simulated telerobotic tasks. (a) In the 3-D tracking task, the operator must keep the ball inside the box. The box moves in an unpredictable 3-D trajectory; the position of the ball is controlled by the operator. (b) In the pick-and-place task, the operator puts a ball inside its correspondingly sized box. The operator controls the position of the gripper with joysticks.

$y_x(i)$ for $i = 0, \cdots, N-1$ was defined as

$$E_x = \frac{1}{R}\sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(y_x(i) - p_x(i))^2} \qquad (1)$$

where N was the total number of data points, R was the RMS error between the target trajectory $p_x(i)$ and its mean value, $\bar{p}_x$, which was given as

$$R = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(p_x(i) - \bar{p}_x)^2}. \qquad (2)$$

A RMS error of zero signified perfect tracking, while an error of one was equivalent to leaving the cursor at the mean position $(\bar{p}_x, \bar{p}_y, \bar{p}_z)$ of the target trajectory for the duration of the trial. A random input produced an error of approximately 1.4 and perfect tracking that legs by 180 degrees produced an error of 2. We chose the normalized RMS error as the performance measure because it represented how well the subjects were achieving their task goals. Frequency response measures, such as those used by Tharp *et al.* [13], are more suited to investigate the biological limitations of human performance.

Pick-and-Place: The workspace for the pick-and-place task had the same dimensions as the tracking task. Four balls with radii of 10 cm, 8 cm, 6 cm and 4 cm appeared at random locations in a virtual workspace. Four boxes, with similar sizes, were randomly placed on the floor of the workspace. Their positions were computationally checked prior to display to ensure that boxes were not inside one another. The subjects picked up each ball with the simulated manipulator and placed or dropped it inside the box of corresponding size. A single trial for each subject consisted of five instances of the scene described above for a total of 20 balls per trial. We measured performance by recording the time required to grasp one ball and put it in its proper box. The subjects could pick up the balls in any order.

Only the simplest interactions between the objects in the display were modelled. To close the manipulator gripper and grab a ball, the center of the ball must be located between the fingers of the gripper. To open the gripper and place the ball in the box, the center of the ball must be above the appropriate box. When the gripper was opened, the ball fell into the box as though a magnet pulled the ball into the box, similarly to the "magnetic latch" used in previous experiments [11], [25]. If the ball was not properly aligned, the gripper would not open and the ball would not be released. The manipulator gripper is opened and closed with the foot pedal switch. No other properties, such as collision detection, dynamics, or force feedback, were implemented.

C. Experimental Protocol

We repeated the 3D tracking and pick-and-place tasks for various update rates, display conditions, and delay. Five graduate students were used as subjects. Each subject was trained at both tasks under the control conditions (30 Hz update rate, no delay, polygon rendered, and stereoscopic images) until their performance reached a steady level. All subjects repeated the experiments under identical conditions. The individual experimental protocols are described in more detail in the following sections.

Update Rate Experiments: For the tracking experiments, the subjects performed each task for ten different update rates (30, 10, 5, 3, 2, 1.5, 1, 0.5, 0.25, 0.1 Hz). The values were presented in a constrained pseudorandom sequence; each update rate was repeated three times. Different target trajectories were used for each repetition of a display condition. Two subjects were considered experienced and three others inexperienced. For the pick-and-place task, three experienced subjects completed the task at seven different update rates (30 , 10, 3, 2, 1.5, 1, 0.25 Hz). Each subject completed one trial per update rate. The update rates were presented in a constrained pseudorandom sequence. Subjects were able to complete all seven trials without a break.

Display Condition Experiments: To study the effect of display conditions on performance, we manipulated two major depth cues, disparity and occlusion, in the display. The update rate was fixed at 30 Hz so that it would not limit performance. One or both of these cues can be eliminated from the display if the update rate is too slow for efficient manipulation. All four possible display conditions were used. The simplest condition consisted of objects rendered as line drawings and presented bioptically. Two conditions included either stereo presentation or polygon rendered objects, and the final condition included both stereo presentation and polygon rendering.

For the tracking task, each subject completed 40 trials with 10 repetitions of each display condition. Five subjects of varying experience were used in these experiments. Subjects removed the HMD and took a 5- to 10-min break after 20 trials that took approximately 30 min. This break prevented the subjects from becoming fatigued either by the accommodation-vergence conflict inherent in stereo displays or from boredom with the simple task.

The pick-and-place task was shorter in duration. Three experienced subjects each completed one trial for each display condition for a total of four trials per subject. The subjects were able to complete all four trials in approximately 20 min without having to take a rest break.

Delay Experiments: Three experiments with communication delay were performed. The first two experiments used the 3-D tracking task and the last experiment used the pick-and-place task.

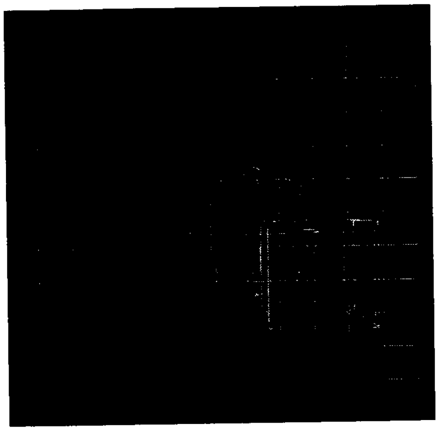
In the first experiment, one way delay was varied from 0 ms to 900 ms in 100 ms increments. Each delay condition was seen three times for a total of 30 trials. The three subjects completed the trials in two sittings, each lasting about 20 min. In the second experiment, we added a wireframe overlay for the ball which showed the instantaneous position of the ball as determined by the joystick position (see Fig. 6). This type of display has been shown to be helpful for other types of tasks. Only two of the three subjects from the first experiment also completed the second experiment. The sequence of delays used in the first experiment was repeated in the second experiment.

In the third experiment, three subjects completed the pick-and-place task with varying amounts of delay. One way delay was varied between 0 ms and 600 ms in 100 ms increments. Each delay condition was presented once, for a total of seven trials. The protocol was modified so that each trial consisted of three scenes for a total of 12 balls per trial.

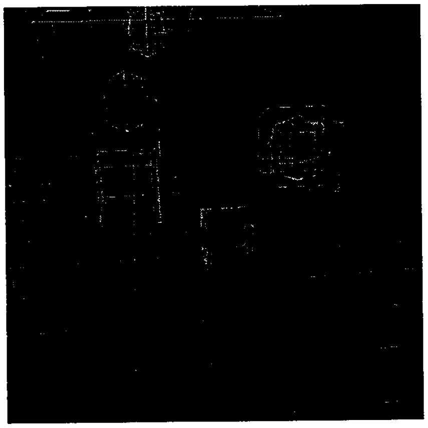
III. Results

A. Update Rate

The performance of the tracking task for two experienced (see Fig. 7(a)) and three inexperienced (see Fig. 7(b)) subjects decreased as update rate decreased. The RMS errors are shown for each axis of motion separately; z-axis motion represents motion in depth. The experienced subjects had extensive practice with both the equipment and the task, while the inexperienced subjects only had a limited training session before the experiment. The experienced subjects easily maintained performance levels down to update rates of 10 Hz. They did not show a marked performance drop until the update rate went below 2 Hz after which RMS error increased rapidly. Inexperienced subjects showed a performance decrease almost immediately with slowed update rate, but there was some indication that they had a performance plateau at rates faster than 10 Hz. In both cases, the RMS error for very low update frequencies leveled off at approximately 1.4.



(a)



(b)

Fig. 5. Monoscopic views of the wireframe rendered virtual environment. (a) 3-D tracking. (b) pick-and-place.

Three experienced subjects exhibit almost identical results for the pick-and-place task when compared to the tracking results (see Fig. 8). The completion times increased slightly when the update rate fell below 10 Hz, but showed a dramatic increase below 1.5 Hz. Also, the standard error of the means increased an order of magnitude when the update rate was 0.25 Hz.

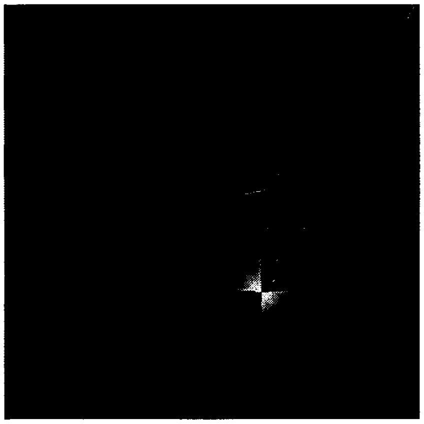
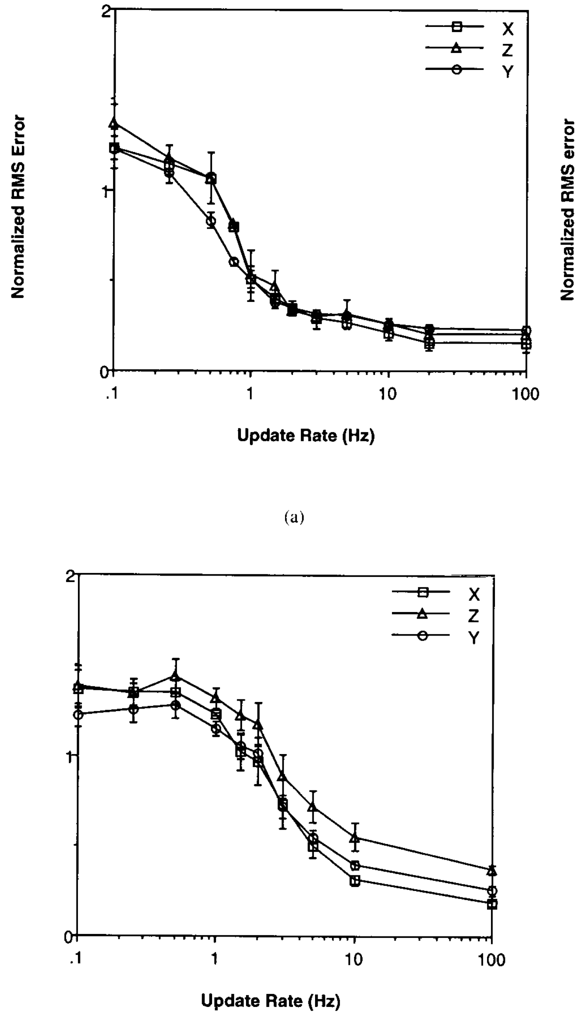


Fig. 6. Monoscopic view of the tracking task with a predictive overlay. The overlay indicates the position of the ball specified by the current joystick positions. The ball will move to the position of the overlay after the delay time (D).



(a)

(b)

Fig. 7. Effect of update rate on tracking performance. (a) The two experienced subjects can maintain performance down to a 2-Hz update rate. (b) The three inexperienced subjects can maintain performance only to 10 Hz.

B. Display Conditions

The average performance for the two tasks is plotted as a function of the depth cues present in the scene (see Fig. 9). For the tracking task (see Fig. 9(a)), each data point is the average of 50 trials (ten trials per subject). The error bars represent the standard error of the mean. The performance is measured as the normalized RMS error described above. The box, circle and triangle symbols represent the RMS errors along the x-, y- and z-axes, respectively.

The average RMS errors measured along the x- and y-axes are relatively constant for all four display conditions, with a slightly higher error when neither stereo nor occlusion are present. As expected, the average RMS error along the z-axis (depth) is the largest for all display conditions. A higher stereoacuity threshold, which is used for detecting motion in depth, in comparison to the one arc-minute acuity threshold, which is used for detection in the display plane, may account for the difference in the average error. The lowest average RMS error occurred when both cues were present; the highest RMS error occurred when neither cue was present. When only occlusion is present, the error is only slightly larger than the two conditions when disparity is present. This suggests that occlusion is almost as effective as a cue for the depth judgements in this type of task. The z-axis error when neither cue is present is approximately two times larger than the RMS error for other conditions.

The results of the pick-and-place task show the same effects as the z-axis error of the tracking task (see Fig. 9, right). Each point is the average completion time of three subjects. Mean completion time was largest when neither cue was present and almost constant for the other three conditions. The standard error for this display condition is also much higher than the other three conditions. It may be more difficult for the operator to make the relative depth judgements for certain positions. The completion time when occlusion is present is slightly longer than when stereo is present, providing further evidence that occlusion can be a very useful cue for making depth judgements. When both cues are present, the completion time is fastest.

C. Delay

As delay increases there is a decrease in performance for the three experienced subjects (see Fig. 10(a)) The data show an approximate linear increase in error with increased delay up to a level that indicates the control is no longer correlated with the target. Once the delay reached 400 ms, the error was approximately 1, a level where the subjects could do just as well by leaving the ball at the center of the telerobotic workspace! At 800 ms delay, the RMS error appears to plateau at a value around 1.5. At this point, the operator is performing as well as if they were randomly moving the ball.

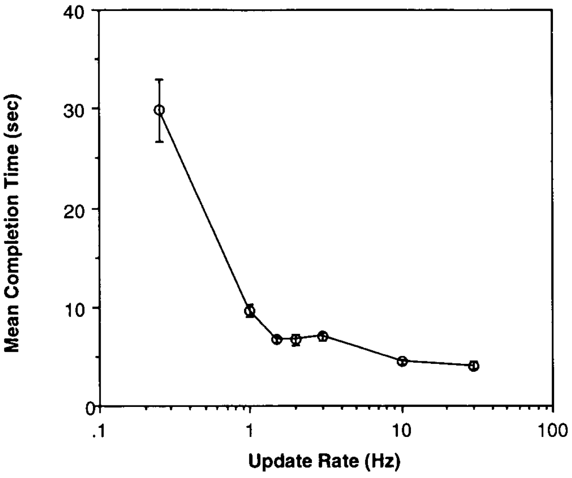


Fig. 8. Effect of update rate on pick-and-place performance. The data represent the average of three experienced subjects. Performance is maintained down to 1.5 Hz. Error bars are the standard error of the mean.

The condition with the predictive overlay shows a steeper rise, reaching a z-axis RMS error of 1 at 300 ms delay (see Fig. 10, right). The RMS error begins to level at 600 ms delay at a value of approximately 1.5. This suggests that the overlay had a negative influence on performance. If the subjects attempted to track the box using the overlay, the ball would follow at a distance proportional to the delay time and the velocity of the box. Thus, the RMS error would increase as the delay time increased. By comparison, the subjects might be able to track the box slightly better by predicting its trajectory.

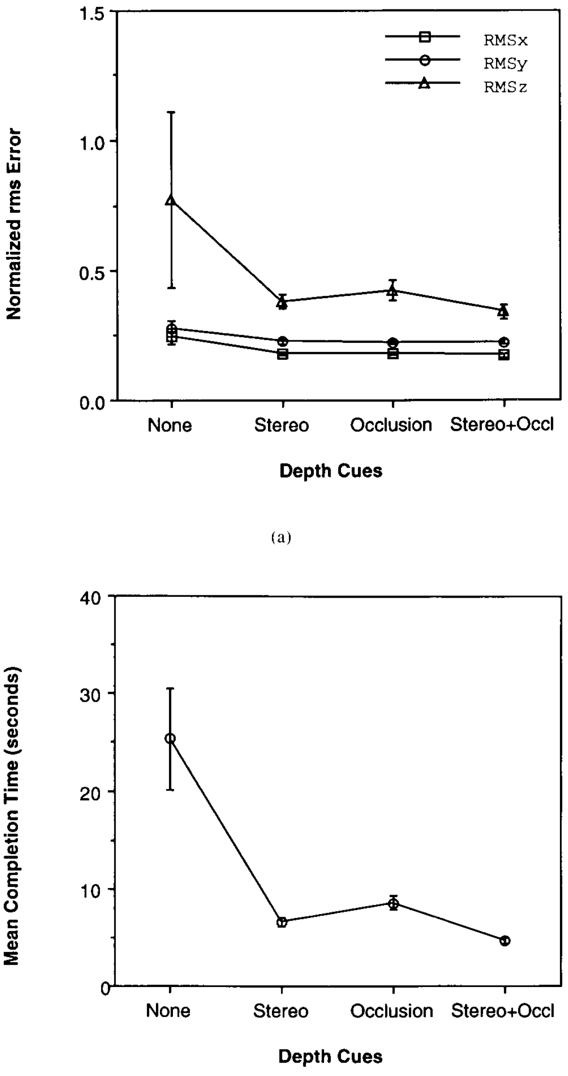
The results of the various delays on pick-and-place performance are very similar to the changes in tracking performance (see Fig. 11). The delays shows an immediate and proportional effect on the completion time. After a 500 ms delay, the average completion time is about three times longer than the completion time with no delay. There is no evidence of any plateau, nor should it be expected.

IV. DISCUSSION

A. Update Rate

Our main result is establishing a necessary minimum updated rate of 2 Hz for experienced subjects and 10 Hz for inexperienced subjects. The 2 Hz limit for experienced subjects is applicable to both tasks. This limit is somewhat lower than the 5.6 Hz found by Ranadive [7] but the difference may be attributed to the higher resolution monitors used in our experiments. Differences in the complexity of the tasks may also account for the difference.

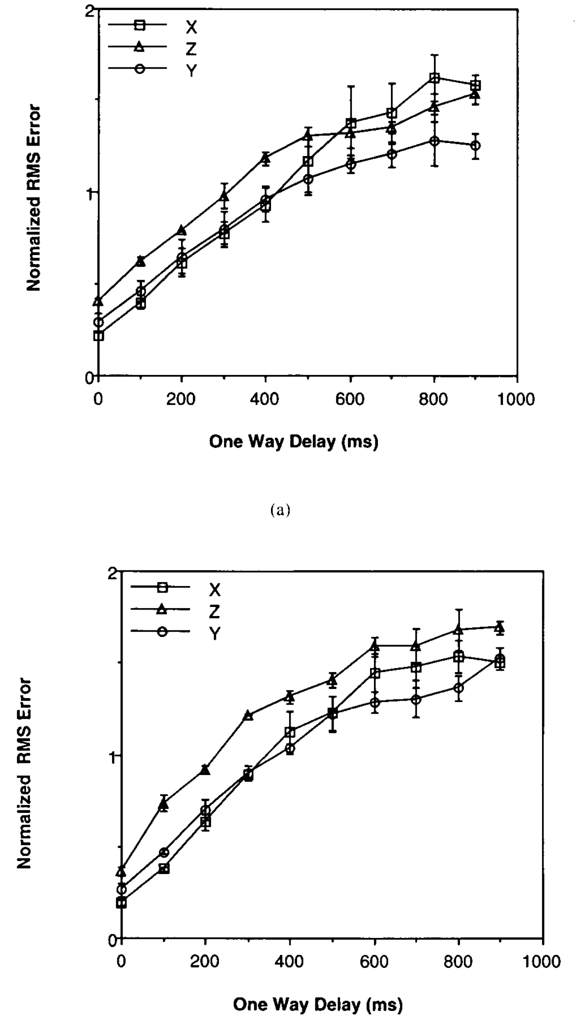
An interesting aspect of the update rate data is that faster update rates do not necessarily lead to better performance. Clearly for experienced subjects and to some extent for inexperienced subjects, increasing the update rate beyond 10 Hz produced only a small increase in performance. This mean that using limited bandwidth communication channels for transmitting video signals might still be practical even if full video update speed can not be obtained. It is even more important for systems incorporating graphics because these systems might use the extra rendering time available in a slow update condition to produce graphical enhancements to the display. Graphical enhancements to the display, such as reference lines or predictors, have already been shown to aid operator performance dramatically [26], [27]. Indeed, graphical top-down model enhancements improve performance even when added as superimposed lines onto a video image [26].



(a)

(b)

Fig. 9. Effect of display conditions on performance. Absence of occlusion and stereo strongly degraded performance; line rendering and absence of occlusion alone had a very small effect. (a) Average RMS Error for the five subjects along each axis. Error bars represent the standard error of the mean. □ = x-axis error, ○= y-axis error, △ = z-axis error. (b) The average mean completion time of the pick-and-place task for three subjects. Each data point is the average time to grab one ball and put it into its respective box. Error bars represent the standard error of the mean.



(a)

(b)

Fig. 10. Effect of delay on tracking performance. One way delay represents the time to send a command signal from the local to remote site. (a) Averaged delay response. Communication delay between a local and remote site has an immediate effect on tracking performance. (b) Delay response with overlay. Adding the overlay causes the tracking performance to degrade at a slightly faster rate than in the previous experiment. Error bars represent the standard error of the mean.

B. Display Conditions

Other graphical enhancements might take the form of polygon rendered objects with shading or stereo images. For our two simulated tasks, we have found that either disparity or occlusion provide enough information about the relative depth of objects to perform the task with nearly equivalent performance. In the case of occlusion, subjects determine the relative depth of the objects by noting whether the box blocks the view of the ball or the manipulator covers part of the ball. For disparity, the directional information can be obtained at disparity discontinuities that occur when lines from one object overlap the lines of the other. The subject is be able to discern which lines belong to which object by a mental model, such as a $2\frac{1}{2}$D sketch [28] or the quasi-3-D appearance of the Necker cube [29]. When both cues are present, performance is the best for both tasks. Undoubtedly, there is some cooperation between the cues that produces better depth judgements. However, since the improvement is small, other considerations such as maintaining the update rate should take precedence over realism in the display. We have planned more experiments to examine the contribution of other depth cues and graphical enhancements on telemanipulation. We will also examine the interaction of the display conditions with slower update rates for other effects on performance.

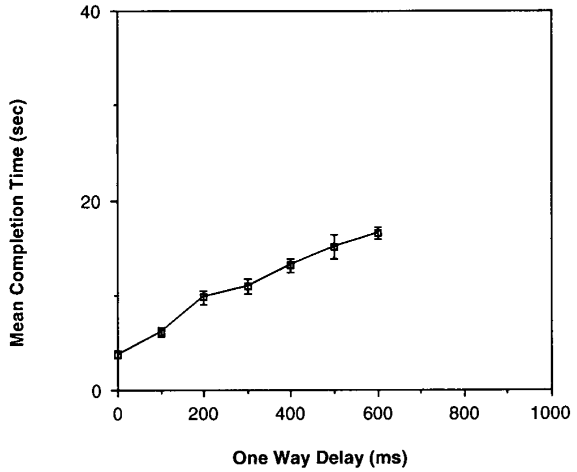


Fig. 11. Effect of delay on pick-and-place performance. One way delay represents the time to send a command signal from the local to remote site. Communication delay between the local and remote site has an immediate effect on pick-and-place performance. Error bars are the standard error of the mean.

C. Delay

Delay is one of the most difficult problems confronting telemanipulation. It can not be simply engineered out of the system. If we wish to accomplish various remote tasks by direct manual control, we may have to adapt to or accept the inherent delays involved and their consequent decrements in performance. Our data show that the HMD, like other displays, does not solve the problems encountered with delays and if not carefully tested may even exacerbate them. Implementing remote site autonomous control may be the best solution for space teleoperation. For example, active compliance control might be more effective in compensating for delay and avoiding serious collisions. Results from a JPL study indicate that peg-in-hole task performance under time-delayed conditions is better with compliance control [30], [31].

Our first attempt at a useful predictive display made the task more difficult for the subjects. We found that the most difficult parameter to control in the presence of delay was the view direction of the display. It was very difficult to coordinate move-and-wait strategies for the head and the hands. Perhaps a better system would include a local loop from the human to the display. In a completely modelled system, we can use the view direction immediately because it is measured at the "home" site. This might be implemented in the fashion of the "virtual dome", suggested by Hirose *et al.* [20] or by a virtual environment telerobotic system [32]. In the future, we plan to test a system that controls the display based upon immediate operator head motions and delay only that data that must come from the remote site. Ideally, the performance will only reflected degradations from the delay of the command signals to the manipulator, not from delayed head tracking.

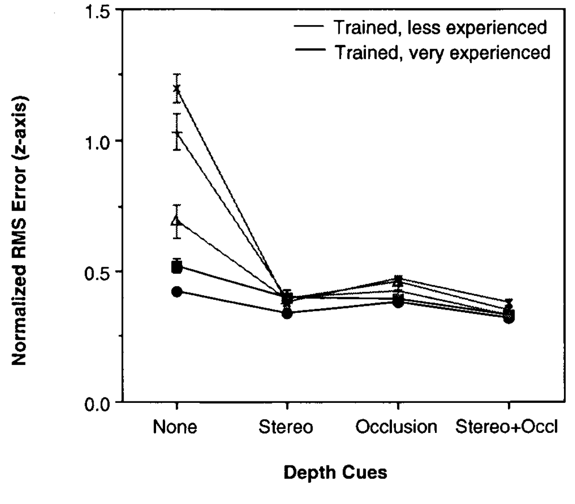
D. Physical Interactions Between Objects

The omission of realistic physical interactions between objects (i.e., objects in our simulation can pass through each other) could affect how depth cue are used in these tasks. In the pick-and-place task, occlusion is used by moving the manipulator through the objects, thereby determining their relative depth. Of course, this strategy is not possible in a real world situation. Attempting to ignore the physical constraints would result in damaged equipment and lost time! Depth from stereo, on the other hand, can be determined without any contact of this sort. So while occlusion is theoretically a useful cue for this type of task, it may not be usable in practice and is not necessary for the display. Experiments with "real" objects would be desirable to determine if occlusion is still a useful cue.

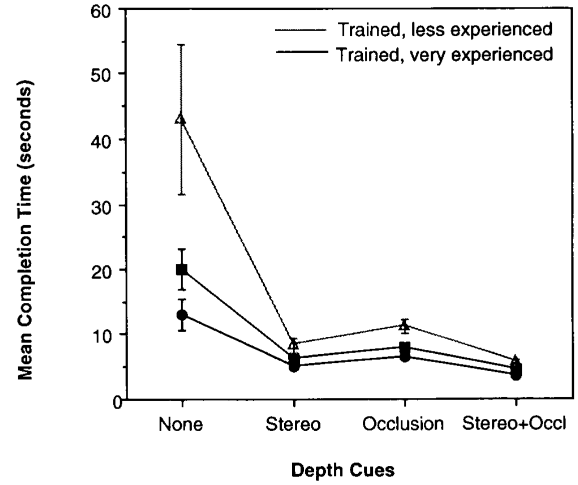
Graphical overlays such as the predictor used in the delay experiments will always be able to pass through "real" objects that are displayed in a video image. However, previous experiments have shown that the predictive display is not helpful for determining impending collisions without implementing collision detection algorithms [21]. Collision detection would be a desirable addition to the overlay, however, current algorithms are probably too slow to maintain a suitable display update rate [33], [34]. The compliance control mentioned above would be more computationally efficient for dealing with collisions between manipulator and objects. The overlay might flash or change color when the manipulator is in contact with another object.

E. Operator Experience

The effect of training and experience on task performance is an interesting result. More experienced subjects are better able to adapt to degraded conditions. For the update rate experiments, there is a noticeable shift in the performance curve. Experienced subjects may be using their familiarity with and knowledge of the task to direct their movements and improve performance. Plotting the results of the display condition experiment by subject, we see that the more experienced subjects suffer a smaller degradation in performance for the condition with no depth cues (see Fig. 12). For these conditions, the more experienced subjects may be using a perspective depth cue. In this case, subjects are probably using the relative sizes of the objects to complete the task. We can conclude that one result of experience is that subjects are able to utilize more of the information in a display than a less experienced operator.



(a)



(b)

Fig. 12. Effect of operator experience for the different display conditions. In both tasks, less experienced subjects had worse performance when stereo and occlusion cues were not present. Performance levels are comparable for the other conditions. (a) 3-D tracking task. (b) Pick- and-place task.

V. CONCLUSION

We have found, by examining some of the timing considerations for HMD performance, that there are some parameters of the display, such as update rate, that can be degraded without affecting performance. The extra time can be used for graphical enhancements such as stereo display or more realistic modeling of the remote environment. We have also found that some solutions to problems, such as a predictive display to compensate for delay may in fact do more harm than good. The HMD has the potential, with the correct properties, to allow the operator to use their predictive ability more efficiently and therefore compensate somewhat for delays in the system. Clearly, more research needs to be done examining the costs and benefits of each HMD parameter. Only then can an intelligent display be designed to meet the requirements of particular tasks.

REFERENCES

[1] C. L. Hanson, T. M. Longridge, R. Barrete, B. Welch, and R. Kruk, "Fiber-optic helmet-mounted-display for full visual flight simulation," *SID Digest*, pp. 112–115, 1984.

[2] M. Hirose, T. Myoi, H. Amari, K. Inamura, and L. W. Stark, "Development of visual 3D virtual environment for control software," in *Human Machine Interfaces for Teleoperators and Virtual Environments*, N. I. Durlach, T. B. Sheridan, and S. R. Ellis, Eds., vol. CP10071, NASA, 1990, pp. 120–124.

[3] F. P. Brooks, H. Fuchs, and W. Robinett, "Five applications using head-mounted displays," *Computer Graphics*, July 1991. Report of the *Tomorrow's Realities* venue at SIGGRAPH '91.

[4] S. S. Fisher, M. W. McGreevy, J. Humphries, and W. Robinett, "Virtual environment display system," in *Proc. of the ACM Workshop on Interactive 3D Graphics*, Chapel Hill, NC, Oct. 1986.

[5] W. S. Kim, A. Liu, K. Matsunaga, and L. Stark, "A helmet mounted display for telerobotics," in *Proc.1988 IEEE Computer Society COMPCON*, San Francisco, CA, pp. 543–547, 1988.

[6] M. J. Wells and M. Venturino, "Performance and head movements using a helmet-mounted display with different sized fields-of-view," *Opt. Eng.*, vol. 29, pp. 870–877, 1990.

[7] V. Ranadive and T. B. Sheridan, "Video framerate, resolution and grayscale tradeoffs for undersea manipulator control," in *Proc. 17th Annu. Conf. Manual Contr.*, Los Angeles, CA, 1981, pp. 77–85.

[8] R. E. Cole and B. H. Kishimoto, "Remote operator performance using bandwidth-limited TV displays: A review and proposal," Tech. Rep. 379, Naval Ocean Systems Center, 1981.

[9] L. Stark, F. Tendick, W. S. Kim, *et al.*, "Telerobotics: Problems and research needs," *IEEE Trans. Aero. Elec. Sys.*, vol. 24, pp. 542–551, Sept. 1988.

[10] R. Pepper and R. E. Cole, "Display system variables affecting operator performance in undersea vehicles and work systems," Tech. Rep. 269, Naval Ocean Systems Center, 1978.

[11] W. S. Kim, F. Tendick, and L. Stark, "Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator," *IEEE J. Robotics Automat.*, vol. 3, pp. 418–425, 1987.

[12] W. S. Kim, S. R. Ellis, M. E. Tyler, B. Hannaford, and L. Stark, "Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks," *IEEE Trans. Syst., Man, Cybern.*, vol. 17, pp. 61–72, 1987.

[13] G. Tharp, A. Liu, H. Yamashita, L. Stark, B. Wong, and J. Dee, "A helmet mounted display to adapt the telerobotic environment to human vision," in *Proc. 3rd Annu. Workshop Space Operations, Automat. and Robotics*, Houston, TX, July 1989, pp. 477–481.

[14] S. T. Bryson and S. S. Fisher, "Defining, modeling and measuring system lag in virtual environments," in *Stereoscopic displays and applications*, J. O. Merritt and S. S. Fisher, Ed., vol. 1256, pp. 98–109, SPIE, 1990.

[15] T. B. Sheridan and W. R. Ferrell, "Remote manipulation control with transmission delay," *IEEE Trans. Human Factors Electron.*, vol. HFE-4, pp. 25–28, 1963.

[16] L. Stark, W. S. Kim, and F. Tendick, *et al.*, "Telerobotics: Display, control, and communication problems," *IEEE J. Rob. Autom.*, vol. 3, pp. 67–75, 1987.

[17] J. L. Adams, "An investigation of the effects of time lag due to long transmission distances upon remote control," Tech. Rep. TN-1351, NASA, 1962.

[18] J. M. Leslie, "Effects of time delay in the visual feedback loop of a man-machine system," Tech. Rep. CR-560, NASA, 1966.

[19] R. Held, A. Efstathiou, and M. Greene, "Adaptation to displaced and delayed visual feedback from the hand," *J. Exp. Psych.*, vol. 72, no. 6, pp. 887–891, 1966.

[20] M. Hirose, K. Hirota, R. Kijima, M. Kanno, K. Hayakawa, and K. Yokoyama, "A study of synthetic visual sensation through artificial reality," in *Proc. . Seventh Symp. Human Interface*, Kyoto, Japan, pp. 675–682, Oct. 1991.

[21] M. V. Noyes and T. B. Sheridan, "A novel predictor for telemanipulation through a time delay," in *Proc. 20th Annual Conf. Manual Contr.*, Moffett Field, CA, 1984.

[22] A. K. Bejczy, W. S. Kim, and S. C. Venema, "The phantom robot: predictive displays for teleoperation with time delay," in *Proc. 1990 IEEE Int. Conf. Robotics Automat.*, Cincinnati, OH, May 1990, pp. 546–551.

[23] M. K. Duffy, "A head monitor system using the search coil method," Master's thesis, Dept. Elec. Eng., Univ. Calif., Berkeley, 1985.

[24] W. S. Kim, S. R. Ellis, M. Tyler, and L. Stark, "Visual enhancements for telerobotics," in *Proc. 1985 IEEE Int. Conf. Syst., Man, Cybern.*, Tuscon, AZ, 1985, pp. 819–823.

[25] A. Liu and L. Stark "The effect of monocular target blur on simulated telerobotic manipulation," *IEEE Trans. Robotics Automat.*, vol. 7, pp. 372–376, 1991.

[26] W. S. Kim, M. Takeda, and L. Stark, "On-the-screen visual enhancements for a telerobotic vision system," in *Proc. 1988 IEEE Int. Conf. Syst., Man, Cybern.*, Beijing, China, Aug. 1988, pp. 126–130.

[27] W. S. Kim and A. K. Bejczy, "Graphics displays for operator aid in telemanipulation," in *Proc. 1991 IEEE Int. Conf. Syst., Man, Cybern.*, Charlottesville, VA, Oct. 1991.

[28] D. Marr, *Vision*. New York: W. H. Freeman, 1984.

[29] S. R. Ellis and L. W. Stark, "Eye movements during the viewing of necker cubes," *Perception*, vol. 7, pp. 575–581, 1978.

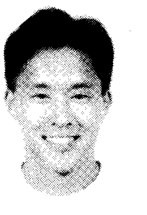
[30] Z. Szakaly, W. S. Kim, and A. K. Bejczy, "Force-reflective teleoperated system with shared and compliant control capabilities," in *Proc. NASA Conf Space Telerobotics*, G. Rodriguez and H. Seraji, Eds., Pasadena, CA, Jet Propulsion Lab., Jan. 1989, pp. 145–155.

[31] W. S. Kim, B. Hannaford, and A. K. Bejczy, "Force-reflection and shared compliant control in operating telemanipulators with time delay," *IEEE Trans. Robotics Automat.*, vol. 8, pp. 176–185, Apr. 1992.

[32] A. M. Liu, "Depth cues for telerobotic tasks in virtual environments," Ph.D. dissertation, Joint Bioengineering Graduate Group, Univ. Calif., Berkeley and San Francisco, 1992.

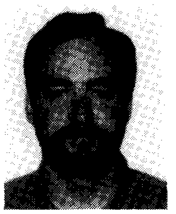
[33] S. Stifter, "Collision detection in the robot simulation system SMART," *Int. J. Adv. Manufacturing Tech.*, vol. 7, no. 5, pp. 277–284, 1988.

[34] C. A. Shaffer, "Real-time robot arm collision detection for telerobotics," *Computers & Electrical Engineering*, vol. 17, no. 3, pp. 205–215, 1991.



Andrew Liu received the B.S. degree in biomedical engineering from The Johns Hopkins University, Baltimore, MD, in 1986 and the Ph.D. from the Joint Program in Bioengineering at the Berkeley and San Francisco campuses of the University of California, in 1992.

At Berkeley, his dissertation research focused on the study of human depth perception in virtual environments and the evaluation of virtual environment interfaces for telerobotic manipulation. He completed a postdoctoral appointment at the University of California at Berkeley, a joint project with the NASA-Ames Research Center, studying telepresence and its relationship to task performance in virtual environments. Presently, he is a research scientist at the Nissan Cambridge Basic Research Laboratory, Cambridge, MA, where he is a project manager for the testing and validation of the 240SX driving simulator, which will be used in basic vision research. His current interest are human perception, visual displays, and virtual environments.



Gregory K. Tharp received the B.S. in mechanical engineering from University of California, Berkeley, holds the B.A. degree in economics from the University of California, Santa Cruz, and received the M.S. degree in mechanical engineering from the University of California, Berkeley, in 1989.

After a one-year research staff appointment at University of California, Berkeley, he worked as a consultant for Fujita Research for three years, studying Virtual Environment and Telepresent Display Systems for remote manipulation. He has recently joined the Man Machine Human Factors Group in the Robotic Systems and Advanced Computer Technology Section of the Jet Propulsion Laboratory in Pasadena, California. At JPL, he is working on the Operator Interface and the Image Processing Systems of the Remote Surface Inspection Project.

Lloyd C. French received the B.S. and M.S. degrees in mechanical engineering in 1989 and 1993, respectively, from the University of California at Berkeley, where he is a Ph.D. candidate in the area of mechanical engineering.

He has worked with the Man–Machine Human Factors Group (now called Robotics Systems and Advanced Computer Technology Section) at the Jet Propulsion Laboratory/NASA in Pasadena, California. This group researched human performance in manipulating robots within workspaces. He built and tested a six-degree of freedom force reflecting hand controller (FRHC). He and the group used the controller device to study human performance under various control modes such as compliance and delayed feed back. He assisted in designing pick and place testbed experiments to study usefulness of telerobotic technology for space station and satellite repair and maintenance. He continues to do research at the Telerobotics Unit at the University of California on see-through head mounted display technology.

Mr. French is a member of the Minority Engineering Program (MEP) and the Black Graduate Engineering and Scientist Students (BGESS) at the University of California at Berkeley and an associate member of the ASME.

Stephen Lai, biography and photograph are not available.

Lawrence W. Stark (SM'61–F'70) is a Professor of Engineering Science and of Physiological Optics at the University of California, Berkeley, and Professor of Neurology (Neuro-Ophthalmology) at the University of California, San Francisco. His recent research in telerobotics explores visual display interfaces between man and nonautonomous robots and control problems introduced by communication delays. Image compression by modeling is similar to the role of cognitive models in human visual perception, and robotic and human neurological control have interesting parallels.