N90-22946

VISUAL ENHANCEMENTS IN PICK-AND-PLACE TASKS: HUMAN OPERATORS CONTROLLING A SIMULATED CYLINDRICAL MANIPULATOR

Won S. Kim, Frank Tendick, and Lawrence Stark
University of California
Berkeley, California

ABSTRACT

A visual display system serves as an important human/machine interface for efficient tele-operations. However, careful consideration is necessary to display three-dimensional information on a two-dimensional screen effectively. A teleoperation simulator is constructed with a vector-display system, joysticks, and a simulated cylindrical manipulator in order to evaluate various display conditions quantitatively. Pick-and-place tasks are performed, and mean completion times are used as a performance measure. Two experiments are performed. First, effects of variation of perspective parameters on a human operator's pick-and-place performance with monoscopic perspective display are investigated. Then, visual enhancements of monoscopic perspective display by adding a grid and reference lines are investigated and compared with visual enhancements of stereoscopic display. The results indicate that stereoscopic display does generally permit superior pick-and-place performance, while monoscopic display can allow equivalent performance when it is defined with appropriate perspective parameter values and provided with adequate visual enhancements. Mean-completion-time results of pick-and-place experiments for various display conditions shown in this paper are observed to be quite similar to normalized root-mean-square error results of manual tracking experiments reported previously.

INTRODUCTION

Visual display systems serve as an important human/machine interface for efficient teleopera-tions in space, underwater, and in radioactive environments.[1-4] Closed-circuit television systems, presenting two-dimensional (2-D) images captured by remote video cameras, have been commonly used for these visual displays. As technology evolves from manually controlled teleoperations to sensor/computer-aided advanced teleoperations[5,6] or telerobotics,[7-11] graphics displays have been drawing attention as a means to provide an enhanced human/machine interface. A graphic display can present an abstract portrayal of the working environment or state of the control system based on sensor signals and a data base.[2,12] A force-torque display[13] and a "smart" display[14] are exam-ples of graphic displays developed for efficient teleoperations.

There are two types of visual displays: monoscopic and stereoscopic. The stereoscopic dis-play provides two slightly different perspective views for the human operator's right and left eyes. A stereoscopic view enables the human to perceive depth by providing a distinct binocular depth cue called stereo disparity. Some earlier studies with television displays showed that stereoscopic displays, as compared to monoscopic displays, did not provide significant advantage in performing some telemanipulation tasks.[15-17] Careful recent studies,[18,19] however, indicated that stereo per-formance was superior to mono under most conditions tested, while the amount of improvement

varied with visibility, task, and learning factors. These results showed that the advantage of the stereoscopic television display became pronounced with increased scene complexity and decreased object visibility.

Monoscopic and stereoscopic graphic displays were recently compared by employing three-axis manual tracking tasks.[20,21] Root-mean-square (rms) tracking error was used as a performance measure for quantitative evaluation. Results were consistent with previous television display results, indicating that stereoscopic graphic displays did generally permit superior tracking performance, while monoscopic displays allowed equivalent performance when they were defined with appropriate perspective parameters and provided with adequate visual-enhancement depth cues such as reference lines.

The purpose of our present study is to examine generality or consistency of the above results. A three-axis pick-and-place task, instead of the three-axis manual tracking task, is employed in our present study as a realistic teleoperations task. Two experiments similar to those in reference 21 are performed. In the first experiment, we quantitatively evaluate monoscopic perspective display by investigating individual effects of perspective parameters. Perspective projection alone, however, does not provide sufficient three-dimensional (3-D) depth information for monoscopic display. Thus, a 5-line-by-5-line horizontal grid representing a base plane and a vertical reference line representing vertical separation from the base plane are introduced as two visual-enhancement depth cues. In the second experiment, we investigate effects of these two visual-enhancement depth cues on pick-and-place performance for both monoscopic and stereoscopic displays.

METHODS

In order to evaluate various display conditions quantitatively, a teleoperations simulator is constructed with a vector-display system, joysticks, and a simulated cylindrical manipulator. Figure 1 shows a schematic diagram of the experimental setup, with which three-axis pick-and-place tasks are performed.

Real-Time Simulation of The Manipulator

The Hewlett-Packard 1345A vector-display module is used for real-time dynamic display. It has high resolution (2048 x 2048 addressable data points), and high vector-drawing speed (8194 cm of vectors at 60-Hz refresh rate). It also has a fast vector-updating speed (approximately 10 μsec/vector), communicating with a host computer through a 16-bit parallel I/O port. Two isotonic (displacement) joysticks are employed for the Cartesian position control of the manipulator gripper. An LSI-11/23 computer with the RT-11 operating system is used as a host computer. It performs computations for the simulated manipulator motion and perspective or stereoscopic display, and measures task completion time.

The human operator indicates the desired gripper position of the manipulator in robot base Cartesian coordinates by using three axes of the two joysticks. The computer senses the joystick displacements through 12-bit A/D converters. The joystick gain for each axis is chosen to be 1 so that the full range of the joystick displacement for each axis corresponds to the full movement range of the gripper position for the corresponding axis. The computer transforms the desired

gripper position in Cartesian coordinates to the desired joint angle (θ_1 for the revolute joint 1) and joint slidings ($d2$ and $d3$ for the prismatic joints 2 and 3) by employing the inverse kinematic position transformation. The next two sections describe how to present 3-D information of the manipulator on the 2-D display screen.

Monoscopic Perspective Display

A monoscopic perspective display can be constructed by a perspective projection of an object onto the view plane (projection plane) followed by a mapping of the view plane onto the screen.[22] There are two approaches to obtaining the perspective projection of an object. One is to leave the object stationary and choose a desired viewpoint and a projection plane, called the viewpoint-transformation method. The other approach is to fix the viewpoint and transform the object, called the object-transformation method. These two approaches are mathematically equivalent.[21,23] The latter will be described here.

In order to derive the perspective display formulas based on the object-transformation method, a right-handed XYZ world coordinate system is established. The viewpoint is fixed at the origin $(0, 0, 0)$ and the view plane at the $z = -d$ plane. Perspective projection can be obtained by three transforms: rotation R, translation T, and perspective transform P.

Initially, an object is located so the view reference point of the object is at the origin. Then the object is appropriately rotated and translated to achieve the desired viewing angles and distance. In general, an arbitrary orientation of an object can be described by successive principal-axis rotations about the Y, X, and Z axes.

$$R = Rot(Y, -\theta_1) \, Rot(X, \theta_2) \, Rot(Z, \theta_3) \tag{1}$$

where the yaw, pitch and roll angles are $-\theta_1$, $-\theta_2$, and θ_3, respectively. It can be shown that the yaw and pitch angles used in the object transformation approach are equivalent to the azimuth and elevation angles in the viewpoint-transformation approach.[21]

For simplicity, 4-space homogeneous coordinate transformations are used. The rotation of a point at position (x, y, z) to a new position (x', y', z') can be described by

$$(x', y', z', 1) = (x, y, z, 1) \, R \tag{2}$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} & 0 \\ R_{21} & R_{22} & R_{23} & 0 \\ R_{31} & R_{32} & R_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

From equation (1), each element of the 4 x 4 matrix R can be calculated as
$R_{11} = C_1C_3 - S_1S_2S_3$, $R_{12} = -C_1S_3 - S_1S_2C_3$, $R_{13} = S_1C_2$, $R_{21} = C_2S_3$, $R_{22} = C_2C_3$, $R_{23} = S_2$, $R_{31} = -S_1C_3 - C_1S_2S_3$, $R_{32} = S_1S_3 - C_1S_2C_3$, $R_{33} = C_1C_2$. S_i and C_i denote $\sin \theta_i$ and $\cos \theta_i$, respectively.

After the rotation, the object is translated by D along the negative Z axis.

$$T = \text{Trans}(0, 0, -D)$$ (4)

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -D & 1 \end{bmatrix}.$$ (5)

The length D represents the distance from the viewpoint to the view reference point, called the object distance.

The UV coordinate system is embedded in the view plane. Perspective transformation of a point Q (x, y, z) in the world coordinate to its projection Q_p (u, v) on the view plane can be described by

$$(x', y', z', w) = (x, y, z, 1)\ P$$ (6)

$$(u, v) = (x'/w, y'/w)$$ (7)

where

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1/d \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$ (8)

The symbol d denotes the view plane distance from the viewpoint. Increase of the view plane distance results in uniform magnification of the perspective projection. Thus, d can be specified in terms of the zoom or magnification factor, which can be defined as $M = d/D$. Distance d can also be specified in terms of field-of-view (fov) angle, which is the angle at the viewpoint subtended by the view-plane window. If the view plane window is specified as a square region $(u_{min}, u_{max}, v_{min}, v_{max}) = (-1, 1, -1, 1)$, then the fov angle is related to the view-plane distance by $d = \cot(\text{fov}/2)$. The perspective projection obtained with a wide fov angle is similar to the picture taken by a wide-angle camera lens, and a narrow fov angle is similar to one taken by a tele-photo lens.

After the object is projected onto the view plane, mapping of the view plane onto the physical display screen is performed. Mapping of a point from (u, v) in the UV coordinate to (x_s, y_s) in the screen coordinate can be achieved by appropriate translations and scalings:

$$x_s = \text{VSX}\ u + \text{VCX}$$ (9)

$$y_s = \text{VSY}\ v + \text{VCY}$$ (10)

where VSX and VSY are scaling factors, and VCX and VCY are translation factors.

Stereoscopic Display

The monoscopic display does not give true depth perception. The human brain merely interprets the 2-D monoscopic picture as 3-D space. The stereoscopic display presents two views of an object on the display: one for the right eye, and the other for the left. This pair of pictures is called a stereo pair or a stereogram. The human operator views a stereogram through a stereoscope.[24] Most people can fuse the stereo pair into one 3-D image, perceiving relative depth by the human stereoscopic vision ability. The stereoscope is composed of two converging lenses and a supporting frame (septum) separating right and left views. As illustrated in figure 2, two converging lenses form the image of the stereo pair onto the image plane behind the actual display screen, which can provide fairly correct accommodation and convergence conditions for the human eyes, if the geometrical and optical conditions are appropriately arranged.

In order to obtain the formulas for the stereoscopic display, an XYZ coordinate system is established with its origin in the middle of the two optical centers for the right and left eyes, as depicted in figure 2. The display screen, on which a stereogram is presented, is located at the picture plane (view plane, projection plane) $z = -d$. The two converging lenses of the stereoscope form the virtual image of the stereogram on the image plane $z = -D$. By denoting the focal length of the binocular lens as F, the converging lens formula yields

$$\frac{1}{d} - \frac{1}{D} = \frac{1}{F} \tag{11}$$

When D is infinity, $d = F$. When $D = 40$ cm and $F = 20$ cm, $d = 13.3$ cm.

As in the object-transformation approach used previously for the monoscopic perspective display, the object is initially located so the view-reference point of the object is at the origin. Then the object is appropriately rotated and translated using equations (3) and (5) to achieve the desired viewing angles and distance.

Denoting the interocular distance (IOD) (approximately 5.5 to 6.5 cm), we can express the positions of the two optical centers by $(x_{or}, 0, 0)$ for the right eye and $(x_{ol}, 0, 0)$ for the left eye, where $x_{or} = IOD/2$, and $x_{ol} = -IOD/2$. The projection of a point P (x, y, z) onto the view plane for each eye is formed at the intersection of the projection line with the view plane. By representing the right and left projection points by $P_r (x_r, y_r)$ and $P_l (x_l, y_l)$, respectively, the following equations can be obtained:

$$x_r = x_{or} + (x - x_{or})(-d/z) \tag{12}$$

$$x_l = x_{ol} + (x - x_{ol})(-d/z) \tag{13}$$

$$y_r = Y_l = y (-d/z) \tag{14}$$

Finally, these projection points on the projection plane can be mapped onto the physical screen coordinates by appropriate translations and scalings.

Experimental Procedures

Two sets of experiments were performed, varying perspective parameters and visual enhancement conditions. In both experiments, subjects were seated in front of the display (on which the manipulator, the objects to pick up, and the boxes to place them in were presented) (fig. 3), and the subjects were asked to perform three-axis pick-and-place tasks. The subjects controlled the manipulator using two joysticks to pick up each object with the manipulator gripper and place it in the corresponding box. One hand, using two axes (forward-backward and right-left) of one joystick, controlled the gripper position for the two axes parallel to the horizontal base plane. The other hand, using one axis (forward-backward) of the other joystick, controlled the vertical axis.

Each of the four objects (point targets A, B, C, D) was positioned randomly within the manipulator reach space. Each object position was marked by a tiny diamond and a letter. Picking up an object was accomplished when the manipulator gripper touched the object within the boundary of the error tolerance, defined by a hypothetical cube. The size of the cube was set so that the picking process was neither too easy nor too hard within the range of experimental variation. Accomplishment of picking up an object was indicated by doubling the object letter. Thereafter, the object moved together with the gripper until it was placed in the right box. Placing an object was accomplished by touching the correct box with the gripper, similar to the picking process. After the touch, the object symbol letter became single again, and the object remained in the box, while the gripper was free to move for the next operation.

One run of the pick-and-place task consisted of five sessions of four pick-and-place operations in order from object A to D, totaling 20 pick-and-place operations.

Perspective Parameter Experiment. In this experiment, we investigated the effects of different perspective parameters on the human operator's pick-and-place performance with monoscopic perspective display. The five perspective parameters, azimuth, elevation, roll, fov angle, and object distance were independently varied, keeping the other variables fixed at their nominal values. The nominal perspective parameter values were chosen as elevation = –45°, azimuth = 0°, roll = 0°, fov angle = 12°, and object distance = 40 cm.

Experimental variables were varied as follows: (1) seven elevation angles: 0°, –15°, –30°, –45°, –60°, –75°, and –90°; (2) eight azimuth angles: –135°, –90°, –45°, 0°, 45°, 90°, 135°, and 180°; (3) eight roll angles: –135°, –90°, –45°, 0°, 45°, 90°, 135°, and 180°; (4) five fov angles: 8°, 12°, 24°, 48°, and 64°, (5) four object distances: 30, 40, 80, and 160 cm.

The monoscopic perspective presentation with the nominal perspective parameters is shown in figure 3. Some examples of variations in perspective parameter values used in this experiment are shown in figure 4. In this experiment, a 5-line-by-5-line horizontal grid and vertical reference lines were always presented. The experiment was run with each of the 32 experimental conditions presented in random order. There were two runs of 20 pick-and-place operations per condition for each subject. For the monoscopic conditions, the subjects were seated 40 cm in front of the display screen.

Visual Enhancement Experiment. In this experiment, effects of visual enhancements on the human operator's pick-and-place performance were investigated. The visual-enhancement depth cues used for both monoscopic and stereoscopic displays were a grid and reference lines. Three-axis pick-and-place tasks were performed for four visual-enhancement conditions at each of five different perspective parameter conditions with both monoscopic and stereoscopic displays. The

four visual-enhancement conditions were: GL (presence of both grid and reference line), L (reference line only), G (grid only), and O (neither). The five perspective parameter conditions used were: (1) 0° in elevation, (2) –90° in elevation, (3) nominal perspective parameter values, (4) 45° in azimuth, and (5) 80 cm in object distance.

Monoscopic presentations for the four visual-enhancement conditions with the nominal perspective parameters (condition III) are shown in figure 5. Monoscopic presentations for the five perspective parameter conditions, when both grid and reference lines are presented (condition GL), are shown above the mean completion time plot in figure 8. A stereoscopic presentation with the nominal perspective parameters, when both grid and reference lines are presented, is shown in figure 6. The experiment was run first with each of the 20 monoscopic display conditions presented in random order, then with each of the 20 stereoscopic display conditions presented in random order. There were two runs of 20 pick-and-place operations per condition for each subject.

In the monoscopic display conditions, subjects were seated 40 cm in front of the screen. In the stereoscopic display conditions, subjects were seated 13.3 cm in front of the screen, viewing the stereogram through the stereoscope. The focal length of the converging lens of the stereoscope was 20 cm, and thus the virtual image of the stereogram was formed at 40 cm from the lens (by eq. (11)).

Subjects

Two young adult male subjects with normal stereo vision participated in each of the two experiments. Each subject was trained for at least 5 hr before the experiments to saturate the "learning" effect. During the training period, mean completion times were regularly checked to see whether the subject reached an asymptotic, steady-state, pick-and-place performance. However, during the actual experiment, mean completion times were not checked until all the experimental runs were completed. Each subject repeated the experiment once more in order to examine intra-subject variation as well as inter-subject variation.

EXPERIMENTAL RESULTS

Mean completion time was used as the performance measure in our pick-and-place tasks. Each of the mean completion time data points in figures 7 and 8 is the average obtained from one run of 20 pick-and-place operations.

The experimental results for two subjects with two runs each plotted in figure 7 with mean completion time as the ordinate and perspective parameter values as the abscissa. The effects of elevation, azimuth, roll, fov angle, and object distance are plotted in figure 7 (a), (b), (c), (d), and (e), respectively.

The experimental results for two subjects with two runs each are shown in figure 8. Mean completion time (ordinate) is plotted for the various display conditions (abscissa). The monoscopic display data are marked by squares and dashed lines, and the stereoscopic display data are marked by filled diamonds and solid lines. The five separate columns represent five different per-

spective parameter settings, conditions 1-5. Each column has four different visual-enhancement conditions, GL, L, G, and O.

DISCUSSION

Effects of Perspective Parameters

The mean-completion-time plots of figure 7 show the effects of variation of perspective parameters on pick-and-place performance. Plot (a) shows that as the elevation angle approaches 0° or –90°, mean completion time increases. This is due to the loss of one axis' position information. Performance at –90° elevation was better than at the 0° extreme because the perspective view at –90° elevation made it possible to see some of the height of the reference line if it was not near the center of the projected image. Thus, there was a partial view of the "lost" axis. Plot (b) shows that as the azimuth angle exceeds the range of –45° to +45°, the mean completion time increases markedly. An azimuth angle other than 0° implies rotation of the display reference frame relative to the joystick control axes, thus making the joystick control more difficult compared to the 0° azimuth angle. When the azimuth angle is beyond –45° to +45°, it is difficult for the human operator to compensate. Performance is especially poor when the azimuth angle is about –90° or +90°, even worse than the case when azimuth angle is 180°. At 180° azimuth angle, the human operator uses inversion rather than rotation. Plot (c) shows that change in roll angle produces an effect similar to changing the azimuth angle, because of analogous disorientation. Plots (d) and (e) show that as the fov angle or the object distance increases, and the displayed object picture becomes smaller, task performance degrades.

Effects of Visual Enhancements

The results of the visual-enhancement experiment appear in figure 8 (a) and (b). Monoscopic display results in columns I and II show that when the elevation angle is 0° or –90°, the mean completion times are very long, even with grid or reference line enhancements. This is because position information for one axis is lacking, and the subject must sweep the gripper along that axis until it touches the correct position. At –90° elevation, the reference lines almost disappear. At 0° elevation, the grid appears as a single line. Monoscopic display results in columns III, IV, and V show that by choosing adequate elevation angles, mean completion times can be shortened, and fast pick-and-place performance can be attained with monoscopic perspective display, if reference lines are provided (GL, L). However, the grid alone without the reference line (G) does not appear to shorten completion time.

The stereoscopic display results in figure 8 show that mean completion times with stereoscopic display are short over all visual conditions, regardless of the presence of a grid or reference lines. Especially, stereoscopic display data in columns I and II show that stereoscopic displays maintain fast performance even with extreme elevation angles. Comparable mean completion times between monoscopic and stereoscopic displays in columns III, IV, and V demonstrate that pick-and-place performance with monoscopic perspective displays, if reference lines are provided and suitable perspective parameters are chosen, can be as good as that with stereoscopic displays.

Comparison With Three-Axis Manual Tracking Tasks

It is observed that the mean-completion-time plots obtained from the pick-and-place experiments in this paper are quite similar to the normalized rms tracking error plots obtained from the three-axis manual tracking experiments in reference 2. This strong similarity suggests that the results obtained in this paper are not task-specific, but may be applicable to other tasks.

Choice of Display

There are many kinds of depth cues that a display can provide. Monoscopic display can provide monocular depth cues such as interposition (occlusion), brightness (light and shade), perspective projection (size), and monocular motion parallax. The human operator's knowledge and learning can also provide strong depth information pertaining to a 3-D model of a working environment. Stereoscopic display also provides a distinct binocular depth cue, called stereo disparity or binocular parallax. Consideration of these cues basically explains the experimental results of Pepper, Smith, and Cole.[18] Their results indicated that stereoscopic display performance was superior to monoscopic display performance under most conditions tested, although the amount of improvement varied with task, visibility, and learning factors. For some simple telemanipulation tasks, monocular depth cues and cognitive depth cues from knowledge and learning may be enough for successful and reliable performance, and there will be no advantage in using stereoscopic display.[15] However, for some complex tasks, monocular and cognitive depth cues may be insufficient or unavailable for successful performance with monoscopic display, and the use of stereoscopic display could significantly enhance performance. In our experiments, monocular depth cues were minimized, and target positions were randomly arranged to minimize learning effect. Consequently, our experimental results showed that pick-and-place performance with stereoscopic display was superior to monoscopic display when visual-enhancement depth cues were not presented.

Our results also showed that when reference lines were presented for visual enhancement, monoscopic display performance with adequate perspective parameters was equivalent to stereoscopic display performance. In order to present reference lines on the monoscopic display, 3-D position information of the displayed objects must be available. In a graphic display of current manipulator and camera positions, 3-D position information is normally available via joint position sensors, and reference lines can be easily provided. In a television image display of the working environment, only camera views are normally available for 3-D position information. Under current technology, a machine vision system that extracts 3-D position information of each pixel in real time from a stereo camera view is too difficult to construct,[25] although the human visual system can easily produce a 3-D image from a stereoscopic view. However, a special-purpose, machine-vision system that extracts 3-D position information of only some salient points in real time from a stereo camera view can be built. Then, reference lines for these points can be presented or superimposed on the monoscopic television display for enhanced teleoperation.

CONCLUSION

Results of the perspective parameter experiments indicate that in order to attain good performance with a monoscopic perspective display, adequate parameter values should be chosen. For

example, extreme elevation angles or excessive azimuth angles result in very long mean completion times. Results of the visual-enhancement experiment indicate that the horizontal grid does not appear to improve pick-and-place performance in our task. The vertical reference line, however, was significant in improving performance with monoscopic perspective display. When the mono-scopic display was defined with appropriate perspective parameters and provided with adequate visual-enhancement depth cues such as reference lines, the monoscopic display allowed pick-and-place performance equivalent to that of the stereoscopic display. Stereoscopic display showed short mean completion times over all visual display conditions regardless of the presence of the grid or the reference lines.

Strong similarities were observed between the mean completion time results of the three-axis pick-and-place experiments for various display conditions and the normalized rms error results of the three-axis manual tracking experiments reported previously. This demonstrates that the effects seen are robust and not task-dependent.

Acknowledgment

REFERENCES

[1]E. G. Johnsen and W. R. Corliss, Teleoperators and Human Augmentation, NASA SP-5047, 1967.

[2]E. G. Johnsen and W. R. Corliss, Human Factors Applications in Teleoperator Design and Operation, Wiley-Interscience, 1971.

[3]E. Heer, Remotely Manned Systems: Exploration and Operation in Space, California Institute of Technology, 1973.

[4]J. Vertut and P. Coiffet, Teleoperations and Robotics: Evolution and Development, Robot Technology Volume 3A, Prentice-Hall, 1986.

[5]A. K. Bejczy, "Sensors, Controls, and Man-Machine Interface for Advanced Teleoperation," Science, Vol. 208, No. 4450, pp. 1327-1335, 1980.

[6]J. Vertut, R. Fournier, B. Espiau, and G. Andre, "Sensor-aided and/or Computer-aided Bilateral Teleoperator System (SCATS)," RO.MAN.SY Proc. of the 5th Symposium (also in Theory and Practice of Robots and Manipulators, ed. by A. Morecki, G. Bianchi, and K. Kedzior, pp. 281-292, 1984).

[7]W. R. Ferrell and T. B. Sheridan, "Supervisory Control of Remote Manipulation," IEEE Spectrum, Vol. 4, No. 10, pp. 81-88, 1967.

[8]M. A. Bronez, M. M. Clarke and A. Quinn, "Requirements Development for a Free-Flying Robot—The Robin," Proc. of IEEE Robotics and Automation, Vol. 1, pp. 667-672, 1986.

[9]L. M. Jenkins, "Telerobotic Work System—Space Robotics Application," Proc. of IEEE Robotics and Automation, Vol. 2, pp. 804-806, 1986.

[10]M. D. Montemerlo, "NASA's Automation and Robotics Technology Development Program," Proc. of IEEE Robotics and Automation, Vol. 2, pp. 977-986, 1986.

[11]M. M. Clarke and M. A. Bronez, "Telerobotics for the Space Station," Mech. Eng., Vol. 108, No. 2, pp. 66-72, 1986.

[12]K. Leinemann and E. G. Schlechtendahl, "Computer Graphics Support for Remote Handling Simulation and Operation," Trans. Am. Nucl. Soc., Vol. 46, pp. 771-772, 1984.

[13]A. K. Bejczy, R. S. Dotson, J. W. Brown, and J. L. Lewis, "Manual Control of Manipulator Forces and Torques using Graphic Display," IEEE Proc. of the Int. Conf. on Cybernet. and Society, pp. 691-698, 1982.

[14]A K. Bejczy, J. W. Brown, and J. L. Lewis, "Evaluation of Smart Sensor Displays for Multidimensional Precision Control of Space Shuttle Remote Manipulator," 16th Ann. Contr. Conf. on Manual Contr., pp. 607-627, 1980.

[15]W. N. Kama and R. C. DuMars, "Remote Viewing: A Comparison of Direct Viewing, 2-D and 3-D Television," AMRL-TDR-64-15, Wright-Patterson Air Force Base, 1964.

[16]P. Rupert and D. Wurzburg, "Illumination and Television Considerations in Teleoperator Systems," in Remotely Manned Systems, ed. by E. Heer, pp. 219-228, California Inst. Technol., 1973.

[17]L. A. Freedman, W. H. Crooks, and P. P. Coan, "TV Requirements for Manipulation in Space," Mechanism and Machine Theory, Vol. 12, No. 5, pp. 425-438, 1977.

[18]R. L. Pepper, D. C. Smith, and R. E Cole, "Stereo TV Improves Operator Performance Under Degraded Visibility Conditions," Opt. Eng., Vol. 20, No. 4, pp. 579-585, 1981.

[19]R. L. Pepper, R. E. Cole, E. H. Spain, and J. E. Sigurdson, "Research Issues Involved in Applying Stereoscopic Television to Remotely Operated Vehicles," Proc. SPIE, Vol. 402—Three-Dimensional Imaging, pp. 170-173, 1983.

[20]W. S. Kim, S R. Ellis, M. Tyler, and L. Stark, "Visual Enhancement for Telerobotics," IEEE 1985 Proc. Int. Conf. on Cybern. and Society, pp. 819-823, 1985.

[21]W. S. Kim, S. R. Ellis, B. Hannaford, M. Tyler, and L. Stark, "A Quantitative Evaluation of Perspective and Stereoscopic Displays in Three-Axis Manual Tracking Tasks," IEEE Trans. Syst., Man, Cybern., Vol. SMC-17, No. 1, pp. 61-72, 1987.

[22] J. D. Foley and A. Van Dam, Fundamentals of Interactive Computer Graphics, Addison-Wesley, 1983.

[23]I. Carlbom and J. Paciorek, "Planar Geometric Projections and Viewing Transformation," Computing Surv., Vol. 1, No. 4, pp. 465-502, 1978.

[24]E. F. Pischel and J. J. Pearson, "Image Processing and Display in Three Dimensions," Proc. SPIE, Vol. 528—Digital Image Processing, pp. 23-28, 1985.

[25]M. Drumheller and T. Poggio, "On Parallel Stereo," Proc. of IEEE Int. Conf. on Robotics and Automation, Vol. 3, pp. 1439-1448, 1986.
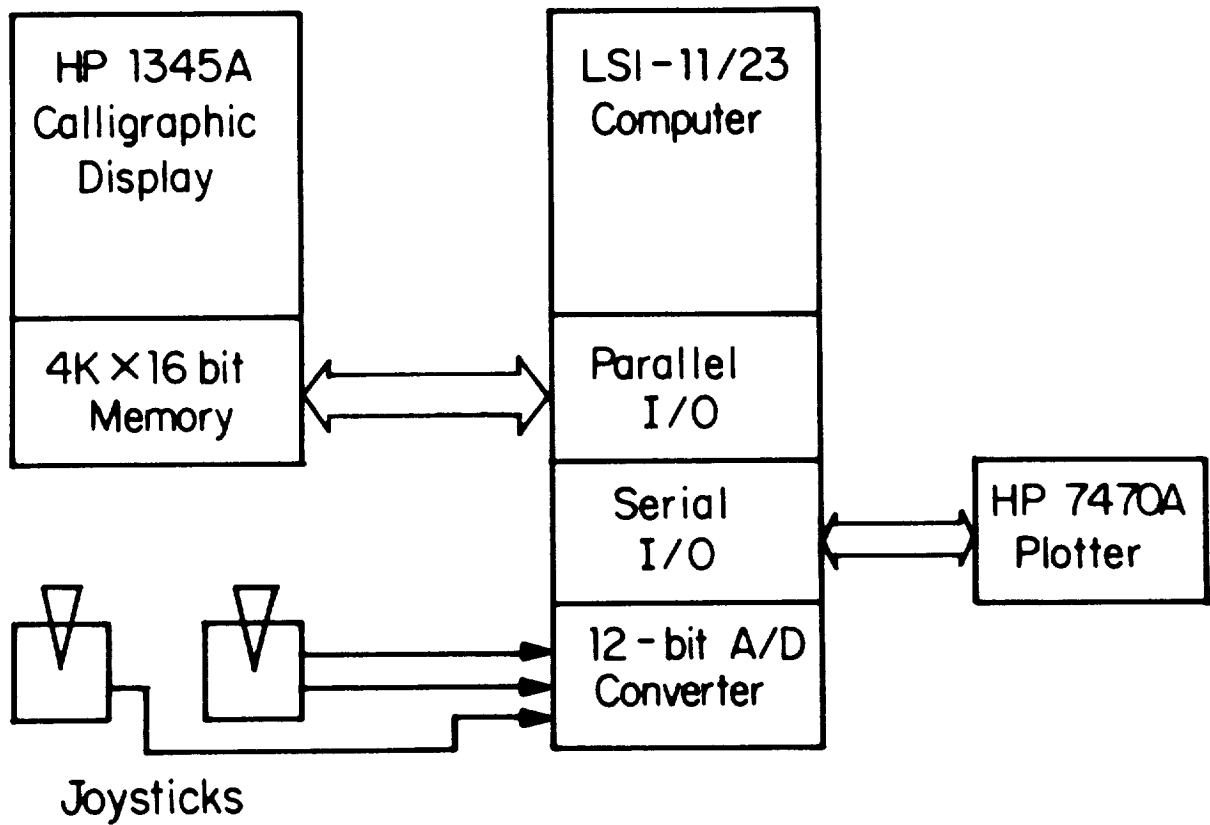
Figure 1.– The experimental setup.

left
retinal
image

right
retinal
image

Z Y

X

P

R

S

Q

retina pupil & crystalline lens

binocular converging lens

picture plane

(display screen)

image plane
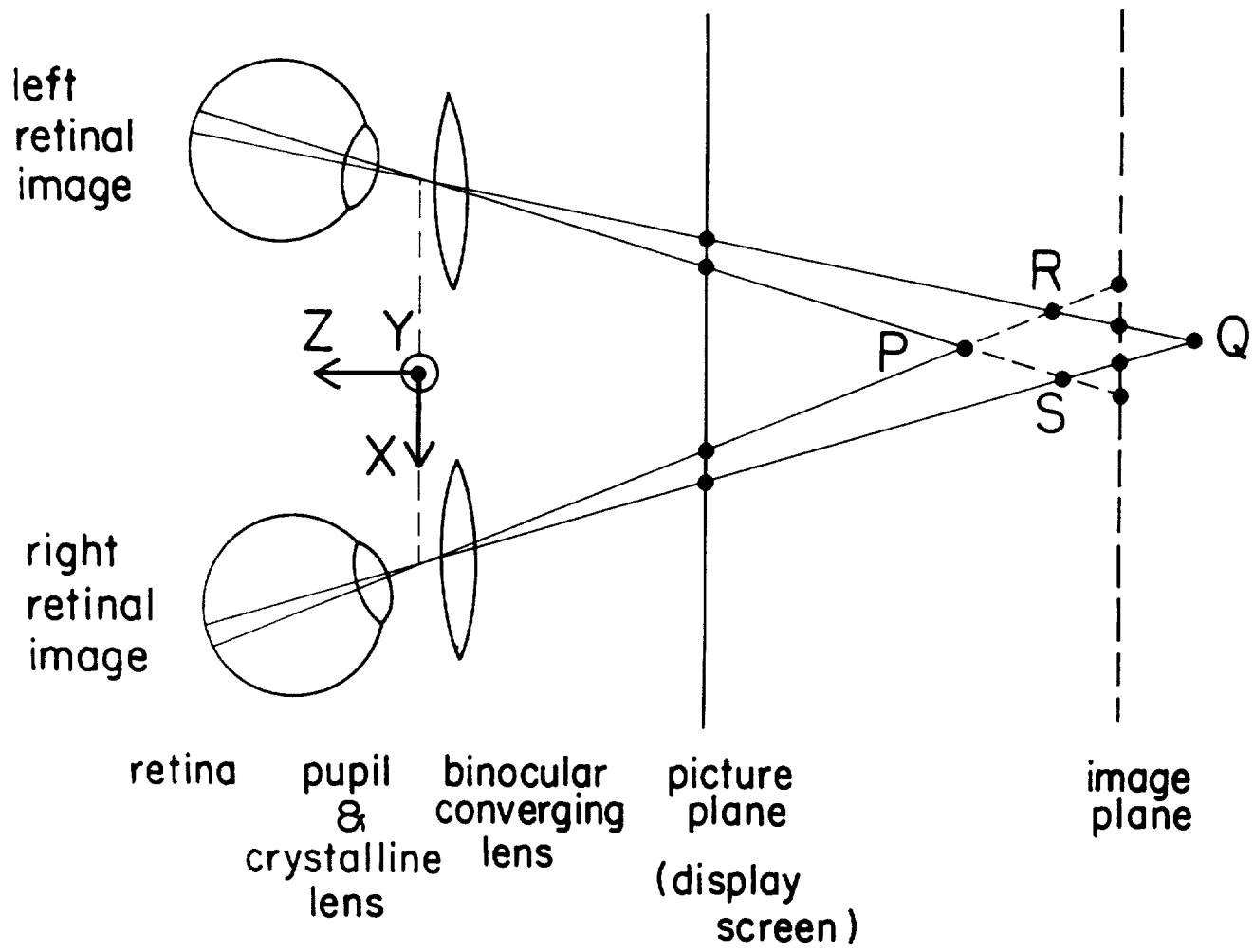
Figure 2.– Stereoscopic display.
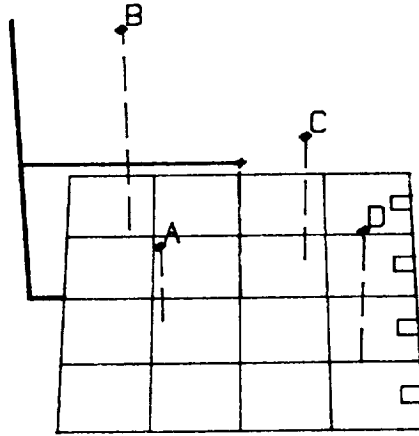
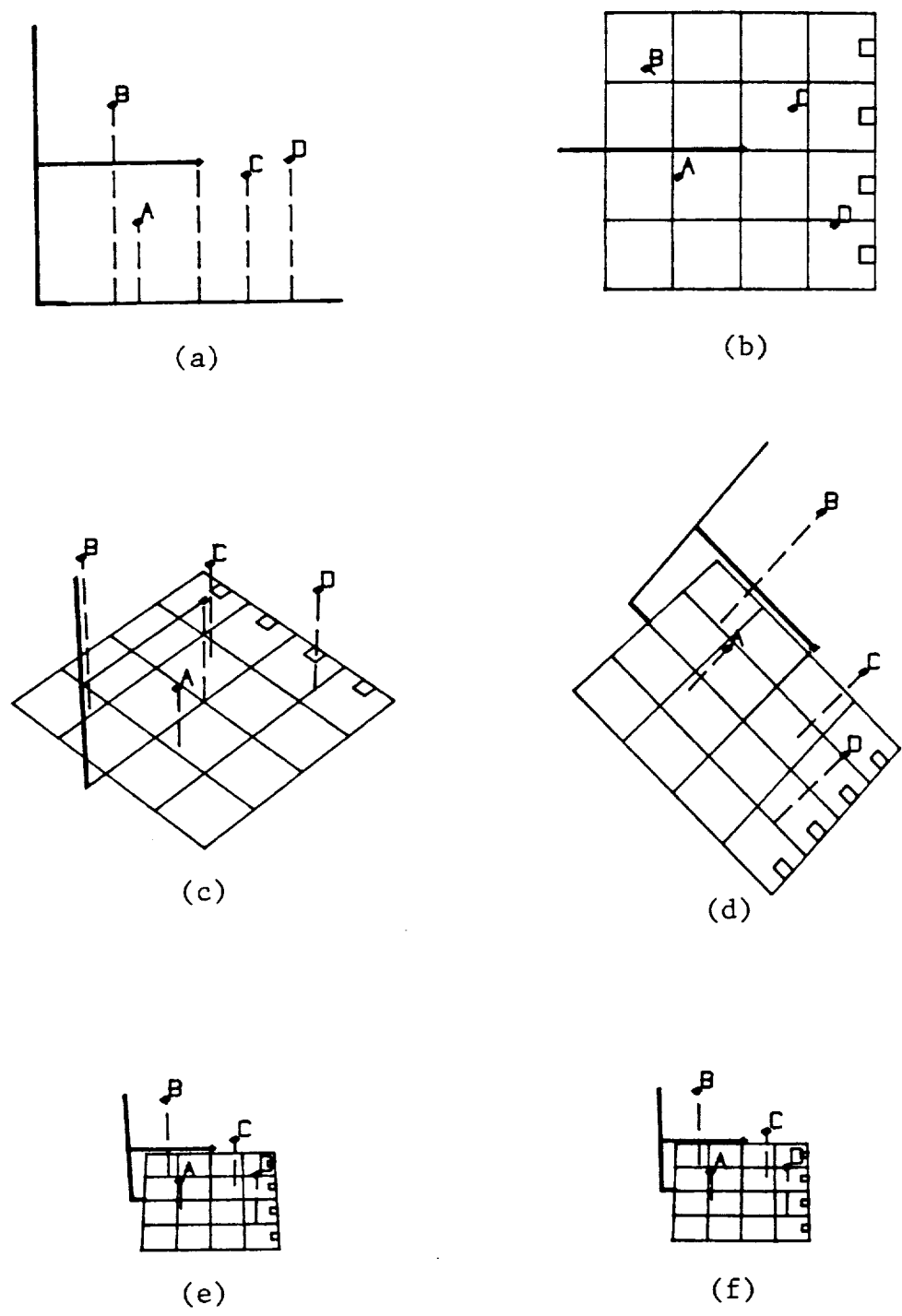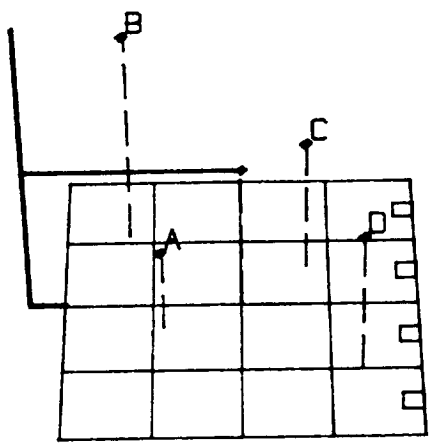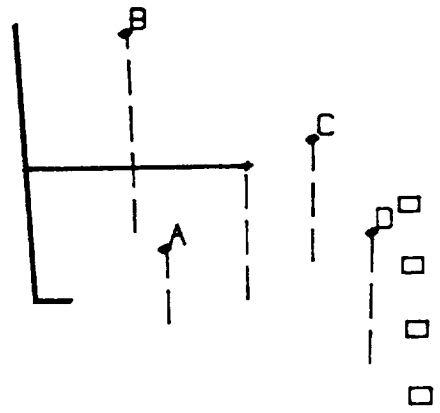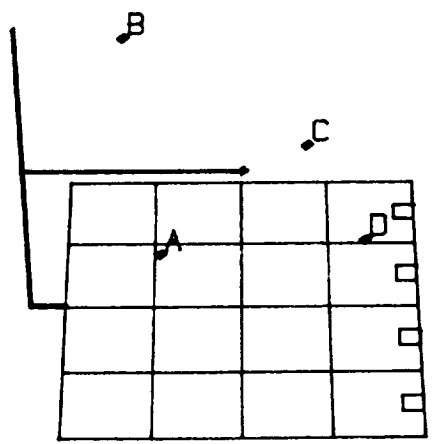Figure 3.– A monoscopic perspective presentation using nominal perspective parameters.

Figure 4.– Examples of various monoscopic perspective presentations with (a) an extreme 0° elevation angle, (b) the other extreme –90° elevation angle, (c) 45° azimuth angle, (d) 45° roll angle, (e) fov angle doubled to 24°, and (f) object distance doubled to 80 cm. A 5-line-by-5-line horizontal grid and vertical reference lines are presented.
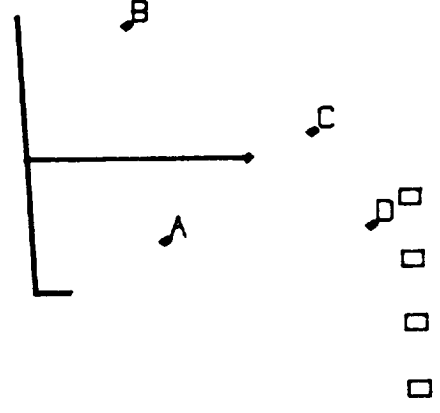
Figure 5.– Monoscopic presentations under four visual-enhancement conditions: (a) GL (presence of both grid and reference line), (b) L (reference line only), (c) G (grid only), and (d) O (neither).
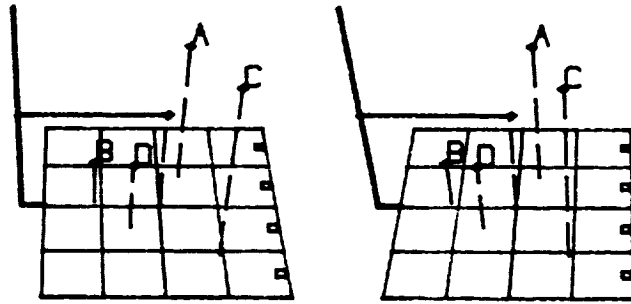
Figure 6.– An example of a stereoscopic presentation.
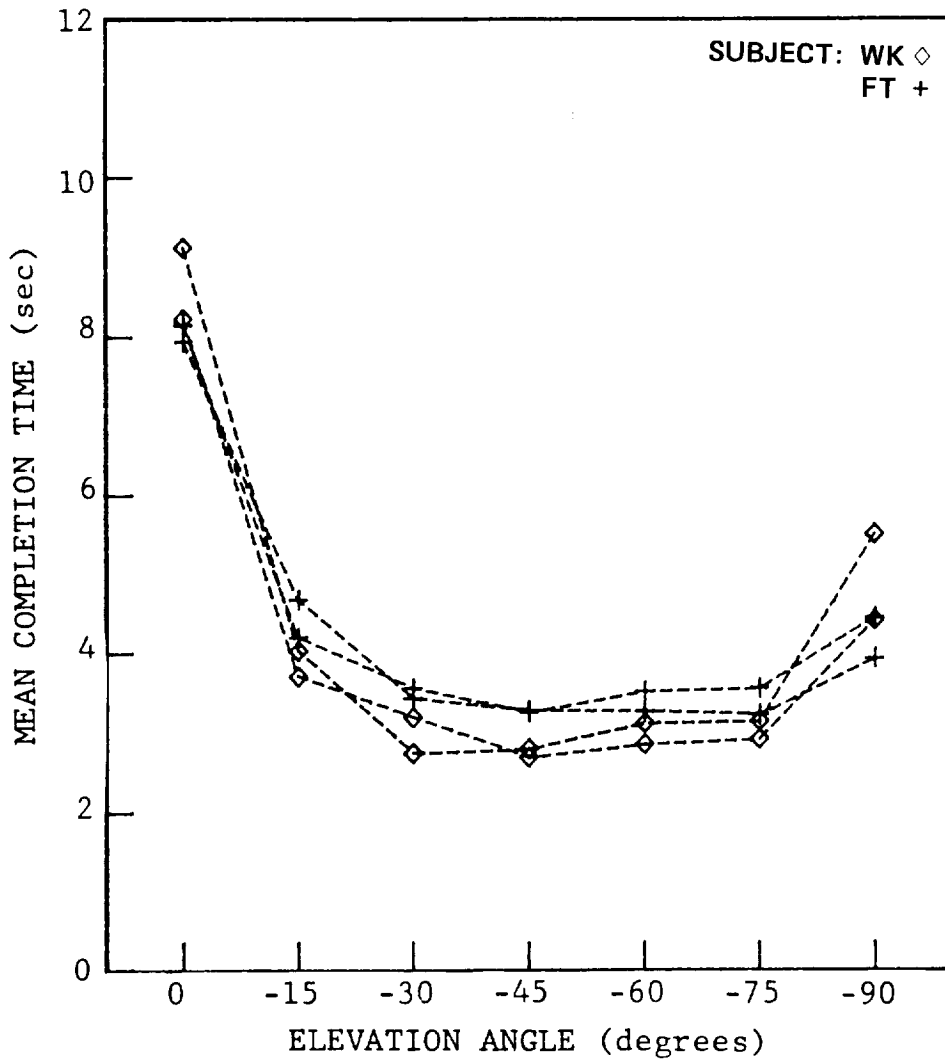
(a) Mean completion time as a function of elevation.

Figure 7.– Perspective parameter experiment. Three-axis pick-and-place performance with various monoscopic perspective displays.
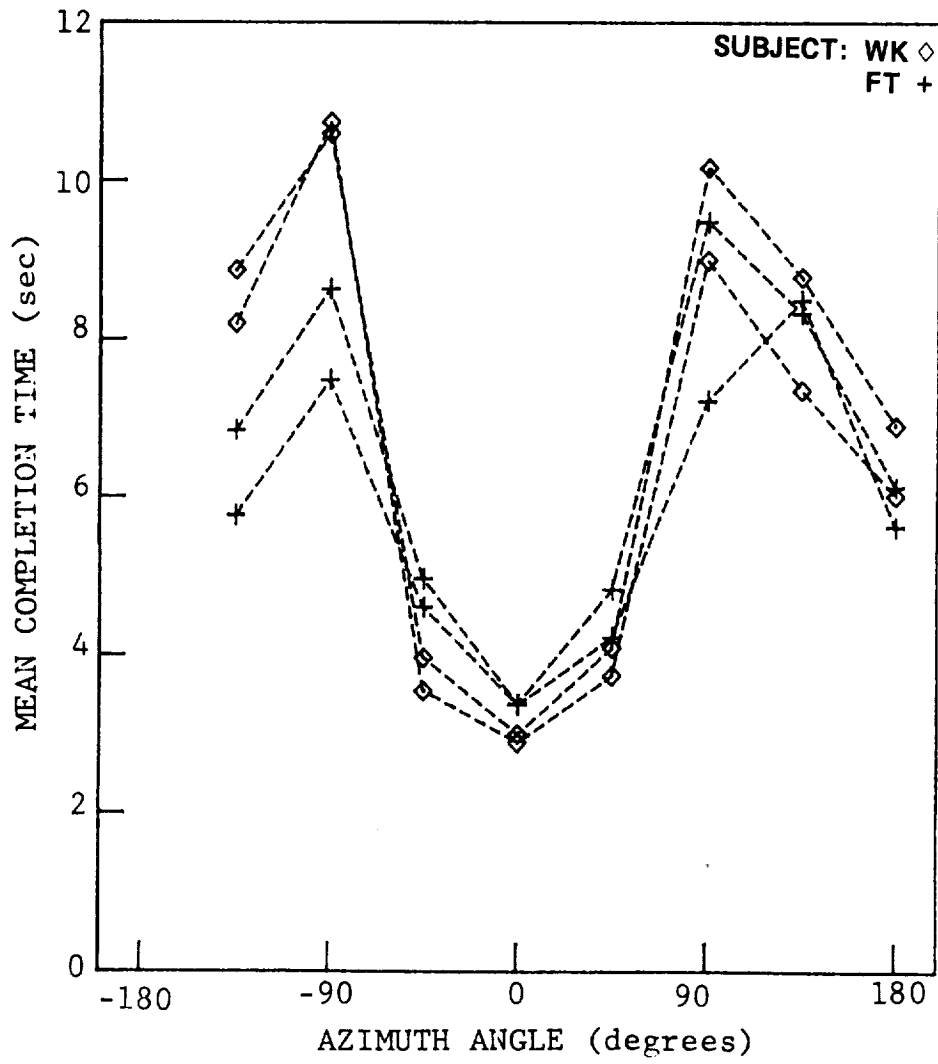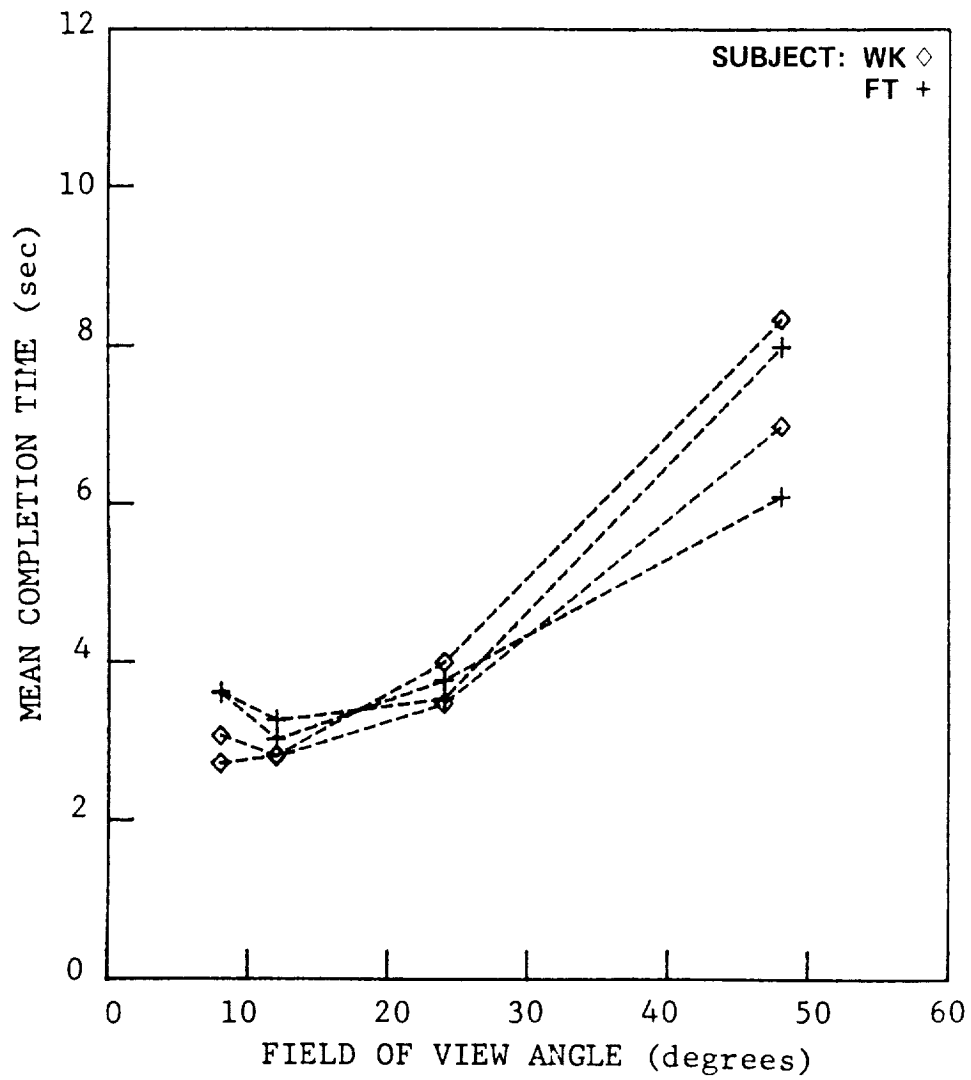
(b) Mean completion time as a function of azimuth.
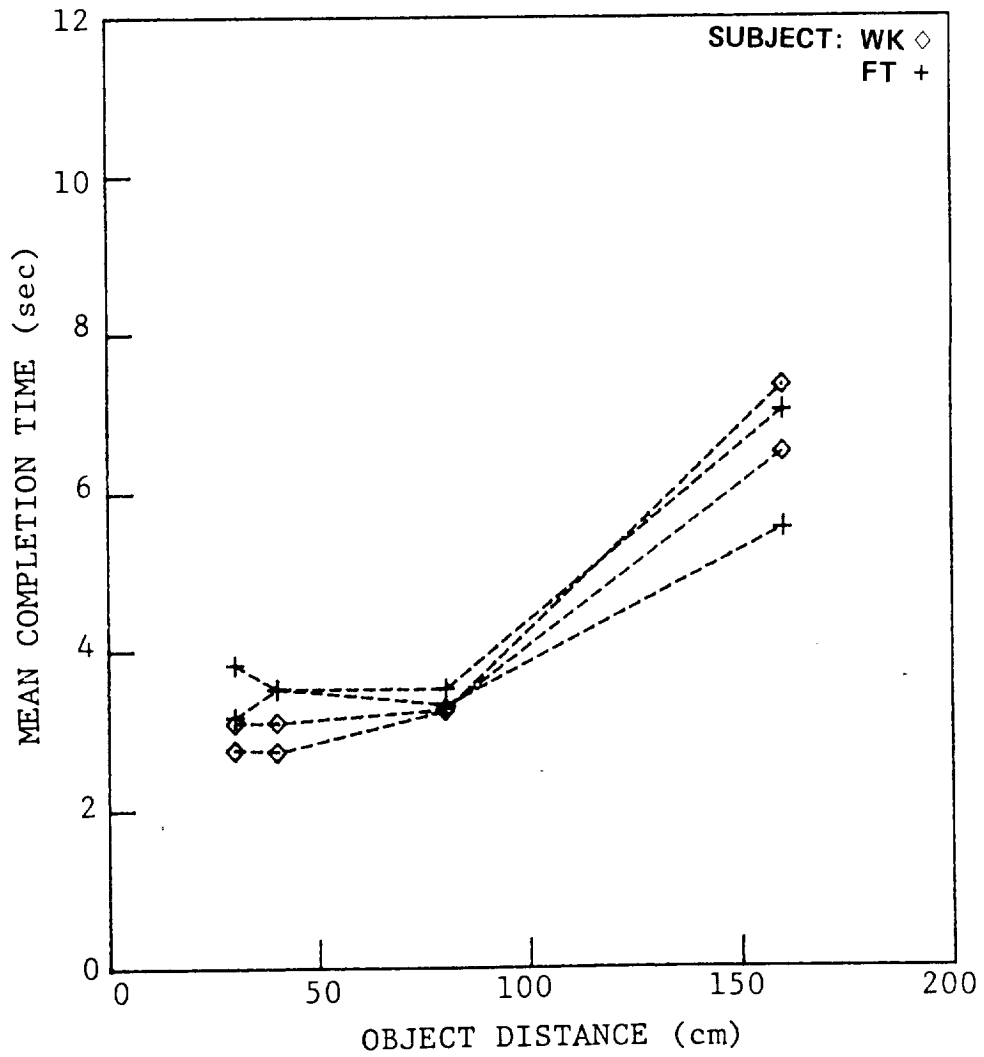
Figure 7.– Continued.
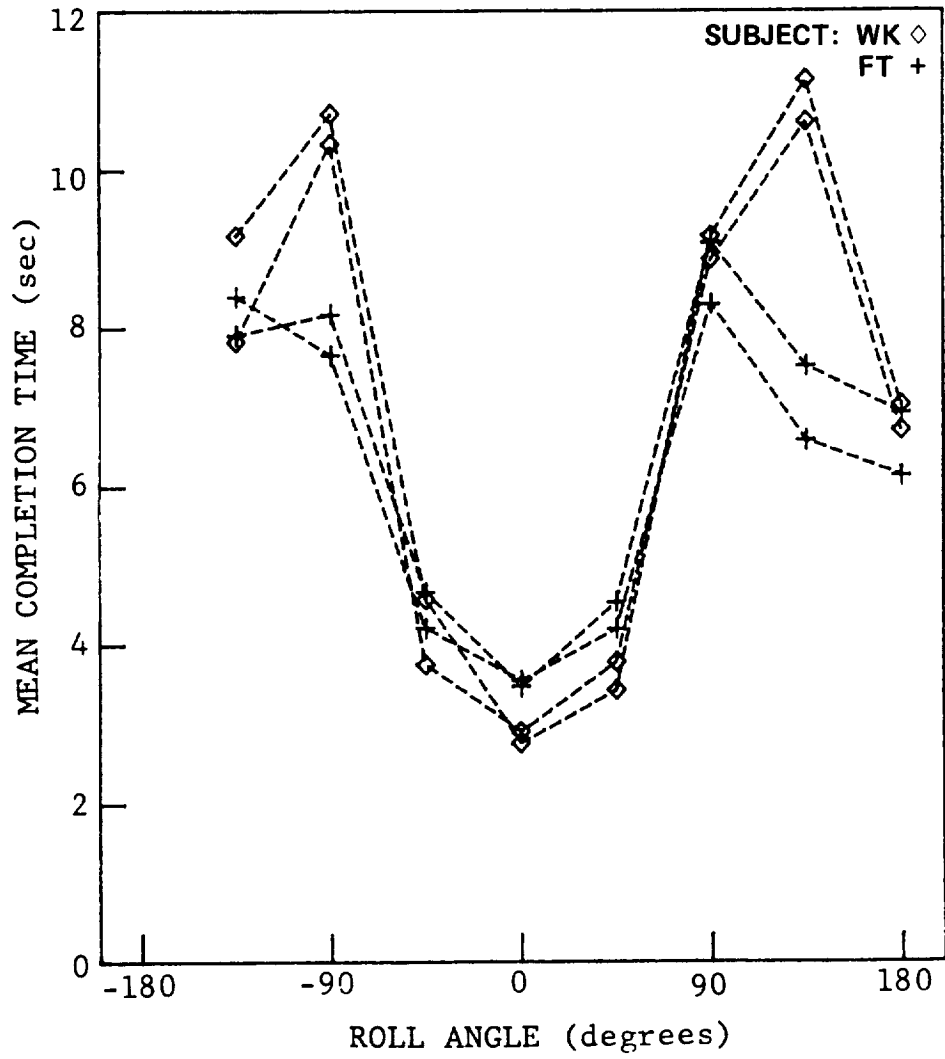
(c) Mean completion time as a function of roll.

Figure 7.– Continued.

(d) Mean completion time as a function of fov angle.

Figure 7.– Continued.

(e) Mean completion time as a function of object distance.
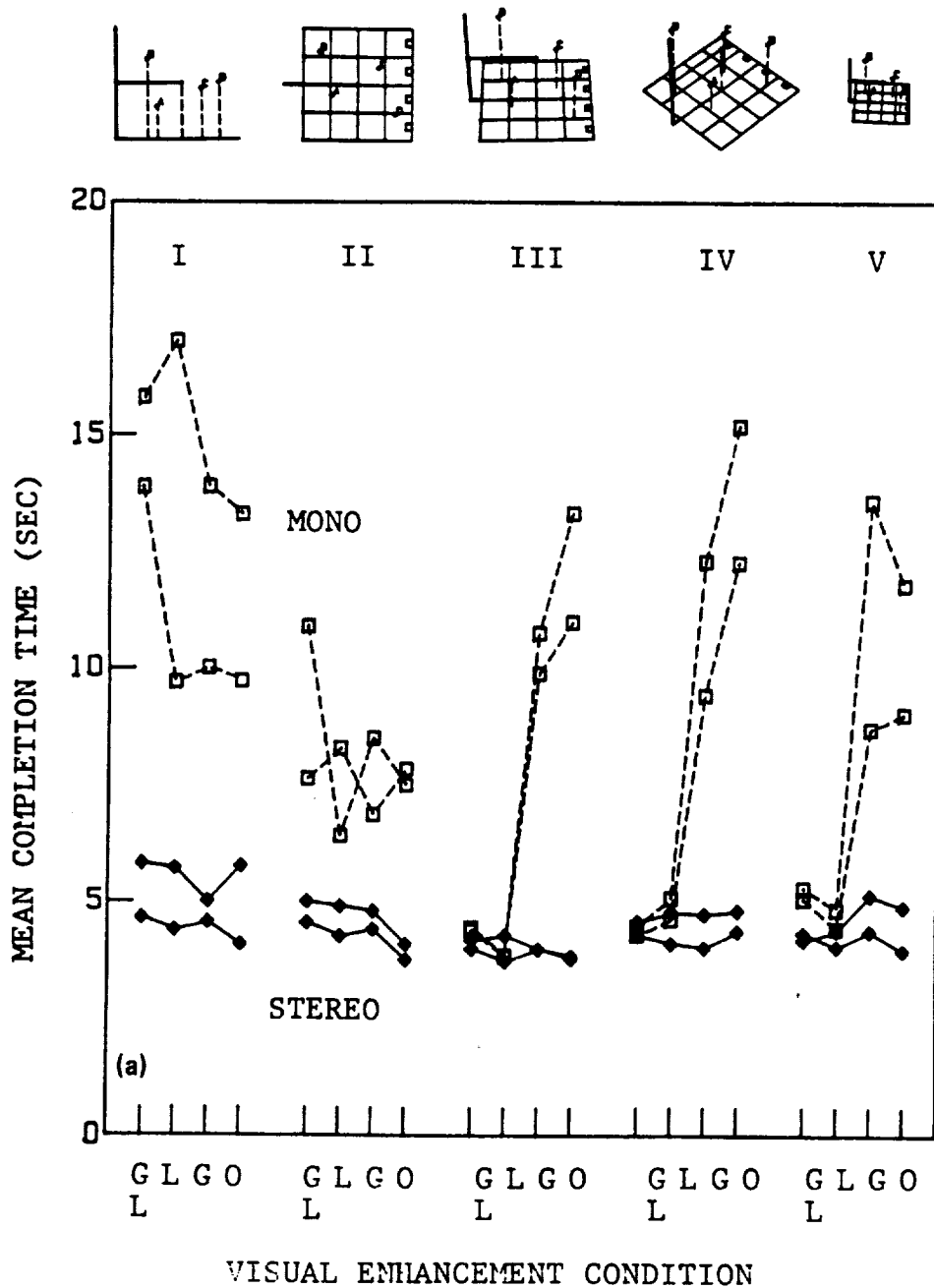
Figure 7.– Concluded.

Figure 8.– Visual-enhancement experiment. Three-axis pick-and-place performance for four visual-enhancement conditions at each of five different perspective parameter conditions with both monoscopic display and stereoscopic display. The monoscopic presentations for the five perspective parameter conditions are shown above the plot (a). Four visual-enhancement conditions are GL (presence of both grid and reference line), L (reference line only), G (grid only), and O (neither). Subjects: (a) WK, (b) MT. Two runs for each subject. In plot (b), confidence intervals at the 95% level are shown about the means.
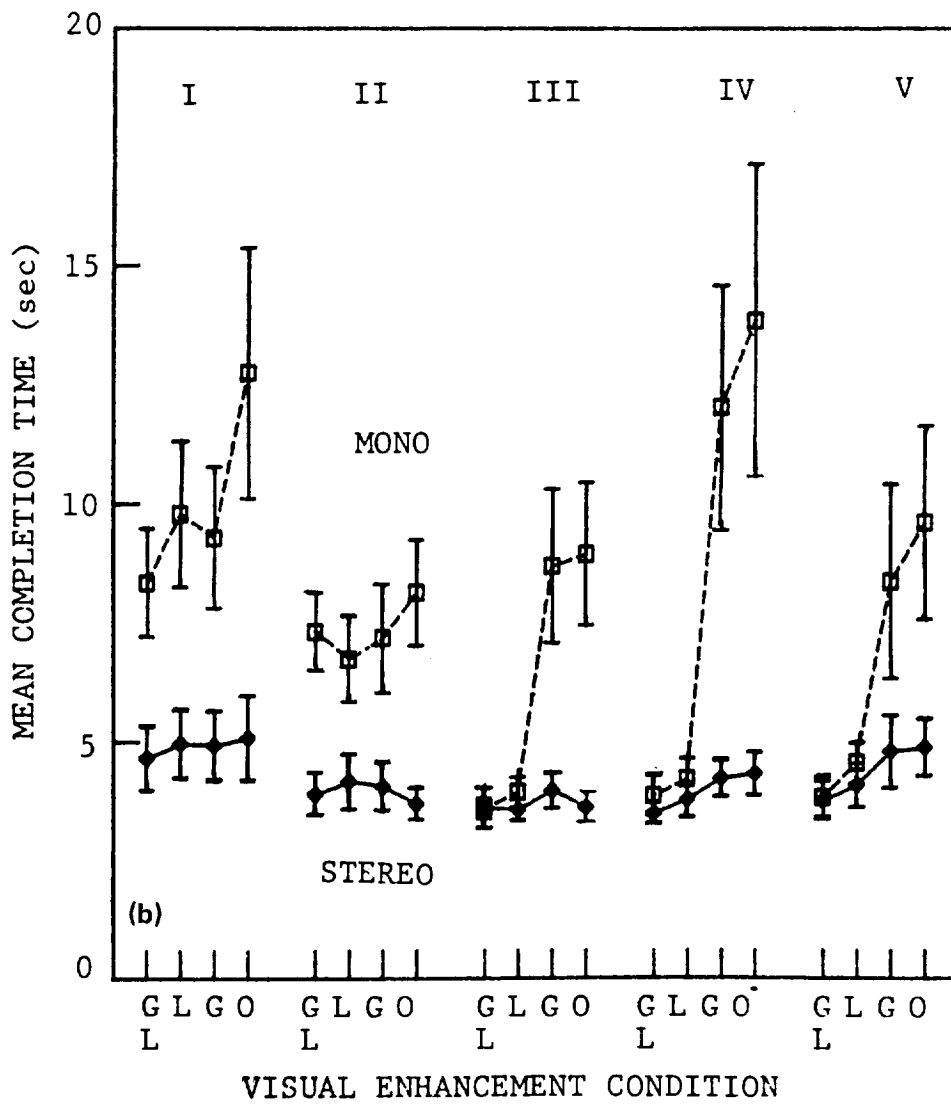
Figure 8.– Concluded.