

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Industrial and Management Systems Engineering
Faculty Publications

Industrial and Management Systems Engineering

2000

Design of Visual Displays for Teleoperation

S. H. Park

Hannam University

Jeffrey C. Woldstad

University of Nebraska - Lincoln, jwoldstad2@unl.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/imsefacpub>



Part of the [Operations Research, Systems Engineering and Industrial Engineering Commons](#)

Park, S. H. and Woldstad, Jeffrey C., "Design of Visual Displays for Teleoperation" (2000). *Industrial and Management Systems Engineering Faculty Publications*. 9.

<http://digitalcommons.unl.edu/imsefacpub/9>

This Article is brought to you for free and open access by the Industrial and Management Systems Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Industrial and Management Systems Engineering Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Design of Visual Displays for Teleoperation

S. H. Park* and J. C. Woldstad†

* School of Industrial and Mechanical Engineering, Hannam University, 133 Ojeong-dong, Taeduk-gu, Taejon 306-791, South Korea

† Department of Industrial Engineering, Texas Tech University, Lubbock, TX 79409-3061, USA

1 Introduction

Human exposures to hostile work environments can be reduced when tele-operated systems are used to remotely control manipulators in these environments. Tele-operation systems were initially developed by the nuclear industry but are currently applied in areas including surgery, construction, mining, warehousing, firefighting, undersea exploration, military operations, and space. The role of human operators in tele-operated systems varies with the level-of-automation of the system. Some systems require only supervisory control from the operator, while many others require direct manual manipulation through a controller. However, even for highly automated systems, human operators play an important role as a backup when the system fails.

Telepresence can be achieved through sensing appropriate information about the tele-operation task environment, and providing this information to the human operator at the remote site (Sheridan 1992). A great deal of human engineering information concerning the design of workstations for telerobotic systems is closely related to meeting telepresence requirements. In this regard, human depth perception (as a third dimension) based on the information provided by the two-dimensional (2D) surface of a video display terminal (VDT) very much determines the quality of a tele-operation system.

2 Depth Perception

A visual display has a basic limitation in that the surface of the display screen is two-dimensional. However, when designing a visual display for tele-operation tasks, it is necessary to represent depth or distance as a third dimension along the line of sight. Depth and/or distance perception are achieved through the combination of several depth cues. The term, cues, has been utilized to formalize the specification of stimulus conditions for space perception (Carr 1935).

In the study of perception, an object and its properties in the world are called distal stimuli. They only stimulate our nervous system by patterns of energy (e.g. light energy, sound pressure wave). We call the patterns

of energy that reach and affect our sense organs proximal stimuli. Therefore, a depth cue is a pattern of proximal stimulation that contains information about the spatial location of distal objects. Depth cues can be classified into two types: monocular and binocular. Monocular depth cues require the activity of a single eye; binocular depth cues require the use of two eyes.

2.1 Monocular Depth Cues

2.1.1 Relative size

The same-sized objects produce smaller tracings when they are at greater distances: this is called the cue of relative size, or size perspective. By comparing the apparent size of a distant object with that of a similar, much closer object, the relative distance of the distant object can be approximated.

2.1.2 Interposition

A nearer (overlapping) object interrupts the outline of a farther (overlapped) object. This is an effective depth cue, but it can only indicate which object is in front, not the distance separating them.

2.1.3 Linear perspective

If the size of a distal object is fixed, the visual angle will be inversely proportional to the distance from the object: this is called linear perspective. A constant distance between points subtends a smaller and smaller angle at the eyes as the points withdraw from the eye. For example, telephone wires appear to approach each other (i.e. the retinal images of the lines converge) as the distance from the eyes increases although they are parallel. Therefore, converging lines are a cue that they are parallel and receding in depth.

2.1.4 Monocular movement parallax

When a subject's eye moves with respect to the environment, or vice versa, there exists a differential angular velocity between the line of sight to an object (fixated) and the line of sight to any other objects. For example, far objects move with the direction of movement, while near objects move against the direction of movement (Gra-

ham 1965). By observing the amount and relative direction that a given image moves on the retina, its distance can be approximated.

2.1.5 Familiar size

We can use objects known size to infer relative depth. For example, we know a man is taller than a boy. However, if they produce the same size of retinal image, then we deduce the man is located farther away than the boy. This is often a weak or ineffective cue (Hochberg 1978).

2.1.6 Light and shadow distribution

Shadows provide some information about the orientation of the objects and their three-dimensional (3D) shapes. Objects may appear to lie at different distances and have different dimensions as combinations of shadow and highlight change (Graham 1965). If objects have a light source from one direction, they will have shadows unique to their shape and orientation.

2.1.7 Gradient of texture-density

A gradient is the rate of some measured property changing over a continuous, extended stimulus. The surface of most objects is likely to be covered with a reasonably uniform texture or pattern. When looking straight ahead at a textured surface, the gradient of texture-density is zero; as the slant increases, the gradient increases. The gradient of texture-density can provide precise and relatively unambiguous information about the distances and sizes of surfaces and objects in the world (Hochberg 1978).

2.2 Binocular Depth Cues

2.2.1 Convergence

The eyes are capable of convergence, in which both eyes turn inward toward the medial plane. A large convergence corresponds to near objects and a slight convergence corresponds to far objects. The ocular muscles control the angle of convergence. The brain receives proprioceptive messages from the ocular muscles about the degree of convergence (Wickens 1992). By analyzing the information received, the brain can approximate the angle of convergence. In this way, convergence may serve as a depth cue. A large convergence may lead to the response "far-off," while a slight convergence may lead to the response "nearer."

2.2.2 Stereopsis

The retinal image of a distal object in the right eye is different from that of the same distal object in the left eye. The difference in the retinal image plays a great role in spatial discrimination. In the 3D world, the view each eye receives is somewhat different because the two eyes see the object from slightly different positions. Differences in these views give two possible depth cues: double image (each eye contributes different image of far object when viewing near, and vice versa) and binocular disparity

(Hochberg 1978). The disparity is the difference between where a target falls on the right eye and the left eye.

The disparity can generate a powerful depth cue. This depth cue can be obtained by taking two photographs of a scene (a stereoscopic picture pair, stereogram), one from the position of each eye (65 mm apart average), and presenting each picture to its appropriate eye, and then viewing such photographs with special devices called stereoscopes. The stereoscope is composed of two converging lenses and a supporting frame that simply separate right and left views.

3 Visual Displays

3.1 Principles of Display Design

3.1.1 Visual momentum

Visual momentum refers to the visual landmarks that film editors generate to reduce visual inconsistency among several scenes when editing a film. The concept of visual momentum can be applied to integrate information of one display into the other displays among multiple displays. For example, a visual interface for a telerobotic system can be composed of multiple 2D displays having different reference frames (e.g. a plan view and side views). In such a case, human operators must integrate information across multiple displays in order to control the robots arm. The concept of visual momentum can be applied by presenting a line originating from the face of the gripper of the robot. The line always aims the direction of the gripper, and therefore visual momentum is provided across the displays having different reference frames.

3.1.2 Object integrality and the principle of proximity compatibility

An object (or integrated) display is a display which integrates multi-dimensional information into one object. For example, the conventional attitude indicator in aviation represents the aircraft's pitch and roll information (2D) in one display. It should be noted that integrating two dimensions into an object will help performance if the information associated with those two dimensions is relevant, but decrease performance if the information is irrelevant.

The integrality of displayed was further expanded by Wickens and his colleagues, introducing the proximity compatibility principle – "To the extent that information sources must be integrated, there will be a benefit to presenting those dimensions in an integrated format (high mental proximity). In contrast, to the extent that information must be treated separately (low mental proximity), the benefits of object displays will be reduced" (Wickens 1992, p. 98). In summary, close proximity can be used to increase performance if parallel processing is guaranteed in the object display. However, it may cause problems and decrease performance if tasks require focused attention on a particular dimension in the display.

3.1.3 Working memory

New information is stored temporarily at a working memory until it is used or stored in a long-term memory. It is well known and accepted that human working memory is limited in size and time. The concept of chunk has been employed to define the working memory limit, where a chunk can be any set of information associated with long-term memory. Based on the concept of chunk, human working memory is limited by 7 ± 2 chunks of information (Miller 1956). When designing visual displays, it is desired to avoid exceeding the limit of 7 ± 2 chunks of information at the level of working memory.

3.1.4 Mental rotation

The time required to compare two visual images increases linearly as the angular disparity between the two images increases. For example, using multiple displays having different reference frames, human operators must mentally rotate a standard alignment in congruence with the alignments of other views. The mental rotation requirements of a task are more demanding if human operators must frequently scan different views.

3.2 Dimensionality

3.2.1 2D displays

A 2D display is an orthogonal-view that provides spatial information about two dimensions. The only depth

cue that can be provided using a 2D display is interposition, which only indicates which object is in front of the others, not the distance separating them. A multiple 2D display can be composed of a plan-view, a front-view, and/or side-views (i.e. all 2D orthogonal views). In the multiple 2D display, 3D spatial information is provided by the combination of at least two views; for example, the plan-view provides spatial information about two dimensions while the side-view or the front-view provides spatial information about a third dimension, such as altitude.

A 2D display does not provide any depth information in a third dimension. However, it provides unambiguous depth information in two dimensions and removes ambiguous pictorial depth cues. The 2D display, if provided with other 2D displays (i.e. forming the multiple 2D display), can provide unambiguous depth information in 3D. As compared with the perspective display, however, one of the main disadvantages of the multiple 2D display may be additional scanning and integrating effort among separate displays.

3.2.2 3D displays

A 3D perspective display can be achieved by projecting an object onto the view plane (projection plane) and then mapping the view plan onto the display screen. The 3D perspective display provides more natural spatial information about 3D environments by creating the appearance of depth along an observer's line of sight, as

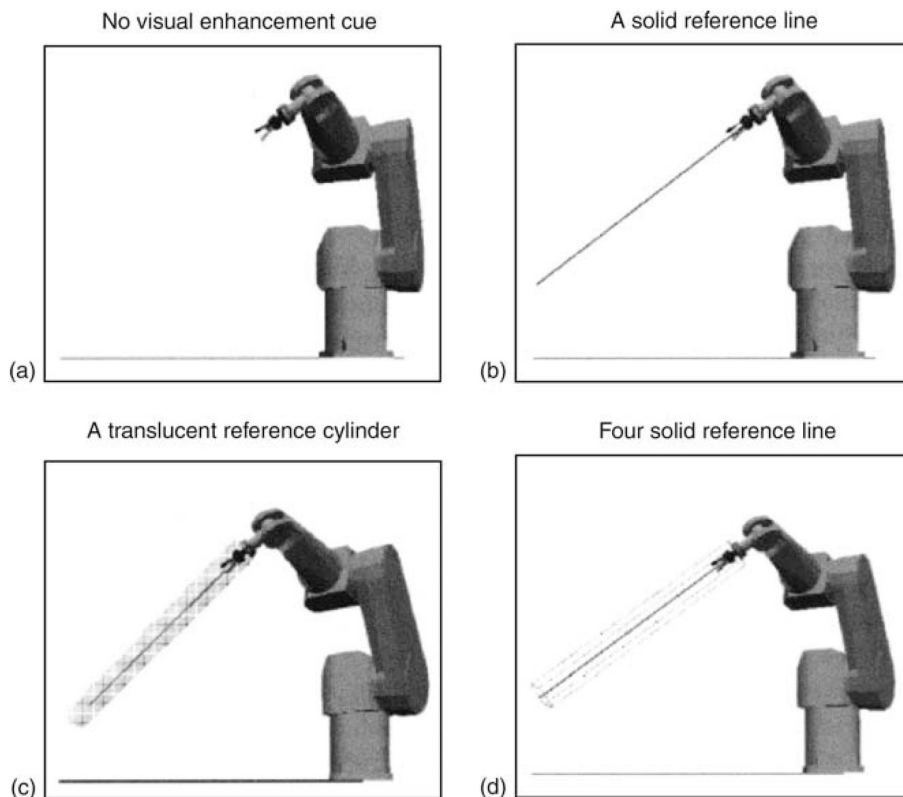


Figure 1. Level of visual enhancement cues for a telerobotic task. (a) No visual enhancement cue; (b) a solid reference line; (c) a translucent reference cylinder; (d) four solid reference line.

compared with the 2D display. Most of the monocular depth cues can be achieved through the 3D perspective display. However, the monocular depth cues are inherently ambiguous. Based on the human stereoscopic vision capability of fusing two retinal images into one image, a stereoscopic display can be generated by adding a depth cue (called stereopsis) to the perspective display.

3.2.3 2D versus 3D displays

It is hard to conclude that 3D displays are superior to 2D displays, or vice versa. A multiple 2D display can improve tele-operator performance as compared with the perspective display if the task requires frequent use of focused attention on a particular dimension. In general, the choice of display format is dependent on the specific task being performed (e.g. gross positioning tasks requiring global attention or dexterous tasks requiring focused attention).

3.3 Use of Visual Aids

Using computer generated visual enhancement cues can enhance the communicative purpose of a visual display. Figure 1 shows several visual enhancement cues that extend from the point of motion (gripper of the robot) and allow the user to orient the manipulator in space. In Figure 1b a solid reference line is placed orthogonal to the face of the gripper of the robot allowing the user to orient and aim the gripper and objects attached to it. The reference line is a useful cue that has been found to help operators orient and position the gripper when performing a telerobotic task. However, it does not provide appropriate depth information. A translucent reference cylinder or four solid lines (Figure 1c and d) are better enhancements. These cues have volumes in 3D space; therefore, they not only aim the orientation of the gripper but also provide depth information by overlapping other objects in the work environment.

4 Control Factors

4.1 Display–Control Correspondence

One-to-one correspondence between display and control devices, ensuring that a control movement and a resulting change in the display are in the same relative location, should be provided as far as possible (Sheridan 1992). In general, lateral displacements $<15^\circ$ can be adapted to by humans almost completely. However, the ability to adapt to mismatches is decreased for angles $>15^\circ$. Any type of inversed feedback is most disruptive.

4.2 Force Feedback

Tele-operation performance can be improved when force feedback is provided. The human body's joint,

muscle, and tendon receptors sense the net reaction force and torque acting on the hand (i.e. resolved force sensing as defined by Sheridan 1992). Force feedback to human operators can be achieved by measuring the force and torque acting on the tele-operator, and then driving motors on the controller which provide the hand of the human operators with the same force and torque. However, there is little information about the effectiveness of this method. As an alternative, display of force feedback to the operator can also be achieved through the use of computer graphic force-torque display.

4.3 Teletouch

Teletouch refers to the remote sensing of differential forces acting on the skin in time and space. Teletouch has been found to support depth perception, provide detailed information about surfaces of objects contacted by grippers, and provide information about surfaces visually occluded by other objects. A major problem associated with teletouch is the difficulty associated with displaying artificially sensed pressure patterns to the skin on the hand, when the hand is at the same time operating the control stick (Sheridan 1992). One useful way to present teletouch information to the operator is the use of visual display method (i.e. visually representing tactile data).

5 Recommendations

- Select display dimensionality based on the given task being performed. Consider a multiple 2D display format if the task requires frequent use of focused attention.
- For 3D perspective displays, provide visual enhancement cues to aid depth perception.
- As a visual enhancement cue, a single line is not sufficient enough to aid depth perception. Design a visual enhancement cue which itself has a volume in 3D space.
- Consider task difficulty when selecting the level of visual enhancement.
- For 3D perspective displays, ensure that the control movement and the resulting change in the display are in the same relative location. Maintain lateral displacements to $<15^\circ$.
- Provide the hand of the human operator with the same force of the manipulator (force feedback). Consider a display of force feedback as an alternative.

References

- Carr, H.A., 1935, *An Introduction to Space Perception* (New York: Longmans, Green).

- Graham, C.H., 1965, Visual space perception. In Graham, C.H. (ed.) *Vision and Visual Perception* (New York: Wiley), pp. 504-47.
- Hochberg J.E., 1978, *Perception* (2nd ed.) (Englewood Cliffs: Prentice-Hall).
- Miller, G.A., 1956, The magical number seven, plus or minus two: some limitations on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Sheridan, T.B., 1992, *Telerobotics, Automation, and Human Supervisory Control* (Cambridge, MA: MIT Press).
- Wickens, C.D., 1992, *Engineering Psychology and Human Performance* (2nd ed.) (New York: HarperCollins).