OPERATOR VISION AIDS FOR SPACE
TELEOPERATION ASSEMBLY AND SERVICING

Thurston L. Brooks
Ilhan Ince
Greg Lee
STX Robotics
4400 Forbes Blvd
Lanham, MD 20706

ABSTRACT

This paper investigates concepts for visual operator aids required for effective telerobotic control. Operator visual aids, as defined here, mean any operational enhancement that improves man-machine control through the visual system. These concepts were derived as part of a study of vision issues for space teleoperation. Extensive literature on teleoperation, robotics, and human factors was surveyed to definitively specify appropriate requirements. This paper presents these visual aids in three general categories of camera/lighting functions, display enhancements, and operator cues.

CAMERA AND LIGHTING FUNCTIONS

Camera and lighting functions were found to take a significant portion of an operators time. Operator aids could make a significant performance improvement in this area. Concepts discussed are: (1) automatic end effector or task tracking; (2) novel camera designs; (3) computer-generated virtual camera views; (4) computer assisted camera/lighting placement; and (5) voice control.

Automatic Tracking of End Effector or Task

The function of any automatic tracking aid is to relieve the operator of having to perform additional tasks associated with camera and lighting movement.

Uhrich [1978] reported the first known attempt to track the end effector automatically while performing a task. The camera pan and tilt motions tracked the end effector in only two manipulator joints of motion, but for limited areas of operation the system kept the end effector centered in the camera view. This simple study showed a subjective improvement in operation (actual experiments were not performed). Uhrich reported that although disorientation problems were anticipated, they were surprised to find that the operators could utilize end effector tracking without disorientation.

However, Brooks [1978, Bejczy 1980] found that automatic tracking of end effector motion could create situations in which information normally available to the operator was masked, resulting in operational problems. Specifically, auto-tracking across backgrounds without features or texture gave the operator the impression that the manipulator was not moving, because the end effector remained at the screen center without apparent motion. Under master-slave manipulation, subjects became frustrated due to differences of stimulus-response compatibility (the operator’s hand was moving but without apparent effect on the end effector). Additionally, under a non-analogic control such as rate, the

OPERATOR VISION AIDS

This paper outlines a number of operator aids that can improve visual performance during teleoperation. The techniques discussed here are specifically related to improving the process of human vision under remote conditions. We have not dealt with the larger field of graphical display of sensory data (such as force/torque, proximity, etc.) except as this data directly improves visual performance. This paper presents these visual aids in three general categories of camera/lighting functions, display enhancements, and operator cues.
operator did not even realize motion was occurring at all! Brooks' auto-tracking scheme is shown in Figure 1.

Another potentially useful operator aid would be to automatically keep the whole manipulator in the field-of-view so that the kinematic configuration could be observed. This concept will be very important as generalized hand controllers and redundant degree-of-freedom arms become more prevalent in teleoperation. Unfortunately, in most task environments, it simply is not possible to simultaneously view the entire arm, and therefore, a graphical 3D representation is believed to be a more practical solution.

In addition to proprioceptive tracking, auto-tracking can also be implemented through machine vision to allow the task motion to be followed [Bejczy 1980]. Using image processing to determine the task/robot relationship, the system could automatically aim the cameras at a point of interest and maintain a fixed relationship of the end effector with the task to relieve the operator of these burdens [Brooks 1979]. Figure 2 illustrates a simple visual method to determine the task-to-telerobot relationship using "labels" that was implemented at JPL [Brooks 1980, 1982 and Bejczy 1980]. Task/object tracking can also involve more sophisticated machine processing such as template-matching or object modeling techniques. Visual servoing and auto-tracking have also been investigated for the OMV, Shuttle umbilical mating and satellite attitude determination [McAnulty 1985, Harrison 1986, Russell 1986, Feddema 1989, KSC 1990 a&b].

New Camera Designs
There are endless possibilities for new and novel camera designs that would improve teleoperation and aid the operator. Two are of interest because they solve recognized problems: glare and acuity.

The first problem is one of specular reflections and glare in the space environment. A specialized camera could help solve this problem through the use of liquid crystal technology. In essence, a CCD camera could have a liquid crystal shade (LCS) with identical resolution placed over the CCD array. Through image processing the system could determine the pixels receiving too much light and the LCS in those areas could suppress or block the offending rays (see Figure 3).

The development of a camera with a foveal view of very high resolution and a surrounding peripheral view of low resolution is another novel camera concept first suggest by Carl Ruoff at JPL.1 Designed properly, a single CCD chip could pack a 1-2° foveal center and greater than a 100° peripheral view on one single NTSC channel (Figure 4). A design like this would permit high resolution and peripheral view without a bandwidth increase.

1 Discussion between T. Brooks and Carl Ruoff at JPL in 1979.
Figure 4: A CCD chip with a high-resolution fovea and a low-resolution peripheral could permit the use of a low NTSC bandwidth with high-acuity.

Preview Scenes
Synthetic camera image generation based on CAD models of the task environment could permit the astronaut to determine if a particular view would meet task performance requirements before committing the resources. These preview scenes give the astronaut "what-if" capability to explore alternative solutions.

Computer Assisted Camera/Lighting Placement
A graphical representation of current state of cameras and lighting could be a useful overview display. A display such as the one shown in Figure 5 would permit the astronaut to quickly determine which camera(s) and light(s) would result in the optimal viewing conditions for a task. The display could present an overhead view illustrating light beams and camera field-of-view based on current or proposed zoom, focus, convergence, light intensity, etc. If a method for entering up-to-date orbit status was available, the system could also show Sun illumination, including reflected light. A perspective view showing expected illumination and visibility from a selected, or desired, camera could also be generated and displayed by simply selecting the desired camera vantage point (Figure 6).

Another option would be to have computer-aided camera/lighting optimization, wherein the computer generates a graphic suggesting the best camera and lighting position for a specified task location. With advanced computer aiding, the astronaut could request the system to optimize lighting and camera views for current orbit trajectories. With advanced computer aiding, the astronaut could simply "point-and-go", leaving the system to select the best camera, set focus, zoom, and convergence, and adjust light levels automatically. Other possibilities include: (1) having the camera system automatically maintain a fixed image size of an object-of-interest on the monitor, (2) allowing the astronaut to specify a desired depth of field, and (3) automatic contrast adjustment through optimization of lighting and camera adjustments.

As a final note of interest, robotic cameras have been used in the broadcast industry for a couple of years now, with reported benefits of smoother panning motions, allowing observers to more easily follow details and maintain perspective [Lehtinen 1990].

Voice Control
Voice control allows the operator to simultaneously maintain hand controller movements, observe CCTV monitors, and issue verbal commands. Befczy [1981] speculates that voice control does not disturb the operator's visual attention or manual control functions, and that it minimizes mental distractions. However, early experiments with voice control in the Manipulator Development Facility at JSC, although deemed successful, indirectly highlighted an important aspect of voice control — its discrete
nature. Voice control is a discrete system input rather than an analog one, and hence, it is best for discrete functions. Voice is quick and efficient for switching a camera to a monitor (e.g., "Camera B3 on Monitor 1"), or homing in on a known location (e.g., "Zoom in on ORU Handle"), but it is not as efficient for panning and tilting functions which results in a number of overshoots or slow motions (e.g., "Pan Left...Faster...Slow...Stop...Pan Right...Slow...Stop").

At NASA, a number of voice control systems have been developed and tested for controlling cameras and lighting. For example, at the Jet Propulsion Laboratories a teleoperator control station has been developed in which cameras and monitors can be selected, zoomed, focused, panned, and tilted by voice control. Johnson Space Center's voice system has been flight tested onboard the Shuttle with some success [Foley 1991]. The JSC voice control system allowed complete hands-free control of CCTV functions:

1) monitor selection
2) camera selection
3) pan, tilt, focus, iris, zoom, and scene track

DISPLAY ENHANCEMENTS

A number of operator display enhancements were discovered to be potentially useful: (1) zone displays, such as imminent collision or indexing limits; (2) predictive displays for temporal and spatial location; (3) stimulus-response reconciliation displays; (4) graphical display of depth cues such as 2D symbolic depth, virtual views, and perspective depth; and (5) view enhancements through image processing and symbolic representations.

Zone Displays

A zone display highlights areas of the visual field that place restrictions on the operator for one reason or another. Two of the more obvious possibilities for restricted access are areas where collisions are likely to occur and locations that the operator cannot reach due to limits of the arm geometry (i.e., workspace). Figure 7, for example, illustrates a workspace restriction display.

Another useful display is one that highlights imminent collisions of the telerobot that can occur with its environment, its carrier (e.g., the RMS), or with itself. It is preferable to prevent all undesirable collisions, but a complete lockout of all contact would render the teleoperator incapable of performing its mission. A display that could warn the astronaut of a collision before it occurs would give the astronaut the option of taking corrective action or proceeding if essential to the mission. An imminent collision display could operate in a number of different modes:

1) The offending object(s) could be "painted" yellow to indicate an impending collision and red to indicate that a collision had occurred.
2) Stick figures of imminent objects could be superimposed on the real-time video to indicate collision status [Crane 1986].
3) All possible collision points in the camera's view could be false colored (e.g., yellow) to indicate that a collision is possible. This is likely to be quite a "busy" display which could be more confusing than helpful.
4) A graphic simulation of the remote manipulator could be false colored in the general area where a collision is expected as shown in Figure 8.
5) A graphic overview of the task environment and remote system could highlight the highest likelihood collision point based on calculated approach velocity and separation distances. This collision display highlights the highest probability collision, not necessarily the closest points. For example, if the arm is moving away from an object it would be impossible to collide with the object even if one were nearly touching it.

Another potentially useful display would highlight the instantaneous workspace of the slave manipulator due to the limits of the master controller's motion (i.e., limits due to indexing of the master when there is not a one-to-one correspondence between master and slave). Indexing is a control strategy in which the operator can re-reference the master hand controller relative to the slave to maintain the hand controller and its movements within an optimum volume. Typically, indexing is done to allow the operator to use the hand controller in a small volume of space while controlling a remote manipulator in a larger volume without resorting to scaling. Unfortunately, once the slave is in position and the hand controller is re-referenced, there is no guarantee that the current location has sufficient freedom to permit the task to be performed without re-indexing in the middle of the task. This can be a nuisance, if not a detriment, if the task requires the operator to index just to move the manipulator a few additional millimeters. A solution to this problem would be to display the achievable slave workspace as a function of the indexed master movements as shown in Figure 9. An indexed-workspace display such as shown in Figure 9 could assist the operator in placing the slave in an optimal position for completing a task with a minimum of re-referencing actions.

Predictive Displays
Predictive displays, as used here, are displays that estimate where objects are or will be. When we speak of where objects are (or should be) at this instant, we use the term spatial predictive displays. When we speak of where objects will be in the future, we use the term temporal predictive displays.

Figure 10 is an example of a spatially predictive display in which a CAD model is superimposed on the actual video image for the purpose of highlighting the location of components when they cannot be seen due to obscuration or insufficient lighting.

Spatially predictive displays can also be used to indicate where the telerobot "thinks" an object is, versus where its true location is. GE has used this technique to allow an operator to correlate a CAD model to the real world and bring the two into alignment [Oxenberg 1988]. In the GE system used at JPL, the operator updates the database by superimposing a wireframe line drawing of an object on a video scene containing the object. The real-world and model are brought into alignment by having the operator match vertices of the model to their equivalent vertices on the image.

Overlay of stick figures on video can also be used to compare planned (computer) versus actual performance. Russell [1986], for example, has suggested superimposing stick figures on live video to aid in satellite docking.

Temporal predictive displays project what will happen at some point of time in the future (Figure 11). Sheridan [1986] first suggested using computer generated images to predict robot response to operator commands under a time delay. Since then, the concept has developed to
Include full dynamic models in which the operator interacts as if actually manipulating the remote object. The operator's position and force trajectories are then fed to the remote robot which performs the task using the human generated trajectories as controller inputs. Reported performance improvements have been significant [Sheridan 1987, 1989].

PROJECTION OF LOCATION AFTER TIME t

Figure 11: Temporally predictive displays project what will happen in the future based on controller actions occurring now.

Stimulus-Response Reconciliation
A well-recognized but perplexing problem, that anyone who has operated a remote manipulator has experienced, is control reversal, resulting in unexpected movement. Control reversal (also known as cross-coupling, stimulus-response mismatch, and orientation-display incompatibility) has been reported in fly-by-wire applications [e.g., Van Cott 1972], teleoperated vehicles [e.g., McGovern 1988], and telemanipulators [e.g., Brooks 1979, Smith 1988].

Brooks [1979] developed a classification of stimulus-response mismatch in an effort to understand the problems observed during experiments. He concluded that cross-coupling was due to mechanical, geometrical, or observational stimulus-response (SR) mismatch. See Figures 12-14 for illustrations of these conditions. The key to understanding SR mismatch is to recognize that it is caused by a disparity between what the operator expects and what s/he observes.

The effects of operator stimulus-response incompatibility have been recognized for many years, but have only recently been receiving research attention. Smith [1990] concludes that a workable approach may be to use "computer-mediated transformation of video images of the telemanipulator to a spatially compliant form before they are displayed to the operator." This solution is currently only possible in a 3D graphical format, but increasing processor speeds will make real-time image manipulation a viable alternative in coming years.

Figure 12: Mechanical Stimulus-Response Mismatch. The operator desires to move to the left, but mechanical cross-coupling in the hand controller causes the manipulator to move differently. The observed and actual motion are the same, but incorrect [Brooks 1979].

Figure 13: Geometrical Stimulus-Response Mismatch. The operator commands a move to the left, but dissimilar geometric relationships between the hand controller and telerobot causes the manipulator to move differently. The observed and actual motion are the same, but not what the operator expected [Brooks 1979].

Figure 14: Observational Stimulus-Response Mismatch. The operator commands a move to the left and the manipulator moves as commanded, but due to the positioning of the camera the observed motion is different than the expected motion [Brooks 1979].
One possible solution, suggested by Crane [1989], is to use screen coordinates as the control frame. The idea being that regardless of manipulator orientation or position, operator commands will be referenced with respect to the screen coordinates that he is currently viewing. Hence, a camera sighted along the end effector axis would move forward whenever the operator pushed the control stick forward in the screen direction. This is equivalent to using an end effector control frame, provided that the camera and end effector have a rigid connection. The true difference between screen coordinates and end effector control becomes clear when one has a camera located at least one or more joints from the end effector. To illustrate (see Figure 15), imagine using an end effector control frame with the camera fixed at the manipulator base, and rotate the end effector 180° about the vertical axis so that the end effector is facing the camera. Now, using an end effector control frame, rightward stick motions cause the end effector to move left in the image. If control were in screen coordinates, a right stick motion would cause a rightward end effector motion. It can be hypothesized that this form of stimulus-response remediation could have disastrous effects when more than one monitor (screen control frame) is used if the operator retains the mental model/relationships from the previous screen. Regardless of this possible shortcoming, this scheme does warrant investigation in a teleoperated setting.

2D Symbolic Depth
The first method is a symbolic representation of various depth cues which are superimposed on the real-time image. An example of such a display is shown in Figure 16. The cross in the middle represents the pitch and yaw of the end effector, while the two converging parallel lines represent range to the task. The astronaut knows that the end effector is properly aligned when the yaw and pitch rectangles are centered on the reticule, and the converging range rectangles decrease to zero. The data for this display can be generated from either a CAD database, proximity sensors [Bejczy 1980], or laser range finders.

![Figure 16](image)
Figure 16: A 2D symbolic depth and alignment display for improving end effector control with only a single camera view.

Virtual Views
A second method of providing depth information is to generate a simulated view as if seen from a virtual camera placed at a useful location near the task (see Figure 17). Through virtual scene generation, the astronaut could see the task/manipulator interface from any angle/location desired, even from within the object looking out if that would aid performance. Obviously, a view orthogonal to the existing video camera view would provide the most accurate information for motions along the camera axis.

Display of Depth thru Graphics
Depth information is fundamental to most teleoperated tasks. However, for many teleoperation tasks, either a needed view will not be available, or available views will not provide appropriate visual depth cues. Fortunately,
However, other viewing angles would also allow the astronaut to "fly" around the task to observe progress from all angles.

**Perspective Depth**

A third method of displaying useful depth information is through the use of perspective grids [Stark 1987, Kim 1987, and Wickens 1990]. As shown in Figure 18, a perspective grid is superimposed on the image, and markers are dropped to the grid to indicate the control point position to the operator. The goal state could also be indicated through the use of some appropriate symbol. Experiments by Kim [1987] have shown that a perspective grid can improve human performance.

![Figure 18: A perspective grid display superimposed on a single camera view to aid the operator.](image)

**View Enhancements**

Through the use of image enhancement techniques, it is possible to aid the astronaut in performance of tasks which might be impossible without such aids. For example, a low-contrast picture could be "stretched" in real-time to bring out details that would otherwise be unobservable [Gonzalez 1977]. Or features within shadows could be expanded to effectively increase the dynamic range of the display. Another possibility would be to delineate edges or surfaces which might be indiscernible due to similar texture or surface reflectivities.

Another enhancement might involve deliberate image distortions to aid in grasping/docking operations. For example, images could be plastically "stretched" in screen dimensions which highlight errors. Yet another enhancement would be to remove image blurring due to task motion.

Image enhancement could also improve specular reflections through dynamic memory of prior scenes. As the sun reflects off thermal blankets during orbit, specular image portions could be removed and lost detail could be filled in by computer memory (acquired when the Sun was in a different position) or computer generated from a CAD data base.

Symbolic enhancements are artificial graphic elements which aid the operator by providing information that might otherwise not readily be available. One such enhancement might be to turn the end effector different colors to indicate gripper status directly to the astronaut rather than through a secondary display, such as a light on the console. For example, a red gripper would mean the object is not firmly grasped, whereas a green gripper would indicate positive capture. Another enhancement would be to indicate problem areas with symbolic colors. For example, a temperature probe in the end effector could be used in conjunction with an ORU CAD data base to indicate problem areas by "painting" a cryogenic cold-spot blue or a hot-spot red (see Figure 19). Symbolic enhancements such as these can be achieved through graphic overlays of the real-time video.

![Figure 19: Symbolic enhancements are artificial graphic overlays that supply information thru direct visual analogy. Here, for example, a hot-spot is highlighted to indicate a fault location.](image)

**OPERATOR CUES**

Another important operator visual aid is represented by a class of visual "cues" used to improve proper orientation and positioning of the end effector when viewing conditions are limited. Figure 20 illustrates a visual cue used for grasping payloads with the Shuttle RMS. These types of cues are typically constructed as an integral part of the payload in a position that can be directly viewed from an end effector mounted camera. Cues help identify size, distance, shape, orientation and location of objects and aid task performance. Cues minimize task time and insure proper execution by providing a pre-existing, logical relationship between end effector, task and operator actions.
Cues typically use high-contrast backgrounds and foreground elements to insure they are readily discernable. In color systems, colors of differing hues can be used. Alignment cues typically use single or multiple stripes or patterned markings. Range cues typically use size on the screen or comparison markings (e.g., the end effector is at the proper range when a marking on the end effector and a marking on the task are the same length).

Another method of visual enhancement for grasping or docking functions would be to generate graphic cues superimposed on the monitor. For example, alignment of the real-time image with a graphic outline or "stick-figure" could indicate successful operation. Figure 21 is a concept in which moire circles are superimposed on the end effector video image to indicate yaw and pitch alignment. Graphic overlay techniques would allow tasks to be performed without the need for mounting or painting special target cues on the satellite.

CONCLUSIONS
This paper has presented a number of different concepts for improving operator performance through visual aids. Operator vision aids were broken down into three primary categories: (1) remote camera and lighting functions, (2) display aids, and (3) operator cues.

In the area of remote camera and lighting, we presented concepts for:

• Automatic tracking of the end effector or task with the camera and lighting.

• New camera designs.

• Simulated camera views to aid in proper camera and lighting placement.

• Computer assisted camera and lighting placement.

• Voice control of camera and lighting functions.

In the area of display aids for the operator, we presented concepts for:

• Displaying restricted zones such as collision and object limits.

• Predictive displays that help the operator identify where something is or will be after the control action is taken.

• Displays that reconcile visual feedback with operator inputs to prevent display/control mismatch.

• Methods for providing the operator with depth information through graphical aids.

• Enhancement techniques that provide the operator information through image modification and symbolic image manipulation.

Finally, visual cues can be used to improve proper orientation and positioning of the end effector when viewing conditions are limited. Cues help identify size, distance, shape, orientation and location of objects, and aid task performance. Cues minimize task time and insure proper execution by providing a pre-existing, logical relationship between end effector, task, and operator actions.

It is hoped that this report condenses and focuses both the current state-of-the-art and the future challenges of applying operator visual aids to teleoperation.
REFERENCES


Brooks, T., Discussion with Paul Heckman of NOSC on Uhrich and Heckman's end effector tracking results. These discussions ultimately led to the development of a 6 DOF end effector tracker at JPL in 1980 in Tony Bejczy's lab. The results were unpublished since statistical experiments were not performed; however, the tracking algorithms were published in Brooks [1979] and Bejczy [1980], 1978.


