# Evaluation of Vision Systems for Teleoperated Land Vehicles

**Dwight P. Miller**

ABSTRACT: This paper describes a pilot study that compares vision systems that might be used to control a teleoperated land vehicle. The study compares three forward-looking vision systems in two modes of driver interaction: (1) actual remote driving and (2) noninteractive video simulation. Remote driving has the advantage of realism but is subject to variability in driving strategies and can be hazardous to equipment. Video simulation provides a more controlled environment in which to compare vision-system parameters, but at the expense of some realism. Results demonstrate that relative differences in performance among the visual systems are generally consistent in the two test modes. A detection-range metric was found to be sensitive enough to demonstrate performance differences viewing large obstacles using black-and-white and color vision systems. Consequently, future experimentation, aimed at optimizing vision-system parameters, will rely to a greater extent on the more cost-effective, video-simulation approach.

## Introduction

Sandia Laboratories is currently working with the U.S. Army Missile Command to develop a portable controller for the Teleoperated Mobile Antiarmor Platform (TMAP). Conceptually, TMAP is a small, highly maneuverable, remotely operated vehicle with the capabilities to perform surveillance and antitank missions at or behind enemy lines. In addition to controller development, mobility testing is being conducted at the Sandia Labs Mobile Vehicle Test Range to study the effects of vision-system characteristics on off-road, remote-driving performance. Factors such as camera placement, video resolution, field of view, color, steering coupling, and the use of multiple cameras are being considered for evaluation. Recent activities include running a pilot study to establish relationships between remote-driving performance and laboratory visual performance tasks using videotapes of off-road driving recorded with the vision systems of interest.

*Background*

Although researchers have been interested in the human factors involved in remotely operated devices for several decades, there is a relative dearth in the reported findings of studies involving the teleoperation of land vehicles [1]. Most of the literature concerning automobile control and its visual requirements address driving tasks performed in a well-defined roadway environment [2]. Visual tasks of interest in this environment usually include previewing roadway features, reading highway signs, monitoring vehicular instrumentation, and checking rearview mirrors. Many of the findings in these studies do not have much relevance to off-road driving, where speeds are much slower, and visual tasks concentrate on the perception of the immediate terrain and its safe negotiation.

Television requirements for teleoperation have been addressed, for the most part, in the literature dealing with remote manipulators in controlled processing environments or in space [3], [4]. Studies in this field have examined the relative benefits of using color, multiple cameras, stereo vision, and different levels of image resolution in tasks involving the remote manipulation of objects in hazardous work environments. In these applications, where the manipulation and placement of solid objects are critical, visual information regarding detail and depth has been found to be important [5]. It is anticipated that the visual requirements of remote, off-road driving tasks have little in common with remote manipulator tasks, and that, except for generalities concerning the inadequacies of television as a substitute for direct viewing, little can be gleaned from this body of literature.

Perhaps the most relevant research to date comes from a series of studies conducted by Ross Pepper and his colleagues at the Naval Ocean Systems Center (NOSC) in Kailua, Hawaii [6]-[9]. Originally concerned with the human performance aspects of underwater teleoperators, NOSC has recently broadened its mission to include land, airborne, and space environments. Recent investigations have involved the remote driving of a jeep-sized vehicle over a slalom course with the aid of a helmet-mounted, head-motion-coupled, stereo viewing system. Remote driving produced course times at least twice as long as when the same vehicle was driven conventionally [10]. Attempts to equate remote- and direct-driving times by degrading direct-driving vision failed, suggesting that the visual processes involved in off-road driving are not yet well understood.

After several years of conducting technology feasibility studies on vision systems for remote driving, several observations have been made by engineers and technicians at Sandia. First, we found that color video seems to have an advantage over black-and-white (B&W) for systems of comparable resolution. Although this contradicts prior findings in the teleoperator literature, it may be a valid concept for remote, off-road driving. Second, increasing the effective horizontal field of view (HFOV), by using multiple cameras or by panning a single camera, seemed to assist local area navigation and turn negotiation. Third, although we could recognize significant differences in video picture quality due to system resolution, the effects of resolution on remote-driving performance are unknown. This study will test some hypotheses based on these observations using controlled experimentation. This research, although formally tied to the TMAP project, can support any land-vehicle application where one human operator must perform off-road remote driving using primarily visual feedback.

*Experimental Approach*

Military-mission scenarios require teleoperated vehicles to be driven from point A to point B quickly and successfully. Therefore, to compare candidate vision systems, the typical approach has been to have operators remotely drive a course and measure the elapsed time, probability of success, or operator workload. Although this is the most realistic of approaches, it has some inherent

Presented at the 1987 International Conference on Systems, Man, and Cybernetics, held in Alexandria, Virginia, October 20-23, 1987. Dr. Miller is a Member of the Technical Staff in the Human Factors Group at Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185. This research was supported by the U.S. Army Missile Command through the Teleoperated Mobile Antiarmor Platform project.

0272-1708/88/0600-0037 $01.00 © 1988 IEEE

drawbacks. First, it is a costly, time-consuming, and sometimes dangerous proposition. Second, since the drivers' experience, abilities, and strategies play important roles in overall performance, experiments evaluating vision-system parameters are susceptible to strong influences from these factors.

In order to examine any but the largest differences in vision systems, we must extract the visual performance tasks from the remote-driving activity. The following visual tasks are essential to remotely driving from point A to point B in undeveloped terrain: (1) Scanning immediate terrain for obstacles or terrain features that may impede progress or damage equipment. (2) Searching distant terrain for landmarks that could be used to assist in navigating from the current location to a desired location. (3) Monitoring pitch and roll attitudes via the vision system and/or instrumentation to maintain vehicle stability. (4) Assessing the size and separation of obstacles to judge vehicle clearance prior to driving between them.

All of these tasks can be simulated in the laboratory using video imagery without the use of an interactive driving task. Missing from this part-task, noninteractive simulation are the responses of the observer to the results of visual tasks, such as turning the steering wheel to avoid an obstacle. Instead, the observer presses response keys and makes verbal reports, telling the experimenter what he sees. This study examines two of the visual tasks referred to in the preceding paragraph (tasks 1 and 3) using three different visual systems in a part-task simulation. Tasks 2 and 4 will be examined in follow-on experimentation. In addition to the simulation, identical visual performance tasks were performed during actual remote driving so that comparisons could be made to evaluate the validity of the video-simulation technique. If comparisons of vision systems prove to be consistent across actual remote-driving and video-simulation conditions, subsequent experimentation could rely, to a greater extent, on video simulation. It is hoped that the development of a valid, inexpensive laboratory-simulation technique will lead to more cost-effective methods of evaluating and comparing numerous combinations of vision-system parameters for use in teleoperated land vehicles.

| | | Technique: Remote Driving | Technique: Simulation |
|---|---|---|---|
| Camera | SS Color | 6 Subjects | 7 Subjects |
| | Color | 6 Subjects | 6 Subjects |
| | B&W | 6 Subjects | 5 Subjects |

Fig. 1. Pilot study experimental design.

## Experimental Method

### *Conditions*

A 3 × 2 factorial design was selected in which three levels of vision-system quality were tested under both remote driving (RD) and video simulation (Sim). Based on previous experience, the following vision systems (cameras) were chosen:

- Poor: Black-and-white, fixed mount (B&W)
- Medium: Color, fixed mount (color)
- Best: Color, steering slaved (SS color)

Nominally, six subjects (Ss) participated in each condition, for a total of 36 (see Fig. 1). All six subjects were screened for average driving experience, corrected 20/20 vision, absence of bifocal or trifocal corrective lenses, and absence of major color-vision anomalies.

To control for variations in the video medium other than the variables of interest, identical video sequences and obstacle orders were used in all six conditions. In fact, all three Sim tapes were recorded simultaneously to ensure that they were identical in terms of lighting conditions, approach angle, speed, and duration. If each subject had been tested under all six conditions, then, due to the richness of the video information and the potential for using location as a memory aid, there was a possibility that the six subjects could have anticipated obstacles on their second and third exposures to the paths. Obviously, this would have invalidated the obstacle detection and identification data. Therefore, different subjects were used in each condition to avoid data contamination resulting from learning effects.

### *Equipment*

Videotapes were made using a Sony DXC-101 charge-coupled-device (CCD) color camera, capable of 320 lines of resolution (poor and medium conditions), and a Canon CI-10 CCD color camera, capable of 300 lines of resolution (best condition). Both cameras were fitted with Cosmicar 16-mm ALC lenses, providing a HFOV of approximately 42 deg, and automatic iris control, which adjusted for instantaneous changes in lighting conditions. The Canon camera was mounted on a turntable that was rotated such that the camera panned left and right in coordination with the vehicle's steering. The full lock-to-lock, steering-slaved effective HFOV was 102 deg, with an instantaneous HFOV of 42 deg. The Sony camera was installed on a fixed Vicon adjustable mount. Both cameras were mounted on the roof of the instrumented 1981 Jeep Cherokee at approximately 6.5 ft above ground level. Both cameras were mounted looking forward, with a 10–11-deg negative pitch, putting the aim points at approximately 26 ft in front of the vehicle's front bumper.

Recording and playback of the videotapes were done using Panasonic PV-4700 VHS videocassette recorders. Panasonic BT-S1900N 19-in. color monitors (350 lines of resolution) were used for all video displays. The B&W video condition was implemented by recording in color and playing back with no color on the monitor. This technique not only obviated the need for a third camera but also allowed a direct comparison between color and B&W video without any differences in resolution contributed by the monitor. Detection/identification runs were controlled by an IBM PC, which also recorded keyboard responses.

***Driving Station*** All experimentation took place at a fixed-base, remote-driving station assembled from three standard-width racks forming a concave console with a 28-in.-high work surface. Subjects in all conditions sat in front of a control pedestal, which was placed directly in front of the console. The controls were mounted on the pedestal, including a steering wheel, brake pedal, throttle pedal, keyboard, and a separate pushbutton for out-of-range responses. Viewing distance to the display monitors was approximately 4 ft.

*Tasks*

Subjects in both the RD and Sim conditions performed identical visual performance tasks. The only difference was that those in the RD condition actually drove the test vehicle via steering, throttle, and brake controls at the remote-driving station, while those in the Sim condition sat at the driving station and observed videotapes recorded from the test vehicle traveling over the identical course at an earlier date.

***Obstacle Detection and Identification*** Subjects were instructed to preview the path of the vehicle for obstacles and terrain features that may impede its progress. The path was delineated by 28-in.-high traffic pylons, separated by 24 ft, and spaced approximately every 40 yd. The bright orange pylons were covered with white paper to improve visibility for the monochrome camera. The obstacles included man-made objects such as railroad ties, gas cans, 55-gal drums, and intersecting roads, as well as natural objects such as bushes, cacti, ditches, berms, and rocks. Upon first detecting obstacles in the far field of the video scene, subjects pressed the "D" key on the IBM PC keyboard. When the obstacles became identifiable in the near field, subjects pressed the "I" key and verbally reported what they saw. Revisions to identifications were permitted, as the data of interest were maximum distances at which obstacles were correctly identifiable using a given visual system. All other keyboard keys were covered in order to simplify responding.

Each subject drove or observed five paths across the desert mesa, traveling at approximately 3 mph, covering a total distance of about 1 mi., and taking about 20 min. Nominally, five examples of each obstacle type (e.g., railroad tie) were placed at random locations within the five paths. Occurrence of natural obstacles was controlled by the selection of paths on the mesa. Subjects performed visual tasks on each path for about 4 min, with 2-min breaks intervening. The data of interest for this task were: (1) the probabilities of correct detection and identification, (2) false alarms, (3) misses, and (4) distances at which correct detections and identifications were made.

***Instrument Monitoring*** In addition to the primary task of searching for obstacles, the six subjects were asked to perform a secondary task as time permitted. The purpose of the additional task was to make the remote-driving task more realistic and to increase the workload for the subjects in the Sim conditions to a level comparable to that of remote driving. The secondary task consisted of visually monitoring three graphic instruments on a second CRT and responding to any one of three variables going out of its acceptable range by holding down a pushbutton for the duration of the excursion. The instruments displayed the actual current *speed*, *pitch*, and *roll* of the vehicle in the RD conditions and the actual recorded data in the Sim conditions. Acceptable ranges were clearly indicated on the moving-pointer gauges as shown in Fig. 2.

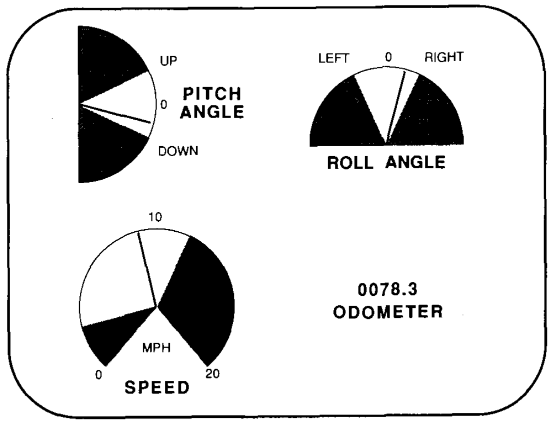


Fig. 2. Graphical displays for secondary task.

In order to acquire subjective comparisons of the three visual systems, subjects were shown 1-min video clips of each and asked to fill out a questionnaire after their obstacle detection/identification runs. In addition to qualitatively comparing the three visual systems, the subjects were asked to comment on the adequacy of the lens angle, the desirability of using color, the desirability of slaving the camera to the vehicle's steering, and several other aspects of the study.

## Results and Discussion

*Questionnaire*

Subjective data from the questionnaire provided consistent insights regarding the desirability of visual-system features. When asked to rate the three cameras on a scale of 1 to 10, the six subjects in both the Sim and RD conditions gave higher ratings for color (6.6) and SS color (6.1) than B&W (4.2) ($F = 20.4, p = 0.0001$). (The $p$ value indicates the probability that the difference was due to chance alone. In the behavioral sciences, $p$ values less than 0.05 are usually considered as being statistically significant; however, larger $p$ values may be cause for explanation.) Since there was no significant difference between the ratings for color and SS color, the effect can be attributed to the color feature alone. Independent evaluations of the features of color and steering slaving resulted in the judgments that color was a more desirable feature (Chi-Squared = 15.7, $p$ = 0.001). No differences were found to exist between the Sim and RD groups. The preference for color was concordant with the experimenters' expectations. However, the limited perceived value of the steering-slaving feature was surprising. Previous experience of the Sandia researchers suggested that the steering-slaved camera would be a significant improvement over a fixed camera in remote-driving applications. In retrospect, two factors may have contributed to this outcome. First, the mounting hardware for the steering-slaved camera introduced some jitter to the video image. Second, the subjects were asked to evaluate the two features in the context of searching for potential obstacles. It is likely that higher ratings would have resulted if the subjects operated and were asked about this feature in a context stressing vehicle maneuverability.

When asked if they would prefer to have a wider or a more narrow-angled lens on the camera, RD subjects responded *wider* and Sim subjects responded *narrower* on a 10-point scale ($F = 7.6, p < 0.001$). There was also a trend that the subjects in the SS color condition preferred a narrower lens, while those in the fixed camera conditions indicated a preference for a slightly wider lens ($F = 2.7, p = 0.1$). These results make logical sense in that subjects who actually performed remote driving would be more sensitized to problems associated with navigating through the course, while Sim subjects would be more concerned with the visibility of detail and, hence, the magnification of potential obstacles. Likewise, subjects using the steering-slaved camera might feel that the effective HFOV was adequate, whereas those using fixed cameras might prefer a wider view of the terrain than that provided in the study.

### *Obstacle Avoidance*

For both the Sim and RD groups, runs 2-5 were analyzed for obstacle detection and identification, as run 1 was treated as a training period. Overall, the subjects correctly detected 75 percent of the 35 obstacles and correctly identified 63 percent. Subjects reported an average of five false alarms (objects, real or imagined, that were not considered to be obstacles by the researchers) per session. No differences were found among the six experimental conditions based on any of these indexes of performance. The search performance metrics did not appear to be sensitive to the vision-system differences used in this study.

In contrast to the search performance metrics, the detection-range measurements proved to be sensitive to differences in performance among the six experimental conditions. Detection range (DR) is the distance from the obstacle at which the subject detected its presence and pressed the response key. As could be expected with such a wide variety of visual targets, DR performance varied greatly across the obstacles. Mean DRs across all conditions ranged from 17 ft for a pile of rocks to 194 ft for the 10-ft-tall cholla cactus. The grand mean DR for all obstacles was 57 ft. Variability among subjects was also large. For example, the subjects' DRs for the large cactus ranged from 31 to 321 ft. Typically, the standard deviation for DRs of a given obstacle was about 50 percent of the mean. The large variability in DR data made analysis using parametric techniques somewhat difficult and the fixed effects of camera and technique somewhat inconsistent.

Looking at the frequency histograms of the DR data, we find positively skewed distributions, that is, the majority of the responses fell in the short distances (from 0 to 50 ft) with decreasing numbers of responses at longer distances (see Fig. 3).

A logarithmic transformation of these data decreases the impact of the few, very long DRs on central-tendency statistics (such as the mean), stabilizes the variance, and results in producing a frequency histogram that is more normally distributed (see Fig. 4). Both of these effects improve the validity of parametric inferential statistics such as an analysis of variance (ANOVA). Several of

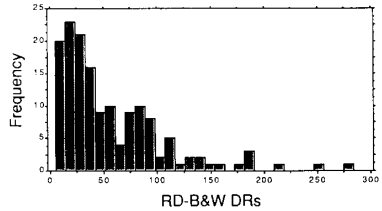


Fig. 3. Frequency distribution of DRs in the RD-B&W condition.

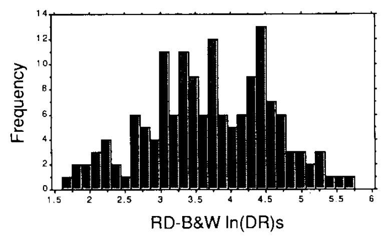


Fig. 4. Frequency distribution of ln(DR)s in the RD-B&W condition.

the results reported subsequently will be based on transformed DR data.

One of the major questions to be answered by this study deals with the ability to obtain data using the simulation technique that are comparable to those acquired during actual remote driving. We know that there were no differences in overall search performance, but how do results based on the more sensitive DR metric compare? Frequency histograms for the three camera conditions look very similar when comparing the Sim and RD distributions. Looking at individual obstacles, however, we find mixed results. The large cactus, referred to earlier, showed an advantage for RD (231 ft) over Sim (161 ft). A similarly large advantage was found for a large cable spool (155 ft versus 112 ft). A large difference was found in the opposite direction for a large tire (66 ft in RD versus 110 ft in Sim). This inconsistency is difficult to explain. For the remaining obstacles, differences in technique were much smaller and had no apparent pattern in direction. When data from all 35 obstacles were combined, the differences canceled each other, producing means of 57 ft (Sim) and 58 ft (RD). A two-factor ANOVA using ln(DR) data yielded no significant difference due to technique ($F$ = 0.375, $p$ = 0.5).

The similarity of performance in the Sim and RD conditions, both in terms of probability of detection and mean DR, suggests that the noninteractive simulation technique can realistically measure visual performance on tasks associated with off-road vehicle teleoperation. It appears that the secondary task imposed upon all six subjects contributed to the success of the method by helping to equalize the workload across the Sim and RD conditions. The mean number of responses per session was 19.4 for Sim conditions and 0.3 for RD conditions ($F$ = 98.3, $p$ = 0.0001), with no differences between cameras.

The second major question is how does the camera used affect obstacle detection performance? Inconsistencies, similar to those mentioned earlier, evidenced themselves when obstacle data were examined individually. For example, the large cactus demonstrated large camera effects (B&W = 150 ft, color = 190 ft, SS color = 242 ft) in one direction, while a cable spool showed a modest increase in DR for color over B&W (103 ft versus 85 ft) but no benefit for SS color (90 ft). This may be due to the fact that the large cactus was at a location on the course where the approach was curved enough for the steering-slaved camera to have a beneficial effect, and the spool was not. However, when obstacles were divided on the basis of a straight or a curved approach, no consistent steering-slaved advantage was found for obstacles on curved approaches.

When data from all 35 obstacles were combined, the large inconsistencies tended to neutralize each other, as was true with the technique differences, but the small differences in the same directions also tended to accumulate. The other half of the two-factor ANOVA using ln(DR) data showed a nearly significant difference due to camera ($F$ = 2.83, $p$ = 0.07). Most of this effect was attributable to the difference between B&W (51.4 ft) and the two color conditions, color (59.9 ft) and SS color (60.3 ft) (see Fig. 5). When data from both color conditions are combined, the B&W-color difference becomes statistically significant ($F$ = 5.66, $p$ = 0.02), demonstrating an average advantage of about 9 ft for color cameras.

Comparing overall DR means in the six experimental conditions, we found that, although not statistically strong ($F$ = 1.74, $p$ = 0.2), there was some indication of an interaction taking place between technique and camera (see Fig. 6). Notice that in Sim conditions, the camera effect monotonically increases from B&W (45.4 ft), to color (58.3 ft), to SS color (63.6 ft). However, in the RD conditions, B&W (56.4 ft) and SS color (56.5 ft) are nearly identical, with color slightly better at 61.5 ft. The consistent color advantage confirmed a priori expectations, but the lack of a steering-slaved advantage

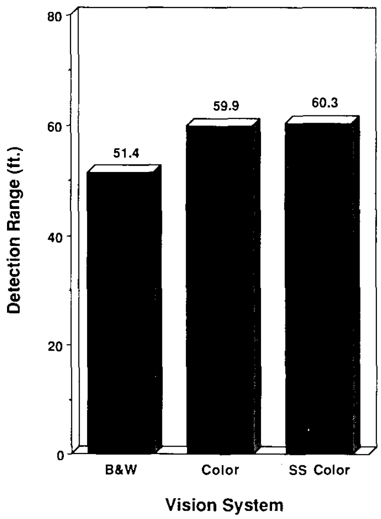


Fig. 5. Detection range means for three vision systems.

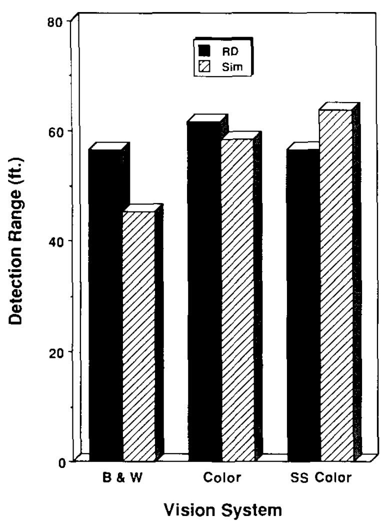


Fig. 6. Detection range means for six experimental conditions.

was surprising. The jitter affecting the SS color camera may have been exacerbated as RD testing continued and as the equipment acquired additional wear. This may explain the lower mean DR for the RD-SS color condition and, hence, the apparent interaction.

As expected, large obstacles were detected at longer mean DRs (90 ft) than smaller obstacles (40 ft), ($F$ = 186.7, $p$ = 0.0001). The color advantage ranged from an average of less than 5 ft for smaller obstacles to an average of about 22 ft for large obstacles. Analyzing performance on large obstacles only, a significant camera effect was attained ($F$ = 3.52, $p$ = 0.04). These results suggest that for subtle differences on vision systems, the use of large obstacles will provide the most sensitive comparisons.

Based on the experimental results of this pilot study, we feel that the following conclusions can be drawn: (1) Unbiased subjective appraisals of the vision systems agree with the performance data and can be useful in evaluating various aspects of video performance. (2) Contrary to the predictions of the researchers, the current implementation of a steering-slaved camera does not provide an advantage in obstacle detection and avoidance on the type of course used in the study. (3) Color is preferred by the subjects and provides a detection-range advantage that increases with obstacle size. (4) The noninteractive simulation technique can provide data comparable to those acquired during actual remote driving.

In summation, the pilot study did not support some hypotheses regarding the relative quality of the three chosen vision systems. The study did validate the notion that purely visual tasks can be extracted from the interactive driving activity and analyzed in a laboratory-based, noninteractive simulation. We think that, with some modifications and refinements, the simulation approach could be a very cost-effective method for future evaluations of numerous parameters of vision systems for teleoperated off-road vehicles.

## References

[1] D. E. McGovern, "Technology Assessment for Command and Control of Teleoperated Vehicles," Sandia Report SAND87-1225, Sandia National Laboratories, Albuquerque, NM, May 1987.

[2] T. W. Forbes, *Human Factors in Highway Traffic Safety Research*, New York: Wiley-Interscience, 1972.

[3] E. W. Johnsen and W. R. Corliss, *Human Factors Applications in Teleoperator Design and Operation*, New York: Wiley-Interscience, 1971.

[4] C. T. Huggins, T. B. Malone, and N. L. Shields, "Evaluation of Human Operator Visual Performance Capability for Teleoperator Missions," in *Remotely Manned Systems* (E. Heer, Ed.), California Institute of Technology, Pasadena, CA, pp. 337–350, 1973.

[5] W. H. Crooks, L. A. Freedman, and P. P. Coan, "Television Systems for Remote Manipulation," in *Proc. Human Factors Society 19th Annual Meeting*, pp. 428–435, 1975.

[6] R. L. Pepper, D. C. Smith, and R. E. Cole, "Stereo TV Improves Operator Performance Under Degraded Visibility Conditions," *Opt. Eng.*, vol. 2, no. 4, pp. 579–585, 1981.

[7] R. E. Cole, R. L. Pepper, and B. L. Pinz, "The Influence of Head Movement Parallax on Perceptual Performance Under Direct- and TV-Displayed Conditions," CA:NOSC TR678, San Diego, CA, May 1981.

[8] E. H. Spain, "A Psychophysical Investigation of the Perception of Depth with Stereoscopic Television Displays," doctoral dissertation, University of Hawaii, May 1984.

[9] E. H. Spain and R. L. Pepper, "Effect of Wide Camera Baselines and Magnification on Visual Performance with a Stereoscopic TV System," *Proc. Psychology in the Department of Defense*, Colorado Springs, CO, Apr. 1986.

[10] R. L. Pepper, "Human Factors in Remote Vehicle Control," in *Proc. Human Factors Society 30th Annual Meeting*, pp. 417–421, 1986.

**Dwight P. Miller** received the B.S. degree in electrical engineering and the B.A. degree in psychology from Lafayette College in 1974, and the M.A. and Ph.D. degrees in experimental psychology from The Ohio State University in 1980. During his graduate studies, he served a one-year internship with Honeywell's Man-Machine Sciences Group in Minneapolis, where he helped design avionics systems and consumer products. He is currently a Member of the Technical Staff in the Human Factors Group at Sandia National Laboratories, Albuquerque, New Mexico, where he supports programs in energy, weapons, safeguards, and microelectronics. His research and design interests are in the teleoperation of land vehicles, automotive human factors, human reliability, architectural ergonomics, and human-computer interaction.