The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Scaling Effects for Streaming Video vs. Static Panorama in Multirobot Search

Prasanna Velagapudi, Huadong Wang, Paul Scerri, Michael Lewis, *Member, IEEE* and Katia Sycara, *Fellow, IEEE*

Abstract— **Camera guided teleoperation has long been the preferred mode for controlling remote robots with other modes such as asynchronous control only used when unavoidable. Because controlling multiple robots places additional demands on the operator we hypothesized that removing the forced pace for reviewing camera video might reduce workload and improve performance. In an earlier experiment participants operated four teams performing a simulated urban search and rescue (USAR) task using a conventional streaming video plus map interface or an experimental interface without streaming video but with the ability to store panoramic images on the map to be viewed at leisure. Operators were more accurate in marking victims on maps using the conventional interface; however, ancillary measures suggested that the asynchronous interface succeeded in reducing temporal demands for switching between robots. This raised the possibility that the asynchronous interface might perform better if teams were larger. In this experiment we evaluate the usefulness of asynchronous video for teams of 4, 8, or 12 robots. Operators in the two conditions were equally successful in finding victims, however, the streaming video maintained its advantage for accuracy in locating victims.**

I. INTRODUCTION

PRACTICAL applications of robotics can be classified by two distinct modes of operation. Terrestrial robotics in tasks such as surveillance, bomb disposal, or pipe inspection has used synchronous realtime control relying on intensive operator interaction usually through some form of teleoperation. Interplanetary and other long distance robotics subject to lags and intermittency in communications have used asynchronous control relying on labor intensive planning of waypoints and activities that are subsequently executed by the robot. In both cases planning and decision making are performed primarily by humans with robots exercising reactive control through obstacle avoidance and safeguards. The near universal choice of synchronous control for situations with reliable, low latency communication suggests a commonly held belief that experientially direct control is more efficient and less error prone. When this implicit position is rarely discussed it is usually justified in terms of "naturalness" or "presence" afforded by control relying on teleoperation. Fong and Thorpe [1] observe that direct control while watching a video feed from vehicle mounted cameras remains the most common form of interaction. The ability to leverage experience with controls for traditionally piloted vehicles appears to heavily influence the appeal for this interaction style.

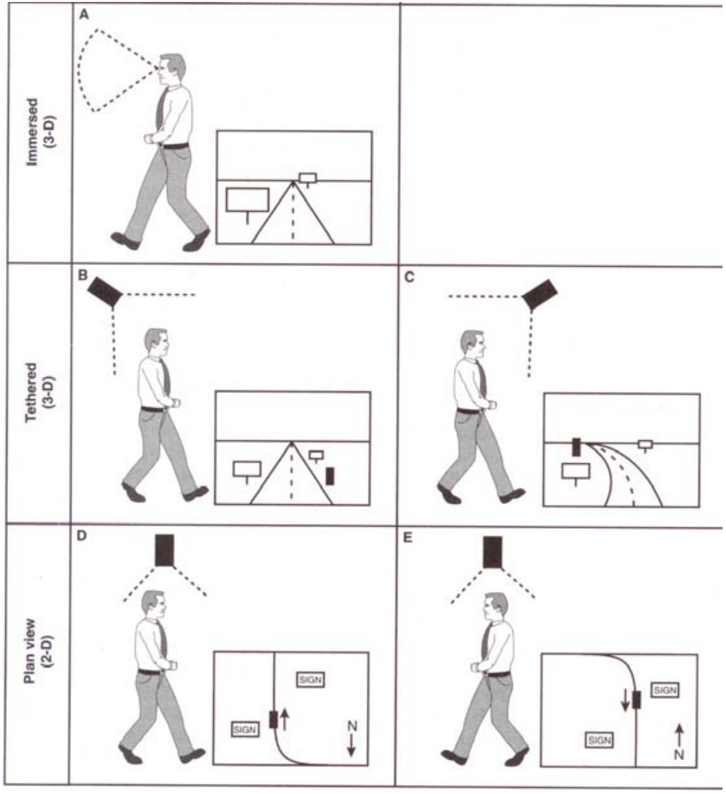

Figure 1. Viewpoints for control from Wickens and Hollands *Engineering Psychology and Human Performance*, 1999.

Control based on platform mounted cameras, however, is no panacea. Wickins & Hollands [2] identify 5 viewpoints used in control, three of them, immersed, tethered, and "plan view" can be associated with the moving platform while 3rd person (tethered) and plan views require fixed cameras. In the immersed or egocentric view (A) the operator views the scene from a camera mounted on the platform. The field of view provided by the video feed is often much narrower than human vision, leading to the experience of viewing the world through a soda straw from a foot or so above the ground. This perceptual impairment leaves the operator

---

Manuscript received March 1, 2009. This research has been sponsored in part by AFOSR FA9550- 07-1-0039 and AFOSR FA FA9550-08-1-0356. Prasanna Velagapudi, Paul Scerri, and Katia Sycara are with the Robotics Institute, Carnegie Mellon University, USA (pkv@cs.cmu.edu, 412-268-8825)

Huadong Wang and Michael Lewis are with the School of Information Sciences, University of Pittsburgh, Pittsburgh, PA 15260 (ml@sis.pitt.edu, 412-624-9426).

978-1-4244-3804-4/09/$25.00 ©2009 IEEE

prone to numerous, well-known operational errors, including disorientation, degradation of situation awareness, failure to recognize hazards, and simply overlooking relevant information [3, 4]. A sloped surface, for example, gives the illusion of being flat when viewed from a camera mounted on a platform traversing that surface [5]. For fixed cameras the operator's ability to survey a scene is limited by the mobility of the robot and his ability to retain viewed regions of the scene in memory as the robot is maneuvered to obtain views of adjacent regions. A pan-tilt-zoom (ptz) camera resolves some of these problems but introduces new ones involving discrepancies between the robots heading and the camera view that frequently lead to operational mishaps [6]. A tethered "camera" (B,C) provides an oblique view of the scene showing both the platform and its 3D environment. A 3rd person fixed view (C) is akin to an operator's view controlling slot cars and has been shown effective in avoiding roll-overs and other teleoperation accidents [4] but can't be used anywhere an operator's view might be obstructed such as within buildings or in rugged terrain. The tethered view (B) in which a camera "follows" an avatar (think Mario Brothers©) is widely favored in virtual environments [7,8] for its ability to show the object being controlled in relation to its environment by showing both the platform and an approximation of the scene that might be viewed from a camera mounted on it. This can be simulated for robotic platforms by mounting a camera on a flexible pole giving the operator a partial view of his platform in the environment [9]. Because of restriction in field of view and the necessity of pointing the camera downward, however, this strategy is of little use for surveying a scene although it can provide a view of the robot's periphery and nearby obstacles that could not be seen otherwise. The exocentric views show a 2 dimensional version of the scene such as might be provided by an overhead camera and cannot be obtained from an onboard camera. This type of "overhead" view can, however, be approximated by a map. For robots equipped with laser range finders, generating a map and localizing the robot on that map provides a method for approximating an exocentric view of the platform. If this view rotates with the robot (heading up) it is a type D plan view. If it remains fixed (North up) it is of type E.

An early comparison at Sandia Laboratory between viewpoints for robot control [4] investigating accidents focused on the most common of these: (A) egocentric from onboard camera and (C) 3rd person. The finding was that all accidents involving rollover occurred under egocentric control while 3rd person control led to bumping and other events resulting from obstructed or distanced views. In current experimental work in remotely controlled robots for urban search and rescue (USAR) robots are typically equipped with both a ptz video camera for viewing the environment and a laser range finder for building a map and localizing the robot on that map. The video feed and map are usually presented in separate windows on the user interface and intended to be used in conjunction. While Casper and Murphy [10] reporting on experiences in searching for victims at the World Trade Center observed that it was very difficult for an operator to handle both navigation and exploration of the environment from video information alone, Yanco and Drury [9] found that first responders using a robot to find victims in a mock environment made little use of the generated map. One possible explanation is that video is simply more attention grabbing than other presentations [11] leading operators to control primarily from the camera while ignoring other information available on their interface. A number of recent studies conducted by Goodrich, Neilsen, and colleagues [12,14,15,16] have attempted to remedy this through an ecological interface that fuses information by embedding the video display within the map. The resulting interface takes the 2D map and extrudes the identified surfaces to derive a 3D version resembling a world filled with cubicles. The robot is located on this map with the video window placed in front of it at the location being viewed. This strategy uses the egocentric camera view and the overhead view from the map to create a synthetic tethered view of the sort found most effective in virtual environments and games [7,8]. The anticipated advantages, however, have been difficult to demonstrate with ecological and conventional interfaces trading advantages across measures. Of particular interest have been comparisons between control based exclusively on maps or videos. In complex environments with little opportunity for preview, maps were found superior [14] in assisting operators to escape from a maze.

When considering such potential advantages and disadvantages of viewpoints it is important to realize that there are two, not one, important subtasks that are likely to engage operators [8]. The escape task and the accidents reviewed at Sandia involved *navigation*, the act of explicitly moving the robot to different locations in the environment. In many applications *search*, the process of acquiring a specific viewpoint—or set of viewpoints—containing a particular object may be of greater concern. While both navigation and search require the robot to move, an important distinction is the focus of the movement. Navigation occurs with respect to the environment at large, while search references a specific object or point within that environment. Switching between these two subtasks may play a major role in undermining situation awareness in teleoperated environments. For example, since search activities move the robot with respect to an object, viewers may lose track of their global position within the environment. Additional maneuvering may be necessary to reorient the operator before navigation can be effectively resumed. Because search relies on moving a viewpoint through the environment to find and better view target objects, it is an inherently egocentric task. This is not necessarily the case for navigation which does not need to identify objects but only to avoid them.

Search, particularly multi-robot search, presents the additional problem of assuring that areas the robot has traversed have been thoroughly searched for targets. This requirement directly conflicts with the navigation task which requires the camera to be pointed in the direction of travel in order to detect and avoid objects and steer toward its goal. When the operator attempts to compromise by choosing a path to traverse and then panning the camera to search as the robot moves he runs both the risk of hitting objects while he is looking away and missing targets as he attends to navigation. For multirobot control these difficulties are accentuated by the need to switch attention among robots multiplying the likelihood that a view containing a target will be missed. In earlier studies [17,18] we have demonstrated that success in search is directly related to the frequency with which the operator shifts attention between robots over a variety of conditions. An additional issue is the operator's confidence that an area has been effectively searched. In our natural environment we move and glance about to construct a representation of our environment that is informed by planning and proprioception that knit together the sequence of views. In controlling a robot we are deprived of these natural bridging cues and have difficulty recognizing as we pan and tilt whether we are resampling old views or missing new ones. The extent of this effect was demonstrated by Pausch [19] who found that participants searching for an object in a virtual room using a headmounted display were twice as fast as when they used a simulated handheld camera. Since even the handheld camera provides many ecological cues we should expect viewing from a moving platform through a ptz camera to be substantially worse.

A. Asynchronous Imagery

To combat these problems of attentive sampling among cameras, incomplete coverage of searched areas, and difficulties in associating camera views with map locations we are investigating the potential of asynchronous control techniques previously used out of necessity in NASA applications as a solution to multi-robot search problems. Due to limited bandwidth and communication lags in interplanetary robotics camera views are closely planned and executed. Rather than transmitting live video and moving the camera about the scene, photographs are taken from a single spot with plans to capture as much of the surrounding scene as possible. These photographs taken with either an omnidirectional overhead camera (camera faces upward to a convex mirror reflecting $360^{\circ)}$ and dewarped [20,21] or stitched together from multiple pictures from a ptz camera [22] provide a panorama guaranteeing complete coverage of the scene from a particular point. If these points are well chosen, a collection of panoramas can cover an area to be searched with greater certainty than imagery captured with a ptz camera during navigation. For the operator searching within a saved panorama the experience is similar to controlling a ptz camera in the actual scene, a property that has been used to improve teleoperation in a low bandwidth high latency application [23]. In our USAR application which requires finding victims and locating them on a map we merge map and camera views as in [15]. The operator directs navigation from the map being generated with panoramas being taken at the last waypoint of a series. The panoramas are stored and accessed through icons showing their locations on the map. The operator can find victims by asynchronously panning through these stored panoramas as time becomes available. When a victim is spotted the operator uses landmarks from the image and corresponding points on the map to record the victim's location. By changing the task from a forced paced one with camera views that must be controlled and searched on multiple robots continuously to a self paced task in which only navigation needs to be controlled in realtime we hoped to provide a control interface that would allow more thorough search with lowered mental workload. The reductions in bandwidth and communications requirements [12] are yet another advantage offered by this approach.

B. Pilot Experiment

In a recent experiment reported in [13] we compared performance for operators controlling 4 robot teams at a simulated USAR task using either streaming or asynchronous video displays. Search performance was somewhat better using the conventional interface with operators marking slightly more victims closer to their actual location at each degree of relaxation. This superiority, however, might have occurred simply because streaming video users had the opportunity to move closer to victims thereby improving their estimates of distance in marking the map. A contrasting observation was that frequency of shifting focus between robots, a practice we have previously found related to search performance [25] was correlated with performance for streaming video participants but not for participants using asynchronous panoramas. Because operators using asynchronous video did not need to constantly switch between camera views to avoid missing victims we hypothesized that for larger team sizes where forced pace search might exceed the operator's attentional capacity asynchronous video might offer an advantage. The present experiment tests this hypothesis.

II. EXPERIMENT

A. USARSim and MrCS

The experiment was conducted in the high fidelity USARSim robotic simulation environment [24] developed as a simulation of urban search and rescue (USAR) robots and environments intended as a research tool for the study of human-robot interaction (HRI) and multi-robot coordination. USARSim is freely available and can be downloaded from www.sourceforge.net/projects/usarsim. USARSim uses Epic Games' UnrealEngine2 to provide a high fidelity simulator at low cost. USARSim supports HRI by accurately rendering user interface elements (particularly camera video), accurately representing robot automation and behavior, and accurately representing the remote environment that links the operator's awareness with the robot's behaviors. MrCS (Multi-robot Control System), a multirobot communications and control infrastructure with accompanying user interface developed for experiments in multirobot control and RoboCup competition [17] was used with appropriate modifications in both experimental conditions. MrCS provides facilities for starting and controlling robots in the simulation, displaying camera and laser output, and supporting inter-robot communication through Machinetta [25] a distributed mutiagent system. The distributed control enables us to scale robot teams from small to large.

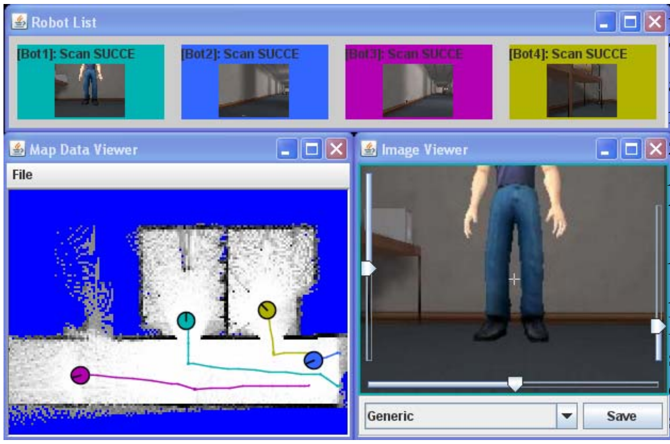

Figure 2. MrCS components for Streaming Video mode

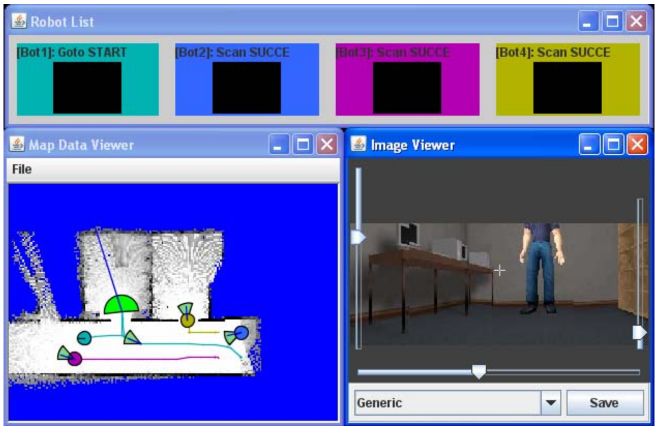

Figure 3. MrCS components for Asynchronous Panorama mode

Figures 2 and 3 show the elements of the MrCS involved in this experiment. In the standard MrCS (Fig. 2) the operator selects the robot to be controlled from the colored thumbnails at the top of the screen that show a slowly updating view from the robot's camera. Streaming video from the *in focus* robot which the operator now controls is displayed on the *Image Viewer*. To view more of the scene the operator uses pan/tilt sliders (not shown) to control the camera. Robots are tasked by assigning waypoints on a heading-up map on the *Mission Panel* (not shown) or through a teleoperation widget (not shown). The current locations and paths of the robots are shown on the *Map Data Viewer*. Although the experimental panoramic interface (Fig. 3) looks much the same it behaves quite differently. Robots are again selected for control from the colored thumbnails which now lack images. Panoramic images are acquired at the terminal point of waypoint sequences. Icons conveying the robot's location and orientation at these points are placed on the map for accessing the panoramas. The operator can then view stored panoramas by selecting an icon and dragging a mouse over the Image Viewer to move the image around or using the mouse's scroll wheel to zoom in and out of the image. The associated icon on the *Map Data Viewer* changes orientation in accordance with the part of the scene being viewed.

B. Method

A large search environment previously used in the 2006 RoboCup Rescue Virtual Robots competition [26] was selected for use in the experiment. The environment consisted of maze like halls with many rooms and obstacles, such as chairs, desks, cabinets, and bricks. Victims were evenly distributed throughout the environments. Robots were started at different locations leading to exploration of different but equivalent areas of the environment. A simpler environment was used for training. The experiment followed a between groups design with participants searching for victims using either panorama or streaming video modes. Participants searched over three trials beginning with 4 robots, then searching with 8, and finally 12. Robots were started from different locations within a large environment making learning from previous trials unlikely.

C. Participants and Procedure

29 paid participants were recruited from the University of Pittsburgh community. None had prior experience with robot control although most were frequent computer users. Approximately a quarter of the participants reported playing computer games for more than one hour per week.

After collecting demographic data the participant read standard instructions on how to control robots via MrCS. In the following 15~20 minute training session, the participant practiced control operations for either the panorama or streaming video mode and tried to find at least one victim in the training environment under the guidance of the experimenter. Participants then began three testing sessions in which they performed the search task controlling 4, 8, and 12 robots.

III. RESULTS

Data were analyzed using a repeated measures ANOVA comparing streaming video performance with that of asynchronous panoramas. On the performance measures, victims found and area covered, the groups showed nearly identical performance with victim identification peaking sharply at 8 robots accompanied by a slightly less dramatic maximum for search coverage (figure 4). Although the number of identified victims did not vary the differences in precision for marking victims observed in the pilot study were found again. For victims marked within 2m, (figure 5) the average number of victims found in the panorama condition was 5.36 using 4 robots, 5.50 for 8 robots, but dropping back to 4.71 when using 12 robots. Participants in the Streaming condition were significantly more successful at this range, $F_{1,29} = 3.563$, $p < .028$, finding 4.8, 7.07 and 4.73 victims respectively.

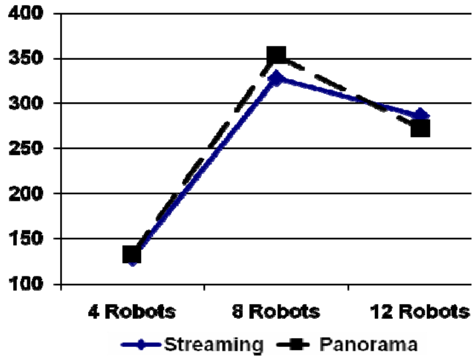

Figure 4. Area Covered

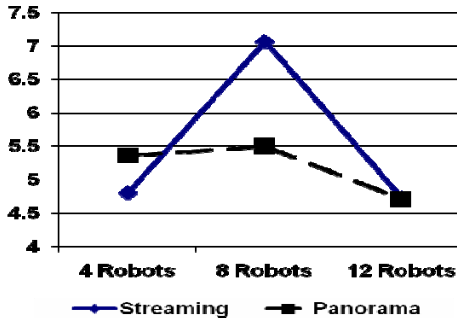

Figure 5. Victims Found as a function of N robots (2 m)

A similar advantage was found for victims marked within 1.5m (figure 6), with the average number of victims found in the panorama condition dropping to 3.64, 3.27 and 2.93 while participants in the streaming condition were more successful, $F_{1,29} = 6.255$, $p < .0025$, finding 4.067, 5.667 and 4.133 victims respectively.

Fan-out [27] is a model-based estimate of the number of robots an operator can control. While Fan-out was conceived as an invariant measure, operators are noticed to adjust their criteria for adequate performance to accommodate the available robots [28]. We interpret Fan-out as a measure of attentional reserves. If Fan-out is greater than the number of robots there are remaining reserves. If Fan-out is less than the number of robots, capacity has already been exceeded.

Fan-out for the panorama conditions increased from 4.1, 7.6 and 11.1 for 4 to 12 robots. Fan-out, however, was uniformly higher in the streaming video condition, $F_{1,29} = 3.355$, $p < .034$, with 4.4, 9.12 and 13.46 victims respectively (figure 7). Participants in the streaming condition appear to have cognitive reserves for controlling additional robots because computed Fan-out remains higher than the number they control. Panorama participants by contrast have Fan-out estimates below the N robots they are being asked to control for 8 and 12 robots.

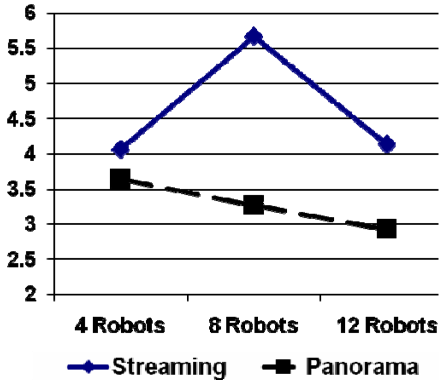

Figure 6. Victims Found as a function of N robots (1.5m)

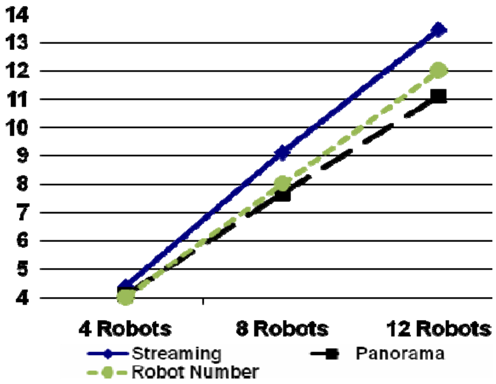

Figure 7. Fan-out as a function of N robots

IV. DISCUSSION

The most unexpected thing about these data is how similar the performance of streaming and asynchronous panorama participants appeared. The tasks themselves seem quite dissimilar. In the panorama condition participants direct their robots by adding waypoints to a map without getting to see the robots' environment directly. Typically they tasked robots sequentially and then went back to look at the panoramas that had been taken. Because panorama participants were unable to see the robot's surrounding except at terminal waypoints, paths needed to be shorter and contain fewer waypoints in order to maintain situation awareness and avoid missing potential victims. Despite fewer waypoints and shorter paths, panorama participants managed to cover the same area as streaming video participants within the same number of missions. Ironically, this greater efficiency may have resulted from the absence of distraction from streaming video [9] and is consistent with [14] in finding maps especially useful for navigating complex environments.

Examination of pauses in the streaming video condition failed to support our hypothesis that these participants would execute additional maneuvers to examine victims. Instead, streaming video participants seemed to follow the same strategy as panorama participants of directing robots to an area just inside the door of each room. This leaves panorama participants' inaccuracy in marking victims unexplained other than through a general loss of situation awareness. This explanation would hold that lacking imagery leading up to the panorama, these participants have less context for judging victim location within the image and must rely on memory and mental transformations.

Panorama participants also showed lower Fan-out perhaps as a result of issuing fewer waypoints for shorter paths leading to more frequent interactions.

While we undertook this study to determine whether asynchronous video might prove beneficial to larger teams we found performance to be essentially equivalent to the use of streaming video at all team sizes with a small sacrifice of accuracy in marking victims. This surprising finding suggests that in applications that may be too bandwidth limited to support streaming video or involve substantial lags, map-based displays with stored panoramas may provide a useful display alternative without seriously compromising performance.

REFERENCES

[1] T. Fong and C. Thorpe, "Vehicle teleoperation interfaces," *Autonomous. Robots*, no. 11, pp. 9–18, 2001.

[2] C. Wickens and J. Hollands, *Engineering Psychology and Human Performance*, Prentice Hall, 1999.

[3] R. Darken, K. Kempster, and B. Peterson, "Effects of streaming video quality of service on spatial comprehension in a reconnaissance task," in *Proc. Meeting I/ITSEC*, Orlando, FL, 2001.

[4] D. E. McGovern, "Experiences and Results in *Teleoperation of Land Vehicles*," Sandia Nat. Labs., Albuquerque, NM, Tech. Rep. SAND 90-0299, 1990.

[5] M. Lewis and J. Wang, "Gravity referenced attitude display for mobile robots : Making sense of what we see, " *Transactions on Systems, Man and Cybernetics Part A*, 37(1), 94-105, 2007.

[6] H. Yanco, M. Baker, R. Casey, B. Keyes, P. Thoren, J. Drury, D. Few, C. Nielsen, D. Bruemmer. "Analysis of human-robot interaction for urban search and rescue, *Proceedings of PERMIS, 2006*.

[7] P. Milgram and J. Ballantyne, "Real world teleoperation via virtual environment modeling," in *Proc. Int. Conf. Artif. Reality Tele-Existence*, Tokyo, 1997.

[8] D. S. Tan, G. G. Robertson, and M. Czerwinski, "Exploring 3D navigation: Combining speed-coupled flying with orbiting," in CHI 2001 *Conf. Human Factors Comput. Syst*., Seattle, WA, 2001.

[9] H. A. Yanco and J. L. Drury. "Where am I?" Acquiring situation awareness using a remote robot platform. In *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, October 2004.

[10] J. Casper and R. R. Murphy. Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center. *IEEE Transactions on Systems, Man, and Cybernetics Part B*, 33(3):367–385, June, 2003.

[11] R. Kubey and M. Csikszentmihalyi. Television addiction is no mere metaphor. *Scientific American*, 286(2):62–68, 2002.

[12] D. J. Bruemmer, D. A. Few, M. C. Walton, R. L. Boring, J. L. Marble, C. W. Nielsen, and J. Garner. Turn off the television: Real-world robotic exploration experiments with a virtual 3-D display. *Proc. HICSS*, Kona, HI, 2005.

[13] P. Velagapudi, J. Wang, H. Wang, P. Scerri, M. Lewis, and K. Sycara. Synchronous vs. Asynchronous Video in Multi-Robot Search, *Proceedings of first International Conference on Advances in Computer-Human Interaction* (ACHI'08), 2008

[14] C. Nielsen and M. Goodrich. Comparing the usefulness of video and map information in navigation tasks. In *Proceedings of the 2006 Human-Robot Interaction Conference*, Salt Lake City, Utah, March 2006.

[15] B. W. Ricks, C. W. Nielsen, and M. A. Goodrich. Ecological displays for robot interaction: A new perspective. In *International Conference on Intelligent Robots and Systems IEEE/RSJ*, Sendai, Japan, 2004.

[16] H. Yanco, M. Baker, R. Casey, B. Keyes, P. Thoren, J. Drury, D. Few, C. Nielsen, D. Bruemmer. "Analysis of human-robot interaction for urban search and rescue, *Proceedings of PERMIS, 2006*.

[17] J. Wang, and M. Lewis, "Human control of cooperating robot teams", *2007 Human-Robot Interaction Conference*, ACM, 2007a.

[18] J. Wang, and M. Lewis, "Assessing coordination overhead in control of robot teams," *Proceedings of the 2007 International Conference on Systems, Man, and Cybernetics*, 2007b

[19] R. Pausch, M. A. Shackelford, and D. Proffitt, "A user study comparing head-mounted and stationary displays," *IEEE Symposium on Research Frontiers in Virtual Reality*, October 25-26, 1993, San Jose, CA.

[20] J.R. Murphy, "Application of Panospheric Imaging to a Teleoperated Lunar Rover," in *Proceedings of the 1995 International Conference on Systems, Man, and Cybernetics*, 1995.

[21] N. Shiroma, N. Sato, Y. Chiu, and F. Matsuno, "Study on effective camera images for mobile robot teleoperation," In *Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication*, Kurashiki, Okayama Japan, 2004

[22] R. Volpe, "Navigation results from desert field tests of the Rocky 7 Mars rover prototype," in *The International Journal of Robotics Research*, 18, pp.669-683, 1999.

[23] M. Fiala, "Pano-presence for teleoperation," *Proc. Intelligent Robots and Systems (IROS 2005)*, 3798-3802, 2005.

[24] M. Lewis, J. Wang, and S. Hughes, "USARsim : Simulation for the Study of Human-Robot Interaction," *Journal of Cognitive Engineering and Decision Making*, 1(1), 98-120, 2007.

[25] P. Scerri, Y. Xu, E. Liao, G. Lai, M. Lewis, and K. Sycara, "Coordinating large groups of wide area search munitions," In D. Grundel, R. Murphey, and P. Pandalos (Ed.), *Recent Developments in Cooperative Control and Optimization*, Singapore: World Scientific, 451-480.

[26] S. Balakirsky, S. Carpin, A. Kleiner, M. Lewis, A. Visser, J. Wang, and V. Zipara, "Toward hetereogeneous robot teams for disaster mitigation: Results and performance metrics from RoboCup Rescue," *Journal of Field Robotics*, 24(11-12), 943-967, 2007.

[27] D. Olsen and S. Wood, "Fan-out: measuring human control of multiple robots", In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2004, ACM Press: Vienna, Austria. p. 231-238, 2004.

[28] H. Wang, M. Lewis, P.Velagapudi, P. Scerri and K. Sycara, K., " How search and its subtasks scale in N robots", *2009 Conference on Human-Robot Interaction* (HRI'09), ACM, 2009.

[29] C. Humphrey, G. Henk, B. Sewell, B. Williams, B. and J. Adams, J., "Assessing the Scalability of a Multiple Robot Interface," *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robotic Interaction, 2007..*