Ecological Interfaces for Improving Mobile Robot Teleoperation

Curtis W. Nielsen, *Member, IEEE*, Michael A. Goodrich, *Senior Member, IEEE*, and Robert W. Ricks

Abstract—Navigation is an essential element of many remote robot operations including search and rescue, reconnaissance, and space exploration. Previous reports on using remote mobile robots suggest that navigation is difficult due to poor situation awareness. It has been recommended by experts in human–robot interaction that interfaces between humans and robots provide more spatial information and better situational context in order to improve an operator's situation awareness. This paper presents an ecological interface paradigm that combines video, map, and robot-pose information into a 3-D mixed-reality display. The ecological paradigm is validated in planar worlds by comparing it against the standard interface paradigm in a series of simulated and real-world user studies. Based on the experiment results, observations in the literature, and working hypotheses, we present a series of principles for presenting information to an operator of a remote robot.

Index Terms—3-D interface, augmented-virtuality, human–robot interaction, information presentation, teleoperation, USAR-Sim, user study.

I. INTRODUCTION

NAVIGATION is an essential element of many remote robot operations including search and rescue, reconnaissance, and space exploration. Such settings provide a unique problem in that the robot operator is distant from the actual robot due to safety or logistical concerns. In order to operate a robot efficiently at remote distances, it is important for the operator to be aware of the environment around the robot so that the operator can give informed accurate instructions to the robot. This awareness of the environment is often referred to as *telepresence* [1], [2] or *situation awareness* [3], [4].

Despite the importance of situation awareness in remote robot operations, experience has shown that operators typically do not demonstrate sufficient awareness of the robot's location and surroundings [5], [6]. Many robots only provide video information to the operator, which creates a sense of trying to understand the environment through a "soda straw" or a "keyhole" [7], [8]. The limited view of the robot's environment makes it difficult for an operator to be aware of the robot's proximity to obstacles [9], [10]. Experiments with robots that have more sensing and operators with more familiarity with the robots have also shown that operators generally have a poor situation awareness [11]–[13].

One likely reason that operators demonstrated poor situation awareness in the previous studies is the way that conventional interfaces, which we refer to as 2-D interfaces, present information to the operator. Conventional 2-D interfaces present related pieces of information in separate parts of the display. This requires the operator to mentally correlate the sets of information, which can result in increased workload, decreased situation awareness, and decreased performance [4], [14]–[16]. From a cognitive perspective, these negative consequences arise because the operator must frequently perform mental rotations between different frames of reference (e.g., side views, map views, perspective views) and must fuse information even if the frames of reference agree.

To improve situation awareness in human–robot systems, Yanco *et al*. recommend 1) using a map; 2) fusing sensor information; 3) minimizing the use of multiple windows; and 4) providing more spatial information to the operator [17]. These recommendations are consistent with observations and recommendations from other researchers involved with human–robot interactions [5], [6], [18], [19].

In this paper, we address the recommendations for better interfaces by presenting an ecological interface paradigm as a means to improve an operator's awareness of a remote mobile robot. The ecological paradigm is based on Gibson's theory of affordances, which claims that information to act appropriately is inherent in the environment. Applying this theory to remote robotics means that an operator's decisions are made based on the operator's perception of the robot's affordances in the remote environment. The notion of effective presentation of information and ability to act on the information is also addressed by Endsley's definition of *situation awareness* [4], and Zahorik and Jenison's definition of *telepresence* [20] The ecological paradigm uses multiple sets of information from the robot to create a 3-D virtual environment that is augmented with real video information. This mixed-reality representation of the remote environment combines video, map, and robot-pose into a single integrated view of the environment. The 3-D interface is used to support the visualization of the relationships between the different sets of information. This representation presents the environment's navigational affordances to the operator and shows how they are related to the robot's current position and orientation. This paper proceeds as follows. Section II discusses previous work on technologies for improving mobile robot teleoperation. Section III presents the ecological interface paradigm and describes the 3-D interface. Section IV presents the summaries from new and previously published user studies that illustrate the usefulness of the 3-D interface in tasks ranging from robot control to environment search. Section V identifies the principles that governed the success of the 3-D interface

technologies in the user studies, while Section VI concludes the paper and summarizes the directions for future work.

II. PREVIOUS WORK

In this section, work related to improving robot teleoperation is presented. We will discuss the approaches based on robot autonomy and intelligence, followed by various modes of user interaction. We then present the notion of situation awareness and show that augmented virtuality can be applied to the human–robot interaction domain to improve the situation awareness of the operator.

A. Autonomy

One method to improve teleoperation is to use autonomy or intelligence on the robot. Some autonomy-based approaches to teleoperation include shared control [2], safeguarded control [21], [22], adjustable autonomy [23]–[26] and mixed initiatives [24], [27], [28]. One limitation of these approaches is that some control of the robot is taken away from the human. This limits the robot to the behaviors and intelligence that have been preprogrammed. There are situations where the operator may know more than the robot, and it is unlikely that the robot would be "designed" to handle every possible situation.

B. User Interaction

Fong observed that there would always be a need for human involvement in vehicle teleoperation despite intelligence on the remote vehicle [29]. Sheridan holds similar thoughts, and used the notion of supervisory control to explain how the human should be "kept in the loop" of the control of the robot [2] regardless of the level of autonomy of the robot.

There are many approaches for interacting with a robot, including gestures [30], [31], haptics [32]–[34], web-based controls [35], [36], and personal digital assistants (PDAs) [37]–[39]. Fong and Murphy addressed the idea of using dialog to reason between an operator and a robot when the human or robot needs more information about a situation [40], [41]. Most of these approaches tend to focus on different ways of interacting with a robot, as opposed to identifying when the approaches could be useful. In comparison, we are interested in helping the operator gain an awareness of the environment around the robot by identifying the information needs of the operator. In similar light, Keskinpala and Adams implemented an interface on a personal digital assistant (PDA) that combined sensed and video information and tested it in comparison to video-only and sensor-only interfaces in a robot control task [38].

C. Situation Awareness

In remote robot tasks, poor situation awareness has been identified as a reason for operator confusion in robot competitions [13], [17] and urban search and rescue training [6]. In fact, for the urban search and rescue domain, Murphy suggests that "more sophisticated mobility and navigation algorithms without an accompanying improvement in situation awareness support can reduce the time spent on a mission by no more than 25 percent" [19].

In her seminal paper, Endsley defines situation awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [4]. Additionally, Dourish and Bellotti define awareness as "...an understanding of the activities of others, which provides a context for your own activity" [42]. When applied to human–robot interactions, these definitions imply that a successful interaction is related to an operator's awareness of the activities and consequences of the robot in a remote environment. Endsley's work has been used throughout many fields of research that involve humans interacting with technology [3], [43], [44] and has been fundamental for exploring the information needs of a human operating a remote robot.

D. Interfaces

To enhance an operator's situation awareness, effort has gone into improving the visual experience afforded by human operators. One method is to use a panospheric camera [45]–[48], which gives a view of the entire region around the robot. An alternative to panospheric cameras is to use multiple cameras [49]–[51]. These approaches may help operators better understand what is all around the robot, but they require fast communications to send the large images with minimal delay. We are restricting attention to robots with a single camera. Other methods, which have been used to improve interfaces for teleoperation, include multisensor, sensor fusion, and adjustable autonomy interfaces [29], [46], [52], [53].

Yet another way to enhance the display for teleoperation is to use virtual reality (VR) to create a sense of presence. For example, Nguyen et al. use a VR interface for robot control by creating a 3-D terrain model of the environment from stereo images in order to present a terrain map of the surrounding landscape to the operator [54]. Moreover, information from the Mars Pathfinder was analyzed with a VR interface [55]. Similar to virtual reality are mixed reality and augmented reality [56], [57], which differ from VR in that the virtual environment is augmented with information from the real world. Milgram developed a system, which overlays a video stream with virtual elements such as range and obstacles with the intent of making the video information more useful to the operator [58]. Virtual reality-based interfaces can use a virtual environment to display information about robots in an intuitive way.

III. ECOLOGICAL PARADIGM

A. Background

Many of the terms used to describe robotic interfaces are defined in different ways by different people [59]. We operationally define teleoperation to be control of a robot, which may be at some distance from the operator [29]. Additionally, we operationally define telepresence as understanding an environment in which one is not physically present. This definition of telepresence is similar to Steuer's definition [60], which allows telepresence to refer to a "real" environment or a "nonexistent virtual world." This definition is less restrictive than Sheridan's definition [1] because one does not have to feel as though one
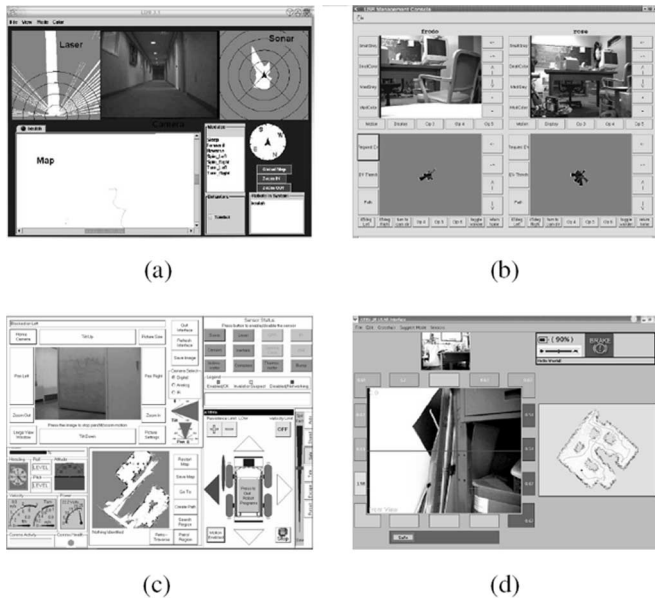
Fig. 1. Interfaces in the standard paradigm present information in separate windows within the display. (a) Our 2-D interface. (b) Adopted from [17]. (c) Adopted from [64]. (d) Adopted from [65].

is physically present at the remote site. Another definition of telepresence is discussed by viewing reality as not "outside" people's mind, but as a social construct based on the relationships between actors and their environments as mediated by artifacts [61]. Similar discussions on definitions exist for virtual presence [62], [63] and situation awareness [3], [4]. Telepresence is important because many believe that increased telepresence will increase performance on various tasks. The real problem with the definitions for telepresence is that they focus on the accuracy with which an environment is presented instead of focusing on communicating effective environmental cues. This has led to the use of displays such as those shown in Fig. 1, which show accurate information from the environment, but the information is presented in a diffused manner rather than in integrated form. The disparate information requires the operator to mentally combine the data into a holistic representation of the environment.

In contrast to the standard interface, our displays are based on Gibson's *ecological* theory of visual perception [66]. Gibson contends that we do not construct our percepts, but that our visual input is rich and we perceive objects and events directly. He claims that the information that an agent needs to act appropriately is inherent in the environment and not based on inferences of perceptions. Affordances embody the correlation between perception and action. In his words, "the *affordances* of the environment are what it *offers* animals, what it *provides* or *furnishes* either for good or ill" (emphasis in original). In other words, affordances eliminate the need to distinguish between real and virtual worlds because valid perception is one that makes successful action in the environment possible [66]. Zahorik and Jenison similarly observed that "presence is tantamount to successfully supported action in the environment" [20]. In order to support action in an environment far from a robot operator, it is important to convey the affordances of the environment

to the operator such that the operator's perceived affordances of the robot in the environment match the environment's true affordances [67].

B. 3-D Interface

Affordances are attractive to the robotics community because they are compatible with the reactive-based robot paradigm, and they simplify the computational complexity and representational issues [68]. With Gibson's ecological approach, successful human–robot interaction implies that the operator should be able to directly perceive the cues from the environment that support the actions of the robot.

To facilitate the operator's perception of the environmental cues and the robot's affordances within the environment, we implement a 3-D augmented virtuality interface. Augmented virtuality is a form of mixed reality [69] that refers to virtual environments, which have been enhanced or augmented by inclusion of real-world images or sensations. Augmented virtuality differs from virtual environments due to the inclusion of real-world images, and it differs from augmented reality (another form of mixed reality) because the basis of augmented virtuality is a virtual environment, as opposed to the real world in augmented reality [70]. In essence, our goal is to design an interface that implements Gibson's theory of perception by facilitating the direct perception of robot affordances. This will be done by supplying the operator with not only a visualization of information from the robot but also an illustration of the relationships between distinct sets of information, and how the information affects the possible actions of the robot.

The framework for the 3-D interface is a virtual environment that is based on a map or sensor readings of the robot's environment. For navigation tasks, the important environment cues are obstacles and open space, which are detected by the robot and saved using range sensors and a simultaneous localization and map-building (SLAM) algorithm. The map of the environment is placed on the floor of the virtual environment, and obstacles in the map are rendered with a heuristically chosen height to illustrate to the operator navigable and impassable areas and to provide depth cues. A 3-D model of the robot is rendered in the virtual environment at the position and orientation of the robot with respect to the map of the environment. The size of the robot model is scaled to match the scale of the virtual environment. The virtual environment is nominally viewed by the operator from a position a short distance above and behind the robot such that some map information is visible on all sides of the robot as illustrated in Fig. 2, but this virtual perspective can be changed as needed for the task. In congruence with Gibson's theory of affordances, this presentation of the robot information allows the operator to immediately perceive the possible actions of the robot within its remote environment.

For exploration tasks, the important environment cues also include video information, as well as the orientation of the camera with respect to the robot and environment. To facilitate the operator's perception of the video information, the video image is displayed in the virtual environment as the information relates to the orientation of the camera on the robot. This is done by
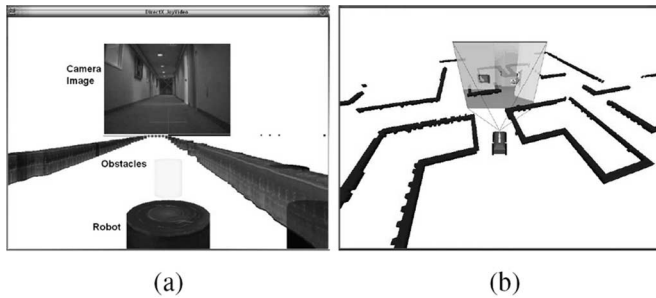
Fig. 2. Ecological paradigm combines information into a single integrated display. (a) Raw range data. (b) Map data.

rendering the video on a panel at a heuristically chosen distance from the robot and an orientation that corresponds with the orientation of the camera on the physical robot such that obstacle information in the video is spatially similar to the corresponding information from the map. As the camera is panned and tilted, the representation of the video moves in 3-D around the model of the robot accordingly.

IV. EXPERIMENTS

To validate the utility of the 3-D interface, it is important to compare its usefulness with that of a traditional 2-D interface. In this section, we summarize a series of user studies, which validate the 3-D interface in remote navigation tasks. The following user studies illustrate progressively more interesting and sophisticated navigation tasks. The tasks compare a prototypical 2-D interface with the 3-D interface and progress from basic robot control to environment search. The progression is best understood by presenting experiments and results from previous conference publications along with unpublished experiments. For each of the experiments, we will discuss the task, the approach for information presentation, the level of autonomy, the experiment design, the dependent measures, and the results. All of the experiments are counter-balanced to minimize learning effects, and the results are significant with $p < 0.05$ according to a double-sided t-test.

A. Robot Control

The most basic skill relevant to performing a search task with a mobile robot is the ability to remotely control the robot along a predetermined path. The purpose of this experiment is to compare how well an operator can perform this task with a traditional 2-D interface and an ecological 3-D interface. The simulated environment was in the form of a maze with a few different paths that could be taken to reach the goal destination. In this section, we summarize the most relevant results from [71].

1) Information Presentation: The operator is shown a representation of the robot in a virtual world of obstacles, which represent range data from the sonar sensors and the laser range-finder. The operator's perspective of the virtual world is from a tethered position, a little above and behind the robot. Included in the display is the most recently received image from the

robot's camera. Time delay is addressed through a *quickening* algorithm, which allows the operator to see the effects of their actions right away. Quickening is accomplished by moving the camera and the robot through the virtual world in response to the measured delay in communications. A precise description of the quickening algorithm and interface technology is provided in [71].

2) Autonomy in Safeguarding: The robot takes initiative to prevent collisions; no map-building.

3) Experiment Design: This experiment was setup as a within-subjects user study where each participant used both the 2-D and 3-D interface to follow the predetermined paths of varying difficulty. Thirty two subjects participated in the experiment with simulated robots and environments, using a home-built simulator that emulated a Pioneer 2 DXe. An additional eight subjects used a real Pioneer 2 DXe robot (with camera, laser range finder, sonar, and in-house control software) in an empty laboratory environment that was filled with cardboard boxes and was more than 700 m from the operator. The display that the test subjects used first and the order of the mazes was chosen randomly, but with the constraint that approximately the same number of people would be included in each group. The operator was informed of the route to follow through visual and audible cues.

4) Dependent Measures: The experiment depended on completion time, number of collisions, and workload (NASA-TLX and behavioral entropy).

5) Results: The results from the experiments found that in simulation, the operators finished the task 15% faster ($\bar{x}_{3D} = 212$ s, $\bar{x}_{2D} = 249$ s, $p = 8.6 \times 10^{-6}$) with 87% fewer collisions ($x_{3D} = 30$, $x_{2D} = 237$, $p = 2.2 \times 10^{-4}$) when using the 3-D interface in comparison to the 2-D prototype interface. Similarly, with the physical robot, the operators finished the task 51% faster ($\bar{x}_{3D} = 270$ s, $\bar{x}_{2D} = 553$ s, $p = 4.5 \times 10^{-3}$) with 93% fewer collisions ($x_{3D} = 6$, $x_{2D} = 83$, $p = 5.5 \times 10^{-3}$) when using the 3-D interface. The workload was also reduced significantly as measured subjectively with NASA-TLX [72] and as measured objectively with behavioral entropy [73]. These results suggest that it was easier, safer, and faster to guide the robot along a predetermined route with the 3-D interface than with the 2-D interface.

B. Spatial Coverage and Navigation

Often, in remote robot exercises, the physical structure of the environment is unknown beforehand and must be discovered by the robot. The purpose of this experiment was to determine how quickly and safely participants could discover the physical structure of an environment using simplified versions of the 2-D and 3-D interfaces. The simulated environment was an open room with various walls and obstacles, which had to be circumnavigated. This navigation-based task included recognizing where the robot had and had not visited and planning routes to unexplored areas.

For this and subsequent experiments, a map of the environment was not provided *a priori*, rather a SLAM algorithm was used by the robot to incrementally build a map of the
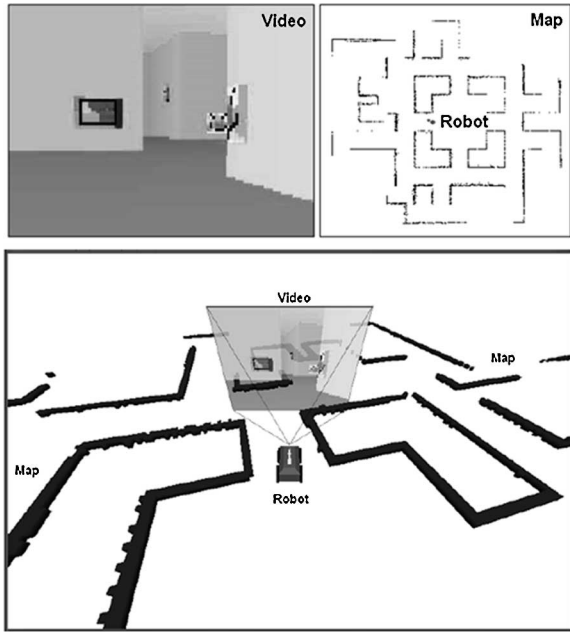
Fig. 3. 2-D prototype interface (top) and the 3-D prototype interface (bottom) used for the map-building experiment.

Fig. 4. Map of one of the mazes used in the sensor usage for navigation experiment.

environment as the robot traversed the environment. For experiments in simulation, the SLAM algorithm is based on perfect information from the simulator, and in real world experiments, we use Konolige's SLAM algorithm [74].

1) Information Presentation: To minimize distracting sets of information, the 2-D and 3-D interfaces were simplified such that only video, map, and robot pose were displayed as shown in Fig. 3. The operator's perspective of the 3-D interface was presented from above and behind the robot such that some of the map information behind the robot was also visible. Time delay was not addressed in this experiment because the simulator was on the same computer as the interface and the communications delay was insignificant.

2) Autonomy in Teleoperation: The robot does not take the initiative to avoid a collision; incremental map-building algorithm.

3) Experiment Design: The experiment was setup as a between-subjects user study where each participant used either the 2-D or the 3-D interface and a home-built robot simulator that emulated the Pioneer 2DXe robot. The experiment took place as a special exhibit in "Cyberville" at the St. Louis Science Center, where participants consisted of visitors from local high schools and colleges. Thirty participants performed the experiment with the 3-D interface and 30 participants used the 2-D interface.

4) Dependent Measures: The experiment depended on completion time, average robot speed, number of collisions, and proximity to obstacles.

5) Results: In this experiment, there were many instances when an operator drove the simulated robot into a wall and was unable to extricate the robot and, therefore, unable to complete the map-building task. Of the participants, 9 (30%) were unable to complete the task with the 3-D interface and 17 (57%) were unable to complete the task with the 2-D interface. Of
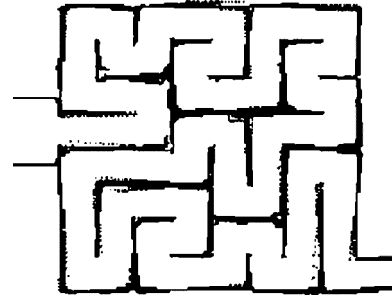
the participants who completed the task, those who used the 3-D interface finished 34% faster ($\bar{x}_{3D} = 178$ s, $\bar{x}_{2D} = 272$ s, $p = 3.4 \times 10^{-4}$) and had 66% fewer collisions ($\bar{x}_{3D} = 5.1$, $\bar{x}_{2D} = 14.9, p = 2.6 \times 10^{-4}$) than those who used the 2-D interface. Since collisions only measure actual impact with obstacles, and not any near misses, the average distance from the robot to the nearest obstacle was also measured. It was found that with the 3-D interface, the average distance to the walls was 16% greater than when the 2-D interface was used ($\bar{x}_{3D} = 0.85$ m, $\bar{x}_{2D} = 0.74$ m, $p = 4.2 \times 10^{-3}$). These results show that operators using the 3-D interface completed the task more efficiently than operators using the 2-D interface.

C. Sensor Usage for Navigation

Anecdotal evidence from pilot studies and the previous user studies revealed that the operators tended to focus a lot of their attention on the video information while driving the robot with the 2-D interface. The goal of this previously published experiment was to test the relative usefulness of the video and map information with 2-D and 3-D interfaces in a navigation task [75]. The task was to get the robot through a maze as fast as possible while avoiding collisions with walls. The simulated maze had 2 m wide hallways, covered a 256 m^2 area, and consisted of a starting location and a single path to the end location. There were six different mazes used for the experiment, and each of them had the same dimensions and the same number of turns (42) and straight portions (22) to minimize the differences in results from different mazes. A map of one of the environments is shown in Fig. 4.

1) Information Presentation: The operator's perspective of the 3-D interface was somewhat higher than the previous studies so that the operator could see more of the maze environment around the robot. Furthermore, depending on the task, different sets of information were presented on the interface (e.g., map-only, video-only, map + video). Time delay was not addressed in this experiment.

2) Autonomy: In simulation: teleoperation, incremental map-building algorithm. In the real world: safeguarding, incremental map-building algorithm.

3) Experiment Design: The experiment was setup as a 2×3 within-subjects user-study, where each operator performed one test with each of the three conditions (map-only, video-only, map + video) for both interfaces (2-D, 3-D). The conditions were presented in a random order with the constraints that the 2-D and

TABLE I
COMPARISON OF THE VARIOUS CONDITIONS IN THE SIMULATION PORTION OF
THE SENSOR USAGE FOR NAVIGATION EXPERIMENT

	Time to Completion (mean/stdev)	Collisions (mean/stdev)
2D map-only	258 / 57	9.8 / 7.8
2D map+video	271 / 55	8.5 / 4.6
2D video-only	366 / 118	19.1 / 10.2
3D map-only	196 / 28	1.3 / 2.2
3D map+video	208 / 34	1.3 / 1.8
3D video-only	351 / 100	22.7 / 14.4

TABLE II
COMPARISON OF THE VARIOUS CONDITIONS IN THE REAL-WORLD PORTION OF
THE SENSOR USAGE FOR NAVIGATION EXPERIMENT

	Time to Completion (mean/stdev)	Collisions (mean/stdev)
2D map-only	319 / 102	46. / 27.9
2D map+video	247 / 54	36.3 / 17.4
video-only	243 / 59	38.8 / 6.3
3D map+video	205 / 47	24.8 / 13.5
3D map-only	227 / 48	28.6 / 20.1

3-D interfaces were used alternately, and the interface conditions were counter-balanced in the order they were used. Twenty four participants performed the experiment in simulation, and 21 participants performed the experiment in the real world.

The simulation portion of this experiment made use of the USARSim simulator [76], which provides more realistic images than the previous in-house simulator and is better for studying the utility of video for navigation. The simulated robot was an ATRV-Jr. The real-world portion of this experiment took place in the halls of the second floor of the Computer Science Department, Brigham Young University. The real-world experiment utilized an ATRV-Jr robot developed by IRobot that implements communications and safeguarding algorithms developed by Idaho National Laboratory [64], [77] and map-building algorithms developed by Stanford Research Institute [74]. The safeguarding algorithm moderates the maximum velocity of the robot through an event-horizon calculation, which estimates the time-to-collision with sensed obstacles [78]. When the robot is too close to an obstacle, movement in the direction of the obstacle is inhibited. Both the real and simulated robot had a pan-tilt-zoom (PTZ) camera, laser range-finder, and sonar sensors.

4) Dependent Measures: The experiment depended on completion time and the number of collisions.

5) Results: For this experiment, we present a summary of the results; for a detailed discussion, refer to [75]. The results from this experiment show that in simulation with the 2-D interface, the operators finished the task fastest with the map-only condition and the slowest with the video-only condition. When the map and video were combined, the performance was faster than the video-only condition but slower than the map-only condition. This suggests that the video was not very helpful and distracted the operator's attention away from the map that was probably the more useful piece of information, at least for this navigation task.

With the 3-D interface, the operators had results similar to the 2-D interface except that when the map and the video information were combined, it did not negatively affect the task completion times. In contrast, the map-only condition and the map + video conditions had similar times to completion and collisions. This suggests that although the video did not have very useful navigational information, it did not adversely affect the navigation of the robot when combined with the map. In summary, the 3-D map-only and 3-D map + video conditions performed the best, followed by the 2-D map-only then the 2-D map + video condition. The worst conditions were the 3-D video and 2-D video conditions, which had comparable results. The results from the simulation experiment are summarized in Table I.

In the real-world portion of the experiment, the participants did not use the 3-D video condition since the interface and results were similar to the 2-D video condition in the simulation portion of the study. In the real-world experiment, the video-only condition supported task completion. By comparison, the 2-D map condition took much longer to complete the task than the video-only condition. When the 2-D map + video condition was used, the completion time was the same as that in the video-only condition. With the 3-D interface, the map information was helpful, and the map-only and video-only conditions had similar completion times. When the 3-D map + video condition was used, the performance was even better compared to when only the video or only the 3-D map was available. In summary, the best condition was the 3-D map + video and the worst condition was the 2-D map. The rest of the conditions (3-D map-only, video-only, 2-D map + video) performed similarly. The results from the real-world portion of the experiment are summarized in Table II.

This experiment suggests that having both map and video available does not mean that they will automatically support each other. One hypothesis is that with the 2-D interface, the different sets of information "compete" for the attention of the operator. This competition resulted in, at best, no improvement in performance when the multiple sets of information were used, and, at worst, an actual decrease in performance. In contrast, with the 3-D interface, the different information sets seemed to complement each other. This synergy led to better performance with both map and video than with only map or only video. This hypothesis of competing and complementary sets of information is an area that needs to be studied further.

By way of comparison, it was found that operators with the 3-D interface and the map-only and map + video conditions completed the tasks on average 23% faster with at least 85% fewer collisions than the 2-D counterparts.

D. Navigation in the Presence of Delay

For the next experiment, we revisit the challenge of communications delay between the operator and the robot. The purpose of this experiment was to compare the effects of minor delay on a navigation task when the 2-D and 3-D interfaces are used. The task was to get the robot through a maze as fast as possible while avoiding collisions with walls. The simulated mazes for this experiment were the same as those in the previous experiment.

1) Information Presentation: The interfaces for this experiment were the same as those in the previous experiment, i.e., video, map, and robot pose were available. Although this

TABLE III
COMPLETION TIMES FOR THE DELAY EXPERIMENT

	Delay Condition		
	0-seconds	0.5-seconds	1-second
2D Interface	302s	422s	578s
3D Interface	221s	311s	466s
% Change	-27%	-26%	-19%
p-values	5.0×10^{-5}	2.4×10^{-4}	2.3×10^{-3}

TABLE IV
AVERAGE VELOCITY FOR THE DELAY EXPERIMENT

	Delay Condition		
	0-seconds	0.5-seconds	1-second
2D Interface	0.46m/s	0.38m/s	0.31m/s
3D Interface	0.60m/s	0.47m/s	0.36m/s
% Change	30%	22%	18%
p-values	3.7×10^{-4}	1.4×10^{-3}	1.9×10^{-1}

TABLE V
AVERAGE COLLISIONS IN THE DELAY EXPERIMENT

	Delay Condition		
	0-seconds	0.5-seconds	1-second
2D Interface	10.6	22.7	38.6
3D Interface	1.7	8.1	28.4
% Change	-84%	-65%	-27%
p-values	9.9×10^{-4}	6.8×10^{-3}	1.2×10^{-1}

Fig. 5. Simulation environments used in the St. Louis Science Center exploration tasks.

experiment compared the effect of minor delay on navigation, no quickening or predictive algorithms were used to support the operator in the presence of time delay. Rather, when the operator issued a command, the representation would not reflect the given command until the delay condition had elapsed.

2) Autonomy: Teleoperation, incremental map-building algorithm.

3) Experiment Design: The experiment was setup as a 2×3 within-subjects user study where each operator performed one test with each of the three delay conditions (0, 0.5, and 1 s) for both interfaces (2-D, 3-D). The conditions were presented in a random order with the constraints that the 2-D and 3-D interfaces were used alternately and the interface conditions were counter-balanced in the order they were used. This experiment was performed with the USARSim simulator since it was anticipated that the communications delay would significantly hinder the operator's ability to maintain control of the robot. The simulator implemented the ATRV-JR robot. Eighteen volunteers participated in the experiment.

4) Dependent Measures: The experiment depended on completion time, number of collisions, and average velocity.

5) Results: The results from this experiment show that the operators were able to finish the task 27%, 26%, and 19% faster with the 3-D interface than with the 2-D interface for delays of 0, 0.5, and 1 s, respectively. In fact, when the 3-D interface had 0.5-s more delay than the 2-D interface, the completion time was about the same. Furthermore, ten participants finished the task faster with the 3-D 0.5-s condition than the 2-D 0-s condition, and six finished the task faster with the 3-D 1-s condition than the 2-D 0.5-s condition. Table III summarizes the average completion time for the various conditions.

The operators averaged faster velocities with the robot when using the 3-D interface in comparison to the 2-D interface as shown in Table IV. It is to be noted that the average velocity with the 3-D interface and 0.5-s delay is similar to the 2-D interface and 0-s delay and, similarly, the 3-D interface with 1-s delay has an average velocity similar to the 2-D interface with 0.5-s delay.

There was also an 84%, 65%, and 27% decrease in collisions with the 3-D interface in comparison to the 2-D interface for the 0-, 0.5-, and 1-s conditions, respectively (see Table V).

These results show that the 3-D interface is consistently better than the 2-D interface across multiple levels of minor delay.

Additionally, the 2-D interface has results similar to the 3-D interface with an additional 0.5-s of delay. This suggests that the operator is better able to anticipate how the robot will respond to commands amidst minor network latency with the 3-D interface than with the 2-D interface. These results are consistent with results from the first experiment, which had 1-s delay. In that experiment, quickening of the robot's position amidst the obstacles was used because the obstacles were based on current sensor readings without a global map, and errors in the estimate could easily be corrected. In the future, it would be valuable to apply quickening to a map-based display.

E. Payload Management and Navigation

The previous experiments focused on navigating the robot through environments. Next, we will summarize experiments where a navigation task is augmented with payload control [79]. Specifically, a PTZ camera is manipulated while navigating the robot. This is a particularly challenging navigation problem because it is often difficult to navigate the robot while operating the camera especially when the video information is not centered in front of the robot.

The purpose of this experiment is to compare the usefulness of a PTZ camera against a stationary camera with both the 2-D and 3-D interfaces. The task for the operator was to drive the robot around a simple maze environment that contained numerous intersections with dead-end hallways as shown in Fig. 5. At the end of some of the hallways were flags that the operator was asked to look for.

1) Information Presentation: The operators used either the 2-D interface or the 3-D interface and either the stationary camera or the PTZ camera. The perspective of the 3-D interface was a little lower than the previous experiments and further behind the robot so that when the camera was moved from side to side, it was still completely visible within the interface and had minimal skew, as would have been observed from a higher or closer perspective. Time delay was not addressed in this experiment.

2) Autonomy: Teleoperation, incremental map-building algorithm.

3) Experiment Design: The experiment was setup as a 2×2 between-subjects user study, where each participant used one of the following conditions: 2-D-PTZ; 2-D-stationary; 3-D-PTZ; 3-D-stationary with our in-house simulator. The simulator implemented the Pioneer 2 DXe robot. The experiment took place as a special exhibit in "Cyberville" at the St. Louis Science Center, where participants consisted of visitors from local high schools and colleges. Forty four volunteers participated in each of the conditions.

4) Dependent Measures: The experiment depended on completion time, average velocity, distance covered by robot, number of collisions, and qualitative robot path differences.

5) Results: The results from the experiment show that with the 2-D interface, on average, the task was finished in the same amount of time, irrespective of whether the PTZ camera or stationary camera was used. With the stationary camera, a common behavior observed with the operators was to move the robot forward and deviate down each dead-end corridor before correcting and continuing along the main hallway. With the PTZ camera, the operators would basically stop the robot at each intersection, and then move the camera to the side to look down the hallway. Once the search was complete, they would recenter the camera and continue along the main path. Despite the different driving styles, the actual time to complete the task did not change because although the actual distance driven with the PTZ camera was smaller, there was an equal decrease in the average velocity.

With the 3-D interface, the task was finished faster with the PTZ camera than with the stationary camera. Even though the operators slowed the navigational speed of the robots with the PTZ camera, they generally did not stop moving the robot, nor did they necessarily recenter the camera before continuing along the path. This meant that less distance was traveled than with the stationary camera, but the average velocity did not drop as much as the change in distance. This resulted in a faster completion time.

On an average, the operators with the 3-D interface finished 27% faster with the stationary camera ($\bar{x}_{3D} = 181$ s, $\bar{x}_{2D} = 249$ s, $p = 4.3 \times 10^{-5}$) and 37% faster with the PTZ camera than operators with the 2-D interface ($\bar{x}_{3D} = 157$ s, $\bar{x}_{2D} = 250$ s, $p = 7.7 \times 10^{-7}$). Additionally, operators with the 3-D interface had 63% fewer collisions with the stationary camera ($\bar{x}_{3D} = 4.11$, $\bar{x}_{2D} = 11.1$, $p = 9.9 \times 10^{-3}$) and 91% fewer collisions with the PTZ camera than with the 2-D interface ($\bar{x}_{3D} = 0.56$, $\bar{x}_{2D} = 6.04$, $p = 1.9 \times 10^{-3}$) [79]. In a related study, it was found that the operators were able to issue 33% more PTZ commands per second with the 3-D interface than with the 2-D interface ($\bar{x}_{3D} = 3.7$ s, $\bar{x}_{2D} = 2.5$ s, $p = 1.0 \times 10^{-3}$) while still completing the task faster [80]. These results suggest that the 3-D interface supports the use of a PTZ camera more than the 2-D interface, at least in planar environments.

F. Environment Search

This final experiment was designed to put everything together into a search and identify task to see how well the 2-D interface
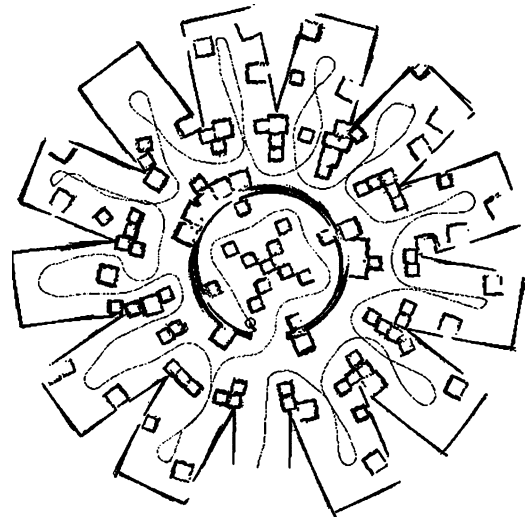

Fig. 6.   Map of the main floor of the simulation environment in the search experiment.

and 3-D interface compared to each other. The task was to explore an environment with the goal of finding and identifying as many things as possible.

1) Information Presentation: The 3-D interface was similar to the previous study. Although there were some communication delays seen in the real-world portions of this experiment, no quickening or predictive algorithms were used to support the operator. Rather, if there was delay, the representation would not change until the delay time had elapsed.

2) Autonomy: In simulation: teleoperation, incremental map-building algorithm. In the real world: safeguarding, incremental map-building algorithm.

3) Experiment Design: This experiment was designed as a 2×2 within-subjects user study, where each operator used both the 2-D and 3-D interfaces with both the USARSim simulator (ATRV-JR simulation) and the real ATRV-JR robot running the INL and SRI software (see Section IV-C). The real-world experiments were performed first, followed by the USARSim experiments. The display that was used first was chosen randomly with the constraint that an equal number of participants would start with each interface. Eighteen participants completed the experiment with both the real and simulated robots.

In simulation, the scenario was the exploration of an underground cave with areas of interest on three separate floors. The arena was shaped like a wheel with spokes (see Fig. 6), and at the end of each of the spokes or hallways, there was a cell that may or may not be occupied. The operators were required to identify if the cell was occupied, and if it was, they were to identify the color of the clothing of the person in the cell. In addition to the cells on the main floor, there were cells and occupants above and below the main floor. To view these other cells, the center of the environment was transparent, which allowed the operators to see above and below the robot's level when the camera was tilted up and down. Fig. 7 shows screen shots of the simulated environment, and Fig. 8 shows a screen shot of the avatars used for the experiment. The participants were given a time limit of
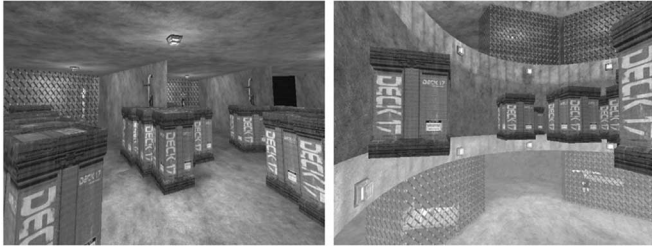
Fig. 7.   Images of the environment used for the simulation experiment.

Fig. 8.   3-D models for victims used in the simulated exploration experiment.

6 min and were asked to characterize as many cells as possible within the time limit.

The real-world portion of this experiment took place on the second floor of the Computer Science building at Brigham Young University. The physical environment was not as complex as the simulated environment, but still required the use of the PTZ camera to see and identify the items to the sides and above and below the center position of the camera. In this case, there were numerous objects of varying sizes hidden among Styrofoam and cardboard piles that were only visible and recognizable by manipulating the camera including the zoom capability. The participants were not given a time limit for the real-world portion of the experiment.

4) Dependent Measures: The experiment depended on the number of collisions, number of objects identified, time to identify, and completion time.

5) Results: The results show that in simulation, the operators were able to find and identify 19% more places with the 3-D interface ($\bar{x}_{3D} = 21.1$, $\bar{x}_{2D} = 18.1$, $p = 1.1 \times 10^{-2}$), and they had 44% fewer collisions ($\bar{x}_{3D} = 4.8$, $\bar{x}_{2D} = 8.6$, $p = 7.4 \times 10^{-3}$) with obstacles than when the 2-D interface was used. With the 3-D interface, three participants identified all the places within 6 min, whereas with the 2-D interface, no one identified all the places within the time limit. In the real-world experiments, there was no significant difference in time to complete the task or the total number of objects found; however, there was a 10% decrease in the average time spent identifying each object ($\bar{x}_{3D} = 40.0\,\text{s}$, $\bar{x}_{2D} = 44.5\,\text{s}$, $p = 2.8 \times 10^{-2}$).

This experiment shows that the 3-D interface supports a search task somewhat better than the 2-D interface. This is probably because the search task has a significant navigational component. One of the problems observed throughout the last

two studies was that it was difficult for many novice users to navigate the robot while controlling a PTZ camera with a joystick. In fact, sometimes it seemed that we were measuring thumb dexterity (for the PTZ controls) as opposed to task performance. An area of research that needs to be addressed in future work is how to navigate the robot while operating the robot's payload, in this case, the PTZ camera.

V. INFORMATION PRESENTATION PRINCIPLES

In an effort to understand why the 3-D interface supported performance better than the 2-D interface, we next present three principles that helped the 3-D interface overcome the previously observed limits to teleoperation to more closely match the theoretical limits on navigation. The principles are 1) present a common reference frame; 2) provide visual support for the correlation between action and response; and 3) allow an adjustable perspective. These principles relate to previous work in human–robot interaction [81], cognitive engineering [82], and situation awareness [83].

A. Common Reference Frame

When using mobile robots, there are often multiple sources of information that could be theoretically integrated to reduce the cognitive processing requirements of the operator. In particular, a mobile robot typically has a camera, range information, and some way of tracking where it has been. To integrate this information into a single display, a common reference frame is required. The common reference frame provides a place to present the different sets of information such that they are displayed in context of each other. In terms of Endsley's three levels of situation awareness [4], the common reference frame aids perception, comprehension, and projection. In the previous user studies, both the robot-centric and map-centric frames of reference were used to present the information to the operator.

1) Robot-Based Reference Frame: The robot itself can be a reference frame because a robot's sensors are physically attached to the robot. This is useful in situations where the robot has no map-building or localization algorithms (such as the experiment in Section IV-A) because the robot provides a context in which size, local navigability, etc., can still be evaluated. The reference frame can be portrayed by displaying an icon of the robot with the different sets of information rendered as they relate to the robot. For example, a laser range-finder typically covers 180° in front of the robot, the information of where the laser-detected obstacles could be presented as barrels placed at the correct distance and orientation from the robot (see Section IV-A). Another example is the use of a pan-tilt camera. If the camera is facing toward the front of the robot, then the video information should be rendered in front of the robot. If the camera is off-center and facing toward the side of the robot, the video should be displayed at the same side of the virtual robot (see Section IV-E). The key is that the information from the robot is displayed in a robot-centric reference frame.

2) Map-Based Reference Frame: There are many situations where a robot-centered frame of reference may not be appropriate. For example, the robot-centered frame of reference will

not be beneficial to represent two or more robots except under the degenerate conditions of them being collinear. Similarly, a robot-centered reference frame may not be useful for long-term path planning. If the robots have map-building and/or localization capabilities, an alternative reference frame could be map-based. With a map as the reference frame, different sets of information may be correlated even though they are not tied to a robot's current set of information. As an example, consider the process of constructing a map of the environment. As laser scans are made over time, the information is often combined with probabilistic map-building algorithms into an occupancy grid-based map [74], [84]. Updates to the map depend not only on the current pose of the robot, but on past poses as well. When the range scans of a room are integrated with the map, the robot can leave the room and the obstacles detected are still recorded because they are stored in relation to the map and not the robot. Mapping was used in all the experiments except the first one (Section IV-B–IV-F).

Another example of where a map can be useful as a common reference frame is with icons or snapshots of the environment. When an operator or a robot identifies a place and records information about it, the reference frame of the map provides a way to store the information as it relates to the map of the environment. Moreover, using a map as the reference frame also supports the use of multiple robots as long as they are localized in the same coordinate system. This means that places or things identified by one robot can have contextual meaning for another robot or an operator who has not previously visited or seen the location.

3) Reference-Frame Hierarchy: One advantage of reference frames is that they can be hierarchical. At one level, the information related to a single robot can be displayed from a robot-centric reference frame. At another level, the robot-based information from multiple robots can be presented in a map-based reference frame, which shows the spatial relationships between entities. Other reference frames include object-centered (something interesting in the environment, such as a landmark), manipulator-centered (improvised explosive device (IED) disposal), camera-centered (especially with a PTZ), and operator-centered (proprioception, sky-up, left and right). In the map-based reference frame, each robot still maintains and presents its own robot-centric information, but now the groups of individual robot-centric reference frames are collocated into a larger reference frame.

Another frame of reference could also be used wherein multiple maps are discovered and populated by entities from physically distinct regions. These maps could be correlated into a single larger reference frame (i.e., global positioning system (GPS) or interior maps of different buildings in a city). The common reference frame is simply a way to combine multiple sources of information into a single representation.

4) 2-D and 3-D Reference Frames: Both traditional 2-D interfaces and the 3-D interface support a common reference frame between the robot pose and obstacles by illustrating the map of the environment. However, that is the extent of the common reference frame with the 2-D interface since video, camera pose, and operator perspective are not presented in the same reference
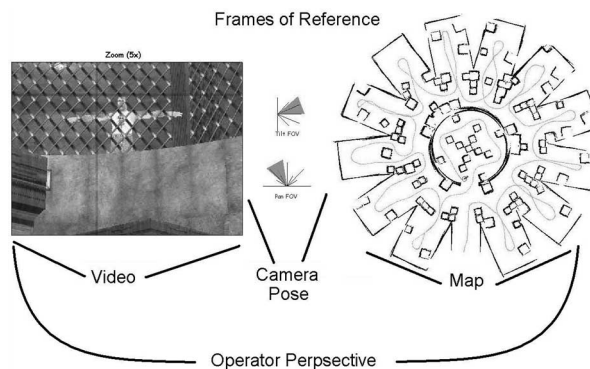

Fig. 9. Four reference frames of the information displayed in a 2-D interface: video, camera pose, map, and operator perspective.
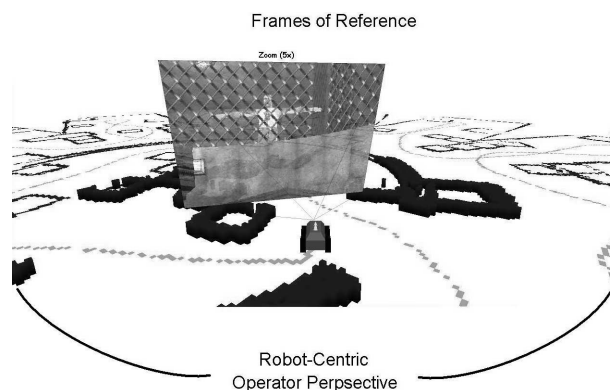

Fig. 10. Reference frames of the information displayed in a 3-D interface: robot-centric and operator perspective (which are both the same).

frame as the map or the robot. In fact, Fig. 9 illustrates that with the 2-D interface, there are at least four different frames of reference from which information is presented to the operator. Specifically, video is presented from the front of the camera, the tilt angle is presented from the right side of the robot, the pan angle is presented from above the robot, the map is presented from a "north-up" perspective, and the operator perspective is a conglomeration of the previous reference frames.

In contrast, the 3-D interface presents the video, camera pose, and user perspective in the same reference frame as the map and the robot pose, as illustrated in Fig. 10.

The multiple reference frames in the 2-D interface require more cognitive processing than the single reference frame in the 3-D interface because the operator must mentally rotate the distinct reference frames into a single reference frame to understand the meaning of the different sets of information [85]. With the 3-D interface, the work of combining the reference frames is supported by the interface which, in turn, reduces the cognitive requirements on the operator.

B. Correlation of Action and Response

Another principle to reduce cognitive workload is to maintain a correlation with commands issued by the operator and the expected result of those commands as observed by the movement

of the robot and changes in the interface. In terms of Endsley's three levels of situation awareness [4], the correlation of action and response affects the operator's ability to project or predict how the robot will respond to commands.

An operator's expected response depends on his or her mental model of how commands translate into robot movement and how robot movement changes the information on the interface. When an operator moves the joystick forward, the general expectation, with both the 2-D and the 3-D interface, is that the robot will move forward. However, the expectation of how the interface will change to illustrate the robot's new position is different for both interfaces. In particular, an operator's expectation of the change in video and the change in the map can lead to confusion when using the 2-D interface.

1) Change in Video: One expectation of operators is how the video will change as the robot is driven forward. In the 2-D interface, the naïve expectation is that the robot will appear to travel "into" the video when moving forward. With the 3-D interface, the expectation is that the robot will travel "into" the virtual environment. Both of these expectations are correct if the camera is in front of the robot. However, when the camera is off-center, an operator with the 2-D interface still might expect the robot to move "into" the video when, in reality, the video moves sideways, which does not match the expectation and can be confusing [17]. With the 2-D interface, the operator's expectation matches the observed change in the interface only when the camera is directly in front of the robot. In contrast, with the 3-D interface, the operator expects the robot to move into the virtual environment regardless of the orientation of the camera, which is the visual response that happens.

2) Change in Map: Another expectation of the operator is how the robot icon on the map will change as the robot is driven forward. With the 2-D interface, the naïve expectation is that the robot will travel up (north) on the map when the joystick is pressed forward. With the 3-D interface, the expectation is that the robot will travel forward with respect to the current orientation of the map. Both of these expectations are correct if the robot is heading "up" with respect to the map. When the robot is heading in a direction other than north, or up, an operator with the 2-D interface would still have the same naïve expectation; however, the robot icon will move in the direction in which the robot is heading, which rarely coincides with "up." This can be particularly confusing when turn commands are issued, because the way in which the turn command affects the robot icon on the map will change on the basis of the global orientation of the robot, which changes throughout the turn command [82], [86].

With the 2-D interface, different sets of information that could be related are displayed in an unnatural presentation from different perspectives. This requires mental rotations by the operator to orient the sets of information into the same frame of reference. The mental rotations required to understand the relationships between the sets of information result in increased mental workload. With the 3-D interface, the information is presented in a spatially natural representation, which does not require mental rotations to understand the information. Future work could address if the workload from mental rotations is affected by operator perspectives of either north-up maps or forward-up maps.

3) Change in Camera Tilt: One area of operator expectation that is difficult to match is the operator's mental model of how the interface should change when a camera is tilted up or down. To control the camera tilt in previous experiments, the point of view (POV) hat on top of the joystick was used, the problem is that some operators prefer to tilt the camera up by pressing "up" on the POV and others prefer to tilt the camera up by pressing "down" on the POV. This observation illustrates the fact that sometimes the mental model of the operator is based on preferences and not the manner in which information is presented. To increase the usability of an interface, some features should be adjustable by the user. Alternatively, different control devices, those that support a less-ambiguous mental mapping from human action to robot response, could be used.

4) Cognitive Workload: The advantage of the 3-D interface is that the operator has a robot-centric perspective of the environment because the viewpoint through which the virtual environment is observed is tethered to the robot. This means that the operator issues commands as they relate to the robot, and the expected results match the actual results. Since the operator's perspective of the environment is robot-centric, there is minimal cognitive workload to correctly anticipate how the interface will change as the robot responds to commands.

The problem with the 2-D interface is that the operator either has a map-centric perspective or a video-centered perspective of the robot that must be translated to a robot-centric perspective in order to issue correct commands to the robot. The need for explicit translation of perspectives results in a higher cognitive workload to anticipate and verify the robot's response to commands.

Additionally, the 2-D interface can be frustrating because it may seem that the same actions in the same situations lead to different results. The reason for this is that the most prominent areas of the interface are the video and the map, which generally have a consistent appearance. The orientation of the robot and the camera, on the other hand, are less prominently displayed even though they significantly affect how displayed information will change as the robot is moved.

If the orientation of the robot or the camera is neglected or misinterpreted, it can lead to errors in robot navigation. Navigational errors increase cognitive workload because the operator must determine why the actual response did not match his or her expected response. For this reason, a novice operator can be frustrated that the robot does different things when it appears that the same information is present and the same action is performed.

C. Adjustable Perspective

Although sets of information may be displayed in a common reference frame, the information may not always be visible or useful because of the perspective through which the operator views the information. Therefore, the final principle that we discuss for reducing cognitive workload is to use an adjustable perspective. An adjustable perspective is one where the
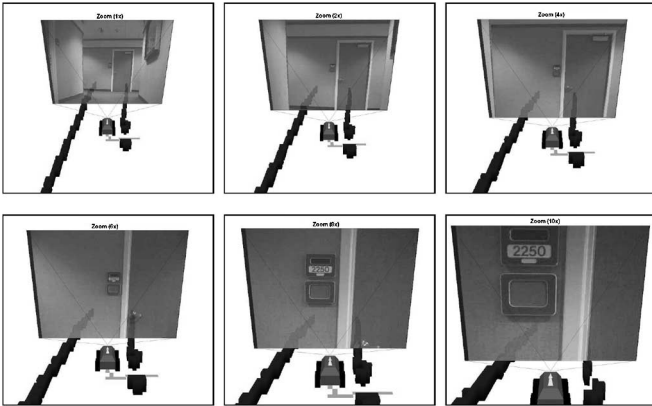
Fig. 11. 3-D representation of the level of zoom with a PTZ camera. The appearance of zoom is affected by adjusting the operator's perspective of the environment. On the top row from left to right, the zoom levels are $1\times$, $2\times$, and $4\times$. On the bottom row from left to right, the zoom levels are $6\times$, $8\times$, and $10\times$.

operator controls the changes, and an adaptive perspective is one that is controlled automatically by an algorithm. Video games tend to use adaptive perspectives that change to avoid obstacles. An adjustable perspective can aid all three levels of Endsley's situation awareness [4] because it can be used to 1) visualize the required information (perception); 2) support the operator in different tasks (comprehension); and 3) maintain awareness when switching perspectives (projection).

1) Visualization: One advantage of an adjustable perspective is that it can be changed depending on the information that the operator needs to "see." For example, if there is too much information in a display, the perspective can shrink to eliminate extra information and focus on the information of interest. Similarly, if there is some information that is outside of the visible area of the display, then the perspective can be enlarged to allow the visibility of more information.

Visualizing just the right amount of information can have a lower cognitive workload than either observing too much or too little of the environment. When there is too little information in the display, the operator is left with the responsibility to remember the previously seen information. When there is too much information in the display, the operator has the responsibility to find and interpret the necessary information. Determining the best visualization, however, comes at a cost to the operator since he or she must think about choosing the right perspective. The ability to zoom in and out is a common feature of most 2-D and 3-D maps, but in 2-D interfaces, the map is usually the only part of the interface with an adjustable perspective, and as the zoom level changes, the relationships between the map and other sets of information also change.

One issue that deserves further work with an adjustable or an adaptable interface is the use of the zoom feature on a PTZ camera. The challenge is to simultaneously inform the user of an increase in detail and a decrease in the field of view. One approach would be to show the increase in detail by making the video larger, but this gives the illusion of an increased field of view. On the other hand, making the video smaller shows a decreased field of view, but also gives the illusion of decreased detail. One possible solution with the 3-D interface is to provide

a perspective of the robot and environment a distance above and behind the robot, and when the camera is zoomed in, the virtual perspective moves forward, which gives the impression that the field of view is smaller (less of the environment is visible) and the level of detail is increased (the video appears larger) [87]. Fig. 11 shows how the interface might be adjusted. Such software should be tested to determine whether or not it actually helps the operator.

2) Changing Tasks: Another advantage of an adjustable perspective is that the perspective through which an operator views a robot in its environment can influence the performance on a particular task. For example, direct teleoperation is usually performed better with a more egocentric perspective, while spatial reasoning and planning tasks are performed better with a more exocentric perspective [16], [82]. When the perspective of the interface is not adjusted to match the requirements of a task, the cognitive workload on the operator is increased because the operator must mentally adjust the perceived information to match the requirements of the task. The kinds of 2-D interfaces that we studied tacitly present an adjustable perspective in so much as many different perspectives are visible at the same time and the operator can switch between them. The problem is not that the interfaces do not allow adjusting the perspective, but that they neither present an integrated perspective nor the ability to adjust the integrated perspective.

3) Maintain Awareness: Often, robots are versatile and can be used to accomplish multiple tasks; thus, it is reasonable to anticipate that an operator would change tasks while a robot is in operation. To facilitate this change, an adjustable perspective can be used to create a smooth transition between one perspective and another. A smooth transition between perspectives has the advantage of allowing the operator to maintain situational context as the perspective changes, which reduces the cognitive workload by reducing the need to acquire the new situational information from scratch [88], [89]. Some instances where a smooth transition might be useful include switching between egocentric and exocentric perspectives, information sources (GPS-, map-, or robot-based), map representations (occupancy-grid, topological), video sources (cameras in different locations, different types of camera), or switching between multiple vehicles.

In the user studies presented previously, a different perspective was used for many of the 3-D interfaces because there were different requirements for the tasks, and the information sometimes needed to be viewed differently. In comparison, the 2-D interface always had the same perspective because conventional 2-D interfaces do not provide an adjustable perspective.

VI. CONCLUSION

In order to improve remote robot teleoperation, an ecological interface paradigm was presented based on Gibson's notion of affordances. The goal of this approach was to provide the operator with appropriate information such that the observed affordances of the remote robot matched the actual affordances, thereby facilitating the operator's ability to perceive, comprehend, and project the state of the robot. To accomplish this task,

a 3-D augmented-virtuality interface was presented that integrates a map, robot pose, video, and camera pose into a single display that illustrates the relationships between the different sets of information.

To validate the utility of the 3-D interface in comparison to conventional 2-D interfaces, a series of user studies was performed and summarized. The results from the user studies show that the 3-D interface improves 1) robot control; 2) map-building speed; 3) robustness in the presence of delay; 4) robustness to distracting sets of information; 5) awareness of the camera orientation with respect to the robot; and 6) the ability to perform search tasks while navigating the robot.

Subjectively, the participants preferred the 3-D interface to the 2-D interface and felt that they did better, were less frustrated, and better able to anticipate how the robot would respond to their commands. The ability of the operator to stay further away from obstacles with the 3-D interface is a strong indication of the operator's navigational awareness. There is a much lower rate of "accidentally" bumping into a wall because the operator is more aware of the robot's proximity to obstacles, and the operator does a better job of maintaining a safety cushion between the robot and the walls in the environment.

From a design perspective, three principles were discussed that ultimately led to the success of the 3-D interface. The principles are: 1) present a common reference frame; 2) provide visual support for the correlation of action and response; and 3) allow an adjustable perspective. These principles facilitated the use of the 3-D interface by helping to reduce the cognitive processing required to interpret the information from the robot and make decisions.

VII. FUTURE WORK

In the current implementation of the 3-D interface, the map is obtained from a laser range-finder that scans a plane of the environment a few inches off the ground. This approach works particularly well for planar worlds, which generally limit the work to indoor environments. In order to apply the research to an outdoor environment, we will look at approaches for measuring and representing terrain (e.g., an outdoor trail). One of the main challenges of presenting a visualization of terrain is that it will necessarily increase the cognitive workload on the operator, because there will be more information displayed in the interface since terrain information is available at every place in the environment. A solution will be determined by answering the question of how much information is required to give the operator sufficient awareness with a minimal effect on the operator's cognitive workload.

A second area of work is to make the interface adjustable or adaptive based on the role of the operator using the interface. For example, in a search and rescue operation, there may be one operator who is in charge of moving the robot while another is in charge of searching the environment. Further, consider the director of the search operation who may not be in charge of operating a robot but may require information about what has been explored, what has been found, and how resources are being used. Each individual may require different sets of infor-

mation to adequately perform his task. If too much information is provided, then the cognitive workload to understand the required information for a particular task will lead to decreased performance. Similarly, too little information will also lead to decreased performance. Therefore, it would be useful to find a satisfying balance between the information needs of multiple operators performing different tasks.

Lastly, it would be interesting to study how and when robot intelligence might help an operator accomplish a task with a robot in comparison to having a robot with no intelligence. Following such a path could enable the comparison of how the interface and the robot intelligence can be combined to improve robot usability.

REFERENCES

[1] T. B. Sheridan, "Musings on telepresence and virtual presence," *Presence: Teleoper., Virtual Environ.*, vol. 1, no. 1, pp. 120–125, 1992.
[2] T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control.* Cambridge, MA: MIT Press, 1992.
[3] A. A. Nofi, "Defining and measuring shared situation awareness," *Center Naval Anal.*, Alexandria, VA, Tech. Rep. CRM D0002895, Nov. 2000.
[4] M. R. Endsley, "Design and evaluation for situation awareness enhancement," in *Proc. Hum. Factors Soc. 32nd Annu. Meet.*, Santa Monica, CA, 1988, pp. 97–101.
[5] J. Casper and R. R. Murphy, "Human–robot interactions during the robot-assisted urban search and rescue response at the world trade center," *IEEE Trans. Syst., Man, Cybern. B*, vol. 33, no. 3, pp. 367–385, Jun. 2003.
[6] J. L. Burke, R. R. Murphy, M. D. Coovert, and D. L. Riddle, "Moonlight in Miami: A field study of human–robot interaction in the context of an urban search and rescue disaster response training exercise," *Hum.–Comput. Interact.*, vol. 19, pp. 85–116, 2004.
[7] D. D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human–robot coordination in future operations," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, no. 2, pp. 210–218, May 2004.
[8] D. Woods and J. Watts, "How not to have to navigate through too many displays," in *Handbook of Human–Computer Interaction,* 2nd ed., M. Helander, T. Landauer, and P. Prabhu, Eds. Amsterdam, The Netherlands: Elsevier Science, 1997, pp. 1177–1201.
[9] P. L. Alfano and G. F. Michel, "Restricting the field of view: Perceptual and performance effects," *Percept. Mot. Skills*, vol. 70, no. 1, pp. 35–45, 1990.
[10] K. Arthur, "Effects of field of view on performance with head-mounted displays" Ph.D. dissertation, Dept. Comput. Sci., Univ. North Carolina, Chapel Hill, 2000.
[11] H. A. Yanco and J. L. Drury, ""Where am I?" Acquiring situation awareness using a remote robot platform," in *Proc. IEEE Conf. Syst., Man, Cybern.*, Oct. 2004, vol. 3, pp. 2835–2840.
[12] A. Jacoff, E. Messina, and J. Evans, "A reference test course for autonomous mobile robots," in *Proc. SPIE-AeroSense Conf.*, Orlando, FL, Apr. 2001.
[13] J. L. Drury, J. Scholtz, and H. A. Yanco, "Awareness in human-robot interactions," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Washington, DC, Oct. 2003.
[14] B. P. DeJong, J. Colgate, and M. A. Peshkin, "Improving teleoperation: Reducing mental rotations and translations," in *Proc. IEEE Int. Conf. Robot. Autom.*, New Orleans, LA, Apr. 2004.
[15] J. D. Lee, B. Caven, S. Haake, and T. L. Brown, "Speech-based interaction with in-vehicle computers: The effect of speech-based email on driver's attention to the roadway," *Hum. Factors*, vol. 43, pp. 631–640, 2001.

[16] J. Scholtz, "Human–robot interactions: Creating synergistic cyber forces," presented at the 2002 NRL Workshop Multirobot Syst., Washington, DC, Mar.

[17] H. A. Yanco, J. L. Drury, and J. Scholtz, "Beyond usability evaluation: Analysis of human–robot interaction at a major robotics competition," *J. Hum.–Comput. Interact.*, vol. 19, no. 1 and 2, pp. 117–149, 2004.

[18] J. L. Burke, R. R. Murphy, E. Rogers, V. J. Lumelsky, and J. Scholtz, "Final report for the DARPA/NSF interdisciplinary study on human–robot interaction," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, no. 2, pp. 103–112, May 2004.

[19] R. R. Murphy, "Humans, robots, rubble, and research," *Interactions*, vol. 12, no. 2, pp. 37–39, Mar./Apr. 2005.

[20] P. Zahorik and R. L. Jenison, "Presence as being-in-the-world," *Presence*, vol. 7, no. 1, pp. 78–89, Feb. 1998.

[21] T. Fong, C. Thorpe, and C. Baur, "A safeguarded teleoperation controller," in *Proc. IEEE Int. Conf. Adv. Robot.*, Budapest, Hungary, Aug. 2001.

[22] E. Krotkov, R. Simmons, F. Cozman, and S. Koenig, "Safeguarded teleoperation for lunar rovers: From human factors to field trials," in *Proc. IEEE Planetary Rover Technol. Syst. Workshop*, Minneapolis, MN, Apr. 1996.

[23] J. Bradshaw, M. Sierhuis, A. Acquisti, P. Fetovich, R. Hoffman, R. Jeffers, D. Prescott, N. Suri, A. Uszok, and R. Hoff, "Adjustable autonomy and human-agent teamwork in practice: An interim report on space applications," in *Agent Autonomy*. Norwell, MA: Kluwer, 2002, pp. 9–38.

[24] D. J. Bruemmer, D. Dudenhoeffer, and J. Marble, "Dynamic autonomy for urban search and rescue," presented at the 2002 AAAI Mobile Robot Workshop, Edmonton, Canada, Aug.

[25] M. Goodrich, D. Olsen, Jr., J. Crandall, and T. Palmer, "Experiments in adjustable autonomy," in *Proc. IJCAI-01 Workshop Auton., Delegation, Control: Interact. Auton. Agents*, 2001, pp. 1624–1629.

[26] P. Scerri, D. Pynadath, and M. Tambe, "Adjustable autonomy in real-world multi-agent environments," in *Proc. 5th Int. Conf. Auton. Agents*, Montreal, Canada, 2001.

[27] M. Hearst, "Mixed-initiative interaction—Trends and controversies," *IEEE Intell. Syst.*, vol. 14, no. 5, pp. 14–23, Sep./Oct. 1999.

[28] P. Kroft and C. Wickens, "Displaying multi-domain graphical database information: An evaluation of scanning, clutter, display size, and user activity," *Inf. Des. J.*, vol. 11, no. 1, pp. 44–52, 2002.

[29] T. W. Fong and C. Thorpe, "Vehicle teleoperation interfaces," *Auton. Robots*, vol. 11, no. 1, pp. 9–18, Jul. 2001.

[30] S. Iba, J. Weghe, C. Paredis, and P. Khosla, "An architecture for gesture based control of mombile robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Kyongju, Korea, Oct.1999, pp. 851–857.

[31] S. Waldherr, R. Romero, and S. Thrun, "A gesture based interface for human–robot interaction," *Auton. Robots*, vol. 9, no. 2, pp. 151–173, 2000.

[32] N. Diolaiti and C. Meichiorri, "Teleoperation of a mobile robot through haptic feedback," in *Proc. IEEE Int. Workshop Haptic Virtual Environ. Appl.*, Ottawa, ON, Canada, Nov. 2002, pp. 67–72.

[33] V. Kulyukin, C. Gharpure, and C. Pentico, "Robots as interfaces to haptic and locomotor spaces," in *Proc. ACM/IEEE Int. Conf. Hum.–Robot Interact.*, Arlington, VA, 2007, pp. 325–331.

[34] S. Lee, G. S. Sukhatme, G. J. Kim, and C.-M. Park, "Haptic control of a mobile robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Lausanne, Switzerland, 2002, pp. 2867–2874.

[35] L. Yu, P. Tsui, Q. Zhou, and H. Hu, "A web-based telerobotic system for research and education at Essex," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatron.*, Como, Italy, Jul. 2001.

[36] D. Schulz, W. Burgard, D. Fox, S. Thrun, and A. Creemers, "Web interfaces for mobile robots in public places," *IEEE Robot. Autom. Mag.*, vol. 7, no. 1, pp. 48–56, Mar. 2000.

[37] G. Chronis and M. Skubic, "Robot navigation using qualitative landmark states from sketched route maps," in *Proc. IEEE 2004 Int. Conf. Robot. Autom.*, New Orleans, LA, pp. 1530–1535.

[38] H. Keskinpala, J. Adams, and K. Kawamura, "A PDA-based human robotic interface," in *Proc. Int. Conf. Syst., Man, Cybern.*, Washington, DC, 2003.

[39] M. Skubic, D. Perznowski, S. Blisard, A. Schultz, W. Adams, M. Bugajska, and D. Brock, "Spatial language for human–robot dialogs," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, no. 2, pp. 154–167, May 2004.

[40] T. Fong, C. Thorpe, and C. Baur, "Robot as partner: Vehicle teleoperation with collaborative control," in *Proc. 2002 NRL Workshop Multirobot Syst.*, Washington, DC, Mar.

[41] R. R. Murphy and E. Rogers, "Cooperative assistance for remote robot supervision," *Presence*, vol. 5, no. 2, pp. 224–240, 1996.

[42] P. Dourish and V. Bellotti, "Awareness and coordination in shared workspaces," in *Proc. ACM Conf. Comput.-Supported Coop. Work*, Toronto, ON, Canada, 1992, pp. 107–114.

[43] M. R. Endsley, "Automation and situation awareness," in *Automation and Human Performance: Theory and Applications*, R. Parasuraman and M. Mouloua, Eds. Mahwah, NJ: Lawrence Erlbaum, 1996.

[44] C. Wickens, "Situation awareness and workload in aviation," *Current Directions Psychol. Sci.*, vol. 11, no. 4, pp. 128–133, 2002.

[45] S. Nayar, "Catadioptric omnidirectional camera," Bell Labs., Holmdel, NJ, Tech. Rep., 1997.

[46] G. Thomas, W. D. Robinson, and S. Dow, "Improving the visual experience for mobile robotics," in *Proc. 7th Annu. Iowa Space Grant*, Des Moines, IA, Nov. 1997, pp. 10–20.

[47] G. Thomas, T. Blackmon, M. Sims, and D. Rasmussen, "Video engraving for virtual environments," in *Proc. Electron. Imaging: Sci. Technol.*, San Jose, CA, Feb. 1997.

[48] K. Yamazawa, Y. Yagi, and M. Yachida, "Obstacle avoidance with omnidirectional image sensor hyperomni vision," in *Proc. IEEE Int. Conf. Robot. Autom.*, Nagoya, Japan, 1995, pp. 1062–1067.

[49] S. Hughes, J. Manojlovich, M. Lewis, and J. Gennari, "Camera control and decoupled motion for teleoperation," in *Proc. 2003 IEEE Int. Conf. Syst., Man, Cybern.*, Washington, DC, Oct. 5–8.

[50] B. Keyes, R. Casey, H. Yanco, B. Maxwell, and Y. Georglev, "Camera placement and multi-camera fusion for remote robot operation," in *Proc. IEEE Int. Workshop Safety, Security, Rescue Robot.*, Gaithersburg, MD, Aug. 2006.

[51] M. G. Voshell and D. D. Woods, *Breaking the keyhole in human–robot coordination: Method and evaluation*. Ohio State Univ., Athens, OH, Tech. Rep., 2005.

[52] G. Terrien, T. Fong, C. Thorpe, and C. Baur, "Remote driving with a multi-sensor user interface," presented at the SAE 30th Int. Conf. Environ. Syst., Toulouse, France, 2000.

[53] T. W. Fong, C. Thorpe, and C. Baur, "Advanced interfaces for vehicle teleoperation: Collaborative control, sensor fusion displays, and remote driving tools," *Auton. Robots*, vol. 11, no. 1, pp. 77–85, Jul. 2001.

[54] L. A. Nguyen, M. Bualat, L. J. Edwards, L. Flueckiger, C. Neveu, K. Schwehr, M. D. Wagner, and E. Zbinden, "Virtual reality interfaces for visualization and control of remote vehicles," *Auton. Robots*, vol. 11, no. 1, pp. 59–68, Jul. 2001.

[55] C. Stoker, E. Zbinden, T. Blackmon, B. Kanefsky *et al.*, "Analyzing pathfinder data using virtual reality and superresolved imaging," *J. Geophys. Res.—Planets*, vol. 104, no. E4, pp. 8889–8906, 1999.

[56] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. E77-D, no. 12, pp. 1321–1329, 1994.

[57] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoper. Virtual Environ.*, vol. 6, no. 4, pp. 355–385, Aug. 1997.

[58] P. Milgram, S. Zhai, D. Drascic, and J. Grodski, "Applications of augmented reality for human–robot communication," in *Proc. Int. Conf. Intell. Robots Syst.*, Yokohama, Japan, Jul. 1993, pp. 1467–1472.

[59] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using presence questionnaires in reality," *Presence*, vol. 9, no. 5, pp. 497–503, Oct. 2000.

[60] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *J. Commun.*, vol. 42, no. 2, pp. 73–93, 1992.

[61] G. Mantovani and G. Riva, "Real presence: How different ontologies generate different criteria for presence, telepresence, and virtual presence," *Presence: Teleoper. Virtual Environ.*, vol. 8, no. 5, pp. 538–548, Oct. 1999.

[62] W. Sadowsky and K. Stanney, "Measuring and managing presence in virtual environments," in *Handbook of Virtual Environments Technology*, K. Stanney, Ed. Hillsdale, NJ: Lawrence Erlbaum, 2002.

[63] K. Stanney, R. Mourant, and R. Kennedy, "Human factors issues in virtual environments: A review of the literature," *Presence: Teleoper. Virtual Environ.*, vol. 7, no. 1, pp. 327–351, 1998.

[64] D. J. Bruemmer, J. L. Marble, D. A. Few, R. L. Boring, M. C. Walton, and C. W. Nielsen, "Shared understanding for collaborative control," *IEEE Trans. Syst., Man, Cybern. A*, vol. 35, no. 4, pp. 494–504, Jul. 2005.

[65] M. Baker, R. Casey, B. Keyes, and H. A. Yanco, "Improved interfaces for human–robot interaction in urban search and rescue," in *Proc. IEEE Conf. Syst., Man Cybern.*, The Hauge, The Netherlands, Oct. 2004, pp. 2960–2965.

[66] J. J. Gibson, *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin, 1979.

[67] D. A. Norman, "Affordance, conventions, and design," *Interactions*, vol. 6, no. 3, pp. 38–43, May/Jun. 1999.

[68] R. R. Murphy, "Case studies of applying gibson's ecological approach to mobile robots," *IEEE Trans. Syst., Man, Cybern. A*, vol. 29, no. 1, pp. 105–111, Jan. 1999.

[69] P. Milgram, D. Drascic, J. J. Grodski, A. Restogi, S. Zhai, and C. Zhou, "Merging real and virtual worlds," in *Proc. IMAGINA Conf. 1995*, Monte Carlo, NV, Feb.

[70] D. Drascic and P. Milgram, "Perceptual issues in augmented reality," *Proc SPIE: Stereoscopic Displays Virtual Real. Syst. III*, vol. 2653, pp. 123–134, 1996.

[71] B. W. Ricks, C. W. Nielsen, and M. A. Goodrich, "Ecological displays for robot interaction: A new perspective," in *Proc. Int. Conf. Intell. Robots Syst. IEEE/RSJ*, Sendai, Japan, 2004.

[72] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., North-Holland, Amsterdam, The Netherlands, 1988, pp. 139–183.

[73] O. Nakayama, T. Futami, T. Nakamura, and E. Boer, "Development of a steering entropy method for evaluating driver workload," presented at the Int. Congr. Expo., Detroit, MI, Mar. 1999.

[74] K. Konolige, "Large-scale map-making," in *Proc. Nat. Conf. AAAI*, San Jose, CA, 2004.

[75] C. W. Nielsen and M. A. Goodrich, "Comparing the usefulness of video and map information in navigation tasks," in *Proc. 2006 Hum.–Robot Interact. Conf.*, Salt Lake City, UT.

[76] J. Wang, M. Lewis, and J. Gennari, "A game engine based simulation of the NIST urban search and rescue arenas," in *Proc. 2003 Winter Simul. Conf.*, vol. 1, pp. 1039–1045.

[77] M. C. W. Douglas, A. Few, and D. J. Bruemmer, "Improved human–robot teaming through facilitated initiative," in *Proc. 15th IEEE Int. Symp. Robot Hum. Interact. Commun.*, Hatfield, U.K., Sep. 2006.

[78] E. B. Pacis, H. Everett, N. Farrington, and D. J. Bruemmer, "Enhancing functionality and autonomy in man-portable robots,," in *Proc. SPIE Unmanned Ground Veh. Technol. VI, Defense Security*, Orlando, FL, Apr. 2004, pp. 355–366.

[79] C. W. Nielsen and M. A. Goodrich, "Testing the usefulness of a pan-tilt-zoom (PTZ) camera in human–robot interactions," in *Proc. Hum. Factors Ergon. Soc. 50th Annu. Meet.*, San Francisco, CA, 2006.

[80] C. W. Nielsen, M. A. Goodrich, and R. J. Rupper, "Towards facilitating the use of a pan-tilt camera on a mobile robot," in *Proc. 14th IEEE Int. Workshop Robot Hum. Interact. Commun.*, Nashville, TN, 2005.

[81] M. A. Goodrich and D. R. Olsen, Jr., "Seven principles of efficient interaction," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.* Oct. 5–8, 2003, pp. 3943–3948.

[82] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[83] M. R. Endsley, B. Bolté, and D. G. Jones, *Designing for Situation Awareness*. New York: Taylor & Francis, 2003.

[84] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence New Millennium*, G. Lakemeyer and B. Nebel, Eds. San Mateo, CA: Morgan Kaufmann, 2002.

[85] B. P. DeJong, J. E. Colgate, and M. A. Peshkin, "Improving teleoperation: Reducing mental rotations and translations," in *Proc. Am. Nuclear Soc. 10th Int. Conf. Robot. Remote Syst. Hazard. Environ.*, Gainesville, FL, Mar. 2004.

[86] R. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, pp. 701–703, 1971.

[87] M. A. Goodrich, R. J. Rupper, and C. W. Nielsen, "Perceiving head, shoulders, eyes, and toes in augmented virtuality interfaces for mobile robots," in *Proc. 14th IEEE Int. Workshop Robot Hum. Interact. Commun.*, Nashville, TN, 2005.

[88] J. W. Crandall, M. A. Goodrich, D. R. Olsen, Jr., and C. W. Nielsen, "Validating human–robot interaction schemes in multi-tasking environments," *IEEE Trans. Syst., Man, Cybern. A*, vol. 35, no. 4, pp. 438–449, Jul. 2005.

[89] D. R. Olsen, Jr. and S. B. Wood, "Fan-Out: Measuring human control of multiple robots," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Vienna, Austria, 2004, pp. 231–238.

Curtis W. Nielsen (M'07) received the B.S., M.S., and Ph.D. degrees in computer science from Brigham Young University, Provo, UT, in 1999, 2003, and 2006, respectively.

In 2005, he joined the Idaho National Laboratory, Idaho Falls, as a Principal Research Scientist in the Robotics and Human Systems Group. His research interests include robotics, human–robot interaction, interface design, search and rescue, computer graphics, user studies, and human factors.

Dr. Nielsen received the R&D 100 Award for his work on a robot intelligence kernel in 2006.

Michael A. Goodrich (S'92–M'92–SM'05) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Brigham Young University, Provo, UT, in 1992, 1995, and 1996, respectively.

From 1996 to 1998, he was a Research Associate with Nissan Cambridge Research, Nissan Research and Development, Inc., Cambridge, MA. Since 1998, he has been with the Computer Science Department, Brigham Young University, where he is currently an Associate Professor. His research interests include human–robot interaction, decision theory, multiagent learning, and human-centered engineering.

Robert W. Ricks received the B.S and M.S. degrees in computer science from Brigham Young University, Provo, UT, in 2002 and 2004, respectively.

Since 2004, he has been a Computer Systems Researcher with the United States Department of Defense. His current research interests include artificial intelligence, fuzzy logic, knowledge discovery, and robotics.