INL/CON-06-01388
PREPRINT

Comparing The Usefulness Of Video And Map Information In Navigation Tasks

Human-Robot Interaction

Curtis W. Nielsen
Michael A. Goodrich

March 2006

Comparing the Usefulness of Video and Map Information in Navigation Tasks

Curtis W. Nielsen
Brigham Young University
3361 TMCB
Provo, UT 84601
curtisn@gmail.com

Michael A. Goodrich
Brigham Young University
3361 TMCB
Provo, UT 84601
mike@cs.byu.edu

ABSTRACT

One of the fundamental aspects of robot teleoperation is the ability to successfully navigate a robot through an environment. We define successful navigation to mean that the robot minimizes collisions and arrives at the destination in a timely manner. Often video and map information is presented to a robot operator to aid in navigation tasks. This paper addresses the usefulness of map and video information in a navigation task by comparing a side-by-side (2D) representation and an integrated (3D) representation in both a simulated and a real world study. The results suggest that sometimes video is more helpful than a map and other times a map is more helpful than video. From a design perspective, an integrated representation seems to help navigation more than placing map and video side-by-side.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems—*Human factors, Human information processing*

General Terms

Design, Experimentation, Human factors, Performance

Keywords

HRI, Human Robot Interaction, Information Presentation, Integrated Display, User Studies

1. INTRODUCTION

One of the fundamental aspects of robot teleoperation is the ability to successfully navigate a robot through an environment. We define successful navigation to mean that the robot minimizes collisions with obstacles and arrives at a destination in a timely manner. In order to support an operator in navigational tasks it is important to present navigation-relevant information to the operator. In remote, mobile robot navigation, it is common to use video and/or range information to inform the operator of obstacles and available directions of travel [1, 3, 6, 7, 19].

Both video and range information provide distinct sets of information that have advantages and disadvantages for navigation tasks. For example, a video stream provides a visually rich set of information for interpreting the environment and comprehending obstacles, but it is usually limited by a narrow field of view and it is often difficult to comprehend how the robot's position and orientation relate to the environment. In contrast, range information is typically generated from infra-red sensors, laser range finders, or sonar sensors which detect distances and directions to obstacles, but do not provide more general knowledge about the environment. Advancements in map-building algorithms allow the integration of multiple range scans into maps which help an operator visualize how the robot's position and orientation relate to the environment.

In previous studies we used both video and range information (current readings or a map) to navigate a robot [13, 15]. During the experiments we observed that operators sometimes focused their attention on the map section of the interface and other times focused their attention on the section that contains the video.

These anecdotal observations lead to the question of how useful video and map information are for teleoperation. Although the ways to combine maps and visualization tools have been studied in other domains such as aviation (see, for example [4, 16]) this problem has not been well studied in human-robot operation with occupancy grid maps.

This paper seeks to understand the usefulness of video and map information in navigation by comparing a prototypical 2D interface and a 3D augmented-virtuality interface [13, 15]. Specifically we hypothesized that for navigational tasks the video will hinder performance with the 2D interface, but minimally affect performance with the 3D interface. Further, we hypothesized that map information is more helpful to navigation than video information for both types of interface.

2. MOTIVATION

During the World Trade Center disaster in September 2001, Casper and Murphy used robots to search the rubble for victims [5]. Their robots were primarily operated via a video stream from a camera on the robot. One of their observations was that it was very difficult for an operator to handle both the navigation and the exploration of the environment with only video information.

In a separate study, Yanco and Drury had first responders search a mock environment using a robot that had camera, and map-building capabilities. One of their conclusions is that some participants considered the map useless because they felt it did not help them understand the robot's location [18]. Further, in an analysis of a robot competition, Yanco, Drury and Scholtz observed that many operators demonstrated a lack of awareness of the robot's location and surroundings [19].

Most mobile robot interfaces implement some aspect of video and/or range information to inform the operator of the environment around the robot. Some of these approaches present the information in a 2D, side-by-side approach [1, 3, 19] and others present the information integrated into a single 3D display [12, 7]. In previous work an integrated display was found to be more useful for some navigation tasks in comparison to a side-by-side display [3, 13, 15].

To test the usefulness of map and video information in 2D and 3D interfaces, we next present two user studies: one in simulation and one using a real robot.

3. EXPERIMENT 1

In the first experiment we look at the usefulness of video and map information as aids for navigation with both a side-by-side approach (2D) and an integrated approach (3D). We hypothesized that with 2D interfaces video may negatively influence an operator's ability to perform a navigation task because it does not provide sufficient lateral information and it may draw the operator's attention away from more useful places on the interface such as a map or range information [9]. Furthermore, we hypothesized that with a 3D interface, video information will not hinder navigation when other range information is present. To explore the effect of range and video information on navigation, we assess an operator's ability to navigate a maze environment with two interfaces (2D and 3D) and three conditions for each interface (map-only, video-only, and map+video).

3.1 Framework

For this experiment we used a simulator based on the popular Unreal Tournament game engine as modified by Michael Lewis and colleagues at the University of Pittsburgh [11, 17]. Their modifications were originated with the intent of providing an inexpensive yet realistic simulator for studying urban search and rescue with mobile robots. The Unreal Tournament game engine provides a rich visual environment, which when combined with accurate models of common research robots and the game's physics engine provides for a very good mobile robot simulator [10].

We used the Unreal Tournament level editor to create maze environments that have the appearance of concrete bunkers which are filled with pipes, posters, windows, cabling, and electronic devices to provide a detailed environment for the robot to travel through. Some images of the virtual environment are shown in Figure 1.

The environment we created has seven separate mazes which are designed to explicitly test low-level navigation skills. There is only one path through each maze and no dead-ends, but it takes considerable teleoperation skill to navigate a maze from start to finish without crashing the robot. One of the mazes is used for training and the other 6 mazes are used for testing. The training maze contains a continuous path without an exit so that participants can practice driving the robot as long as desired.

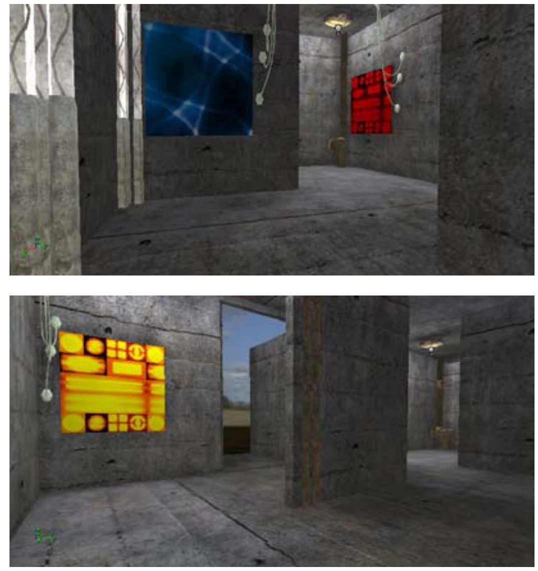

Figure 1: Images from the Unreal Tournament environment used for Experiment 1.

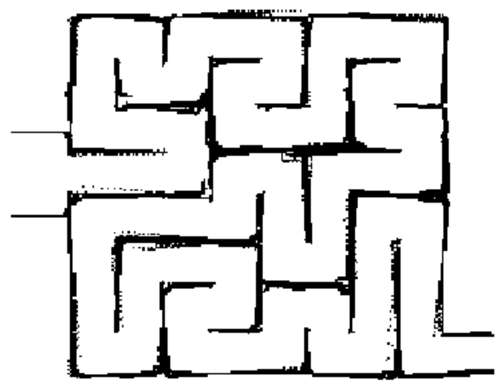

Figure 2: A map of one of the mazes used in Experiment 1.

Each maze is an 8x8 grid where each cell in the grid is 2x2 meters for a total maze area of $256m^2$. Each maze is designed to have 42 turns and 22 straight cells to minimize differences in results from different mazes (see Figure 2). The simulated robot used for this experiment is a model of the ATRV-Jr robot and has a width and length of 0.6 meters.

3.1.1 Procedure

Operators were instructed on how to drive the robot and how to perform the experiment through speakers on a headset, and they were told that their goal was to get the robot out of the maze as quickly as possible without hitting too many walls.

Before testing, operators were given a chance to practice driving the robot with both the 2D and the 3D interfaces. Each interface displayed both map and video information. The operators were asked to drive at least once through

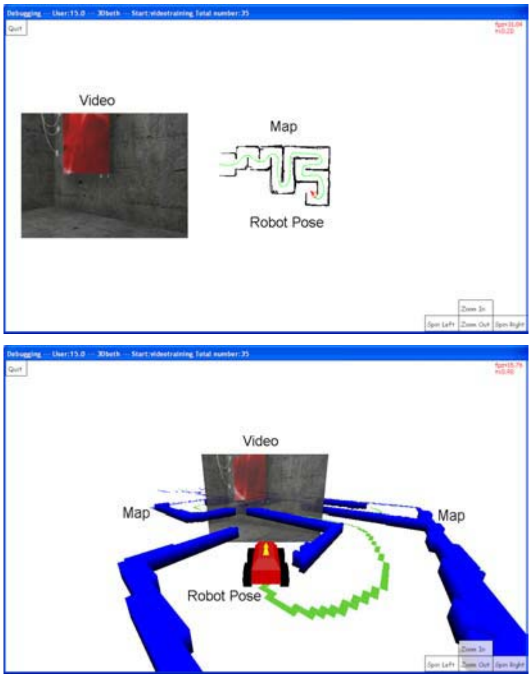

Figure 3: The 2D interface (top) and the 3D interface (bottom) used for Experiment 1.

the training maze to ensure a minimum amount of training. Once an operator had completed the training maze they were asked to continue practicing until they felt comfortable controlling the robot with the interface (most participants stopped training at this point). Following each training session and each experiment, participants were given a questionnaire to evaluate their performance. The purpose of the questionnaires after the training sessions was to familiarize the operators with the questions we would ask after each experiment.

Once training was complete, each participant was asked if they had any questions and they were told that the experiments would be very similar to the training, except that there would be an exit to the maze and that they would have different sets of information visible on the interface for each test. In particular, participants were given conditions of *video-only*, *map-only*, and *map+video* for both the 2D and 3D interfaces. For testing, we used a within-subjects counter-balanced design where each operator performed one test with each of the six conditions which were presented in a random order with the constraints that the 2D and 3D interfaces were used alternately and the conditions were counter balanced on which order they were used. The interfaces for the map+video conditions are shown in Figure 3.

3.2 Results

Twenty-four participants were paid to navigate a simulated robot with six different conditions of information presentation. Participants were recruited from the Brigham Young University community with most subjects enrolled as students. Two participants terminated the experiment prior to completion of the six conditions, but completed portions of the experiment were used for our analysis. Throughout

	2D Interface	3D Interface
Map-only	258	196
Video-only	366	351
% Change	42%	79%
p	$7.8e^{-4}$	$1.6e^{-7}$

Table 1: Time to completion in Experiment 1.

	2D Interface	3D Interface
Map-only	9.83	1.25
Video-only	19.10	22.71
% Change	94%	18x
p	$1.3e^{-3}$	$1.3e^{-6}$

Table 2: Number of collisions in Experiment 1.

the discussion of the results significance was obtained with a paired, two-tailed t-test with $n = 24$ samples unless otherwise specified.

3.2.1 Map-only vs. Video-only

The results indicate that the video-only condition took significantly longer than the map-only condition for both the 2D(42%) and the 3D(79%) interfaces (see Table 1). Additionally, there were nearly twice as many collisions with the video-only condition in 2D than with the map-only condition and there were eighteen times the collisions with the 3D video-only condition in comparison to the 3D map-only condition (see Table 2). The 2D video-only and 3D video-only conditions both had similar (not statistically different) results as measured by time to completion and the number of collisions. This is as we expected because the 3D and 2D interfaces present the video-only condition similarly.

3.2.2 Map+video

We found that with both the 2D and 3D interfaces, the map+video condition had results that were most similar to the map-only condition in comparison to the video-only condition (see Table 3). In particular we found that, on average, there were exactly the same number of collisions with the 3D interface for the map-only and map+video conditions and that there was no significant difference between the 2D map-only and map+video conditions. Figure 4 shows the average number of collisions for each of the six conditions.

On average there was an insignificant change in time to completion when video information was added to map information for both the 2D and 3D interfaces. However, we noticed a learning effect that took place with the 2D map-

	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 3: Comparison of the map+video condition to the map-only and video-only conditions in the simulation experiment.

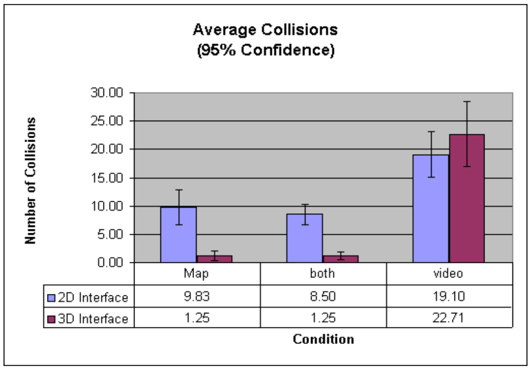

Figure 4: Number of collisions in Experiment 1.

	First	Second	% Change	p
2D map-only	278	238	-14%	0.0953
2D map+video	269	273	1.7%	0.849
%change	-3.1%	15%		
p	0.736	0.0909		

Table 4: Time to completion in 2D after adjusting for learning.

.

only condition and the 3D map+video condition. In particular, the participants that used the 2D map-only condition *after* the 2D map+video condition finished the task 14% faster than the participants that used the 2D map-only condition *before* the 2D map+video condition ($\bar{x}_{2Dmap1} = 278$, $\bar{x}_{2Dmap2} = 238$, $p = .0953$, $n = 12$, unpaired t-test, see Table 4).

Similarly, the participants that used the 3D map+video condition *after* the 3D map-only condition finished the task 15% faster than those that used the 3D map+video condition *before* the 3D map-only condition ($\bar{x}_{3Dmap+video1} = 225$, $\bar{x}_{3Dmap+video2} = 191$, $p = .0115$, $n = 12$, unpaired t-test, see Table 5). We did not notice a learning effect between any of the other conditions.

When we compare the set of experiments in 2D where the map-only and map+video conditions were used first (Table 4), we find that adding video to the map has an insignificant effect. However in the set of experiments where the map-only and map+video conditions were used second, we find the time to completion of the task increases by 14.8% with the map+video condition in comparison to the map-only condition, which suggests that after accounting for learning, adding video to the map hurts navigation by increasing the time it takes an operator to navigate the robot out of a maze.

	First	Second	% Change	p
3D map-only	195	196	0.32%	0.961
3D map+video	225	191	-15%	0.0115
% Change	-15%	-2.7%		
p	0.0357	0.626		

Table 5: Time to completion in 3D after adjusting for learning.

.

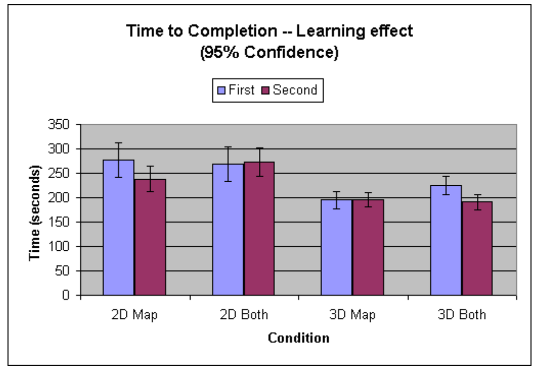

Figure 5: Time to completion after adjusting for learning in Experiment 1.

When we compare the set of experiments in 3D where the map-only and map+video conditions are used first (Table 5), we find that adding video to the map increases the time to completion by 15.2%. However, in the set of experiments where the map-only and map+video conditions are used second, we find the difference in the time to complete the task is insignificant, which suggests that after accounting for learning, adding video to the map in the 3D interface does not affect the time it takes to navigate the robot out of the maze. A summary of the time to completion measurements when considering the learning effect is shown in Figure 5.

3.3 Discussion

These results suggest that video can hurt navigation when the video does not contain sufficient navigational cues and video and map information are placed side-by-side. Even when map information is present and more useful than video for navigating, a novice operator's attention tends to be drawn towards the video, which, in this case, negatively affects their ability to navigate. These results make sense in light of research done by Kubey and Csikszentmihalyi which has shown that television draws attention because of the constantly changing visual scene [9]. It is interesting that even though it took longer to navigate, there were not more collisions with the 2D map+video condition than the 2D map-only condition, which implies that operators were not bumping into walls more, just moving slower through the maze.

4. EXPERIMENT 2

Experiment 1 provided an initial analysis of the usefulness of video and map information for performing navigation tasks with a remote, mobile robot in simulation. It is also useful to verify that the results and conclusions in simulation carry over and are applicable to environments and robots in the real world. For this purpose we have designed the second experiment to compare the usefulness of video and map information when navigating a robot in the real world. We hypothesized that the results would be similar to the results in simulation.

4.1 Framework

For this experiment we converted part of the second floor

Figure 6: Images of the environment and the robot used for Experiment 2.

of the Computer Science Department at Brigham Young University into an obstacle course for our robot to travel through. The normal hallway width in the building is 2 meters and we used cardboard boxes, Styrofoam packing, and other obstacles to create a 50 meter course which has a minimum width of 1.2 meters. Figure 6 shows images of the robot and the two hallways used in the experiment.

4.1.1 The Robot

The robot used for the experiment is an ATRV-Jr which is approximately 0.6 meters in width and 0.7 meters in length (see Figure 6) . The robot uses artificial intelligence algorithms developed at the Idaho National Laboratory (INL) to safeguard it from colliding with walls and obstacles as it is teleoperated [2, 3]. Additionally, the robot uses a map-building algorithm developed by Konolige at the Stanford Research Institute (SRI) to represent the environment and localize the robot within the map [8].

The robot is controlled with a Microsoft Sidewinder 2 joystick[1] and range and video information from the robot are presented to the operator via our 3D interface [13, 14]. The 3D interface is integrated with the INL base station which handles the communication of movement commands and general information between the operator and the robot via radio modems. Live video from the robot is transmitted to the interface via 802.11b wireless Ethernet. The interfaces used for this experiment have been modified from the previous experiment by including icons which indicate where the robot's intelligence identifies obstacles that might interfere with robot movement. The interfaces used for this experiment are shown in Figure 7.

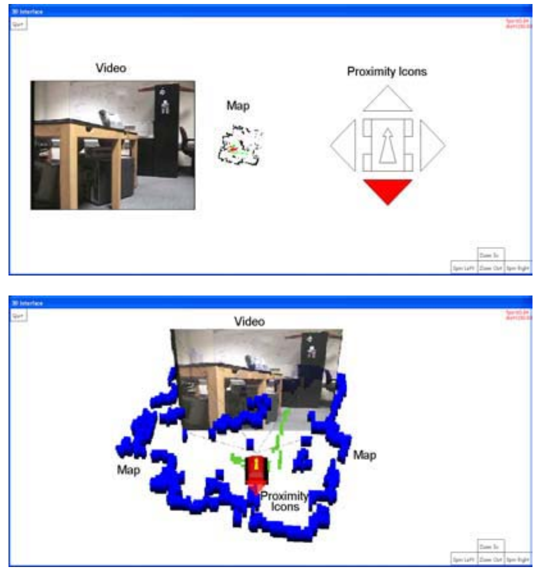

Figure 7: The 2D interface (top) and 3D interface (bottom) used for Experiment 2.

4.1.2 Procedure

Before using the real robot, operators were trained to drive the robot with the Unreal Tournament training maze used in the first experiment. While training, operators drove the simulated robot with a joystick for a few minutes with each of the five conditions that they would be tested on (2D map-only, 2D map+video, video-only[2], 3D map-only, and 3D map+video). Upon completion of the training, the operators were moved to a different base station which was communicating with the real robot.

For testing, we used a within-subjects counter-balanced design where each operator used all five conditions in a pseudo-random order with the constraints that the 2D and 3D interfaces were used alternately and the conditions were counter-balanced on the order in which they were used. The experiment was setup such that an operator would drive the robot through the obstacle course with one condition, then at the end of the course an assistant would change the condition, turn the robot around, reset the map information, and start the next test. After every two runs the robot was plugged in for three to five minutes to keep the batteries charged.

4.2 Results

Twenty-one participants were paid to navigate the ATRV-Jr robot with five different conditions of information presentation. Participants were recruited from the Brigham Young University community with most subjects enrolled as students. The first three participants were used as part of a

[1]The INL base station did not support the steering wheel used in Experiment 1

[2]We did not compare 2D and 3D video-only conditions because in the previous experiment the video-only condition had similar results for both the 2D and 3D interfaces.

	2D Interface	3D Interface
Map-only	46.9	28.6
Video-only	38.8	38.8
% Change	-17%	35%
p	.663	$2.4e^{-2}$

Table 6: Number of times the robot took initiative to protect itself in Experiment 2.

	2D Interface	3D Interface
Map-only	319	227
Video-only	243	243
% Change	-24%	7.2%
p	$1.6e^{-3}$	.599

Table 7: Time to completion in Experiment 2.

pilot study to determine a sufficient complexity of the obstacle course and to determine how best to use the robot while maintaining a sufficiently high charge on the batteries, therefore, there results were not included as part of the analysis. Additionally, the robot's responsiveness to commands was adversely affected by low batteries in eleven of the testing conditions (out of 90) therefore, this data was also discarded.

One of the differences between this experiment and the previous is that the real robot has intelligence on board to protect itself from hitting obstacles. For each test we recorded the number of times the robot acted to protect itself and discuss these results as *robot initiative*. Statistical significance was determined using a paired, two-tailed t-test with $n = 18$ samples except as otherwise noted.

4.2.1 Map-only vs. Video-only

With the 3D interface, there was not a significant difference in the time to completion with the map-only and video-only conditions, however, the robot took initiative to protect itself nearly twice as much with the video-only condition than with the map-only condition ($\bar{x}_{map} = 18.7$, $\bar{x}_{video} = 36.6$, $p = .0378$, see Table 6).

With the 2D interface, there was not a significant difference in the times the robot took initiative to protect itself with the map-only and video-only conditions, however, there was a significant difference in the time to complete the task. In fact, the results were opposite those from the simulated experiment. In particular it was 24% faster to use the video-only condition as opposed to the map-only condition ($\bar{x}_{map} = 319s$, $\bar{x}_{video} = 243s$, $p = 1.6e^{-3}$, see Table 7).

Most likely the reason these results differ from the previous experiment is that the environment in the second experiment provided more navigational cues visible in the video than the environment in the simulation experiment. In the simulation environment it was often the case that the video image was filled by a wall and none of the edges of the wall were visible. Moreover, the path through the simulation maze doubled back on itself numerous times, so the operator could not see very far in front of the robot. In contrast, for this second experiment, the edges of obstacles were nearly always visible through the camera and the operator could see future parts of the map as most obstacles were shorter than the height of the camera and there was only one 90 degree turn in the environment.

	Time to Completion (mean/stdev)	Robot Initiative (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

Table 8: Comparison of the map+video condition to the map-only and video-only conditions in the real-world experiment.

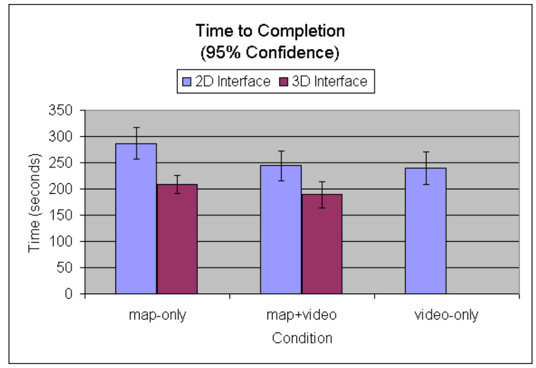

Figure 8: Time to completion for the five conditions in Experiment 2.

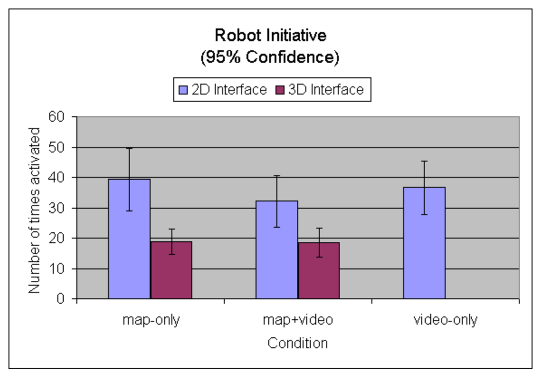

Figure 9: Number of times the robot took initiative to protect ltlself for the five conditions in Experiment 2.

4.2.2 Map+video

When map and video information were combined with the 2D interface, we found the results to be similar to the video-only condition with negligible difference in the time to completion and the number of collisions (see Table 8 and Figures 8 and 9).

When map and video information were combined with the 3D interface, the number of collisions are nearly identical to the map-only condition but we found that operators finished the obstacle course 9.6% faster with the map+video condition in comparison to the map-only condition ($\bar{x}_{map+video} = 205s$, $\bar{x}_{video} = 227s$, $p = 4.6e^{-2}$, see Figure 8).

This result is interesting because it suggests that when useful navigational information is available in both the map and the video, the 3D interface supports the complementary nature of the information and can lead to an improved performance over the individual parts. In contrast, performance with the 2D interface seems to be constrained by the best one can do with an individual set of information.

4.3 Discussion

To determine an ordering for the conditions, we define one condition to be better than another if both of the categories (time to completion and robot initative) are at least non-significantly different and one of the categories is significantly better. Conditions are considered equivalent if there is no statistical difference in either cateogry of analysis.

According to this criteria we found that when using the 3D interface, the map+video condition is better than the map-only condition (because the task took less time), and the map-only condition is better than the video-only condition (because there were fewer instances of robot initiative). These results suggest that when there is useful navigational information, in both the map and the video sets of information, integrating the information can yield better results than using either map or video individually. Furthermore, when using the 2D interface, the map+video and the video-only conditions are similar and are both better than the map-only condition (because the task took less time).

Interestingly, these results are different from our simulation studies where we found the video-only condition to be significantly worse than the other conditions. One complaint among participants with the 2D interface was that the map was too small (although it was the same relative size as the previous experiment) and that it was difficult to correlate the direction of the joystick movement with how the robot would move because the robot icon in the map was not always heading towards the top of the interface. Further, the map+video condition had results most similar to the video-only condition because the video tends to "pull" an operator's attention and hold it more than the map [9]. This assertion is further supported by the questions following the experiments, where operators claimed that most of their time was spent focused on the video.

5. CONCLUSION

Mobile robot navigation depends on the ability to see and comprehend information in the environment surrounding the robot. Typically information from the environment is presented to the operator via range and/or video, however, the manner in which this information is presented to an operator may affect navigational performance.

We have shown that video is helpful in environments where there are navigational cues in the video information, but video can diminish performance when there are minimal navigational cues. Furthermore, when video and map information are placed side-by-side they tend to compete for the operator's attention whereas when video and map information are integrated, they tend to complement each other and improve overall performance.

For design purposes, integrating maps with video in a 3D perspective seems much better than presenting map and video side-by-side in a 2D perspective. Most likely this is because the maps are always visible, even if the operator pays too much attention to the video. These results are consistent with previous results [13, 15].

In the future we plan to look at how delay affects navigation with both the 2D and 3D interfaces. Additionally we plan to look at exploration tasks using different interfaces and different sets of information.

6. REFERENCES

[1] M. Baker, R. Casey, B. Keyes, and H. A. Yanco. Improved interfaces for human-robot interaction in urban search and rescue. In *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, October 2004.

[2] D. J. Bruemmer, J. L. Marble, D. Dudenhoeffer, M. Anderson, and M. McKay. Mixed-initiative control for remote characterization of hazardous environments. In *Proceedings of the Hawaii International Conference on System Sciences*, Waikoloa, Hawaii, January 2003.

[3] D. J. Bruemmer, J. L. Marble, D. A. Few, R. L. Boring, M. C. Walton, and C. W. Nielsen. Let rover take over: A study of mixed-initiative control for remote robotic search and detection. *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans*, 35(4):494–504, July 2005.

[4] G. L. Calhoun, M. H. Draper, M. F. Abernathy, F. Delgado, and M. Patzek. Synthetic vision system for improving unmanned aerial vehicle operator situation awareness. In J. G. Verly, editor, *Proceedings of SPIE Vol. 5802. p. 219-230, Enhanced and Synthetic Vision 2005*. May 2005.

[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 Transactions on Systems, Man, and Cybernetics Part B*, 33(3):367–385, June, 2003.

[6] T. W. Fong and C. Thorpe. Vehicle teleoperation interfaces. *Autonomous Robots*, 11(1):9–18, July 2001.

[7] T. W. Fong, C. Thorpe, and C. Baur. Advanced interfaces for vehicle teleoperation: Collaborative control, sensor fusion displays, and remote driving tools. *Autonomous Robots*, 11(1):77–85, July 2001.

[8] K. Konolige. Large-scale map-making. In *Proceedings of the National Conference on AI (AAAI)*, San Jose, CA, 2004.

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

[10] M. Lewis and J. Jacobson. Game engines in research. *Communications of the ACM*, 45(1):27–48, 2002.

[11] M. Lewis, K. Sycara, and I. Nourbakhsh. Developing a testbed for studying human-robot interaction in urban search and rescue. In *10th International Conference on Human-Computer Interaction*, Crete, Greece, 2003.

[12] R. Meier, T. Fong, C. Thorpe, and C. Baur. A sensor fusion based user interface for vehicle teleoperation. In *International conference on field and service robotics (FSR)*, 1999.

[13] C. W. Nielsen, M. A. Goodrich, and R. J. Rupper. Towards facilitating the use of a pan-tilt camera on a mobile robot. In *Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, Nashville, TN, 2005.

[14] C. W. Nielsen, B. Ricks, M. A. Goodrich, D. J. Bruemmer, D. A. Few, and M. C. Walton. Snapshots for semantic maps. In *Proceedings of the 2004 IEEE Conference on Systems, Man, and Cybernetics*, The Hague, The Netherlands, 2004.

[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] L. C. Thomas and C. D. Wickens. Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, Santa Monica, CA, 2004.

[17] J. Wang, M. Lewis, and J. Gennari. A game engine based simulation of the NIST urban search and rescue arenas. In *Proceedings of the 2003 Winter Simulation Conference*, 2003.

[18] 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.

[19] H. A. Yanco, J. L. Drury, and J. Scholtz. Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Journal of Human-Computer Interaction*, 19(1 and 2):117–149, 2004.