

# Comparing Situation Awareness for Two Unmanned Aerial Vehicle Human Interface Approaches

Jill L. Drury, Justin Richer, Nathan Rackliffe  
The MITRE Corporation  
202 Burlington Road  
Bedford, MA 01730 USA  
{jldrury, jricher, nrackliffe}@mitre.org

Michael A. Goodrich  
Computer Science Dept., 3368 TMCB  
Brigham Young University  
Provo, UT 84602 USA  
mike@cs.byu.edu

**Abstract**—Our goal is to improve the design of human-Unmanned Aerial Vehicle (UAV) interaction so operators can have better situation awareness (SA) of conditions pertaining to the UAVs. We developed a UAV interaction design approach that uses pre-loaded terrain data to augment real-time video data sensed by the UAVs. We hypothesized that augmentation of the video in this manner would provide better SA than a video stream alone. To test the hypothesis, we performed a counterbalanced within-subjects experiment in which the independent variable was video presentation approach. Our results show an increase in comprehension of 3D spatial relationships between the UAV and points on the earth when experiment participants were given an augmented video presentation, as evidenced by a statistically significant difference in participants' mapping accuracy. We believe our results will generalize to situations beyond UAVs to those situations in which people must monitor and comprehend real-time, map-based information.

## I. INTRODUCTION

Consider an aircraft that has crashed in a remote area, or hikers who have become lost in thousands of acres of back country. The National Aeronautics and Space Administration (NASA) Goddard Space Flight Center Search and Rescue Mission Office is considering using unmanned aerial vehicles (UAVs) to search for downed aircraft (NASA, 2006). County rescue organizations are experimenting with UAVs for rural search-and-rescue.

UAVs are promising for search tasks, yet people often have difficulty controlling UAVs and interpreting the data they send to the ground, as evidenced by the fact that UAVs suffer more mishaps per 1,000 flight hours than manned aircraft. More than half of these mishaps have been attributed to problems with human-systems integration (Tveryanas et al., 2005). If the promise of UAVs is to be fully realized, UAV human interface designs need to be improved.

When asked about ways to improve human-system integration, UAV operators repeatedly point to a lack of situation awareness: "Piloting ... is an intensely visual task. Gone are the large field of regard, the subtle 'seat of the pants' inputs and numerous clues which allow...better SA" (Draper, 2005).

Situation awareness (SA) was defined by Endsley (1988) 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. The goal of our research is to improve the design of human-UAV interaction so operators can have better SA of conditions pertaining to the UAVs as well as the activities of distributed (human) team members.

Before developing interaction designs we observed operators working with UAVs in realistic military exercises. As a result, we developed a detailed definition of what it means for UAV teams, in particular, to have SA (Drury et al., 2006). Our observations of UAV operators also led us to determine that the UAV interface designs of current-generation UAVs lack critical contextual information. We felt that providing a narrow field of view via the live video camera only, sometimes called a "soda straw" view because of the analogy of looking through a narrow pipe, was not the optimal way to provide SA.

In an attempt to address shortcomings in SA, we developed a UAV interaction design that uses pre-loaded terrain data to augment real-time video data sensed by the UAVs. (Note that Cooper and Goodrich (2006) developed this type of interface for a small handheld controller.) We felt that augmentation of the video in this manner, which draws on concepts in Drascic and Milgram (1996) and is called the Augmented Virtuality Interface (AVI), would provide better SA than a video stream alone. In particular, we hypothesized that this approach would improve the UAV SA component we identified as comprehension of "3D spatial relationships between the UAV and points on the earth." In other words, we believed that operators would have a better understanding of where the aircraft was with respect to locations on the ground using AVI rather than unaugmented video.

To test this hypothesis, we performed a within-subjects experiment in which the independent variable was video presentation approach. A description of this experiment is contained in section 3, after a presentation of related

literature. Experiment results are contained in section 4 followed by conclusions in section 5.

## II. RELATED LITERATURE

Others have used an augmented virtuality approach for human interfaces to robots: Nielsen et al. (2005), Ricks et al. (2004), and Calhoun et al. (2005). Calhoun et al. term their approach “Picture in Picture.” Quigley et al. (2004) used a “chase plane” perspective to control the UAV on a Personal Digital Assistant (PDA).

There are a few studies that have explicitly examined robot (airborne or ground-based) operators’ SA. Drury et al. (2006) evaluated problems encountered by trainee UAV operators and found that all of them could be at least partially attributed to missing or suboptimal SA. Yanco and Drury (2004) found that search and rescue workers participating in their ground-based robot experiment spent, on average, approximately 30% of the time solely trying to gain or maintain SA, which chiefly consisted of understanding their remote robot’s location, surroundings, and status. In another ground-based robot experiment, Burke et al. (2004) found that “operators spent significantly more time gathering information about the state of the robot and the state of the environment than they did navigating the robot” (p. 86). They reported that 24% of operators’ communications with each other concerned the robots’ state, 14% concerned the robots’ location (“robot situatedness”), and 13% concerned the robots’ surroundings (the “state of the environment”). Clearly, designing interfaces to provide SA in a form that can be more quickly grasped would free up time for activities beyond those undertaken solely to gain or maintain SA.

We needed a way to determine whether our new design would, in fact, provide improved SA. There is a very extensive literature on measuring SA<sup>1</sup>, including whole books (e.g., (Endsley and Garland, 2000)). One means of measuring SA is to focus on examining how well a task is performed, implying that better SA leads to better task outcomes; this is called an *implicit performance measure*. Another class of SA measurement techniques is termed *subjective measures*: these measures pertain to people’s self-assessment of SA. We employed measures from both of these classes, as described in the next section.

## III. EXPERIMENT METHODOLOGY

### A. Overview

Experiment participants performed a search and rescue task in which they were asked to find lost hunters while the UAV flew autonomously between pre-loaded waypoints. They marked hunters’ locations on a topographical map of a type often given to rural search-and-rescue workers. Each participant performed

this task using both interfaces and we examined the differences in positional accuracy with one interface versus the other. We hypothesized that participants would map more accurately with the Augmented Virtuality Interface.

### B. Test Environment Description

Four movie files were loaded onto the same laptop computer: one each depicting what an operator would see during a UAV search sweep. Two show an aircraft avatar and simulated video inset into pre-loaded terrain data and two show only a stationery window fed by a simulated video stream. Four total movie files were required because we needed two sets of positions for lost hunters to avoid participants applying knowledge of hunters’ positions from the first run to the second (two video presentations and two hunter patterns yielded four movie files). Movies were used instead of interactive interfaces because we wished to eliminate the training associated with directing the aircraft and, more importantly, we wished the aircraft to fly the exact same pattern each time so each participant could be guaranteed the same amount of time when hunters were in view. Prior to each run participants were given a pen and a paper topographical map similar to what first responders might use in a rescue situation.

Figure 1 depicts the AVI interface. The center of this screen shows a silhouette of the UAV from behind that changes attitude in real time as the aircraft flies through the virtual environment. The video display is in the inset box. The video is geo-referenced to the pre-loaded map data, meaning that it appears approximately on top of the map area to which it refers. The video inset box changes orientation, becoming trapezoidal and tilting as the viewing angle is changed by the user (in a functional interface) and as the aircraft attitude changes.

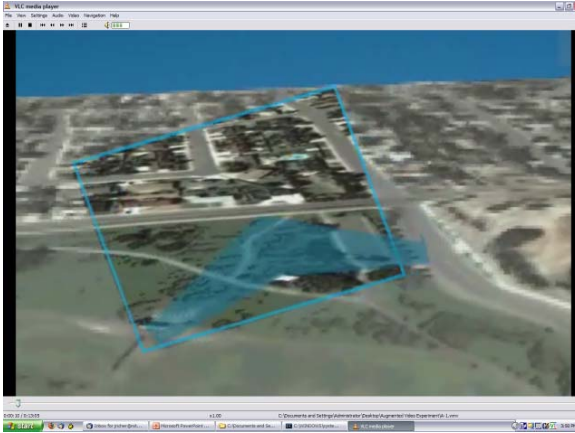
Figure 2 depicts the video stream used as the alternative interface. The video is shown in a stationery window of the same size as the video presentation in the AVI display.

### C. Experiment Participants

In recognition of the fact that organizations such as the Air Force are currently training non-pilots and people without prior search-and-rescue experience to perform UAV surveillance and search tasks, we chose participants who were computer-savvy in general but did not seek out specialists in either piloting or search-and-rescue. Accordingly, twelve people from a local high-technology company participated in the experiment: seven men and five women with a wide distribution of ages (two each in their twenties and thirties, five in their forties, two in their fifties and one person was over sixty). All considered themselves to have at least moderate computer expertise. Six had not used robots previously and the rest had used robots at least once, but primarily robotic toys (no participant claimed extensive experience with robots). Six had operated remote controlled cars or aircraft previously, but primarily in the context of using their children’s toys. Eight play video games at least occasionally while four do not. Of the twelve participants,

---

<sup>1</sup> See, for example: Brickman et al., 1999; Durso et al., 1995; Endsley, 1988; Endsley et al., 1998; Fracker, 1991; McGuinness, 1999; McGuinness and Ebbage, 2002; Scholtz et al., 2004; Taylor, 1990; and Vidulich et al., 1991.



**Fig. 1. Augmented Reality Interface.**

The center of this screen shows a transparent silhouette of the UAV from behind that changes attitude in real time as the aircraft flies through the virtual environment. The video display is in the inset box. The video is geo-referenced to the pre-loaded map data, meaning that it appears on top of the map area to which it refers.

one was a sailplane pilot twenty years ago and one pilots ultralight and powered paraglider aircraft.

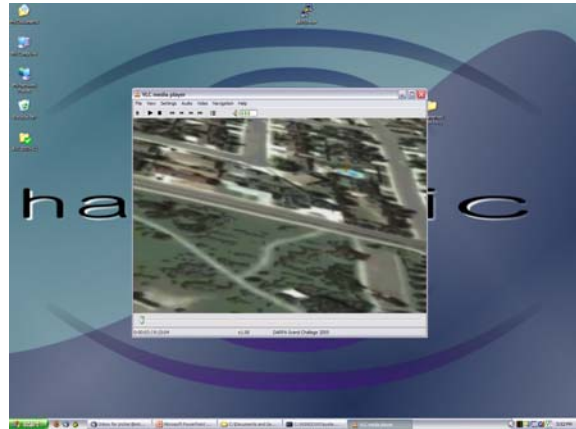
#### D. Experiment Procedure

Each participant was welcomed and received an explanation of the experiment methodology. Participants then answered a set of demographic questions so we could understand their knowledge of computers, video games, robots, flying, and UAVs.<sup>2</sup>

Participants then received training regarding the first interface via pre-scripted materials. They were shown a snapshot with a hunter in view so they would know what to look for. Next they performed a task using the first interface. The task was presented to participants as:

“We’ve heard that several hunters are lost and a UAV is being used to search for them. Your job is to be a sensor operator for this rural search-and-rescue mission. The sensor in this case is a video camera and it is fixed to the UAV (you can’t ‘steer’ the camera except by steering the whole UAV). To simplify training, we have pre-loaded a flight plan so that the aircraft will fly between the preloaded waypoints autonomously. In other words, you will not be directing the aircraft or camera. You will be looking at the information provided by the video camera and noting where the hunters are, as depicted by their blaze orange jackets. When you see each hunter, place a mark on this paper map indicating where you think the hunter is along with a label indicating whether it’s the first hunter you see, second, etc. Don’t worry if you’re not completely accurate in placing each hunter on the map; the point is to use the paper annotations as rough guidance to a

<sup>2</sup> This experiment was approved by the MITRE Institutional Review Board, which did not require participants to sign Informed Consent Forms.



**Fig. 2. Non-augmented video presentation.**

The video is shown in a stationary window of the same size as the video presentation in the augmented display.

team of rescuers who will use 4-wheel drive vehicles to get into the general area where they should be searching. While you are performing this task, please ‘think aloud’<sup>3</sup>: in other words, say what you are thinking while you are performing this task. If you become quiet, I will likely prompt you to say what you are thinking.”

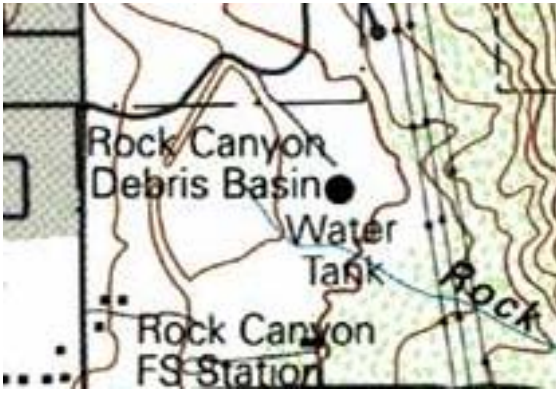
The first task lasted approximately 13 minutes and was followed by a post-run questionnaire. Next, participants received training on the second interface and we began the movie representing the next interface, asking the participant to perform the same search task using the second interface. (The order of interface presentation was alternated for counterbalancing.) Finally, participants answered post-run and post-experiment questions. They were thanked and paid their regular salary for their time but received no other remuneration. We spent approximately one hour total with each participant.

#### E. Data Collection and Measures

This experiment yielded three sources of data. First, participants marked hunter positions on paper maps. A fragment of a sample map is shown in figure 3. They placed a number next to each mark to indicate the order in which the hunters were found. The observer made a note of the time at which each hunter was found because the time and sequence allowed us to determine which hunter was being referenced and thus we could look up the true position of the hunter. We compared the ground truth position with the position marked by participants to obtain the difference, which we measured in millimeters. By performing this comparison, we obtained an implicit measure of SA: the better the accuracy, the better the implied awareness.

The second source of data consisted of post-test questionnaires and other comments given to us after the runs. After each run, participants were asked two Likert scale questions: “I knew at all times where the UAV was

<sup>3</sup> Ericsson and Simon, 1980.



**Fig. 3. Fragment of map.**

Figure 3 illustrates the type of map given to participants to be marked with hunter locations during the experiment.

located in relation to the hunters I found” and “This interface helped me to perform the search task,” where 1 corresponded to “strongly disagree” and 7 corresponded to “strongly agree.” The first question, in particular, pertains to a subjective assessment of participants’ SA. After the second run and post-run questionnaire, participants were asked two final Likert scale questions: “I prefer the first interface I worked with to the second interface” and “The first interface was more suited to the tasks I performed than the second interface.”

Finally, participants were audiotaped, capturing their comments as they performed the tasks. The audiotapes captured comments that indicated participants’ degree of certainty regarding hunter placement, such as when they said, “I have no idea where I am” or “I’m sure this is where this hunter is located.”

#### *F. Mitigating Threats to Validity*

We counterbalanced the order of the interfaces and the two patterns of hunter locations, so that there were four different combinations. The simulated aircraft flew the same search path with each pattern of hunter locations. Because all participants saw the same two hunter patterns on the same aircraft flight path they had exactly the same opportunity to see and mark hunters. The simulated video stream was the same for both interfaces for a particular hunter pattern and was presented at the same size; the sole difference was in how the video was presented.

Different people obviously have different skills in map-reading and spatial orientation. By designing the experiment to be within-subject, individuals’ map-related skills cancel out. It did not matter whether participants had well-developed or poorly-developed map-based skills; what mattered was whether they were able to map hunter locations more accurately with one interface versus the other.

All training and explanatory information given to the participants was completely scripted so each participant received the same information. We did not answer questions during the experiment conduct unless doing so would not affect the results.

## IV. RESULTS

We analyzed differences in participants’ mapping accuracy between the AVI versus video presentation when both are compared to ground truth. It soon became apparent that we had to handle two special cases: missed hunters and multiple locations assigned to the same hunter.

We counted the instances that participants marked two or more different locations for the same hunter. On average, participants assigned multiple locations to 1.4 hunters when using AVI and 2.3 hunters when using the video presentation (note that this difference is not statistically significant). We took all the markings a participant made for the same hunter, calculated the difference between each mark and the ground truth position, and averaged the differences to assign a single accuracy value for that hunter for that participant.

Participants missed an average of 3.11 hunters when using AVI and 3.33 hunters when using video (again, this difference is not statistically significant). Originally we had planned to compute overall accuracy figures by taking into account only the hunters that had been marked. We decided that this would not show a true picture of participants’ spatial knowledge of the environment, however. Participants often declined to mark hunters when they were not sure of their positions; thus if we counted only those that they marked it would skew the accuracy numbers. We assumed that if participants made a complete guess regarding hunters’ positions they would be inaccurate by, on average, half of the width of the map, or 125mm. In fact, participants sometimes marked hunter locations that were off by as much as 183mm, and a significant fraction were off by over 100mm; so we felt it was reasonable to assign a standard value of 125mm inaccuracy for hunters that were not marked.

Given the methodology just described to handle multiple and missed hunter markings, we found that participants’ marked positions were off by an average of 54mm when using AVI and 66mm when using unaugmented video. This difference is statistically significant;  $p < 0.05$  when using a paired t-test with  $df = 11$ . The results can be seen in table 1.

We also analyzed the results of the Likert scale questions asked of participants post-run and post-experiment. Participants felt they had a better understanding of the UAV’s location with respect to the hunters when using AVI versus video (3.9 versus 2.7 on a Likert scale of 1 to 7,  $p < 0.003$ ). Similarly, they felt that AVI helped them to perform the search task more than video (3.9 versus 3.0,  $p < 0.003$ ). Finally, participants preferred the AVI interface to the video interface (5.8 versus 2.2,  $p < 0.0009$ ).

Several participants remarked that they thought the video in the AVI interface moved slower than in the video-only interface. In fact, the video speed was exactly the same in both cases. We believe the AVI video movement seemed slower as a result of the pre-loaded terrain data surrounding the video enlarging the virtual field of view.

TABLE I  
RESULTS OF ACCURACY ANALYSIS

Participant	Missed -AVI*	Mult-AVI#	Ave-AVI	Missed -Video	Mult-Video	Ave-Video
1	2	0	27.55	1	4	38.33
2	4	1	55.50	2	3	42.17
3	1	1	19.67	2	0	38.45
4	3	1	51.92	5	2	84.73
5	7	0	83.82	12	0	125.00
6	3	2	62.00	3	0	58.82
7	3	3	57.55	4	4	100.00
8	4	4	91.91	0	7	91.92
9	1	5	81.67	1	5	73.18
10	0	0	13.00	2	1	32.91
11	2	0	32.55	1	0	27.83
12	4	0	67.42	7	1	81.73
Aves	3.11	1.42	53.71	3.33	2.25	66.26
Std dev	1.83	1.73	25.89	3.61	2.34	31.12

\*Number of hunters not marked when using the AVI interface

#Number of hunters that were assigned multiple locations when using the AVI interface

Participants commented directly on the enlarged field of view, saying “it’s easier to recognize where the UAV is relative to the entire search space.” Another participant noted that the augmented terrain data “was distracting at first but definitely helped orient me.” We also received positive feedback regarding the “chase plane” view of the aircraft avatar: “seeing attitude of the plane was useful.”

Some participants wanted additional support for the search task, such as “bookkeeping support” (meaning, automated help in numbering sightings) and the “compass direction UAV was flying.”

## V. CONCLUSIONS

This experiment yields empirical confirmation that providing contextual information via pre-loaded terrain data, as well as a transparent avatar in a “chase plane” view, aids the SA of UAV operators. Specifically, this design approach helps operators by assisting comprehension of 3D spatial relationships between the UAV and points on the earth.

Note that this experiment focused on the visualization aspects of the interface only and not the input mechanisms that would normally be used when interacting with a UAV. But because participants performed a generic search task and did not interact with the interface in any way specific to a UAV, we feel that the results will be applicable to other domains that require people to monitor real-time, map-based data. An example of a different situation in which the results may apply is when security personnel monitor inputs from remote, ground-based robots roving the grounds of an industrial plant.

## ACKNOWLEDGEMENTS

This work was supported in part by the United States Air Force Electronic Systems Center and performed under MITRE Mission Oriented Investigation and

Experimentation (MOIE) Project 03057531 of contract 19628-94-C0001.

## REFERENCES

- [1] B. J. Brickman., L. J. Hettinger, D. K. Stautberg, M. W. Haas, M. A. Vidulich, and R. L. Shaw. “The global implicit measurement of situation awareness: implications for design and adaptive interface technologies.” In M.W. Scerbo & M. Mouloua (Eds.), *Automation Technology and Human Performance: Current Research and Trends*. Mahwah, NJ: Lawrence Erlbaum Associates, 1999.
- [2] 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.” *Human-Computer Interaction* 19(1-2), pp. 85 – 116.
- [3] 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 *Proc. SPIE Vol. 5802, Enhanced and Synthetic Vision 2005*, J. G. Verly, Ed., May, 2005, pp. 219 – 230.
- [4] J. L. Cooper and M. A. Goodrich. “Integrating critical interface elements for intuitive single-display aviation control of UAVs.” In *Proceedings of SPIE DSS06 - Defense and Security Symposium*, Kissimmee, FL, USA. April 17-21, 2006.
- [5] M. H. Draper, G. L. Calhoun, M. J. Patzek., and G. L. Feitshans. “UAV Human Factors Research within AFRL/HEC.” Presentation to 2<sup>nd</sup> Annual Human Factors of UAV Workshop, Mesa, AZ, May 2005.
- [6] D. Drascic and P. Milgram. “Perceptual Issues in Augmented Reality.” In *Proceedings of SPIE Vol.2653: Stereoscopic Displays and Virtual Reality Systems III*, San Jose, CA, 1996.
- [7] J. L. Drury, L. Riek, and N. Rackliffe. “A Decomposition of UAV-Related Situation Awareness.” In *Proceedings of the Human-Robot Interaction 2006 Conference*, March 2006.
- [8] F. T. Durso, T. R. Truitt, C. A. Hackworth, J. M. Crutchfield, D. Nikolic, P. M. Moertl, D. D. Ohrt, and C. A. Manning. “Expertise and chess: comparing situation awareness methodologies.” In *Proceedings of the International Conference on Situation Awareness*, Daytona Beach, FL., 1995.
- [9] M. R. Endsley and D. J. Garland, eds. *Situation Awareness Analysis and Measurement*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2000.
- [10] M. R. Endsley. “Design and evaluation for situation awareness enhancement.” In *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA, Human Factors Society, 1988.
- [11] M. R. Endsley, S. J. Selcon, T. D. Hardiman, and D. G. Croft. “A comparative analysis of SAGAT and SART for evaluations of situation awareness.” In *Proceedings of the 42nd annual meeting of the Human Factors and Ergonomics Society*, October 1998.
- [12] K. A. Ericsson and H.A. Simon. “Verbal Reports as Data.” *Psychological Review*, Vol. 87, pp. 215 – 251, 1980.
- [13] M. L. Fracker. “Measures of situation awareness: an experimental evaluation.” AL-TR0191-0127, Armstrong Laboratory, Wright-Patterson AFB, Ohio, 1991.
- [14] B. McGuinness. “Situational awareness and the CREW awareness rating scale (CARS).” In *Proceedings of the 1999 Avionics Conference*, Heathrow, 17 – 18 November. ERA Technology Report 99-0815 (paper 4.3), 1999.
- [15] B. McGuinness and L. Ebbage. “Assessing human factors in command and control: workload and situational awareness metrics.” In *Proceedings of the 2002 Command and Control Research and Technology Symposium*, Monterey, CA., 2002.
- [16] NASA Goddard Space Flight Center. Web page for search and rescue: <http://searchandrescue.gsfc.nasa.gov/techdevelopment/sar2.htm>, accessed April 2006.

- [17] 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 IEEE International Workshop on Robots and Human Interactive Communications*, Nashville, TN, 2005.
- [18] M. Quigley, M. A. Goodrich, and R. W. Beard. Semi-Autonomous Human-UAV Interfaces for Fixed-Wing Mini-UAVs. *Proceedings of IROS 2004*. Sept 28-Oct 2, 2004, Sendai, Japan.
- [19] B. Ricks, C. W. Nielsen, and M. A. Goodrich. "Ecological Displays for Robot Interaction: A New Perspective." In *Proceedings of IROS 2004*, Sept 28-Oct 2, 2004, Sendai, Japan.
- [20] J. Scholtz, B. Antonishek, and J. Young. "Evaluation of a Human-Robot Interface: Development of a Situational Awareness Methodology". In *Proceedings of the Hawaii International Conference on Systems Sciences*, January 2004.
- [21] R. M. Taylor. "Situational awareness rating technique (SART): The development of a tool for aircrew systems design." In *Situational Awareness in Aerospace Operations (AGARD-CP-478)* pp. 3/1 – 3/17. Neuilly-sur-Seine, France: NATO-AGARD, 1990.
- [22] A. P. Tvaryanas, B. T. Thompson, and S. H. Constable. "US Military Unmanned Aerial Vehicle Mishaps: Assessment of the Role of Human Factors using HFACS." 311<sup>th</sup> Performance Enhancement Directorate, US Air Force, Brooks AFB, TX, 2005.
- [23] M. A. Vidulich, F. G. Ward, and J. Schueren. "Using the subjective workload dominance (SWORD) technique for projective workload assessment." *Human Factors*, Vol. 33, No. 6, pp. 677 – 692, 1991.
- [24] H. A. Yanco and J. 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.