The Effect of Bandwidth on Operator Control of an Unmanned Ground Vehicle

Jon French
Micro Analysis and Design
Orlando, FL
jfrench@maad.com

Thomas G. Ghirardelli
Jennifer Swoboda
AMSRL-HR-SE
Aberdeen Proving Ground, MD
tghirard@arl.army.mil, jcrouch@arl.army.mil

ABSTRACT

Unmanned ground vehicles (UGV) are a critical conceptual capability for Future Combat Systems where they will be used for reconnaissance, surveillance and target acquisition. The objective of this study was to ascertain the effect of the relative loss of signal strength or bandwidth, represented by 2, 4, 8 and 16 frames per second (fps) on a UGV operator's workload and situational awareness. A secondary objective was to measure the effect of different modes of control (joystick, voice command, joystick and voice and completely autonomous or passive control) on UGV control. The UGV was supplied by ARL/CISD and the tests were conducted on a field at Aberdeen Proving Ground. Participants were 22 soldiers recruited from the 16th Ordnance Battalion at Aberdeen Proving Ground. Each was randomly assigned to one of the four control device between-group conditions and experienced all four levels of the bandwidth within group conditions. Reduced signal strength had a significant impact on operator performance. The average time to navigate the course increased for 2 and 4 fps compared to 8 and 16 fps. Voice command was more difficult than the other modes of control. Visual and cognitive workload scores also were affected by bandwidth with 8 and 16 fps resulting in the best scores. Finally, NASA TLX workload was improved for 4, 8 and 16 fps compared to 2 fps. The modern battlefield poses significant challenges for UGV operations due to loss of signal strength from a variety of sources such as distance, obstacles, multiple UGVs on the same channel or electronic jamming. These results suggest that optimal bandwidth for UGV operation should remain above 8 fps. Also, the passive or completely autonomous condition provided no advantages in terms of target acquisition or workload. These results will provide empirical detail for a computer model of UGV operator performance.

ABOUT THE AUTHORS

Jonathan French is the Lead Behavioral Neuroscientist for Micro Analysis & Design, Inc. He received an M.S. and PhD in Physiological Psychology at Colorado State University. He completed postdoctoral training at Cornell University Medical College and worked as a research pharmacologist at the Parke-Davis Co and the University of Michigan. Dr French served for 11 years as a civilian scientist for the US Air Force Research Laboratory at Brooks AFB in the area of sustained operations. He is the author of over 50 articles on Human Factors issues in military and civilian settings.

Thomas G. Ghirardelli is a Research Psychologist at the U.S. Army Research Laboratory, Human Research and Engineering Directorate. He also holds an appointment at Johns Hopkins University as a Lecturer in the Krieger School of Arts and Sciences. His research is focused on the human cognitive and perceptual factors associated with target acquisition and human-system interfaces. Previously, he held a Postdoctoral Fellowship at the National Institutes of Health, conducting basic research in cognitive neuroscience. He received his A.B. in Psychology from Davidson College in 1986, an M.S. degree in Experimental Psychology from Villanova University in 1992, and M.A. and Ph.D. degrees in Experimental Psychology from Johns Hopkins in 1994 and 1996 respectively.

Jennifer Swoboda is a Research Psychologist at the U.S. Army Research Laboratory, Human Research and Engineering Directorate. Her current work focus is on the development of command and control human performance models in order to understand the effects of information quality and uncertainty on operator decision making. She also conducts research in the area of multi-tasking. She received her B.A. in Psychology from Washington College in 1991.

The Effect of Bandwidth on Operator Control of an Unmanned Ground Vehicle

Jon French
Micro Analysis and Design
Orlando, FL
jfrench@maad.com

Thomas G. Ghirardelli
Jennifer Swoboda
AMSRL-HR-SE
Aberdeen Proving Ground, MD
tghirard@arl.army.mil, jcrouch@arl.army.mil

INTRODUCTION

Unmanned ground vehicles (UGV) are planned to be a critical capability for the US Army's Future Combat Systems. For example, they could provide remote sensing for reconnaissance, surveillance and target acquisition in hazardous or difficult environments. In addition, they could serve as re-supply or equipment carriers. DARPA recently hypothesized that reconnaissance missions were the most likely to benefit substantially from mobile robot technology, particularly in urban settings (Krotkov, et al., 1999). They are envisioned to provide increased situational awareness and understanding for commanders as well as that of the Objective Force Warrior.

Most current UGVs use traditional computer interface technologies to navigate. However, there is a need for empirical data to help select the optimal control technologies and conditions of use. One objective of this study was to measure the effect of different modes of control of an unmanned ground vehicle (UGV) on situational awareness and workload The different modes of control considered were 1) a standard joystick controller, 2) a voice control device that uses a limited vocabulary of commands, 3) a combination of voice control and joystick, and 4) a completely autonomous or passive condition.

The modern battlefield poses significant challenges for the UGV due to loss of signal strength from a variety of sources such as distance, obstacles, or electronic jamming. Accordingly, in order to make this analysis of operator UGV control interfaces operationally meaningful, we examined simulated degradation of signal strength at 2, 4, 8 and 16 frames per second (fps).

METHODS

Participants were assigned randomly to one of the four control device between-group conditions (joystick, voice, voice and joystick, and autonomous) and experienced all four levels of the video quality conditions. Participants were trained in the operation of the computer and control of the UGV using the control device they would use in the experiment. Participants controlled the UGV through a course under each of the four video quality conditions (2, 4, 8, 16 fps) with a 10-minute break in between each. At the end of each video quality condition, participants completed questionnaires designed to assess the workload associated with each condition.

Participants

Participants were 24 soldiers (2 female) recruited from the US Army Ordnance Center and School (16th Ordnance Battalion) at Aberdeen Proving Ground, MD. Two participants (both male) were dropped from the data analysis because they failed to complete all the conditions within the experimental session and thus represented incomplete data. These failures were partly the result of mechanical difficulty with the UGV and partly the result of the participant making too many errors. Mean age of the included participants was 20.5 years. Median age was 19 years, with an inter-quartile range of 18 to 21 years. The participants all had visual acuity correctable to 20/30 in both eyes and normal color vision and were screened by questionnaire for medication that might interfere with their vision or cognitive abilities. The participants held ranks from E-1 to E-3 and came from the range of military occupational specialties (MOS) represented by the Ordnance Center and School.

Apparatus

The UGV (RWI, Model #ATRV) was modified in house at ARL-CISD and the dimensions of the main chassis were 46 (L) x 32 (W) x 28 (H) inches (Figure 1). The UGV was controlled by the different control devices connected to the ground station, which consisted of a computer, monitor, control interfaces, and RF transceiver.

There were two different control devices: 1) a standard joystick controller (Microsoft, SideWinder Precision 2, and 2) a voice control device that uses commercial voice-recognition software and a limited vocabulary of pre-programmed commands to control movement speed, direction, and distance. A third control condition consisted of the combination of the two and participants were asked to attempt to divide their time using the joystick and voice controls about equally. Participants were tested individually while seated in front of a computer. These three conditions represented the active control condition. Finally, a videotape was made of an expert driving the UGV around the course which was shown to participants in a fourth condition, a completely autonomous or passive operation condition. This group viewed the UGV video individually in a quiet laboratory at the ARL HRED facility on Aberdeen Proving Ground. A frame grabber was used to present the video image in this condition to 2, 4, 8 and 16 fps.

For driving a high-resolution color charge coupled device camera (Chugai, Model #CCD-Z14) with a maximum field of view of 48.6 degrees visual angle was used. The camera was center mounted on top of the vehicle, tilted slightly downward, and fixed in a straightforward position. The image from the camera was transmitted as an analog signal via radio to the control station where it was converted to a digital image and presented on a second computer monitor in black and white. The different video quality conditions (i.e., frame rates) were achieved by the subsampling of frames from the video image at four different rates, 16, 8, 4, and 2 fps.

Figure 1. Unmanned ground vehicle

Four courses on an otherwise empty grassy field at Aberdeen Proving Ground were plotted. Wooden stakes connected by 3.25-inch yellow plastic tape approximately 18 inches off the ground marked the course boundaries on both sides. One course served as a training course. This training was designed to familiarize them with the operation of the particular mode of control they were assigned as well as the tasks they were to perform in the experiment, including examples of the same targets they were to identify. They gained practical experience by actually driving the vehicle over terrain similar to that in the experiment on a course comparable to that used in testing. This generally took no longer than 20 minutes for participants to achieve the criterion of navigating the training course with no navigation errors. The participants and ground station were inside a trailer parked near the course. During all training and testing, participants were not able to directly view the UGV.

The remaining three courses were used for data collection. One course consisted of sharp box turns, one of sharp S-turns and one of diagonal tracks. Each course was approximately 140 feet long. The targets consisted of four human-sized olive drab plywood silhouettes, an ammunition can, and an inert antitank mine. Two of the silhouettes were fitted with a U.S. Army load-bearing vest (LBV) and were designated as soldiers or combatants. The two silhouettes without the LBV were identified as civilians or refugees. The targets were placed in predetermined positions around the course for each run.

Procedures

Each participant experienced four levels of the video quality in four separate runs of the course. A run consisted of controlling the UGV through all three different courses. The order of presentation of the video quality conditions was varied randomly.

While navigating, participants were asked questions designed to assess their SA and were asked to identify the target objects. The SA questions were designed to assess basic awareness. For example, participants were asked how many right (or left) turns they had made on that course and were asked to estimate how long they had been on that course. Participants answered the SA questions aloud and indicated to the experimenter when they identified the targets by saying the name of the target. The experimenter recorded the accuracy of their responses to the SA questions and targets.

The time each target was identified was also recorded using a stopwatch. Another experimenter measured the time it took to navigate the course and the number of course errors (e.g., the number of times the vehicle made contact with the tape or stakes marking the outside border of the course).

Participants in the passive condition (n=6) were seated in a laboratory and viewed the video on the same computer used in the active conditions. They performed the same SA and target identification tasks as in the active conditions. The video was presented under each of the same video quality conditions (2, 4, 8, 16 fps) and were shown the same target objects. Before viewing the video, these participants received training designed to familiarize them with the tasks but did not engage in any remote driving. Participants in this condition had no control over where the UGV went.

At the completion of each video quality condition, participants in both the active and passive conditions completed questionnaires designed to assess the workload associated with the task in each condition. We collected subjective measures for visual, auditory, cognitive, and psychomotor (VACP) components of workload (Aldrich, Szabo, & Bierbaum, 1989) and also using the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart & Staveland, 1988).

RESULTS

Navigation

For each participant in the active conditions, there were two means to assess navigation ability. The time to complete each run and the number of errors on each run. The navigation duration and course error data were submitted to a 3 (Mode of Control) x 4 (Video Quality) Mixed Analysis of Variance (ANOVA) with Mode of Control as the between-subjects factor and Video Quality (frame rate) as the within-subjects factor. There were only three levels of control device because these measures did not apply to the passive condition where participants did not control the UGV. The Greenhouse-Geisser adjustment was used for all ANOVA to account for the small sample size. A paired comparison t-test was used for all *post hoc* comparisons with $p < .05$ set as the significance level.

For the navigation duration data shown in Figure 2, we found a significant main effect of mode of control [$F(2, 43) = 33.6$, $p < .001$] and of video quality [$F(2, 102) = 16.47$, $p < .001$] but no interaction. *Post hoc* tests revealed that the main effect comparison for Voice control differed from Joystick or Joystick plus Voice conditions. *Post* hoc tests on the effect of video quality revealed significant differences between 2 and compared to 4, 8 and 16 fps and between 4, 8 fps compared to 16 fps. These results are shown in Figure 2. No effects were found for navigation error data

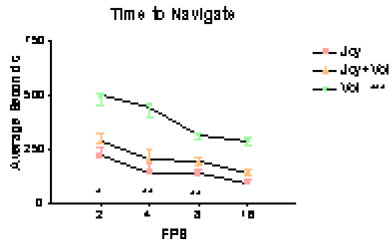

Figure 2. Time to navigate the UGV around the course shown for the control devices across all frames per second (FPS). FPS 2 is different from FPS 4, 8 and 16 (*) and FPS 4 and 8 are different from FPS 16 (**). Voice is different from Joy Stick and Joystick + Voice (***).

Target Identification

The number of targets correctly identified as each participant navigated the UGV around the course in the active control conditions or as s/he viewed the video in the passive condition were recorded. The percentage of targets correctly identified by each participant were analyzed with a 4 (Mode of Control) x 4 (Video Quality) Mixed ANOVA. No significant effects of Video Quality, Mode of Control nor for the interaction were found.

Situation Awareness

The accuracy of each participant's responses to the various questions designed to assess SA was recorded. The proportion of SA questions answered correctly by each participant were calculated and submitted to the same ANOVA. No significant effects of Video Quality, Mode of Control, nor the interaction were found.

Workload

Figure 3 shows the mean workload ratings for the Visual component of the VCAP profile as a function of mode of control and video quality. The scores were evaluated in the same ANOVA. A significant effect of Video Quality on workload [$F(2, 28) = 13.1$, $p < .001$] but no effect of Mode of Control and no significant interaction were found. *Post hoc* tests on the main effect of Video Quality revealed significantly greater workload at 2 fps, compared to 4, 8 and 16 fps and 4 fps differed from 8 and 16 fps.

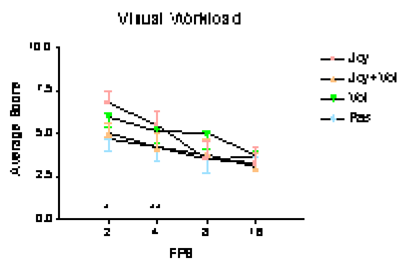

Figure 3. Visual workload effort for teleoperating the UGV around the course shown for the control devices across all frames per second (FPS). FPS 2 differed from FPS 4, 8 and 16 (*) and FPS 4 and 8 differed from 16 (**).

The workload scores for cognitive data were also evaluated using the same ANOVA. A significant effect of Video Quality was found [F(2, 37) < .002] as shown in Figure 4. No effects were found for Mode of Control nor significant interaction. *Post hoc* tests on the main effect of Video Quality revealed a significant difference between 2 fps with 8 and 16 fps. We found no effects of video quality, mode of control nor any interaction for the auditory or psychomotor subscales.

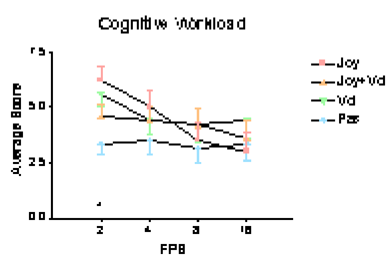

Figure 4. Cognitive workload effort for teleoperating the UGV around the course shown for the control devices across all frames per second (FPS). FPS 2 differed from FPS 8 and 16 (*).

The test of workload using the NASA TLX score was significant for Video quality [F(3, 30) < .03]. The 4, 8 and 16 fps were different from the 2 fps.

In order to assess if the amount of training was adequate the number of test runs from first to the last were plotted for each mode of control. These results are shown in Figure 5. It can be seen that for the Joystick and combination conditions, level of training was adequate as navigation time did not differ from the first exposure to the test track to the last. However, training may not have been adequate for the Voice condition since navigation time decreased until about the sixth exposure and then seemed to increase again.

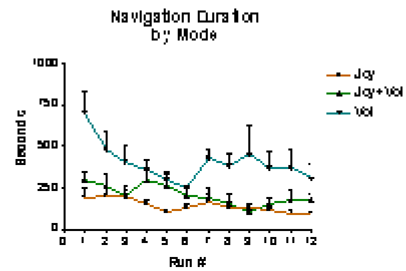

Figure 5. Navigation duration for teleoperating the UGV around the course shown for the control devices from the first runs to the last.

DISCUSSION

The primary impediment to UGV navigation was reduced signal strength, simulated by frame rate degradation. Reduced signal strength resulted in significant increases in mean time to navigate, visual, cognitive and NASA TLX workload measures. Signal strength reductions are a constant threat to UGV operations in the field and must be considered during UGV mission execution. Perhaps some control devices are more impacted by signal degradation such as Voice control.

Mode of control had an effect on navigation of the UGV. Participants in the Voice condition took significantly longer to navigate than those in the other two active control conditions. Participants in the Joystick condition did not differ from those in the Joystick + Voice condition, however.

Voice control over the UGV was not as effective in the current experiment as was Joystick or Joystick and Voice control. There are a few possible explanations for this. It may be that the course used was not conducive to voice control since there were numerous sharp turns and adjustments required for navigation. Voice control may be more appropriate to open space navigation such as traversing great distances. The training was effective as shown by Figure 5 for Joystick but not for Voice control so perhaps more training would have made Voice as effective as Joystick. The fact that the voice commands themselves were difficult to remember lends credence to this possibility and the use of a more "natural language" structure could have made a difference. Such a language base is under development at ARL/CISD and will be implemented soon.

No evidence for an effect of mode of control on target identification performance nor on our measures of SA were found. Perhaps our measures were not sensitive enough to detect differences. The passive or completely autonomous condition did not seem to do any better nor worse than the active viewing conditions. However, if the UGV ever encountered difficulties under operational conditions, it would certainly be decidedly worse.

Video quality's effect on performance was most noticeable in the lowest frame rate, 2 fps. This is not a surprising result but provides evidence for the importance of having sufficient bandwidth available for teleoperation. The 16 fps condition provided ideal viewing condtions for teleoperation of the UGV. Our results suggest that in order for control conditions to be as effective as the ideal, no less than 8 fps should be permitted in the field.

The results of these experiments will be used to refine the development of high fidelity computer models of UGV operations. These models will be used to test future innovations in design concepts far more timely and economically than would be possible empirically. These models will be used to extend empirical findings of human performance and to address subsequent questions about UGV operations across a wider variety of tasks and situations and will be integrated with other models undergoing development for the Future Combat Systems concept. Future experiments will examine similar issues related to control of UGV platforms, including control of multiple UGVs, during different types of video degradation and conditions of less-than-perfect autonomy.

REFERENCES

Aldrich, T.B., Szabo, S.M., & Bierbaum, C.R. (1989). The development and application of models to predict operator workload during system design. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton and L.van Breda (Eds) *Applications of Human Performance Models to System Design* (pp. 65-80). New York: Plenum Press.

Hart, S.G., & Staveland, L. (1988). Development of the NASA Task Load Index (TLX): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload* (pp. 139-183). Amsterdam: North-Holland.

Krotkov, E., & Blitch, J. (1999). Tactical Mobile Program. *The International Journal of Robotics Research, 18 (7),* 769-776.