Improving Tele-robotic Navigation through Augmented Reality Devices

Richard T. Stone[1,2], Ann Bisantz[2], James Llinas[2], Victor Paquet[2]

[1] Iowa State University, Ames IA
[2] University at Buffalo, the State University of New York, Amherst NY

Navigation is an essential element in many telerobotic operations especially those that involve time pressure and complex terrains. This study demonstrates that properly designed augmented reality interfaces can significantly aid human operators in the task of land based tele-robotic navigation. Participants in this study were assigned to one of three experimental conditions (control, sonification interface, visual interface). All participants were trained in tele-robotic landmine detection and navigation, the groups performed these tasks first with augmentation (session one) and then without (session two/transfer task). The results show that participants in the sonification interface group outperformed the visual interface and control group in terms of primary task sensitivity, the number of sectors cover and the number of sectors repeated. The results further indicated that a well designed AR navigation interface could serve as an effective training tool.

INTRODUCTION

To some degree all tele-robotic systems rely on human supervision and control to conduct their intended operations. Often the human operator is responsible for monitoring information feeds related to a specific task while at the same navigating, if not actively maneuvering the robot.

Navigation is an important aspect of tele-robotic control which requires that the user knows where the robot is, how it will get to a new location and how to deal with the environmental factors encountered along the way (Steinfeld et al., 2006). Past work has revealed that the operators tended to focus their attention resources on navigation related video information while driving a robot using standard 2d feedback system (Nielsen et al., 2007). Alternative means of presenting or further defining navigation information could potentially increase maneuvering accuracy and free up resources that could be useful in performing a primary tele-robotic task.

One means of enhancing navigation information is through the use of augmented reality (AR) interfaces. Introducing AR interfaces as a means of conveying information has been shown to be effective in increasing accuracy in task performance (Kok-Lim et al. 2003, Bristow et al., 2002, Boud et al., 1999).

AR interfaces have also been shown to be effective instructional technologies that lead to increased learning (Hamza-Lup et al. 2004, Schwald et al. 2003). The importance of this learning aspect cannot be understated. In the event that an AR device were to fail during a tele-robotic missions, the human operator must still be ability to demonstrate a level of navigational competency to completing a task.

The goal of the study described in this article is to demonstrate the benefits of using a properly designed augmented reality interface for enhancing a human operator's ability to accurately navigate a tele-robotic system in the context of a target detection task.

Interface Design

As part of a larger study the domain of tele-robotic navigation in the area of landmine detection was studied using archival review, interviews and direct observation. The general task of tele-robotic landmine detection is often visually dominated. In many cases an operator may have to scan as many as five visual inputs. A major factor linked to successful field navigation was an operator's ability to accurately track a robot's relative position when in a search area. Specifically, the ability to understand were in a 2D localization a robot was in relation to the search area is related to how accurately an human operator could maintain a search pattern which does not lead to missing part of the search area or accidentally repeating parts that were already inspected.

The accurate use of relative position information related to path duplication and path deviation from and intended search area were determined to be factors related to superior navigation technique. As such this information was thought to be a good candidate for augmentation.

Properly designing a AR interface requires that the human cognitive and physiological needs are carefully considered in the context of the task and environment they will be working in. For this reason a human centered method for the design of augmented reality devices was used and the auditory system was identified as a primary candidate for augmentation (Stone, 2008). After reviewing the domain it was determined that using the auditory sensory system to convey relative position information would not create any cross or intra-modal interference. The sonification AR interface used for this device was capable of create a background sonification that indicated if the robot was repeating a previously inspected area (high pitch tone), was outside of the search area (low pitch tone) or operating within parameters (no tone). Using auditory warning has been shown to be effective for conveying time-relevant or continuously changing information that may require immediate action as is the case in our study (Welch and Warren, 1986).

Our analysis revealed that using a visual AR interface would further divide an already heavily taxed visual resource. On the surface however a visual overlay would appear to be a viable option. To test this we created a visual AR interface. This interface would visually overlay that displays text to indicate an area as being repeated "Repeating Area" or that the robot is outside of the search area "Outside Search Area", and when the robot is within parameters no message is displayed.

Experimental Materials and Setting

In addition to the AR devices a landmine/UXO detection robot and landmine field were constructed and utilized. The robot was controlled by and responsive to a control unit that had capabilities mirroring those of current tele-robotic systems used in military and civilian ordinance operations.

The robot was operated in a large scale landmine area (266 sq feet), setup in an indoor laboratory. This area included aspects of a real world minefield and included trees, grass, rocks, geomorphic obstacles, and different types of soil. Vines and branches were used to obscure participants' view of the field, such that aspects of the field were visible, but participants did not have a clear view of the minefield or items to be detected.

Between sessions the landmine field, mine and distracter placements were altered, to create a different configuration. This was necessary to ensure that participants did not memorize the terrain (which could otherwise have affected navigation outcomes in session two).

METHOD

Hypotheses

In this study it was hypothesized that properly designed AR interfaces focused on enhancing task critical navigation information would, (1) increase a human operator's ability to effectively cover more ground while avoiding path duplication and (2) that the resulting navigation control will lead to more accurate primary task performance (in this case more accurate detection of landmine). It is further hypothesized (3) that participants who were trained using a properly designed AR interfaces would demonstrate improved performance when compared to participants who were trained without augmented devices. Finally, it was hypothesized (4) that a properly designed AR interface will outperform an AR interface that is less appropriately designed even when the same information is being augmented.

Participants

There were 21 participants in this portion of the study (12 male and 9 female). The participants ranged in age from 21 to 24 years of age, all had (20/20) or corrected to normal vision, normal hearing and no prior tele-robotic training.

Independent and Dependent Variables

There were two independent variables: interface condition at three levels (control, sonification interface and visual interface); and session at two levels (session 1, using augmentation, and session 2, without the use of augmentation). The control group did not utilize augmented interfaces in either condition. Interface condition was a between-subject condition, and session was a within-subjects condition.

Dependent measures reported here include, the number of sectors covered (the number of 1 sq foot sectors successfully driven over and inspected), the number of sectors repeated (the number of 1 sq foot sectors that were driven over more than once) and sensitivity (ratio of hits to false alarms reported as d').

Experimental Protocol

The 21 participants were randomly assigned to one of three experimental conditions (control, sonification interface, visual interface), and performed the task twice, in two sessions. The first session assessed how the use of augmented reality interfaces (sonification and visual) affected performance. In the second session (conducted 7 days later for each subject) participants performed the

same basic task but without any augmented reality devices. The results of this session made it possible to assess the potential for the augmented reality interfaces to be used as training technologies.

Session one was divided into three phases. Phase one involved screenings to ensure normal vision and hearing, followed by informed consent. In phase two, participants were trained in the principles of tele-robotic landmine detection and search/navigation techniques. During this period participants learned how to operate the robot. Participants assigned to an augmented group (sonification or visual) were familiarized with the use of their assigned augmented reality interface in addition to the standard interface. In the third phase, participants performed the task of landmine detection (using their assigned interface) in the landmine area. The participants were given a maximum of 900 seconds to clear the field (note that none of the groups in this part of the study were found to have any significant difference in the time taken to complete the course).

In the second session, participants performed the task of landmine detection, with the non-augmented interface (this was the same interface used by members of the control group in both sessions). Following this, participants were formally debriefed and given the opportunity to ask questions.

RESULTS

The results from the experiments are discussed for session one and session two as they relate to the number of sectors covered, the number of sectors repeated, and sensitivity (d').

Number of Sectors Covered

A mixed effects ANOVA showed a main effect of experimental condition on the number of sectors covered ($p < 0.000$), a main effect of session ($p = 0.028$) but no session*experimental condition interaction ($p = 0.458$). Based on these results, main effects of interface conditions were tested within each session, separately. There were main effects of interface condition for session one ($p = 0.027$) and session two ($p < 0.001$). The mean sectors covered by experimental condition are shown in Figure 1 for both sessions.

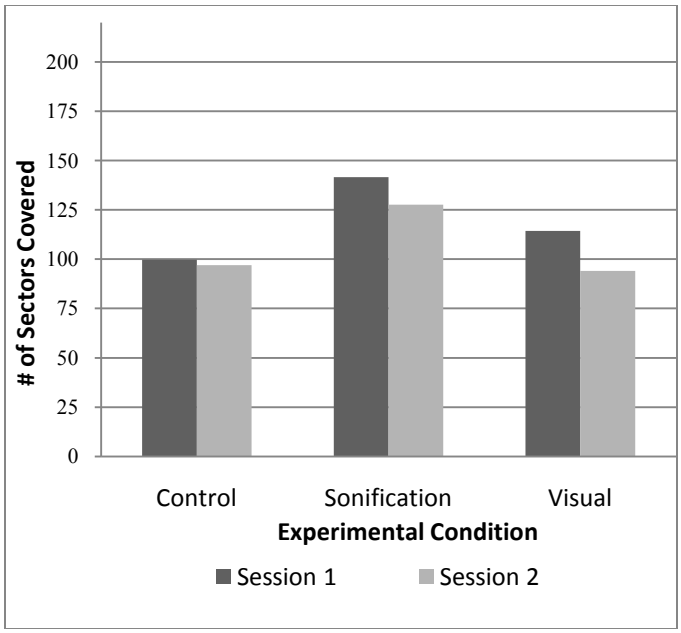

Figure 1: Sectors Covered by Experimental Conditions across session.

For session one, Tukey's test revealed that the number of sectors covered by the participants using the sonification interface (*M* Sectors= 142, $p = 0.024$) was significantly higher than the control group (*M* Sectors= 100). The participants using the visual interface (*M* Sectors = 114, $p = 0.632$) were not found to have covered a significantly larger area than did participants in the control group.

Likewise, participants assigned to the sonification interface ($p = 0.143$) did not demonstrated a significantly higher number of sectors covered when compared to participants in the visual condition but did so descriptively.

For session two, Tukey's test revealed that the number of sectors covered by the participants using the sonification interface (*M* Sectors= 128, $p = 0.014$) was significantly higher than the control group (*M* Sectors= 97). The participants using the visual interface (*M* Sectors = 94, $p = 0.154$) were not found to have covered a significantly larger area than did participants in the control group.

Similar to the results from the first session participants assigned to the sonification interface ($p = 0.303$) did not significantly outperformed those assigned the visual interface in term of area covered given a CI of 90%.

These findings show that the area covered by an operator can be significantly increased through the use of an appropriately selected augmented reality interfaces (i.e. the sonification interface). Importantly, the use of the sonification device in session one led to improved performance in the second session, even though participants were no longer using the augmented

interface, suggesting that the interfaces had a positive training effect.

Number of Sectors Repeated

A mixed effects ANOVA showed a main effect of experimental condition on the number of sectors repeated ($p < 0.01$), a main effect of session ($p = 0.05$) but no session*experimental condition interaction ($p = 0.901$). Based on these results, main effects of interface conditions were tested within each session, separately. There were main effects of interface condition for session one ($p = 0.023$) and session two ($p < 0.006$). The mean sectors repeated by experimental condition are shown in Figure 2 for both sessions.

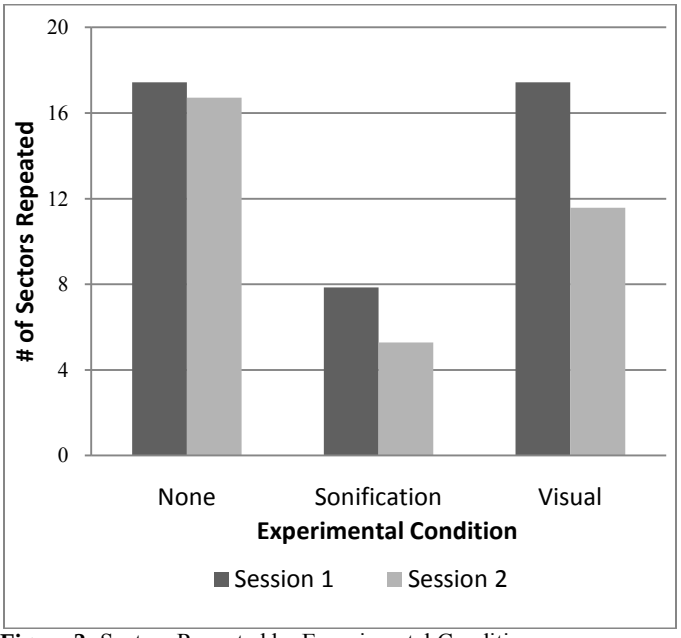

Figure 2: Sectors Repeated by Experimental Conditions across session.

For session one, Tukey's test revealed that the number of sectors repeated by the participants using the sonification interface (*M* Sectors= 7.9, $p = 0.100$) was significantly lower than the control group (*M* Sectors= 17.5). The participants using the visual interface (*M* Sectors = 17.4, $p = 0.733$) were not found to have repeated a significantly lower number of sectors than did participants in the control group.

Participants assigned to the sonification interface ($p = 0.022$) had a lower number of sectors repeated when compared to participants using the visual interface.

For session two, Tukey's test revealed that the number of sectors repeated by the participants using the sonification interface (*M* Sectors= 5.2, $p = 0.014$) was significantly lower than the control group (*M* Sectors= 16.7). The participants using the visual interface (*M* Sectors = 11.5, $p = 0.997$) were not found to have repeated a significantly smaller number of sectors than did participants in the control group.

Similar to the results from the first session, participants assigned to the sonification interface ($p = 0.012$) repeated a significantly lower number of sectors than participants assigned the visual interface.

Similar to sector coverage, these findings show that the area repeated by an operator can be significantly decreased through the use of an appropriately selected augmented reality interfaces (i.e. the sonification interface).

An appropriately selected AR device when used as training technologies has been shown to lead to operators that are less likely to repeat an area already traversed in a fixed period of time than participants trained using standard control systems or less appropriate iterations of augmented reality.

Sensitivity (d')

For the sensitivity (d') measure, higher scores indicate a greater ability to discriminate landmines (Targets) from harmless objects (Non-Targets) (see Figure 3).

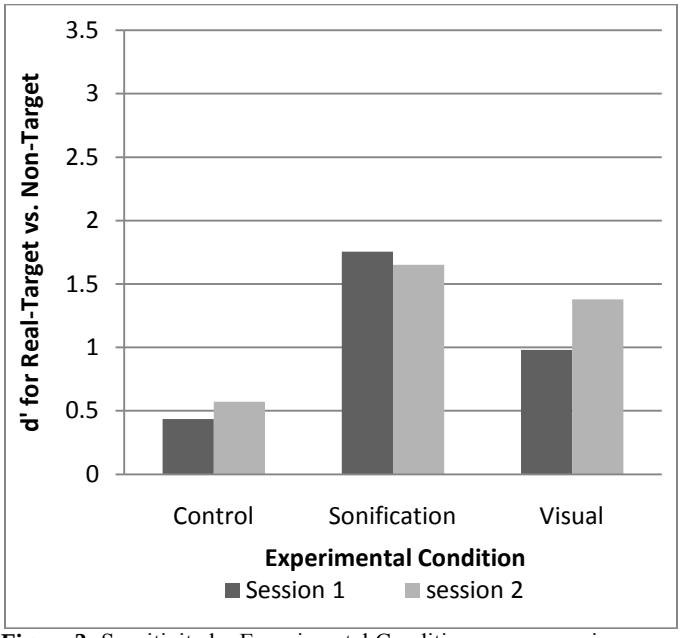

Figure 3: Sensitivity by Experimental Conditions across session.

In the first session participants using the sonification interface (d' = 1.75) had the greatest ability to discriminate landmine (targets) from harmless objects (Non-targets). The participants using the visual interface (d' = 0.98) were considerably more likely to have false alarms. However participants in the control group (d' = 0.44) were found to have the worst ability to discriminate between objects.

The sensitivity (d') measure appeared to be greatly affected by the increase navigation capabilities for participants in the sonification interface condition when

compared to either the control or visual interface conditions.

In the second session participants using the sonification interface (d' = 1.65) continued demonstrated the greatest ability to discriminate targets from non-targets. The participants using the visual interface (d' = 1.38) were less distinctively differentiated but still less accurate by comparison. The participants in the control group (d' = 0.57) again were found to have the worst ability to discriminate between objects.

DISCUSSION AND CONCLUSION

These results indicate that the accuracy of the participant's ability to perform their primary task (target identification) was clearly affected by the navigation related augmentations. Video analysis showed that participants in both augmented groups were much more likely to execute and maintain focused search patterns over the course of their missions than were participants in the control group. The participants using the sonification interface were particularly adept at maintaining a consistent speed, this speed control was certainly a factor in their ability to accurately interpret sensory information, which in turn lead to increased sensitivity.

In the second session the participants in both augmented groups continued to execute and maintain search patterns that were very similar to those demonstrated in session one. The members in the sonification group were however less consistent in their speed control.

The results do indicate that the sonification interface was successful in relaying valuable navigation relation information that lead to an increase in sensitivity. This same device once removed has been shown to be a viable training technology.

This study demonstrated that properly designed sonification AR interfaces successfully transmitted critical navigation information to human operators and this resulted in an increase in ground coverage (the number of sectors covered) and a decrease in path duplication (the number of sectors repeated) and that this resulted in a increase in navigation control which lead to more accurate primary task performance (sensitivity d').

The results of the second session showed that the sonification AR interface could be used as an effective training technology, participants in this group continued to demonstrate superior navigation skills and primary task sensitivity when compared to participants who were trained without augmented devices.

The properly designed sonification AR interface was shown to consistently outperform it less appropriate counterpart the visual AR interface even though they both augment the same relative position information. AR interfaces can be powerful aids for direct navigation or training application, they are however best designed with the consideration given first to the human system interaction.

REFERENCES

Boud, A. C., Haniff, D. J., Baber, C., & Steiner, S. J. (1999, July 14-26). Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks. Paper presented at the IEEE International Conference on Information Visualization., London, UK.

Bristow, H., Baber, C., Cross, J., & Woolley, S. (2002). Evaluating contextual information for wearable computing. Paper presented at the Proceedings of the 6th International Symposium on Wearable Computing, Los Alamitos, CA.

Brown, D. G., Stripling, R., & Coyne, J. T. (2006). Augmented Reality for Urban Skills Training. Paper presented at the Proceedings of the IEEE Virtual Reality Conference (VR'06), Orlando, FL.

Hamza-Lup, F. G., Santhanam, A., Fidopiastis, C., Davenport, P., & Rolland, J. P. (2004, Nov 2-5th). A Distributed Augmented Reality System for Medical Training and Simulation. Paper presented at the IEEE VR, Arlington, VA.

KoK-Lim, L., Llie, A., Welch, G., & Lastra, A. (2003). Combining head-mounted and projector-based displays for surgical training. Paper presented at the Virtual Reality, 2003. , Chapel Hill, NC.

Nielsen, C., Goodrich, M., Ricks, R. (2007). Ecological interfaces for improving mobile robot teleoperation. IEEE Transactions on Robotics, 23(5).

Schwald, B., & Laval, B. (2003). An Augmented Reality System for Training and Assistance to Maintenace in the Industrial Context. Journal of WSCG, 11(3), 425-432.

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A.,and Coodrich, C.,(2006) "Common metrics of human-robot interaction," *In Proceedings of HRI2006 1st annual conference on Human-Robot Interaction*, Salt Lake City, UT, USA, March 2-3 2006. ACM.

Stone, R.T., Augmented Multisensory Interface Design: Performance Enhancement Capabilities and Training Potential. (Doctoral dissertation, University at Buffalo, the State University of New York, 2008)