Evaluation of Multimodal Displays for Waypoint Navigation

Andre Garcia, Victor Finomore, Gregory Burnett, Andres Calvo, Carryl Baldwin,
& Chris Brill

ABSTRACT— Point to point navigation is a critical and demanding task for dismounted operators, especially while traversing hostile terrains. Visual displays such as a compass, maps, and global positioning systems have been the ubiquitous means of navigation and have proven to be effective; however, these tools require visual attention in an already visually demanding environment. Multiple resource theory proposes that time-sharing between two tasks with the same sensory modality can produce interference and the use of separate perceptual channels can be more effective. Since the dismounted operator's visual attention is already engaged in examining their environment for potential threats, the use of auditory and/or tactile displays could supplement the visual display to provide enhanced waypoint information. The reduction of the operator's visual load to allow for undivided attention to their environment could allow for more efficient and secure navigations through novel environments. The study investigated the effectiveness of waypoint navigation with the use of a visual map, spatialized auditory and tactile displays in a virtual environment. In addition to performance data, as measured by root mean squared error and time of completion, the participants completed usability and perceived mental workload questionnaires pertaining to the utility of the different displays.

Index Terms — Multimodal Display, Situational Awareness for Human-Computer Systems, Tactical Operations.

INTRODUCTION

Waypoint navigation is a critical task carried out by dismounted soldiers, especially when navigating through novel environments with potential threats (Mitchell, Samms, Glumm, Krausman, Brelsford, & Garrett, 2004). This endeavor is made even more complicated by unidentifiable threats such as improvised explosive devices, suicide bombers, or potential enemy threats hidden throughout civilian crowds. Dismounted soldiers need to maintain awareness of their operating environment and constantly scan their immediate surrounding for signs of potential threats. Navigating at night or in crowded environments with degraded or cluttered visual information and loud noises increase the workload and stress of the operator where the margin for error is already extremely small.

In these dangerous environments, the soldiers should have their “eyes-up” and “ears-out” scanning the environment for critical signals. Visual displays such as a compass, maps, and global positioning systems have been the ubiquitous means of navigation and have proven to be effective. However, these tools require visual attention in an already visually demanding environment. Multiple resource theory proposes time-sharing between two tasks appealing to the same sensory modality can produce interference, whereas decoupling multiple tasks from the same modality to different modalities can increase performance (Wickens, 1984). Thus supplementing the visual display with directional cues from other modalities could free the operator’s eyes to scan the environment instead of being heads down on the display with attention diverted from identifying potential threats. Two main technologies investigated in this study were spatial audio and tactile displays.

¹Manuscript received November 4, 2011.

- A. Garcia is a graduate student at George Mason University, Fairfax, VA 22030 USA (e-mail: agarciaf@gmu.edu).
- V. S. Finomore, Jr., is with the Air Force Research Laboratory’s 711 Human Performance Wing, Battlespace Acoustics Branch, Wright-Patterson AFB, OH 45433 USA (937-904-7123; e-mail: victor.finomore@wpafb.af.mil).
- G. M. Burnett is with the Air Force Research Laboratory’s 711 Human Performance Wing, Battlespace Acoustics Branch, Wright-Patterson AFB, OH 45433 USA (937-904-7956; e-mail: gregory.burnett@wpafb.af.mil).
- C. Baldwin is a professor at George Mason University, Fairfax, VA 22030 USA (703-993-4653; e-mail: cbaldwi4@gmu.edu).
- A. Calvo is a contractor with Ball Aerospace working at the Air Force Research Laboratory’s 711 Human Performance Wing, Battlespace Acoustics Branch, Wright-Patterson AFB, OH 45433 USA (937-904-7956; e-mail: andres.calvo.ctr@wpafb.af.mil).
- C. Brill is professor at Old Dominion University, Norfolk VA USA (757-683-4242; e-mail: jcbri11@odu.edu).

Spatial audio displays were first developed in the late 1980s and are defined as binaural headphone-based audio display systems that attempt to recreate the directional cues that allow human listeners to identify the location of sound sources in the real world (Wenzel, 1992). These directional cues are captured in the Head-Related Transfer Function (HRTF), which characterizes the physical acoustic transformations that occur as sound from a distant source interacts with a listener's head, torso, and pinnae. Information about the azimuthal location of a source is primarily contained in the Interaural Time Differences (ITD) and Interaural level differences (ILD) associated with the HRTF. These cues are easily recreated in a spatial audio display by playing back recorded sounds created using a "dummy head" microphone or through synthesis using digital audio processing. Although spatial audio displays have long been discussed as a potential interface, only recently have they approached the point where they might be practical for use by dismounted individuals for ground navigation. A number of researchers have demonstrated the potential utility of spatial audio cues for ground navigation (Ephrem, et al., 2008; Gunther, Kazman, & MacGregor, 2004; Loomis, 1998).

An alternative display investigated in this study was based on vibrotactile technology in which eight evenly-spaced linear actuators, called tactors, are embedded into a belt and worn around the participants' torso. These tactors vibrate against the skin to provide a physical stimulus used to direct the participants' attention. Research has shown that tactile technology is an effective way to relay preset forms of information for improving situation awareness (Brill, Mouloua, Hancock, & Terrence, 2004; Brill Terrence, Downs, Gilson, Mouloua, & Hancock, 2004, Rupert, 2000), facilitating battlefield communication (Brill & Gilson, 2006; Brill, Terrence, Stafford, & Gilson, 2006), presenting spatial orientation cues (Rupert, 2000; Terrence, Brill, & Gilson, 2005), and for navigation (e.g., Vap Erp & Van Veen, 2004; Van Erp, 2007).

The goal of this experiment is to explore the effectiveness of spatial audio and vibrotactile displays to quickly and accurately direct dismounted operators in a waypoint navigation task as compared to visual map navigation.

METHOD

Participants. Seven paid participants, 4 men and 3 women, ranging in age from 20 – 27 ($M = 23.5$) years old served as route-finders in this study. All participants were tested and passed for normal hearing as well as color vision and acuity.

Apparatus. The experiment was conducted in the Battlefield Airman Machine-Interface Laboratory located at Wright Patterson Air Force Base, Air Force Research Laboratory. Participants were situated in front of a 55-inch screen with a game controller to control their movement through the virtual urban environment, as seen in Figure 1. Visual directions were provided on a paper map whereas the spatial auditory tones were presented over stereo headphones, and the tactile cues were presented from an eight-tactor vibrotactile display belt from Engineering Acoustics, Inc. (EAI).

Figure 1. Participant navigating in virtual environment

Procedure. A within-subject design was employed with three display modality conditions (spatial audio, visual, and tactile). The task was for participants to navigate through a series of waypoints to reach their destination in a virtual urban environment. For each display modality condition, participants traversed seven unique routes, which were approximately 6-8 minutes long and ranged from 5-10 waypoints, as seen in Figure 2. The order of routes and modality condition were randomized via a Latin Square design. A training scenario was used so that participants could familiarize themselves with the equipment and scenarios. During training, a cross-modal matching procedure was used to equate perceived loudness across the auditory and tactile modalities.

Figure 2. Map of virtual environment with start and end points designated with green and red circles and nine waypoint designated by white Xs.

The visual condition consisted of a paper map, similar to Figure 2, with the route drawn on it. In the auditory condition, participants were presented with pulses of broad-bandwidth noise once every three seconds indicating the azimuthal direction of the next waypoint from the perspective of the center of the screen. For the tactile condition, the participants wore the EAI eight-tactor array on their torso. The eight-tactors were arranged in a manner consistent with compass points, with “north” tactor aligned with the naval and the “south” tactor resting on the spine. As in the Spatial Audio condition, participants received a 250 Hz sinusoid vibrotactile pulse every three seconds at the location of the waypoint. In all conditions, participants received an auditory tone once a waypoint was acquired.

RESULTS

Navigation Error. Mean root mean squared (RMS) error scores (Jagacinski & Flach, 2003) were calculated for all of the participants as seen in Figure 3.

Figure 3. Mean root mean squared error scores in meters for all three display modality conditions. Error bars are standard error.

Data from Figure 3 were tested for statistical significance by means of a within subjects analysis of variance (ANOVA). A significant main effect was found for modality condition, $F(2, 12) = 5.60, p < .05$. A post hoc test revealed that participants made significantly fewer navigation errors in the Tactile ($M = 6.39$) condition than Visual condition ($M = 10.46$), however, no other significant difference were found between the three conditions, $p > .05$. It is important to note that, although it was not significantly different, participants generally made fewer errors in the Auditory than Visual Condition.

Completion Time. Figure 4 displays the median completion times for all participants.

Figure 4. Median completion time for all three display modalities. Error bars are standard error.

A similar ANOVA was performed on the data in Figure 4 and revealed a statistically significant main effect for display modality, $F(2, 12) = 7.52, p < .05$. Post hoc tests found that participants completed the routes more quickly in the Tactile ($M = 335.94$) and Auditory ($M = 349.12$) conditions, as compared to the visual conditions ($M = 414.43$). No differences in route completion times were observed between the Tactile and Auditory Conditions.

DISCUSSION

This study proved to be a productive pilot study for larger ongoing efforts investigating the use of multimodal displays for presenting critical information to the dismounted operator as a means of increasing situation awareness and reducing workload. Future studies will increase the navigator's workload by adding visual threats and radio communication as one would experience in actual operation. Additionally we are also interested in investigating the effects of

combinations of congruent displays to present information (Merlo, Duley, & Hancock, 2010) as well as individual differences in navigation style and cognitive map construction (Furukawa, Baldwin, & Carpenter, 2004).

While these results are preliminary, they strongly suggest the utility and feasibility offloading traditional visual information pertaining to land navigation to auditory and/or tactile displays. This would allow the operator to keep their “eye out” to scan their immediate environment for potential threats.

CONCLUSION

It is critical that dismounted operators accurately navigate dynamic paths in unrecognizable and potential hostile environments while remaining alert of their surroundings. Accordingly, this study evaluated the effectiveness of advanced multimodal displays used for land navigation through a series of waypoints in a virtual urban environment. Navigational errors and completion times were used to evaluate the effectiveness of spatial audio and tactile navigation display as compared to navigating with paper map. It was found that participants made fewer navigational errors with the tactile display than with spatial audio displays or traditional paper maps. Further, they completed the routes more quickly with the tactile and spatial audio displays. These findings suggest that both the spatial audio and haptic displays provide adequate information for participants to find their next waypoint.

REFERENCES

- [1] Brill, J. C., & Gilson, R. D. (2006). Tactile technology for covert communications. *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA.
- [2] Brill, J. C., Gilson, R. D., Mouloua, M., Hancock, P. A., & Terrence, P. I. (2004). Increasing situation awareness of dismounted soldiers via directional cueing. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human Performance, Situation Awareness, and Automation: Current Research and Trends* (Vol. 1, p. 130-132). Mahwah, NJ: Lawrence Erlbaum Publishers, Inc.
- [3] Brill, J. C., Terrence, P. I., Downs, J. L., Gilson, R. D., Hancock, P. A., & Mouloua, M. (2004). Search space reduction via multi-sensory directional cueing. *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society*, New Orleans, LA.
- [4] Brill, J. C., Terrence, P. I., Stafford, S., & Gilson, R. D. (2006). A wireless tactile communication system for conveying U.S. Army arm-hand signals. *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA.
- [5] Ephrem, Finomore, Gilikey, Newton, Romigh, Simpson, Brungart, & Cowgill, 2008. An “Audio Annotation” technique for intuitive communication in spatial tasks. *Proceeding for the Human Factors and Ergonomics Society*.
- [6] Gunther, R. Kazman, R. & MacGregor, C. (2004). Using 3D Sound as a Navigational Aid in Virtual Environments. *Behaviour & Information Technology*, 23, 435-446.
- [7] Jagacinski, R.J., & Flach, J.M. (2003). *Control theory for humans: Quantitative approaches to modeling performance*. Mahwah, NJ: Erlbaum.
- [8] Loomis, J.M., Golledge, R.G., & Klatzky, R.L. (1998). Navigation System for the Blind: Auditory Display Modes And Guidance. *Presence*, 7, 193-203.
- [9] Merlo, J.L., Duley, A.R., & Hancock, P.A. (2010). Cross-modal congruency benefits for combined tactile and visual signaling. *American Journal of Psychology*, 123 (4), 413-424.
- [10] Mitchell, D. K.; Samms, C.; Glumm, M.; Krausman, A.; Brelsford, M.; Garrett, L. *Improved Performance Research Integration Tool (IMPRINT) Model Analyses in Support of the Situational Understanding as an Enabler for Unit of Action Maneuver Team Soldiers Science and Technology Objective (STO) in Support of Future Combat Systems (FCS)*; ARLTR-3405; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2004.
- [11] Rupert, A.H. (2000). Tactile situation awareness system: proprioceptive prostheses for sensory deficiencies. *Aviation, Space, and Environmental Medicine*, 71(9 Suppl), A92-99.
- [12] Terrence, P. I., Brill, J. C., & Gilson, R. D. (2005). Body orientation and the perception of spatial auditory and tactile cues. *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL.
- [13] Van Erp, J.B.F. (2007). *Tactile displays for navigation and orientation: perception and behavior*. Mostert & Van Onderen, Leiden, The Netherlands.
- [14] Van Erp, J. & Van Veen, H. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F*, 247-256.
- [15] Wenzel, E.M. (1992) Localization in Virtual Acoustic Displays. *Presence: Teleoperators and Virtual Environments*, 1, 80 – 107.
- [16] Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds), *Varieties of attention* (pp. 63-102). New York: Academic Press.