Effect of Beacon Sounds on Navigation Performance in a Virtual Reality Environment

Bruce N. Walker and Jeff Lindsay

Sonification Lab, School of Psychology
Georgia Institute of Technology
654 Cherry Street NW
Atlanta, GA 30332-0170
bruce.walker@psych.gatech.edu gte457e@prism.gatech.edu

ABSTRACT

Navigating without vision can be quite a daunting task. This study details an auditory interface to deal with this situation and then describes initial validation experiments. Ss were assigned to conditions with differing auditory beacons and assessed on how quickly and efficiently they were able to navigate through a series of virtual environments. The effect of beacon type was explored as well as practice effects due to multiple trials with the interface. The two experiments differed on the size of the beacon capture radius. Results show a main effect of beacon type on performance in the first study and a trend towards significance in the second study. A strong practice effect was also found across all conditions.

1. INTRODUCTION

Moving through varied and complex environments every day is something that most people do with ease. If, however, the input from the visual system is unavailable (e.g. damage to the optic nerves or smoke in a burning building), navigating and avoiding obstacles become much more demanding. For learning a geographic environment from primary sources in the environment, vision is the most effective sense [1]. It is therefore highly desirable to develop a navigation aid for use where visual input has become unavailable. Two studies reported here investigated different classes of sounds in an auditory interface for navigation.

Prior Investigation and Existing Interfaces

Tran, Letowski, and Abouchacra [2] have investigated some issues related to the use of acoustic beacons in navigation tasks. They studied the effect of beacon types on localization and navigation, the effect of real environments compared to virtual environments, and the qualitative aspects of various types of acoustic beacons.

First, Tran et al. [2] determined that a virtual 3D environment is adequate for azimuthal localization (i.e. localization in the horizontal plane). The system used in their study allowed for accuracy of between 7-15 degrees azimuth in the frontal direction. A more accurate localization accuracy would of course be preferable, and is also likely obtainable if user-specific head related transfer functions are used to create the virtual environment [2]. Next, they studied the different beacon sounds. The 10 beacons used in their studies ranged from pure tones to complex sounds, including both speech and non-speech. They found that beacon type had a significant effect on both number of localization errors made (or, accuracy) and the user's comfort level. Based on their findings, they suggested that any acoustic beacon intended for use in navigation tasks should be a wide-band non-speech sound, with a proper balance between low-and high-frequency energy to make the beacon pleasant and easy to localize. The sonar sound was found to be the most appropriate beacon overall when presented at a r epetitions per second. Another interesting result was that a user's rating of the quality of a beacon was highly correlated with the localization performance level achieved while using that beacon, indicating that subjective user rating of beacon quality could be a useful metric for selecting auditory beacons.

One of the reasons [2] evaluated both speech and non-speech beacons is that virtually all of the existing navigation systems use speech sounds to guide the listener. One such system is the Personal Guidance System (PGS) [3,4]. The PGS interface consists of a virtual 3D auditory environment where a computer creates spatialized speech beacons such that the perceived location of the beacon is at the place that the semantic content in the beacon refers (e.g. "Doorway here" as an auditory beacon). Loomis [4] had previously found that a "simple" virtual 3D auditory environment has the potential to provide navigation information to a visually impaired traveler, so their system, and others following it (e.g., [5]) have used speech beacons and spoken directions for navigation. However, given the findings [2] that speech as an auditory beacon is harder to localize in a virtual environment than non-speech beacons, and human factors principles suggesting avoidance of the speech channel when not absolutely necessary (e.g., [6]), our own auditory navigation projects have focused on non-speech audio. We first set out to test relative speed and efficiency of navigation using our system, with different non-speech beacon sounds. We also sought to study just how precisely the listeners could maneuver through a path, by varying the way that a listener interacts with the beacons.

Experimental Validation

Based on the limits of existing research in this area, it is clear that there are a number of fundamental research questions that still need to be addressed before higher order aspects of auditory interfaces can begin to be investigated. The studies reported here represent the initial findings in this forthcoming line of investigation. We went back to the basics and examined the most basic component of the proposed interface: the nature of the navigation beacon sounds. Tran [2] found that certain beacon types (non-speech sonar pings) were more localizable than others; the objective of this first study was to replicate and extend that finding within the context of our VR test interface.

2. METHOD

Participants

Participants consisted of Georgia Institute of Technology undergraduates with reported normal vision. All of the participants were screened for hearing problems prior to their participation. The participants were compensated by receiving course credit. In Study 1 there were 21 participants (16 male, 5 female, mean age=20.4, range 18 to 30). In Study 2 there were 18 participants (9 male, 9 female, mean age=21.2, range 18 to 28).

Design

The independent variable being manipulated in both studies was the beacon sound a participant was assigned to use for navigation through all the virtual environments. The dependent measures being recorded were the participants' position in three dimensional space with regard to time and the beacon they were currently navigating towards. The studies used a mixed design, with participants randomly assigned to one of three beacon sounds. Each condition had a set of three repeated measures resulting from navigating through three different environments. The only difference between the two studies was the capture radius of the beacons (described below). The materials and procedures for both studies were otherwise identical.

Apparatus

Three beacon sounds with the same approximate frequency (1 kHz), but of widely varying timbre characteristics were used. The first sound was a 1 s burst of broadband noise centered on 1 kHz. This particular sound had the most broad spectrum of the three. The second beacon was a 1 s pure sine wave with a frequency of 1 kHz. This was the most narrow spectrum. The third beacon sound was a 1 s sonar pulse, similar to the sound that Tran el al. (2000) found to be one of the best sounds for use as a navigation beacon.

Our VR environments were constructed using the Simple Virtual Environments (SVE) software package developed by the College of Computing at the Georgia Institute of Technology [7]. SVE was run on a Dell Optiplex (1.7Mhz, 528mb RAM). The beacon sounds were played through closed ear headphones. To change direction participants rotated on the spot where they were standing. They used two buttons on a joystick to control forward and backward movement in the VR environments (they did not actually walk around). Their orientation within the environment was tracked by an Intersense 2 head-mounted tracking cube attached to the headphones.

There were three different VR environments (maps) that participants were asked to navigate. Each environment is essentially a large empty room with four walls. The principle difference between each environment was the location of the beacons in the environment and hence the path the participants were asked to follow. The locations of the beacons were hard-coded into each environment, with five beacons in the first and ten in each of the two subsequent environments. The SVE software was modified to enable it to log the participants' current location in the environment (in terms of X, Y, and Z coordinates), their current head orientation (angular pitch, yaw, and roll), and the beacon they are currently tracking towards. All of this data was recorded approximately every 2 ms.

In addition to the perceptual aspects of the beacon sounds, we were interested in examining how precisely a listener can navigate such an auditory environment. One way to examine this is to vary the capture radius of the waypoints. To explain, if a user is moving from one map waypoint to the next, it is unlikely that the person's X,Y,Z location, as determined by the VR, will ever exactly coincide with the explicitly defined waypoint. They will typically come very close, which should constitute a successful arrival at the waypoint (and the activation of the next waypoint beacon), but will not technically arrive. To prevent this VR artifact from interfering with their navigation, a "capture radius" is defined around the waypoint, so that when the user enters the capture radius the system considers that a successful arrival. A larger radius is easier to navigate, but if the user needs to be constrained to a tighter path, such as in a corridor or on a sidewalk, a smaller capture radius is required. It remained unknown how the capture radius would affect movement, and whether it would interact with beacon type.

Other than the use of different participants, the two studies reported here, Study 1 and Study 2, differ only in the capture radius of the beacons (and 5 m and 30 cm, respectively).

Procedure

Each participant was randomly assigned to use one of the three beacon sound types. Participants in each sound condition navigated through each of the three maps using a single beacon type throughout. For example, participants in the noise condition navigated the three maps using only the noise beacon.

The experimenter explained the temporal aspects of the beacons, namely that tempo is mapped to distance from a beacon and that the sound is spatialized to indicate the beacons' relative direction. Participants were also warned about potential front-back confusions that would sometimes occur with artificially spatialized sounds and non-individualized HRTFs. Participants were then shown how to use a combination of body movements and buttons on the joystick in order to move and turn in the VR environment.

After making sure that participants did not have any questions about the beacons or how to navigate within the VR environment they were asked to move through the three maps one after the other. The order in which the maps were presented was the same across all conditions. There was a brief pause between maps in order to save the data and load the next map. Once participants finished navigating all of the maps they were asked to rate the ease of navigating each environment on a scale from 1 to 5 (1 being very hard and 5 being very easy). Following completion of the third map, the experimenter explained the purpose of the study, answered any questions, and thanked the participant.

3. RESULTS

A two way mixed factors multivariate analysis of variance (MANOVA) was run on the data from each study. There were two major dependent variables of interest: overall map completion rate (operationalized as meters of the map traversed per unit number of seconds) and efficiency (the ratio of the scheduled path to the actual traveled path). We define the "scheduled path" as the length of a map, calculated by adding the Cartesian length of each segment. That is, if the person walked precisely from waypoint to waypoint to waypoint, they would travel a certain distance over the course of a map. However, the participants actually traveled a less linear path that was usually (but not always!) longer than the scheduled path. The çextraé distance they traveled can be considered as wasted time and effort, and we predicted that comparing the path the person was supposed to travel, compared to the path they actually followed, would be a useful metric of the effectiveness of the different beacon sounds and capture radii.

Study 1 Results

Figures 1 and 2 present the results of Study 1 for movement rate and path efficiency, respectively. Across the three maps, participants showed a significant improvement in performance due, presumably, to practice. The main effect of map was significant for both movement rate, $F(2,18)=52.864$, $p=.00$, and path length efficiency, $F(2,18)=5.592$, $p<.01$. In addition, collapsing across maps, participants performed best with the noise burst beacon, followed by the pure tone, then the sonar pulses. For movement rate the main effect of beacon type was significant, $F(2,18)=2.948$, $p<.08$. The multivariate interaction between map and beacon type and beacon type alone tended toward traditional levels of significance, $F(4,16)=2.069$, $p=.13$ and $F(2,18)=2.960$, $p<.08$ respectively.

It is helpful to consider Figures 3 and 4 for a graphical depiction of paths with noise and pure tone beacons. Clearly the noise beacon (Figure 1) resulted in more successful navigation.

Study 2 Results

Figures 5 and 6 present the results of Study 2 for movement rate and path efficiency, respectively. The results are very similar. Across the three maps, participants again showed a significant practice effect. The main effect of map was significant for both movement rate, $F(2,15)=72.751$, $p=.00$, and path length efficiency, $F(2,15)=9.904$, $p=.00$. Across maps, participants again performed best with the noise burst beacon, followed by the pure tone, then the sonar pulses. The multivariate interaction between map and beacon type and beacon type alone tended toward traditional levels of significance, $F(4,13)=2.068$, $p=.14$, and $F(2,15)=2.019$, $p=.17$ respectively.

Additional Comments for Both Studies

This level of significance is not the usual statistical standard, but there is a definite trend towards significance at a more stringent alpha. These results should be viewed in light of the number of participants in each study. It is believed that running more participants will make the trends found even more evident.

The subjective measures for the first study showed that 71% of participants felt that navigating using the interface was easy to begin with or became easier as they progressed through the maps. In the second study, however, only 44% of participants indicated a similar perception of ease. This difference is likely due to the smaller capture radii of the beacons in the second study making navigation more difficult.

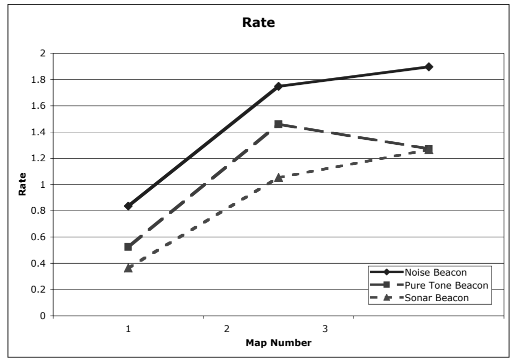

Figure 1. *Movement rate sorted by beacon type for Study 1.*

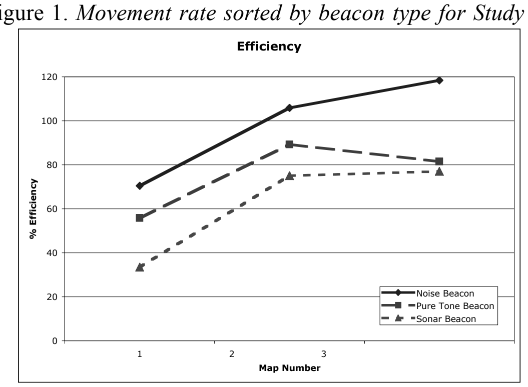

Figure 2. *Efficiency sorted by beacon type for Study 2*

Figure 3. *Paths traveled by participants using the noise beacon on Map 2 in Study 1*

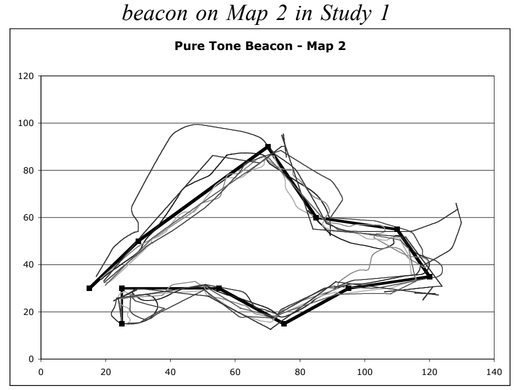

Figure 4. *Paths traveled by participants using the pure tone beacon on Map 2 in Study 1*

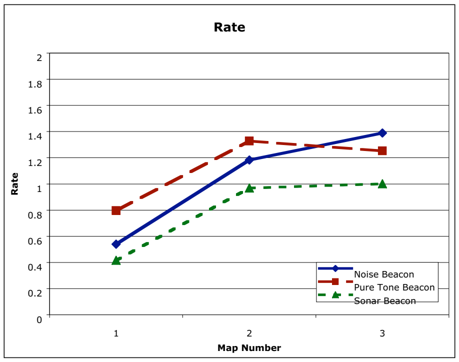

Figure 5. *Movement rate sorted by beacon type for Study 2.*

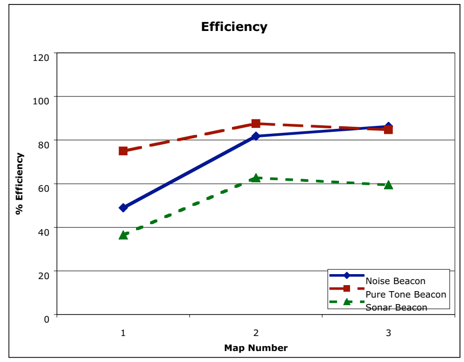

Figure 6. *Efficiency sorted by beacon type for Study 2.*

4. DISCUSSION

Across these two studies there are three main points to be made. First, practice has a major effect on performance, which is not surprising, given that none of the participants had experienced an auditory wayfinding system before. Thus it is critical to examine performance longitudinally when evaluating auditory display designs. The second point is that different beacon sounds lead to markedly different performance. The predicted best performer, the sonar pulse, was actually the worst for navigation in these studies. This reinforces the point first made by Walker and Kramer [8] that it is crucial to empirically test any sound design, and not rely on the intuitions of the designer, or the results of one prior study. It is interesting to note, as well, that the noise burst was not necessarily the best performer at the outset, but it seems that with practice it allowed listeners the best cues for effective navigation. Clearly this sort of practice by sound type interaction is important to consider in further experiments and future systems. And finally, when comparing the two studies, here, the larger capture radius in Study 1 resulted in faster times and shorter path lengths. The smaller radius in Study 2 was still large enough to support navigation, but it was not as effective. Clearly this is one example of how issues beyond the acoustic design of the sounds need to be considered in any effective real-world auditory navigation system.

5. REFERENCES

[1] Jacobson, R.D. (1998). 'Cognitive mapping without sight: Four preliminary studies of spatial learning.' *Journal of Environmental Psychology*, 18, 289-305.

[2] Tran, T.V., Letowski, T., and Abouchacra, K.S. (2000). 'Evaluation of acoustic beacon characteristics for navigation tasks'. *Ergonomics*, 43 (6), 807-827.

[3] Loomis, J.M, Golledge, R.G., Klatzky, R.L., Speigle, J.M., and Tietz, J. (1994). 'Personal guidance system for the visually impaired'. *Proceedings of the First Annual International ACM/SIGCAPH Conference on Assistive Technologies*, Marina del Rey, CA.

[4] Loomis, J.M., Hebert, C., and Cicinelli, J.G. (1990). 'Active localization of virtual sounds'. *Journal of the Acoustical Society of America*, 88, 1757-1764.

[5] Helal, A., Moore, S.,and Ramachandran, B., "Drishti: An Integrated Navigation System for Visually Impaired and Disabled", Fifth International Symposium on Wearable Computers (ISWC'01), 2001.

[6] Salvendy, G., *Handbook of Human Factors and Ergonomics*, 2nd edition, Wiley, New York, 1997.

[7] GVU Virtual Environments, SVE Toolkit. (1997, April 8). Retreived November 24, 2000, from http://www.cc.gatech.edu/gvu/virtual/SVE/

[8] Walker, B.N., & Kramer, G. 'Mappings and metaphors in auditory displays: An experimental assessment.' In S. Frysinger & G. Kramer (Eds.), *Proceedings of the Third International Conference on Auditory Display, ICAD 96*, ICAD, Palo Alto, CA, 1996, p.71-74.