NAVIGATION PERFORMANCE IN A VIRTUAL ENVIRONMENT WITH BONEPHONES

Bruce N. Walker and Jeffrey Lindsay

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

ABSTRACT

Audio navigation interfaces have traditionally been studied (and implemented) using headphones. However, many potential users (especially those with visual impairments) are hesitant to adopt these emerging wayfinding technologies if doing so requires them to reduce their ability to hear environmental sounds by wearing headphones. In this study we examined the performance of the SWAN audio navigation interface using bone-conduction headphones ("bonephones"), which do not cover the ear. Bonephones enabled all participants to complete the navigation tasks with good efficiencies, though not immediately as effective as regular headphones. Given the functional success here, and considering that the spatialization routines were not optimized for bonephones (this essentially represents a worst-case scenario), the prospects are excellent for more widespread usage of bone conduction for auditory navigation, and likely for many other auditory displays.

1. INTRODUCTION

Audio navigation interfaces have usually been studied using headphones, for reasons including cost and the ease of obtaining off the shelf hardware designed to work with them. However, many potential users are hesitant to adopt emerging audio navigation technologies if doing so requires them to impair their ability to hear environmental sounds by wearing headphones. If vision is unavailable, audition is the most reliable distal sense remaining. However, even a very effective navigation system is unlikely to be adopted if users are forced to choose between the information the system provides and all other external auditory cues. For this reason, we have begun investigating the effectiveness of bone-conduction headphones as an alternative display device. Specifically, we have examined the effects on performance of these alternative display devices using the System for Wearable Auditory Navigation (SWAN) [1].

1.1. SWAN System and Prior Work

The SWAN system has been described in detail elsewhere by Walker and Lindsay [2, 3]. Briefly, the SWAN is an auditory interface composed of spatialized, non-speech auditory icons and earcons that aid users in navigation and awareness of features in the environment. Sounds in SWAN are classified as beacon sounds, object sounds, and surface transition sounds.

Beacon sounds are used for navigation, indicating the path the user is to follow to reach the desired destination. These sounds are placed (virtually) at waypoints along a route to the destination the user has selected. The sound is spatialized, appearing to emanate from the direction of the waypoint. As a user approaches a waypoint, the tempo of the beacon sound increases. When the user reaches the waypoint, the current beacon sound ceases and the beacon for the next waypoint becomes audible. Using this trail of beacon sounds the SWAN is able to guide users through their environment.

Object sounds indicate features in the environment that could potentially be of interest (e.g., a water fountain) or hazardous (e.g., a table blocking the hallway). Surface transition sounds denote changes in the surface the user is walking on (e.g., transition from carpet to tile). These can often indicate important boundaries (e.g., transition from sidewalk to street).

Previously, Walker and Lindsay's work with the SWAN has focused on beacon sound design and how user interaction with the sounds is affected by their display parameters [1]. They have examined what types of sounds result in good performance when used as auditory beacons. Using the metrics of path efficiency (how closely a user follows the prescribed path) and time efficiency (how quickly a user travels the prescribed path), Walker and Lindsay confirmed that auditory beacon sounds have a significant effect on users' efficiencies, and that broad spectrum sounds, such as a pink noise burst, which are more easily localized, result in better performance. They have also investigated the impact of user interaction with the beacons. Specifically, Walker and Lindsay [1] studied the effect of varying the capture radius of auditory beacons, where capture radius is how close to a beacon's location the user must achieve before the system will consider the user to have reached the beacon. Their findings indicate that a capture radius that is very large or very small (i.e., greater than 9ft. or only a few inches) results in decreased performance compared to a medium size capture radius. In addition, users' behaviors when interacting with a large capture radius (e.g., 'cutting corners') or small capture radius (e.g., overshooting the beacon) raise potential real world safety concerns. Similar real world safety concerns over diminishing users' ability to hear with conventional headphones are a chief motivation for investigation into the use of alternative display technologies such as bone-conduction headphones. However, these studies were done with conventional headphones, and there is a need to study performance in the SWAN with alternative output devices.

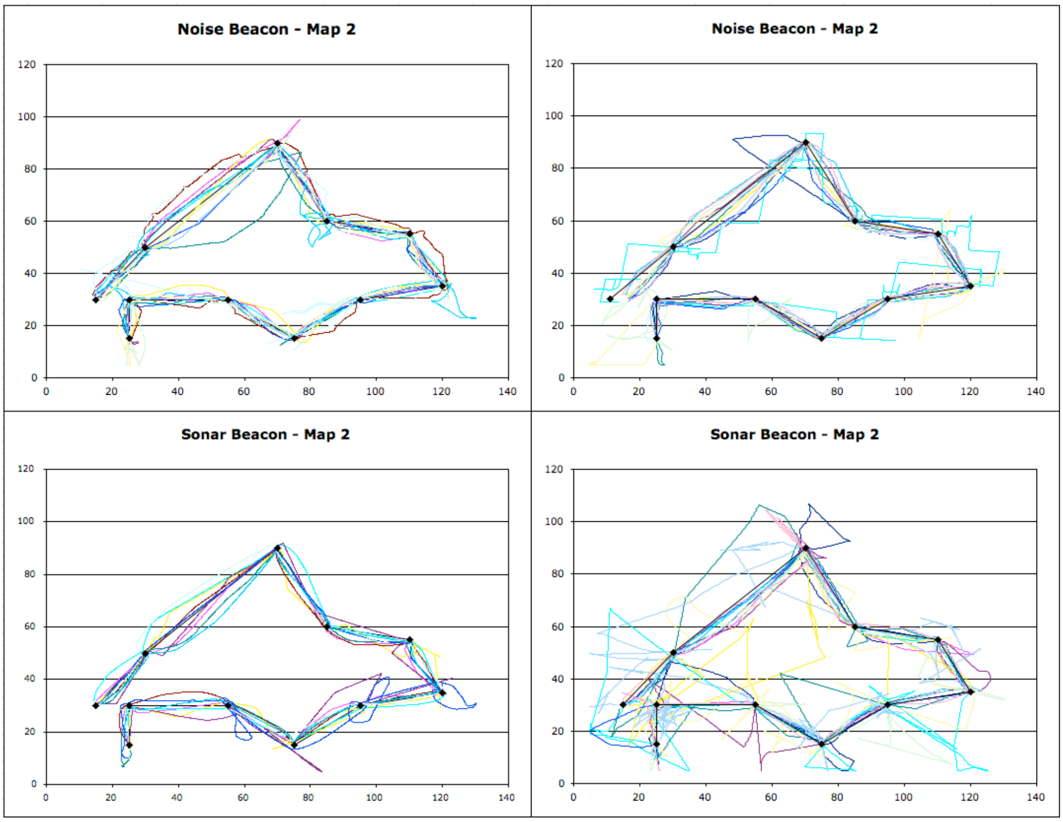

Figure 1. *Raw movement data for Path 2 with headphones (from previous studies, left panels) and bonephones (from the present study, right panels). Top panels indicate noise beacons and bottom panels indicate sonar beacons. Bone phones supported navigation, but errors were larger, especially for the sonar beacon. Note that HRTFs were not optimized for the bonephones, which puts them at an obvious disadvantage. This could be overcome with psychoacoustical research currently underway.*

1.2. Bonephones

Bone-conduction headphones are similar to conventional headphones in that both have vibrating bodies that generate pressure waves we perceive as sounds. The difference between them is their primary medium for transmission of these waves. Typical headphones transmit these waves through the air, whereas bone-conduction headphones send the sound through the bones in the skull directly to the cochlea. The advantages of bone-conduction headphones for use with the SWAN are that they do not obstruct the pinnae or ear canal and they are small and relatively discrete. Relatively new as a display device, the psychoacoustical properties of bone-conduction headphones have not been well explored, though some recent work has begun in this area (e.g., [4, 5]). Beyond the perceptual aspects, it is important to investigate whether these devices can enable a listener to complete auditory tasks traditionally accomplished using headphones. Thus, in this study we looked at performance of the SWAN audio navigation interface using bonephones.

2. METHODS

2.1. Participants

Participants were sighted undergraduates at the Georgia Institute of Technology. The 28 volunteers (18 male, 12 female; mean age 19.8, range 18 to 24) reported normal or corrected-to-normal hearing, and received course credit for participating.

2.2. Apparatus

The experimental apparatus was similar to that used in Walker and Lindsay [1]. As in previous studies, the SWAN virtual reality (VR) testing environment was used. This VR environment was built using the Simple Virtual Environments (SVE) software developed at the Georgia Institute of Technology [6]. Sounds were spatialized using OpenAL calls to an external Soundblaster Extigy sound card, with a non-individualized Head Related Transfer Function (HRTF). Sounds were output through Temco binaural (stereo) bone-conduction headphones. In order to move through the VR space, participants used a modified joystick with only two buttons: pressing one moved them straight forward; pressing the other moved them straight backward. To turn or rotate in the VR, participants rotated in place (where they were standing); their real orientation was noted by a head-mounted tracking device (InterSense InertiaCube2), which was translated into rotation in the VR world.

2.3. Procedure

Participants were asked to navigate a series of three paths in the VR. The first path consisted of five waypoints, and the other paths each contained ten waypoints. Throughout the experiment each participant's position in the VR (X, Y, and Z coordinates), their head orientation (pitch, yaw, and roll), and their current waypoint were logged approximately every 200 milliseconds. Each participant was assigned to use one of two beacon sounds. Once assigned, the beacon sound did not change for a

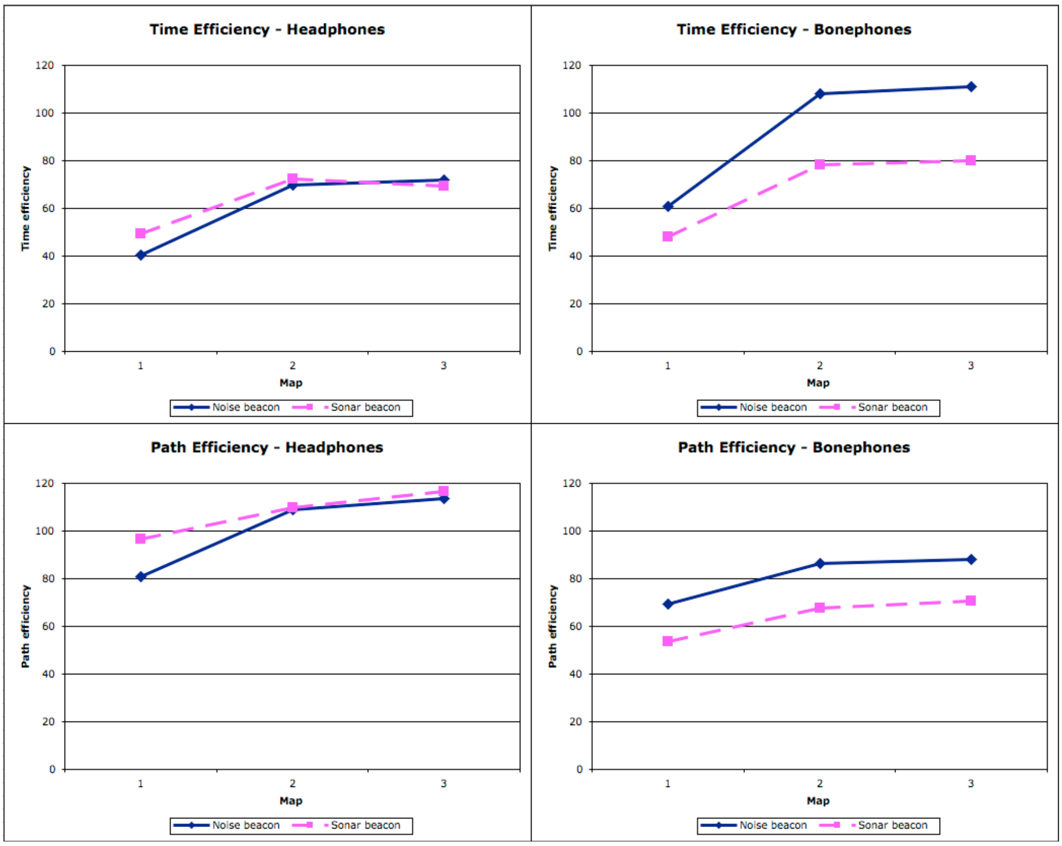

Figure 2. *Time efficiencies (top panels) and path efficiencies (bottom panels) for headphones (from previous studies, left panels) and bonephones (from this study, right panels). Solid lines indicate noise beacons and dashed lines indicate sonar beacons. Performance with bonephones was more efficient in terms of time, but less efficient in terms of path length. This reflects that the speed-accuracy tradeoff was different for the two output devices, but does not imply that one is "better" or "worse" than the other.*

participant across all three maps. Both of the sounds were approximately 1 second long and spectrally centered around 1kHz. The first sound was a burst of pink noise and the second was a sonar 'ping.'

3. RESULTS

The raw movement data for each participant in the present study (using bonephones) are presented in Figure 1, alongside data obtained in prior investigations using regular headphones [1]. As evident in the figure (right panels), participants completed the paths with the bonephones, but the movement paths were generally more erratic than with headphones. This is not surprising, given that the audio spatialization routines were not optimized for this novel presentation hardware.

The data collected here were then processed by computing the magnitude of the movement vector for a participant across each position measurement and obtaining an overall time measurement. These measurements were then used with the planned (or optimal/shortest) path length and the constant movement rate to normalize the measurements to account for the differing lengths of the paths. The results of this process are referred to as path efficiency and time efficiency (see Figure 2). Path efficiency is a percentage measure of the distance traveled by a participant in relation to the length of the planned path. Similarly, time efficiency is a percentage measure of the time taken by a participant to complete navigation of a path in relation to the shortest possible time assuming the path was followed perfectly. It is important to note that by traveling a shorter path than the planned path (e.g., cutting corners), it is possible for participants to achieve a percentage efficiency that is greater than 100 (likewise for time efficiency). The comparison involved only headphone data with the same beacon sounds and capture radius used in this study. In Figure 2 it can be seen that the bonephones resulted in faster but less accurate performance than headphones (this supports the more erratic raw movement traces shown in Figure 1). Thus, the speed-accuracy tradeoff was different for the two output devices. This does not, in itself, suggest which, if either, of the devices is "better", overall.

To consider the two dependent measures (path efficiency and time efficiency) together, a multivariate analysis of variance was performed with the between-subjects independent variables beacon sound type and headphone type, and the within-subjects independent variable practice (i.e., path). The results of this analysis showed a significant multivariate interaction of practice and headphone type, $F(4, 43) = 12.26$, $p < .001$, Wilk's Lambda = .467, a significant interaction of practice and beacon sound, $F(4, 43) = 2.70$, $p < .05$, Wilk's Lambda = .799, as well as significant main effects of practice, $F(4, 43) = 32.61$, $p < .001$, Wilk's Lambda = .248 and headphone type, $F(2, 45) = 73.07$, $p < .001$, Wilk's Lambda = .235. The interaction of beacon sound and headphone type was marginally significant, $F(2, 45) = 2.72$, $p < .10$, Wilk's Lambda = .892.

These significant effects lead us to examine the effects for each of the dependent variables. For all the multivariate cases mentioned above, both dependent measures were significant with one exception. For the interaction of practice and headphone type there was only a significant effect on time efficiency, $F(1, 46) = 5.494$, $p<.05$, but not on path efficiency. This can been seen in Figure 2 and the implications of this are discussed below.

4. DISCUSSION

When examining the efficacy of bone-conduction headphones for use with an auditory navigation interface, there are several important conclusions. The results of this study indicate that it is indeed possible to navigate using bone-conduction headphones. This is a relatively subjective question, but given the performance results, it seems reasonable to conclude that navigation using the SWAN with bone-conduction headphones is certainly viable. All participants did complete each path they were given, indicating that they were able to perform the navigation task.

In terms of comparative performance, as can be seen in Figure 2, participants deviated more from the path when using the bone-conduction headphones that when using traditional headphones (see path efficiency measure). However, participants using the bone-conduction headphones had a better overall time efficiency than those using the headphones. This is essentially a speed accuracy trade-off.

When considering these results, it is important to note that the spatialization algorithms (i.e., the HRTFs) built into the sound card were not optimized for sound conduction through bone, but rather sound conduction through air. It is likely that by determining the appropriate “bone related transfer function” (BRTF), the sound localization performance, and therefore the navigation performance, would be considerably increased. Beginnings of this research are underway, and reported by Walker and Stanley [5]. Future investigation into the characteristics of spatializing sounds with these devices is a fertile direction for future research. This is encouraging despite what prevailing opinion may have been. However, further research is still required.

Nevertheless, the take home message from this initial study is that bone conduction headphones are likely to provide an effective alternative to headphones, wherever access to ambient sounds must be maintained, or in other situations where covering the ears is inappropriate. Wayfinding interfaces that rely on auditory cues, such as the SWAN, are excellent candidates for bonephones, and bonephones seem ready to make contributions to the utility of the system, and the safety of the users.

5. REFERENCES

[1] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," *Human Factors*, in press.

[2] B. N. Walker and J. Lindsay, "Effect of beacon sounds on navigation performance in a virtual reality environment," Proceedings of the Ninth International Conference on Auditory Display ICAD2003, Boston, MA, 2003.

[3] B. N. Walker and J. Lindsay, "Auditory navigation performance is affected by waypoint capture radius," Proceedings of the Tenth International Conference on Auditory Display ICAD2004, Sydney, Australia, 2004.

[4] B. N. Walker, R. Stanley, N. Iyer, D. S. Brungart, and B. D. Simpson, "Evaluation of bone-conduction headsets for use in multitalker communication environments," Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society HFES 2005, Orlando, FL, 2005.

[5] B. N. Walker and R. Stanley, "Thresholds of audibility for bone-conduction headsets," Proceedings of the International Conference on Auditory Display ICAD2005, Limerick, Ireland, 2005.

[6] GVU Virtual Environments Group, "SVE Toolkit," vol. 2000, 1997.