

DEVELOPMENT AND EVALUATION OF A SYSTEM FOR WEARABLE AUDIO NAVIGATION

Bruce N. Walker and Jeffrey Lindsay
Sonification Lab
School of Psychology
Georgia Institute of Technology
Atlanta, Georgia, USA

If it is not possible to use vision when navigating through one's surroundings, moving safely and effectively becomes much harder. In such cases, non-speech audio cues can serve as navigation beacons, as well as denote features in the environment relevant to the user. This paper outlines and summarizes the development and evaluation of a System for Wearable Audio Navigation (SWAN), including an overview of completed, ongoing, and future research relating to the sounds used, the human-system interaction, output hardware, divided attention, and task effects.

INTRODUCTION

When vision is unavailable for navigating the environment, moving safely and effectively can become very difficult. In these cases, automated navigation aides can greatly assist by providing information that is otherwise difficult or impossible to attain without vision. The development of such an aide has been the focus of research by Walker and Lindsay with the System for Wearable Audio Navigation (SWAN) (in press). The purpose of the present paper is to summarize the ongoing development and evaluation of the SWAN, and discuss issues relevant to audio navigation.

Importance of Non-visual Navigation

According to the World Health Organization (Resnikoff et al., 2004), the number of visually impaired individuals worldwide is estimated at 161 million. In the United States alone there are an estimated 11 million visually impaired. These numbers are expected to increase in our aging population due to age related vision impairments such as macular degeneration. Clearly there is a significant population who face the task of navigating with diminished vision. Even with training, this is no easy feat and there is some information that is difficult to obtain even with mobility training and an assistive device such as a cane. Thus, a navigation aide that can augment a user's soundscape and provide information such as their current location, the optimal route to their destination, and nearby features (e.g., a water fountain or restroom) could be extremely beneficial to persons with visual impairments.

In addition to those with vision loss, persons with normal vision may be denied the use of vision due to environmental factors (e.g., a firefighter in a smoky building) or because vision is being used for a more immediate task (e.g., tracking an enemy plane during a dogfight). In such instances a navigation aide that utilizes sound may be able to improve performance by supplying data that is not available visually.

Why Audition?

In cases where vision is not an option, audition is a good alternative. It is one of the oldest distal senses and people generally are very good at recognizing and localizing complex sounds. Outside of vision, audition produces the most precise localization of any other sensory pathways, and is therefore a logical choice to convey spatial information.

Prior Investigation

While there is a body of theoretical work that has bearing on creating an auditory interface for navigation, there is relatively little applied research in creating such an interface with non-speech audio. Tran, Letowski, and Abouchacra (2000) have investigated what types of sounds might be good as beacon sounds. They examined the localizability of a set of 10 diverse sounds (including speech sounds) and users' subjective preferences for the sounds. They found that a broad spectrum sound was most easy to localize, and that sounds preferred by users also tended to be the best beacon sounds. Their results also indicate that speech is most likely not an optimal beacon sound to use. Despite this, many of the interfaces

that have been designed to date have relied on speech as the primary display method.

One of the earliest auditory navigation systems is the Personal Guidance System (PGS) (Loomis, Gollidge, Klatzky, Speigle, & Tietz, 1994). The PGS is composed of a 3D auditory environment where spatialized speech beacons are placed so that the sound is perceived to come from the same location as the object it represents. Spoken directions and speech beacons guide the user (e.g., repeating “Left, left, left...”). Another more recent system is the Drishti system (Helal, Moore, & Ramachandran, 2001). Drishti is similar to PGS in that it uses synthesized speech beacons. However, it also has a more complex mapping system that factors in user preferences and environmental features. While systems such as these have had some success, our belief based on the work of Tran et al. (2000) is that non-speech audio allows for increased performance and improved multitasking in an auditory navigation interface.

SWAN

The SWAN system has 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 should follow to reach the desired destination. These sounds are placed (virtually) at waypoints along a route from the user’s current location 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.

Object sounds and surface transition sounds provide users with information about the environment as they move along the path of beacon sounds. Object sounds indicate features in the environment that could potentially be of interest (e.g., a water fountain or restroom) or hazardous (e.g., a table blocking the hallway). Surface transitions are sounds that denote changes in the surface the user is walking on (e.g., transition from carpet to tile) or important boundaries (e.g., transition from sidewalk to street).

The SWAN is designed to be used on a wearable device. Various types of sensors (e.g., GPS) gather information about the user’s location and surroundings, which is then displayed via the SWAN audio interface. In addition to the wearable version, there is a virtual

environment version for interface development and indoor testing.

Research Questions

Our focus has always been to shape the SWAN interface through empirical evaluation, as well as to increase what is known about auditory navigation interfaces in general. When examining the most basic facets of such an interface, there are three important questions to be answered:

1. What types of sounds are appropriate for use in such an interface (i.e., for beacons, objects, etc.)?
2. How do users interact with the sounds in the system (i.e., how do factors such as when beacon sounds change impact performance)?
3. What hardware is most appropriate for such an interface?

COMPLETED RESEARCH

Walker and Lindsay have conducted a series of studies (in press) to investigate these questions. The initial experiment was primarily to (a) verify that navigation with the SWAN interface was possible and (b) examine what beacon sounds lead to the best performance. For safety reasons and ease of prototyping, a virtual reality (VR) version of the SWAN was created for these preliminary experiments. Using a subset of beacons similar to those used by Tran et al. (2000), Walker and Lindsay evaluated sighted undergraduates’ performance in terms of time taken to complete preset paths as well as how closely users followed those paths. The results indicated that the SWAN interface was intuitive and effective for navigation, and that broad spectrum sounds (in this case a pink noise burst and a sonar ping) are most appropriate as beacons.

The next investigation examined the effects of a beacon’s capture radius on performance. Capture radius is the proximity to a beacon’s location a user must achieve before the system will consider the user to have reached the beacon. Walker and Lindsay (in press) found that the actual size of the capture radius can have a very significant effect upon performance when using the SWAN. If a capture radius is too large (i.e., 3 meters) this can lead to users ‘cutting corners’ because the beacons change too far from where they are located. This can actually lead to shorter times and distances traveled, but there are safety concerns with a beacon sound that causes users to deviate from the planned path. Conversely, Walker and Lindsay found that if the capture radius was extremely small, users would often

walk right past it, but not close enough to elicit the shift to the next beacon sound. This resulted in users frequently having to retrace their steps, sometimes more than once, in order to get close enough to the beacon to enter the capture radius. Users in this condition traveled more slowly and farther because of these ‘overshoots.’ Walker and Lindsay found that a ‘human-scale’ capture radius (i.e., about 1.5 meters) is effective as well as safe.

Walker and Lindsay have also conducted some preliminary studies with visually impaired participants. While their preliminary results indicate that such users are able to use the SWAN, a larger sample size is required to make a more definitive assessment. One of the most valuable aspects of this testing has been the subjective feedback provided by these participants. One of the most often made comments was that no matter how good the system was, they would not give up access to other auditory stimuli by wearing headphones.

In light of this, Walker and Lindsay have begun testing performance using bone-conduction headphones (bonephones). These headphones do not cover the ear, but instead are located just behind the ear and transmit sound directly through the skull to the cochlea. While it was initially uncertain that these devices could even produce spatialized sound, Walker, Stanley, Iyer, Simpson, and Brungart (2005) have found that it is indeed possible to produce lateralized sound using the bonephones. Further, Walker and Lindsay (2005) have found that participants using bonephones are able to navigate using the SWAN nearly as well as participants using conventional headphones. It is important to consider this finding in light of the fact that the spatialization algorithms used were not optimized for bonephones, and we hypothesize that with a “bone-related transfer function” the navigation performance could equal performance with headphones.

ONGOING AND FUTURE STUDIES

While Walker and Lindsay have begun examining the most fundamental issues with non-speech auditory navigation aides, there remain many aspects that have yet to be investigated and are ripe areas for future research:

1. How does a multitasking/divided attention paradigm affect performance?
2. What aspects of the system are affected by allowing users to customize the interface?
3. What sounds are most appropriate for labeling objects/aspects of the environment?
4. How is gait/movement affected by use of the wearable system?

Ongoing work by Walker and Lindsay is beginning to look into these questions. For example, in one study we examined performance in a divided attention task, similar to walking down the street, navigating along a path, and periodically having to respond to a spoken query. Performance on the primary navigation task was slowed by the secondary speech comprehension task, but participants were still fully able to continue walking and navigating. This points to the effectiveness of a non-speech beacon. The detailed analysis of video from this study are being completed, and gait changes, speed, efficiency, and other metric will be presented and discussed. Also, the results of studies of the most effective and acceptable sounds for labeling the features in the environment will be presented. These new findings, in combination with the other results discussed here, delineate an ongoing program of research to determine the most effective attributes of the audio interface for the SWAN audio navigation system.

REFERENCES

- Helal, A., Moore, S., & Ramachandran, B. (2001). Drishti: An Integrated Navigation System for Visually Impaired and Disabled. *Proceedings of the Fifth International Symposium on Wearable Computers (ISWC'01)*, Zurich.
- Loomis, J. M., Golledge, R. G., Klatzky, R. L., Speigle, J. M., & Tietz, J. (1994). Personal guidance system for the visually impaired. *Proceedings of the Proceedings of the First Annual International ACM/SIGCAPH Conference on Assistive Technologies*, Marina del Rey, CA.
- Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., et al. (2004). Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization*, 82(11), 844-851.
- Tran, T. V., Letowski, T., & Abouchacra, K. S. (2000). Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics*, 43(6), 807-827.
- Walker, B. N., & Lindsay, J. (2005). Auditory navigation performance using bone-conduction headphones. *Proceedings of the International Conference on Auditory Display*, Limerick, Ireland.
- Walker, B. N., & Lindsay, J. (in press). Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice. *Human Factors*.
- Walker, B. N., Stanley, R., Iyer, N., Simpson, B. D., & Brungart, D. S. (2005). Evaluation of bone-conduction headsets for use in multitaler communication environments. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL.